The impact of medium energy electrons on mesospheric chemistry and dynamics

Master Thesis in Space Physics

by

Héctor Daniel Zúñiga López

Supervisors:

Hilde Nesse Tyssøy

Christine Smith-Johnsen

Department of Physics and Technology

University of Bergen

October 2021

Acknowledgements

First of all, I would like to thank my supervisor, Hilde Nesse Tyssøy, for always being welcoming and helping me to get through this thesis. I would also like to thank my co-supervisor, Christine Smith-Johnsen, for always finding the time to give me all the data I've asked for, and to help me.

Thanks also to the rest of the EPP group, who have probably heard about my thesis too many times, for the nice times in the different group activities, such as workshops and meetings. Specially thanks to Ville Maliniemi, who took the time to help us have a clearer picture on the results of this thesis.

I am also grateful to my roommates Angela, Veronika, Rami, Sindre, Malo, Lorena and Luisa for making my stay in Bergen a nicer experience. I am also grateful for my friends back home, Juanda, Tetty, Pau, Pepe, and my ex-basketball team, whom I hope I can see soon, just for checking on me from time to time and make me remember they are still there. Special thanks to Judit for always being around, both in Bergen and back home.

Last, I wish to thank my parents, Patricio and Paqui, who always worry about me and who thankfully could come and visit me last summer. Also to my siblings, Patricio and Andrea, for checking on me and wishing the best for me. And to my beautiful, yet annoying nieces, Lucía, Ariadna and Aroha, whom I can not wait to see again.

Abstract

The energy transfer from the solar wind to the Earth's magnetosphere fuels the Energetic Particle Precipitation (EPP). EPP refers to highly energetic electrons and protons that are accelerated into the atmosphere, mainly in the polar regions. These energetic particles ionize the Earth's atmosphere throughout the mesosphere and lower thermosphere (MLT), which leads to production nitric oxides (NO_x) and hydrogen oxides (HO_x) which both are associated with an ozone (O_3) loss. Auroral electrons (<30 keV) can penetrate to altitudes around the mesopause, while the medium energy electrons (MEE) (30 keV–1000 keV) can reach the lower mesosphere. During polar winter, NO_x is quite long-lived, and it can be transported deeper into the atmosphere, all the way to the stratosphere. In theory, EPP induced change in stratospheric O_3 has the potential to modulate the strength of the stratospheric polar vortex and the polar regional surface temperatures. The dynamical impact of the mesospheric O_3 reduction is, however, unclear and unresolved in the current hypothesis.

To fill this knowledge gap, the chemistry-climate model WACCM (Whole Atmosphere Community Climate Model) version 6 in the specified dynamics mode for the year 2010 is run with and without MEE ionization rates. How well WACCM reproduces the transport of EPP produced NO_x and the impact of MEE on the mesospheric dynamics in WACCM are the two aspects of the EPP impact addressed in this thesis. In comparison to observations from SOFIE (Solar Occultation For Ice Experiment) on board the AIM (Aeronomy of Ice in the Mesosphere) satellite, the production and transport of NO_x in WACCM6 is studied. The comparison show that WACCM overestimates NO_x during quiet geomagnetic times, while the response to geomagnetic activity is underestimated. Nevertheless, the speed of the estimated transport correspond well with observations. The second aspect of this study, demonstrate, however, that the MEE ionization rates themselves has the potential of modulating the residual circulation and NO_x transport. During the Northern Hemispheric winter, even weak ionization rates is able to modulate the mesospheric signal of a sudden stratospheric warming event. The MEE modulation is found in both the NO_x volume mixing ratio (VMR), temperatures and zonal winds. Furthermore, the induced changes manifest themselves also in the Southern Hemispheric summer, whereas a potential inter-hemispheric coupling link is discussed.

Contents

Chapter 1

Introduction

Near-Earth space is a stormy place where the magnetosphere is constantly being ripped and teared by the solar wind. Guided by the Earth's magnetic field, part this energy is deposited into the atmosphere as energetic electron precipitation (EEP). EEP is the main source of nitric oxides (NO_x) and hydrogen oxides (HO_x) production in the polar mesosphere and lower thermosphere (MLT) region (called the direct EEP effect [*Randall et al.*, 2007]). Auroral electrons (<30 keV) originate in the plasma sheet and they penetrate to altitudes around the mesopause. Medium energy electrons (MEE) (30 - 1000 keV) originate in the radiation belts, where they are accelerated to higher energies, and can penetrate deeper into the mesosphere. During the polar winter, when there is no sunlight, NO_x can have a long lifetime. The winter polar vortex prevents NO_x from moving to lower latitudes, keeping it away from sunlight, and the circulation can transport NO_x all the way down to the stratosphere (called the indirect EEP effect [*Randall et al.*, 2007]). There, NO_x can destroy ozone (O_3), and the changes in O_3 can impact both temperature and dynamics of the atmosphere [*Seppälä et al.*, 2013]. On the other hand, HO_x is very reactive and has a lifetime of only hours in the mesosphere [*Brasseur and Solomon*, 2005], thus it is only able to destroy O_3 locally.

Observational studies show that the relation between geomagnetic activity (proxy for EEP) and stratospheric/ tropospheric winter circulation has been significant at least since the 1960s [*Maliniemi et al.*, 2016]. In order to verify the complex chain of reactions coupling space to the atmospheric dynamics, chemistry-climate models need to correctly simulate NO_x and HO_x, as they can change the ozone abundance, leading to temperature and dynamical changes. This means that it needs to correctly simulate the production due to EEP, as well as the transport processes that can bring NO_x to the stratosphere. It has been a long outstanding challenge to understand the NO deficit found in the chemistry climate model Whole Atmosphere Community Climate Model (WACCM) in the lower mesosphere and upper stratosphere. Weak vertical transport, inadequate input or a missing chemical

scheme has been suggested as possible pitfalls. Recently, a more sophisticated chemical scheme has been added to WACCM *Verronen et al.* [2016], to improve on the NO_x production and losses in the mesosphere. Uncertainties remain, however, in respect to the models' capability of reproducing the NO density throughout the lower thermosphere. A comparison of WACCM with eight years of satellite observations in the southern hemisphere found a displacement in maximum altitude, overestimated background levels, and underestimated short-term variability of NO density in the thermosphere [*Hendrickx et al.*, 2018]. The latter is supported by an event study, which found high pre-storm background levels, but a relative weak EEP response in the NO densities in the lower thermosphere compared to the observations [*Smith-Johnsen et al.*, 2018]. Hence, it is unclear if the deficiencies in the model are related to inadequate input, missing chemical schemes in the lower thermosphere, and/or inadequate transport. Either way, the amount of NO that reaches the stratosphere will be underestimated and a potential EPP-effect deeper into the atmosphere will be underrated in the current model setup. Hence, the initial objectives of this thesis were:

- Identify the discrepancy between standard WACCM model runs and observations of NO in the Mesosphere ans Lower Thermosphere region.
- Identify to which degree the discrepancy is due to inadequate transport and/or the estimated EEP ionization rates.

To do so, two WACCM runs in the specified dynamics mode for the year 2010 were studied, one including both auroral and MEE forcing (the MEE run) and one including only auroral forcing (the noMEE run). Hence, when compared to observations from SOFIE (Solar Occultation For Ice Experiment) on board the AIM (Aeronomy of Ice in the Mesosphere) satellite, the production and transport of NO_x in WACCM6 can be disentangled. By serendipity, however, it became evident that it was not only the NO_x production that separated the two runs. The temperature and winds, and thereby the transport, changed in the model when including the MEE ionization rates. This added a new dimension to the current study expanding the objectives to also include the fundamental question:

- Can MEE change the temperature and the dynamics of the atmosphere without changing the stratosphere?

Chapter 2

Theory

The main focus of this thesis is to understand how energetic particle precipitation affects the Earth's atmosphere. That is why the goal of this chapter is to understand how ionized particles from the Sun are ejected into Space and after interacting with the Earth's magnetosphere, they precipitate into the atmosphere.

Section 2 gives a brief description of the Sun's structure, and the different phenomena that eject solar plasma into space. Section 2.2 describes the Earth's magnetosphere and how it interacts with the persistent plasma from the Sun. In addition, it introduces reconnection and the Dungey cycle, leading to trapped particles in the magnetosphere, and how they can be accelerated and scattered, eventually precipitating into the Earth's atmosphere. In section 2.3, different geomagnetic indices, that quantify geomagnetic activity, are introduced, as well as different types of geomagnetic storms. The Earth's atmospheric structure, the general circulation and the response to particle precipitation is explained in section 2.4. Finally, section 2.5 gives an overview of the recent publications relevant for the results and discussion of this thesis.

2.1 The Sun

The Sun is classified as a main-sequence star, or more specifically a yellow dwarf star, located at the center of our Solar System. It is the main source of energy for the Earth, and it drives the seasons, weather and climate, as well as near-Earth plasma dynamics leading to the visible auroras at polar latitudes. Most of the energy coming from the Sun is primarily solar radiation. Furthermore, energy in the form of ionized solar plasma can also reach near-Earth space and enter the Earth's atmosphere.

2.1.1 Solar structure

The Sun is a hot ball of gas, and its mass consists of about 70.6% hydrogen, 27.4% helium, and the rest is a combination of heavier elements. Since it is not a rigid body, it does not rotate at the same rate everywhere, leading to a latitudinal dependency on the rotation rate, which is about 25 days at the equator, and about 36 days at the poles. The solar magnetic field is produced by the solar dynamo, resembling a dipole magnetic field. Due to the difference in rotation rate, the originally poloidal field lines twist and eventually form toroidal field lines emerging from the surface, as seen in Figure 2.1. As the field becomes more twisted, solar activity increases until it peaks, reaching the solar maximum. After that, the magnetic field relaxes back to a dipole magnetic field with inverted polarity, associated with the solar minimum. This polarity reversal happens approximately every 11 years and is closely linked with the solar activity cycle. When it comes to solar magnetic configuration, a complete solar cycle takes 22 years to go back to its initial state.

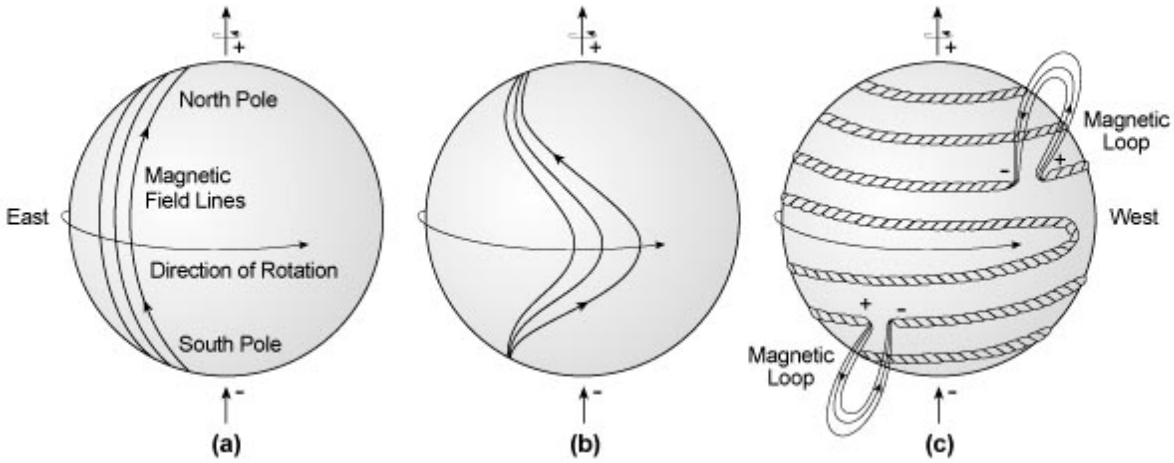

Figure 2.1: Sun's magnetic field evolution due to the difference in rotation rate. Adapted from [*Babcock*, 1961]

The internal structure of the Sun is divided in three layers. The deepest layer of the Sun is the core, which lies at the center of the Sun itself. This layer extends approximately to $0.25R_\odot$, where $R_\odot$ refers to the Solar radii, and it is where the energy production of the Sun takes place. The second layer is the Radiative zone, which extends from approximately $0.25R_\odot$ to $0.75R_\odot$, and as its name suggests, it is where the energy is carried outward by radiation. As we move away from the core, the temperature of the Sun drops, which results in the plasma being more opaque, and convection is the dominating process of energy transport. This outermost layer, between $0.75R_\odot$ and $1R_\odot$, is called the Convective zone. Inside this layer hot bubbles of plasma are displaced upwards, towards the Sun's surface, while colder plasma sinks.

The Sun's visible surface is called the photosphere. It is a 500km thick region with a temperature of about 5500K, and is the lowest layer of the Sun's atmosphere. Above the photosphere, the chromosphere is found, which is about 2000km thick. In this layer, the temperature increases with altitude, reaching about 25000K at the boundary with the transition region. The transition region is where the plasma is fully ionized, and acts as a junction between the chromosphere and the corona. The corona is the outermost region of the Sun's atmosphere and it consists of hot plasma. Here, the temperature increases with altitude, reaching temperatures of about 2 million K.

2.1.2 The solar wind

The solar wind is a continuous radial outflow of fully ionized plasma from the Sun. It travels at supersonic speeds ranging from around 300 km/s for low speeds, to about 800 km/s for high speeds. It consists mainly of electrons and protons and generally has $T_e \approx 10^5$K and $n_e \approx 5\text{cm}^{-3}$ near the Earth [*Baumjohann and Treumann*, 2012]. It is the result of the corona's high temperature and electron density. At its base, the corona reaches 1 million K, resulting in almost half of the electrons having sufficient thermal velocities to escape the Sun's gravity. This is not the case for the heavier ions, which then creates an electric field that accelerates the protons away from the Sun, causing the solar corona expansion [*Parks*, 2004].

Coronal holes are dark in appearance when looking at the Sun with e.g. an X-ray camera. They are colder and less dense than the surrounding corona because particles are escaping at a greater rate. These regions are usually located at high latitudes, where the Sun's magnetic field lines are open and the solar wind can escape with relatively fast velocities. During the declining phase of the solar cycle, they can extend to mid and low latitudes.

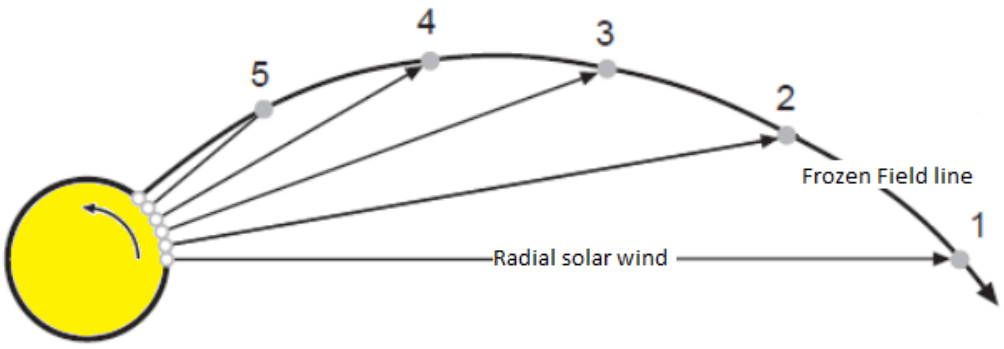

Figure 2.2: Interplanetary magnetic field evolution. The solar wind is ejected radially from the sun, and as the sun rotates it follows a spiral.

The interplanetary magnetic field (IMF) is the Sun's magnetic field, carried out by the solar wind plasma. The IMF is frozen into the solar wind plasma, which is flowing radially outward from the Sun. However, due to the solar rotation, the IMF will form an Archimedean spiral as seen in Figure 2.2, known as the Parker spiral. Near the Earth, the IMF has an average magnitude of about 5 nT [*Parker*, 1958].

Sunspots are regions in the photosphere that appear darker than their surroundings. These regions occur as a consequence of the twisting of the Sun's magnetic field, when concentrations of magnetic field flux emerge from the surface of the Sun, as seen in Figure 2.1. This prevents the convection of plasma on the surface, resulting in a colder plasma in comparison to the surroundings. Sunspots usually come in pairs with opposite polarity. During solar minimum, they form at relatively high latitudes, but they move towards lower latitudes when approaching to solar maximum, as can be seen in Figure 2.3. After solar maximum, the Sun's magnetic field will have inverted its polarity.

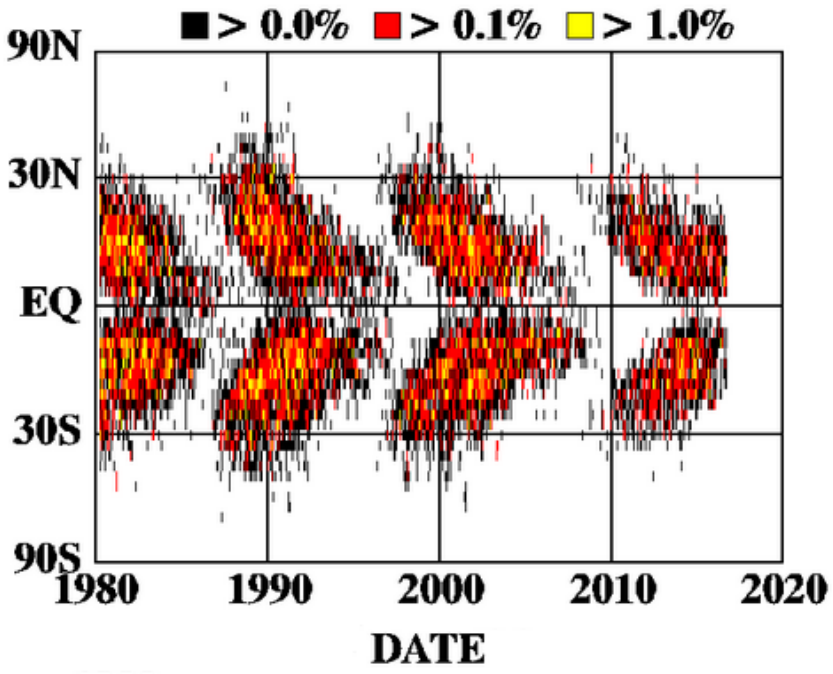

Figure 2.3: Butterfly diagram. The color is the percentage of sunspot area in equal area latitude strips.(NASA)

Sunspots are closely connected to solar flares and coronal mass ejections(CME). These events occur when the magnetic field lines, of opposite polarity, that emerge from the surface reconnect, releasing electromagnetic energy and radiation into space, known as solar flares. Sometimes, the magnetic reconnection also accelerates plasma from the corona into space, hence the name coronal mass ejection. Since these two phenomena are connected to sunspots, they peak during solar maximum.

2.2 The magnetosphere

The Earth's magnetic field is generated in its iron core. The near-Earth space where the geomagnetic field is stronger than the IMF is called the magnetosphere. The shape of the magnetosphere and its internal dynamics is the result of the interaction between the geomagnetic field and the IMF.

2.2.1 Geomagnetic field

The source of the Earth's magnetic field is thought to be the electric currents generated by the rotation motion of convection currents of molten iron, an electrically conducting fluid, inside the Earth's core. This is the dynamo theory, and it proposes a mechanism, in which rotating celestial bodies can maintain a magnetic field over astronomical scales. This field can be approximated to a dipole field, whose axis is tilted 9.69^{o} from the rotational axis of the Earth [*Laundal and Richmond*, 2017]. Additionally, the magnetic North pole does not correspond with the geographic North pole, and so, the Earth's dipole points southwards, meaning that the magnetic field lines in the magnetosphere will point northward, as represented in Figure 2.4.

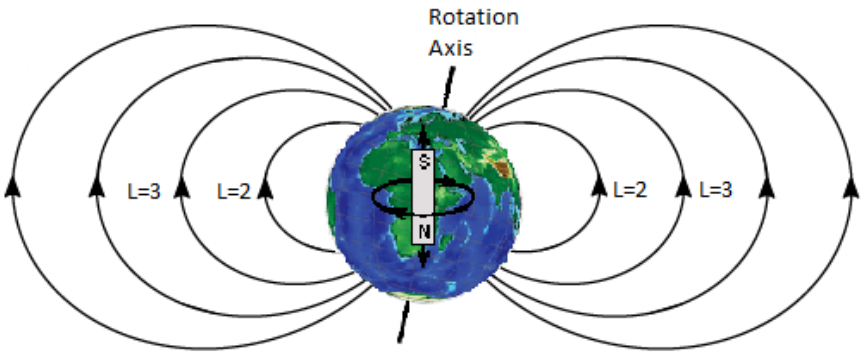

Figure 2.4: Representation of the earth's magnetic dipole field, with magnetic field lines corresponding to different L values. The magnetic axis are tilted from the rotational axis.

The dipole model approximation can be used within a few Earth radius, while more complex models are needed for a more accurate representation of the geomagnetic field if we move further from the Earth. The dipole model is also useful to define the L-shell, which includes a set of magnetic field lines that cross the Earth's magnetic equator at a certain L value, as can be seen in Figure 2.4. This L value is the distance, in Earth radius, from the center of the Earth, to the point where the field line crosses the equator, defined as:

$$L = \frac{r}{R_E} \tag{2.1}$$

where r is the radial distance from the center of the Earth and R_E is the Earth's radii.

2.2.2 Interaction with the Solar Wind

When the supersonic and superalfvénic solar wind encounters the Earth's magnetosphere, which acts as a shield against it, a bow shock is formed as the solar wind is slowed down to subsonic speeds and is deflected around the magnetosphere by this collisionless shock wave. The denser and more compressed solar wind between the bow shock and the magnetosphere is called magnetosheath, and the boundary between the magnetosphere and the magnetosheath is the magnetopause.

When encountering the Earth's magnetic field, electrons and protons will be deflected in opposite directions, giving rise to the current at the magnetopause. The kinetic pressure of the solar wind shapes the Earth's dipole magnetic field, compressing it on the day-side and stretching it on the night-side, where it forms a long magnetotail. A schematic illustration can be seen in Figure 2.5

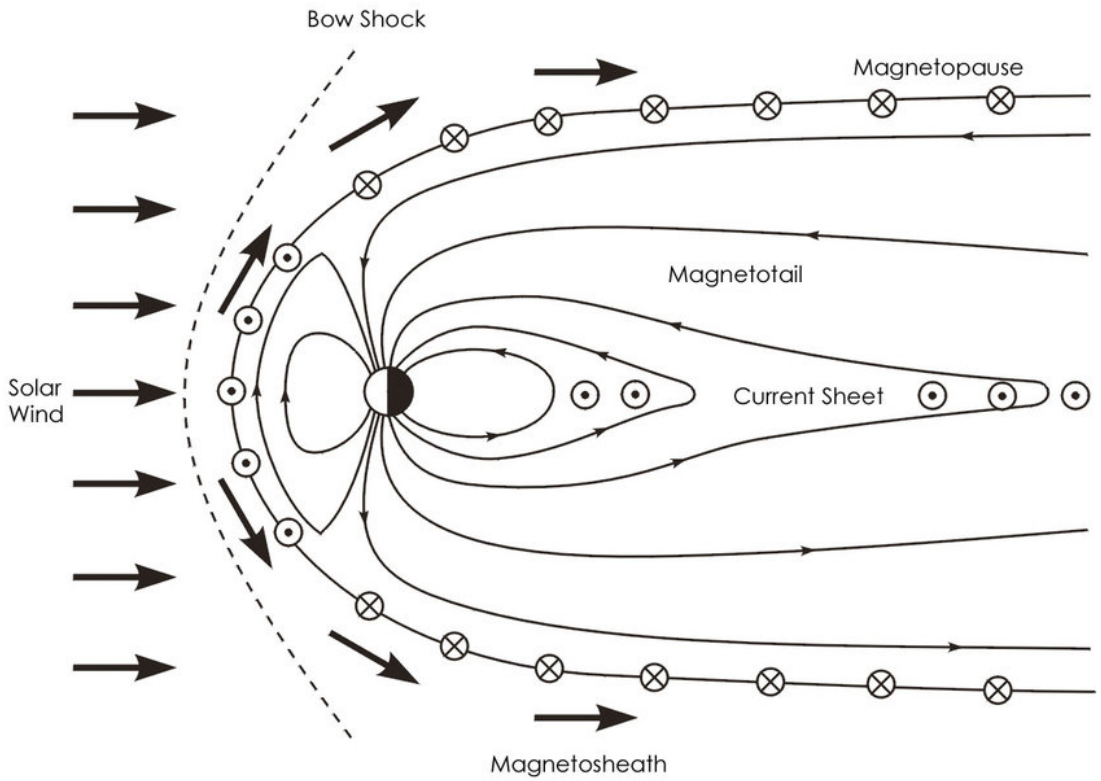

Figure 2.5: Illustration of Bow shock, the solar wind flow forming the magnetosheath, the magnetopause and the currents that arise. [*Hughes*, 1995]

Even if the magnetosphere deflects the solar wind, plasma from the solar wind can enter into the Earth's magnetosphere thanks to magnetic reconnection. The Dungey Cycle is a model that explains how the magnetic reconnection between the IMF from the solar wind and the Earth's magnetic field drives the magnetosphere dynamics [*Dungey*, 1961].

As we know, all magnetic field lines are closed, but in order to easily explain this cycle, the concept of "open" and "closed" magnetic field lines will be conveniently introduced. Magnetic field lines with both ends at the Earth's surface will be referred to as closed magnetic field lines. On the other hand, magnetic field lines with one end at the Earth's surface but the other end connected to the solar wind will be referred to as open magnetic field lines.

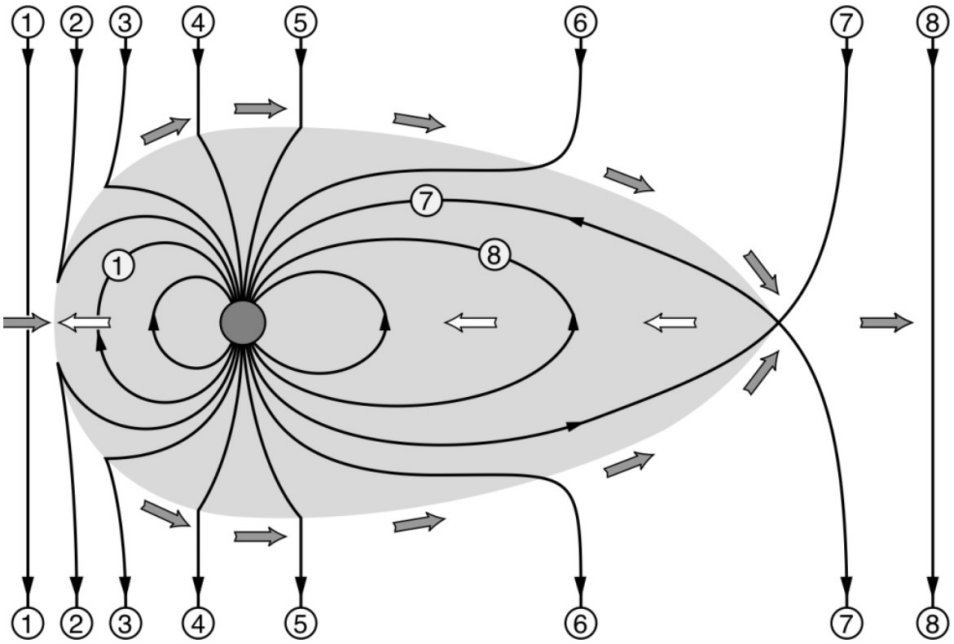

Figure 2.6: Illustration of the Dungey Cycle for a southward IMF. The magnetic field lines from the IMF reconnect with the geomagnetic field lines, and they are displaced to the night-side due to the solar wind motion. In the nightside the field lines stretch until oppositely directed field lines meet and reconnect in the tail. After merging, the field lines return towards the Earth, eventually being transported back to the dayside. This process allows plasma from the solar wind to enter the earth's magnetic field. Taken from [*Baumjohann and Treumann*, 2012].

The Dungey cycle, which is represented in Figure 2.6, begins when two oppositely directed field lines reconnect. The Earth's magnetic field is northward directed, so this happens when an IMF with a southward component reaches the Earth (number 1 in Figure 2.6). When this happens, the field line from the IMF and the closed Earth's field line reconnect, forming two open field lines, with one end on Earth and the other on the solar wind (number 2).

The footpoints on Earth of these field lines lie in the ionosphere, where conductivity is high, impeding field line mobility due to friction with the ions. Outside of the magnetosphere, the field lines are still frozen into the solar wind, so they are dragged towards the night-side by the motion of the solar wind, which results in the bending of the field line. This is transported downwards by Alfvén waves until the magnetic force due to the

curvature exceeds the frictional force that the field lines are subjected to in the ionosphere, and they move towards the nightside (from 3 to 6).

The magnetic field lines keep stretching on the nightside as the solar wind moves away from the Earth until oppositely directed lines meet again in the tail and reconnect (7). After merging together again, the IMF and the geomagnetic field will be detached, and plasma from the solar wind will enter into the magnetosphere (8). The curvature at the reconnection point of the closed field lines will force the field lines towards the Earth, relaxing them back into a dipole structure, and eventually transporting them back to the dayside, completing the cycle.

This is a simple model used to describe magnetic reconnection in the Earth's magnetosphere. In reality, only a component of the incoming IMF is southward and antiparallel with the Earth's magnetic field, and is referred to as the reconnection component. However, the IMF doesn't need to be southward to be able to reconnect, as reconnection is still possible even for northward IMF at high latitudes.

2.2.3 Energetic Particle Precipitation

The plasma in the magnetosphere is grouped in different plasma regions. The plasmasphere is a region around Earth that contains cold, dense plasma. The outer boundary is called the plasmapause, and its position lies between 4 and 7 R_E, depending on the geomagnetic activity [*Chappell et al.*, 1970].

The radiation belts, also known as Van Allen belts, consists of two separated regions of energetic electrons and protons trapped in the Earth's magnetic field. The inner radiation belt extends from 1.2 to 3 R_E [*Ganushkina et al.*, 2011] and is embedded in the plasmasphere. It contains mostly protons with energies ranging from 10 to 50 MeV. The outer radiation belt extends from around 3 to 7 R_E [*Ganushkina et al.*, 2011], and it consists mainly of high energy electrons, ranging from 0.1 MeV to energies over 10 MeV.

The particles trapped in the radiation belts are governed by the Earth's magnetic field. The particles gyrate around the magnetic field lines, bounce along the field line between both hemispheres, and drift around the Earth. Electrons and protons drift in opposite directions around the Earth, electrons drift eastward while protons drift westward, giving rise to the ring current. The bounce motion of the particles is due to the converging magnetic field when they get closer to the poles, causing the particles to convert all of their velocity perpendicular to the magnetic field, and eventually mirroring.

If the mirroring point is sufficiently close to the Earth, particles can reach the upper atmosphere, where they collide with neutrals and deposit their energy. The loss cone contains all the pitch angles at which particles reach the upper atmosphere and can be lost, which is referred to as particle precipitation. The pitch angle is the angle formed between the particle velocity and the magnetic field line that it bounces along, as seen in Figure 2.7. Magnetospheric processes such as convection or wave-particle interaction

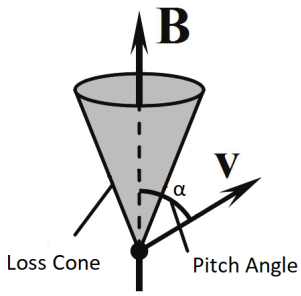

Figure 2.7: Representation of the losscone in grey. The pitch angle α is the angle between the magnetic field B and the velocity of the particle v

can change the particle's energy and pitch angles, resulting in a continuous diffusion of particles into the loss cone.

The plasma sheet is a sheet-like region where the magnetotail plasma is concentrated. Low energy particles from this region precipitate at high latitudes into the Earth's atmosphere and produce aurora. The different regions are represented in Figure 2.8. Auroral electrons precipitate from the plasma sheet, and their energy typically ranges between 1 and 30 KeV. Medium Energy Electrons (MEE) precipitate from the radiation belt, where they are accelerated to higher energies, ranging from 30 to 1000 KeV.

2.3 Geomagnetic storms

A geomagnetic storm is a major disturbance of the Earth's magnetic field that happens when the reconnection rate with the IMF in the solar wind is strong, enhancing the energy transfer from the solar wind to the Earth's magnetosphere. This happens when the IMF has a southward component and the enhanced reconnection rate leads to more particles being injected into the ring current, making it stronger. The magnetic field resulting from the ring current is opposite to the Earth's magnetic field, so when it gets stronger, it effectively weakens the Earth's magnetic field.

The drivers of geomagnetic storms are usually Coronal Mass Ejections(CME), associated with sunspots, or Corotating Interacting Regions (CIR), associated with coronal holes. As mentioned in Section 2.1.2, CMEs occur more often during solar maximum, while the coronal holes can extend to mid and low solar latitudes in the declining phase of the solar cycle where the associated high-speed solar wind will point towards the Earth. Storms are generally divided into three phases. The first phase is the onset, when increased solar wind dynamic pressure compresses the day-side magnetosphere. Following is the main phase, when particles are injected into the inner magnetosphere. Finally, there is a recovery face, when the injection of particles slows, and the ring current falls back to its normal levels [*Tsurutani et al.*, 2006]. To quantify geomagnetic storms strength, different indices are used.

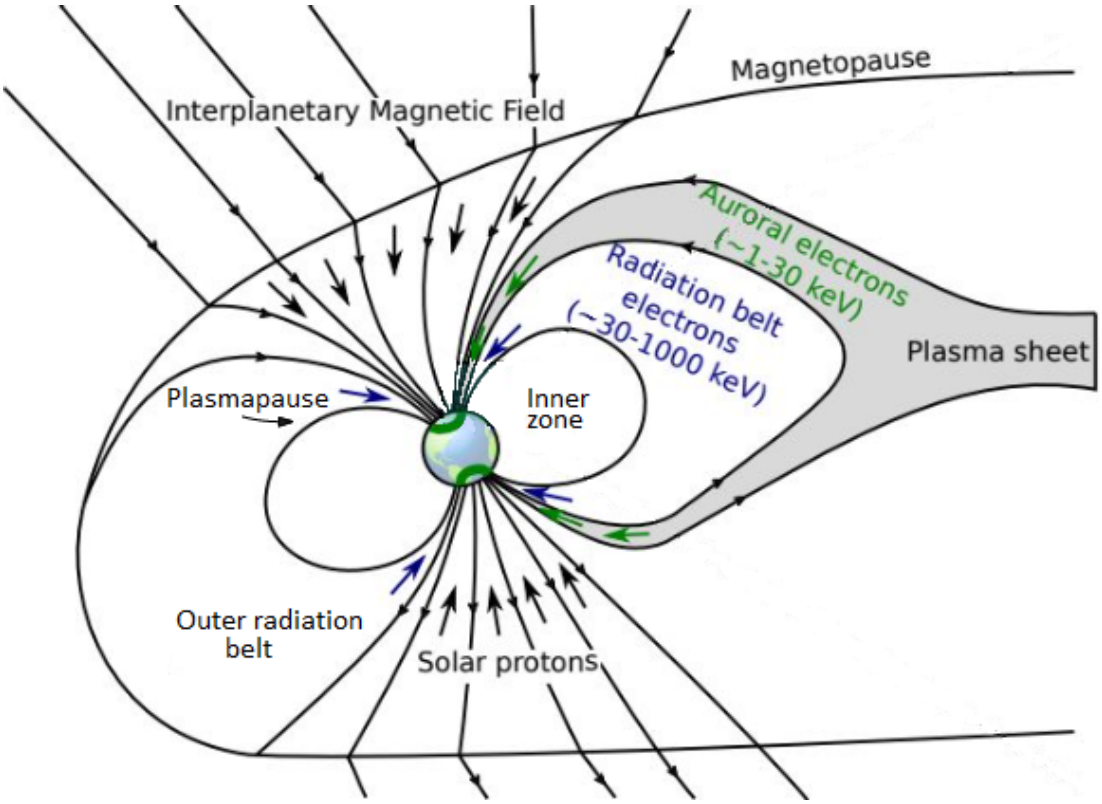

Figure 2.8: Representation of the plasma regions in the magnetosphere, modified from [*Thorne*, 1980]. Solar protons (black arrows) can reach the polar cap directly from the solar wind. Auroral electrons (green arrows) can precipitate at high latitudes from the plasma sheet, while radiation belt electrons (blue arrows), or MEE, precipitate from the radiation belt.

Substorms are brief disturbances of the Earth's magnetic field that cause energy from the magnetotail to be released into the high latitude ionosphere. They last for only a few hours and are only observable in high latitude regions. Usually, several substorms are observed during a geomagnetic storm, but they also happen relatively frequently during quiet (non-storm) periods.

2.3.1 Kp-index and Ap-index

The Kp-index is the global geomagnetic activity index and it is derived from three-hour measurements of the horizontal Earth's magnetic field component, taken by different observatories located in sub-auroral latitudes. Each station reports a K-index depending on the geomagnetic activity measured at its location.

The K-index is a three-hour quasi-logarithmic index of the geomagnetic activity compared to a calm day at the observatory location. The global Kp-index is obtained with an algorithm that combines the K-index values from the different stations together, and its value ranges from 0, meaning there's little geomagnetic activity, to 9, corresponding to

high geomagnetic activity [*Bartels et al.*, 1939].

The Ap-index is the daily average level of geomagnetic activity. Since the K-index is non-linear, every three-hour K-index is converted back into a linear scale, the a-index. The Ap-index is then calculated from the average of the 8 a-index values that are retrieved each day, and higher Ap values correspond to higher geomagnetic activity.

2.3.2 AE-index

The AE-index is the Auroral Electrojet index, and it is a measure of the global electrojet activity in the auroral zone. The AE-index is derived from observations, done by 13 observatories located in the auroral zone in the Northern Hemisphere (geomagnetic latitudes 60-70^{o}), of geomagnetic variations in the horizontal component of Earth's magnetic field. The data from each station is normalized for each month by averaging all the data from the station on the five most quiet days. This base value is then subtracted from the one-minute data obtained at the station and among the data from all the stations, at each given time (UT), the largest and smallest values are selected [*Kyoto-University*, 2021].

The AU-index is defined by the largest value selected, and expresses the strongest current intensity of the eastward auroral electrojet. The AL-index is defined by the smallest value selected, and expresses the same for the westward auroral electrojet. The AE-index represents the overall activity of the electrojets and is defined by the difference between the AU and AL index: $AE = AU - AL$ [*Kyoto-University*, 2021].

2.4 The atmosphere

The atmosphere is the layer formed by the gases surrounding the Earth, that are trapped by its gravity. These gases are not stationary in the atmosphere. The atmospheric circulation will transport energy as well as chemical constituents. When particles precipitate into the atmosphere, these gases will be ionized, dissociated, and excited, triggering different chemical reactions.

2.4.1 Structure

The Earth's atmosphere is divided in a series of layers defined by their temperature profile. A schematic view of these layers can be seen in Figure 2.9

The temperature and location of the atmospheric layers vary with latitude and season. The lowest layer is called *troposphere*, where temperature decreases with altitude until reaching a minimum, the *tropopause*. At the equator it is located near 16 km, while in polar regions it only reaches around 8 km. Above the tropopause, temperature increases with altitude due to the UV absorption by the ozone layer that is located approximately between 10 and 50 km. This layer is called the *stratosphere* and it extends until around 50 km, where the maximum temperature is reached in the *stratopause*. At higher altitudes, in

the *mesosphere*, the temperature decreases again up to 85-100 km, where the mesopause is found. The layer above the *mesopause* is called the *thermosphere*, and the temperature increases rapidly with altitude due to X-rays and UV radiation from the Sun being absorbed. The amount of absorbed energy varies with Sun activity, and so does the top of the thermosphere [*Brasseur and Solomon*, 2005].

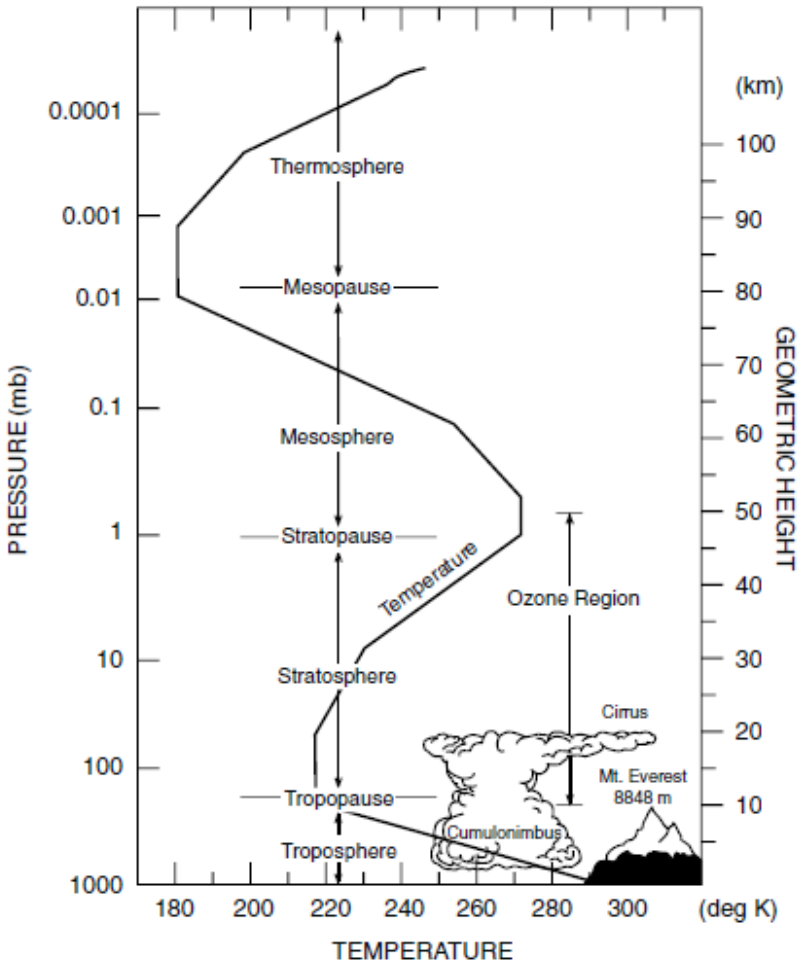

Figure 2.9: Schematic representation of the atmospheric layers, defined by the thermal structure of the atmosphere. Modified from [*Brasseur and Solomon*, 2005]

This classification of layers is not unique, and other characteristics, such as the chemical composition, could be used to define the atmosphere's layers. Up to 100 km, the major constituents of the atmosphere are N_2 and O_2, which make up about 80% and 20% of the density, respectively. This means that the molecular weight varies little with altitude, consisting of a homogeneous mixture, and hence the name *homosphere*. Above this altitude, the gases separate by molecular diffusion, leading to a separation of heavier and lighter molecules, which form the *heterosphere*.

The middle atmosphere is a region extending from approximately 10 to 100km, and it includes the stratosphere and the mesosphere. For this work, the focus is on the transition region between the middle atmosphere and the lower thermosphere. This region is known as the MLT, which stands for mesosphere and lower thermosphere. This region can be directly impacted by particle precipitation, and the interaction of its dynamics with chemistry is of particular interest.

2.4.2 General Circulation

The radiative heating rate at the summer mesopause is large, but the temperatures observed are lower than the ones in the winter hemisphere, where there is less radiative heating. Furthermore, the tropical tropopause is much colder than the one at high latitudes, even if there's no big variation in radiative heating. These cases show that the atmosphere is not in radiative balance, and that also other dynamical processes are important for the temperature profile of the atmosphere [*Brasseur and Solomon*, 2005].

When an air parcel is displaced adiabatically, it expands and cools when being displaced upwards, while when being displaced downwards, it compresses and heats. This implies that upward motions are characteristic of both the summer mesopause and the tropical tropopause [*Brasseur and Solomon*, 2005].

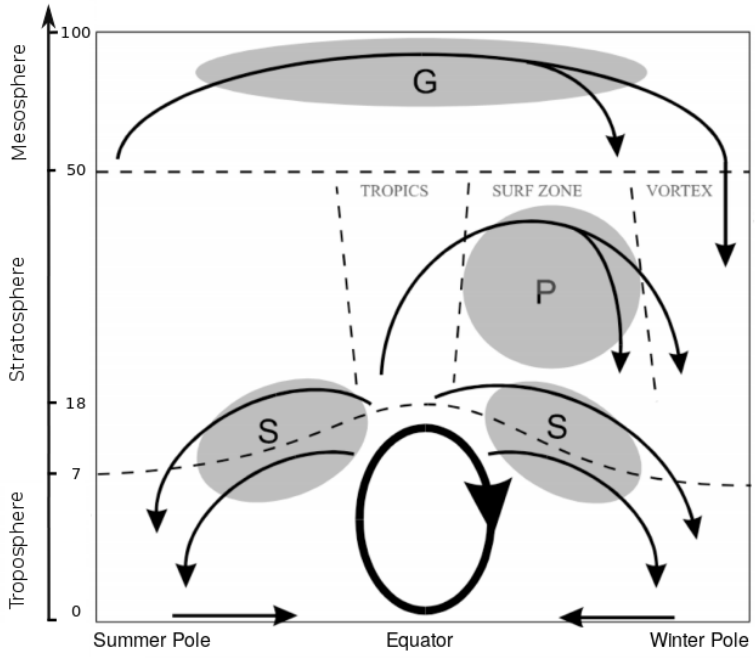

Figure 2.10: Brewer-Dobson circulation. Summer to winter pole circulation in the mesosphere, driven by gravity waves. Stratospheric circulation by planetary waves. Circulation around the tropopause by synoptic waves. Modified from [*Plumb*, 2002]

Additionally, the highest concentrations of ozone are found near the poles despite that most of it is produced in the tropical stratosphere. This stratospheric transport from the tropics to the pole was discovered after observations on the distribution of ozone and water vapor by *Brewer* [1949] and *Dobson et al.* [1929].

A schematic diagram of the atmospheric circulation, also called the Brewer-Dobson circulation, is shown in Figure 2.10. The circulation is characterized by the transport of air parcels from the summer pole to the winter pole in the mesosphere, from the tropics to the winter pole in the stratosphere, and from the tropics to the poles around the tropopause. These circulations are driven by planetary waves, gravity waves and synoptic waves, respectively.

One of the most important features in atmospheric dynamics is the zonal wind, which is the longitudinal component of the wind. These winds are mostly derived from the observed temperature profile of the atmosphere, although local values can be obtained by radar measurements. The longitudinally averaged zonal wind profile is presented in Figure 2.11, where the northern hemisphere corresponds to the winter hemisphere, since the data is from January.

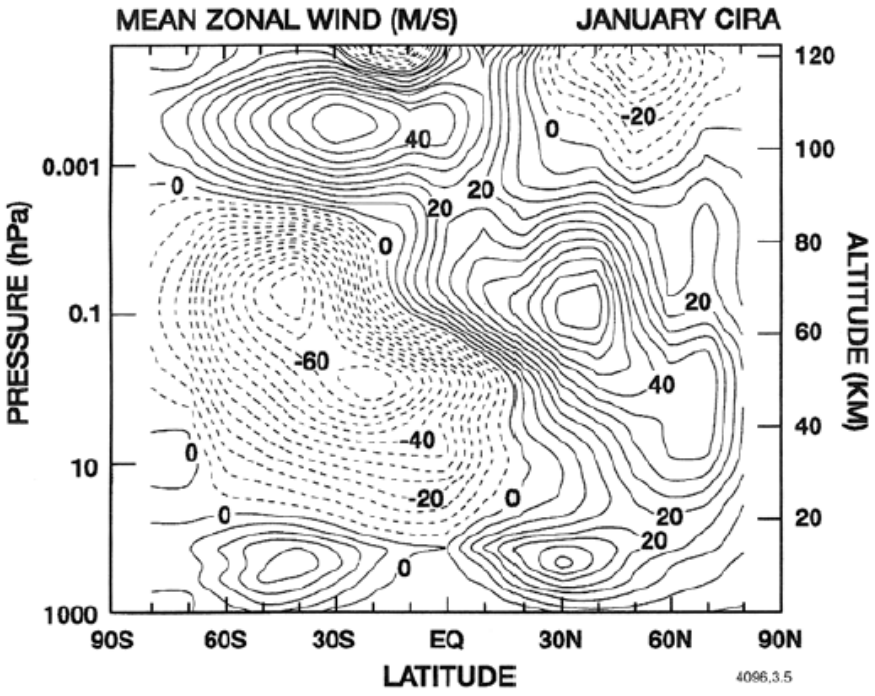

Figure 2.11: Longitudinally averaged zonal wind(m/s) for January, from the surface up to 120 km. Negative winds blow westwards, while positive blow eastwards. Stratospheric and lower mesospheric winds are eastward during winter and westward during summer. Above 90 km, a reversal of the winds is observed [*Brasseur and Solomon*, 2005].

In the lower stratosphere, the zonal circulation is an extension of the tropospheric flow, characterized by two jets centered at 30-40^{o} latitude, and presenting wind speeds almost twice as large in winter than in summer. These jets of strong winds form the *polar vortex*, which isolates the polar region from low latitudes. For most parts of the stratosphere and the mesosphere, the mean zonal winds are mainly eastward in the winter hemisphere, since the temperature gradient increases from pole to equator, and eastward in the summer hemisphere, since temperature gradient increases from the equator to the pole. These winds reach their maximum speed at around 70 km, and above this altitude, the zonal flow becomes less intense up until 90 km, where its direction reverses [*Brasseur and Solomon*, 2005].

2.4.3 Forcing from below: Waves

Atmospheric waves are propagating disturbances through the atmosphere, whose acceleration is balanced by a restoring force. This force can be buoyancy in a stably stratified fluid, which opposes vertical displacements, or the Coriolis force, which opposes horizontal displacements.

Gravity waves are oscillations with horizontal wavelengths typically ranging from 10 to 1000 km, that arise in a stably stratified fluid when air parcels are being displaced vertically. These waves are produced by airflow over mountains, called orographic gravity waves, or by other non-orographic sources, such as thunderstorms or instabilities [*Brasseur and Solomon*, 2005].

The propagation of these waves through the atmosphere is governed by the wind distribution and thermal structure of the atmosphere, which varies with season. When the phase speed of the propagating wave is equal to the zonal wind speed, the wave is absorbed by the zonal wind. Hence, the zonal wind acts as a filter for the propagating gravity waves. Due to the strong filtering in the stratosphere, gravity waves in the mesosphere are mainly westward in the winter hemisphere, since the zonal wind in the winter stratosphere is eastward. Similarly, the gravity waves are eastward in the summer hemisphere, since the zonal wind in the stratosphere is westward. Figure 2.12 shows the filtering by the zonal wind in both winter and summer.

For kinetic energy to remain constant:

$$E = \frac{1}{2}\rho_o |v'(z)|^2 \tag{2.2}$$

and the amplitude of the waves have to grow as:

$$|v'(z)| = A[\rho_o(z)]^{-\frac{1}{2}} \tag{2.3}$$

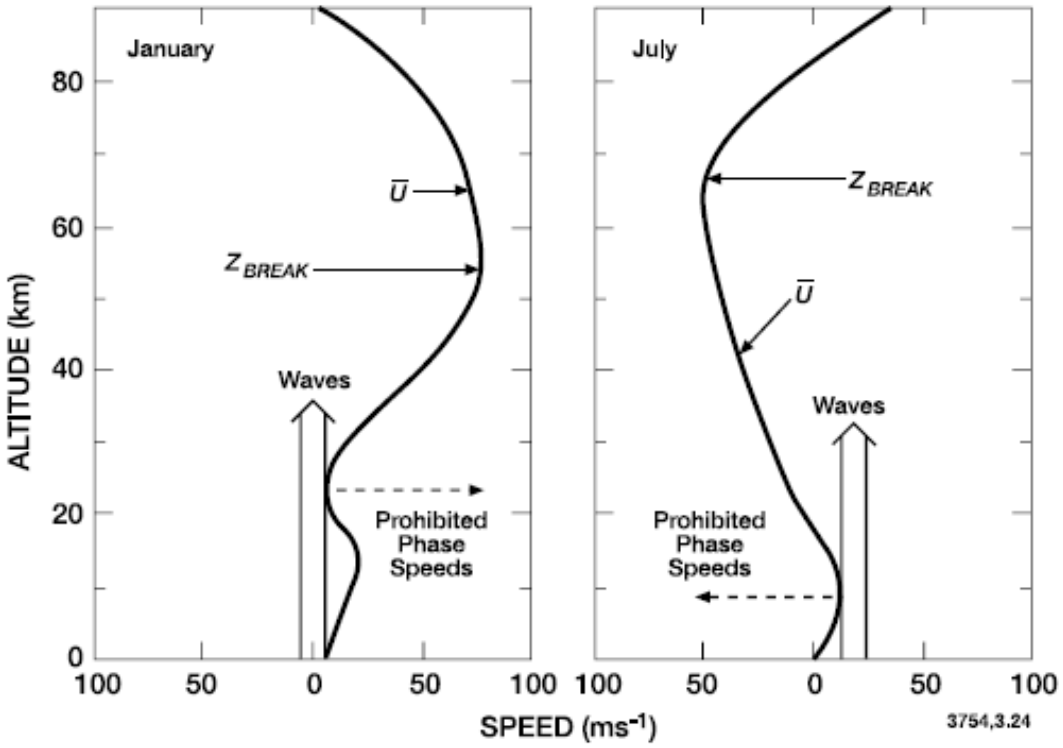

Figure 2.12: Altitde profiles of the mean zonal winds for winter(left) and summer(right). The zonal winds act as a filter for gravity waves propagation, and the permitted phase speeds are shown, as well as the breaking level. [*Brasseur and Solomon*, 2005]

In the absence of dissipation, the amplitude of the wave grows as the inverse square of the density, which means that the wave amplitude grows exponentially with height. At some point, the wave amplitude will grow so large that its temperature perturbation will produce a superadiabatic lapse rate and become convectively unstable, and the wave is said to break.

When the waves break, they deposit their energy and momentum into the background wind, and together with the Coriolis force, they drive the mesospheric circulation from the summer pole to the winter pole seen in figure 2.10. The mass balance will subsequently cause upwelling of air masses in the summer hemisphere, and downwelling in the winter hemisphere.

2.4.4 Forcing from above: Particle precipitation

As seen in section 2.2.3, auroral electrons precipitate into the Earth's atmosphere from the plasma sheet where they are accelerated to energies typically ranging from approximately 1 to 30 keV. They enter the atmosphere at latitudes corresponding to the auroral oval, and deposit their energy in the lower thermosphere and upper mesosphere, at altitudes between 85 and 130 km, as seen in figure 2.13.

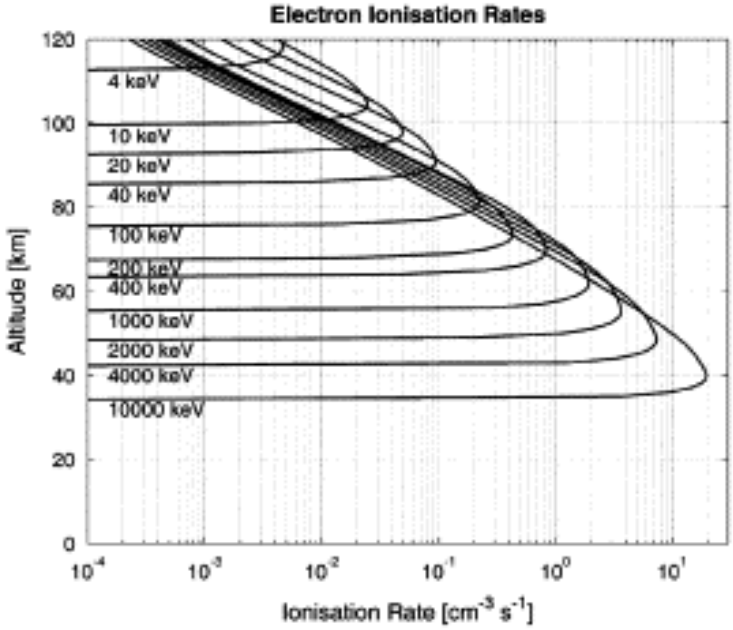

Figure 2.13: Ionization rate for monoenergetic beams of electrons between 4 keV and 10 MeV in the atmosphere. [*Turunen et al.*, 2009]

MEE precipitate into the Earth's atmosphere from the radiation belts. During geomagnetic storms they can be accelerated up to MeV and precipitate in the mesosphere and upper stratosphere [*Turunen et al.*, 2009], to altitudes between 55 and 85 km, as shown in figure 2.13. Electron precipitation of energies >300 keV peak during the main phase of the storm, while for energies > 1MeV it peaks during the recovery phase, having an effect on the atmospheric chemistry with a delay of some days from the peak of the storm [*Horne et al.*, 2009].

When these electrons reach the atmosphere, they increase the local production of HO_x and NO_x, which is considered the direct effect of EEP [*Randall et al.*, 2007]. EEP alters the O_3 density, since both NO_x and HO_x destroy ozone through catalytic reactions. HO_x molecules have a short lifetime corresponding to hours in the middle mesosphere [*Brasseur and Solomon*, 2005]. NO_x has a lifetime of about one day under sunlit conditions in the mesosphere [*Brasseur and Solomon*, 2005]. In the polar winter darkness, however, the molecules can exist for months. Hence, NO_x enhancements due to EEP can be observed at lower altitudes, far from where the molecules were initially formed. This is due to the wave-driven circulation, causing NO_x to be transported across the atmospheric layers from e.g. the lower thermosphere to the upper stratosphere. This is often referred to as the dynamical indirect effect of EEP[*Randall et al.*, 2007].

During winter, NO_x will be transported by the downward residual transport to altitudes where it can deplete stratospheric ozone. The ozone depletion can potentially change the local temperature, leading to changes in the zonal winds. As a result, the filtering of

the waves propagating vertically is also altered by the new dynamics, which will change the momentum deposition of the waves. Changing the momentum deposition affects the dynamics of the atmosphere which again will feed back into temperatures, winds, wave propagation, and transport of the atmospheric gasses.

[*Seppälä et al.*, 2013] suggested that geomagnetic activity (measured by the Ap index) can drive changes in NH winter stratospheric dynamics, specifically the strengthening of the polar vortex, resulting in more planetary waves being refracted equatorward when the geomagnetic activity is higher than average. Especially under solar maximum, the enhanced UV and ozone interaction warms the upper stratosphere, enhancing the equator to pole temperature gradient and strengthening the polar vortex. Under the westerly phase of the Quasi-Biennial Oscillation (QBO-W), the vortex is also strengthened, enabling more planetary waves to be refracted equatorward.

An increase of EEP is associated with a stratospheric polar temperature response in winter, corresponding to a warming of the upper polar stratosphere and a cooling signal below. *Baumgaertner et al.* [2011] suggested that this warming signal is a response to ozone depletion (decrease in ozone radiative cooling), and the cooling arises from dynamical heating due to slowing down of the Brewer-Dobson circulation, which is associated with less upward Eliassen-Palm (EP) flux and more waves reflected toward the equator.

2.4.5 Transport of species

The transport of atmospheric species can be caused by three separate processes: advection, eddy diffusion and molecular diffusion. Outside the polar vortex, NO_x is destroyed by sunlight, so horizontal transport can not explain the large enhancements in stratospheric NO_x.

An important process for the transport of trace species is the large-scale net motion of air, the displacement of air parcels. To be effective, the timescale for the displacement must be short compared to the timescale for the air parcel to mix with the surrounding atmosphere or for the tracer to be created or destroyed by photochemical processes. The transformed Eulerian mean (TEM) circulation approximates the net air parcel displacements, and describes the bulk motion of large air masses [*Smith et al.*, 2011].

Diffusive processes can change the concentration of trace species within an air parcel. Eddy diffusion represents turbulent diffusion generated by wave breaking. Molecular diffusion acts to diffuse any trace species with a significant vertical gradient in mixing ratio and to impose a net vertical drift based on molecular mass [*Smith et al.*, 2011].

NO_x is frequently produced by EEP in the polar lower thermosphere, and transported across the mesopause. Based on the Brewer-Dobson circulation presented in Section 2.4.2, there is no clear advective downward transport from the thermosphere to the mesosphere on the daily mean circulation. Still, molecular diffusion and advection are the dominant processes in the lower thermosphere when it comes to NO_x transport. This is because

the intradiurnal variability is high, probably caused by atmospheric tides, and it strongly affects the downward transport of NO_x. The impact of molecular diffusion decreases rapidly with height, while the impact of advection increases [*Meraner and Schmidt*, 2016].

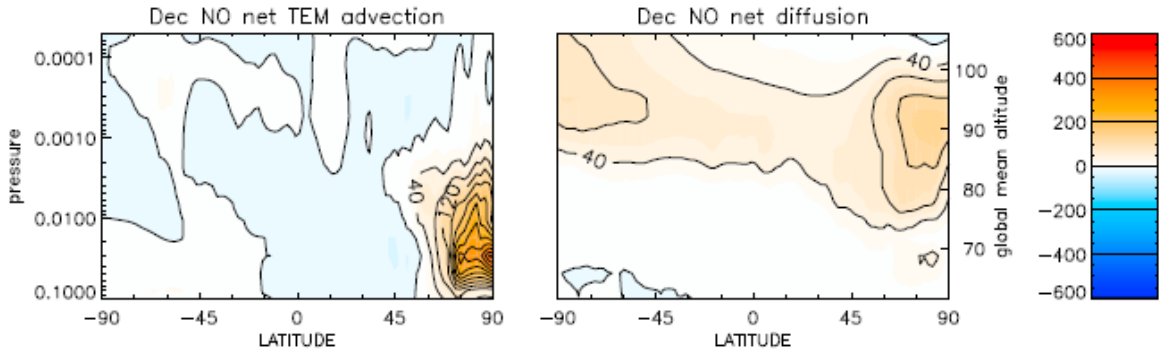

Figure 2.14: Mixing ratio tendencies for NO during December due to the TEM circulation (left panel) and to the sum of eddy and molecular diffusion (right panel). Units are change in mixing ratio as a percent of the long term mean mixing ratio over all latitudes and months at each pressure level, with contour intervals of 40% [*Smith et al.*, 2011].

2.4.6 Sudden Stratospheric Warmings

In winter and spring, large temperature increases are occasionally observed in the lower stratosphere, within few days, accompanied by cooling in the upper stratosphere and mesosphere. Under these circumstances, the polar night jet decreases in strength, and sometimes it even reverses. This sudden warming in the lower stratosphere is known as sudden stratospheric warmings (SSWs) [*Labitzke*, 1981].

SSWs are caused by strong planetary wave drag decelerating the flow. As a result of the decelerated flow, the propagation of planetary waves is modified and will produce a stronger deceleration of the zonal wind at a lower altitude. An SSW is considered to be major when the zonal wind reverses in the polar stratosphere [*Limpasuvan et al.*, 2012].

During SSWs the stratosphere warms and the mesosphere cools. As a consequence, the stratopause displaces to a lower altitude and, after breaking down, it reforms at an anomalously high altitude [*Meraner et al.*, 2016]. When this happens, the polar vortex breaks down, and NO_x is no longer confined in the polar night, leading to NO_x in the stratosphere being destroyed by photo dissociation. Furthermore, around the time of the SSW, the circulation presents upwelling in the high latitude lower mesosphere, preventing NO_x from being transported down. After the SSW, the polar vortex reforms often resulting in extreme winter conditions for the associated downwelling. During this dynamically active period, the mesospheric downwelling is stronger by about a factor of two, and the poleward flow at high latitudes occurs at lower altitudes [*Smith et al.*, 2011]. This leads to an unusually strong descent of NO_x into the stratosphere. The NO_x amount which descends

from the thermosphere to the stratosphere can be 50 times higher than average [*Randall et al.*, 2009]. A comparison between the climatology after an SSW and the climatology of a normal winter can be seen in figure 2.15.

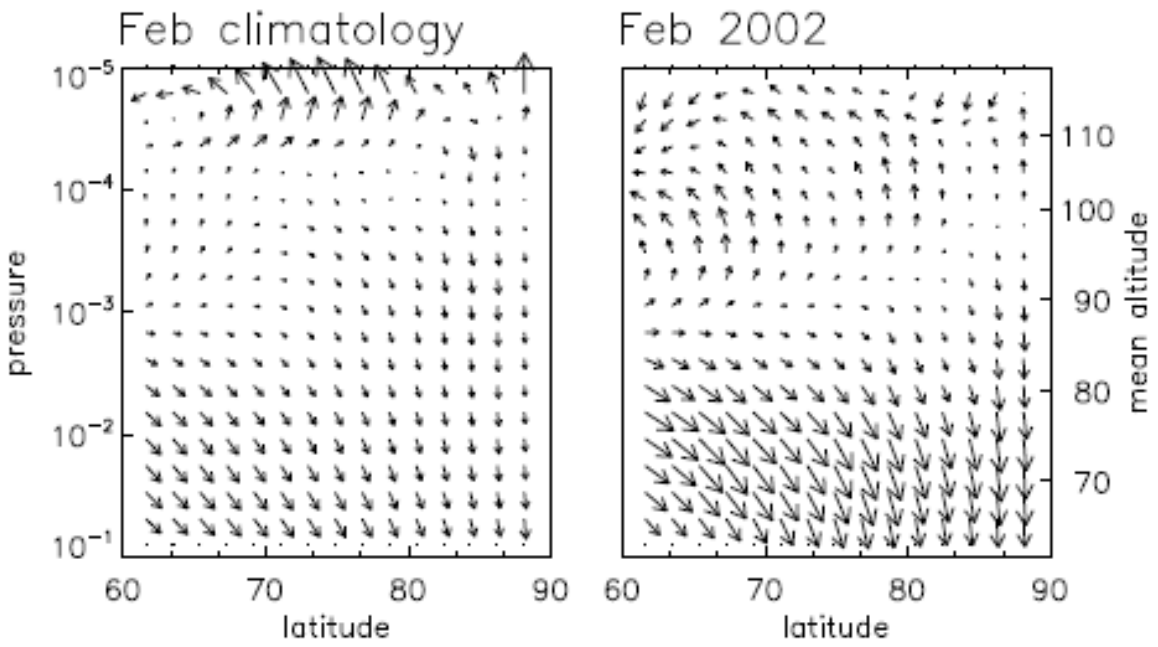

Figure 2.15: Mean climatologal circulation for the month of February in the northern hemisphere (left) and after a sudden stratospheric warming(right). [*Smith et al.*, 2011]

The earlier an SSW occurs in winter, the more NO_x is transported to the stratosphere, because of two main reasons: 1) The earlier in the winter season it occurs, the more negative the vertical wind is following the event. 2) The earlier it occurs in the season, the longer the NO_x descends, and the less NO_x is mixed laterally (due to the breakdown of the polar vortex) to latitudes where it gets destroyed by sunlight [*Holt et al.*, 2013]. Furthermore, the transport of NO_x depends not only on the timing of the SSW but also on the timing of NO_x production by EEP, and the disturbed dynamics at that time and location.

On average, the EEP indirect effect is larger in winters with an SSW event than in years with no event. However, the EEP indirect effect in dynamically calm winters can exceed that of years with a January, February or March SSW event. In winters with a major SSW event occurring late in the winter season, the normal descend is interrupted, and the enhanced descent after the event is too small and too late in the season to make up for it, while the steady NO_x descent is not interrupted in dynamically calm years [*Holt et al.*, 2013].

With weaker gravity wave sources, i.e. smaller amplitudes, more NO_x is transported after SSW to the mesosphere and the elevated stratopause descends more slowly to its

climatological altitude. A deposition of momentum at a higher altitude extends the downwelling branch of the meridional circulation to a higher level and the descent of NO_x is enhanced [*Meraner et al.*, 2016].

2.4.7 Interhemispheric coupling

Interhemispheric coupling manifests as a correlation between dynamical activity in the winter stratosphere and in the polar upper mesosphere of the summer hemisphere. Anomalously warmer temperatures around the polar summer mesopause coincide with anomalously warmer temperatures in the polar winter stratosphere. Modelling studies by *Becker et al.* [2004] and *Becker and Fritts* [2006] support that changes in gravity wave activity in the summer hemisphere play a role in the interhemispheric link.

Karlsson and Becker [2016] argue that climatological hemispheric differences in the mean temperatures of the northern and southern summer mesosphere can be attributed to differences in the average state of the winter mesosphere due to differences in gravity wave activity there.

The vertical coupling within the winter stratosphere and mesosphere is also relevant. Strong wave forcing in the stratosphere leads to poleward flow and warming of the polar stratosphere, which is accompanied by cooling in the upper stratosphere and mesosphere. The warming and cooling are associated with downwelling and upwelling, respectively, from the circulation driven by wave breaking.

Figure 2.16 shows the correlation between the EP flux divergence, averaged over 60-70^{o}N and 3-0.3 hPA, and temperature, the transformed Eulerian-mean meridional wind (w*), the EP flux divergence (delF) and the gravity wave drag (GW) in the winter stratosphere, for the period from 1st December to 28th February. The EP flux divergence accounts for planetary wave drag, being more negative for perturbed conditions, and wave forcing in the winter high latitude is followed closely by temperature perturbations extending across the equator to about 60^{o} or more in the summer hemisphere [*Smith et al.*, 2020].

The lags in figure 2.16 indicate the days by which the EP flux divergence in the winter stratosphere leads to the respective quantity differences at the respective point. For the temperature, the correlation maximizes for lags of 2-4 days. Dynamical activity in the winter hemisphere leads to perturbations in the mesosphere that are opposite to those in the stratosphere, and figure 2.16 shows negative correlations for the winter stratosphere, while positive for the winter mesosphere. Since the EP flux divergence is negative, this translate in a warmer winter stratosphere and colder winter mesosphere. Correlations of opposite sign extend to the summer hemisphere, giving a colder summer stratosphere and a warmer mesosphere.

Temperature perturbations persist for weeks following perturbations in EP flux divergence, even after the wave forcing perturbations decay. The types of wave events that

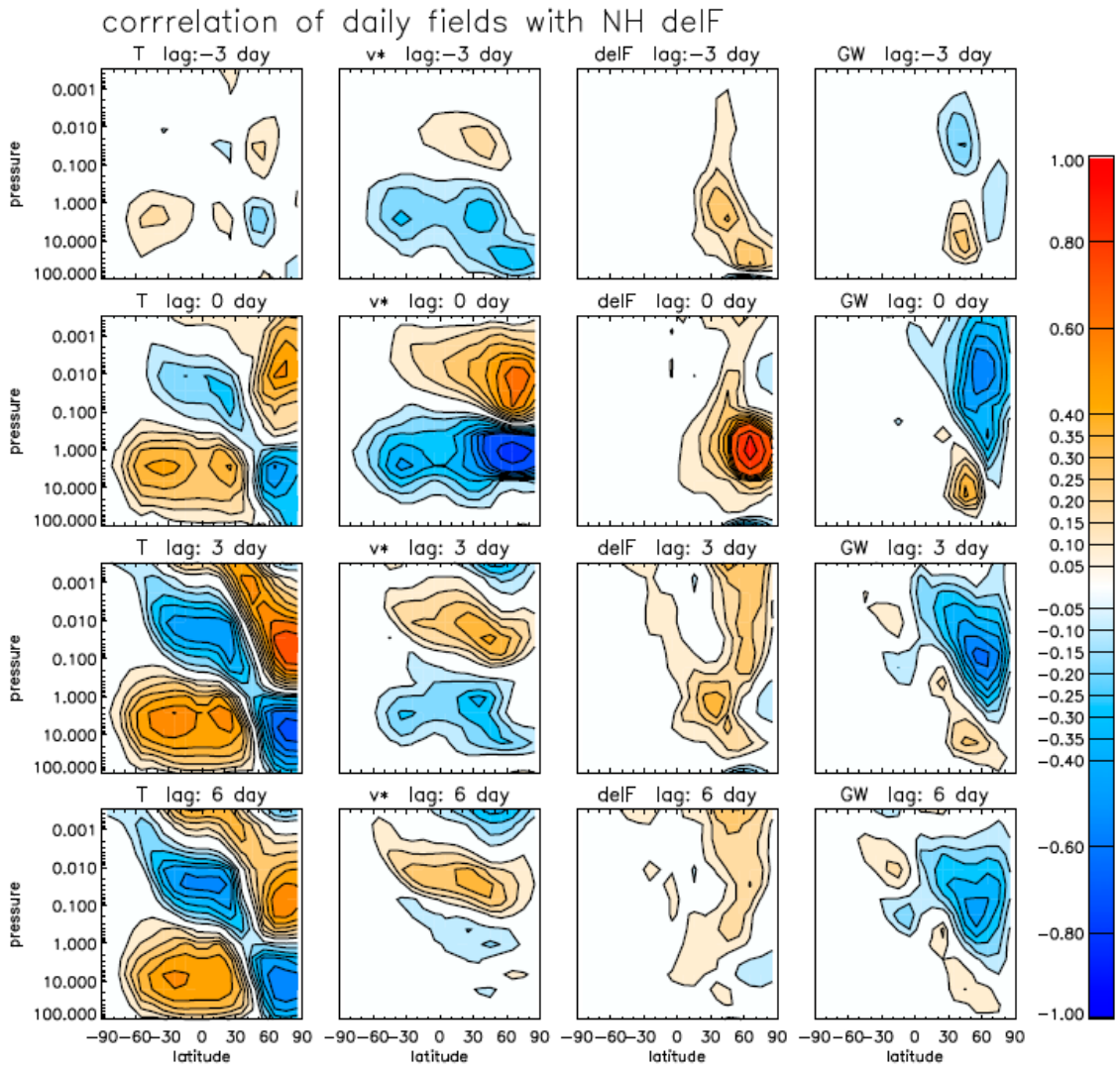

Figure 2.16: Lag correlations between the EP flux divergence, averaged over 60-70^{o}N and 3-0.3hPA, in the winter stratosphere and four dynamical variables for the period from 1st December to 28th February. The dynamical variables are temperature (T), the transformed Eulerian-mean meridional wind (w*), the EP flux divergence (delF) and the gravity wave drag (GW) [*Smith et al.*, 2020].

lead to a global circulation and temperature response are associated with global temperature perturbations that persist for tens of days. Temperature amplitudes in the summer hemisphere range from 1-3 K on average, reaching 5 K for the largest events.

For the TEM meridional wind, the correlation maximizes for a lag of zero days. In the mesosphere, the EP flux divergence and the TEM meridional wind are correlated, which translates into a weakening of the meridional circulation there, while in the stratosphere there's a strengthening. These perturbations propagate well into high latitudes of the

Southern hemisphere.

The gravity wave drag is strengthened in the mesosphere, but as shown in Figure 2.16 this drag is only in the Northern hemisphere and it does not propagate into the southern hemisphere.

The mechanism first proposed by *Karlsson et al.* [2009] to explain how interhemispheric coupling happens, has later been simplified by the findings by *Smith et al.* [2020]:

1. EP flux divergence from planetary waves causes large perturbations to the circulation, temperature and zonal wind of the winter stratosphere.
2. Filtering by the perturbed zonal wind affects the propagation of planetary and gravity waves. This changes the energy deposition when waves dissipate, leading to perturbations of opposite sign in the winter middle mesosphere.
3. Mass balance gives residual circulation cells extending across the equator to high latitudes of the summer hemisphere. Gravity wave perturbations are associated with the coupling between the stratosphere and mesosphere in the winter midlatitudes, however, they don't seem to play a role in extending the response to the summer hemisphere in WACCM6.

Based on the rapid response of the meridional circulation to perturbations in the winter middle atmosphere, together with the weak response of waves everywhere but the midlatitude winter, the climatological interhemispheric coupling in WACCM is driven by a mean circulation response: summer response is due to the circulation induced to restore zonal-mean balance to the atmosphere. This interpretation is consistent with winter perturbations that can extend to about 60^{o} into the summer stratosphere and farther toward the summer pole in the mesosphere [*Smith et al.*, 2020].

2.5 Recent Research

[*Hendrickx et al.*, 2018]: Study WACCM4, without D-region chemistry and MEE, in comparison to NO observations from SOFIE in the southern hemisphere. General features of NO are well represented by WACCM when validated by SOFIE. Above the mesopause, WACCM NO is a factor 2 higher in concentration during quiet times. The maximum NO production is consistently 5 km higher in WACCM than SOFIE, resulting in a NO reservoir too high up in WACCM. The descent rate (at 80-100 km) is however the same in WACCM and SOFIE, but the impact of transport is stronger in SOFIE than in WACCM, where too little is transported down.

[*Smith-Johnsen et al.*, 2018]: Study both WACCM4 and WACCM4-D, with and without MEE, in comparison to NO observations by SOFIE. MEE and D-region chemistry are both needed, but still not enough. In the thermosphere, the WACCM background NO is too high during quiet times and and too low in response to geomagnetic activity. At 90 km almost nothing is produced during the storm. The direct production seems okay in the

mesosphere, but the indirect effect suffers from underestimate above. Transport across mesopause is not enough.

[*Smith-Johnsen et al.*, submitted to JGR]: When the amplitude of the non-orographic gravity waves is changed in WACCM, it has an effect in both the temperature of the atmosphere, as well as in the NO transport. Figure 2.17 shows the temperature differences between a run with increased amplitude gravity waves and a control run, as well as the mesopause altitude for both runs, the purple line corresponding to the control run and the orange line corresponding to the increased amplitude run. It can be seen that when the amplitude is increased, the temperatures below the mesopause decrease, while they increase above the mesopause, during the summer. This results in a lower mesopause, which transitions earlier into an altitude typical of winter. The contrary happens on the winter to summer transition, when the mesopause for the increased amplitude run transitions later in time to an altitude typical of summer.

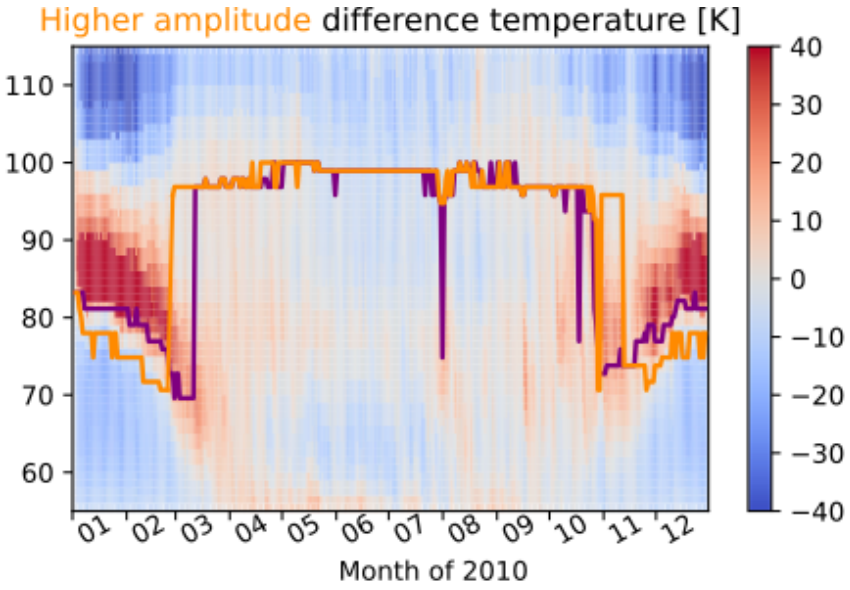

Figure 2.17: Change in temperature [K] due to an increase on the gravity wave amplitude by a factor of 5. The purple line is the mesopause altitude for run with increase wave amplitude, and the purple is the altitude for the control run. Modified from [*Smith-Johnsen et al.*, submitted to JGR]

[*Salminen et al.*, 2020]: During winters when the Quasi-Biennial Oscillation (QBO) phase is easterly, SSWs occur more commonly when AP value is below average (lower EEP) than when it is higher than average. EEP strengthens the polar vortex in the QBO-E phase [*Maliniemi et al.*, 2013], and makes it less vulnerable to planetary waves. On the contrary, low AP values(low EEP) lead to a weaker polar vortex, which is more vulnerable to the increased planetary wave activity of the QBO-E phase, allowing SSW to occur more often.

[*Asikainen et al.*, 2020]: An increase of EEP is associated with a polar vortex enhancement and the corresponding warming(cooling) of the upper(lower) polar stratosphere in winter. The EEP-related response arises from winters when an SSW occurs, while in winters without an SSW event, zonal winds do not have an appreciable response to EEP. The EEP-related signal takes place before the SSW occurs, indicating that the atmospheric conditions preceding the SSW allow the EEP effect to be observed. Stronger planetary wave convergence in the stratosphere, observed during pre-SSW periods, are more likely allowing the EEP responses to be dynamically amplified.

Chapter 3

Methods

To study the transport of NO across the mesopause, both space borne observations and chemistry-climate model simulations are used in this thesis. The Solar Occultation For Ice Experiments (SOFIE) instrument on board the Aeronomy of Ice in the Mesosphere (AIM) satellite is presented in Section 3.1. It provides measurements of the NO composition throughout the mesosphere and lower thermosphere. The Whole Atmosphere Community Climate Model (WACCM) is presented in Section 3.2. Here, the EEP parametrization in WACCM is given special emphasis, as it enables a theoretical study of the impact of MEE precipitation on the mesospheric chemistry and dynamics.

3.1 SOFIE

The AIM satellite was launched on April 25, 2007, and began observations on 14 May, 2007. The aim of the mission is to study the variability in polar mesospheric clouds, by measuring different properties of the mesosphere and lower thermosphere. The satellite is in a near circular 600km sun synchronous orbit with an orbit period of 96 minutes (15 orbits per day). SOFIE is one of the three instruments on board of the AIM satellite [*Russell et al.*, 2009].

SOFIE obtains measurements through solar occultation, which uses the Sun as a source of radiation and measures the change in signal as the Sun rises or sets behind the limb of the atmosphere. This is accomplished by monitoring solar intensity as the satellite enters or exits the Earth's shadow. The measurements are used to retrieve vertical profiles of temperature, five trace gases (NO, CO_2, CH_4, H_2O and O_3) and Polar Mesospheric Clouds (PMC). The measurements are limited to latitudes between 65° and 85°, depending on the year, for both hemispheres [*Russell et al.*, 2009]. Figure 3.1 shows the predicted latitude of SOFIE measurements throughout one year.

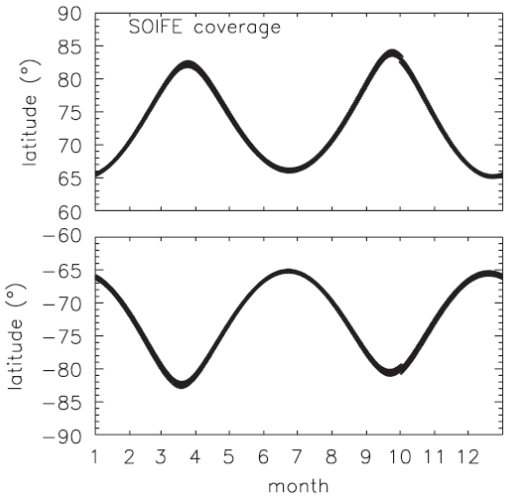

Figure 3.1: Predicted SOFIE measurement latitude. From [*Gordley et al.*, 2009].

SOFIE provides 15 measurements per day in the southern hemisphere, taken during local sunrise, and 15 measurements in the northern hemisphere, taken during local sunset. It measures two times per orbit, one sunset and one sunrise, which means that consecutive sunrises or sunsets are separated by 96 min in time or $\sim 24^{\text{o}}$ in longitude. Because AIM is in a retrograde orbit, SOFIE sunset (sunrise) occur near the time of local sunrise (sunset) [*Gordley et al.*, 2009]. A representation of SOFIE and the location where it takes measurements is shown in figure 3.2.

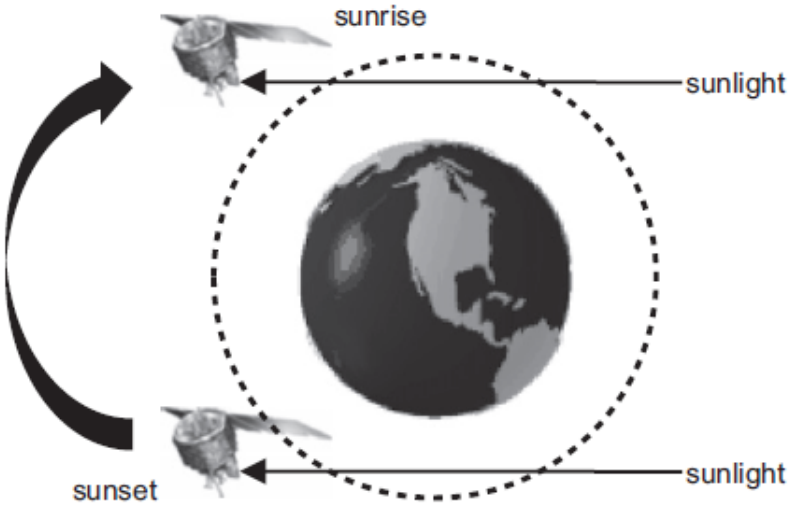

Figure 3.2: Representation of solar occultation measurements, taken during spacecraft sunrise and sunset. From [*Gordley et al.*, 2009].

SOFIE measure the solar occultation in 16 spectral bands, whereas the 5.32μ absorption line is caused by NO molecules. NO measurements are retrieved up to 150 km, with a vertical resolution of 0.2 km. In this study, daily averaged data for both NO volume mixing ratio (VMR) and number density are used. Additionally, a 2km low pass filter is applied to vertically smooth the NO data, as well as an empirical correction to the NO VMR as described by *Gómez-Ramírez et al.* [2013].

3.2 WACCM

WACCM is a global chemistry climate model developed by the National Center for Atmospheric Research (NCAR). WACCM is a configuration of the atmospheric component of the Community Earth System Model (CESM), in which the atmosphere spans from the surface to the lower thermosphere ($\sim$ 140km). WACCM includes interactive chemistry and dynamics integrated in the Community Atmosphere Model (CAM) physics. It includes heterogeneous chemistry that can lead to the development of the ozone hole and ion chemistry to simulate the ionosphere.

WACCM has the same physics as in CAM4 [*Neale et al.*, 2013], but adding essential processes to reproduce the observed mean meridional circulation in the stratosphere and mesosphere, as well as the distribution of minor constituents. These processes include the parametrization of non-orographic waves, molecular diffusion, energetic particle precipitation, non-local thermodynamic equilibrium radiative transfer and a quasi-biennial oscillation is imposed in the winds of the tropical stratosphere [*Hurrell et al.*, 2013].

The model version used in this work is WACCM6, which is a major update of the atmosphere modeling capability of CESM, with enhanced physical, chemical and aerosol parametrizations. WACCM6 extends from the surface up to about 6×10^{-6} hPa ($\sim$ 140km geometric height), with 88 pressure levels and a horizontal resolution of 1^{o} latitude by 1^{o} longitude. WACCM6 is able to reproduce the observed climatology of temperatures, winds and trace constituents in the middle atmosphere, as well as to reproduce stratospheric variability from SSWs [*Gettelman et al.*, 2019].

In the Specified Dynamics (SD) version of WACCM, wind and temperatures are nudged with reanalysis data from NASA Global Modeling and Assimilation Office's Modern-Era Retrospective Analysis for Research and Applications (MERRA) [*Rienecker et al.*, 2011], by the method described in *Kunz et al.* [2011]. The model is nudged from the surface up to $\sim$ 50 km, with a transition region from $\sim$ 50 to $\sim$ 60 km, and is free running above $\sim$ 60 km. The WACCM-SD version is effective for reducing climate noise, reducing biases in the winds and temperatures, as well as reproducing the chemical response to specific events.

WACCM-D is a variant in WACCM which includes extra chemistry in the D-region. The model is based on a simplification of the Sodankylä Ion and Neutral Chemistry (SIC) one dimensional model of the D-region chemistry. WACCM-D includes 20 positive ions

and 21 negative ions, which enable 307 reactions with the aim to reproduce the observed EPP effects in the mesosphere and upper stratosphere [*Verronen et al.*, 2016].

The upper boundary of WACCM6 is the three-dimensional nitric oxide empirical model (NOEM) in the lower thermosphere. The model is based on NO observations between 97.5 km and 150 km done by the Student NO Experiment satellite (SNOE) , which operated in the years 1998-2000. The model is parametrized by the Kp-index and the 10.7cm solar radio flux (F10.7) [*Marsh et al.*, 2004].

WACCM6 includes ionization rate forcing from both auroral electrons and MEE. The ionization rates are parametrized independently for both sources, which makes it possible to include the EEP effect of only one of the sources on the run, as well as to include both sources.

The intensity of auroral electron precipitation (<30 keV) is parametrized by the Kp-index. The Kp index is used to calculate the hemispheric power (HP), which is the total estimated power deposited by the energetic particles. The HP is used to determine the energy-flux distribution of the electrons. The energy spectrum is a Maxwellian distribution with a fixed characteristic energy of 2 keV. From the energy-flux spectrum the ionization rate as a function of altitude is calculated [*Roble et al.*, 1987]. Due to the fixed characteristic energy the resulting ionization rate profile always peaks at around 110km. In general, the energy deposition from auroral electrons is limited to altitudes above 95km. Furthermore, the auroral oval coordinates are scaled by the IMF B_z component.

WACCM6 is the first version of the climate model that by default includes MEE. The MEE precipitation accounts for the radiation belt driven EEP, with energies ranging from 30 to 1000 keV. The model is based on data from the National Oceanic and Atmospheric Administration (NOAA) Polar Orbiting Environmental Satellites (POES) and the empirically described plasmasphere structure. Both are scaled to the geomagnetic Ap-index. The energy-flux spectrum of the precipitating electrons have a time resolution of one day.The ionization rate due to this energy range is typically found at altitudes between 70 and 110 km, and the rates decrease rapidly at altitudes above and below. The peak of ionization rate is at about 90 km, and the lower altitude limit is seen at 55 km, because the electrons with highest energy (1000 keV) can not penetrate further. This model enables simulation of the EEP impact in the middle atmosphere, where EEP is an important source of HO_x and NO_x [*van de Kamp et al.*, 2016].

The precipitation of medium and high energy electrons into the Earth's atmosphere is linked to the level of geomagnetic activity and geomagnetic storms. The most useful measurement of EEP is currently provided by the NOAA POES, with several satellites at different Sun synchronous polar orbits. This model is based on observations from the 0^{o} detector from the Medium Energy Proton/Electron Detector (MEPED) instrument on board POES, and a correction to remove proton contamination is applied. All electron precipitation fluxes are binned as a function of their L value, with a resolution of 0.5 and a 3h resolution, for all magnetic local times together. The L shell range used is 2.5<L<6, and

it covers the outer radiation belt and the dynamical change of the plasmapause. For each bin, the median electron flux is calculated, and daily fluxes are obtained for each energy channel and each L value. Since the plasma pause responds quickly to a geomagnetic storm, but returns slowly to normal after the storm, the geomagnetic Ap-index is integrated over a certain period of time before the time of interest. Since the model is zonally averaged, it has no information on longitudinal variability. Also, the low flux measurements are subject to noise, causing overestimations for low flux levels [*van de Kamp et al.*, 2016].

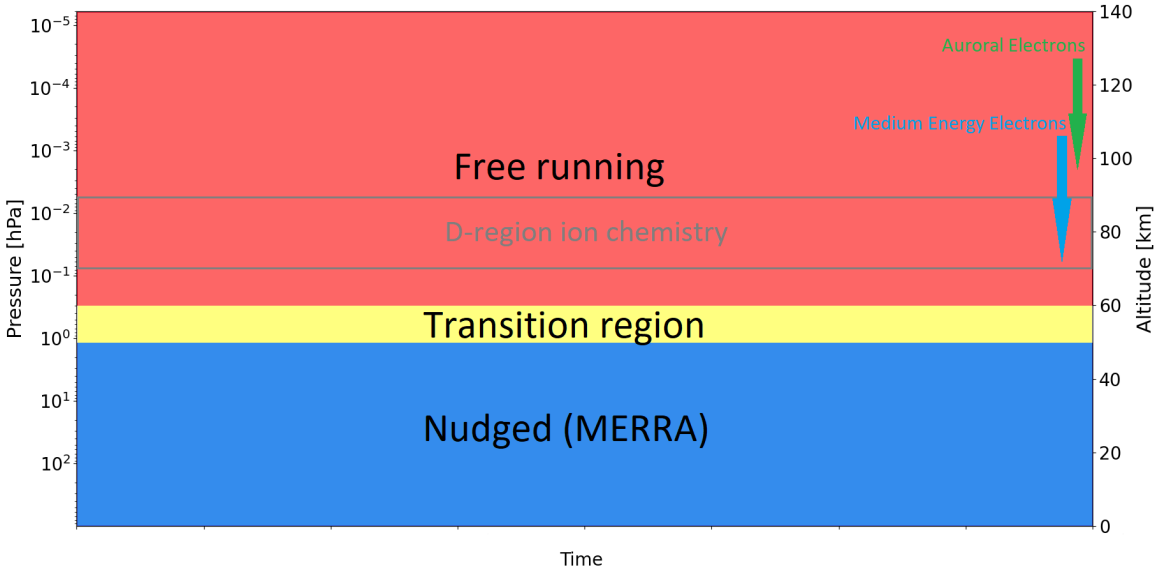

Figure 3.3: Representation of the WACCM6 version used. The arrows represent the typical ionization altitudes for the auroral electrons and MEE.

3.3 The geomagnetic activity in 2010

In order to study WACCM's production in response to geomagnetic events and the associated transport, the periods of interest are those where the geomagnetic events can be clearly differentiated. During solar maximum, the geomagnetic activity is at its highest, and before one event finishes, the next one has already started. The year 2010 corresponds to solar minimum, and for the majority of the year the geomagnetic activity is very low. Thanks to that, geomagnetic events are easily differentiated, making this year a great choice of study.

Furthermore, two previously performed WACCM runs from 2010 were available. This two runs only differ in the EEP forcing. For the first run, the only ionization source for the atmosphere is auroral electrons, while for the second run, both auroral electrons and MEE are used as a ionization source. This makes the runs perfect to disentangle direct production and transport below $\sim$ 100 km, as auroral electrons mainly affect NO_x production above

this altitude. On the downside, however, the wave forcing was not included in the saved parameters and are unavailable for this study.

Chapter 4

Results

In this chapter, an analysis of the data is performed. The data consists of atmospheric observations provided by the SOFIE satellite, as well as two WACCM runs for the year of 2010. The two WACCM runs were performed using the version with improved chemistry in the D-region, WACCM-D, as well as the Specified Dynamics version, WACCM-SD, meaning that the data is nudged with reanalysis data up to 50km, and free-running above. The WACCM data includes one run were the only ionization input is from auroral electrons, and a second one where the ionization input is from both auroral electrons and medium energetic electrons.

Section 4.1 gives a comparison between the observations taken by SOFIE and the data computed in WACCM. The comparison is between the run including MEE and SOFIE, but the run without MEE is used with the aim to disentangle the transport and the direct production of NO by MEE in WACCM. In section 4.2, both WACCM runs are compared, in order to study the differences that arise from including the MEE in the WACCM runs.

4.1 SOFIE vs WACCM

The aim of this section is to compare SOFIE measurements with WACCM simulations, and see how well it is reproducing NO production and transport.

Figure 4.1 shows a line plot including both the NO VMR at 110 km altitude from SOFIE observations and from a WACCM run that includes MEE, for the entire year 2010 in the southern hemisphere. The WACCM NO VMR is taken at the exact same latitude where SOFIE is taking measurements for every day of the year. The time periods selected for this study are highlighted with a red box. This figure shows that the background level of NO is higher in WACCM than in SOFIE, meaning that WACCM is overestimating the level of NO during quiet geomagnetic times. On the contrary, at times where the NO peaks

in SOFIE, due to a geomagnetic event, the response is higher in SOFIE. This is evident in some peaks being higher than in WACCM, even if the NO in SOFIE before the event is at lower levels due to the overestimation of WACCM during quiet times. This means that WACCM is underestimating the NO production during geomagnetic events. Therefore, the NO response to geomagnetic activity in WACCM seems to be inadequately represented.

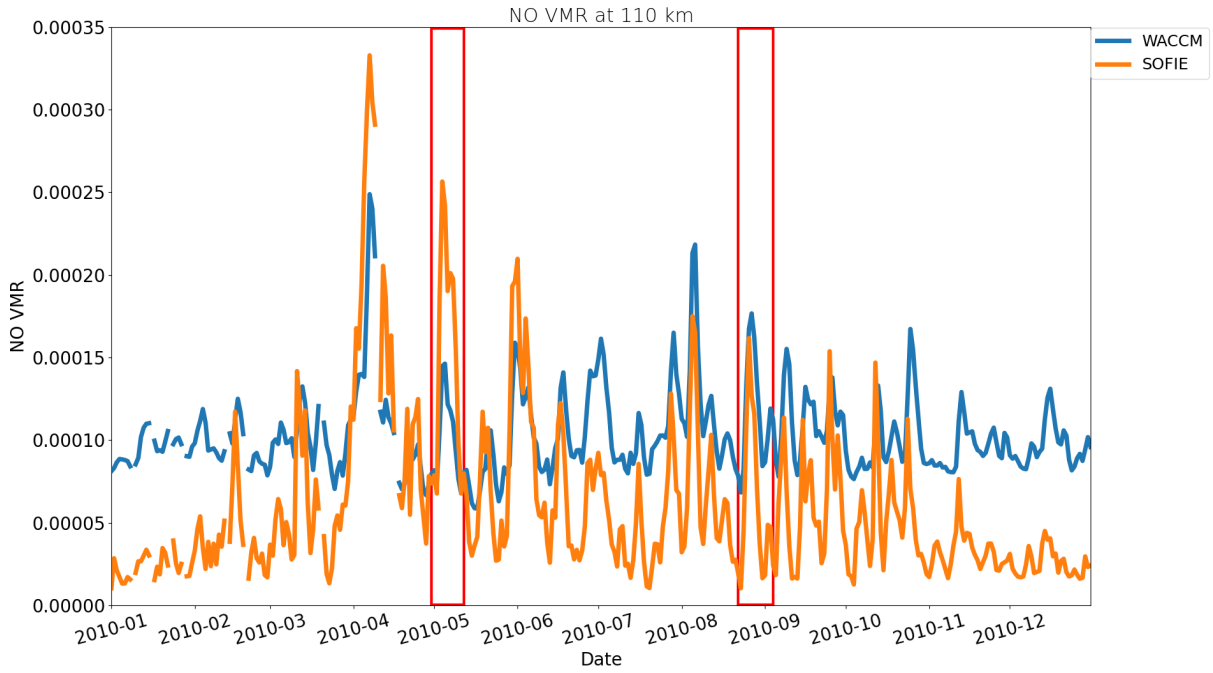

Figure 4.1: NO VMR for SOFIE (orange) and WACCM (blue) at 110 km altitude (lower thermosphere) in the southern hemisphere for the year 2010. The red boxes highlight the two studied periods.

Since the background level of NO in WACCM is too high, it's not straight forward to quantify how much WACCM is underestimating NO production during geomagnetic events. To solve this problem as well as removing the seasonal trend, a 30-day running mean is subtracted from both datasets.

Figure 4.2 shows the same plot as before, but after the 30 day running mean is subtracted from the data, to effectively get rid of the background NO. Without the background NO, the NO levels from observations and WACCM are closer. During quiet times, these levels are about the same, while during geomagnetic events, it can be seen that WACCM is still underestimating the level of NO at the peak for almost all of the events during the year 2010, while it is overestimating the NO level before the event starts and after it ends. Therefore, even after the background NO is subtracted, it looks like NO response to geomagnetic activity is not well reproduced in WACCM.

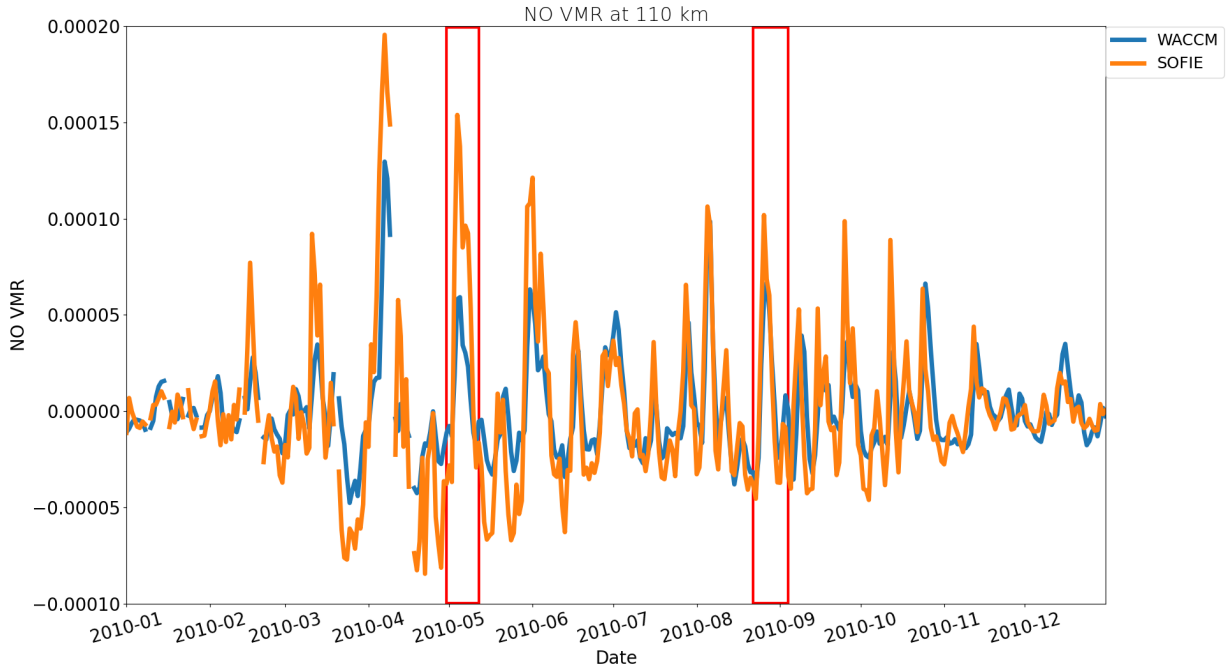

Figure 4.2: NO VMR for SOFIE (orange) and WACCM (blue) in the southern hemisphere for the year 2010. Same as Figure 4.1, but with a 30 day running mean subtracted from both datasets to remove the different background levels between the model and observations. The red boxes highlight the two studied periods.

Overall, the NO response in WACCM to geomagnetic events has a less steep gradient compared to the observations, both when it increases and when it decreases. It can also be seen that for most of the events WACCM is peaking later in time than SOFIE is, suggesting that the response to geomagnetic activity may be delayed in WACCM. At the altitude range of 110km, the NO abundance is mostly governed by production due to the auroral electron ionization, meaning that WACCM is underestimating this production. At lower altitudes, there will be two sources for the NO, local production at the specific altitude, as well as the NO that's being transported from higher altitudes.

Figure 4.3 shows NO VMR at 90km without performing the 30-day running mean. It can be seen that, for the summer months, NO VMR in WACCM is a bit higher than in SOFIE, while in figure 4.1, the background NO during summer was considerably higher in WACCM. This might imply that transport is not accurately represented in WACCM. During times when events happen, it can be seen that before the event start, and after it finishes (at local minimums), the level of NO is roughly the same for both, but during the event itself, WACCM's NO increase is not as steep as it is in SOFIE, showing again that there's also a lack on the direct local production of NO, characteristic of the steep increase at the beginning of events.

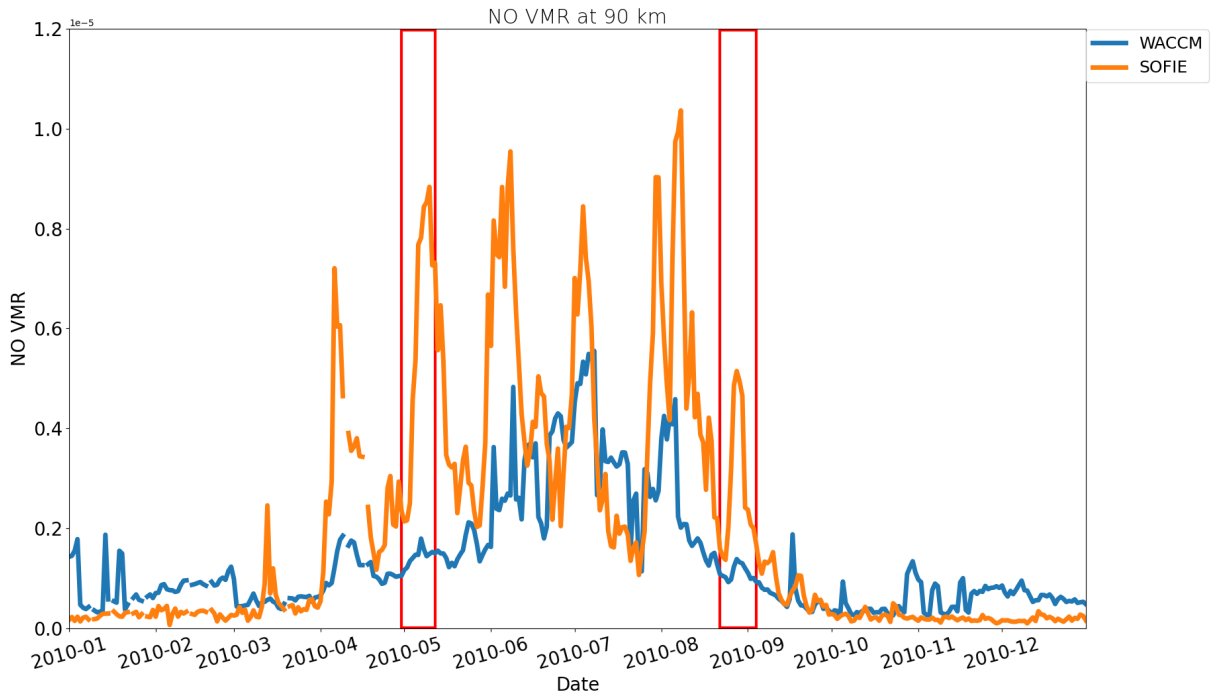

Figure 4.3: NO VMR for SOFIE (orange) and WACCM (blue) at 90 km altitude (upper mesosphere). The red boxes highlight the two studied periods.

In figure 4.4 the NO VMR at 90km, after performing the 30-day running mean, for both WACCM and SOFIE is plotted. At this altitude, NO differences between WACCM and SOFIE have grown larger than the ones seen in figure 4.1 at 100 km. Again, during quiet times they are about the same, while during events, WACCM is still greatly underestimating the NO response, potentially resulting from both a lack in transport and a lack in local production at this altitude.

To further study this discrepancies between WACCM and SOFIE, and to disentangle the role of transport from direct production, it might be more relevant to study individual geomagnetic events instead of the whole year data. In order to have a strong enough NO response and be able to track it, the interesting events need to reach a certain geomagnetic activity threshold. Furthermore, since two geomagnetic events can happen in short periods of time, the signal of one event would be affected by the other event. To avoid this, the goal is to find isolated events where the main production period and subsequent transport is readily identified.

In figure 4.5 the AE-index for the year 2010 is presented. The geomagnetic events giving a daily AE-index higher than 300 nT are highlighted. Furthermore, the blue highlighted events are considered isolated events, which means that they fulfill the condition that the AE-index stays under 170 nT for 3 days before the event starts, as well as for 3 days after the event drops below 170 nT. Hence, based on this criteria, the red highlighted events correspond to compound events.

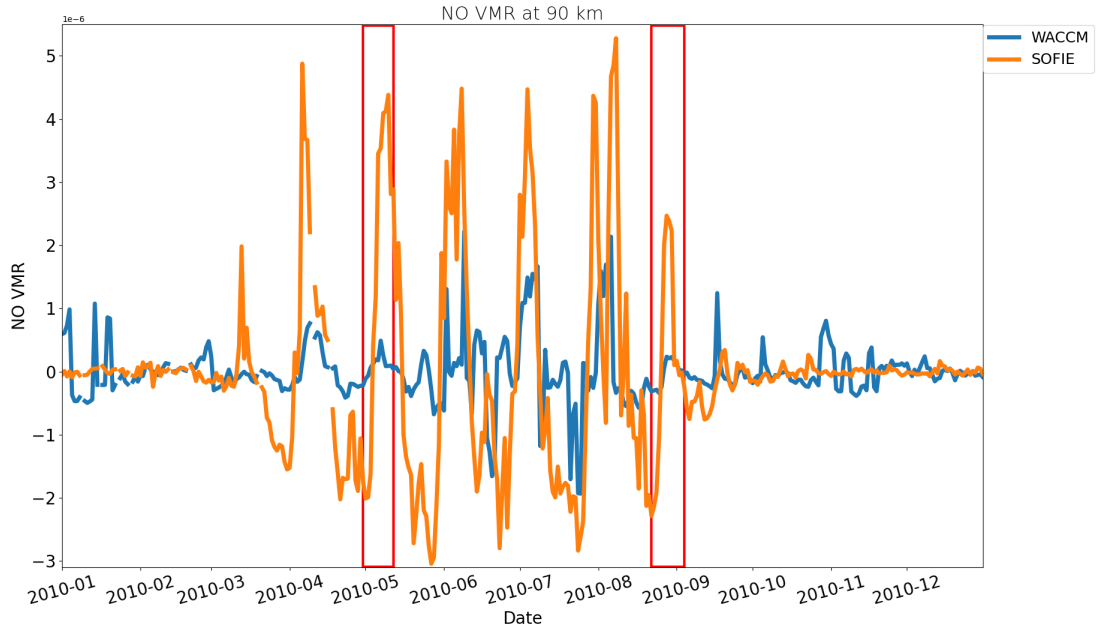

Figure 4.4: NO VMR for SOFIE (orange) and WACCM (blue) at 90 km altitude (upper mesosphere), after subtracting a 30 day running mean. Red boxes highlight the two studied periods.

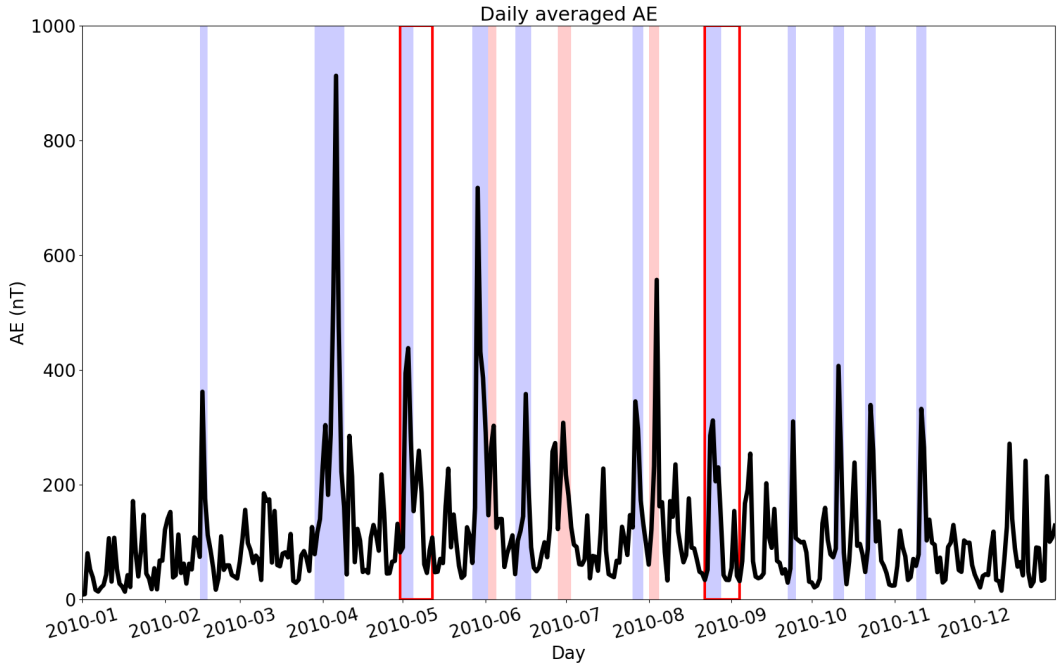

Figure 4.5: Daily averaged AE index for 2010. The blue shaded regions mark isolated events (reaching AE>300 nT with a period of AE<170 nT three days before and after) and the red regions are non isolated events (reaching AE>300 nT, but do not have AE<170 nT three days before and after). The red boxes highlight the two studied periods.

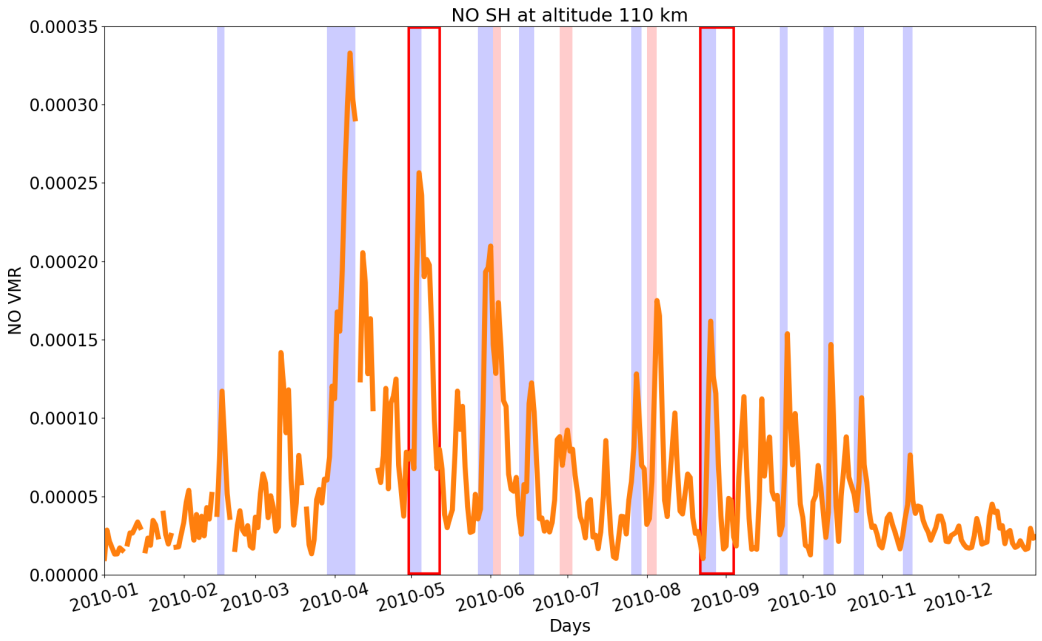

Figure 4.6: NO VMR from SOFIE at 110 km (lower thermosphere), shaded regions are events as defined by AE in the previous figure. The red boxes highlight the two studied periods.

The same color coding is applied on the NO VMR observation from SOFIE in the southern hemisphere in Figure 4.6. The event with the strongest NO response, around the beginning of April is lacking data in the SOFIE observations, so this event is discarded from the study. There is, however, two other events with a high NO response, where the behaviour of NO during and after the main production period can be studied in both WACCM and SOFIE after a geomagnetic storm. These events are the one at the beginning of May, and the one at the end of August, both marked with a red box.

As seen in figure 4.4, at lower altitudes, the NO response in SOFIE can be up to 4 times the NO response in WACCM, making it difficult to make a comparison of both datasets in the same plot. Due to this fact, a more analytical approach will be taken in the comparison, having the figures from both datasets in their own scale. All the data used in this section, both from WACCM and SOFIE, has been normalized using the 30-day running mean to remove the background NO.

4.1.1 August 2010

In figure 4.7 the NO VMR from SOFIE is plotted for two different altitudes. This plot will serve as a reference for the comparisons and the timing of the August event, which corresponds to a CIR. The event takes place from the 23th of August until the 30th of August, reaching a maximum AE of 315 on the 25th of August.

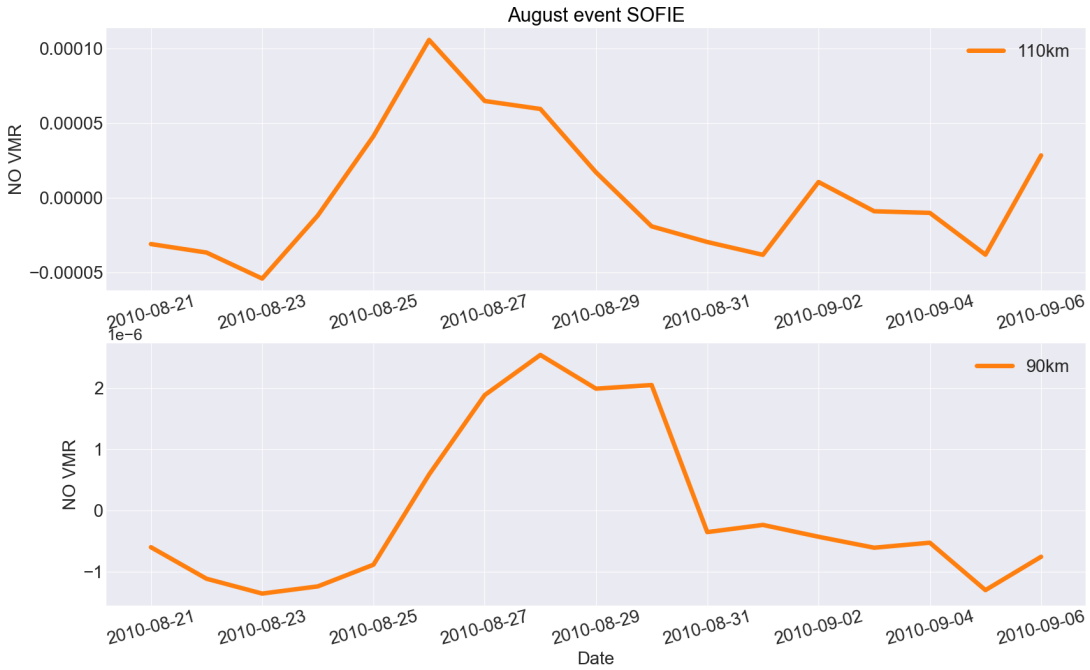

Figure 4.7: NO VMR from SOFIE at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in August 2010

In figure 4.8, the NO VMR from the WACCM MEE run for two different altitudes can be seen. At 110km, SOFIE data builds up for three days before peaking on the 26th. The decay is slower than the rise time but already evident the first day after it peaks. In WACCM, NO peaks the same day, but instead of rapidly declining afterwards, it remains at about the same level until the 28th, when it starts to decay. This means that NO lingers more at high altitudes in WACCM than in SOFIE, which can be caused due to weaker downward transport, not enough NO being destroyed, or that NO transported from above 100 km is replenish the loss.

At 90km, NO in SOFIE peaks on the 28th, two days after the observed peak at 110km. Here, NO lingers at the same level, until the 30th, when it abruptly decreases. In WACCM, it also peaks on the 28th and lingers at the same level until the 30th, but it doesn't abruptly decrease, instead it steadily decreases until the 2nd, which again, implies that NO is either being destroyed or less is being transported.

At 90km the NO VMR is almost one order of magnitude less in the WACCM MEE run in comparison to the data from SOFIE, while at 110km it was only 1.6 times less. The NO at 90 km in both the observations and the model peaks at the same time, which implies that the downwelling rate is similar in both datasets. Hence, the increasing discrepancy in the VMR level with altitude can be attributed to the smaller NO reservoir found at 110 km in WACCM compared to the SOFIE observations, to less NO being transported down in WACCM, and to a weak local production at 90 km.

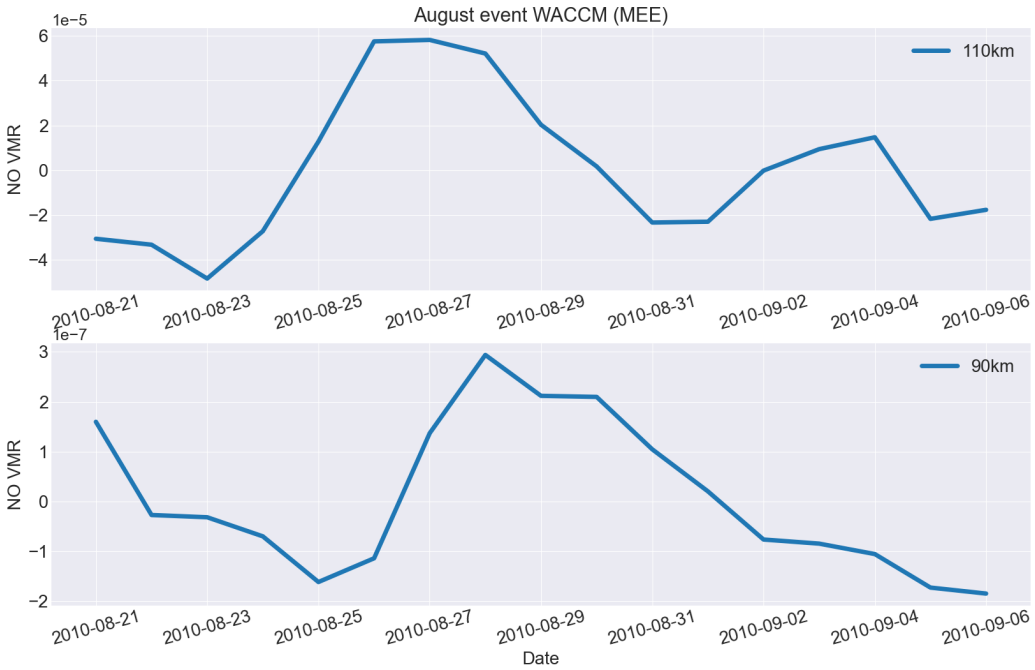

Figure 4.8: NO VMR from WACCM (with medium energy electrons included) at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in August 2010.

It is difficult to pinpoint the relative importance of production and transport in respect to NO level, but if MEE are excluded, we can isolate the contribution due to transport. The production rate should be the same above 100 km, while the direct production rate should be effectively zero below these altitudes, so it is possible to study cases in WACCM, where transport is the only process driving the NO variability below 100 km.

Figure 4.9 shows the NO VMR from the WACCM noMEE run for the same two altitudes as before. At 110km, it can be seen that NO is peaking on the 26th, and it remains on the same level until the 28th, that it starts decreasing. For this case though, it can be seen that the NO VMR is about 1.3 times less (1exp-5) than in the case with MEE, despite that MEE should not produce significant amounts of NO at this altitude.

At 90km, NO peaks on the 30th, meaning that's the day where most production from 100 and above reaches 90 km due to transport. This peak in transport coincides with a period, from the 29th to the 30th, when NO stays at about the same level after it started decreasing, both in the MEE run and in SOFIE. This shows that transport speed is well represented in WACCM, but it is lacking in terms of how much NO is transported. It also shows that even if the timing of direct production by MEE on the 28th is well reproduced, it is also lacking in terms of quantity, as SOFIE NO VMR is one order of magnitude greater than WACCM's at 90km.

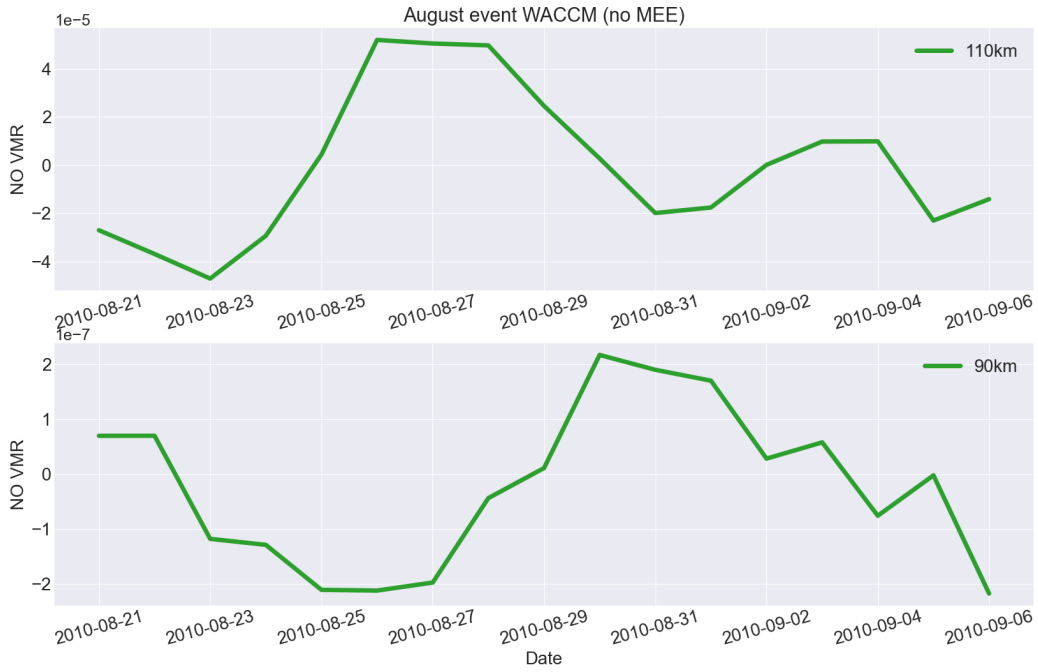

Figure 4.9: NO VMR from WACCM (without medium energy electrons included) at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in August 2010.

May 2010

Figure 4.10 shows the NO VMR from SOFIE for two different altitudes. Again, this plot will serve as a reference for the comparisons and the timing of the May event, which consists of a CME followed by a CIR. The event takes place from the 30th of April until the 5th of May and has a maximum AE of 440 peaking on the 3rd of May.

In figure 4.11, the NO VMR from the WACCM MEE run for two different altitudes is shown. At 110km, SOFIE NO VMR data peaks on the 4th of May, and the following decrease is again already evident on the following day. In WACCM though, the NO VMR peaks the day after, on the 5th, but the NO remains at about the same level since the 4th. As seen before, NO linger for longer in WACCM, meaning that the production peak is on the 4th, and the small increase on the 5th is presumably due to transport, before it starts to decay. After it starts decaying, it can be seen in SOFIE that NO is produced again, having a second peak on the 7th of May, before it starts to decrease towards pre-storm levels. For WACCM though, the second peak does not happen, only decreasing the speed of the decay. Again, this shows that NO production in response to geomagnetic activity is being underestimated by WACCM compared to the SOFIE observations.

At 90km, NO VMR in the SOFIE observations seems to have some production until the 5th of May, where it slows down, but the next day NO increases again and peaks on the 6-7th, presumably due to an indirect NO transported from above. NO in SOFIE remains

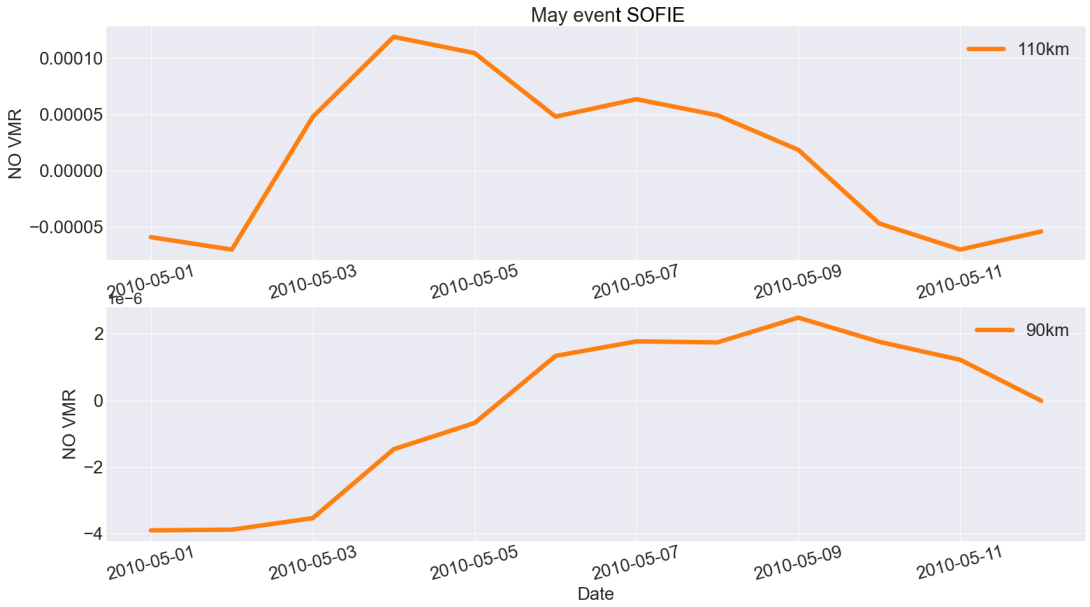

Figure 4.10: NO VMR from SOFIE at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in May 2010.

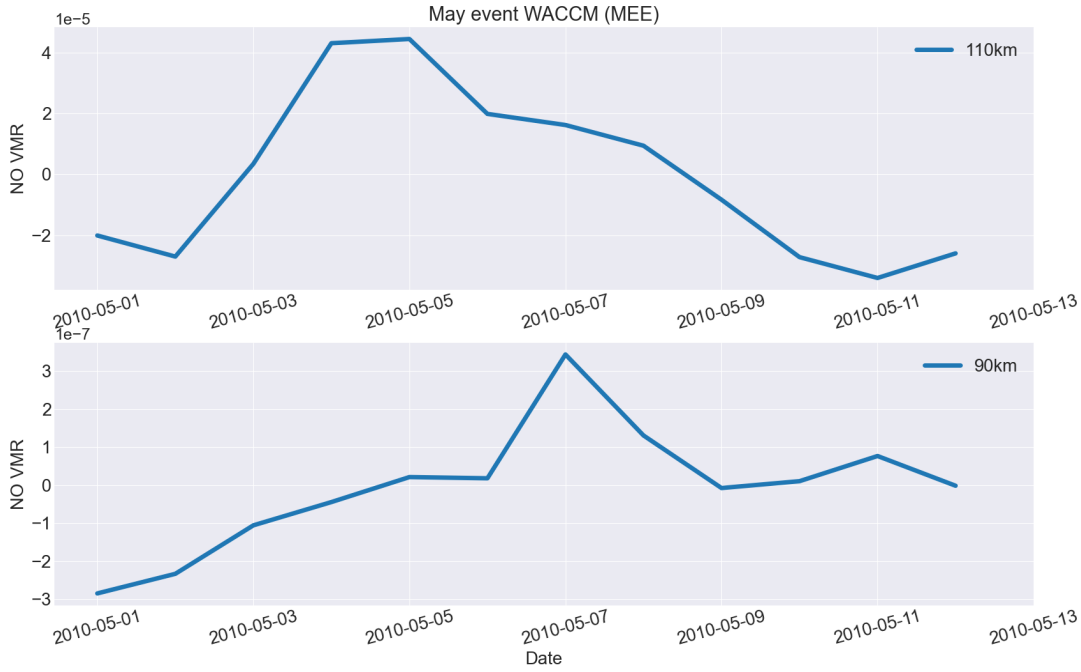

Figure 4.11: NO VMR from WACCM (with medium energy electrons included) at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in May 2010.

around peak level until the 11th, presumably due to the second NO peak on the 7th of May seen at 110 km. In WACCM, it also peaks on the 7th, but it already starts decreasing on the 8th. The fact that in WACCM NO does not stay at peak levels for longer, is probably because the second peak on the 7th at 110km is not as pronounced in WACCM, and does not have the same impact. The NO enhancement in WACCM at 90 km increases and decreases within 3 days.

Again, the NO VMR level is one order of magnitude less in the WACCM MEE run in comparison to the data from SOFIE at 90km, while at 110km it is around 2.3 times less. This increased discrepancy between the two datasets might again be attributed to the weaker NO reservoir in the lower thermosphere in the model compared to the observations, less NO being transported down in WACCM, as well as the direct production at 90 km not being sufficiently strong in WACCM.

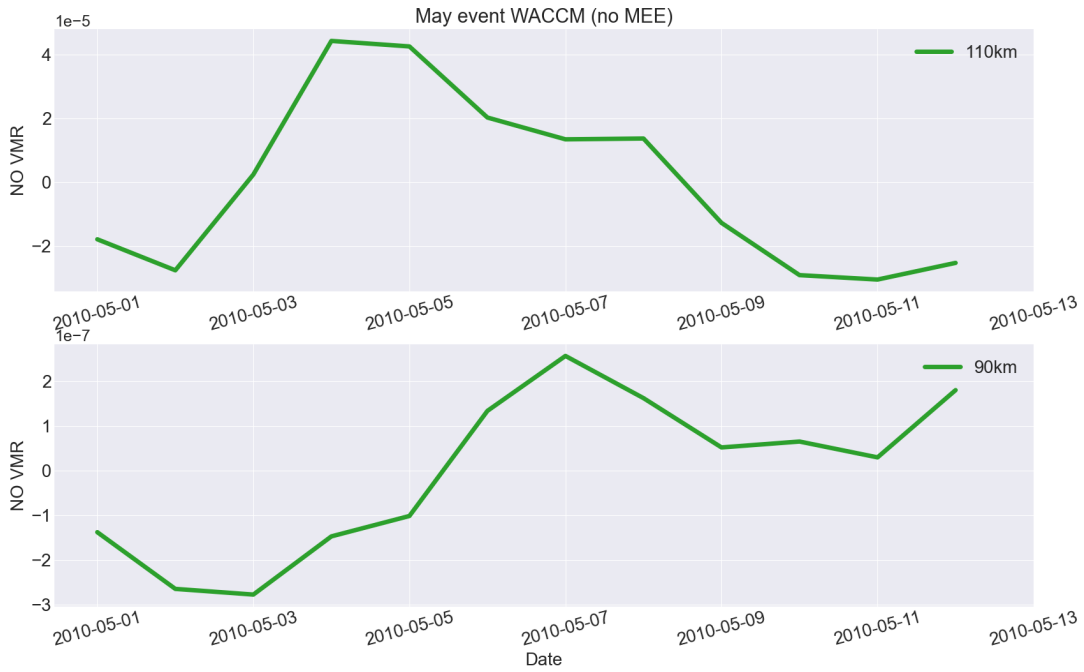

Figure 4.12: NO VMR from WACCM (without medium energy electrons included) at 110 km (lower thermosphere) and 90 km (upper mesosphere) for an event starting in May 2010.

Figure 4.12 is the NO VMR from the WACCM noMEE run for the same two altitudes. At 110km, it can be seen that NO is peaking on the 4th of May, and it remains on the same level until the 5th, then it starts decreasing. For this case though, the second NO production peak, seen on the 7th in SOFIE, is able to stop NO decreasing between the 7th and the 8th, before decreasing to pre-storm levels. As mentioned before, the MEE ionization rates are not expected to have much impact on the NO production rate at 110 km, as seen in this case.

At 90km, NO VMR peaks on the 7th of May, confirming when most production from

100 km and above reaches 90 km due to transport. This peak in transport coincides with the peak in the noMEE run, but has a lower value. Before the peak though, on the 5th the NO VMR in the noMEE run is about 1×10^{-7} lower than in the MEE run. This is also the difference at the peak in transport on the 7th, meaning that the direct production from the MEE which seemingly occurred on the 5th makes up for this difference. For this event, the peak of transported NO from above is greater than the peak due to direct production, explaining most of NO variability found at 90km, while the contrary is seen in the August event, where direct production at 90km exceeds the level of transported NO.

4.1.2 Summary

In summary, based on the NO VMR comparison between WACCM with and without MEE and observations from the satellite instrument SOFIE, some key differences are found:

In the lower thermosphere:

- NO level is too high in WACCM both during summer and during quiet geomagnetic times, while it is too low during the biggest geomagnetic events in winter.
- Subtracting the background NO in both the estimated and observed datasets improves the NO level in WACCM during summer, but during geomagnetic active times WACCM shows less variability than SOFIE.

In the upper mesosphere:

- NO level in WACCM is about the same level as the observations during summer, but it is greatly underestimated during high geomagnetic activity.
- Subtracting the background NO slightly improves it during summer, but during geomagnetic events in winter WACCM show less variability, and greatly underestimates the levels of NO.

Overall, subtracting background NO improves the estimated NO in the thermosphere a lot, meaning that local thermospheric production is correctly modelled, only slightly underestimated during active geomagnetic times. When subtracting the background NO in the upper mesosphere, the summer NO response is slightly overestimated in WACCM. Meaning that local mesospheric production is slightly overestimated during summer. During geomagnetic events in winter, NO in WACCM is greatly underestimated, meaning that both local production and transport could be a problem.

To examine the transport and local production in WACCM, two geomagnetic storms are studied, one in the early winter during May, and the other one in August. To disentangle transport from local production, two WACCM runs are studied, one including MEE and one without MEE. The comparison shows that:

- Local production in the lower thermosphere is slightly less in WACCM than in SOFIE, but NO remains close to the peak level for around two days, while in SOFIE

it starts decreasing after only a day. This could be due to insufficient downward transport of NO from 100 km in the WACCM runs. It is, however, also possible that the slow decay can be attributed to the upper boundary conditions, meaning that loss of NO at 100 km due to transport, is being replenished by NO transported from above 100 km.

- WACCM underestimates the local production by MEE in the upper mesosphere, but the timing of the peak coincides in both WACCM and SOFIE.
- The timing of the peaks at 90 km indicate that the speed of the estimated transport corresponds well with the observations, but the amount of NO being transported down seem to be insufficient, as the difference of NO between WACCM and SOFIE is only about a factor of 2 at 110 km, and it increases to more than one order of magnitude at 90 km. The deficit in transport could also be attributed to the smaller NO reservoir at 110 km in WACCM.

4.2 MEE vs noMEE

In this section, the effect of adding MEE in the WACCM runs is studied. As stated before, auroral electrons deposit their energy at altitudes above the mesopause, while MEE have larger energies, so they penetrate deeper into the atmosphere, producing NO and OH at altitudes throughout the mesosphere, typically peaking between 90km and 70km [*Turunen et al.*, 2009; *van de Kamp et al.*, 2016]. This will lead to an overall increase in production of NO in the mesosphere, and subsequently, more NO will be transported down, since there will be a larger total reservoir. Furthermore, the MEE NO production will be directly embedded in the pole-to-pole branch of the residual transport, seen in figure 2.10. Hence, the molecules do not need to cross the mesopause barrier as the NO produced in the lower thermosphere by auroral electrons.

For all the figures in this section, there will be two subplots. The line plot in the upper panel corresponds to the daily resolved ion-pair production rate (IPR) due to the MEE, which are parametrized by the Ap-index, as described in section 3.2. It indicates the timing and strength of the precipitation events. The lower panel shows the corresponding WACCM data. Here, the data has been averaged over the latitude band 60ºS to 70ºS or 60ºN to 70ºN, depending on which hemisphere is being studied. The altitude is given in pressure levels with units hectopascal (hPa), and the lower end of the axis corresponds to $\sim 50km$, where WACCM is nudged, while the upper end of the axis corresponds to $\sim 105km$. Since the model is nudged below $\sim 50km$, the dynamics in respect to wave forcing and filtering is the same in both runs below this altitude. The time axis is given as monthly ticks on the format YYYY-MM or as YYYY-MM-DD when the resolution is sufficient to resolve days.

Most of the color plots represent differences between the MEE run and the noMEE run, both relative and absolute depending on the variable being studied. For the relative difference the following expression is used :

$$R_{diff} = \frac{Run_{MEE} - Run_{noMEE}}{Run_{noMEE}} \tag{4.1}$$

These plots use a red/blue color scale, where red indicates more quantity in the MEE run and blue means the opposite, more in the noMEE run. The plots also include two colored lines that represent the altitudes of the mesopause for both runs, defined as the temperature minimum at the boundary between the mesosphere and thermosphere. The blue and green lines marks the minimum temperature, assumed to correspond to the mesopause altitude, for the MEE run and noMEE run, respectively.

4.2.1 Northern Hemisphere

Figure 4.13 shows the NO VMR from the MEE run in the northern hemisphere for the year 2010. The NO VMR increase with altitude during winter, while during summer the NO VMR minimum is confined between roughly 10^{-1} and 10^{-2} hPa. Therefore, the NO VMR throughout the mesosphere and the lower thermosphere is higher during winter than during summer. As mentioned in section 3.3, 2010 corresponds to solar minimum, and geomagnetic activity is low when no geomagnetic events are present. This can be seen in the upper plot, which corresponds to the IPR by MEE.

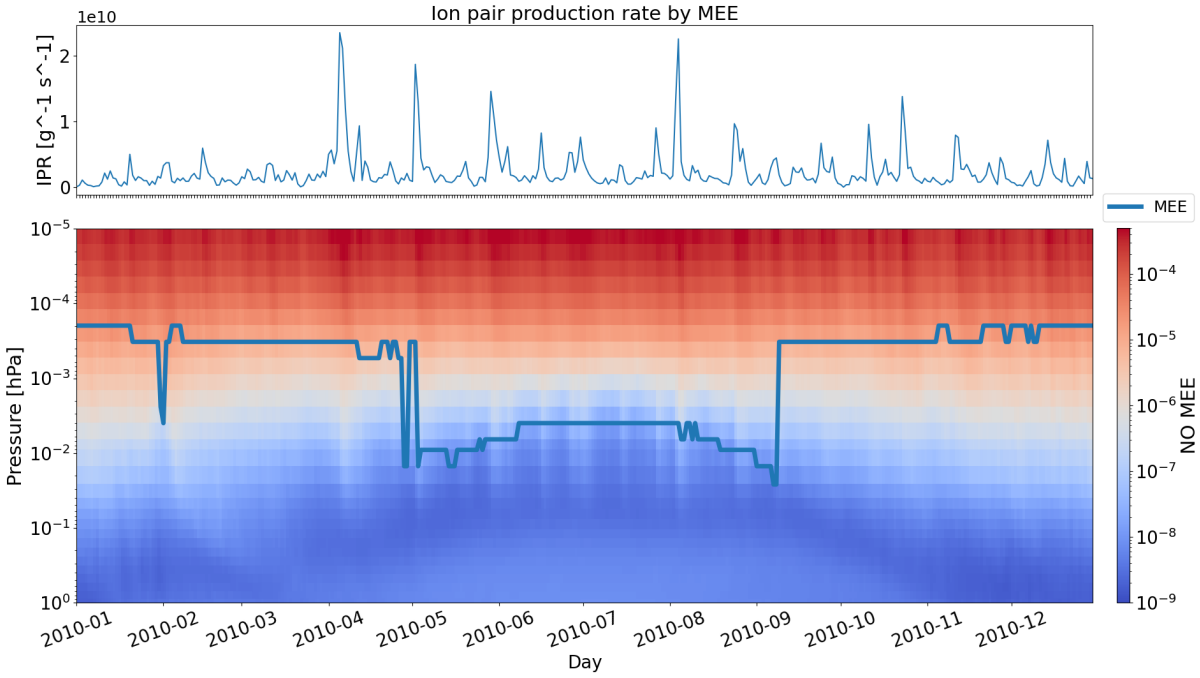

Figure 4.13: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the NO VMR in the northern hemisphere for the year 2010. The data is averaged over the latitude band 60 - 70ºN with a daily resolution. The blue line represents the mesopause altitude (coldest altitude).

Figure 4.14 shows the NO relative difference in the northern hemisphere for the entire year of 2010. The data from both WACCM runs have been 5-day averaged. The location of the mesopause shows an abrupt seasonal mesopause shift around the 1st of May in both the MEE and noMEE case. As seen in figure 4.13, in general there is more NO throughout the mesosphere and lower thermosphere during winter compared to summer. As a result of the small amount of NO in the mesosphere during summer, the added MEE produced NO is relatively stronger during summer compared to winter. The summertime NO enhancement is, however, not transported all the way down to the stratosphere, due to strong photo dissociation. During the winter though, it can be seen during the months of October and November that the MEE produced NO is being transported further down

into the atmosphere, thanks to the polar vortex keeping it in the polar night, safe from sunlight. On the contrary, in the beginning of the year, there's only one time period where there is significantly more NO in the MEE run. Right after, follows a period where there's more NO in the noMEE run around the 1st of February, which is counter-intuitive. Despite relative weak ionization rates at the start of the year, it is still expected that the MEE run have a higher NO production rate throughout the mesosphere. Having more NO, while having a lower production is unexpected, implying that it is more than just the NO production rate that separate the two model runs. It raises the question: Does the MEE ionization change the dynamics of the mesosphere and lower thermosphere in WACCM?

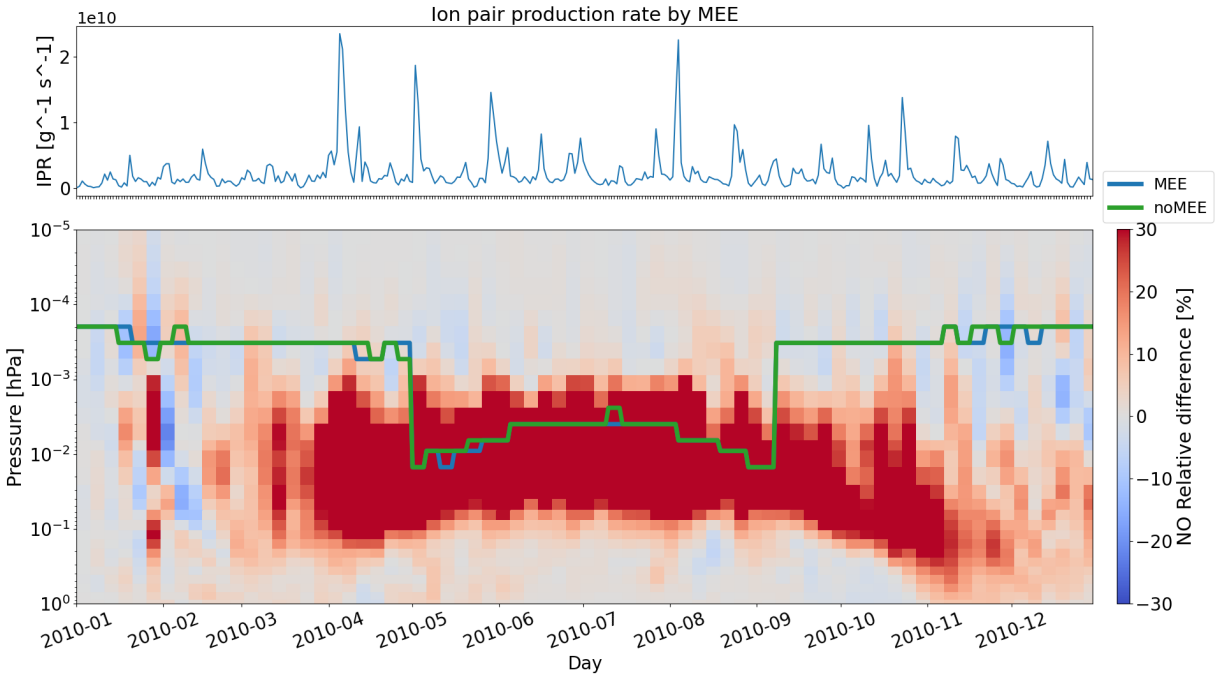

Figure 4.14: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the relative difference in NO VMR, in the northern hemisphere, between the MEE run and the noMEE run for the year 2010. The color red represents more NO being present in the MEE run, while blue represents more NO being present in the noMEE run. The data is averaged over the latitude band 60 - 70^{o}N with a five day resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.15 is the same plot as in figure 4.14 for the first three months of the year now with daily averaged data, instead of 5-day averaged. Around the 25th of January, there's an NO increase in the MEE run , that is being transported down during the next few days, evident by the red feature. Immediately after, a similar blue cell follows, which reflect a relative increase in NO in the noMEE run. Contrary to expectations, less NO production in the noMEE case, is resulting in more NO. The only process that can make this possible is a change in the dynamics, such as a change in the strength/timing of the

residual transport. This kind of profile hints that NO is being transported downwards with a time delay between both runs, happening earlier in the MEE run. Since precipitation events happen at the same time in both runs, this downward transport has to be due to the dynamics of the atmosphere. For this to take place, the chemical changes induced by the MEE are somewhat changing the dynamics. Thus, the zonal wind (U), temperature (T), O_3 and OH will be examined in more detail to understand more the dynamical changes occurring between both runs.

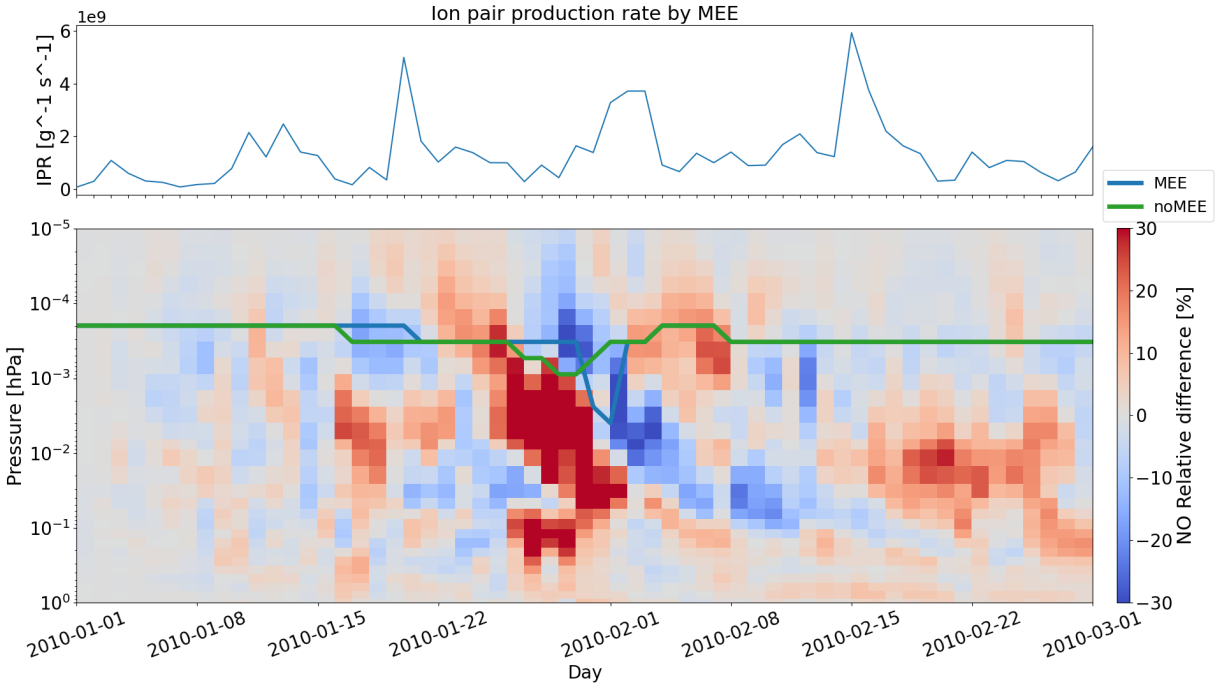

Figure 4.15: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the relative difference in NO VMR, in the northern hemisphere, between the MEE run and the noMEE run for the first two months of the year 2010. The color red represents more NO being present in the MEE run, while blue represents more NO being present in the noMEE run . The data is averaged over the latitude band 60 - 70^{o}N and with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.16 and figure 4.17 show the corresponding zonal winds in the no MEE run and the MEE run, respectively. Therefore, the color bar is not showing differences, but rather wind speeds. The blue color show negative wind speeds, which translates into westward winds. The red color show positive wind speeds corresponding to eastward winds.

It can be seen in figure 4.16 that, for the beginning of the year, it presents a normal winter climatology, with winds being eastward up until almost 10^{-2} hPa, and turning westward above that. This is the case until the 18th of January, when the winds start reversing above 10^{-2} hPa and slowing down below, due to a major SSW with central

date on the 26th of January 2010 [*Dörnbrack et al.*, 2012]. The zonal wind completely reverses from the 23rd of January until the 2nd of February, becoming westward below 10^{-2} hPa. After the SSW, the winds start to get back to normal, presenting the normal winter climatology, but with weaker winds.

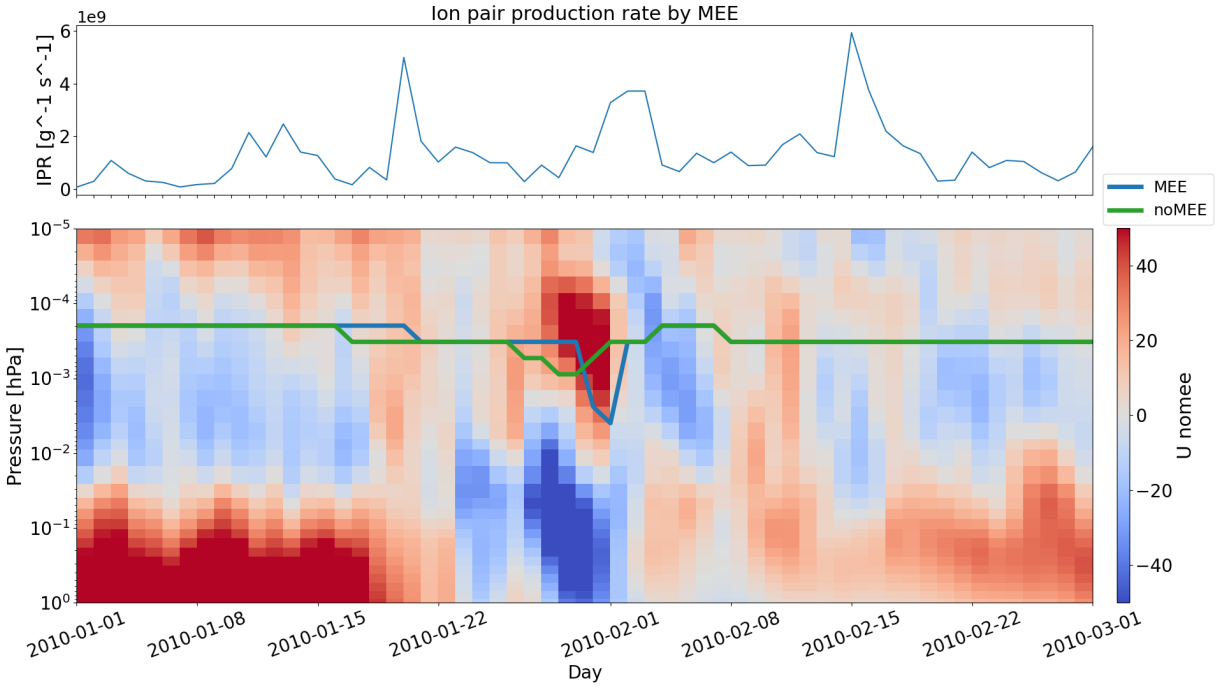

Figure 4.16: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the zonal wind velocity for the noMEE run for the first two months of the year 2010, in the northern hemisphere. The red color represents positive velocities, associated with eastward winds, while blue represents negative velocities, associated with westward winds. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.17, which includes MEE, shows that it presents a normal winter climatology until the 18th, when the winds start slowing down below 10^{-2} hPa. Even if this happens the same day as in the noMEE case, it can be seen that around the mesopause altitude, the winds don't reverse until the 19th, while in the noMEE case, they already reversed on the 17th. Again, the zonal wind completely reverses on the 23rd of January, becoming westward below 10^{-2} hPa. On the 26th though, the winds turn eastward again, before reversing into westward from the 28th of January until the 2nd of February. Thus, in the middle of the SSW, the winds reverse again to normal winter conditions in the MEE run. Hence, the zonal wind behaves differently during the SSW in the cases with and without MEE, despite having the same waves entering the mesosphere. Hence, such a difference are presumably due to a difference in the wave drag in the two runs. This leads to an overall weaker wind reversal from the 28th of January until the 2nd of January in the MEE run.

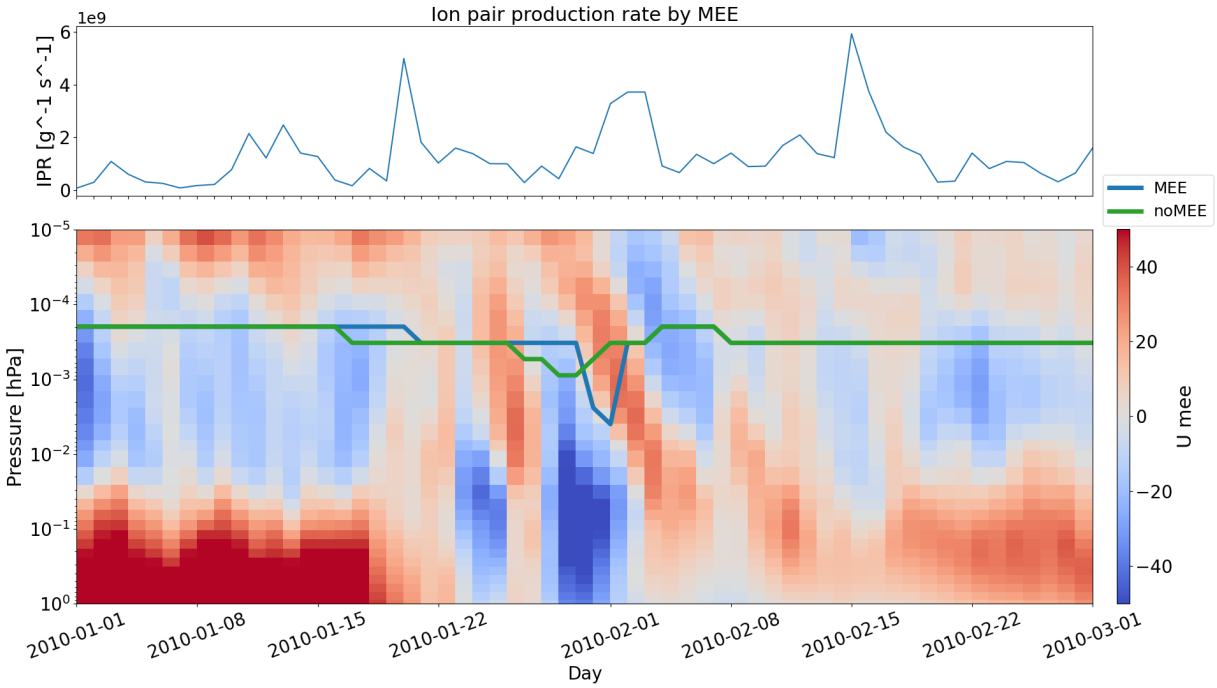

Figure 4.17: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the zonal wind velocity for the MEE run for the first two months of the year 2010, in the northern hemisphere. The red color represents positive velocities, associated with eastward winds, while blue represents negative velocities, associated with westward winds. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.18 shows the absolute difference in the zonal wind velocity between the MEE and noMEE run. Zonal wind discrepancies of more than 20 m/s occur from the 26th of January, when the difference in the SSW between both runs are found, until the 6th of February, 4 days after the SSW ends. Before the SSW takes place, from the 15th of January until the 20th of January, differences between both runs up to 20 m/s are also found between 3×10^{-3} and 10^{-4} hPa. For the whole period, the zonal wind is more eastward in the MEE run. Before the 15th of January and after the 10th of February, the wind differences are less than 10 m/s.

During a SSW, NO greatly decreases in the lower mesosphere, since transport is prevented [*Smith et al.*, 2011]. After the SSW, the downward transport of NO is often enhanced when the winds are getting back to normal winter conditions, as seen in section 2.4.6. According to Figure 4.17, winds start getting back to normal on the 26th of January for the MEE run. On the other hand, for the noMEE run, the wind reversal still persists, preventing NO from being transported downwards. The NO differences in figure 4.15 show a larger amount of NO in the MEE run around the 26th of January, consistent with a

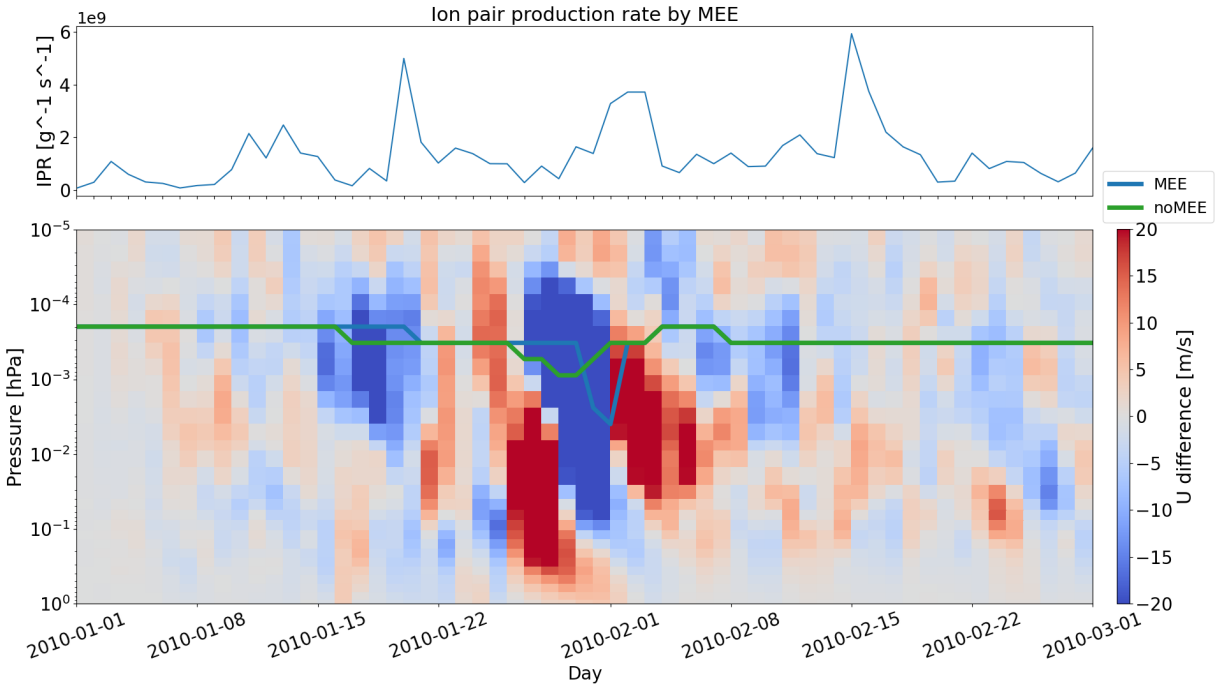

Figure 4.18: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the absolute difference in the zonal wind velocity between the MEE run and the noMEE run for the first two months of the year 2010, in the northern hemisphere. The red color represents more eastward winds in the MEE run, while blue represents more westward winds. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

downward transport associated with the respective wind reversal. It can also be seen that, from the 27th of January until the 1st off February, around the mesopause altitude, the noMEE run has a bigger NO reservoir, since during a SSW, NO is prevented from being transported downward. The SSW in the noMEE run lasted until February 1st and had more persistent and stronger zonal winds. When the winds get back to normal winter conditions, the NO transport will be enhanced, and together with the bigger NO reservoir seen around the mesopause altitude, it may lead to the NO "increase" seen in the noMEE run alongside the associated downward transport.

There is also a temperature difference between the MEE and noMEE case associated with the SSW as demonstrated in figure 4.19. On the 26th of January the temperatures are more than 5K warmer between 10^{-2} hPa and 5×10^{-4} hPa, in the MEE compared to the noMEE case. This is likely due to the adiabatic warming associated with the air parcels being displaced downwards. An opposite, but weaker temperature pattern emerge on the 1st of February between 10^{-1} and 2×10^{-3} hPa, coinciding with the end of the SSW.

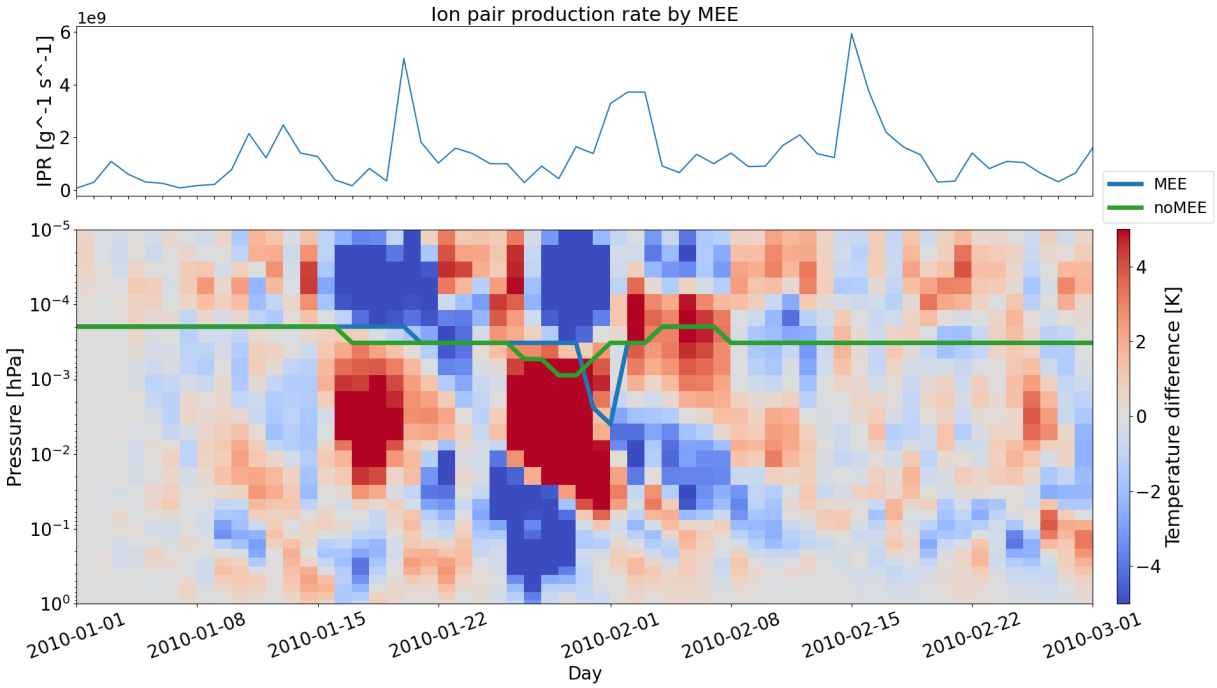

Figure 4.19: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the temperature difference, in the northern hemisphere, between the MEE run and the noMEE run for the first two months of the year 2010. The color red represents warmer temperatures in the MEE run, while blue represents colder temperatures in the MEE run. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.19 also shows temperature differences greater than 5K even before the SSW happens. This differences can be seen from the 16th of January until the 20th reflecting a warmer temperature in the MEE run below the mesopause, between 2×10^{-2} and 3×10^{-4} hPa, while warmer temperatures in the noMEE run above the mesopause, from 2×10^{-4} to 10^{-5} hPa. The dipole-like signal (colder above and warmer below the mesopause for the MEE run), implies that there might be an altitude shift in the dynamics between both runs, presumably on the altitude where gravity waves deposit their energy. Waves depositing their momentum and energy at slightly different altitudes, will result on different wave forcing. Figure 4.18 confirms that the winds are up to 20 m/s more westward from the 15th of January until the 20th of January between 3×10^{-3} and 10^{-4} hPa in the MEE run, potentially resulting from the forcing due to the waves. As mentioned in 3.3, the gravity wave forcing is not available for the respective runs, and this hypothesis cannot be verified at this point.

Focusing on the period before the 16th of January there's overall warmer temperature on the MEE run for the altitudes below 0.01 hPa. These warmer temperatures translate

in a less steep temperature gradient in the MEE run. A weaker temperature gradient will allow gravity waves to reach higher altitudes before their amplitudes produce a super adiabatic lapse rate where they become convectively unstable and "break".

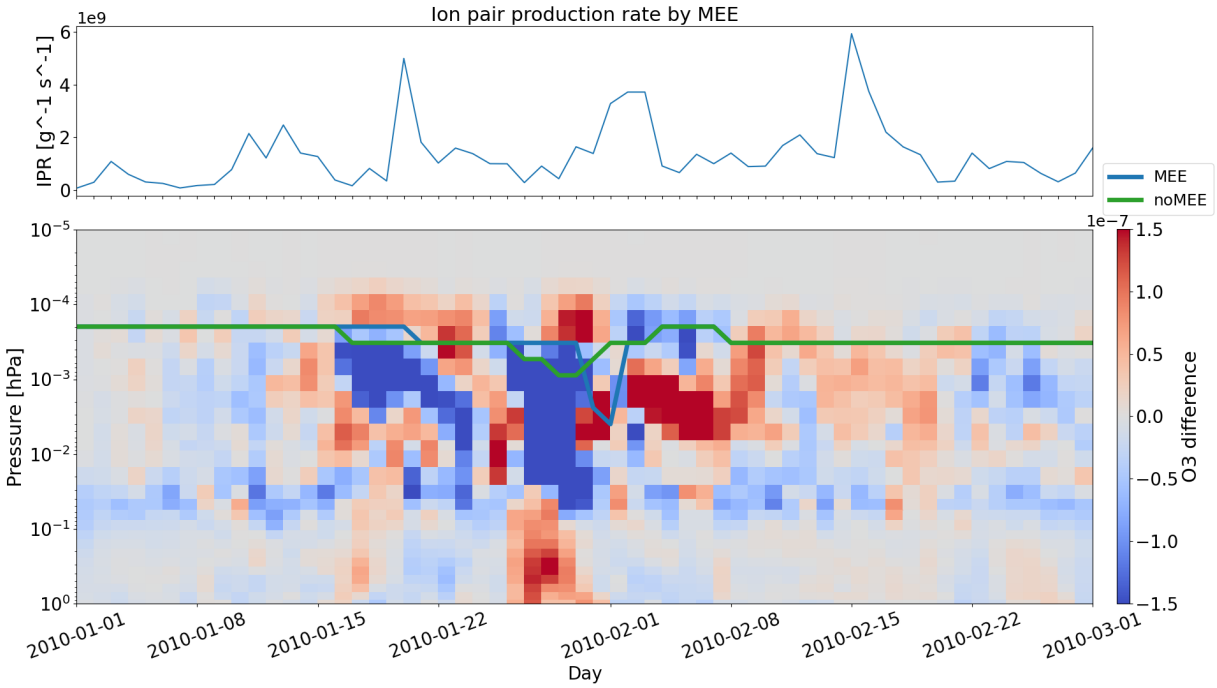

Figure 4.20: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the O_3 VMR difference, in the northern hemisphere, between the MEE run and the noMEE run for the first two months of the year 2010. The color red represents more O_3 being present in the MEE run, while blue represents more O_3 being present in the noMEE run. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Figure 4.20 shows the ozone difference between the MEE and noMEE run. In an altitude band between 10^{-1} and 2×10^{-2} hPa, there is consistently less O_3 in the MEE run throughout most of the year. This is readily evident from the beginning of the year, also before the 16th of January, despite the fact that no significant MEE events have taken place.

January at the northern hemisphere corresponds to wintertime conditions which implies that a decrease in O_3 will lead to enhanced temperatures, since O_3 works as a cooling agent during winter, due to the absence of UV [*Brasseur and Solomon*, 2005]. Indeed, figure 4.19 shows that between 10^{-1} and 10^{-2} hPa, there are about 1-2 K warmer temperatures in the MEE run, corresponding to the altitudes were the O_3 differences are found.

There is, however, no clear evidence to pinpoint where this O_3 deficit in the MEE case comes from. Despite the low geomagnetic activity, the MEE ionization will not be zero. Hence, it will produce molecules that are destroying O_3, also during this period of time. NO can deplete O_3, but figure 4.15 do not show a pronounced difference in the NO for the period of time and the altitude where the O_3 differences are found. However, the O_3 loss catalytic cycle due to NO_x is effective only below $\sim$ 0.5 hPa [*Lary*, 1997].

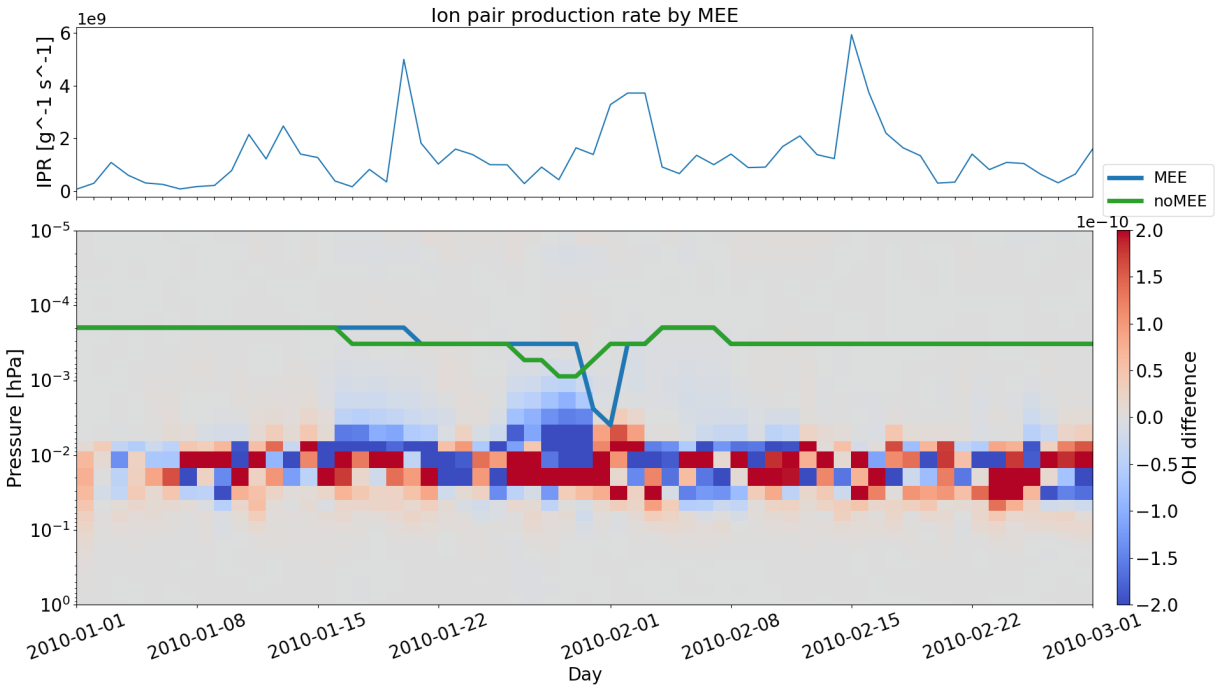

Figure 4.21: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the OH VMR difference, in the northern hemisphere, between the MEE run and the noMEE run for the first two months of the year 2010. The color red represents more OH being present in the MEE run, while blue represents more OH being present in the noMEE run. The data is averaged over the latitude band 60 - 70ºN with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

On the contrary, O_3 loss catalytic cycle due to OH is also effective in the upper and middle mesosphere [*Lary*, 1997], and OH is being produced in the mesosphere by MEE. The relative OH VMR change between the MEE and noMEE runs is plotted in figure 4.21. As expected, the differences are confined to a narrow altitude band. This band is found between 5×10^{-2} and 5×10^{-3} hPa, partly coinciding with the altitude where O_3 is being depleted in the MEE run. There is, however, not a clear increase in the OH abundance in the MEE case compared to the noMEE case.

Regardless, there's evidence that including MEE is effectively changing the dynamics of the atmosphere in the northern hemisphere. These changes potentially come from an

increased O_3 depletion in the MEE run, which effectively changes the temperature of the atmosphere, which subsequently might affect the gravity wave propagation, and consequently lead to different dynamics in both runs. The increased O_3 depletion is during a period of time with low geomagnetic activity, in fact one of the lowest geomagnetic activity for the whole year 2010. Hence, these changes are not induced by a large geomagnetic storm, but rather from unstable atmospheric periods pre-SSW allowing EEP responses to be dynamically amplified.

Hypothetically, the different wave forcing in both runs results in a different SSW. In the MEE run, the SSW appear weaker and break mid-way, as if something was preventing it from happening. This behaviour is not observed in the noMEE run. Hence, based on the two WACCM simulations it appears that even a small amount of MEE ionization has the potential of tipping the scale in a dynamically unstable atmosphere.

4.2.2 Southern Hemisphere

For the southern hemisphere, there is a stronger polar vortex due to the lack of large-scale, planetary waves, which means that the southern hemisphere presents more stable dynamics, and lower chances of having SSWs. Nevertheless, previous observations and model simulations suggest that the two hemispheres are dynamically coupled [*Smith et al.*, 2020], which launches the question: Will the MEE induced dynamical differences found during the northern hemisphere SSW manifests themselves in the southern hemisphere via interhemispheric coupling?

Figure 4.22 shows the relative difference of NO VMR between the MEE and noMEE runs for the southern hemisphere. Again, the data is averaged over 5 days and over the latitude band from 60^{o} to 70^{o}. Already from the very start of the year, despite no major precipitation events, there's a significant amount of NO produced by the MEE. The high percentage change is due to the low level of NO during summertime. This NO abundance in the MEE run is not being transported all the way down to the stratosphere, as NO is destroyed by sunlight, and the concentration in both runs will equalize before reaching the stratosphere. On the contrary, during winter time, the added NO abundance in the MEE run reaches the stratosphere. This is because during winter time, the polar vortex prevents NO from moving to lower latitudes, effectively isolating it from sunlight and making it long lived enough so that it can reach the stratosphere.

Focusing on the mesopause location, illustrated by the blue and green line, the seasonal shift from summer to winter occurs approximately ten days later in the MEE run compared to the no MEE run. This delay suggests, again, a change in the dynamics of the atmosphere between the runs, presumably as a result from the changes produced by the MEE in the mesosphere. On the contrary, the seasonal shift of the mesopause from winter to summer happens at the same time at the end of October, apparently not being affected by the MEE.

Figure 4.23 shows the temperature difference for the first four months of 2010. The

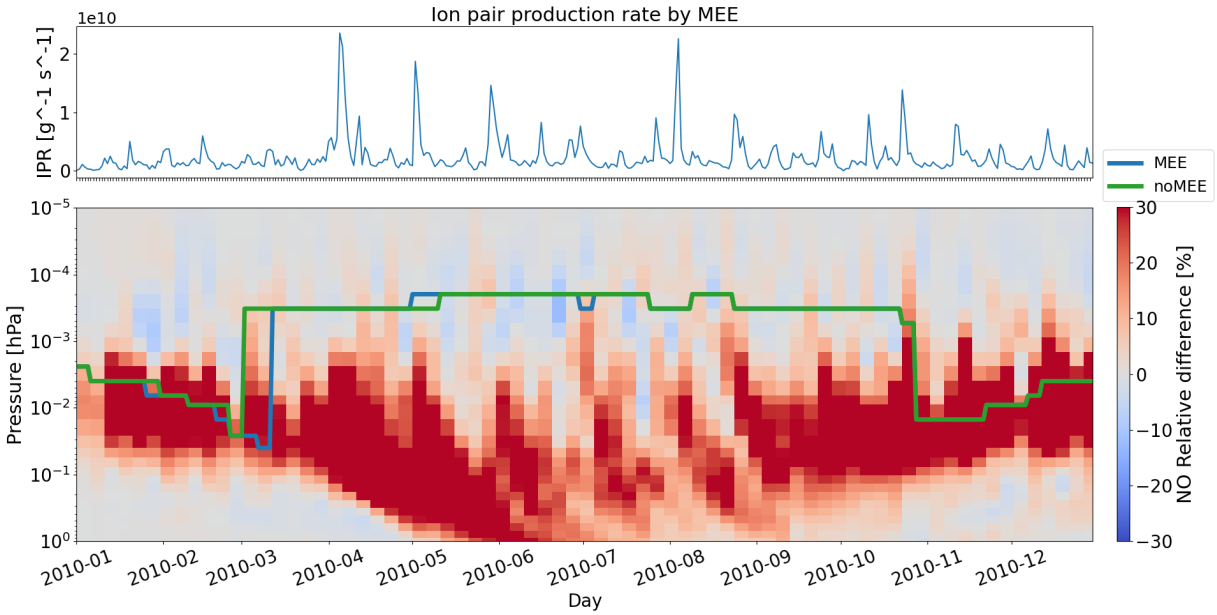

Figure 4.22: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the relative difference in NO VMR, in the southern hemisphere, between the MEE run and the noMEE run for the year 2010. The color red represents more NO being present in the MEE run, while blue represents more NO being present in the noMEE run. The data is averaged over the latitude band 60 - 70^{o}S with five days resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

data is now daily-averages in order to identify the mesopause shift more accurately. The seasonal transition happens on the 2nd of March on the noMEE run, while it happens seven days later, on the 9th of March, for the MEE run. At the altitude just below the mesopause, there's colder temperatures in the MEE run from the beginning of the year and until the time of the seasonal transition, when the mesopause altitude shifts. The persistently colder temperatures below the mesopause in the MEE run compared to the noMEE run imply that MEE have induced a dynamical difference between the two model runs.

There's only one small period of time, from the 10th to the 15th of February when this is not the case, and the temperatures are slightly higher in the MEE run. This seems to be linked to warmer temperatures that appear in the lower thermosphere from 10^{-4} hPa to 10^{-5} hPa on the 28th of January, that are displaced downward over time, resulting in warmer differences just below the mesopause. The warmer temperatures appearing between 10^{-4} hPa and 10^{-5} hPa coincide in the time with colder temperatures appearing on the 28th of January in the lower thermosphere at the northern hemisphere, due to the differences on the SSW between both runs, as seen in figure 4.19. This dipole-like signal of colder temperatures in the northern hemisphere and warmer temperatures in the southern

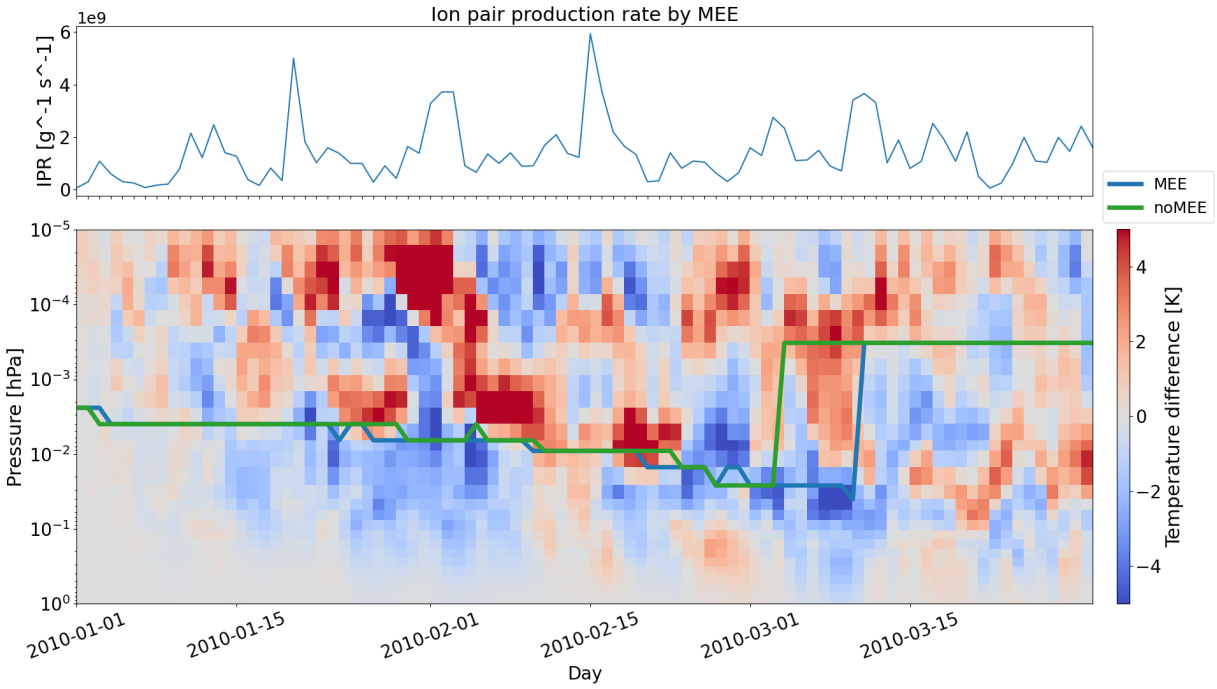

Figure 4.23: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the temperature difference, in the southern hemisphere, between the MEE run and the noMEE run for the first three months of the year 2010. The color red represents warmer temperatures in the MEE run, while blue represents colder temperatures in the MEE run. The data is averaged over the latitude band 60 - 70^{o}S with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

hemisphere, both appearing on the same day at the lower thermosphere, seem to have similar characteristics as temperature anomalies resulting from interhemispheric coupling in *Smith et al.* [2020].

Figure 4.24 shows the seasonal shift from winter to summer on October. From around the 8th of October, colder temperatures just below the mesopause become a persistent trend in the MEE run. This results in the mesopause altitude shift to lower altitudes two days earlier, on the 26th of October, in the MEE run compared to the noMEE run. Nevertheless, the mesopause in both runs reaches a stable winter altitude, at 10^{-2} hPa, around the same day, on the 29th of October. This difference in the mesopause shift did not appear when the data was daily averaged over 5 days. It seems again, that including MEE makes the region below the mesopause colder, changing the time of the seasonal transition, which happens later for the MEE run in March, and sooner in October, effectively shortening the winter time.

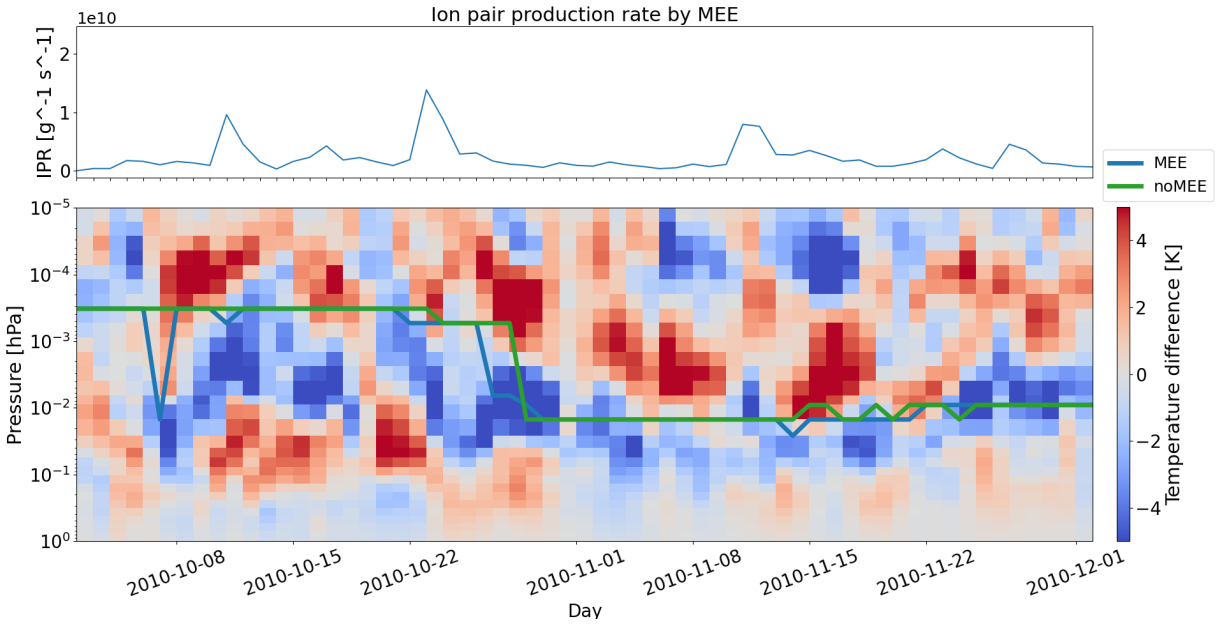

Figure 4.24: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the temperature difference, in the southern hemisphere, between the MEE run and the noMEE run from October to December of the year 2010. The color red represents warmer temperatures in the MEE run, while blue represents colder temperatures in the MEE run. The data is averaged over the latitude band 60 - 70^{o}S with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Chapter 5

Discussion

This chapter will further address the transport of NO based on the data from SOFIE observations and WACCM simulations. A particular focus will be on the requirements for the MEE precipitation to affect the dynamics of the atmosphere, the relation of the dynamical changes in the northern hemisphere and the southern hemisphere, the difference in the seasonal transition between the WACCM runs and finally, a brief discussion about WACCM itself.

5.1 NO_x production and transport in WACCM

The long-lived EPP-enhanced NO_x-gas will be transported both horizontally and vertically. Hence, to reproduce the effect of EPP in current chemistry-climate general circulation models, it is not sufficient to apply a realistic energy input as the chemical productions and losses as well as transport need to be adequately represented. Recently, an improved chemical scheme have been added to WACCM [*Verronen et al.*, 2016], but uncertainties remain in respect to the models' capability of reproducing the NO density throughout the lower thermosphere and mesosphere [*Randall et al.*, 2015]. A comparison of WACCM with eight years of satellite observations in the southern hemisphere found a displacement in maximum altitude, overestimated background levels, and underestimated short-term variability of NO density in the thermosphere [*Hendrickx et al.*, 2018]. The latter is supported by an event study, which finds high pre-storm background levels, but a relative weak response in the NO densities in the lower thermosphere compared to the observations [*Smith-Johnsen et al.*, 2018]. Consequently, the amount of NO that reaches the stratosphere will be underestimated, and a potential EEP-effect deeper into the atmosphere will be underrated.

In section 4.1, two events were studied in order to compare the NO production and transport in WACCM and SOFIE. The transport of NO seemed to be insufficient in WACCM, but the timing of the NO transported at 90 km coincided with SOFIE, meaning

that the speed at which it is transported down coincides in both WACCM and SOFIE. This is consistent with the findings of *Hendrickx et al.* [2018].

Figure 5.1 shows the onset of the NO enhancement, the first NO peak most likely associated with direct EEP impact, and the secondary peak/plateau most likely caused by subsequent transport during the August event for both SOFIE and WACCM at different altitudes, ranging from 85 km to 110 km. The onset of the storm happens one day later in WACCM for most of the altitudes, but the peak of the storm coincides in time for both WACCM and SOFIE at almost all altitudes. The only altitude were this is not the case is at 85km, where there's a difference of only 1 day. For the determination of the peak in transport, increase in NO after the initial production peak, and a slowing of the declining phase after the production peak are used as a reference in the SOFIE data and in the WACCM data above 95 km. Below 95 km, the peak in the noMEE run is used as the reference for the peak in transport in WACCM. These peaks associated with transport coincide in time for both SOFIE and WACCM, showing that the speed of the transport is well reproduced in WACCM, even if the amount of NO transported downward is insufficient.

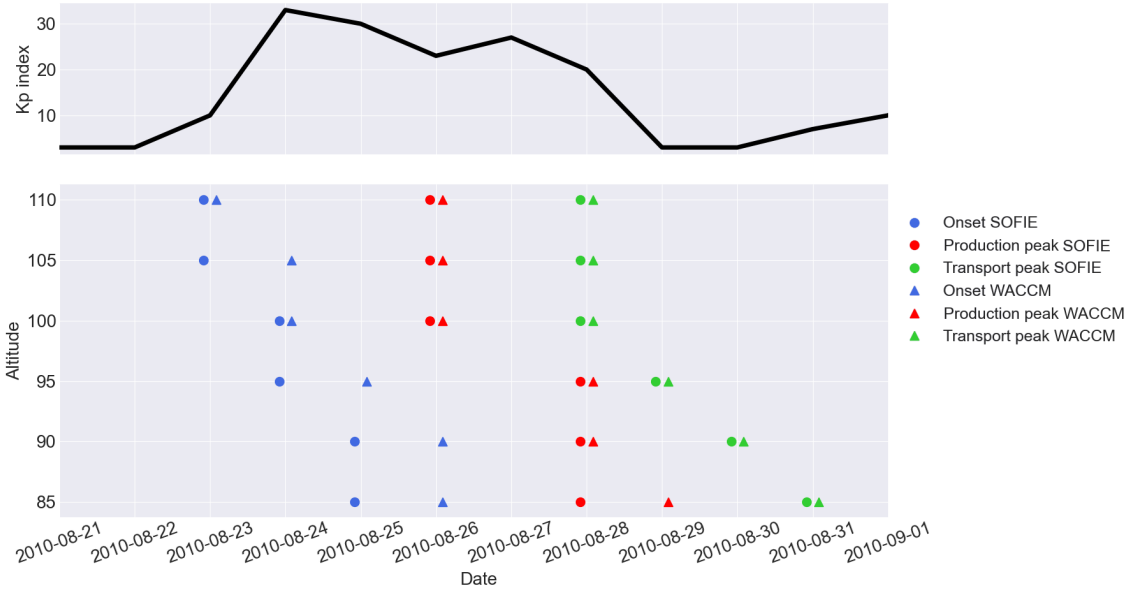

Figure 5.1: Representation of the timing of NO production and transport for both SOFIE and WACCM for the event in late August.

Figure 5.2 shows the similar features for the May event. Again, the timing of the onset happens one day later for two of the altitudes studied. Regardless, the peaks due to direct production in WACCM and SOFIE coincide in time for all the altitudes except at 100 km, where the peak is delayed one day in WACCM. For the peak/plateau associated with transport, there's one day delay in WACCM compared to SOFIE at 100 km and 95 km. Nevertheless, for both SOFIE and WACCM it takes 3 days for the production peak at 110

km to be transported down to 85 km, showing again a good representation of the transport speed in WACCM. The one day difference found associate with transport at 95 and 100 km, could reflect that the NO peak associated with direct production at 100km also occurs one day later in WACCM, affecting the timing of the transport at 95km.

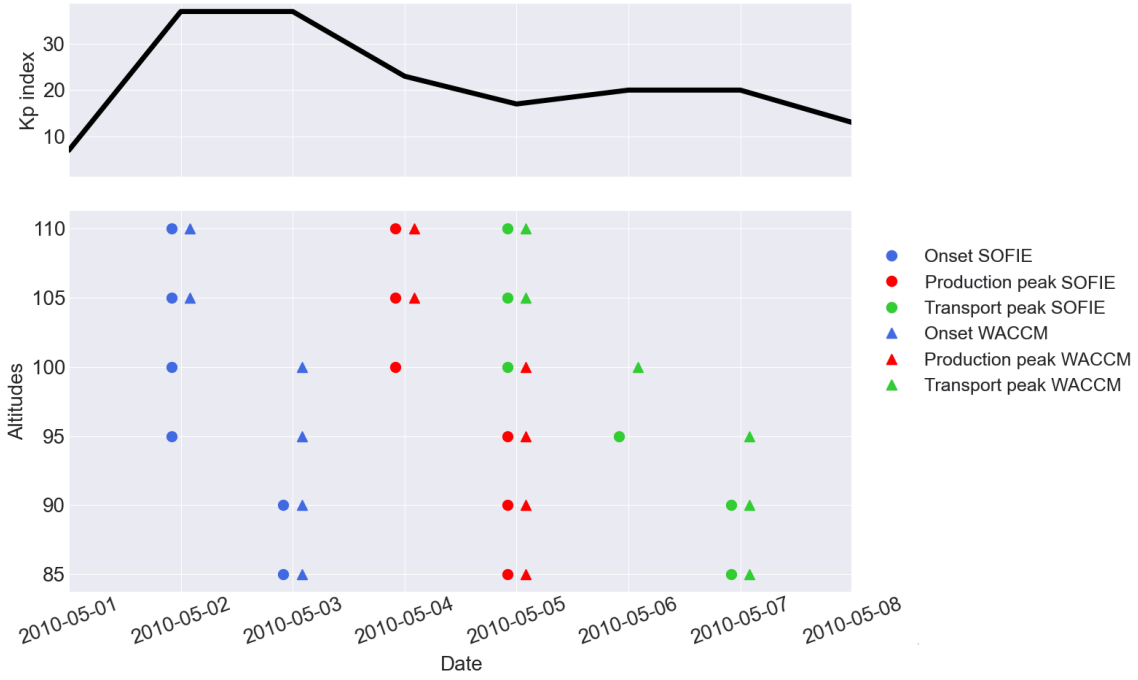

Figure 5.2: Representation of the timing of NO production and transport for both SOFIE and WACCM for the even in early May.

5.2 Unstable atmospheric mean-flow conditions

The chain of reactions involved in the existing theories on how EEP can impact temperature and wind in the atmosphere, are mainly concerned with stratospheric dynamics [*Seppälä et al.*, 2013; *Asikainen et al.*, 2020]. The role of the MEE direct O_3 effect is unclear in the given hypothesis. Furthermore, *Meraner and Schmidt* [2018], rejected the possibility of mesospheric O_3 reduction playing a major role for atmospheric dynamics based on a model study which artificially introduced a steady ozone reduction in the mesosphere and stratosphere separately. In contrast, as shown in section 4.2, MEE changed the dynamics of the atmosphere during the period with the lowest geomagnetic activity for the year 2010. Hence, it appears that even weak MEE ionization has the potential of changing temperature and winds in the mesosphere. One hypothesis, is that EEP can affect the dynamics in periods before SSWs, which are characterized with strong stratospheric planetary wave forcing and a weak polar vortex, i.e. unstable conditions, as found by *Asikainen et al.* [2020] for the stratosphere. Additionally, *Salminen et al.* [2020] found that during a QBO-

E phase, as it was in 2010, low AP values (or low EEP) allow SSW to occur more often. Including the MEE, will increase the EEP, which seemingly prevents SSW from happening in this study case.

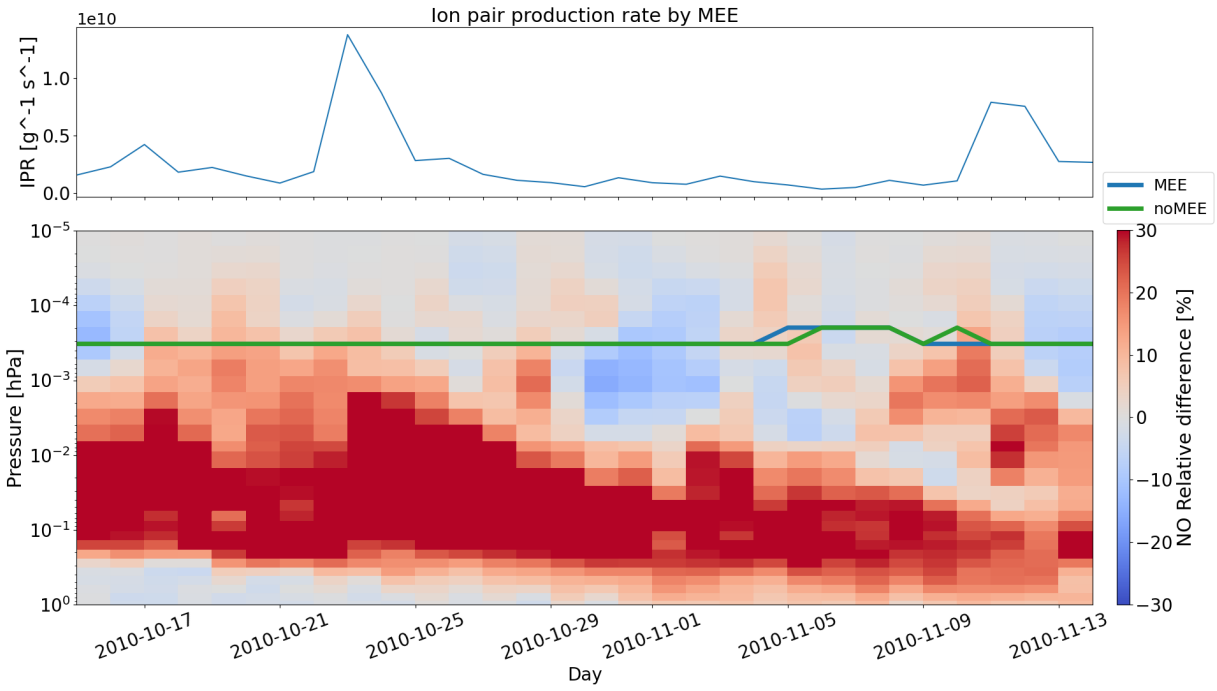

Figure 5.3: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the relative difference in NO VMR, in the northern hemisphere, between the MEE run and the noMEE run from the 15th of October to the 15th of November. The color red represents more NO being present in the MEE run, on the contrary, blue represents more NO being present in the noMEE run. The data is averaged over the latitude band 60 - 70^{o}N and with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

To test this hypothesis, the period with the highest geomagnetic activity during winter in the northern hemisphere will be presented. This event happens on the 20th of October and the IPR is one order of magnitude higher than the IPR before the 15th of January.

Figure 5.3 shows the NO difference between the MEE and noMEE runs for the October event. On the 23rd of October, when the IPR event peaks, NO enhancements of more than 30% occur between 2×10^{-1} and 10−3 hPa. The following days, the NO enhancement at the highest altitudes are being transported down. On the 28th, the NO enhancements of up to 30% are found only up to 10^{-2} hPa. The NO enhancements of more than 30% also occur at the lowest altitudes, from 2×10^{-1} to 2×10^{-2} hPa. This enhancement exist already before the IPR event and are likely due to the increased NO production from previous events. On the contrary, figure4.15 show no such NO enhancements during the quiet times of the beginning of the year, from the 1st of January until the 15th. As expected, NO

enhancements due to the MEE are higher in the October event, since the IPR is almost one order of magnitude higher.

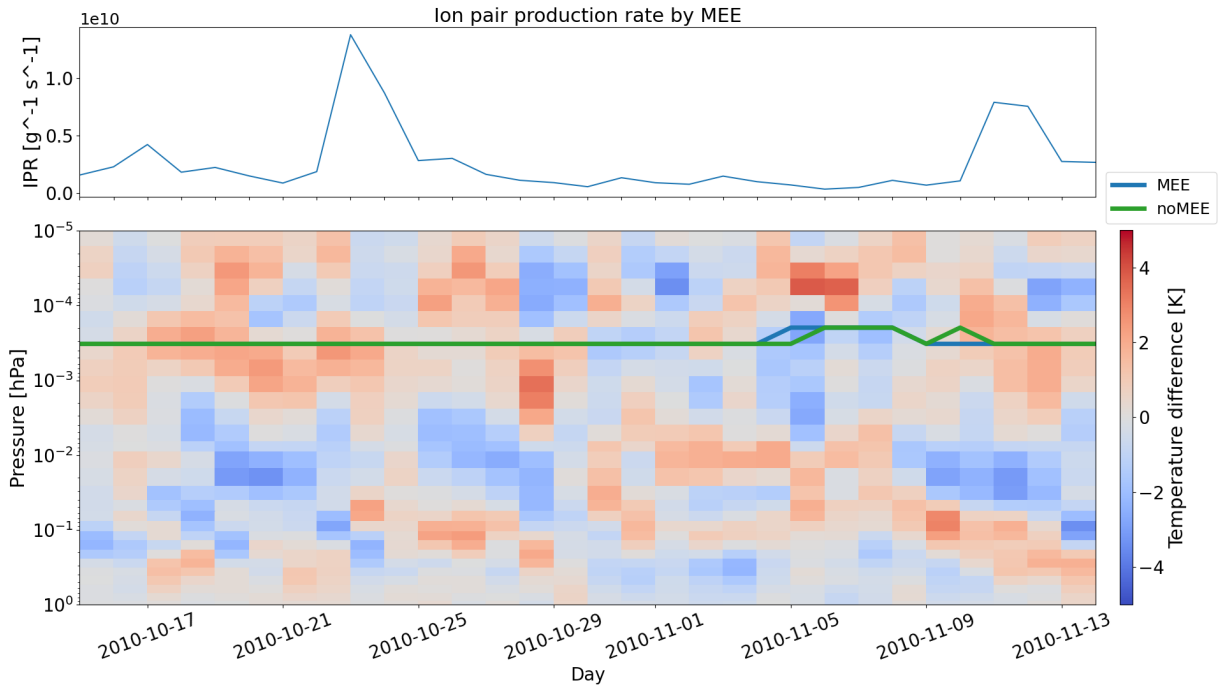

Figure 5.4: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the temperature difference, in the northern hemisphere, between the MEE run and the noMEE run from the 15th of October to the 15th of November. The color red represents warmer temperatures in the MEE run, on the contrary, blue represents colder temperatures in the MEE run. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

The temperature difference for the MEE and noMEE run for the October event is shown in figure 5.4. From the 23rd to the 28th of October the maximum temperature differences across the mesosphere reach about ±2 K. These temperature differences are comparable to the ones seen in figure 4.19 before the 15th of January, where maximum temperature differences across the mesosphere reach also about ±2 K. The initial small temperature difference in January, led to differences of more than 5 K in the pre-SSW period. The temperature differences in October do not precede any major temperature differences.

The same can be seen with the difference in the zonal winds for both events. Figure 5.5 shows that for the October event, the wind differences are about 5 m/s, and the event do not trigger any big difference in the zonal winds. On the contrary, figure 4.18 shows that from the 15th of January until the 20th, wind differences between 15 and 20 m/s, meaning that MEE had a bigger impact on the dynamics in January than in October, even when the IPR due to the MEE was one order of magnitude higher during the October event.

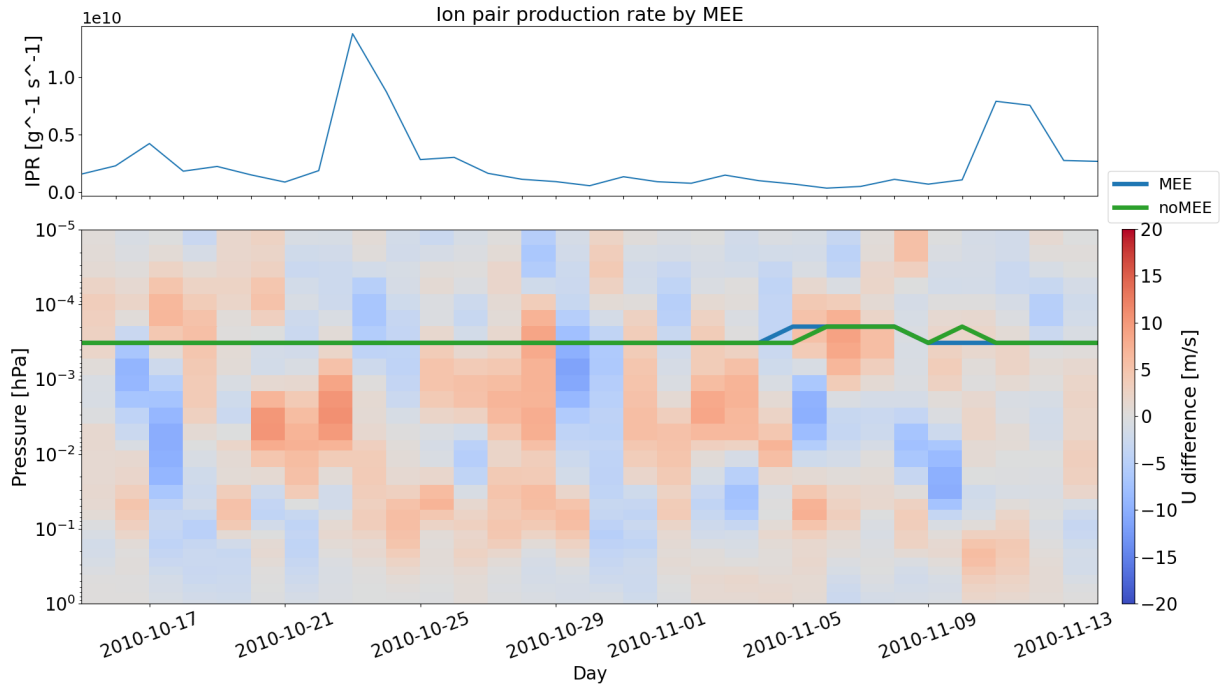

Figure 5.5: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the absolute difference in the zonal wind velocity between the MEE run and the noMEE run from the 15th of October to the 15th of November, in the northern hemisphere. The red color represents more eastward winds in the MEE run, while blue represents more westward winds. The data is averaged over the latitude band 60 - 70^{o}N with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

5.3 Interhemispheric Coupling

Smith et al. [2020] explored the interhemispheric coupling associated to dynamically active winter conditions and the associated response on the opposite hemisphere using WACCM simulations. They suggested that the rapid circulation response in the summer hemisphere in both the stratosphere and the mesosphere are consistent with a circulation change induced to restore zonal-mean balance to the atmosphere. To evaluate a potential interhemispheric link, Figure 5.6 shows the temperature difference between the MEE and noMEE runs across all latitudes for the 26 th of January. This is the first day that the dipole-like temperature differences are seen in figure 4.19. The dipole like feature extend from the northern hemisphere pole to 50^{circ} N latitude, where it is shifted to higher latitudes following the mesopause altitude. Comparing this plot with figure 2.16 similarities are found. The pattern in the northern latitudes of warmer temperatures (positive correlations in figure 2.16) at the higher altitudes, and colder temperatures (negative correlations in figure 2.16) at the lower altitudes, are matched by temperatures of opposite sign extending from the winter hemisphere to the summer hemisphere. For day 0, the correlations are in figure

2.16 not so strong, and it can be seen in figure 5.6 that even if the temperature differences at the higher latitudes are of about 5 K, the warmer temperatures extending to the summer hemisphere are of about 2 K at altitudes up to 10^{-1} hPa, and the colder temperatures between 10^{-2} and 10^{-3} hPa are also about 2K. These temperatures only reach at about 50^{o}S and 0^{o} in latitude, respectively.

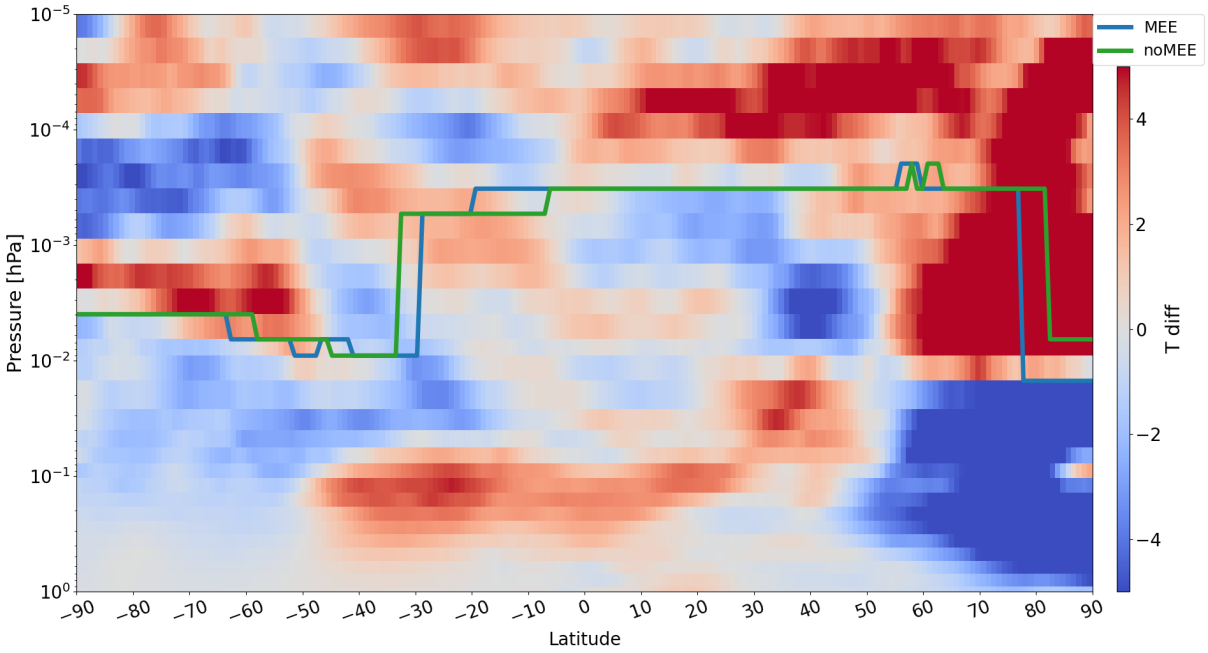

Figure 5.6: Temperature difference between the MEE run and the noMEE run for the 26th of January. The red color represents warmer temperatures in the MEE run, while blue represents colder temperatures. The data is plotted for each pressure level across all latitudes with a daily resolution. The blue line represents the mesopause altitude (coldest altitude) for the MEE run, while the green line represents the mesopause for the noMEE run.

Figure 5.7 shows the latitudinal temperature difference between the MEE and noMEE run three days later, the 29th of January. In figure 2.16 the correlation maximizes at about three days. Analogue, Figure 5.7 shows that the colder temperatures in the upper mesosphere extend to the summer hemisphere all the way to 80^{o}S, accompanied by warmer temperatures in the high latitudes of the northern hemisphere. These colder temperatures are confined between 10^{-1} and 10^{-3} hPa for the midlatitudes, while for the high latitudes they are found between 10^{-1} and 10^{-2} hPa, as well as between 10^{-3} and 10^{-4} hPa. The upper limit in Figure 2.16 is roughly 3×10^{-4} hPa, and for the high latitudes in the northern hemisphere a negative correlation is present above roughly 4×10^{-4} hPa. Figure 5.7 show colder temperatures in the northern hemisphere's high latitudes above 4×10^{-4} hPa, which would correspond to negative correlations in figure 2.16. These colder temperatures are also accompanied by warmer temperatures extending all the way to 90^{o}S above 10^{-4} hPa.

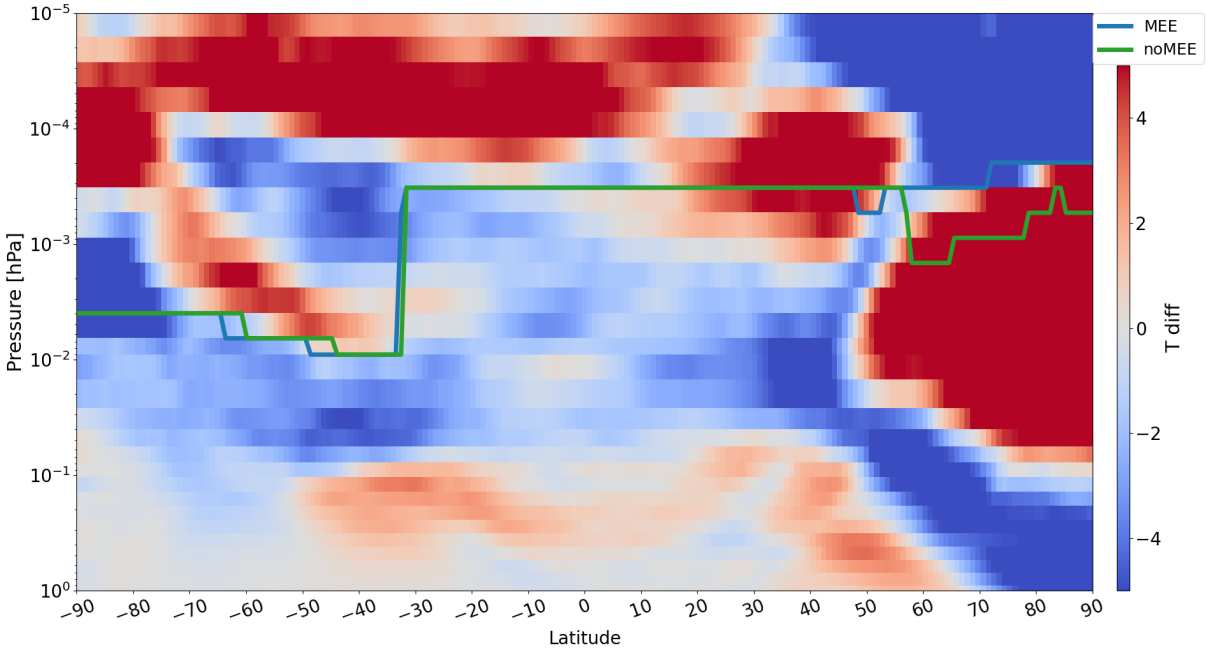

Figure 5.7: Temperature difference between the MEE run the noMEE run for the 29th of January. The red color represents warmer temperatures in the MEE run, while blue represents colder temperatures. The data is plotted for each pressure level across all latitudes with a daily resolution. The blue line represents the mesopause altitude (coldest altitude) for the MEE run, while the green line represents the mesopause for the noMEE run.

Figure 5.8 shows the latitudinal temperature differences between the MEE and noMEE run for the 4th of February. Here, large temperature differences associated with the different SSW characteristics in the two runs are now confined from 70ºN to 90ºN and from roughly 10^{-1} to 10^{-5} hPa. It can also be seen that the middle latitudes do not show any sign of interhemispheric coupling. However, the temperature differences in the southern hemisphere are still persisting from 40ºS to 90ºS. Colder temperatures in the MEE run can be seen between 3×10^{-2} to 2×10^{-3} hPa, while warmer temperatures can be seen between 2×10^{-3} to 3×10^{-5} hPa. This is in good agreement with *Smith et al.* [2020], where it is found that the temperature perturbations persist for tens of days.

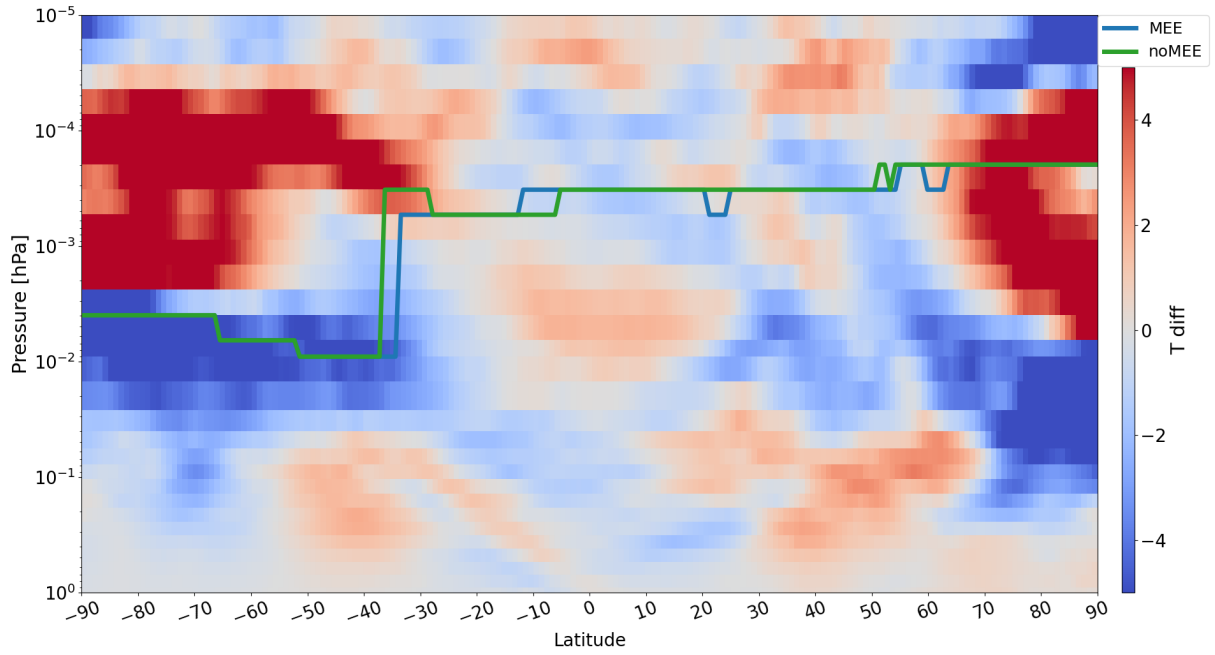

Figure 5.8: Temperature difference between the MEE run and the noMEE run for the 4th of February. The red color represents warmer temperatures in the MEE run, while blue represents colder temperatures. The data is plotted for each pressure level across all latitudes with a daily resolution. The blue line represents the mesopause altitude (coldest altitude) for the MEE run, while the green line represents the mesopause for the noMEE run.

5.4 Seasonal transition

The onset of the seasonal shift in the southern hemisphere is closely related to the transition of winds in the stratosphere [*Lübken et al.*, 2015]. Hence, as the respective WACCM runs are nudged by reanalysis data, the discrepancies found in terms of the seasonal shift in the southern hemisphere are related to dynamical changes induced in the mesosphere when including the MEE IPR. The temperature differences plotted in figure 4.23 show persistently colder temperatures below the mesopause. This can have an impact on the dynamics of the atmosphere. Colder temperatures near the mesopause region would mean that the temperature gradient is steeper in the MEE run, since the mesosphere gets colder with altitude.The temperature gradient change will have an effect on the wave breaking level, forcing the waves to break at a different altitude on the MEE run.

Figure 5.9 shows the velocity difference in the zonal wind between the MEE and noMEE runs. In the period from the 3rd of February until the 3rd of March, when the mesopause moves to a higher altitude, the winds are persistently more eastward just below the mesopause in the MEE run. Also for roughly the same period, the winds are more westerly above the mesopause for the MEE run. Since the gravity wave forcing from the

stratosphere are the same in both runs, the change in zonal wind supports the idea of a change in the wave forcing.

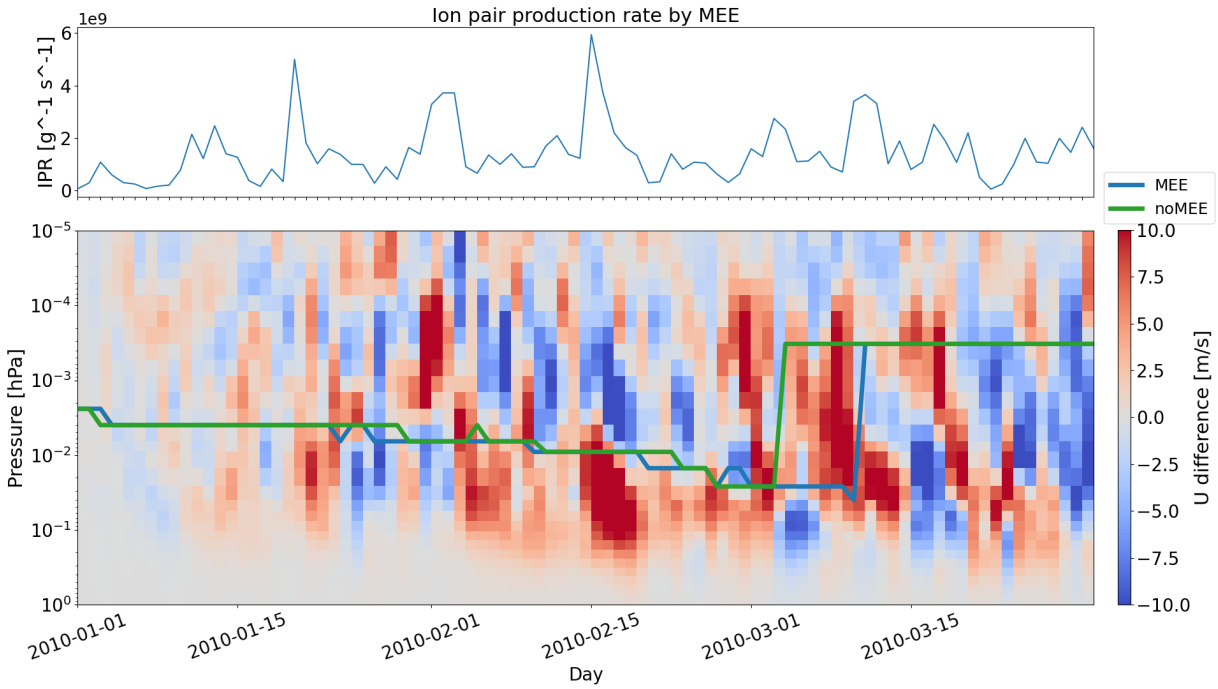

Figure 5.9: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the absolute difference in the zonal wind velocity between the MEE run and the noMEE run for the first three months of the year 2010, in the southern hemisphere. The red color represents more eastward winds in the MEE run, while blue represents more westward winds. The data is averaged over the latitude band 60 - 70^{o}S with a daily resolution. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Smith-Johnsen et al. [submitted to JGR] shows that a change in the amplitude of the non-orographic gravity waves had an impact in the mesosphere/lower thermosphere temperature as well as in the timing of the mesopause seasonal shift. For an increased amplitude, the waves will break at a lower level, and it has an associated cooling of the atmosphere below the mesopause and a warming above the mesopause, as seen in figure 2.17. In figure 4.23, the cooling below the mesopause can be seen, but the warming above the mesopause is not as persistent. Furthermore, *Smith-Johnsen et al.* [submitted to JGR] shows that increasing the amplitudes of the waves displaces the mesopause to a lower altitude before the seasonal shift happens, which is sooner than in the control run. Figure 4.23 shows, however, that the altitude of the mesopause stays at about the same level for both runs, and the seasonal shift for the MEE run happens later in time.

This could be due to the fact that in *Smith-Johnsen et al.* [submitted to JGR], the amplitude was increased for a factor of 5, causing temperature changes up to 40 K. For

the study in this thesis though, the temperature changes are only about 5 K, making the mesopause shift less dramatic in current study. Another difference is that in *Smith-Johnsen et al.* [submitted to JGR], the altitude of wave breaking changes from day 1, while for the case in this thesis, the wave altitude seem to not be changed until the 2nd of February. Hence, the lack of a mesopause altitude shift can be due to a shorter duration of the forcing and smaller temperature changes compared to *Smith-Johnsen et al.* [submitted to JGR]. Furthermore, figure 2.17 shows that the warmer temperatures above the mesopause follows the mesopause altitude. For the case of study in this thesis, these warmer temperatures are not present at the time of the seasonal shift, and the temperatures are colder at the mesopause altitude.

5.5 WACCM

WACCM6 is the model used for this study. It is a useful tool to disentangle the impact of chemical production versus transport, as well as the impact due to different production sources as discussed in the previous sections. It does, however, also present some limitations for the study. In this section, limitations of WACCM6 are discussed.

5.5.1 Nudging

As mentioned in section 3.2, the runs in this thesis are made with the Specified Dynamics version of WACCM6. This means that winds and temperatures are nudged by reanalysis data from the surface up to 50 km, which roughly corresponds to the lower limit of the plots presented, 10^{-1} hPa. Above this point, there's a transition region up to 60 km, and the model is free running from 60 km to the upper limit of WACCM, which is the upper limit of the plots at 10^{-5} hPa.

Nudging is a good way to make sure that the dynamics of both runs are the same in the stratosphere, meaning that wave propagation is the same up to 50 km, and any changes found between the runs, in both dynamics and chemistry, is mainly due to changes in the mesosphere. That said, for the studied period there is a SSW causing wind reversals in the northern hemisphere. The SSW is different in both runs, and as can be seen in figure 4.17, the winds in the MEE run turn back to normal on the 25th of January, as if the SSW was ending there. Two days later though, the wind reversal linked to a SSW can be seen again, until the SSW ends for both runs on the 2nd of January.

This is presumably due to the fact that both runs have the same dynamics in the stratosphere, and for this period of the year there's a SSW. That means that even if the wind reversal is stopping in the mesosphere, the wave filtering in the stratosphere is still influenced by the SSW, and eventually this wave forcing causes the wind reversal to be forced again. So, if the nudging of the stratosphere was not present, could the MEE have completely stopped the SSW?

5.5.2 Noise

Figure 5.10 shows the relative difference in OH between both runs. Figure 4.21 showed the absolute difference, and it could be seen that the biggest OH changes were confined into a narrow altitude band, ranging from 5×10^{-2} to 5×10^{-3} hPa. The differences do not show a clear signal of OH being more abundant in one of the runs, so could it be noise?

In figure 5.10, the OH relative differences for this altitude band are about 10 %. The data used is daily-averaged, while the lifetime of OH in the mesosphere is only about hours. Could it imply that we are looking for changes that are not seen in the daily resolved parameters?

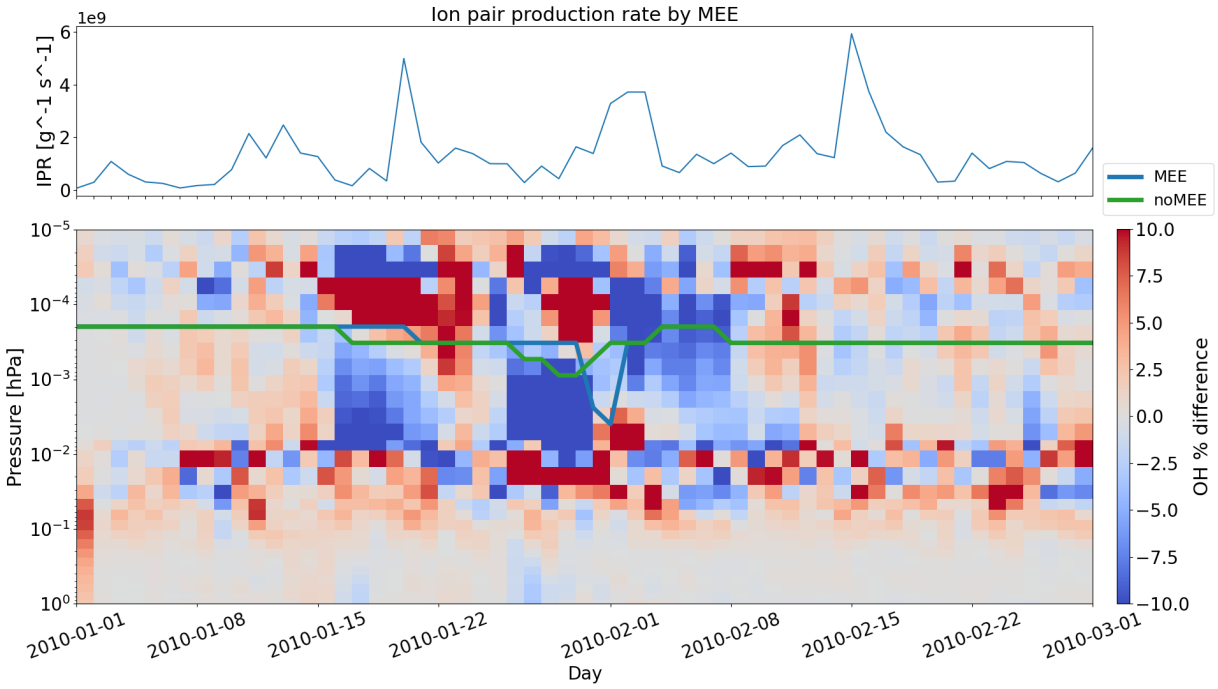

Figure 5.10: The upper plot corresponds to the ion production rate by the MEE. The lower plot shows the OH relative difference, in the northern hemisphere, between the MEE run and the noMEE run for the first two months of the year 2010. The color red represents more OH being present in the MEE run, on the contrary, blue represents more OH being present in the noMEE run. The data is averaged over the latitude band 60 - 70ºN and it is also daily-averaged. The two colored lines represent the mesopause altitude (coldest altitude) for both runs, the blue line corresponds the MEE run, and the green line to the noMEE run.

Chapter 6

Conclusions

The goal of this thesis was to study the effects that EEP have in WACCM. The main focus was set in the production and transport of NO. To do so, two WACCM6 runs were used, both of them presented extra chemistry in the D-region, as well as a nudged stratosphere with reanalysis data. The only difference between the runs was the EEP ionization sources; one run included both auroral electrons and MEE (MEE run), while the other run only included auroral electrons (noMEE run). Observations from the instrument SOFIE were used in order to compare the NO production and transport.

In the first study, the NO VMR in the MEE run from WACCM and SOFIE observations are compared. WACCM overestimates NO during quiet geomagnetic times, while the response to geomagnetic activity is underestimated. To remove the background NO, which is too high in WACCM, as well as the seasonal trend, a 30-day running mean is subtracted from both datasets. This improves NO during quiet geomagnetic times, but it is still underestimating the NO response to geomagnetic activity, specially at lower altitudes. To further analyse the production and transport, the two strongest geomagnetic events with available data are studied. The events show that NO is slightly underestimated in WACCM and lingers for a longer period at high altitudes, maybe due to the upper boundary conditions or to insufficient transport across the mesopause. At lower altitudes, NO is underestimated, maybe due to weak transport or to the underestimated NO reservoir at high altitudes. Regardless, the speed of the estimated transport correspond well with observations. In the discussion, further evidence of the similar speed in transport between WACCM and SOFIE is presented, following the same two events from 100 km to 85 km.

In the second study, the two separate WACCM runs are compared. A time period with more NO throughout the mesosphere in the noMEE run was found in the northern hemisphere. This was counter-intuitive since the MEE run should produce more NO in the mesosphere. This period of time was found to correspond to a SSW, and dynamical differences were found between the SSW in the MEE run and the noMEE run. Seemingly,

temperature differences in the mesosphere due to MEE, changed the wave forcing in the MEE run, leading to the dynamical differences in the SSW. These differences stopped the wind reversal in the middle of the SSW in the MEE run, leading to the enhanced transport of NO sooner than in the noMEE run. Despite that the differences in the SSW are found in the northern hemisphere, evidence that they propagate into the southern hemisphere is found in section 5.3. These perturbations in the southern hemisphere persist for days after their first appearance.

These differences in the SSW, happen in a period with low geomagnetic activity, but with unstable atmospheric mean-flow conditions leading to the SSW. The EEP dynamical impact is enhanced by the unstable atmospheric conditions. To test this, the biggest geomagnetic event during winter in the northern hemisphere is studied, and it is found that the dynamics of the atmosphere do not change.

In the southern hemisphere, it was found that the seasonal shift of the mesopause happens with seven days delay in the MEE compared to the noMEE run. Persistently colder temperatures just below the mesopause are seen in the MEE run, leading to the mesopause shifting later in time.

The existing theories on how EEP can affect the dynamics of the atmosphere are mainly lined with stratospheric dynamics. The findings on this thesis suggest that EEP can change the dynamics of the atmosphere without reaching the stratosphere. In this study, the stratosphere was nudged with reanalysis data, and EEP only affected the mesosphere and lower thermosphere. Regardless, evidence is found that adding MEE changed the dynamics of the atmosphere, meaning that the enhanced EEP in the mesosphere lead to the dynamical changes.

These changes were possible due to unstable atmospheric wind-flow conditions, as they happened in a period with low geomagnetic activity. On the contrary, the biggest geomagnetic event in winter did not have any impact in the dynamics during stable atmospheric conditions. The temperature differences caused by the dynamical changes in the northern hemisphere high latitudes propagate all the way to the southern hemisphere high latitudes and persist for days.

In summary, this thesis has shown that:

- NO VMR in WACCM are overestimated during quiet geomagnetic times.
- NO VMR response in WACCM to geomagnetic events are underestimated throughout the upper mesosphere lower thermosphere region.
- The NO transport speed is similar in WACCM and SOFIE.
- MEE can directly change the dynamics of the mesosphere without reaching the stratosphere.
- Unstable atmospheric conditions, as the period pre-SSW, allow MEE to have a bigger dynamical impact.

- Differences in the dynamics due to the MEE propagate into the opposite hemisphere.
- Seasonal transition of the mesopause is affected by MEE.

Chapter 7

Future Work

The results of this thesis, show that adding MEE in WACCM runs can have an effect on the dynamics of the mesosphere even without changing the dynamics in the stratosphere. However, the chain of events is not completely understood, partly due to the fact that the wave forcing parameter was not available. Furthermore, only one case have been studied in this work.

Therefore, a future investigation should include an ensemble of WACCM runs with and without MEE from the same year would ensure that the features found in this work are persistent. Furthermore, having access to the wave forcing parameter in those runs would also help to completely understand how the mechanism leading to the dynamical changes works.

The next step would be to identify periods of SSW from other years and study how MEE affect those cases. This could also be helpful in understanding how strong is the link between geomagnetic activity and dynamical perturbations during unstable atmospheric conditions. These periods could also be used to study if the seasonal transition happens later in the MEE run throughout the different years, and understand why this happens.

Ideally, artificially scaled runs could be use to change the ionization rates during unstable atmospheric conditions and see if the impact on the mesospheric dynamics is proportional to the forcing. Also, runs without using the specified dynamics version, i.e. without nudging the stratosphere, could be used to see if MEE would impact the SSW signature also in the stratosphere.

In addition, the same study could be done in other atmospheric models and see how they respond to it. For example in SOCOL (SOlar Climate Ozone Links) in ECHAM5, which is a chemistry-climate model covering from the surface up to the mesopause. In this model, MEE ionization is directly represented, while the thermospheric NO is parametrized and forced through the model top boundary, leaving the possibility to "turn off" auroral

electrons ionization.

In summary:

- Bigger ensemble of WACCM runs with and without MEE and study the same cases to make sure that the features are persistent throughout all the runs.
- Find SSW from other years and study the effect that MEE has in WACCM.
- Study without nudging and see if MEE can impact the SSW also in the stratosphere.
- Ideal, scaled runs to investigate how the dynamics will respond to e.g. twice as strong ionization rates.
- Same study in other models, such as SOCOL, where thermospheric NO is parametrized.

Acronyms

AIM Aeronomy of Ice in the Mesosphere.

CAM Community Atmosphere Model.

CESM Community Earth System Model.

CIR Corotating Interacting Regions.

CME Coronal Mass Ejections.

EEP Energetic Electron Precipitation.

EP Eliassen-Palm.

EPP Energetic Particle Precipitation.

HP Hemispheric Power.

IMF Interplanetary Magnetic Field.

IPR Ion-pair Production Rate.

MEE Medium Energy Electrons.

MEPED Medium Energy Proton/Electron Detector.

MERRA Modern-Era Retrospective Analysis for Research and Applications.

MLT Mesosphere and Lower Thermosphere.

NCAR National Center for Atmospheric Research.

NOAA National Oceanic and Atmospheric Administration.

PMC Polar Mesospheric Clouds.

POES Polar Orbiting Environmental Satellites.

QBO Quasi-Biennial Oscillation.

SIC Sodankylä Ion and Neutral Chemistry.

SNOE Student NO Experiment.

SOCOL SOlar Climate Ozone Links.

SOFIE Solar Occultation For Ice Experiment.

SSW Sudden Stratospheric Warming.

TEM Transformed Eulerian Mean.

VMR Volume Mixing Ratio.

WACCM Whole Atmosphere Community Climate Model.

Bibliography

Asikainen, T., A. Salminen, V. Maliniemi, and K. Mursula (2020), Influence of enhanced planetary wave activity on the polar vortex enhancement related to energetic electron precipitation, *Journal of Geophysical Research: Atmospheres*, *125*(9), e2019JD032,137, doi:https://doi.org/10.1029/2019JD032137, e2019JD032137 10.1029/2019JD032137. 27, 65

Babcock, H. (1961), The Topology of the Sun's Magnetic Field and the 22-YEAR Cycle., *apj*, *133*, 572, doi:10.1086/147060. 4

Bartels, J., N. H. Heck, and H. F. Johnston (1939), The three-hour-range index measuring geomagnetic activity, *Terrestrial Magnetism and Atmospheric Electricity*, *44*(4), 411–454, doi:https://doi.org/10.1029/TE044i004p00411. 13

Baumgaertner, A. J. G., A. Seppälä, P. Jöckel, and M. A. Clilverd (2011), Geomagnetic activity related no$_x$ enhancements and polar surface air temperature variability in a chemistry climate model: modulation of the nam index, *Atmospheric Chemistry and Physics*, *11*(9), 4521–4531, doi:10.5194/acp-11-4521-2011. 20

Baumjohann, W., and R. Treumann (2012), *Basic Space Plasma Physics - Revised Edition*, doi:10.1142/P850. 5, 9

Becker, E., and D. C. Fritts (2006), Enhanced gravity-wave activity and interhemispheric coupling during the macwave/midas northern summer program 2002, *Annales Geophysicae*, *24*(4), 1175–1188, doi:10.5194/angeo-24-1175-2006. 23

Becker, E., A. Müllemann, F.-J. Lübken, H. Körnich, P. Hoffmann, and M. Rapp (2004), High rossby-wave activity in austral winter 2002: Modulation of the general circulation of the mlt during the macwave/midas northern summer program, *Geophysical Research Letters*, *31*(24), doi:https://doi.org/10.1029/2004GL019615. 23

Brasseur, G., and S. Solomon (2005), *Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere*, doi:10.1007/1-4020-3824-0. 1, 14, 15, 16, 17, 18, 19, 56

Brewer, A. W. (1949), Evidence for a world circulation provided by the measurements of helium and water vapour distribution in the stratosphere, *Quarterly Journal of the Royal Meteorological Society*, *75*(326), 351–363, doi:https://doi.org/10.1002/qj.49707532603. 16

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, *Journal of Geophysical Research*, *75*, 50–56. 10

Dobson, G. M. B., D. N. Harrison, and J. Lawrence (1929), Measurements of the amount of ozone in the earth’s atmosphere and its relation to other geophysical conditions.—part iii, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *122*(790), 456–486, doi:10.1098/rspa.1929.0034. 16

Dörnbrack, A., M. C. Pitts, L. R. Poole, Y. J. Orsolini, K. Nishii, and H. Nakamura (2012), The 2009–2010 arctic stratospheric winter – general evolution, mountain waves and predictability of an operational weather forecast model, *Atmospheric Chemistry and Physics*, *12*(8), 3659–3675, doi:10.5194/acp-12-3659-2012. 52

Dungey, J. W. (1961), Interplanetary magnetic field and the auroral zones, *Phys. Rev. Lett.*, *6*, 47–48, doi:10.1103/PhysRevLett.6.47. 8

Ganushkina, N. Y., I. Dandouras, Y. Y. Shprits, and J. Cao (2011), Locations of boundaries of outer and inner radiation belts as observed by cluster and double star, *Journal of Geophysical Research: Space Physics*, *116*(A9), doi:https://doi.org/10.1029/2010JA016376. 10

Gettelman, A., M. J. Mills, R. R. Kinnison, D. E.and Garcia, A. K. Smith, D. R. Marsh, S. Tilmes, F. Vitt, C. G. Bardeen, J. McInerny, H.-L. Liu, S. C. Solomon, L. M. Polvani, L. K. Emmons, J.-F. Lamarque, J. H. Richter, A. S. Glanville, J. T. Bacmeister, A. S. Phillips, R. B. Neale, I. R. Simpson, A. K. DuVivier, A. Hodzic, and W. J. Randel (2019), The whole atmosphere community climate model version 6 (WACCM6), *J. Geophys. Res. Atmos.*, *124*, 12,380–12,403, doi:10.1029/2019JD030943. 31

Gómez-Ramírez, D., J. W. C. McNabb, J. M. Russell, M. E. Hervig, L. E. Deaver, G. Paxton, and P. F. Bernath (2013), Empirical correction of thermal responses in the solar occultation for ice experiment nitric oxide measurements and initial data validation results, *Appl. Opt.*, *52*(13), 2950–2959, doi:10.1364/AO.52.002950. 31

Gordley, L. L., M. E. Hervig, C. Fish, J. M. Russell, S. Bailey, J. Cook, S. Hansen, A. Shumway, G. Paxton, L. Deaver, T. Marshall, J. Burton, B. Magill, C. Brown, E. Thompson, and J. Kemp (2009), The solar occultation for ice experiment, *Journal Atmospheric and Solar-Terrestrial Physics*, *71*, doi:10.1016/j.jastp.2008.07.012. 30

Hendrickx, K., L. Megner, D. R. Marsh, and C. Smith-Johnsen (2018), Production and transport mechanisms of no in the polar upper mesosphere and lower thermosphere in observations and models, *Atmospheric Chemistry and Physics*, *18*(12), 9075–9089, doi:10.5194/acp-18-9075-2018. 2, 25, 63, 64

Holt, L. A., C. E. Randall, E. D. Peck, D. R. Marsh, A. K. Smith, and V. L. Harvey (2013), The influence of major sudden stratospheric warming and elevated stratopause events on the effects of energetic particle precipitation in waccm, *Journal of Geophysical Research: Atmospheres*, *118*(20), 11,636–11,646, doi:https://doi.org/10.1002/2013JD020294. 22

Horne, R., M. M. Lam, and J. Green (2009), Energetic electron precipitation from the outer radiation belt during geomagnetic storms, *Geophys. Res. Lett*, *36*, doi:10.1029/2009GL040236. 19

Hughes, W. J. (1995), The magnetopause, magnetotail, and magnetic reconnection, *Introduction to Space Physics*, pp. 227–287. 8

Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall (2013), The community earth system model: A framework for collaborative research, *Bulletin of the American Meteorological Society*, *94*, doi:10.1175/BAMS-D-12-00121.1. 31

Karlsson, B., and E. Becker (2016), How does interhemispheric coupling contribute to cool down the summer polar mesosphere?, *Journal of Climate*, *29*(24), 8807 – 8821, doi:10.1175/JCLI-D-16-0231.1. 23

Karlsson, B., C. McLandress, and T. G. Shepherd (2009), Inter-hemispheric mesospheric coupling in a comprehensive middle atmosphere model, *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*(3), 518–530, doi:https://doi.org/10.1016/j.jastp.2008.08.006, global Perspectives on the Aeronomy of the Summer Mesopause Region. 25

Kunz, A., L. L. Pan, P. Konopka, D. E. Kinnison, and S. Tilmes (2011), Chemical and dynamical discontinuity at the extratropical tropopause based on start08 and waccm analyses, *Journal of Geophysical Research: Atmospheres*, *116*(D24), doi:https://doi.org/10.1029/2011JD016686. 31

Kyoto-University (2021), Auroral electrojet indices, `http://wdc.kugi.kyoto-u.ac.jp/aedir/ae2/onAEindex.html`, last accessed on 2021-08-11. 13

Labitzke, K. (1981), Stratospheric-mesospheric midwinter disturbances: A summary of observed characteristics, *Journal of Geophysical Research: Oceans*, *86*(C10), 9665–9678, doi:https://doi.org/10.1029/JC086iC10p09665. 21

Lary, D. J. (1997), Catalytic destruction of stratospheric ozone, *Journal of Geophysical Research: Atmospheres*, *102*(D17), 21,515–21,526, doi:https://doi.org/10.1029/97JD00912. 57

Laundal, K. M., and A. D. Richmond (2017), Magnetic coordinate systems, *Space Science Reviews*, *206*, 27–59, doi:10.1007/s11214-016-0275-y. 7

Limpasuvan, V., J. Richter, Y. Orsolini, F. Stordal, and O.-K. Kvissel (2012), The roles of planetary and gravity waves during a major stratospheric sudden warming as characterized in waccm, *Journal of Atmospheric and Solar-Terrestrial Physics*, *78-79*, doi:10.1016/j.jastp.2011.03.004. 21

Lübken, F.-J., J. Höffner, T. P. Viehl, E. Becker, R. Latteck, B. Kaifler, D. J. Murphy, and R. J. Morris (2015), Winter/summer transition in the antarctic mesopause region, *Journal of Geophysical Research: Atmospheres*, *120*(24), 12,394–12,409, doi:https://doi.org/10.1002/2015JD023928. 71

Maliniemi, V., T. Asikainen, K. Mursula, and A. Seppälä (2013), Qbo-dependent relation between electron precipitation and wintertime surface temperature, *Journal of Geophysical Research: Atmospheres*, *118*(12), 6302–6310, doi:https://doi.org/10.1002/jgrd.50518. 26

Maliniemi, V., T. Asikainen, and K. Mursula (2016), Effect of geomagnetic activity on the northern annular mode: Qbo dependence and the holton-tan relationship, *Journal of Geophysical Research: Atmospheres*, *121*(17), 10,043–10,055, doi:https://doi.org/10.1002/2015JD024460. 1

Marsh, D. R., S. C. Solomon, and A. E. Reynolds (2004), Empirical model of nitric oxide in the lower thermosphere, *Journal of Geophysical Research: Space Physics*, *109*(A7), doi:https://doi.org/10.1029/2003JA010199. 32

Meraner, K., and H. Schmidt (2016), Transport of nitrogen oxides through the winter mesopause in hammonia, *Journal of Geophysical Research: Atmospheres*, *121*(6), 2556–2570, doi:https://doi.org/10.1002/2015JD024136. 21

Meraner, K., and H. Schmidt (2018), Climate impact of idealized winter polar mesospheric and stratospheric ozone losses as caused by energetic particle precipitation, *Atmospheric Chemistry and Physics*, *18*(2), 1079–1089, doi:10.5194/acp-18-1079-2018. 65

Meraner, K., H. Schmidt, E. Manzini, B. Funke, and A. Gardini (2016), Sensitivity of simulated mesospheric transport of nitrogen oxides to parameterized gravity waves, *Journal of Geophysical Research: Atmospheres*, *121*(20), 12,045–12,061, doi:https://doi.org/10.1002/2016JD025012. 21, 23

Neale, R. B., J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang (2013), The mean climate of the community atmosphere model (cam4) in forced sst

and fully coupled experiments, *Journal of Climate*, *26*(14), 5150 – 5168, doi:10.1175/JCLI-D-12-00236.1. 31

Parker, E. N. (1958), Dynamics of the Interplanetary Gas and Magnetic Fields, *Astrophysical Journal*, *128*, 664. 6

Parks, G. K. (2004), *Physics of space plasmas. An introduction.*, Advanced Book Program. 5

Plumb, R. (2002), Stratospheric transport, *Journal of the Meteorological Society of Japan*, *80*, doi:10.2151/jmsj.80.793. 15

Randall, C. E., V. L. Harvey, C. S. Singleton, S. M. Bailey, P. F. Bernath, M. V. Codrescu, H. Nakajima, , and J. M. R. III (2007), Energetic particle precipitation effects on the southern hemisphere stratosphere in 1992-2005, *Journal of Geophysical Research - Atmospheres*, *112*, doi:10.1029/2006JD007696. 1, 19

Randall, C. E., V. L. Harvey, D. E. Siskind, J. France, P. F. Bernath, C. D. Boone, and K. A. Walker (2009), Nox descent in the arctic middle atmosphere in early 2009, *Geophysical Research Letters*, *36*(18), doi:https://doi.org/10.1029/2009GL039706. 22

Randall, C. E., V. L. Harvey, L. A. Holt, D. R. Marsh, D. Kinnison, B. Funke, and P. F. Bernath (2015), Simulation of energetic particle precipitation effects during the 2003–2004 arctic winter, *Journal of Geophysical Research: Space Physics*, *120*(6), 5035–5048, doi:https://doi.org/10.1002/2015JA021196. 63

Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, S. Bloom, J. Chen, D. Collins, A. Conaty, A. da Silva, W. Gu, J. Joiner, R. D. Koster, R. Lucchesi, A. Molod, T. Owens, S. Pawson, P. Pegion, C. R. Redder, R. Reichle, F. R. Robertson, A. G. Ruddick, M. Sienkiewicz, and J. Woollen (2011), Merra: Nasa's modern-era retrospective analysis for research and applications, *Journal of Climate*, *24*(14), 3624 – 3648, doi:10.1175/JCLI-D-11-00015.1. 31

Roble, R. G., E. C. Ridley, and R. E. Dickinson (1987), On the global mean structure of the thermosphere, *Journal of Geophysical Research: Space Physics*, *92*(A8), 8745–8758, doi:https://doi.org/10.1029/JA092iA08p08745. 32

Russell, J., S. Bailey, L. Gordley, D. Rusch, M. Horányi, M. Hervig, G. Thomas, C. Randall, D. Siskind, M. Stevens, M. Summers, M. Taylor, C. Englert, P. Espy, W. McClintock, and A. Merkel (2009), The aeronomy of ice in the mesosphere (aim) mission: Overview and early science results, *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*, 289–299, doi:10.1016/j.jastp.2008.08.011. 29

Salminen, A., T. Asikainen, V. Maliniemi, and K. Mursula (2020), Dependence of sudden stratospheric warmings on internal and external drivers, *Geophysical Research Letters*,

47(5), e2019GL086,444, doi:https://doi.org/10.1029/2019GL086444, e2019GL086444 10.1029/2019GL086444. 26, 65

Seppälä, A., H. Lu, M. A. Clilverd, and C. J. Rodger (2013), Geomagnetic activity signatures in wintertime stratosphere wind, temperature, and wave response, *Journal of Geophysical Research: Atmospheres*, *118*(5), 2169–2183, doi:https://doi.org/10.1002/jgrd.50236. 1, 20, 65

Smith, A. K., R. R. Garcia, D. R. Marsh, and J. H. Richter (2011), Waccm simulations of the mean circulation and trace species transport in the winter mesosphere, *Journal of Geophysical Research: Atmospheres*, *116*(D20), doi:https://doi.org/10.1029/2011JD016083. 20, 21, 22, 53

Smith, A. K., N. M. Pedatella, and Z. K. Mullen (2020), Interhemispheric coupling mechanisms in the middle atmosphere of waccm6, *Journal of the Atmospheric Sciences*, *77*(3), 1101 – 1118, doi:10.1175/JAS-D-19-0253.1. 23, 24, 25, 58, 60, 68, 70

Smith-Johnsen, C., D. Marsh, Y. Orsolini, H. Nesse Tyssøy, K. Hendrickx, M. Sandanger, L.-K. Glesnes Ødegaard, and F. Stordal (2018), Nitric oxide response to the april 2010 electron precipitation event: Using waccm and waccm-d with and without medium-energy electrons, *J. Geophys. Res.*, *123*, 5232–5245, doi:10.1029/2018JA025418. 2, 25, 63

Smith-Johnsen, C., D. R. Marsh, A. K. Smith, H. Nesse Tyssøy, and V. Maliniemi (submitted to JGR), Mesospheric nitric oxide transport in waccm, *J. Geophys. Res. Space Physics*. 26, 72, 73

Thorne, R. M. (1980), The importance of energetic particle precipitation on the chemical composition of the middle atmosphere, *pure and applied geophysics*, doi:10.1007/BF01586448. 12

Tsurutani, B. T., W. D. Gonzalez, A. L. C. Gonzalez, F. L. Guarnieri, N. Gopalswamy, M. Grande, Y. Kamide, Y. Kasahara, G. Lu, I. Mann, R. McPherron, F. Soraas, and V. Vasyliunas (2006), Corotating solar wind streams and recurrent geomagnetic activity: A review, *Journal of Geophysical Research: Space Physics*, *111*(A7), doi:https://doi.org/10.1029/2005JA011273. 11

Turunen, E., P. T. Verronen, A. Seppälä, C. J. Rodger, M. A. Clilverd, J. Tamminen, C.-F. Enell, and T. Ulich (2009), Impact of different energies of precipitating particles on nox generation in the middle and upper atmosphere during geomagnetic storms, *Journal of Atmospheric and Solar-Terrestrial Physics*, *71*(10), 1176–1189, doi:https://doi.org/10.1016/j.jastp.2008.07.005, high Speed Solar Wind Streams and Geospace Interactions. 19, 48

van de Kamp, M., A. Seppälä, M. A. Clilverd, C. J. Rodger, P. T. Verronen, and I. C. Whittaker (2016), A model providing long-term data sets of energetic electron precipitation

during geomagnetic storms, *J. Geophys. Res. Atmos.*, *121*, doi:10.1002/2015JD024212. 32, 33, 48

Verronen, P. T., M. E. Andersson, D. R. Marsh, T. Kovács, and J. M. C. Plane (2016), Waccm-d—whole atmosphere community climate model with d-region ion chemistry, *Journal of Advances in Modeling Earth Systems*, *8*(2), 954–975, doi:10.1002/2015MS000592. 2, 32, 63