Energetic particle induced intra-seasonal variability of ozone inside the Antarctic polar vortex observed in satellite data

T. Fytterer1, M. G. Mlynczak2, H. Nieder1, K. Pérot3, M. Sinnhuber1, G. Stiller1, and J. Urban3,†

1Institute for Meteorology and Climate Research, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany
2Atmospheric Sciences Division, NASA Langley Research Center, Hampton, VA, USA
3Department of Earth and Space Sciences, Chalmers University of Technology, Göteborg, Sweden

†deceased, 14 August 2014

Correspondence to: T. Fytterer (tilo.fytterer@kit.edu)

Received: 19 October 2014 – Published in Atmos. Chem. Phys. Discuss.: 11 December 2014
Revised: 19 February 2015 – Accepted: 6 March 2015 – Published: 24 March 2015

Abstract. Measurements from 2002 to 2011 by three independent satellite instruments, namely MIPAS, SABER, and SMR on board the ENVISAT, TIMED, and Odin satellites are used to investigate the intra-seasonal variability of stratospheric and mesospheric O3 volume mixing ratio (vmr) inside the Antarctic polar vortex due to solar and geomagnetic activity. In this study, we individually analysed the relative O3 vmr variations between maximum and minimum conditions of a number of solar and geomagnetic indices (F10.7 cm solar radio flux, Ap index, ≥2 MeV electron flux). The indices are 26-day averages centred at 1 April, 1 May, and 1 June while O3 is based on 26-day running means from 1 April to 1 November at altitudes from 20 to 70 km. During solar quiet time from 2005 to 2010, the composite of all three instruments reveals an apparent negative O3 signal associated to the geomagnetic activity (Ap index) around 1 April, on average reaching amplitudes between −5 and −10 % of the respective O3 background. The O3 response exceeds the significance level of 95 % and propagates downwards throughout the polar winter from the stratopause down to ∼25 km. These observed results are in good qualitative agreement with the O3 vmr pattern simulated with a three-dimensional chemistry-transport model, which includes particle impact ionisation.

1 Introduction

Energetic particles (keV–MeV), mainly originating from the sun but also from the Earth’s magnetospheric radiation belts and the aurora region, penetrate the atmosphere down to mesospheric and stratospheric regions, depending on their energy. The particles are guided by the Earth’s magnetic field lines and therefore mostly precipitate at auroral and radiation belt areas (~55–70° geomagnetic latitudes), depositing energy and directly influencing the chemical composition of the stratosphere and mesosphere. Due to the air compounds, precipitating particles mainly produce large abundances of O+2 as well as N(D) and N+2. N(D) and N+2 lead to increased concentrations of odd nitrogen (NOx = N + NO + NO2) through a number of reactions, including dissociative recombination of N+2 and ion-neutral chemistry with species of the oxygen family (e.g. Rusch et al., 1981). Additionally, O2+ and water vapour initialise chain reactions associated with water cluster ion formation and accompanied recombination reactions, which eventually lead to the production of odd hydrogen (HOx = H + OH + HO2; e.g. Solomon et al., 1981).

Both HOx and NOx play an important role in destroying O3 in the mesosphere and stratosphere (e.g. Lary, 1997). However, HOx is short-lived (seconds–hours) and therefore more important near its source region in the mesosphere, while NOx has a relatively long lifetime (days–months), at least during night-time conditions. Consequently, NOx can be transported downwards inside the polar vortex (e.g. Solomon et al., 1982) from the upper mesosphere/lower thermosphere down to the stratosphere, resulting in stratospheric...
O₃ depletion through catalytic chemical reactions in combination with solar radiation. Thus, energetic particle precipitation (EPP) indirectly affects O₃ during polar winter. Since O₃ is the major radiative heating source in the stratosphere, variations of this gas will also influence the stratospheric temperature field and eventually lead to altered atmospheric dynamics. However, the atmospheric response to EPP is not fully understood so far. The current knowledge is discussed in more detail by Sinnhuber et al. (2012).

Observations of the EPP indirect effect on stratospheric polar O₃ are relatively rare, at least compared to other latitudes, due to a lack of long-term O₃ measurements in these regions. However, a hint for this mechanism was presented by Randall et al. (1998) which analysed the Polar Ozone and Aerosol Measurement instrument data, revealing a close anticorrelation between NO₂ and O₃ mixing ratios in winter/spring from 1994 to 1996 in the Antarctic stratosphere (∼25–35 km). They suggested that the relationship cannot originate from downwards transported O₃-deficient air but is due to photochemical destruction of O₃ by NO₂. Further observations from several satellite instruments from 1992 to 2005 show that the stratospheric NOₓ enhancement in the Southern Hemisphere is caused by EPP (Randall et al., 2007). More recent satellite observations from 2002 to 2012 reported by Funke et al. (2014) reveal that particle induced NOₓ is indeed transported downwards to the middle stratosphere at polar latitudes, while further model studies suggest that the subsiding of NOₓ leads to strongly reduced stratospheric O₃ concentrations (∼30%) down to altitudes ∼30 km (e.g. Reddmann et al., 2010). Thus, it appears promising to search for a link between EPP and O₃ in actual data sets, because the downwards propagating signal of the EPP indirect effect on stratospheric and mesospheric O₃ throughout the polar winter has not been explicitly observed so far. Note that NOₓ can be only transported downwards inside a stable large-scale dynamical structure, which provides sufficient subsidence and prevents NOₓ removal/dilution by horizontal transport. These conditions are found primarily inside the Antarctic polar vortex, because the Arctic vortex is strongly disrupted by planetary waves, leading to its weakening or temporary breakdown. This large dynamical variability eventually causes high variations in O₃ volume mixing ratios (vmr), superposing the EPP indirect effect.

Therefore our study is focused on O₃ vmr observations inside the Antarctic polar vortex from ∼20–70 km, derived from environmental satellite/Michelson Interferometer for passive atmospheric sounding (ENVISAT/MIPAS), thermosphere ionosphere mesosphere energetics and dynamics/sounding of the atmosphere using broadband emission radiometer (TIMED/SABER), and Odin/sub-millimetre radiometer (SMR) measurements. The intra-seasonal variability of the O₃ vmr values has been investigated and the relation to a number of solar and geomagnetic indices, namely the F10.7 cm solar radio flux, the Ap index, and the ≥ 2 MeV electron flux is analysed.

2 Data analysis and numerical modelling

2.1 Approximation of the Antarctic polar vortex

The position and the extension of the Antarctic polar vortex were estimated by using the gradient of the potential vorticity (PV) on isentropic surfaces (Nash et al., 1996). Assuming a dry atmosphere at altitudes ≥20 km, the PV was calculated from temperature, pressure, relative vorticity, and the corresponding latitude taken from ERA-Interim (https://ecaccaccess.ecmwf.int/ecmwf), the latest version of global atmospheric reanalysis data produced by the European Centre for Medium-Range Weather Forecasts (ECMWF). The reference pressure was set to 1000 hPa and the gravitational constant was considered to be dependent on latitude and height. The PV was calculated for all height intervals between 20 and 70 km which were adapted from the MIPAS retrieval grid (see Sect. 2.2.1). Note that ERA-Interim data are primarily model-driven at mesospheric altitudes but the individual PV results look reasonable at each height interval. As an example, Fig. 1 shows the PV, depending on time and equivalent latitude (EQL), during the Antarctic winter 2011 at ∼40 km. The EQLs assigned to an individual PV isoline enclose the same area as the geographical latitudes of equivalent values. However, this area is located around an estimate of the vortex centre position, rather than around the geographical pole. In general, the EQL of the strongest PV gradient indicates the estimated location of the vortex edge; however, in most cases, there are at least two locations revealing gradients of similar magnitude. Therefore, Nash et al. (1996) also considered the zonal wind to locate the real vortex edge, but here we added a visual analysis instead of the zonal wind to divide the Southern Hemisphere into three non-overlapping zones: deep inside the Antarctic polar vortex (CORE) and the corresponding outermost edge (EDGE), covering all EQLs poleward of the respective borders, as well as an area not influenced by the vortex (OUTSIDE), which extends from the equator to the respective OUTSIDE border. In this study we will consider the EDGE region as the Antarctic vortex area, but the CORE and OUTSIDE region are still necessary to determine whether the observed features inside the EDGE zone are actually originating from the vortex itself. The limits of the three regions of each height interval revealed no strong variation from 2002 to 2011, therefore holding for every winter (Table 1). Note that the ECMWF ERA-Interim data only cover heights up to ∼63 km. However, considering the behaviour of the Antarctic vortex at altitudes between 60 and 70 km (Preusse et al., 2009, their Fig. 2a), it seems reasonable to assume that the estimated limits of the three regions at ∼63 km are also valid up to 70 km.
Table 1. Limits, derived from the potential vorticity \((10^{-6} \text{ K m}^2 \text{s}^{-1} \text{kg}^{-1})\), of the southern hemispheric regions CORE, EDGE, and OUTSIDE at the individual heights. The altitudes are adapted from the MIPAS retrieval grid and the shown potential vorticity values hold for 2002–2011.

<table>
<thead>
<tr>
<th>Nominal height (km)</th>
<th>CORE</th>
<th>EDGE</th>
<th>OUTSIDE</th>
<th>Nominal height (km)</th>
<th>CORE</th>
<th>EDGE</th>
<th>OUTSIDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>−60</td>
<td>−50</td>
<td>−30</td>
<td>37</td>
<td>−3600</td>
<td>−1900</td>
<td>−1000</td>
</tr>
<tr>
<td>21</td>
<td>−70</td>
<td>−50</td>
<td>−30</td>
<td>38</td>
<td>−6000</td>
<td>−2500</td>
<td>−1500</td>
</tr>
<tr>
<td>22</td>
<td>−90</td>
<td>−70</td>
<td>−40</td>
<td>39</td>
<td>−6000</td>
<td>−3000</td>
<td>−2000</td>
</tr>
<tr>
<td>23</td>
<td>−150</td>
<td>−90</td>
<td>−40</td>
<td>40</td>
<td>−9000</td>
<td>−3500</td>
<td>−2000</td>
</tr>
<tr>
<td>24</td>
<td>−200</td>
<td>−100</td>
<td>−60</td>
<td>41</td>
<td>−9000</td>
<td>−4000</td>
<td>−2000</td>
</tr>
<tr>
<td>25</td>
<td>−180</td>
<td>−110</td>
<td>−60</td>
<td>42</td>
<td>−9000</td>
<td>−4000</td>
<td>−2000</td>
</tr>
<tr>
<td>26</td>
<td>−280</td>
<td>−160</td>
<td>−100</td>
<td>43</td>
<td>−15000</td>
<td>−5000</td>
<td>−3000</td>
</tr>
<tr>
<td>27</td>
<td>−360</td>
<td>−220</td>
<td>−120</td>
<td>44</td>
<td>−15000</td>
<td>−5500</td>
<td>−3000</td>
</tr>
<tr>
<td>28</td>
<td>−600</td>
<td>−250</td>
<td>−120</td>
<td>46</td>
<td>−22000</td>
<td>−8000</td>
<td>−4000</td>
</tr>
<tr>
<td>29</td>
<td>−900</td>
<td>−300</td>
<td>−150</td>
<td>48</td>
<td>−18000</td>
<td>−10000</td>
<td>−2000</td>
</tr>
<tr>
<td>30</td>
<td>−800</td>
<td>−400</td>
<td>−200</td>
<td>50</td>
<td>−36000</td>
<td>−12000</td>
<td>−4000</td>
</tr>
<tr>
<td>31</td>
<td>−1000</td>
<td>−400</td>
<td>−200</td>
<td>52</td>
<td>−32000</td>
<td>−16000</td>
<td>−4000</td>
</tr>
<tr>
<td>32</td>
<td>−1600</td>
<td>−600</td>
<td>−300</td>
<td>54</td>
<td>−36000</td>
<td>−16000</td>
<td>−4000</td>
</tr>
<tr>
<td>33</td>
<td>−2000</td>
<td>−800</td>
<td>−400</td>
<td>56</td>
<td>−60000</td>
<td>−30000</td>
<td>−5000</td>
</tr>
<tr>
<td>34</td>
<td>−1800</td>
<td>−1100</td>
<td>−400</td>
<td>58</td>
<td>−60000</td>
<td>−30000</td>
<td>−10000</td>
</tr>
<tr>
<td>35</td>
<td>−2800</td>
<td>−1400</td>
<td>−800</td>
<td>60</td>
<td>−60000</td>
<td>−30000</td>
<td>−10000</td>
</tr>
<tr>
<td>36</td>
<td>−4000</td>
<td>−1600</td>
<td>−1000</td>
<td>62−70</td>
<td>−180000</td>
<td>−90000</td>
<td>−30000</td>
</tr>
</tbody>
</table>

Figure 1. Potential vorticity \((10^{-6} \text{ K m}^2 \text{s}^{-1} \text{kg}^{-1})\) at \(\sim 40\text{ km}\) during the Antarctic winter 2011 as a function of time and equivalent latitude. The thresholds of the regions OUTSIDE (dotted line), EDGE (solid line) and CORE (dashed line) are included. Potential vorticity was calculated from ECMWF Era-Interim data.

### 2.2 Ozone measurements

#### 2.2.1 MIPAS

MIPAS (Fischer et al., 2008) was a limb sounder on board ENVISAT, which had a sun-synchronous orbit. The main advantages of MIPAS measurements are the global coverage from 87° S to 89° N and the availability of observations during both day and night, crossing the equator at \(\sim 10:00\ \text{LT}\) and \(\sim 22:00\ \text{LT}\), respectively. MIPAS was a Fourier transform infrared (4.15–14.6 \(\mu\)m) emission spectrometer, allowing simultaneous observations of several atmospheric trace gases, including \(\text{O}_3\). MIPAS was operational from July 2002 to April 2012, but due to an instrument failure in March 2004, the entire observation period is divided into two subintervals from July 2002 to March 2004 and January 2005–April 2012 (referred to as P1 and P2 here, respectively). During P1 an almost continuous time series is available, while larger data gaps are present during P2 before October 2006. Here, we use the complete data set of the most frequent observation mode (nominal mode), covering the altitudes from the upper troposphere up to \(\sim 70\text{ km}\) at the poles which was derived from the MIPAS level-2 research processor developed by IMK/IAA. Details of the retrievals are described in von Clarmann et al. (2003), Glatthor et al. (2006), and von Clarmann et al. (2009). Note that the number of tangent heights is constant during P1 (17) and P2 (23), and that the actually available altitudes (cloud contaminated observations are disregarded) only slightly differ from day to day. The corresponding vertical resolution becomes coarser at higher altitudes (independent of the geographical location), increasing from 3.5 to 8 km (Steck et al., 2007) and from 2.5 to 5 km (Eckert et al., 2014) in P1 and P2, respectively. However, the retrieval grid in all MIPAS \(\text{O}_3\) data versions used here (V30_O3_9, V5R_O3_220, V5R_O3_221) is independent of the tangent heights, with a grid width of 1 km below 44 and 2 km above. During P1/P2 \(\text{O}_3\) was measured at two different wavelength intervals, ranging from 9.0–9.4/9.6–9.7 \(\mu\)m and 12.5–13.5/12.7–13.2 \(\mu\)m in particular. However, the full spectral ranges were not used, but sub-intervals (microwindows). These were selected to minimise the computing time and to optimise the relation between the measurement-noise induced random error and other errors. These other errors...
originating, among further error sources, from spectral contributions of further atmospheric constituents of unknown abundances. It should also be noted that there is a bias in MIPAS $O_3$ data between the two periods, which was estimated using a multi-linear parametric trend model (Eckert et al., 2014). To accept an $O_3$ data point, the recommended filter criteria for MIPAS $O_3$ data were applied by using an averaging kernel diagonal value > 0.03 as well as the visibility flag equal to 1 which indicates spectral available data.

At least 10 accepted data points inside the Antarctic polar vortex at a certain grid level were required to calculate the arithmetic average of one day, while at least 13 days were arithmetically averaged to a 26-day running mean from 1 April to 1 November, repeating this algorithm for each height interval and all years from 2002 to 2011. The time interval of 26 days was chosen to minimise a possible influence of the 27-day cycle of the sun, also ensuring that each time interval includes only one 27-day solar rotation maximum at most. The analysis was repeated for $NO_2$ (V5R_NO2_220, V5R_NO2_221) and the corresponding retrieval is described in Funke et al. (2005 and 2011).

### 2.2.2 SABER

The SABER instrument on board the TIMED Satellite has been nearly continuously operating since January 2002, measuring vertical profiles of several atmospheric parameters and minor constituents (e.g. $O_3$) from the surface up to altitudes > 100 km. The SABER measurements are governed by a periodic quasi 60-day cycle, each time changing from the Southern Hemisphere mode ($83^\circ S–52^\circ N$) to the Northern Hemisphere mode ($52^\circ S–83^\circ N$) and vice versa. Note that the “switching day” is only varying a few days from year to year. To consider both day and night $O_3$ observations, SABER Level 2A Ozone96 data v2.0 and v1.07 (http://saber.gats-inc.com/custom.php, Rong et al., 2009) measured at $\sim 9.6 \mu m$ are used. However, v1.07 was only used to fill v2.0 data gaps, which seemed reasonable because the data fit quite well the results of the performed analysis during the respective periods (15–31 May, 7–31 August, not shown here). Consequently, the combined dataset of both versions shows no larger data gap and the measurements of both versions were restricted to values $< 20 \text{ ppm}$ to exclude outliers. Comparisons with the results of an increased threshold to $< 100 \text{ ppm}$ revealed only minor differences (not shown here).

The investigated height interval, ranging from 20 to 70 km, is divided in 38 non-overlapping subintervals and binned at the same altitudes as MIPAS data. The algorithm used to calculate the running means is also identical to the one applied for the “accepted” MIPAS data points. However, SABER needs approximately 60 days to cover all local times, leading to a quasi 60-day wave like oscillation in $O_3$ if 26-day running means are used. This behaviour becomes evident at altitudes > 50 km, where the averaging interval was consequently extended from 26 to 60 days. Note that the calculation of the 60-day running means required at least 30 days.

### 2.2.3 SMR

The Odin satellite mission started in February 2001 and is a joint project between Sweden, Canada, France and Finland (Murtagh et al., 2002). Odin was launched into a sun-synchronous polar orbit, carrying the SMR instrument and nominally covering the latitude range from 82.5$^\circ S$ to 82.5$^\circ N$. The SMR makes vertical profile measurements during both day and night, while passing the equator at $\sim 06:00/18:00$ LT in the descending/ascending node. The $O_3$ data were extracted from the Odin/SMR Level 2 data product, version 2.0 (http://odin.rss.chalmers.se/, Urban et al., 2005), only using measurements of the frequency band centred around $\sim 544.6 \text{ GHz}$, providing vertical $O_3$ profiles in the $\sim 15–70 \text{ km}$ altitude range. The filtering criteria used for SMR is the measurement response, which corresponds to the sum of the rows of the averaging kernel matrix. The profiles characterised by a measurement response lower than 0.9 are not reliable enough, and are therefore excluded. The algorithm to calculate the 26-day running means is identical to the one applied to MIPAS data. Note that Odin/SMR was a two-discipline satellite until April 2007, switching between atmospheric (aeronomy mode) and astronomy observations, and is entirely dedicated to aeronomy since this date. Consequently, measurements in the relevant mode are roughly performed one day out of three before April 2007 and every other day afterwards. However, the calculation of the 26-day running means is still possible because the data gaps occur in a regular way, so they do not essentially worsen the 26-day averages. The vertical resolution of the data version used here is better than 3 km below 45 km, but increases to 5–6 (50–60 km) and 7–10 km (60–70 km), leading to noisy results at altitudes > 50 km compared to the other two instruments.

### 2.3 Solar data and geomagnetic indices

The data of the indices were obtained from two different websites provided by the National Geophysical Data Center. In detail the flux of the 10.7 cm radio emission from the sun (F10.7) and the geomagnetic Ap index (Ap), commonly used proxies for solar variation and geomagnetic activity, respectively, were downloaded from http://spidr.ngdc.noaa.gov/spidr/. The $\geq 2 \text{ MeV}$ electron flux (2 MeV), including the flux of all electrons with energy levels above 2 MeV, was measured by the Geostationary Operational Environmental Satellites (GOES) and the corresponding time series were downloaded from ftp://ftp.ngdc.noaa.gov/STP/SOLAR_DATA/SATELLITE_ENVIRONMENT/Daily_Fluences/. Note that the 2 MeV data set also considers contamination effects on the electron detectors on the spacecrafts due to protons $> 32 \text{ MeV}$. Furthermore the 2 MeV data are obtained from...
geostationary satellites which perform in-situ measurements in the radiation belts and consequently do not directly provide observations of precipitating particles. However, it is very likely that there is at least a positive relation between 2 MeV and precipitating relativistic radiation belt particles. Thus, the 2 MeV is not used as a proxy of precipitating particles but as an indicator of the influence from the magnetosphere. Precipitating particle integral fluxes in polar regions are observed by sun-synchronous Polar orbiting Operational Environmental Satellite (POES) detectors and the corresponding data correlate better with geomagnetic indices than the GOES electron fluxes (Sinnhuber et al., 2011). However, the respective measurements of the POES instruments tend to underestimate the fluxes from ground-based observations during weak geomagnetic activity (Rodger et al., 2013). Since this study focus on 2002–2011 and an essential part of this time interval overlaps with low geomagnetic activity, GOES data and Ap are used instead of POES measurements. The time series of all data sets are based on daily values, which were arithmetically averaged to 26-day means centred at 1 April, 1 May, and 1 June. The means were separately calculated for each index for the individual years from 2002 to 2011; however, 2 MeV data are only available until 2010.

2.4 Numerical modelling

The three-dimensional chemistry and transport model (3dCTM; Sinnhuber et al., 2012, Appendix 1) used here is based on the Bremen 3dCTM (e.g. Wissing et al., 2010), extending on 47 pressure levels from the tropopause up to the lower thermosphere (∼10–140 km) with a latitude/longitude resolution of 2.5° × 3.75°. The model was recently updated with a variable H2 and O2 distribution, leading to proper HOx and consequently night time O3 values at altitudes > 60 km (see Sect. 1). The 3dCTM is driven by meteorological data obtained from simulations of the three-dimensional dynamical model LIMA (Berger, 2008) and the advection is calculated by applying the second-order moments scheme reported by Prather (1986). In the stratosphere, a family approach for the chemical families: Ox (O + O(1D) + O3), NOx (N + NO + NO2), HOx (H + OH + HO2), BrOx (Br + BrO), ClOx (Cl + ClO + 2Cl2O2), and CHOx (CH3+ CH2O2+ CH2OH + CH3O + HCO) is used, but was not used for O3, HOx, and NOx in the mesosphere/lower thermosphere region.

In this study the 3dCTM was used to investigate the impact of precipitating particles on O3 inside the Antarctic polar vortex at altitudes from 20 to 70 km. After a multi-year two-dimensional model spin-up, two simulations from 2003 to 2009 were performed. The first run (base run) does not consider any energetic particles, while the second run (EP run) includes ionisation effects by both protons and electrons, using the ionisation rates provided by the Atmospheric Ionisation Module Osnabrück (AIMOS; Wissing and Kallenrode, 2009). The resulting NOx production per created ion pair includes various ionic and neutral reactions depending on the atmospheric background state (Nieder et al., 2014). Simple parameterisations are used for the production of HOx (Solomon et al., 1981) and O (Porter et al., 1976). Note that heterogeneous chemistry was not included, which only becomes important during spring in the lower stratosphere. Both model runs considered constant solar minimum conditions (F10.7 = 70 × 10−22 W m−2 Hz−1) to exclude O3 variations due to solar activity. The obtained O3 model results of both runs were separately selected according to the vertical MIPAS retrieval grid for direct comparisons to the observations, repeating the described algorithm to calculate the 26-day running means. Finally, in order to derive the O3 vmr variations solely originating from precipitating particles, the obtained averages of the base run were subtracted from the corresponding O3 values of the EP run. The results were divided by the arithmetic mean of both runs and eventually multiplied by 100 %.

3 Results and discussion

3.1 Satellite observations

3.1.1 O3 response from 2002 to 2011

The 26-day O3 vmr averages from 2002 to 2011 of each altitude-time interval (1 April–1 November, 20–70 km) were individually grouped into years of high and low index activity. For this purpose the index median of the corresponding time series of the 26-day average of an index (F10.7, Ap, 2 MeV) centred around 1 April was calculated, only including years of actually available O3 observations. Therefore the median of an index time series works as a threshold, dividing the entire time interval from 2002 to 2011 in years of high (above the median) and low (below the median) index activity. Note that the classification of the years does not only depend on the chosen index, but due to data gaps also on the considered height-time interval as well as the instrument used. Afterwards the arithmetic O3 mean of the years of low index activity was subtracted from the O3 mean of the years of high index activity, eventually dividing this absolute O3 difference by the arithmetic O3 average of the entire observation period and multiplying the results by 100 % for more handy values. Thus the calculated relative O3 difference (referred to as O3 amplitude here) represents the impact of the respective index on the O3 background. To reduce the measurement noise of the individual instruments, the results of all three instruments were merged by simply calculating the arithmetic average but only if the corresponding O3 amplitude of all three instruments was available. Note that due to the major sudden stratospheric warming centred around 27 September (Azeem et al., 2010) the O3 observations from 1 September to 1 November 2002 were excluded. In contrast, the solar proton event in the end of October 2003 (Jackman...
et al., 2005) was neglected due to its late occurrence. The performed analyses with \( \text{O}_3 \) observations, considering the indices from 1 May and 1 June (not shown here), revealed no essential differences compared to 1 April or the structures became less obvious. Comparisons with earlier periods are not reasonable because the vortex first builds up in April. Therefore the focus is set on the \( \text{O}_3 \) response to indices centred around 1 April. The \( \text{O}_3 \) amplitude was calculated for all three regions (CORE, EDGE, and OUTSIDE) which were introduced in Sect. 2.1. The corresponding results reveal that the pattern found inside the EDGE region are fairly similar and less noisy compared to the features observed in the CORE area (not shown here). In contrast the \( \text{O}_3 \) amplitudes outside the Antarctic polar vortex are fundamentally different. An example for the \( \text{O}_3 \) response associated to 1 April Ap in the EDGE and the OUTSIDE region derived from MIPAS measurements is presented in Fig. 2, showing considerably disagreeing structures and essentially weaker amplitudes, especially below 50 km. Thus comparison between the individual regions of the Southern Hemisphere ensures, that the pattern found in the EDGE region are actually originating from the Antarctic polar vortex.

Figure 3 displays the corresponding results of the \( \text{O}_3 \) amplitude from 2002 to 2011, but only for values above the significance level of 95\% while shaded areas show regions between the significance level of 95 and 99\%. The significance was calculated according to a Student’s \( t \) test, based on the error of the mean of the 26-day running \( \text{O}_3 \) means and assuming the worst case scenario of absolute error propagation. The MIPAS \( \text{O}_3 \) measurements (left column) reveal a high negative response to Ap (upper row) in early Antarctic winter > 60 km, on average ranging around \(-10\%\). Further striking negative \( \text{O}_3 \) amplitudes occur in July between 30 and 40 km as well as around 1 October at \( \sim 30 \) km, at least weakly indicating the downward transport of the Ap signal in stratospheric \( \text{O}_3 \) due to NO\(_x\) predicted by model studies (e.g. Reddmann et al., 2010). In contrast, a positive \( \text{O}_3 \) amplitude is found at the beginning of the winter between \( \sim 25 \) and \( \sim 55 \) km (\( \sim 10\%\)), as well as at altitudes < 30 km throughout the winter (up to \( \sim 20\% \) in October at \( \sim 20 \) km) and above the indicated subsiding layer of negative amplitudes. But considering that most of these features drop below the significance level of 95\% by combining the data of all three instruments (right column), a more detailed investigation of these patterns is not reasonable. However, the results of the merged data set show a well pronounced subsiding negative Ap signal from \( \sim 50 \) km in June down to \( \sim 25 \) km in October, which is disrupted in August, while the generally positive structures below 30 km are also still present.

The \( \text{O}_3 \) response to 2 MeV (middle row) derived from MIPAS observations also indicates a downwelling of negative \( \text{O}_3 \) amplitudes, descending from \( \sim 60 \) km in June down to \( \sim 30 \) km in late August. Additionally, the MIPAS \( \text{O}_3 \) response to 2 MeV in early winter is reversed compared to the corresponding influence from Ap on \( \text{O}_3 \), which does not originate from missing 2 MeV data from 2011. Strong positive \( \text{O}_3 \) amplitudes are generally observed throughout the winter below 30 km, exceeding values of \( \sim 20\% \) in April and October, as well as during October between 30 and 50 km where the maximum amplitude is lower (\( \sim 10\% \)). The positive features can be validated with the composite results even if they are damped in the region below 30 km. However, this is not the case for the negative response, except for a small area in June in the lower mesosphere. Considering that the Ap responds to lower particle energy levels compared to 2 MeV and that the behaviour of both indices is essentially different from 2002 to 2010 (see Fig. 4), the different \( \text{O}_3 \) amplitudes associated to Ap and 2 MeV are still reasonable.

The \( \text{O}_3 \) response to F10.7 (lowermost row) is fairly similar between MIPAS and the merged measurements, and both also agree with the respective pattern observed for Ap, including the indicated downwelling of negative \( \text{O}_3 \) amplitudes during midwinter from 50 to 25 km. The composite \( \text{O}_3 \) shows strong positive amplitudes in May > 55 km which originate...
Figure 3. \( O_3 \) amplitude (see Sect. 3.1.1 for definition) inside the Antarctic polar vortex between years of high index values and years of low index values, namely \( Ap \) index (upper row), \( \geq 2 \) MeV electron flux (middle) as well as \( F10.7 \) cm solar radio flux (lowermost row) centred around 1 April, derived from MIPAS (left column) and composite (MIPAS+SMR+SABER, right column) observations from 2002 to 2011. Shown are only values above the significance level of 95 %. Additionally, regions between the significance level of 95 and 99 % are shaded in black or white, according to a Student’s \( t \) test.

from SMR measurements and are most likely due to the low vertical resolution of the SMR instrument at these altitudes (see Sect. 2.2.3). The high agreement between the results of \( Ap/O_3 \) and \( F10.7/O_3 \) might originate from the coupling of both indices during solar maximum years (Gray et al., 2010, their Fig. 1). In order to investigate a possible cross-correlation between solar radiation and geomagnetic disturbances, the analysis was repeated for years of moderate solar activity, only including 2005–2010 (Fig. 4). Similar analyses to extract a more distinct solar signal during times of approximately constant geomagnetic activity were not reasonable, because the respective years of nearly constant \( Ap \) values (2002, 2005, 2006, 2008, 2010) do not provide a sufficient amount of data in MIPAS and SMR measurements.

3.1.2 \( O_3 \) behaviour during solar minimum activity (2005–2010)

Figure 5 displays the obtained \( O_3 \) amplitudes for solar quiet times (2005–2010) associated to 1 April \( Ap \), again only showing values above 95 % significance level and shading the area of regions between 95 and 99 %. The MIPAS \( O_3 \) response to \( Ap \) indicates a subsiding negative signal (\( \sim -10 \) to \( -15 \) %), starting in late June slightly below 50 km and prop-
agitating downwards to ~25 km throughout the winter. However, the middle part of the downwelling between late July and late September is below the significance level of 95% and therefore not shown here. Furthermore, the hinted subsidence is closely surrounded by well pronounced positive O₃ amplitudes, especially below ~30 km which maximise in September (>20%). There is also a negative structure centred at 1 June at ~60 km, which cannot be caused by NOₓ but most likely results from HOₓ formation (see Sect. 1). Considering the composite results, the downwelling Ap signal in O₃ becomes apparent and robust but slightly weaker (~−10%) while the positive features are also damped but still present. The mesospheric response is generally weak and the high positive O₃ amplitudes in May are again caused by the SMR measurements.

The 2 MeV impact on MIPAS O₃ shows generally agreeing features with the influence of the geomagnetic activity and is also of similar magnitude, however, the downwelling negative signal is hinted to already start in late May at ~55 km. In contrast to the O₃ response to Ap, the downwards propagating 2 MeV signal is less robust and can be only guessed in the composite O₃ amplitude, while the positive structures (~10–15%) in August below 30 km and in September between 30 and 50 km are still present. In general, the 2 MeV features are less obvious in the O₃ composite, except for the positive O₃ amplitudes above the hinted downward transport. Nevertheless, the agreement between Ap and 2 MeV pattern is quite strong, in MIPAS observations in particular, although both parameters are only indirectly related to O₃. However, the O₃ structure associated to both indices is far too similar and additionally found in all three instruments to be a coincidence, even if the descending O₃ response to 2 MeV is weaker. Since Ap represents lower particle energy levels compared to the 2 MeV and both indices are only moderately correlated (see Fig. 4), the similar results strongly indicate a related source mechanism, suggesting solar wind variability.

Considering the entire process, that energetic particles produce NOₓ, which eventually destroys stratospheric O₃, the Ap impact observed in O₃ (see Fig. 5) is expected to be reversed in NOₓ, at least in the stratosphere. In order to investigate this in more detail, the analysis was repeated for 1 April Ap and NOₓ. Here NOₓ is represented only by NO₂ from MIPAS observations, because the respective NO measurements are quite noisy compared to NO₂, especially below 30 km. This is still reasonable because NO is converted to NO₂ during night and therefore NO₂ is the major fraction of NOₓ inside the Antarctic polar vortex. The corresponding results include the years 2005–2010 and are displayed in Fig. 6, supporting that the stratospheric O₃ depletion can be indeed associated to the catalytic NOₓ/O₃ cycle. The Ap signal in NO₂ is stronger by the factor of 2–5, compared to the respective O₃ amplitudes. The sharp gradient in mid July originates from 2005 NO₂ data, which are not available afterwards. However, the general structure of the subsiding Ap signal in NO₂ is still similar with and without 2005 observations. Note that the essentially smaller NO₂ amplitudes in October below the significance level of 95% are not in conflict with the respective well pronounced negative O₃ response, because the latter one results from an accumulation effect from the NO₂ above. Furthermore, large negative NO₂ amplitudes throughout the entire winter below ~30 km are observed, matching the high positive O₃ response to Ap. A possible reason for this behaviour might be that NO₂ is stored in reservoir species, like ClONO₂, HNO₃, and N₂O₅, due to reactions with ClO, OH, and NO₂, respectively. However, N₂O₅ is converted to HNO₃ via water ion cluster chemistry (López-Puertas et al., 2005, their reactions 1 and 8–12), which was also investigated with respect to EPP for conditions without solar proton events by Stiller et al. (2005). These reactions eventually lead to lower NOₓ concentrations, consequently slowing down the catalytic O₃ depletion. Based on the corresponding MIPAS
climatologies (not shown here), HNO3 is more important until mid July, while ClONO2 is dominating afterwards and its influence becomes essentially crucial in spring due to heterogeneous chemistry which has taken place before. This suggested NO2-ClONO2 mechanism is supported by Whole Atmosphere Community Climate Model results reported by Jackman et al. (2009, their Figs. 6 and 7), who simulated the impact of the SPE in July 2000 on stratospheric O3 and NO3 (= NOx + NO3 + N2O5 + HNO3 + HO2NO2 + ClONO2 + BrONO2).

Furthermore, the positive O3 amplitudes below \( \sim 30 \) km could be also partly explained by the self healing effect of O3 (Jackman and McPeters, 1985). Altitude regions of reduced O3 will lead to increased solar UV radiation in the layers directly below. This is accompanied by a higher production of atomic oxygen and would consequently increase the formation of O3. However, this proposed mechanism would only have an additional effect, contributing to the formation of O3 in the atmospheric layer right below the subsidence, but cannot account for the entire region. Note that this layer is also present throughout the entire winter, and thus an influence from the vortex above is unlikely but any further investigations are beyond the scope of this study.

Additionally, the area of high positive Ap/O3 structure between 35 and 50 km from August to September cannot be completely explained by the NOx/O3 cycle. In detail, the respective Ap influence of NO2 is close to 0 and consequently well below the 95 %, while the respective MIPAS ClONO2 amplitude (not shown here) reveals positive values, which are also mostly below the 95 % significance level. These results are at least not in conflict with a higher O3 amplitude. Furthermore, this positive Ap impact on O3 is essentially less visible in the composite results than in MIPAS data, and a corresponding composite analysis for Ap/NO2 is necessary for a more detailed investigation. But this is not possible due to non-existing NO2 measurements from SABER and SMR. Thus no definite explanation can be given at this stage and...
this feature is a subject of a future work. However, it should be pointed out that this structure does not harm the underlying mechanism proposed to explain the identified negative O$_3$ amplitude and subsequent downward transport.

### 3.2 Comparison with 3dCTM

The simulated O$_3$ amplitude between the EP run and base run, representing high and low geomagnetic activity, respectively, is displayed in Fig. 7. Note that the modelled O$_3$ amplitude is also referred to as O$_3$ amplitude here, which is justified because “observed” and “modelled” O$_3$ amplitude still hold the same physical meaning, even if the calculation algorithm is slightly different. It is reasonable to investigate the complete simulated time interval from 2003 to 2009, because the model runs represent solar minimum conditions similar to the years 2005–2010. The results reveal apparent negative O$_3$ amplitudes propagating downwards throughout the winter with maximum negative values during midwinter between 45 and 60 km. The subsidence shows larger negative O$_3$ amplitudes compared to the measurements and is also much broader, which might be due to the constant F10.7 and the prescribed dynamics, both reducing the inter-annual variability of O$_3$. Furthermore, we performed an on/off experiment, while in reality the EPP indirect effect is a persistent feature. Below 30 km the observed high positive O$_3$ amplitudes associated with Ap are only indicated in the model results by essentially weaker and additionally negative amplitudes. However, the model amplitudes are at least less negative compared to the values above. The second positive region above the downwelling is completely missing. Further note that the strong positive response during late winter/early spring below 30 km might not be reproduced by the model due to missing heterogeneous chemistry. The proposed self healing effect of O$_3$ (see Sect. 3.1.2) was also tested, using O$_1^D$ as a proxy for the O$_3$ photolysis rate in the Lyman-alpha band and calculating the O$_1^D$ amplitude (not shown here). However, the expected positive response directly below the downwelling is only partly visible and even below the 67 % significance level.

The qualitative agreement between model results and observations in the stratosphere suggests that the subsiding Ap signal found in O$_3$ is actually originating from particle precipitation. However, the simulated downwelling starts at altitudes > 60 km while observations reveal no obvious structures in the mesosphere, possibly caused by satellite sampling. As already stated in Sect 3.1.2, the mesospheric behaviour cannot be caused by NO$_x$, because the NO$_x$/O$_3$ cycle is not efficiently working at these altitudes. Thus the O$_3$ depletion > 50 km could be accounted to OH production, which is most likely overestimated in the model and consequently leads to an increased O$_3$ depletion not observed by the satellite instruments.

### 4 Conclusions

We have investigated the O$_3$ behaviour inside the Antarctic polar vortex from 2002 to 2011, observed by three independent satellite based instruments ENVISAT/MIPAS, Odin/SMR, and TIMED/SABER. These O$_3$ vmr measurements, based on 26-day running means from 1 April to 1 November covering altitudes from 20 to 70 km, were individually grouped into high and low index activity according to the 26-day averages centred around 1 April, 1 May, and 1 June of different solar and geomagnetic indices (F10.7, Ap, 2 MeV). After minimising the direct influence of the solar radiation by only considering the period of solar minimum activity from 2005 to 2010 we found a negative O$_3$ response caused by geomagnetic activity (Ap) from 1 April in all three instruments, ranging from −5 to −10 % and propagating downwards throughout the Antarctic winter from ~50 km down to ~25 km. This subsiding negative signal in O$_3$ is above the significance level of 95 % and overlaps with the corresponding positive NO$_2$ response to 1 April Ap, supporting that NO$_x$ is indeed the cause of the O$_3$ depletion. We could also show that the high positive O$_3$ response below 30 km, which is present during the entire winter, is in agreement with respective negative NO$_2$ structures. The cause of the NO$_2$ behaviour is possibly related to the formation of the reservoir species ClONO$_2$ and HNO$_3$, slowing down the catalytic destruction of O$_3$ by Cl. The O$_3$ pattern induced by the magnetosphere (2 MeV) from 1 April are similar but weaker, compared to the respective geomagnetic activity, still suggesting a related source mechanism between 2 MeV and Ap like solar wind variability. The composite observations of all three instruments are in good qualitative agreement with 3dCTM simulation, revealing similar O$_3$ pattern induced by the geomagnetic activity from 1 April while the simulated O$_3$ response is larger but still in the same order of magnitude.

However, we have to point out that the validity of the subsiding O$_3$ depletion associated to geomagnetic activity and
Acknowledgements. T. Fytterer, H. Nieder, and M. Sinnhuber gratefully acknowledge funding by the Helmholtz Association of German Research Centres (HGF), grant VH-NG-624. The authors also acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology. Odin is a Swedish-led satellite project funded jointly by the Swedish National Space Board (SNSB), the Canadian Space Agency (CSA), the National Technology Agency of Finland (Tekes), the Centre National d’Études Spatiales (CNES) in France and the third party mission program of the European Space Agency (ESA). We further like to thank the ERA-Interim for free provision of data and related support.

References


