Effect of energetic electron precipitation on the northern polar vortex: Explaining the QBO modulation via control of meridional circulation

A. Salminen[1], T. Asikainen[1], V. Maliniemi[1], K. Mursula[1]

[1]ReSoLVE Centre of Excellence, Space Climate Research Unit, University of Oulu, Finland

Key Points:

- Energetic electron precipitation (EEP) causes ozone loss, modifies temperatures and strengthens northern polar vortex in winter.
- EEP effect on polar vortex is significantly larger in QBO-E than in QBO-W phase when using QBO at 30 hPa lagged by 6 months.
- Our results suggest that QBO modulated meridional circulation affects the EEP response.

Corresponding author: Antti Salminen, `antti.salminen@oulu.fi`

Abstract

Energetic electron precipitation (EEP) affects the high-latitude middle atmosphere by producing NO_X compounds which destroy ozone. Earlier studies have shown that in the wintertime polar stratosphere increased EEP enhances the westerly wind surrounding the pole, the polar vortex. This EEP effect has been found to depend on the Quasi-Biennial Oscillation (QBO) of equatorial winds, but the mechanism behind this modulation has so far remained unresolved. In this study we examine the atmospheric effect of EEP and its modulation by QBO using the corrected electron flux measurements by NOAA/POES-satellites and the ERA-Interim reanalysis data of zonal wind, temperature and ozone in winter months of 1980-2016. We verify the EEP-related strengthening of the polar vortex, warming (cooling) in the upper (lower) stratosphere and a reduction of ozone mass mixing ratio in the polar stratosphere. We also verify that the EEP effect is stronger and more significant especially in late winter, when the QBO at 30 hPa is easterly. We find here that the difference in the EEP effect between the two QBO phases is largest using a roughly 6-month lag for QBO. We demonstrate that ozone mass mixing ratio in the lower polar stratosphere, a proxy for the strength of Brewer-Dobson circulation, is also larger during QBO-E than QBO-W, with the difference maximizing when the QBO is lagged by 6 months. Our findings indicate that the modulation of the Brewer-Dobson circulation by QBO controls how the EEP affects the polar vortex.

1 Introduction

Energetic particles precipitate into the Earth's atmosphere continuously. Energetic particle precipitation (EPP) excites and ionizes neutral atoms and molecules in the mesosphere and upper stratosphere at high latitudes. EPP can form reactive odd nitrogen NO_X which depletes ozone catalytically [*Crutzen et al.*, 1975]. During polar winter the NO_X molecules have a relatively long lifetime to descend from the mesosphere down to the lower stratosphere [*Funke et al.*, 2005, 2014]. Even energetic electron precipitation (EEP), which has its highest ionization rates in the lower thermosphere and upper mesosphere [*Turunen et al.*, 2009], can then affect the stratosphere via transported NO_X. This is the so-called EEP indirect effect [*Randall et al.*, 2007].

The polar stratosphere does not receive solar UV-radiation during winter, and as a result cools radiatively compared to the lower sunlit latitudes. The resulting meridional temperature gradient leads to the formation of westerly thermal wind, the polar vortex, which surrounds the polar region. Many studies based on models [e.g., *Rozanov et al.*, 2005; *Baumgaertner et al.*, 2011; *Arsenovic et al.*, 2016] and observations [e.g., *Lu et al.*, 2008; *Seppälä et al.*, 2013] have suggested that in northern hemisphere the wintertime polar vortex might be enhanced by EEP activity likely via EEP induced ozone loss. The ozone loss related to descending EEP-NO_X (NO_X generated by EEP) is mostly restricted to the mesosphere and upper stratosphere [*Sinnhuber et al.*, 2018]. In the polar region the upper stratospheric ozone loss leads to a radiative warming in the dark polar winter and to a radiative cooling in spring after polar sunrise [*Sinnhuber et al.*, 2018; *Meraner and Schmidt*, 2018]. Modeling studies suggest that these thermal changes are associated with an enhanced polar vortex and cooling of the lower polar stratosphere, which may be related to reduced downwelling of air in the lower stratosphere [e.g., *Baumgaertner et al.*, 2011; *Arsenovic et al.*, 2016]. The descent of NO_X into the stratosphere [*Funke et al.*, 2005, 2014] and associated ozone depletion [*Damiani et al.*, 2016] occur in both hemispheres, but are more variable in the northern hemisphere because of stronger planetary wave activity. The dynamical effects on the southern polar vortex are much weaker, less significant and even opposite to those in the northern hemisphere [*Lu et al.*, 2008; *Arsenovic et al.*, 2016; *Tomikawa*, 2017]. The variations of polar vortex are also seen in the Northern Annular Mode (NAM) and North Atlantic Oscillation (NAO) surface climate modes [*Baldwin and Dunkerton*, 2001], and EEP (or geomagnetic activity as EEP proxy) has been found to correlate with these modes in several studies [*Palamara and Bryant*, 2004; *Baumgaertner et al.*, 2011; *Maliniemi et al.*, 2013, 2016].

The northern polar vortex is greatly affected by planetary waves, which originate in the troposphere and can propagate upward into the stratosphere if the background wind is westerly and not too strong [*Charney and Drazin*, 1961]. In the stratosphere the planetary waves decelerate the westerly wind and move air in meridional direction. This establishes the so-called Brewer-Dobson circulation which also transports trace gases, e.g. ozone, from the equatorial stratosphere to higher latitudes [*Butchart*, 2014]. Transported ozone accumulates in the lower stratosphere of the winter pole. It has been shown that a majority of inter-annual variations polar total ozone are related to variations of the Brewer-Dobson circulation [*Salby and Callaghan*, 2002]. By continuity, the circulation consists of upwelling in the equator and downwelling in the polar region. Therefore, the polar stratosphere is adiabatically warmed by the Brewer-Dobson circulation while the tropical region is cooled.

An equatorial zonal wind mode, the quasi-biennial oscillation (QBO), affects the propagation of planetary waves and, therefore, the polar vortex. *Holton and Tan* [1980] found that the polar vortex is weaker if the QBO wind at 50 hPa pressure level is easterly. Since planetary waves can propagate only if the background wind is westerly, the QBO affects directly how these waves are guided in the stratosphere. If the QBO wind is easterly, planetary waves are diverted towards higher latitudes. Therefore, the polar vortex is more disturbed and weaker in the easterly QBO phase than in the westerly phase. This is the so-called Holton-Tan effect. The QBO wind pattern also forms its own meridional circulation cell near the equator which affects the Brewer-Dobson circulation so that the ascent rate in the equatorial stratosphere is larger in the easterly phase [*Flury et al.*, 2013].

The QBO also modulates the atmospheric effects of other factors. *Palamara and Bryant* [2004] and *Maliniemi et al.* [2013] found that the EEP effect on the NAO/NAM is stronger in the easterly phase of QBO determined at 30 hPa level (QBO30). However, *Lu et al.* [2008] found that the EEP effect on the polar vortex is stronger in May in the westerly phase of deseasonalized QBO determined at 50 hPa level (QBO50) and *Seppälä et al.* [2013] found the same dependency for December. *Maliniemi et al.* [2016] studied QBO30 and QBO50 modulation of the EEP-NAM relation and concluded that QBO30 modulation has been significant since the beginning of 20th century and stronger than QBO50 modulation, which has been present only since 1970s.

Since energetic electron precipitation is driven by solar wind, it does not follow the 11-year sunspot cycle. EEP activity is mostly controlled by high-speed solar wind streams [e.g., *Meredith et al.*, 2011; *Asikainen and Ruopsa*, 2016]. High-speed streams originate from solar coronal holes which appear at low solar latitudes most commonly in the declining phase of the sunspot cycle [*Bame et al.*, 1976; *Mursula et al.*, 2015]. As a result, EEP activity peaks during the declining phase, a few years after the sunspot maximum. The varying solar irradiance is another solar factor which causes variability in the atmosphere. Solar irradiance follows the sunspot cycle and, e.g., UV-radiation varies 4-8% in wavelengths of 150-250 nm (near the O_2 photolysis region) during the solar cycle [*Floyd et al.*, 2003]. *Labitzke and Van Loon* [1988] found that the northern polar vortex is colder and stronger in solar maxima compared to minima if the QBO phase is easterly while the opposite is true in the westerly QBO phase.

The aim of this study is to clarify how the EEP affects the northern winter stratosphere and troposphere, and how this effect is modulated by the QBO. We use linear regression analysis to estimate the EEP effect on ozone, temperature and zonal wind. Regressions are also calculated separately for the two QBO phases in order to study how the QBO modulates the EEP effect. The paper is organized as follows: In Section 2 we describe the data and methods used. In Section 3 the results of regression analysis without QBO separation are presented. In Section 4 we show the results when data are divided according to QBO phase. Section 5 discusses the QBO lag and in Section 6 we give our conclusions.

2 Data and methods

We used energetic electron fluxes measured by the MEPED instrument onboard the NOAA/POES-satellites flying at 800-900 km altitude from 1979 onwards, as a measure for EEP activity. MEPED instrument detects electrons in three integral energy channels with two telescopes, one directed along the zenith (0° telescope), the other perpendicular to zenith (90° telescope). In this study we used the lowest energy channel which measures the flux of electrons with energies above 30 keV. In inter-annual timescales this flux correlates well with geomagnetic Ap and AE indices (cc=0.86 for Ap and cc=0.87 for AE, $p < 10^{-11}$), which are often used as a proxy for precipitation of auroral electrons (energies 1-30 keV). Therefore, the EEP flux above 30 keV also works as a simultaneous proxy for auroral electron precipitation. MEPED electron flux measurements suffer, e.g., from proton contamination and inhomogeneity caused by two instrument versions. We used the electron fluxes corrected and homogenized by *Asikainen and Mursula* [2013]. MEPED telescopes are differently orientated in NOAA-15 and newer satellites compared to NOAA-12 and older satellites. This causes another inhomogeneity around 1998. We corrected this inhomogeneity by using the overlapping measurements of NOAA-12 and NOAA-15 in 1998-2001 to scale NOAA-15 fluxes to the level of previous measurements. The fluxes were computed as monthly averages of the two telescopes over corrected geomagnetic latitudes of 40°-90°.

To study the atmospheric response, we used the ERA-Interim reanalysis data provided by European Centre for Medium-Range Weather Forecasts (ECMWF) [*Dee et al.*, 2011]. The ERA-Interim data are available from 1979 onwards. The studied variables are ozone (mass mixing ratio), temperature (K) and zonal wind (m/s). We used monthly and zonally averaged values at latitudes 0°-90°N with a 2.5° resolution and at 37 pressure levels from the surface (1000 hPa) to upper stratosphere (1 hPa). As a QBO phase index, we used the zonal wind at 30 hPa averaged over the latitude range of 10°S-10°N and all longitudes.

The ozone assimilation in ERA-Interim reanalysis is known to be problematic [*Dee et al.*, 2011; *Dragani*, 2011]. Ozone observations from several different satellite missions have been assimilated into ERA-Interim [*Dragani*, 2011]. The ozone is assimilated in such a way that the model dynamics affects the ozone estimates but the assimilated ozone does not affect the dynamical variables like temperature and wind. The assimilated ozone measurements mostly come from the sunlit parts of the atmosphere, which is why the high-latitude ozone estimates during polar night can have a relatively large uncertainty [*Dragani*, 2011]. These polar night estimates of ozone at high-latitudes are thus largely based on ERA-Interim model dynamics and the assimilated data from the edges of the polar night region, which may contain information about the variability of ozone due to NOx. Note, however, that the lack of ozone measurements is mostly a problem in December and January, while in February and March a significant part of the polar vortex region is already covered by the ozone measurements.

We considered winters from 1979/1980 to 2015/2016, but we excluded those with an exceptionally large stratospheric sudden warming (SSW) or volcanic activity. Winters 1984/1985 and 2003/2004 were excluded since an unusually early and long-lasting SSW occurred during these two winters [*Manney et al.*, 2005]. The mean January zonal wind at 60°N-90°N at 50 hPa was easterly only in these winters of all the winters considered. We will discuss later in Section 4 the impact that these two excluded winters have on the results. We also performed analyses by excluding other strong SSWs in the considered time period (1987/1988, 2005/2006, 2008/2009, 2012/2013), and the results remained essentially the same. We also excluded winters 1982/1983 and 1991/1992 since these winters follow the large volcanic eruptions of El Chichón and Mount Pinatubo.

We computed linear regressions in which EEP is the explaining variable and one of the atmospheric variables (ozone mass mixing ratio, temperature or zonal wind) is the response variable. EEP and atmospheric data were first detrended. We used the LOWESS (locally weighted scatterplot smoothing) method [*Cleveland*, 1979] with a 31-year window to estimate the local trend. This smooth trend models, e.g., the varying long-term ozone trend (decline until 2000 and leveling off after that) better than a simple linear trend. Regressions were computed separately for each latitude-pressure level grid-box and for each of the four winter months (December-March). We used the Cochrane-Orcutt method [*Cochrane and Orcutt*, 1949] to calculate the regressions. In this method the residual term is modeled as an autoregressive AR(1) process, which can incorporate variability not only due to random uncorrelated noise but also due to other, omitted factors (e.g., solar UV irradiance, ENSO, QBO etc.). The regression is calculated as follows. We first calculate the normal linear regression as

$$Y_t = \alpha + \beta X_t + \epsilon_t, \tag{1}$$

in which X_t is the explaining variable (EEP), Y_t is the response variable (one of the atmospheric variables in the given grid-box), α is the constant term, β is the regression coefficient for explaining variable, and ϵ_t is the residual term. Unlike in normal regression, where the residual is assumed to be uncorrelated white noise, the residual term here is modeled as $\epsilon_t = \rho\epsilon_{t-1} + e_t$ where ρ is the lag-1 autocorrelation of the residual and e_t is normally distributed white noise part of the residual. After the first regression of Eq. 1 we can calculate the residuals and obtain a first estimate of their lag-1 autocorrelation. The regression equation 1 can then be rewritten as

$$Y_t - \rho Y_{t-1} = \alpha(1-\rho) + \beta(X_t - \rho X_{t-1}) + e_t, \tag{2}$$

where the remaining residual term is white noise. We then compute the regression parameters α and β from Eq. 2 and recalculate the residuals. From these new residuals we can then obtain a new estimate for ρ. Computing the regression parameters from Eq. 2 and then ρ is continued iteratively until ρ converges. We then obtain the final value for regression coefficient β. We then multiply β with the standard deviation of the whole EEP time-series to get ΔY, i.e., the response in the atmospheric variable Y to an increase of one standard deviation in EEP. The significance of the response is estimated with the two-tailed Student's t-test, which is appropriate for Cochrane-Orcutt-regression (Eq. 2) due to the way autocorrelation of residuals is incorporated into the model. We note that for regular regression based on solving Eq. 1 Student's t-test would be inappropriate if the residuals are correlated. In addition, we tested validity of our responses by performing Monte-Carlo bootstrap simulations and obtained similar significant responses.

Since other forcings can also affect the northern wintertime atmosphere, we tested the robustness of the obtained EEP responses by calculating multiple linear regressions with the same Cochrane-Orcutt regression method in which explaining variables included EEP, sunspot number (as a proxy for varying solar UV irradiance), QBO, ENSO index and volcanic aerosols. The EEP responses in this case remained essentially the same as those obtained from regressions without the additional explaining variables. This is expected since effects of possible other factors are implicitly modeled by the autoregressive residual term. Furthermore, the EEP activity does not have significant correlations with the sunspot number or with the other explanatory variables mentioned above. Figure 1 shows the time series of February EEP, sunspot number, geomagnetic Ap index (not used here, but often in other studies as a proxy for EEP-NO_X) and QBO30. For February values, the correlation coefficient between EEP and sunspot number is 0.24 (p=0.16), while for other studied months correlation is even smaller. Correlation between EEP and sunspot number is even smaller for detrended data. Thus, these two solar variables are nearly independent of each other. In Figure 1 one can also see a long-term decline in EEP which does not affect the regression results as we have subtracted 31-year smoothly varying trends from all the variables, including EEP. Additionally, we tested if the EEP response is affected by anthropogenic ozone loss which is largest during cold winters. If the temperature decreases low enough in the polar lower stratosphere, polar stratospheric clouds (PSC) form and accelerate chemical ozone loss. If this occurs at the same time with high EEP the ozone loss may be incorrectly attributed to EEP. We tested this by performing the analysis by excluding the winters with exceptionally cold lower polar stratosphere in February (1987/1988, 1992/1993, 1995/1996, 1996/1997, 1999/2000, 2004/2005 and 2010/2011) and the EEP responses remained fundamentally the same.

We calculated regressions also separately in the two QBO phases, easterly (QBO-E) and westerly (QBO-W). Months when the QBO phase index was negative were classified to the easterly phase, and months with positive QBO to westerly phase. Months of one winter could end up to the two different QBO phases. However, the results are almost identical even if whole winters instead of separate months are sectioned into the two QBO phases. We also examined different lags of QBO and how they affect the QBO modulation of EEP effects (Section 5).

3 EEP effect on atmospheric variables

Figure 2 presents the EEP effects, obtained from the linear regression analysis, on zonal wind, temperature and ozone (rows, separately) in December to March (columns)

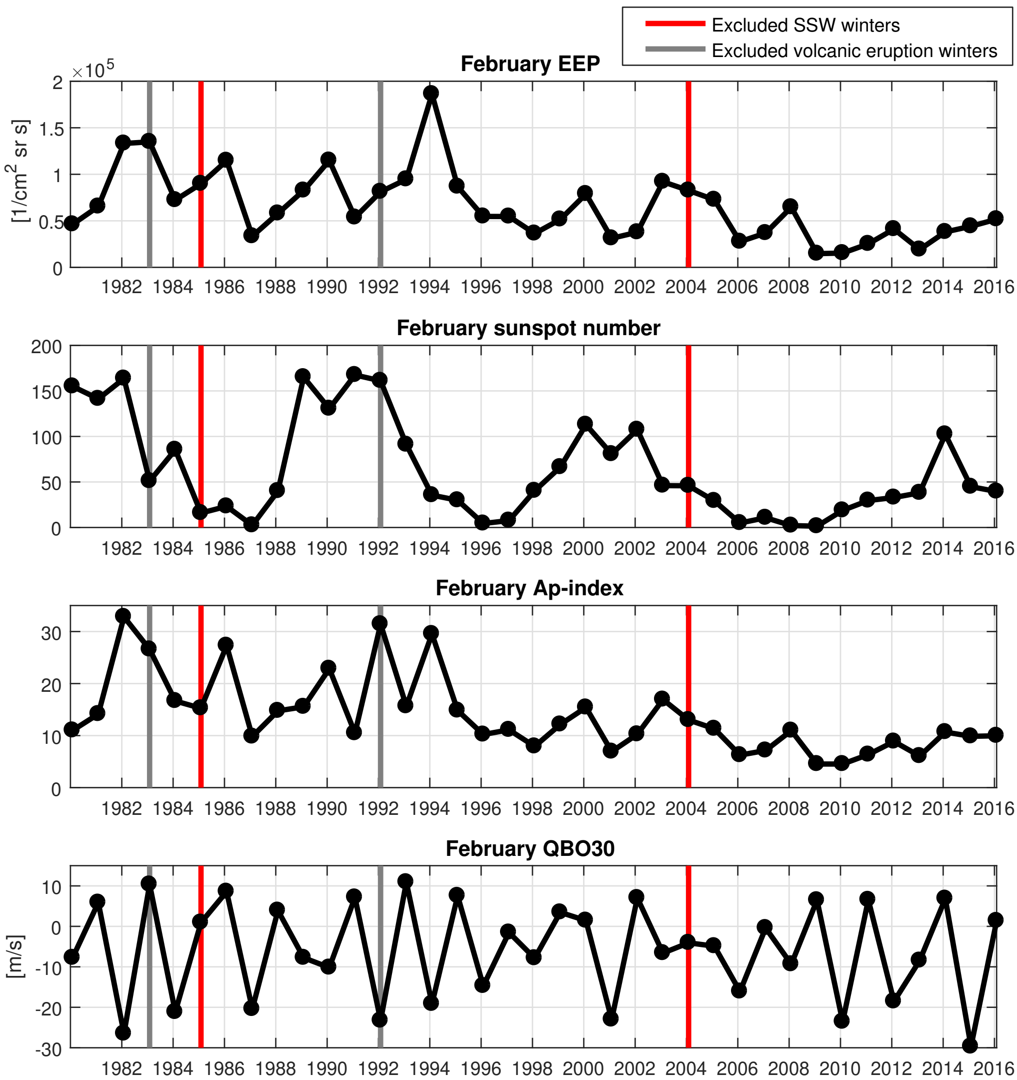


Figure 1. Time-series of February EEP (1. plot), sunspot number (2. plot), Ap-index (3. plot) and QBO30 (4. plot). Excluded SSW winters are marked with red lines and excluded winters after volcanic eruptions with grey line.

in the northern hemisphere. We used here a 1-month lag for EEP in February and March in order to account for the descent time of NO_X from the upper atmosphere [*Funke et al.*, 2014]. No EEP lag was used for December and January. We also tested other EEP lags between 0-2 months for all the winter months but found that although the responses were quite similar for all lags, they were strongest and most significant when using the above mentioned, quite plausible lags. Figure 2 (upper row) shows that an increase in EEP is related to a statistically significant strengthening of the polar vortex in every month. The westerly wind in the stratosphere strengthens at latitudes poleward of 50°N and this response seems to descend down from December to March. The strongest response is seen in January and February. There is a corresponding positive and significant zonal wind response even in the troposphere. Note also that at mid-latitudes (30°-50°) the zonal wind is weakened, probably because planetary waves are not diverted toward higher latitudes, but remain at low to mid-latitudes and slow down the generally westerly directed winds.

Temperature (Fig. 2, second row) displays a statistically significant decrease as a response to increased EEP in the lower stratosphere in December-March, while in the upper stratosphere the temperature increases in January-March. Ozone (Fig. 2, third row) shows a statistically significant decrease in the polar lower stratosphere in December-March. Note that the region of strongest ozone reduction moves systematically downward from December (about 10-20 hPa) to March (about 50-200 hPa) in concert with the descending wind and temperature responses. Because ERA-Interim ozone variations in the polar region in early winter largely result from dynamical changes in the model these EEP-related variations of ozone in the lower stratosphere (10-200 hPa) are probably mostly due to EEP-induced dynamical changes and not due to chemical loss. In February and March a weaker, but statistically significant, ozone decrease is also seen in the upper stratosphere (1-10 hPa). There is also a significant ozone increase at 10 hPa at low to mid-latitudes. This is consistent with observations by *Limpasuvan et al.* [2005], who showed that a stronger polar vortex at this altitude in December to February inhibits horizontal meridional circulation at high-latitudes. This would naturally lead to enhancement of ozone concentration at lower latitudes at this altitude.

The responses depicted in Fig. 2 can be explained in terms of the EEP-NO_X related ozone loss and resulting chemical and dynamical changes. The warming of the polar mesosphere and upper stratosphere (above 10 hPa) in mid-winter results mostly from reduction of radiative cooling [*Sinnhuber et al.*, 2018] and is associated to chemical ozone loss by NO_X. Figure 2 does not conclusively depict all the details of the resulting dynamical changes in the polar stratosphere. The exact mechanisms associated to the EEP effect thus remain subjects of future studies. However, at monthly time scales it is expected that the stratospheric winds and circulation readjust to maintain the thermal wind shear balance after the ozone loss induced warming. This readjustment is evidently associated with changes in planetary wave refraction and meridional circulation and leads to enhance the polar vortex below the region of initial warming in the upper stratosphere. As shown by *Limpasuvan et al.* [2005] a vortex enhancement is associated with more planetary waves being refracted from the lower stratosphere into the upper stratosphere. This leads to anomalous wave convergence (divergence) in the upper (lower) stratosphere. The anomalous convergence leads to enhanced downwelling and anomalous adiabatic heating in the upper stratosphere (further enhancing the heating there). The anomalous divergence in the lower stratosphere leads to weakened downwelling and anomalous adiabatic cooling. Such a dynamical cooling of the stratosphere associated with enhanced polar vortex has also been suggested by other studies [e.g. *Baumgaertner et al.*, 2011; *Arsenovic et al.*, 2016]. Reduced downwelling also results in a dynamical reduction of ozone in the polar lower stratosphere below the altitude of ozone mass mixing ratio maximum. These results are in accordance with previous modeling studies [e.g., *Rozanov et al.*, 2005; *Baumgaertner et al.*, 2011; *Arsenovic et al.*, 2016]. Results of Fig. 2 are also in agreement with the observational study of *Seppälä et al.* [2013], even though *Seppälä et al.* [2013] also excludes other major mid-winter SSW years and covers a partly different time period (1957-2008). However, *Seppälä et al.* [2013] results are weaker and statistically somewhat less significant than ours.

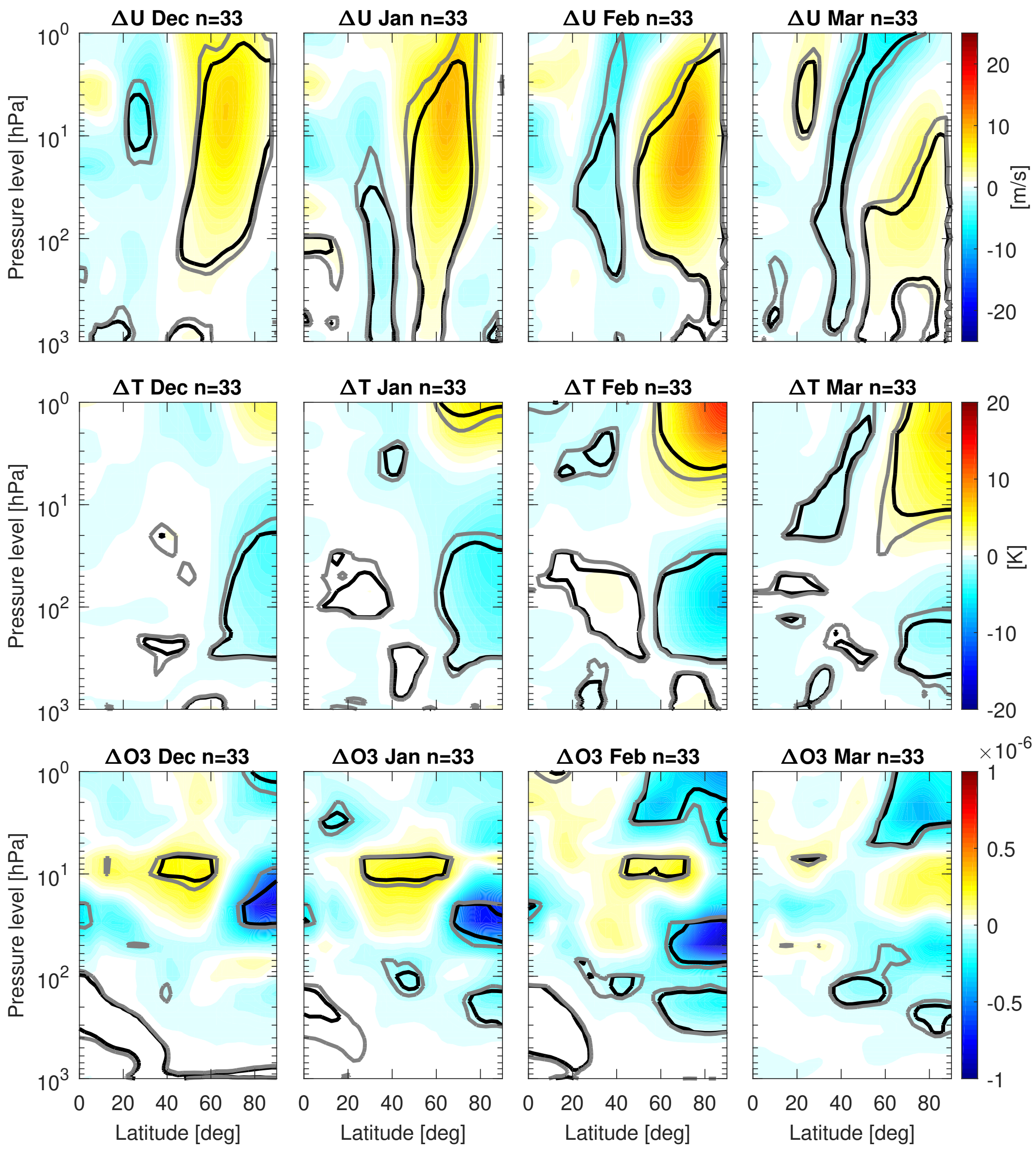


Figure 2. EEP responses in zonal wind (first row), temperature (second row) and ozone (third row) in December-March (columns 1-4). Responses are obtained from linear regression and correspond to a change of one standard deviation in EEP. Yellow and red colors correspond to positive response and blue to negative response. A positive response in zonal wind equals to a strengthening of westerly wind. Black contours correspond to 95% significance level and grey contours to 90% level.

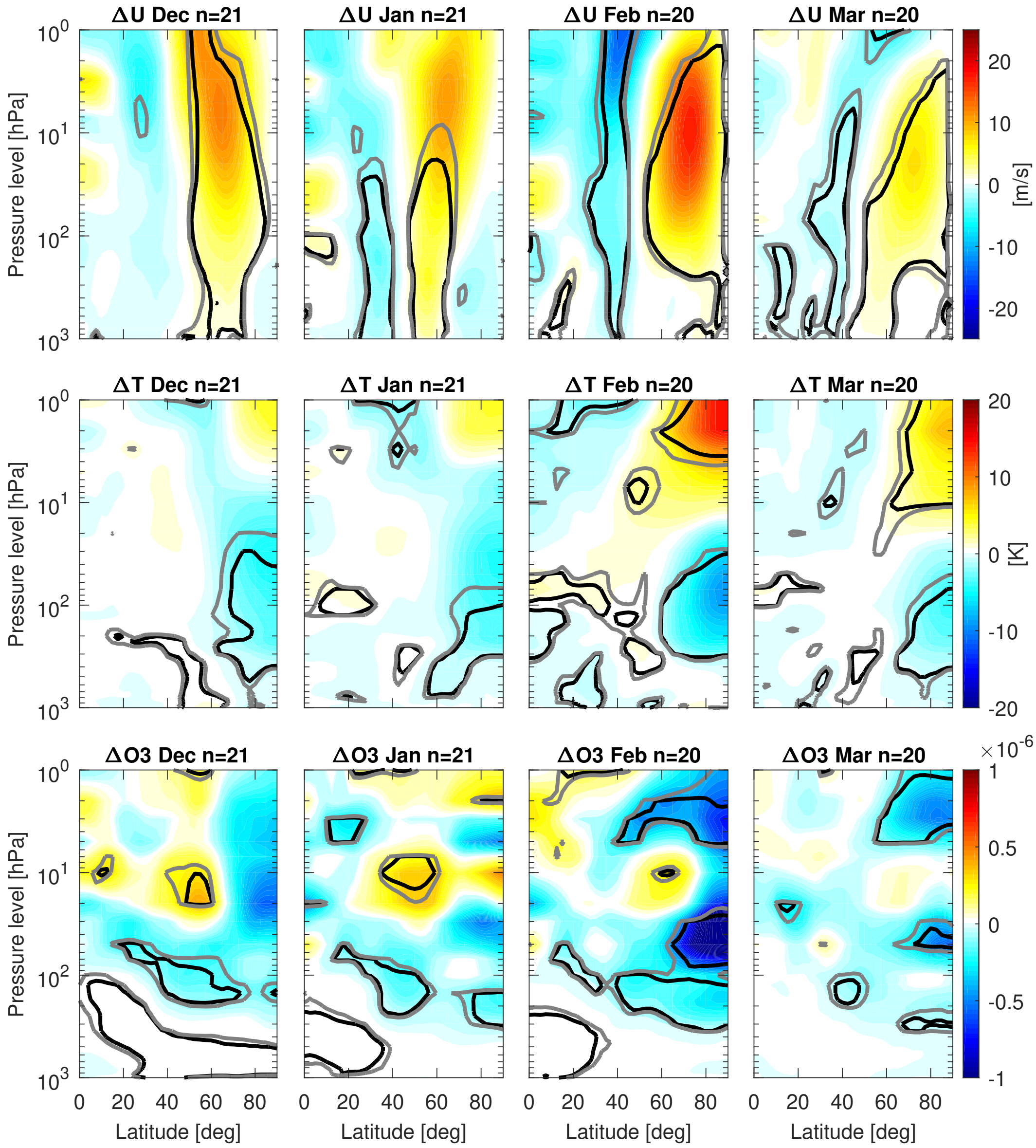
4 EEP effect separately in the two QBO phases

Figures 3 and 4 show the EEP effects on the same three atmospheric variables for the easterly and westerly QBO phase respectively. We used a 1-month lag for EEP in February and March, as in the previous case. Moreover, we used a lag of 6 months in the QBO when dividing months to the two QBO phases. With this lag the EEP responses in the easterly QBO phase are strongest, although similar responses also appear if we use 0-5 or 7-8 months lags for the QBO (see Section 5). In QBO-E (Figure 3) the EEP responses are quite similar to those in the case of no QBO phase separation (Figure 2), including the strengthening of the polar vortex and similar temperature and ozone patterns. Overall, there is some more variability in the response from one month to another in Fig. 3 than Fig. 2, with largest responses in all three parameters in February. The polar vortex is even more strongly and systematically enhanced in Fig. 3 than in Fig. 2, especially in February. Even the decrease of zonal winds at mid-latitudes is somewhat more expressed for QBO-E, especially in February. Figure 3 shows the same reduction of the altitude of the main ozone response as seen in Fig. 2.

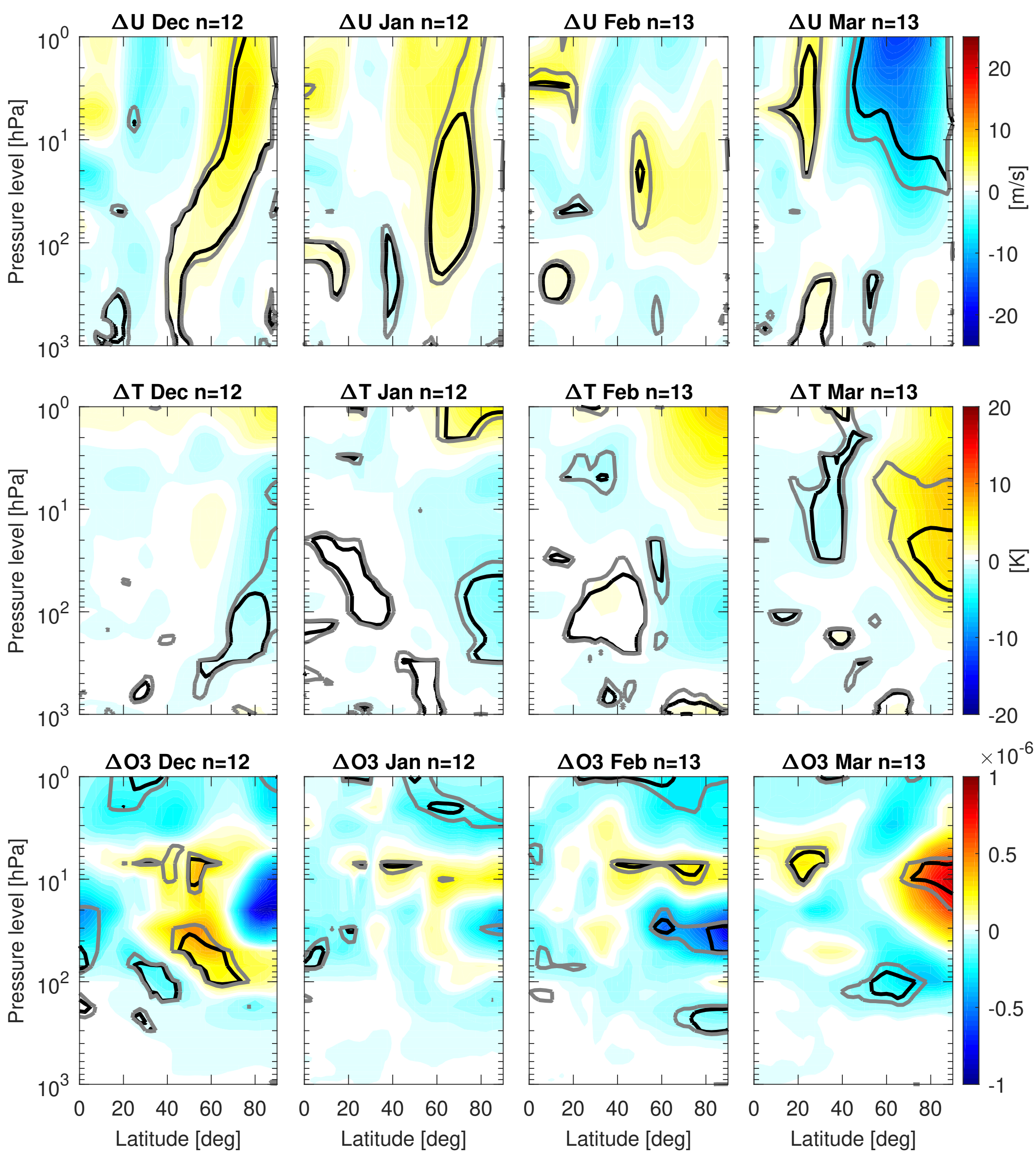
In QBO-W (Fig. 4) the polar vortex strengthens as a response to EEP only in December and January, but in February this response vanishes and in March the polar vortex even weakens in the upper polar stratosphere. Temperature and ozone responses in QBO-W are overall rather weak and spatially limited. Temperature decreases in the polar lower stratosphere in December and January, but increases in the upper stratosphere in January and at 50 hPa - 10 hPa in March. Ozone decreases in the polar lower stratosphere in February, but increases in March, corresponding to the warming in the same region. It is clear that the EEP responses are larger and statistically more significant in the QBO-E phase than in the QBO-W, especially in late winter. Note that even though QBO-E phase typically has somewhat more data points than QBO-W phase this has been automatically taken into account in the estimation of statistical significance and does not compromise the significance of the results. Note also that, on one hand, we do not find any strong responses to EEP in either QBO phase, which would be statistically insignificant. Such signals would obviously be spurious, and might result, e.g., from a too small number of datapoints. On the other hand, we do find weaker responses in QBO-W than in QBO-E, which still are statistically significant. Also this indicates that the number of data points in both QBO phases is sufficient to reveal the significance of the found responses. We also tested whether the differences in the EEP responses between the two QBO phases are statistically significant. Using both the Student's t-test and Monte-Carlo bootstrapping we found these differences to be significant throughout the winter and especially in February and March.

These results agree with the studies by *Palamara and Bryant* [2004] and *Maliniemi et al.* [2013, 2016] which concluded that the EEP effect on the NAO/NAM is larger in the QBO-E phase. *Lu et al.* [2008] found that the EEP effect in May is stronger if the de-seasonalized QBO at 50 hPa is westerly. We did not study the spring season here, but for March our findings do not agree with *Lu et al.* [2008]. This may be due to different QBO definitions, time periods, and the fact that *Lu et al.* [2008] excluded all major SSW years from their analysis. *Seppälä et al.* [2013] used a combination of ERA-40 and ERA-Interim re-analyses to study stratospheric response to geomagnetic activity. They also studied the QBO (de-seasonalized at 50 hPa) dependence of the zonal wind response to geomagnetic activity in December and found quite similar patterns for both QBO phases as here. However, their positive zonal wind responses were somewhat stronger and more significant in QBO-W phase. This difference is largely produced by the different treatment of the QBO phase and the fact that *Seppälä et al.* [2013] also included years with large SSWs into their QBO phase composites. *Seppälä et al.* [2013] did not show QBO phase dependent results for late winter. However, we found (not shown here) that using de-seasonalized QBO at 50 hPa with lags of 0-6 months, the EEP-related signals are much stronger in February and March for QBO-E than for QBO-W, similarly to the 6-month lagged 30 hPa QBO used here.

Figure 5a shows a scatter plot between the February average zonal wind and January EEP in QBO-E phase (indicating also the major SSW and volcanic years, which were excluded from the analysis). Figures 5b-d show standardized time series of January EEP together with February average zonal wind (b), temperature (c) and ozone (d) in QBO-E winters. The averages have been computed over the regions where the observed signals are strongest and statistically most significant in Figure 3. Note that in Figs. 5c-d the sign of the EEP time series has been inverted to allow a better comparison with temperature and ozone. The correlation coefficients between EEP and U,T and O3 are 0.64 ($p=0.002$), -0.49 ($p=0.03$) and -0.42 ($p=0.05$) respectively. This further corroborates the validity of the signals observed with regression.



362 **Figure 3.** EEP responses in zonal wind, temperature and ozone in December-March in the QBO-E phase.
363 QBO is taken with a 6 month lag. All notations are as in Figure 2.



364 **Figure 4.** Same as Figure 3, but for the QBO-W phase.

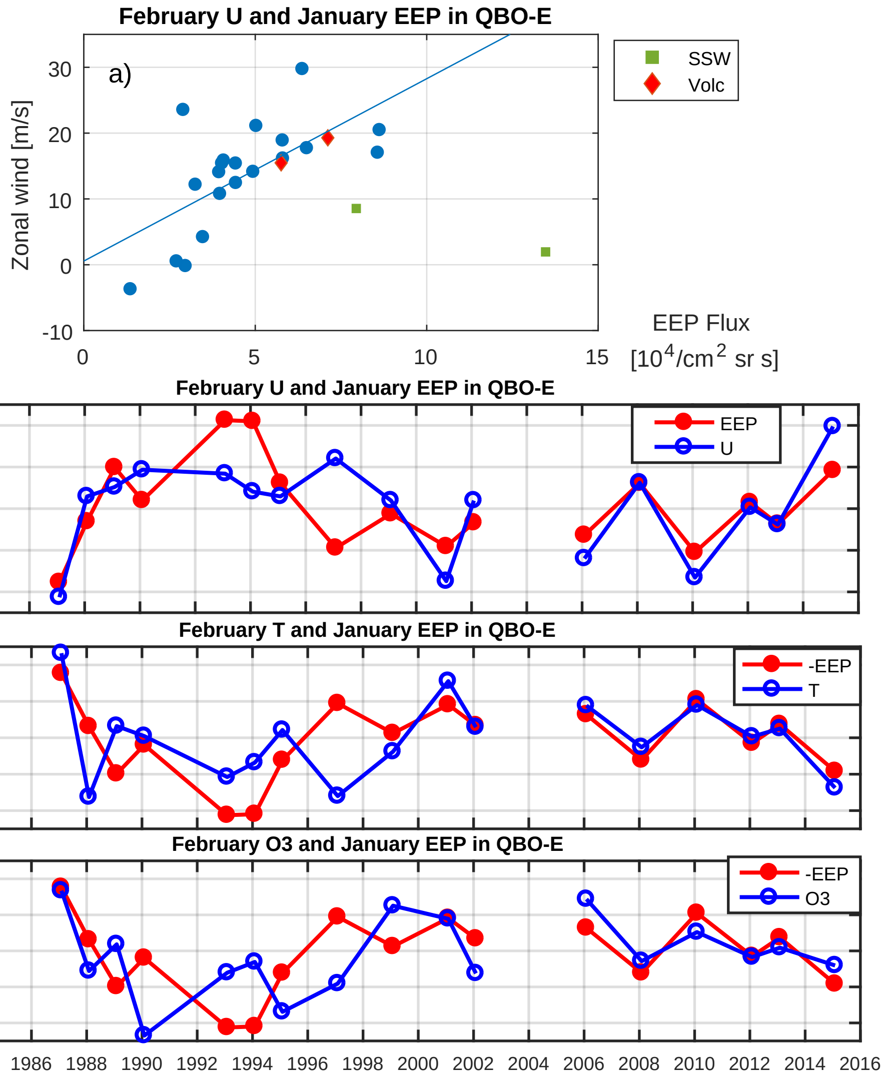


Figure 5. a) Scatter plot of February zonal wind (average over 60°-90°N and 30-100 hPa) and January EEP in QBO-E phase. Years with major SSWs (green squares) and volcanic eruptions (red diamonds) are also indicated. The blue line shows the regression fit to the datapoints excluding volcanic and SSW years. Panels b-d show the standardized time-series of January EEP together with February zonal wind (b), negative January EEP and February polar temperature (60°-90°N, 50-100 hPa) (c), and negative January EEP and February polar ozone (70°-90°N, 30-70 hPa) (d) in QBO-E phase. QBO is taken with a 6 month lag.

5 Dependence on QBO lag

We found that the responses to EEP vary notably in both QBO phases if we use lagged QBO values in separating the months to the two phases. Figures 6 and 7 present the EEP effect on 30 hPa zonal wind as a function of latitude and QBO lag in the QBO-E and QBO-W phase, respectively. Positive lag values correspond to QBO values of previous months, while negative lags correspond to QBO values of following months. The EEP effects in QBO-E (Fig. 6) are largest and significant in every month (Dec-Mar) if the QBO lag is about 6 months. In QBO-W (Fig. 7) the responses are significant with a 6-month QBO lag only in December-January. In both QBO phases the response patterns recur after 25-30 months which corresponds to the QBO period. Moreover, the responses in QBO-E phase with lags of, e.g., 0-10 months are seen also in QBO-W with lags of 15-25 months, half a QBO period later/earlier.

QBO is known to affect the Brewer-Dobson circulation [*Flury et al.*, 2013] which brings air including, e.g., ozone to polar stratosphere. Therefore, QBO-caused variations in the Brewer-Dobson circulation should affect the amount of polar ozone, as previously reported, e.g., by *Salby and Callaghan* [2002]. Figure 8 shows the difference in polar ozone mass mixing ratio (60°-90°N) between the QBO-E and QBO-W winter months with different QBO lags (0-9 months; different panels). The mass mixing ratio is significantly higher in the lower stratosphere, where the ozone density is highest, in QBO-E winter months when the QBO is lagged by 3-9 months. As the lag increases from 3 to 9, the positive difference moves forward in time with the corresponding lag time. With a 6-7 -month lag there is significantly more ozone over the whole winter (Dec-Mar) for the QBO-E phase compared to QBO-W. (We note that we did not obtain similarly large and systematic differences in ozone between the two QBO phases by using QBO from altitudes lower than 30 hPa with shorter lags.)

The Brewer-Dobson circulation takes at least 6 months to transport ozone from the equator to latitudes close the polar region [*Birner and Bönisch*, 2011]. The QBO-related secondary circulation cell in the tropics affects the large-scale circulation so that the ascent rates in the equatorial stratosphere are highest in QBO-E phase [*Flury et al.*, 2013]. Therefore, the QBO in summer affects how much ozone is transported to the pole for next winter. Together Figures 6, 7 and 8 imply that the EEP effects to zonal wind are largest when the transport of ozone to the polar stratosphere by meridional circulation is strongest.

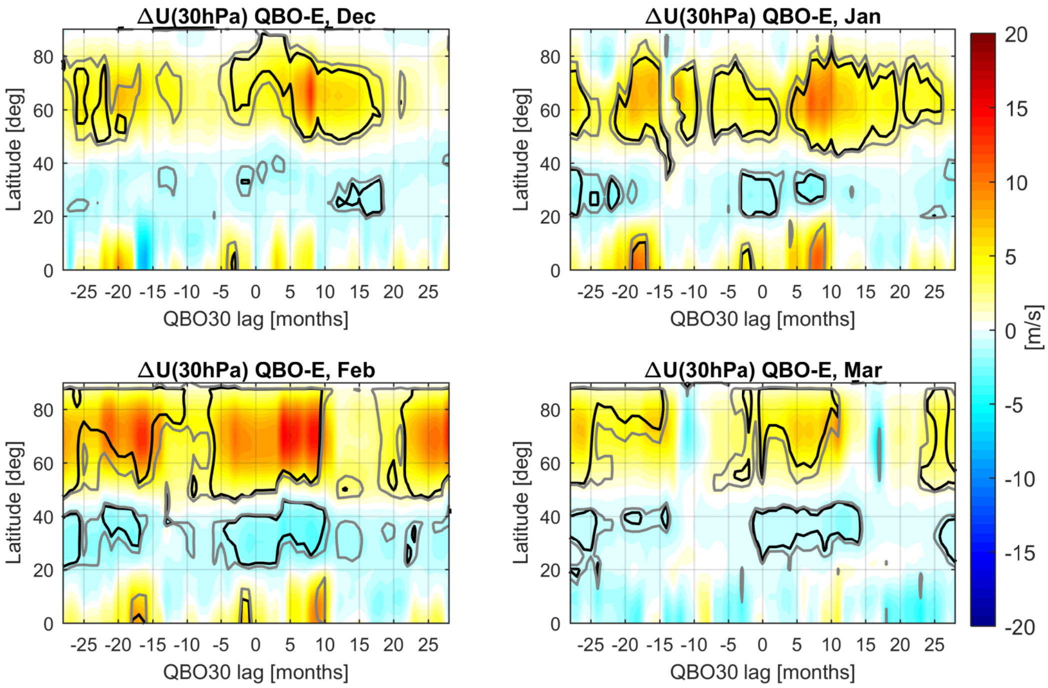


Figure 6. EEP effect on zonal wind at 30 hPa in QBO-E phase (QBO at 30 hPa) in December-March as a function of latitude and QBO lag in December-March. All notations are as in Figure 2.

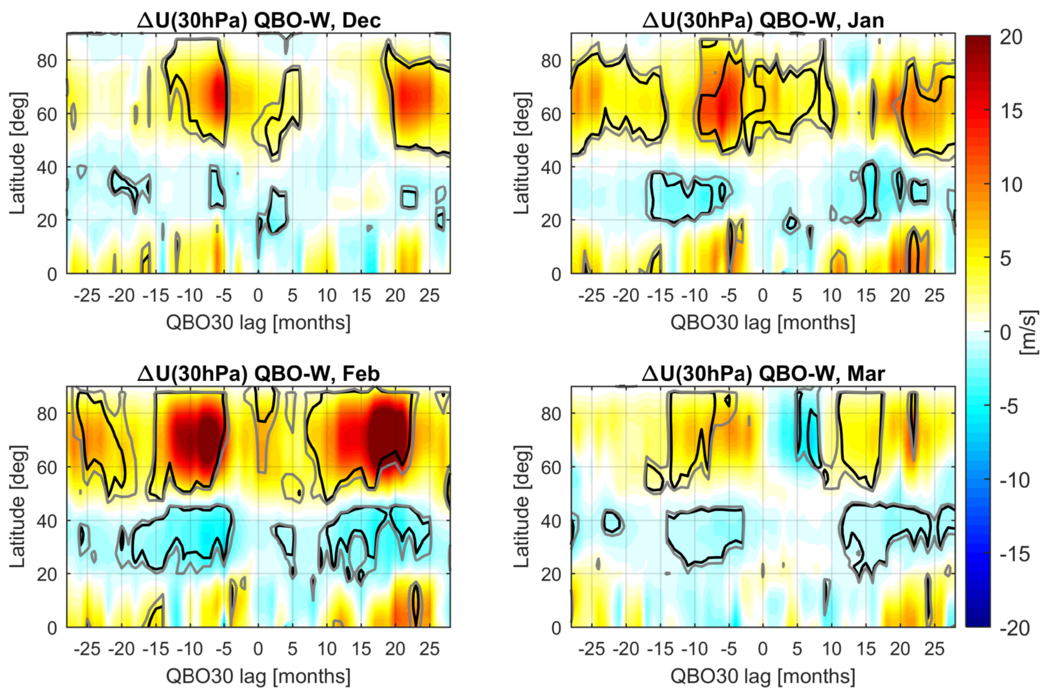
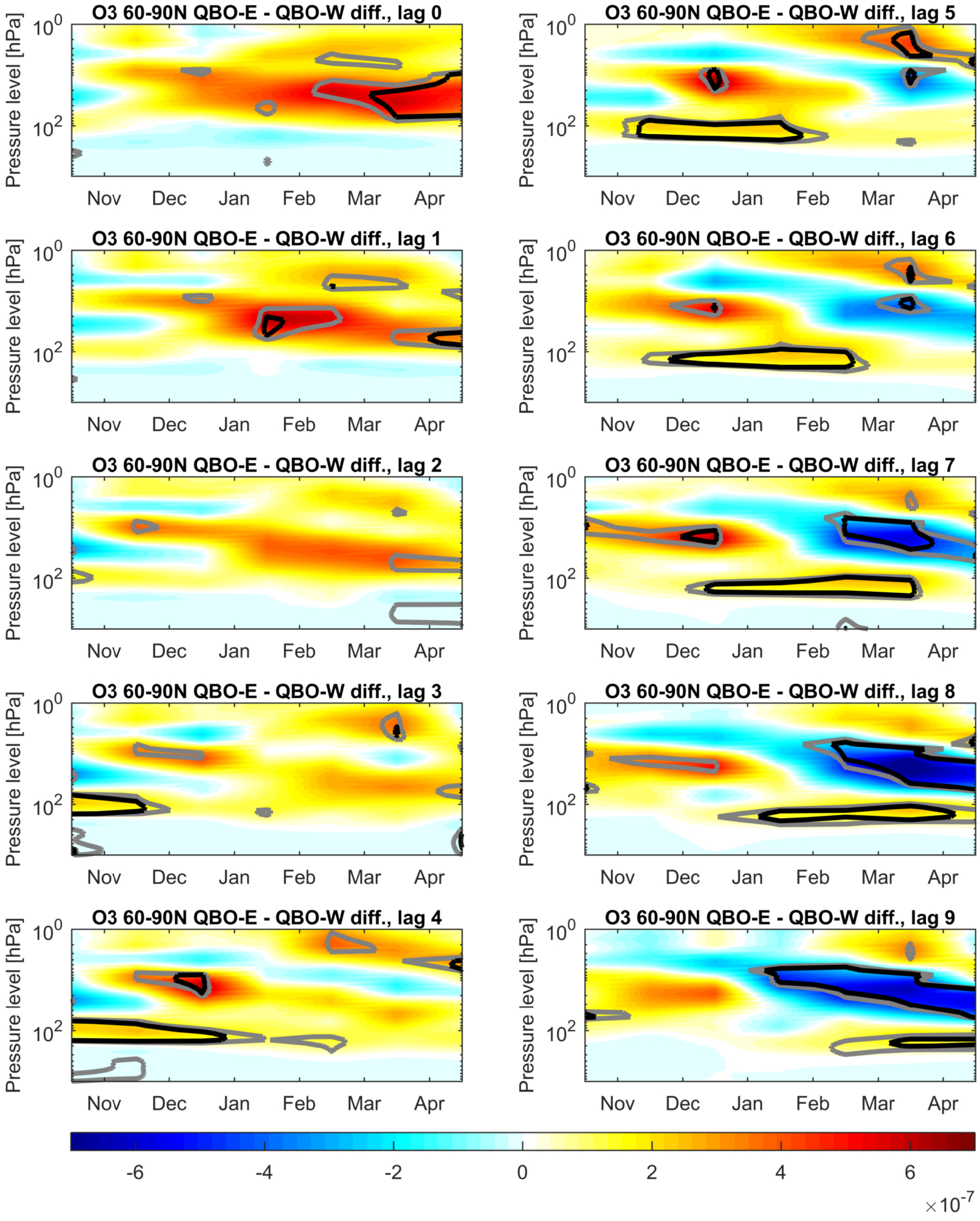


Figure 7. Same as Figure 6, but for the QBO-W phase.



406 **Figure 8.** Difference in polar ozone mass mixing ratio between the QBO-E and QBO-W phases as a func-
407 tion of height and time (November-April). QBO is taken with a lag of 0-9 months, corresponding plots from
408 the top left to bottom right. All notations are as in Figure 2.

6 Discussion and Conclusions

The results obtained here strongly suggest that the energetic electron precipitation affects the northern polar middle atmosphere, and the effect is seen in stratospheric zonal wind, temperature and ozone in winter (December-March). Increased EEP is related to a significant strengthening of the polar vortex, warming in the upper polar stratosphere (in late winter) and cooling in the lower polar stratosphere, and ozone reduction both in the lower and upper (late winter) stratosphere.

In late winter in the upper stratosphere the EEP-associated ozone loss is most likely chemical loss related to EEP-NO_X, which typically descend during the dark polar winter into the upper polar stratosphere at 1-10 hPa [*Funke et al.*, 2014]. In early winter (Dec-Jan) the EEP-associated chemical ozone loss is probably restricted mostly to mesosphere above 1 hPa and thus remains unseen in ERA-Interim data. In mid-winter ozone loss in mesosphere/upper stratosphere leads to a warming there since ozone acts there as a radiative cooler [*Sinnhuber et al.*, 2018]. In monthly time scale this warming is expected to lead into readjustment of zonal winds in order to fulfill the thermal wind shear balance. It is also associated with changes in wave propagation and meridional circulation. As shown by *Limpasuvan et al.* [2005] a vortex enhancement is associated with more planetary waves being refracted from the lower stratosphere into the upper stratosphere. This leads to anomalous wave convergence (divergence) in the upper (lower) stratosphere. The anomalous divergence in the lower stratosphere leads to weakened downwelling and anomalous adiabatic cooling. Our findings are consistent with this interpretation. The weakened downwelling also leads to reduction of ozone in the lower stratosphere, which is also observed here. It is most likely that, by these dynamical processes, the polar vortex can be enhanced even in the early winter before EEP-NO_X effect has reached the upper stratosphere. Similar EEP related variations in the atmosphere have been found in previous studies [e.g., *Palamara and Bryant*, 2004; *Rozanov et al.*, 2005; *Baumgaertner et al.*, 2011; *Maliniemi et al.*, 2013; *Seppälä et al.*, 2013; *Arsenovic et al.*, 2016].

Earlier studies [e.g., *Langematz et al.*, 2003; *Sinnhuber et al.*, 2018; *Meraner and Schmidt*, 2018] have suggested that ozone loss warms radiatively the upper stratosphere in mid-winter. Many studies [e.g., *Langematz et al.*, 2003; *Baumgaertner et al.*, 2011] have also given indications that the lower stratosphere cools dynamically after the initial ozone depletion in the upper stratosphere. *Arsenovic et al.* [2016] found a significant EEP-related ozone loss in the wintertime polar stratosphere, which was associated with warming in the mesosphere/upper stratosphere and cooling in the lower stratosphere. They suggested that the ozone reduction by EEP cools the mesosphere radiatively in sunlit regions, which enhances the polar vortex. The associated enhanced downwelling in the mesosphere would then lead to adiabatic warming in the mesosphere/upper stratosphere, and the weakening of downwelling in the lower stratosphere would in turn lead to relative adiabatic cooling.

Meraner and Schmidt [2018] studied the effect of mesospheric and stratospheric polar ozone loss in a simplified model experiment. They found similar temperature and zonal wind responses to ozone loss, but much weaker and statistically insignificant. Their result deviates from many observational and modeling studies, which have identified statistically significant signals [*Rozanov et al.*, 2005; *Baumgaertner et al.*, 2011; *Maliniemi et al.*, 2013; *Seppälä et al.*, 2013; *Arsenovic et al.*, 2016]. It is so far unclear what causes these differences. Model simulations by *Meraner and Schmidt* [2018] covered a 150-year period which is longer than in previous studies and may contribute to the differences. However, another possible cause might be that *Meraner and Schmidt* [2018] studied prescribed constant ozone loss in the polar mesosphere or in upper stratosphere in all months of the year. In fact, the descending EEP-NO_X affects both regions in winter [*Sinnhuber et al.*, 2018], and the combined effect of mesospheric and stratospheric ozone loss may not be just a sum of separate effects studied by *Meraner and Schmidt* [2018]. For example, ozone loss in the mesosphere allows more solar UV-radiation to penetrate into the stratosphere and affects radiative response of stratosphere.

We found that the EEP responses are strengthened if only the QBO-E winter months are considered. On the other hand, in QBO-W months these effects are weaker and less significant, especially in February-March. *Palamara and Bryant* [2004] and *Maliniemi et al.* [2013] found that the correlation between geomagnetic activity (proxy of EEP activity) and NAO was apparent when the QBO at 30 hPa was easterly, which agrees with our results. We showed here that there is significantly more ozone in the polar lower stratosphere in QBO-E winter months than in QBO-W if the QBO is taken with a roughly 6-month lag, which agrees well with the 6-12 months time it takes for ozone to be transported from the equator to polar latitudes [*Birner and Bönisch*, 2011]. Accordingly, the amount of ozone transported to the polar stratosphere can be used as an indicator for the strength of meridional circulation [*Salby and Callaghan*, 2002]. Studying different QBO lags we found that the EEP responses of zonal wind (and associated temperature and ozone changes) are also largest in QBO-E with the roughly 6-month QBO lag. This strongly suggests that the meridional circulation is involved in the observed QBO-modulation of the EEP-effect on polar stratosphere.

One possibility by which QBO may modulate the EEP effect on polar stratosphere is that the enhanced meridional circulation in QBO-E transports NO_X and ozone more efficiently down into the mesosphere and upper stratosphere. Since the rate of ozone loss is proportional to the concentrations of both NO_X and ozone, the ozone loss and associated radiative and dynamical responses would be larger during stronger meridional circulation of the QBO-E. Another important consequence of stronger meridional circulation in QBO-E is that the polar stratosphere is warmer on average than in QBO-W. Easterly QBO winds are known to channel more planetary wave activity into the polar stratosphere [*Holton and Tan*, 1980]. Both of these factors will make the polar vortex more susceptible to wave-mean-flow-interaction [*Andrews*, 1985; *Scott and Polvani*, 2004], which allows the EEP-induced dynamical effects to efficiently propagate downwards in the stratosphere.

The atmospheric responses to EEP were here found to be strong and statistically significant. However, as these results are based on a relatively short and special (very high solar activity) period of time we do not know how large the EEP effect and its QBO modulation may have been in other time periods in the past (e.g. some 100 years ago when solar activity was much weaker). Thus, the exact mechanism of the QBO modulation of EEP effects require further research, and these issues should be carefully considered in future studies of the polar stratosphere. Also, to fully understand the atmospheric variations driven by Sun-related effects, both the EEP (due to solar wind) and solar UV-radiation should be included in this context.

Acknowledgments

We acknowledge the financial support by the Academy of Finland to the ReSoLVE Centre of Excellence (project no. 307411). We thank ECMWF for providing the ERA-Interim reanalysis data (https://www.ecmwf.int/en/forecasts/datasets/archive-datasets/reanalysis-datasets/era-interim) and NOAA/NGDC for providing the original MEPED energetic electron flux measurements (https://www.ngdc.noaa.gov/stp/satellite/poes/index.html).

References

Andrews, D. G. (1985), Wave–mean-flow interaction in the middle atmosphere, in *Adv. Geophys.*, vol. 28, pp. 249–275, Elsevier, doi:10.1016/S0065-2687(08)60226-5.

Arsenovic, P., E. Rozanov, A. Stenke, B. Funke, J. M. Wissing, K. Mursula, F. Tummon, and T. Peter (2016), The influence of Middle Range Energy Electrons on atmospheric chemistry and regional climate, *J. Atmos. Sol.-Terr. Phys.*, *149*, 180–190, doi: 10.1016/j.jastp.2016.04.008.

Asikainen, T., and K. Mursula (2013), Correcting the NOAA/MEPED energetic electron fluxes for detector efficiency and proton contamination, *J. Geophys. Res. Space*, *118*(10), 6500–6510, doi:10.1002/jgra.50584.

Asikainen, T., and M. Ruopsa (2016), Solar wind drivers of energetic electron precipitation, *J. Geophys. Res. Space*, *121*(3), 2209–2225, doi:10.1002/2015JA022215.

Baldwin, M. P., and T. J. Dunkerton (2001), Stratospheric harbingers of anomalous weather regimes, *Science*, *294*(5542), 581–584, doi:10.1126/science.1063315.

Bame, S., J. Asbridge, W. Feldman, and J. Gosling (1976), Solar cycle evolution of high-speed solar wind streams, *Astrophys. J.*, *207*, 977–980, doi:10.1086/154566.

Baumgaertner, A. J., A. Seppälä, P. Jöckel, and M. A. Clilverd (2011), Geomagnetic activity related NOx enhancements and polar surface air temperature variability in a chemistry climate model: Modulation of the NAM index, *Atmos. Chem. Phys.*, *11*(9), 4521–4531, doi:10.5194/acp-11-4521-2011.

Birner, T., and H. Bönisch (2011), Residual circulation trajectories and transit times into the extratropical lowermost stratosphere, *Atmos. Chem. Phys.*, *11*(2), 817–827, doi: 10.5194/acp-11-817-2011.

Butchart, N. (2014), The brewer-dobson circulation, *Rev. Geophys.*, *52*(2), 157–184, doi: 10.1002/2013RG000448.

Charney, J. G., and P. G. Drazin (1961), Propagation of planetary-scale disturbances from the lower into the upper atmosphere, *J. Geophys. Res.*, *66*(1), 83–109, doi: 10.1029/JZ066i001p00083.

Cleveland, W. S. (1979), Robust locally weighted regression and smoothing scatterplots, *J. Am. Stat. Assoc.*, *74*(368), 829–836, doi:10.2307/2286407.

Cochrane, D., and G. H. Orcutt (1949), Application of least squares regression to relationships containing auto-correlated error terms, *J. Am. Stat. Assoc.*, *44*(245), 32–61, doi:10.1080/01621459.1949.10483290.

Crutzen, P. J., I. S. Isaksen, and G. C. Reid (1975), Solar proton events: Stratospheric sources of nitric oxide, *Science*, *189*(4201), 457–459, doi:10.1126/science.189.4201.457.

Damiani, A., B. Funke, M. LÃşpez Puertas, M. L. Santee, R. R. Cordero, and S. Watanabe (2016), Energetic particle precipitation: A major driver of the ozone budget in the antarctic upper stratosphere, *Geophysical Research Letters*, *43*(7), 3554–3562, doi:10.1002/2016GL068279.

Dee, D., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Q. J. R. Meteorol. Soc.*, *137*(656), 553–597, doi:10.1002/qj.828.

Dragani, R. (2011), On the quality of the ERA-Interim ozone reanalyses: Comparisons with satellite data, *Q. J. R. Meteorol. Soc.*, *137*(658), 1312–1326, doi:10.1002/qj.821.

Floyd, L., J. Cook, L. Herring, and P. Crane (2003), SUSIM's 11-year observational record of the solar UV irradiance, *Adv. Space Res.*, *31*(9), 2111–2120, doi:10.1016/S0273-1177(03)00148-0.

Flury, T., D. L. Wu, and W. Read (2013), Variability in the speed of the Brewer-Dobson circulation as observed by Aura/MLS, *Atmos. Chem. Phys.*, *13*(9), 4563–4575, doi:10.5194/acp-13-4563-2013.

Funke, B., M. López-Puertas, S. Gil-López, T. Von Clarmann, G. Stiller, H. Fischer, and S. Kellmann (2005), Downward transport of upper atmospheric NOx into the polar stratosphere and lower mesosphere during the Antarctic 2003 and Arctic 2002/2003 winters, *J. Geophys. Res.: Atmos.*, *110*(D24), doi:10.1029/2005JD006463.

Funke, B., M. López-Puertas, G. Stiller, and T. Clarmann (2014), Mesospheric and stratospheric NOy produced by energetic particle precipitation during 2002–2012, *J. Geophys. Res.: Atmos.*, *119*(7), 4429–4446, doi:10.1002/2013JD021404.

Holton, J. R., and H.-C. Tan (1980), The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb, *J. Atmos. Sci.*, *37*(10), 2200–2208.

Labitzke, K., and H. Van Loon (1988), Associations between the 11-year solar cycle, the QBO and the atmosphere. Part I: the troposphere and stratosphere in the northern hemisphere in winter, *J. Atmos. Terr. Phys.*, *50*(3), 197–206, doi:10.1016/0021-9169(88)90068-2.

Langematz, U., M. Kunze, K. Krüger, K. Labitzke, and G. L. Roff (2003), Thermal and dynamical changes of the stratosphere since 1979 and their link to ozone and CO2 changes, *J. Geophys. Res.: Atmos.*, *108*(D1), doi:10.1029/2002JD002069.

Limpasuvan, V., D. L. Hartmann, D. W. J. Thompson, K. Jeev, and Y. L. Yung (2005), Stratosphere-troposphere evolution during polar vortex intensification, *J. Geophys. Res.: Atmos.*, *110*(D24), doi:10.1029/2005JD006302.

Lu, H., M. A. Clilverd, A. Seppälä, and L. L. Hood (2008), Geomagnetic perturbations on stratospheric circulation in late winter and spring, *J. Geophys. Res.: Atmos.*, *113*(D16), doi:10.1029/2007JD008915.

Maliniemi, V., T. Asikainen, K. Mursula, and A. Seppälä (2013), QBO-dependent relation between electron precipitation and wintertime surface temperature, *J. Geophys. Res.: Atmos.*, *118*(12), 6302–6310, doi:10.1002/jgrd.50518.

Maliniemi, V., T. Asikainen, and K. Mursula (2016), Effect of geomagnetic activity on the northern annular mode: QBO dependence and the Holton-Tan relationship, *J. Geophys. Res.: Atmos.*, *121*(17), doi:10.1002/2015JD024460.

Manney, G. L., K. Krüger, J. L. Sabutis, S. A. Sena, and S. Pawson (2005), The remarkable 2003–2004 winter and other recent warm winters in the arctic stratosphere since the late 1990s, *J. Geophys. Res.: Atmos.*, *110*(D4), doi:10.1029/2004JD005367.

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

Meredith, N. P., R. B. Horne, M. M. Lam, M. H. Denton, J. E. Borovsky, and J. C. Green (2011), Energetic electron precipitation during high-speed solar wind stream driven storms, *J. Geophys. Res. Space*, *116*(A5), doi:10.1029/2010JA016293.

Mursula, K., R. Lukianova, and L. Holappa (2015), Occurrence of high-speed solar wind streams over the Grand Modern Maximum, *Astrophys. J.*, *801*(1), 30, doi:10.1088/0004-637X/801/1/30.

Palamara, D., and E. Bryant (2004), Geomagnetic activity forcing of the northern annular mode via the stratosphere, *Ann. Geophys.*, *22*, 725–731, doi:10.1088/0004-637X/801/1/30.

Randall, C., V. Harvey, C. Singleton, S. Bailey, P. Bernath, M. Codrescu, H. Nakajima, and J. Russell (2007), Energetic particle precipitation effects on the Southern Hemisphere stratosphere in 1992–2005, *J. Geophys. Res.: Atmos.*, *112*(D8), doi:

10.1029/2006JD007696.

Rozanov, E., L. Callis, M. Schlesinger, F. Yang, N. Andronova, and V. Zubov (2005), Atmospheric response to NOy source due to energetic electron precipitation, *Geophys. Res. Lett.*, *32*(14), doi:10.1029/2005GL023041.

Salby, M. L., and P. F. Callaghan (2002), Interannual Changes of the Stratospheric Circulation: Relationship to Ozone and Tropospheric Structure, *J. Clim.*, *15*(24), 3673–3685, doi:10.1175/1520-0442(2003)015<3673:ICOTSC>2.0.CO;2.

Scott, R. K., and L. M. Polvani (2004), Stratospheric control of upward wave flux near the tropopause, *Geophys. Res. Lett.*, *31*(2), doi:10.1029/2003GL017965.

Seppälä, A., H. Lu, M. Clilverd, and C. Rodger (2013), Geomagnetic activity signatures in wintertime stratosphere wind, temperature, and wave response, *J. Geophys. Res.: Atmos.*, *118*(5), 2169–2183, doi:10.1002/jgrd.50236.

Sinnhuber, M., U. Berger, B. Funke, H. Nieder, T. Reddmann, G. Stiller, S. Versick, T. von Clarmann, and J. M. Wissing (2018), NOy production, ozone loss and changes in net radiative heating due to energetic particle precipitation in 2002–2010, *Atmos. Chem. Phys.*, *18*(2), 1115, doi:10.5194/acp-18-1115-2018.

Tomikawa, Y. (2017), Response of the middle atmosphere in the southern hemisphere to energetic particle precipitation in the latest reanalysis data, *SOLA*, *13*(Special_Edition), 1–7.

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, *J. Atmos. Sol.-Terr. Phys.*, *71*(10), 1176–1189, doi:10.1016/j.jastp.2008.07.005.