Simulation of Energetic Particle Precipitation Effects During the 2003-2004 Arctic Winter

C. E. Randall
V. L. Harvey
L. A. Holt
D. R. Marsh
D. Kinnison

Follow this and additional works at: http://digitalcommons.odu.edu/chemistry_fac_pubs

Part of the Astrophysics and Astronomy Commons, Atmospheric Sciences Commons, Geophysics and Seismology Commons, and the Physical Chemistry Commons

Repository Citation
http://digitalcommons.odu.edu/chemistry_fac_pubs/67

Original Publication Citation
Simulation of energetic particle precipitation effects during the 2003–2004 Arctic winter

C. E. Randall¹-², V. L. Harvey², L. A. Holt³, D. R. Marsh⁴, D. Kinnison⁴, B. Funke⁵, and P. F. Bernath⁶

¹Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado, USA, ²Laboratory for Atmospheric and Space Physics, University of Colorado Boulder, Boulder, Colorado, USA, ³NorthWest Research Associates, Boulder, Colorado, USA, ⁴Atmospheric Chemistry Observations and Modeling, National Center for Atmospheric Research, Boulder, Colorado, USA, ⁵Solar System Department, Instituto de Astrofisica de Andalucia, Granada, Spain, ⁶Department of Chemistry and Biochemistry, Old Dominion University, Norfolk, Virginia, USA

Abstract

Energetic particle precipitation (EPP) during the 2003–2004 Arctic winter led to the production and subsequent transport of reactive odd nitrogen (NOx = NO + NO2) from the mesosphere and lower thermosphere (MLT) into the stratosphere. This caused NOx enhancements in the polar upper stratosphere in April 2004 that were unprecedented in the satellite record. Simulations of the 2003–2004 Arctic winter with the Whole Atmosphere Community Climate Model using Specified Dynamics (SD-WACCM) are compared to satellite measurements to assess our understanding of the observed NOx enhancements. The comparisons show that SD-WACCM clearly displays the descent of NOx produced by EPP but underestimates the enhancements by at least a factor of four. Comparisons with NO measurements in January and February indicate that SD-WACCM most likely underestimates EPP-induced NO production locally in the mesosphere because it does not include precipitation of high energy electrons. Comparisons with temperature measurements suggest that SD-WACCM does not properly simulate recovery from a sudden stratospheric warming in early January, resulting in insufficient transport from the MLT into the stratosphere. Both of these factors probably contribute to the inability of SD-WACCM to simulate the stratospheric NOx enhancements, although their relative importance is unclear. The work highlights the importance of considering the full spectrum of precipitating electrons in order to fully understand the impact of EPP on the atmosphere. It also suggests a need for high-quality meteorological data and measurements of NOx throughout the polar winter MLT.

1. Introduction

In the Arctic spring of 2004, an enormous influx of reactive odd nitrogen (NOx = NO + NO2) from the mesosphere and lower thermosphere (MLT) was observed to enter the polar stratosphere. NOx mixing ratios in the upper stratosphere increased by a factor of 4, causing localized catalytic reductions in ozone (O3) of more than 60% [Randall et al., 2005]. Never before and never since, in either hemisphere, has such a large influx of NOx been observed. The processes leading to this influx, which were initiated by energetic particle precipitation (EPP) and are described more below, probably occur routinely, although not simultaneously as in 2004. To estimate the likelihood that EPP effects similar to those in 2004 occurred before the satellite observational record, or will occur in the future, it is important that the underlying mechanisms be understood and simulated.

When the springtime 2004 stratospheric NOx enhancements were first reported, it was clear that the excess NOx had been produced by solar protons and/or precipitating electrons and then had been transported down into the stratosphere. Investigators initially speculated that the NOx was produced during the famous 2003 “Halloween” solar storms [Natarajan et al., 2004; Rinsland et al., 2005]. However, Randall et al. [2005] and López-Puertas et al. [2005] suggested that the timing of the influx, in late March and April more than 4 months after the Halloween storms, was also consistent with NOx production occurring later in 2003 or in early 2004.

Investigations following those initial reports have shown that the observational evidence overwhelmingly supports production after, not during, the Halloween storms. The most widely accepted explanation now is that the excess NOx was produced in the MLT by precipitating electrons with energies in the range of ~30 keV to 1 MeV, probably in the January–February timeframe during an extended period of relatively
high geomagnetic activity. The production was most likely not associated with solar protons or higher energy electrons, which would have produced NOx directly in the upper stratosphere or lowermost mesosphere [e.g., Clilverd et al., 2006, 2007, 2009; Funke et al., 2007; López-Puertas et al., 2006; Randall et al., 2009; Semeniuk et al., 2005; Seppälä et al., 2007; Sinnhuber et al., 2014].

Contributing to the extraordinary NOx enhancements were unusual meteorological conditions in the Arctic region in early 2004. It is now understood that the downward transport of NOx produced by EPP (hereafter referred to as EPP-NOx) was enhanced by a dynamically induced increase in the residual circulation. In particular, a prolonged stratospheric warming was followed by reformation of an elevated stratopause near 75–80 km [Manney et al., 2008a], indicative of adiabatic warming from enhanced descent [e.g., Jin et al., 2005; Orsolini et al., 2010; Smith et al., 2009; Winick et al., 2009]. The enhanced descent resulted from a complex dynamical situation that began with a filtering of planetary and gravity wave propagation from the disturbed lower stratosphere, which led to cooling and formation of a very strong vortex in the upper stratosphere and lower mesosphere [Manney et al., 2005; Hauchecorne et al., 2007; Siskind et al., 2007; Thurairajah et al., 2010]. The strong westerly vortex winds then allowed westward propagating gravity waves to propagate up to the mesosphere where, upon breaking, their momentum deposition caused a zonal wind shift and strengthening of the residual circulation. This resulted in enhanced descent in the polar MLT, and thus the remarkable influx of EPP-NOx from the MLT into the stratosphere [Hauchecorne et al., 2007; Siskind et al., 2010; Marsh, 2011].

In spite of the considerable attention that has been paid to the unprecedented EPP effects in 2004, in the 9 years since papers describing the NOx enhancements and accompanying O3 depletions were first published, no one has been able to completely simulate the observed effects. For example, Semeniuk et al. [2005] simulated the 2003–2004 winter using the Canadian Middle Atmosphere Model (CMAM), a general circulation model spanning an altitude range of about 0–95 km. They concluded that electron precipitation was responsible for the descending NOx that was observed in early 2004, since the solar proton events (SPEs) in late 2003 did not produce enough ionization in the upper atmosphere. However, they were unable to simulate the NOx enhancements themselves, concluding that this was due in part to insufficient EPP-NOx production because of their low model top and in part because the climatological meteorology in CMAM did not capture the unusual dynamics that prevailed in 2004 [Jin et al., 2005]. Jackman et al. [2009] used version 3 of the National Center for Atmospheric Research Whole Atmosphere Community Climate Model (WACCM3) to investigate the long-term (>months) effects of extremely large SPEs, including the Halloween storms in 2003. In agreement with Semeniuk et al. [2005], they found that solar protons could not account for the large observed NOx enhancements in March–April of 2004. Auroral electrons were included in the WACCM3 simulations of Jackman et al. [2009], but their effects were not isolated from the effects of solar protons.

In their studies of the 2003–2004 winter a number of authors have used models in which the meteorology in the troposphere and stratosphere is specified from reanalysis data. Using the KArlsruhe Simulation Model of the middle Atmosphere (KASIMA) 3-D mechanistic model, Reddmann et al. [2010] calculated the amount of EPP-NOx entering the stratosphere from July 2002 to March 2004. The meteorology in this model was specified with data from the European Center for Medium-Range Weather Forecasts (ECMWF) operational analyses up to 48 km. The model did not explicitly include particle precipitation, so model NOx above 55 km was overwritten with nighttime NO2 values from the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) when enhancements were observed, which includes 15 January through 5 March 2004. With these constraints, KASIMA reproduced the MIPAS observations of NOx entering the stratosphere later in March 2004 fairly well, and they were able to investigate stratospheric chemistry changes from the EPP-NOx. Because of the stringent model constraints, however, they were unable to test our understanding of, or ability to simulate, the initial production of EPP-NOx or its transport from the MLT. MIPAS experienced an anomaly in late March 2004, so it was not operating in April when the largest stratospheric NOx enhancements were observed [Randall et al., 2005].

Shepherd et al. [2014] used a version of CMAM with continuous incremental nudging to investigate MLT effects of sudden stratospheric warmings (SSWs) in 2003–2004 as well as in 2005–2006 and 2008–2009. In this model version, the meteorology was nudged to closely match reanalysis data below 10 hPa. They were able to capture many of the meteorological characteristics of the MLT, which was not
constrained to meteorology, particularly in 2006 and 2009. However, in 2004 the stratopause reformed at an altitude about 15 km lower than observed. Combined with the fact that EPP was not included in the model except via boundary conditions, the amount of NO\textsubscript{x} descending to the stratosphere in 2004 was underestimated by a factor of 5.

Other studies also used models that required specification of NO\textsubscript{x} at the upper boundary level. For example, Vogel et al. [2008] used the Chemical Lagrangian Model of the Stratosphere, a chemical transport model that is driven by ECMWF winds in the stratosphere and troposphere. NO\textsubscript{x} was prescribed at the upper boundary near 50 km from satellite data, so downward transport from the MLT was not simulated. They found that even with the prescribed NO\textsubscript{x} values, however, their simulated upper stratospheric NO\textsubscript{x} mixing ratios in February–March 2004 were up to nearly 400% too low. They concluded that this was likely caused at least in part by the fact that the satellite data set used for the NO\textsubscript{x} prescription was nighttime NO\textsubscript{2}, which does not include all of the NO\textsubscript{x}.

To summarize, none of the previously published studies that examined particle precipitation effects in the 2003–2004 Arctic winter have simulated the observed NO\textsubscript{x} distributions in a comprehensive and self-consistent manner. For the EPP-NO\textsubscript{x} source they either lacked the appropriate precipitating electron spectra or required prescribed EPP-NO\textsubscript{x} at the upper boundary. In addition, even when the meteorology was specified, the simulations did not satisfactorily reproduce the observations. On the basis of the work just described, it is reasonable to hypothesize that previous simulations did not match the observations primarily because they underestimated EPP-NO\textsubscript{x} production in the MLT and/or they did not adequately capture the dynamical recovery from the prolonged stratospheric warming.

This paper describes a recent simulation of the Arctic 2003–2004 winter with a new version of WACCM that includes specified dynamics in the troposphere and stratosphere. The results show that this model also fails to simulate the very large springtime NO\textsubscript{x} enhancements. Section 2 describes the model, section 3 presents the results, section 4 discusses reasons for model-measurement disagreements, and section 5 summarizes the conclusions.

### 2. Specified Dynamics WACCM

Comprehensive simulations of EPP-induced coupling of different atmospheric regions require calculating EPP-induced ionization, production of reactive chemical constituents, 3-D transport and subsequent chemical reactions, radiative transfer, dynamical forcing, and coupled atmospheric responses. In other words, such simulations require a global, chemistry-climate model that spans an altitude range from the surface through the lower thermosphere. WACCM includes all of these capabilities and was therefore used for the work in this paper.

WACCM is a chemistry-climate general circulation model with a vertical domain extending from the surface to $5.9 \times 10^{-6}$ hPa (~140 km); it thus includes the MLT region of primary interest as the EEP source region in the proposed work. The standard horizontal resolution is 1.9° latitude × 2.5° longitude. The chemistry module in WACCM is interactive with the dynamics through transport and exothermic heating and includes chemistry associated with ion species (O\textsuperscript{+}, NO\textsuperscript{+}, O\textsubscript{2}\textsuperscript{+}, N\textsubscript{2}\textsuperscript{+}, and N\textsuperscript{+}) [Kinnison et al., 2007; Marsh et al., 2007]. Therefore, the neutral and ion species are self-consistently resolved; this is a unique feature among whole atmosphere models, as required for the study described here. WACCM is configured to run as the atmospheric component of the NCAR Community Earth System Model (CESM1), which is a coupled model consisting of atmosphere, ocean, land surface, sea and land ice, and carbon cycle components for simulating past, present, and future climates [Hurrell et al., 2013]. These components are linked through a coupler that exchanges fluxes and state information between them. In the simulation presented here, all components except the atmosphere are prescribed from observations. WACCM simulations of the atmospheric response to solar cycle variations are described in several papers, including those by Marsh et al. [2007], Matthes et al. [2013], Calvo and Marsh [2011], and Chiodo et al. [2012, 2014]. Details of recent centennial-scale coupled simulations using the current version of WACCM and an overview of the model climate can be found in Marsh et al. [2013].

WACCM can operate as a free-running climate model, but it can also incorporate meteorological reanalyses in the troposphere and stratosphere [e.g., Marsh, 2011; Brakebusch et al., 2013]. This is achieved by relaxing the
horizontal winds and temperatures to reanalysis fields and is referred to as Specified Dynamics WACCM (SD-WACCM). The results shown in this work are based on the reference (REF) Chemistry Climate Model Initiative (CCMI) REF-C1SD (SD-WACCM) simulation. [Eyring et al., 2013]. Meteorological fields are taken from the NASA Global Modeling and Assimilation Office (GMAO) Modern-Era Retrospective Analysis for Research and Applications (MERRA) [Rienecker et al., 2011]. The reanalysis fields are applied below 50 km, with a transition from 50–60 km, so that above 60 km the model is free-running. The model is “nudged” to reanalysis fields with a 50 h relaxation time constant. Therefore, in the MLT the chemical-dynamical interactions are consistent, but the forcing from below makes it possible to simulate the effects of particular dynamical events, such as occurred in the Arctic winter of 2003–2004.

The CCMI REF-C1SD simulation is forced at the surface with observed, time-varying greenhouse gases (CH4, N2O, and CO2) and organic halogens (CFCs, HCFCs, and halons). A representation of solar variability and sulfate volcanic loading is also included. The impact of both the quasi-biennial oscillation and El Nino–Southern Oscillation on the mean middle atmosphere circulation is implicit in the reanalysis product. The gravity wave parameterization used in WACCM and its impact on the calculated diffusivity is discussed in Garcia et al. [2014]. For the CCMI REF-C1SD simulation the Prandtl number, which is used to calculate the diffusivity due to gravity waves, is set equal to 4.

WACCM has been used extensively to study the influence of EPP on the composition and dynamics of the atmosphere. Recent publications documenting effects of solar protons in WACCM include those by Damiani et al. [2012], Jackman et al. [2008, 2009, 2011], and Funke et al. [2011]. Auroral ionization is calculated using the empirical oval of Roble and Ridley [1987], which depends on a specified hemispheric power or geomagnetic Kp index. Holt et al. [2013] used WACCM to study the effects of SSWs on the atmospheric response to auroral electron precipitation. WACCM has the capability to incorporate higher energy electron precipitation in addition to auroral electron precipitation, which would result in ionization primarily in the polar mesosphere between about 50 and 90 km. In the model/observation intercomparison study described by Funke et al. [2011], medium energy electron ionization was specified using the Atmospheric Ionization Module Osnabrück (AIMOS) model. The precipitating electron maps used in AIMOS are based on Polar Operational Environmental Satellites Medium Energy Proton and Electron Detector (MEPED) data. However, the MEPED electron data suffer from contamination and inadequate sampling of the loss cone [Andersson et al., 2012; Rodger et al., 2010; Yando et al., 2011]. Funke et al. [2011] found that WACCM simulations of the Halloween storms using AIMOS output from uncorrected MEPED data had NOx production that was unreasonably high, consistent with proton contamination. Because of the MEPED errors, only auroral electron precipitation is included in the work described here; including higher energy electron precipitation using the parameterization of Fang et al. [2008, 2010], and corrected MEPED data is work in progress.

The SD-WACCM data were output at the times and locations of many different satellite measurements for direct comparison to observations. The instruments and data versions include the Halogen Occultation Experiment (HALOE), version 19; Stratospheric Aerosol and Gas Experiment (SAGE) II, version 6.2 and SAGE III, version 3.0; Polar Ozone and Aerosol Measurement (POAM) III, version 6.0; Atmospheric Chemistry Experiment (ACE), version 3; Sounding of the Atmosphere using Broadband Emission Radiometry (SABER), version 2.0; and MIPAS, version V3c. For these comparisons, WACCM data were interpolated to a common altitude scale using geopotential height. An averaging kernel was applied to SD-WACCM for comparison to MIPAS observations, as in Funke et al. [2011].

3. Results

Figure 1 compares NOx from HALOE [Gordley et al., 1996] and NO2 from SAGE II [Cunnold et al., 1991], SAGE III [Rault, 2004], and POAM III [Randall et al., 2002] to SD-WACCM output at the times and locations of these solar occultation measurements. The measurements and model output correspond to the northern hemisphere (NH) latitudes sampled by each instrument in 2004, as shown in the figure, and an altitude of 40 km. Since measurement latitudes change rapidly for midinclination solar occultation instruments such as HALOE and SAGE II, all individual NH measurements from these instruments are shown. Gaps in the data result from times when the instruments did not sample in the NH. Since POAM III and SAGE III measurement latitudes change more slowly, 7 day running means are shown. The satellite data in Figures 1a–1c are the
same as the 2004 data in Figure 1 of Randall et al. [2005], where it was shown that the large NOx and NO2 mixing ratios in April–June were unprecedented and up to 4 times larger than ever before observed at these times and locations. The model and measurements are in excellent agreement prior to March, giving confidence that SD-WACCM accurately represents the background atmosphere. However, once the EPP-NOx enhancements at 40 km become evident in the satellite data (March for POAM III and SAGE III; April for HALOE and SAGE II), the agreement breaks down; SD-WACCM completely misses the enhancements. Although not shown, this SD-WACCM simulation also fails to simulate the consequent O3 reductions described by Randall et al. [2005].

Figure 2 compares Atmospheric Chemistry Experiment (ACE) Fourier Transform Spectrometer (FTS) NOx profiles in the NH in 2004 to SD-WACCM output at the ACE-FTS times and locations in 2004. Included here are all days for which the ACE-FTS measurement latitudes were poleward of 50°N, as shown in the figure by the white symbols. ACE-FTS has an advantage over the instruments in Figure 1, in that unlike POAM III and SAGE III, it measures NOx (not just NO2); and unlike HALOE and SAGE II, it samples high latitudes for long, continuous periods of time. The EPP-NOx enhancements in the ACE data were first described by Rinsland et al. [2005]; they are seen here as the descending “tongue” that begins on 21 February (the first day of valid ACE data) and is still apparent at the end of March when the ACE sampling moved to lower latitudes.
SD-WACCM certainly shows NOx descending from the mesosphere into the stratosphere, but consistent with the other solar occultation comparisons, it drastically underestimates the amounts. In the main part of the tongue of descending EPP-NOx, from about 45–60 km, SD-WACCM mixing ratios are generally less than 20% of the ACE values. SD-WACCM also underestimates NOx at the higher altitudes for most of the time period shown here.

Comparisons between SD-WACCM and MIPAS NOx [e.g., Funke et al., 2014a] are shown in Figure 3. Since MIPAS measures infrared emission, it has global coverage on a daily basis, unlike the solar occultation instruments. The plots in Figure 3 show 3 day running averages of all profiles acquired poleward of 70°N, from 1 November 2003 until 26 March 2004. As mentioned above, MIPAS temporarily stopped operating after 26 March, so the MIPAS data set does not include the largest stratospheric enhancements shown in Figure 1. Nevertheless, the near-global coverage gives a continuous picture of NOx variability throughout the polar winter. Both MIPAS and SD-WACCM show large amounts of descending NOx in November–December and again in January through March. Because the only source of NOx in the polar winter mesosphere is EPP, these tongues of descending NOx are unequivocally attributed to EPP-NOx. The tongue in November–December of 2003 has been attributed to EPP-NOx produced during the SPEs in the late October–December time period [Funke et al., 2011; Jackman et al., 2005; López-Puertas et al., 2005]. A sudden stratospheric warming (SSW) in early January [Manney et al., 2005, 2008a] interrupted the descent of EPP-NOx from these SPEs [López-Puertas et al., 2005], resulting in the infusion of NOx-poor air from lower latitudes. Upon recovery from the SSW, NOx once again descended from the MLT, producing the second tongue seen in Figure 3, as well as the enhancements seen in the ACE, POAM, HALOE, and SAGE data in Figures 1 and 2.

In a study of how well WACCM and other models simulated the initial effects from the Halloween storms, Funke et al. [2011] described comparisons between MIPAS and a previous SD-WACCM simulation, for November 2003. Consistent with that work, Figure 3 shows that in November and December of 2003, SD-WACCM underestimates EPP-NOx in the mesosphere. Figure 3 also shows that SD-WACCM overestimates EPP-NOx in the uppermost stratosphere at the beginning of November, and this overestimate appears to descend in time. Funke et al. [2011] concluded that the differences were likely related to errors in simulated ionization rate profiles but could not rule out errors in the background atmosphere or transport scheme. For the early 2004 time period of primary interest in this paper, the region over which SD-WACCM underestimates NOx relative to MIPAS is larger, extending down to altitudes of 30 km by March. This is consistent with the solar occultation comparisons shown above.

Figure 4 explores the early 2004 differences in more detail. As shown in Figure 3, differences between MIPAS and SD-WACCM were apparent in late 2003, before the SSW and subsequent recovery. In order to focus on the January–March time period, the average NOx profiles on 1 January 2004 from SD-WACCM and MIPAS were subtracted from the respective data; results were similar even if a week-long average centered on 1 January was subtracted. Three-day running averages of these differences, $\Delta$NOx, are presented in Figure 4 to show the change in NOx, relative to 1 January 2004. This figure shows that even after accounting for the differences between SD-WACCM and MIPAS that were already present on 1 January, which as mentioned...
above may have resulted from errors in the Halloween storm ionization rates, SD-WACCM still shows an underestimate in the amount of EPP-NOx that descended from the MLT in January–March. In the central part of the tongue of descending NOx, SD-WACCM underestimates MIPAS by more than an order of magnitude throughout the entire time period.

4. Discussion

Figures 1 through 4 clearly show that SD-WACCM was unable to capture the unprecedented enhancements in stratospheric NOx in the Arctic spring of 2004. This was true even though the model was nudged to meteorological reanalysis data in the troposphere and stratosphere and was therefore constrained to realistically represent the lower atmosphere. The two most likely reasons that SD-WACCM would underestimate the amount of EPP-NOx descending to the stratosphere are (1) not enough NOx production by the precipitating particles and (2) inadequate downward transport of EPP-NOx from the source region to the stratosphere. Both of these possibilities are considered here.

NOx production during SPEs often occurs in the upper stratosphere and mesosphere by high energy solar protons. In the Arctic winter of 2003–2004 SPEs occurred on 26 and 28–29 October; 2–3, 4–5, and 21 November; and 2 December. The first SPE in 2004 did not occur until 11 April, and it was very weak. To illustrate the simulated effects of the 2003 SPEs, Figure 5 shows WACCM NOx in the MLT averaged over the MIPAS measurement locations poleward of 70°N for the October 2003 through March 2004 time frame. This therefore represents an extension of Figure 3b to higher altitudes, even though the MIPAS retrievals themselves do not extend above 70 km. For reference, the Ap index is shown above the NOx contour plot, and the SPEs are denoted in the contour plot with dashed, gray vertical lines. The late October SPEs clearly produced substantial NOx in the mesosphere below ~85 km, and as expected, there is a rough correlation between the Ap index and NOx variations in the thermosphere. However, variations above 85 km from any of the SPEs and associated geomagnetic disturbances were no larger than variations that occurred at many other times in the absence of any SPE. Figure 5 also shows no discernible connection between the NOx enhancements during the SPEs and the second tongue of descending NOx that led to the March–April enhancements. This leads to the conclusion that neither solar protons nor auroral electron precipitation during the SPEs was responsible for the second tongue of descending NOx in Figure 3. This SD-WACCM simulation did not include ionization by precipitating electrons with energies greater than 30 keV; precipitation by these high-energy electrons will hereinafter be referred to as HEP (high-energy electron precipitation). That the NOx variations above 85 km during the SPEs clearly produced substantial NOx in the mesosphere below ~85 km, and as expected, there is a rough correlation between the Ap index and NOx variations in the thermosphere. However, variations above 85 km from any of the SPEs and associated geomagnetic disturbances were no larger than variations that occurred at many other times in the absence of any SPE. Figure 5 also shows no discernible connection between the NOx enhancements during the SPEs and the second tongue of descending NOx that led to the March–April enhancements. This leads to the conclusion that neither solar protons nor auroral electron precipitation during the SPEs was responsible for the second tongue of descending NOx in Figure 3. This SD-WACCM simulation did not include ionization by precipitating electrons with energies greater than 30 keV; precipitation by these high-energy electrons will hereinafter be referred to as HEP (high-energy electron precipitation). That the NOx variations above 85 km during the SPEs were both transient and relatively small suggests that even had SD-WACCM included HEP, the total effects of EPP during the SPEs themselves would not have persisted long enough to be responsible for the second tongue of descending NOx in Figure 3.

As noted in the introduction, and consistent with the explanation of Figure 5 just given, it has previously been considered likely that the NOx eventually observed in the 2004 Arctic springtime stratosphere was produced...
by electron precipitation some time after the 2003 SPEs. In order to exhaustively evaluate NO\textsubscript{x} production by precipitating electrons, one would ideally compare model simulations of NO\textsubscript{x} to observations in the MLT throughout the time period of interest. Unfortunately, the only satellite instruments that measured NO\textsubscript{x} profiles in the MLT during the 2003–2004 winter were ACE and HALOE. The ACE data are only valid after 20 February, however, so they cannot be used to examine EPP-NO\textsubscript{x} production earlier in the year. As shown in Figure 1, HALOE measurement latitudes never reached higher than 56°N before March. Thus, neither data set is ideal for verifying EPP-NO\textsubscript{x} production in SD-WACCM. Nevertheless, Figure 6 compares NO\textsubscript{x} profiles from SD-WACCM to HALOE profiles at the HALOE locations and times. All measurements poleward of 40°N are included, as long as the stated error was less than 50%; they are averaged over the discrete time periods during which this latitude sampling occurred. At the latitudes sampled by HALOE, SD-WACCM actually overestimates NO\textsubscript{x} from ~90 km to 110 km or higher. NO\textsubscript{x} at these latitudes and altitudes is produced by both auroral electron precipitation and solar soft X-rays [Barth et al., 2003]; but under the early 2004 conditions of relatively high geomagnetic activity and low solar input, auroral electron precipitation was likely the main source. Thus, an underestimate of auroral electron precipitation in SD-WACCM is an unlikely explanation for underestimating the amount of NO\textsubscript{x} that descended into the stratosphere.

Sheese et al. [2013] compared retrievals of MLT NO from the Odin Optical Spectrograph and Infrared Imaging System (OSIRIS) and Sub-Millimeter Radiometer (SMR) to SD-WACCM simulations for years 2003–2010. Prior to 2007, OSIRIS sampled the MLT only 1 day out of 10, and SMR only 1 day out of 30, and polar comparisons could only be made at southern latitudes from April through August. Therefore, they were unable to compare their observations to SD-WACCM simulations of the 2003–2004 Arctic winter. In addition, they showed large disagreements between the OSIRIS and SMR NO retrievals. Nevertheless, their comparisons of climatological results over Antarctica showed that relative to both instruments, SD-WACCM generally overestimated NO

![Figure 5](image-url) **Figure 5.** (top) Three-day running averages of the Ap index from 1 October 2003 through March 2004. The red rectangles denote time periods of profiles plotted in Figure 6. (bottom) Contour plot of SD-WACCM NO\textsubscript{x} with CO contours at 5, 10, 20, 30, and 40 ppbv superimposed in white; the black NO\textsubscript{x} contours above the highest level in the color bar increase with altitude by factors of 2. For both NO\textsubscript{x} and CO the plot shows 3 day running averages over the MIPAS measurement locations poleward of 70°N; the white stripes indicate missing MIPAS data. The gray, dashed lines indicate SPEs.

![Figure 6](image-url) **Figure 6.** (a) Average NO\textsubscript{x} profiles from HALOE (solid) and SD-WACCM at the HALOE times and locations (dashed) for the time periods shown; latitudes range from 40°N to 56°N. Missing data correspond to measurements for which the errors are greater than 50%. One-sigma uncertainties in the mean mixing ratios at each altitude (not plotted) are on the order of the widths of the profiles themselves. (b) Ratios of the profiles in Figure 6a.
concentrations in midwinter near the peak of the NO profile (~102–105 km in SD-WACCM, ~95–98 km in OSIRIS, and ~90–100 km in SMR). These results thus support the conclusions from the HALOE comparisons that an underestimate of NO production by auroral electron precipitation probably cannot explain the lack of NOx enhancements in the Arctic 2004 springtime stratosphere.

In contrast to the comparisons above 90 km, Figure 6 shows that relative to HALOE, SD-WACCM significantly underestimates NOx values from 60–85 km in all three time periods. The 25–31 January time period followed a peak in geomagnetic activity (see the top plot of Figure 5). Since this suggests the possibility that HEP could have been significant, the NOx underestimate in this time period is consistent with the lack of HEP in SD-WACCM; however, insufficient transport of EPP-NOx cannot be ruled out as an explanation. There was less geomagnetic activity immediately preceding and during the 11–17 January time period; but as shown from the white CO contours in Figure 5, this time period was dynamically active according to SD-WACCM. That is, CO is a tracer of motion in the polar winter mesosphere [Allen et al., 1999, 2000; Jin et al., 2005], so the sudden drop in the 10 ppmv and 20 ppmv CO contours near 13–14 January, which is accompanied by an increase in simulated NOx, indicates increased descent and/or a sudden decrease in horizontal mixing with lower latitude (lower-mixing-ratio) air after the SSW in early January. Although not shown, the vertical component of the residual circulation (wbar*) in SD-WACCM also indicates an increase in descent in mid-January. Inaccurate simulation of these transport effects could therefore be a likely explanation for the disagreement between SD-WACCM and HALOE at this time. Figure 5 shows that during the 9–17 February time period there was a peak in geomagnetic activity (top panel) as well as increased dynamical activity in SD-WACCM below 85 km, as indicated by the dip in the 5 ppmv and 10 ppmv CO contours. Thus errors in simulating both HEP and transport in the mesosphere would be important in this time period. Without more observations of NOx in the polar winter MLT, more definitive conclusions are precluded.

To examine whether errors in descent rates are responsible for the SD-WACCM underestimate of springtime NOx, Figure 7 compares SABER temperatures [Remsberg et al., 2008] to SD-WACCM simulations of temperature at the SABER times and locations poleward of 70°N. Although temperature is not a tracer of vertical motion, as previously mentioned the 2003–2004 Arctic winter was one in which a prolonged SSW was followed by enhanced descent at typically mesospheric altitudes; the adiabatic warming caused by this led to reformation of the stratopause near 75–80 km in middle to late January [Manney et al., 2008a]. Thus, the presence of a stratopause at these altitudes is indicative of enhanced descent that would effectively transport EPP-NOx from the MLT down to stratospheric altitudes and can be used as a diagnostic of vertical transport in SD-WACCM. Figure 7 shows that the stratopause in SD-WACCM is significantly lower in altitude than in the SABER observations in late January and early February. The difference plot in Figure 7c has a clear
signature of the displaced stratopause: From mid-January to mid-February SD-WACCM temperatures are 5 K to 35 K lower than SABER temperatures between 70 km and 80 km, near the SABER stratopause; they are 5 K to 35 K higher between 50 km and 60 km, near the SD-WACCM stratopause. As shown in Figure 7d, the stratopause in SD-WACCM remains lower in altitude than in SABER until ~20 February, after which the agreement improves significantly.

Manney et al. [2008b] compared temperature profiles from several different data sets, including SABER and the Goddard Earth Observing System version 5 (GEOS-5) at high northern latitudes in early 2006. This time period was similar to early 2004 in that there was a prolonged SSW followed by an elevated stratopause (ES). Since the MERRA data set to which SD-WACCM is nudged is based on the GEOS-5 data, the comparisons between SABER and GEOS-5 are relevant to the comparisons in Figure 7 between SABER and SD-WACCM. Manney et al. [2008b] showed that after the early 2006 SSW, the GEOS-5 stratopause reformed at altitudes lower than indicated by the SABER data and that the GEOS-5 stratopause was warmer than the SABER stratopause. In contrast, GEOS-5 temperatures at altitudes below the reformed stratopause were lower than measured by SABER. The result was that while the stratopause remained elevated, GEOS-5 was warmer than SABER near 55–70 km, but cooler than SABER around 40–55 km. Below 40 km the two data sets were in better agreement; the GEOS-5 data did not extend above ~70 km. The comparisons in Figure 7 are mostly consistent with the results of Manney et al. [2008b] in that the reformed stratopause is warmer and lower in altitude in SD-WACCM than in SABER. Since SD-WACCM is nudged to MERRA below 50 km, with a transition region up to 60 km, errors in the MERRA data probably contribute to the inability of SD-WACCM to reproduce the atmosphere’s recovery from the SSW at higher altitudes. Note that nudging SD-WACCM to the European Center for Medium Range Weather Forecasts (ECMWF) instead of to MERRA would likely yield similar results, since ECMWF also gives a stratopause that is too low in altitude during an ES event [Manney et al., 2008b]. That SD-WACCM is cooler than SABER above 70 km throughout January and most of February indicates that the descent in SD-WACCM above 80 km was not enhanced as much as in the actual atmosphere, which would result in less descent of EPP-NOx from the MLT into the stratosphere.

Figure 8. Average vertical component of the residual circulation (wbar*) poleward of 70°N for (a) January and (b) February of 2004, as inferred from MIPAS (black) and SD-WACCM (gray, dash-dotted). The altitude is given as log pressure altitude.

Descent rates between 70 km, the top altitude for MIPAS, and 50 km, below which SD-WACCM is nudged strongly to meteorological reanalysis data, were approximated by the vertical component of the residual circulation, wbar*. Figure 8 compares the average wbar* profile derived from MIPAS and SD-WACCM data during the months of January and February of 2004, for latitudes poleward of 70°N. For MIPAS, the residual circulation was calculated as in Holt et al. [2012] and Funke et al. [2014b] from MIPAS temperatures and diabatic heating rates. The heating rate calculations rely on trace gas (ozone and water vapor) distributions from MIPAS, which are input to the radiative transfer model MODTRAN [Berk et al., 2006] along with polar winter climatological fields for other trace gases. For SD-WACCM, wbar* was calculated using equation (3.64b) in Brasseur and Solomon [2005], which is based on Andrews and McIntyre [1976]. The MIPAS and SD-WACCM wbar* profiles in Figure 8 are in good agreement throughout most of the altitude range from 50 to 70 km in both January and February. In both data sets, the magnitude of wbar* increases with increasing altitude and also increases with time from January to February. In January the descent rates decrease from ~0.9 km/d at 70 km to ~0 by 50 km; in February they decrease from ~1.5 km/d just below 70 km to ~0.2 km/d at 50 km. The overall conclusion from Figure 8 is that below 70 km the SD-WACCM descent rates are reasonable; this supports the suggestion that the EPP-NOx underestimate in the SD-WACCM simulation of the 2004 Arctic spring is due to too little NOx production from HEP and/or insufficient transport from the MLT.
5. Conclusions

As described in previous work, unprecedented enhancements in stratospheric NOx were observed in the Arctic spring of 2004; these enhancements have been attributed to descent of air with excess NOx produced by EPP [Randall et al., 2005; Natarajan et al., 2004; Rinsland et al., 2005]. Output from SD-WACCM, achemistry climate model nudged to the MERRA meteorological data in the troposphere and stratosphere, was analyzed to assess our understanding of these enhancements. On the basis of comparisons to measurements from many satellite instruments, it was shown that SD-WACCM drastically underestimates the amount of EPP-NOx that descended to the stratosphere in the Arctic winter/spring of 2004. Observations show that due to EPP-NOx descent that probably began in January, NOx at 40 km in early April was enhanced by a factor of 4. SD-WACCM did show EPP-NOx descending from the MLT early in 2004, but far too little was brought down, so the April enhancement was completely absent.

The EPP-NOx underestimate in SD-WACCM is attributed to too little NOx production by HEP and/or insufficient transport from the MLT. In agreement with previous publications, it was shown that neither proton nor electron precipitation during the late 2003 SPEs was responsible for the springtime NOx enhancements; this rules out inaccurate simulation of the SPEs themselves as explanations for the underestimate. Comparisons with HALOE, the only satellite instrument that measured NO near the polar region in the MLT early in 2004, do not indicate an underestimate in SD-WACCM of NO production by precipitating auroral electrons. The HALOE comparisons are, however, consistent with SD-WACCM underestimating EPP-NOx production in the mesosphere by higher energy electron precipitation. Codrescu et al. [1997] showed using the Thermosphere Ionosphere Mesosphere Electrodynamic General Circulation Model that precipitating electrons with energies larger than 30 keV resulted in significant production of NO between 70 and 80 km. The SD-WACCM simulation analyzed here did not include precipitation of these electrons, which almost certainly contributed to the underestimate of simulated EPP-NOx descending to the stratosphere.

Comparisons with SABER temperature measurements show that SD-WACCM did not accurately simulate the atmosphere’s recovery from the prolonged early winter SSW. The elevated stratopause that formed in middle to late January was located near 55–65 km in SD-WACCM, whereas it was located near 70–75 km in SABER, indicating that SD-WACCM underestimated descent rates in the mesosphere at this time. Estimates of descent rates based on calculations of the vertical component of the residual circulation show that SD-WACCM and MIPAS are in relatively good agreement from 50–70 km in January and February of 2004. It has been suggested that WACCM might underestimate descent in the polar upper mesosphere either because the poleward branch of the residual circulation is too low in altitude and/or because the eddy diffusion is not strong enough [Holt et al., 2013; Smith et al., 2011]. With regard to the results shown above, this suggestion could imply that during and after SSWs the dissipation of gravity wave and associated momentum deposition in the SD-WACCM simulation used here are incorrect. Most likely error sources include triggering of the gravity waves, amplitude of the gravity wave source spectrum, spectral shape of the launched waves, or inaccurate filtering caused by errors in the background winds in the mesosphere.

A comprehensive description of Sun-Earth connections requires quantifying the atmospheric processes that indirectly amplify the effects of solar and magnetospheric input. This includes nonlinear feedback between chemical, radiative, and dynamical processes that couple different regions of the atmosphere. The atmospheric response to electron precipitation is a key component of Sun-Earth connections [e.g., Andersson et al., 2014] and provides a natural means of probing the underlying physics. The work described here shows the difficulty that even a state-of-the-art, sophisticated, and highly complex coupled chemistry climate model has in simulating the impacts of EPP. Nevertheless, including EPP in future climate simulations is critical for accurate calculation of solar and magnetospheric effects on the atmosphere and potentially climate. The work here suggests that including the full energy range of precipitating electrons is likely necessary to accurately explain and predict the effects of EPP on the atmosphere. It also shows that even with prescribed meteorology in the troposphere and stratosphere, simulated transport at higher altitudes can have significant errors, pointing to the need for measurements of temperatures and winds in the MLT. Finally, it points to the need for measurements of NOx throughout the polar night from the stratosphere to the lower thermosphere, in order to adequately assess our understanding of EPP-induced coupling between the MLT and other regions of the atmosphere.
Acknowledgments

Data used in this paper are available upon request from C.E. Randall. This work was supported by NASA grants NNX14AH54G, NNX10AQ54G, and NNX09AO4G and by NSF grant AGS 1135432. Support for the ACE Mission is primarily from the Canadian Space Agency. We gratefully acknowledge comments from James M. Russell, III, and the careful and diligent efforts of the science and data processing teams who produced the high-quality HALOE and SABER data sets used in this analysis.

The National Center for Atmospheric Research (NCAR) is sponsored by the U.S. National Science Foundation. WACCM is a component of the Community Earth System Model (CESM), which is supported by the National Science Foundation (NSF) and the Office of Science of the U.S. Department of Energy. Computing resources were provided by NCAR’s Climate Simulation Laboratory, sponsored by NSF and other agencies. This research was enabled by the computational and storage resources of NCAR’s Computational and Information Systems Laboratory.

Alan Rodger thanks Manuel López-Puertas and another reviewer for their assistance in evaluating this paper.

References


