Chapter 10: Polar Processes

Chapter lead authors

Michelle L. Santee	NASA Jet Propulsion Laboratory, California Institute of Technology	USA
Alyn Lambert	NASA Jet Propulsion Laboratory, California Institute of Technology	USA
Gloria L. Manney	(1) NorthWest Research Associates (2) Department of Physics, New Mexico Institute of Mining and Technology	USA

Co-authors

Zachary D. Lawrence	(1) Cooperative Institute for Research in Environmental Sciences (CIRES), Univ. of Colorado (2) NOAA Physical Sciences Laboratory (3) NorthWest Research Associates	USA
Simon Chabrillat	Royal Belgian Institute for Space Aeronomy	Belgium
Lars Hoffmann	Forschungszentrum Jülich GmbH	Germany
Sean P. Palmer	Department of Physics, New Mexico Institute of Mining and Technology	USA
Ken Minschwaner	Department of Physics, New Mexico Institute of Mining and Technology	USA

Abstract. This chapter focuses on microphysical and chemical processes in the winter polar lower stratosphere, such as polar stratospheric cloud (PSC) formation; denitrification and dehydration; heterogeneous chlorine activation and deactivation; and chemical ozone loss. These are "threshold" phenomena that depend critically on meteorological conditions. A range of diagnostics is examined to quantify differences between reanalyses and their impact on polar processing studies, including minimum lower stratospheric temperatures; area and volume of stratospheric air cold enough to support PSC formation; maximum latitudinal gradients in potential vorticity (a measure of the strength of the winter polar vortex); area of the vortex exposed to sunlight each day; vortex break-up dates; and polar cap average diabatic heating rates. For such diagnostics, the degree of agreement between reanalyses is an important direct indicator of the systems' inherent uncertainties, and comparisons to independent measurements are frequently not feasible. For other diagnostics, however, comparisons with atmospheric observations are very valuable. The representation of small-scale temperature and horizontal wind fluctuations and the fidelity of Lagrangian trajectory calculations are evaluated using observations obtained during long-duration superpressure balloon flights launched from Antarctica. Comparisons with satellite measurements of various trace gases and PSCs are made to assess the thermodynamic consistency between reanalysis temperatures and theoretical PSC equilibrium curves. Finally, to explore how the spatially and temporally varying differences between reanalyses interact to affect the conclusions of typical polar processing studies, simulated fields of nitric acid, water vapour, several chlorine species, nitrous oxide, and ozone from a chemistry-transport model driven by the different reanalyses for specific Arctic and Antarctic winters are compared to satellite measurements.

Contents

10.1 Introduction

One of the main research themes in atmospheric science over the past three decades has been the investigation of the chemical and dynamical processes involved in stratospheric ozone depletion, the most severe manifestation of which is the Antarctic ozone hole. In general, the processes controlling polar stratospheric ozone are now well understood (*e.g.*, *WMO*, 2018). In the very cold conditions that prevail inside the lower stratospheric winter polar vortices, water vapour (H_2O) and nitric acid (HNO_3) condense to form polar stratospheric clouds (PSCs). PSC particles and cold sulphate aerosols provide surfaces on which heterogeneous reactions can take place very rapidly, converting chlorine from relatively benign reservoir species such as hydrogen chloride (HCl) and chlorine nitrate ($ClONO_2$) into highly reactive ozone-destroying forms such as chlorine monoxide (ClO). Moreover, sequestration in PSCs substantially reduces gas-phase HNO_3 concentrations, and if solid HNO_3-containing PSC particles grow large enough to undergo appreciable gravitational sedimentation, then HNO_3 can be irreversibly removed from the stratosphere in a process known as denitrification. Similarly, sedimentation of water ice particles leads to dehydration. Severe denitrification and dehydration routinely occur in the cold, isolated Antarctic vortex. Compared to the Antarctic, the Arctic vortex is usually substantially warmer, more dynamically disturbed, smaller, and shorter lived, and thus in a typical year it experiences little or no denitrification or dehydration. Chlorine activation is also typically less intense, extensive, and prolonged in the Arctic than in the Antarctic. Consequently, although the same fundamental processes are at work in the lower stratosphere in both polar regions, in most years chlorine-catalyzed ozone loss is considerably weaker in the Arctic than in the Antarctic.

Lower stratospheric polar processes and chemical ozone loss are "threshold" phenomena that depend critically on stratospheric temperatures and other meteorological and dynamical factors (*e.g.*, winter polar vortex characteristics, breakup dates, *etc.*). Several studies over the years have explored the temperature sensitivity of these processes; for example, *Wegner et al.* (2012) showed that heterogeneous reaction rates on liquid aerosols are doubled for every 1 K in cooling and increase tenfold over a 2-K range around 192 K, and *Solomon et al.* (2015, see also references therein) showed that a 2-K perturbation in temperature applied to heterogeneous chemical reactivities and PSC surface area in a specified-dynamics chemistry climate model induces a change in simulated Arctic column ozone loss of ~ 40 DU.

As in many other Earth system science specialties, atmospheric polar processing studies often rely heavily on global meteorological data sets. Thus it is essential to understand the accuracy and reliability of reanalysis fields in a polar processing context. Differences between reanalyses are likely to have the largest impact on such studies when conditions are marginal, *i.e.*, in the Arctic (in most years) and in the autumn and spring in the Antarctic. As noted, for example, by *Hoffmann et al.* (2017a) and *Lambert and Santee* (2018) and discussed further below, in addition to discrepancies in physical parameters (*e.g.*, temperature, winds) between the various reanalyses, differences in their temporal and/or spatial resolution may also play a role in detailed quantitative studies.

Given the importance of stratospheric temperatures, transport, and mixing for ozone chemistry, a number of studies over the last twenty years have assessed the representativeness of meteorological analyses and reanalyses. We briefly summarize here several studies that carried out comparisons of two or more analyses/reanalyses specifically in a stratospheric polar processing framework. In one of the earliest such studies, *Manney et al.* (1996) examined temperatures, geopotential heights, winds, and potential vorticity (PV) calculated from stratospheric analyses provided by the (then) UK Meteorological Office (UKMO) and the US National Meteorological Center (NMC) in both hemispheres during dynamically active periods, when substantial discrepancies between analyses were likely to be seen. Although both analyses captured the qualitative features and evolution of the large-scale winter stratospheric circulation, differences in their temperatures and polar vortex characteristics implied significant effects on quantitative process studies, especially for the Southern Hemisphere. *Knudsen* (1996) also found substantial biases between observed lower stratospheric temperatures and analyses from UKMO and the European Centre for Medium-range Weather Forecasts (ECMWF). *Knudsen et al.* (2001) assessed the accuracy of analyzed winds from ECMWF, UKMO, and the US National Centers for Environmental Prediction (NCEP) Climate Prediction Center (CPC) by comparing calculated air parcel trajectories based on those analyses with long-duration balloon flights in the Arctic stratospheric vortex. Similarly, *Knudsen et al.* (2002) used independent meteorological measurements from long-duration balloon flights in the Arctic stratospheric vortex to quantify errors in five sets of analyzed temperatures: ECMWF, Met Office (formerly UKMO), the Goddard Space Flight Center Data Assimilation Office (DAO), NCEP/CPC, and NCEP/National Center for Atmospheric Research reanalysis (NCEP-NCAR R1); although some of the analyses showed larger scatter around the balloon values than others, occasional large differences occurred in all of them, particularly during a major sudden stratospheric warming. *Manney et al.* (2003b) compared commonly used meteorological analyses (Met Office, NCEP/CPC, NCEP-NCAR R1, ECMWF, DAO) during two cold Arctic winters, examining not only temperatures (average and minimum values, number of cold days, *etc.*) but also temperature histories along trajectories to assess simulated PSC lifetimes and the overall potential for chlorine activation. They found that discrepancies between analyses arise from differences in both the magnitude and the morphology of wind and temperature fields, such that dissimilarities in dynamical conditions in comparably cold winters may strongly influence the degree of agreement between meteorological data sets.

Following on from that study, *Manney et al.* (2005) investigated an extensive set of diagnostics related to lower stratospheric chemistry, transport, and mixing during the 2002 Antarctic winter, when unusual dynamical activity may have exacerbated the disagreement between meteorological data sets. Comparing four operational products (Met Office, ECMWF, NCEP/CPC, and the NASA Global Modeling and Assimilation Office (GMAO) Goddard Earth Observing System (GEOS-4)), as well as the 40-year reanalysis from ECMWF (ERA-40), NCEP-NCAR R1, and a second NCEP/Department of Energy reanalysis (NCEP-DOE R2), they again found considerable differences; such large disparities undermine confidence in the results from scientific studies based on any of those analyses/reanalyses. In particular, NCEP-NCAR R1, NCEP-DOE R2, and ERA-40 were shown to suffer from substantial deficiencies in their depiction of the magnitude, structure, or evolution of temperatures and/or winds that rendered them unsuitable for detailed studies of lower stratospheric polar processing. *Labitzke and Kunze* (2005) compared stratospheric temperatures over the Arctic from NCEP-NCAR R1 and ERA-40 with an independent data set (historical daily analyses of Northern Hemisphere temperature fields over 100 - 10 hPa produced by hand at FU Berlin); although agreement in the long-term mean temperatures and the trends (1957 - 2001) was generally good, they also found unrealistic behavior in ERA-40, which displayed larger biases in the October to January interval after 1979. *Tilmes et al.* (2006) focused specifically on the volume of air below the temperature threshold for PSC existence (V_{PSC}); they found that, although the general patterns of V_{PSC} evolution were similar for Met Office and ERA-40 reanalyses as well as ECMWF operational analyses and data from FU-Berlin, differences between the two reanalyses were as large as 10 % during their period of overlap (1991 - 1999). *Rieder and Polvani* (2013) also touched on comparisons of V_{PSC}, computing it using temperatures from MERRA, ECMWF Interim Reanalysis (ERA-Interim), and NCEP-NCAR R1. Although the depiction of year-to-year variability was seen to be strongly correlated among the three reanalyses, the magnitude of V_{PSC} varied considerably, with MERRA and ERA-Interim indicating V_{PSC} values roughly 30 % larger than those from NCEP-NCAR R1. Rieder and Polvani reiterated the cautions raised earlier about using NCEP-NCAR R1 in detailed polar processing studies.

A few studies have looked at the impact of differences in meteorological fields on results from chemical transport models (CTMs). *Davies et al.* (2003) investigated the effects of denitrification on ozone depletion in a cold Arctic winter by forcing a 3D CTM incorporating different PSC schemes with both UKMO and ECMWF analyses, finding that the two meteorological data sets led to disparate patterns of modeled PSC formation and denitrification, and consequently also chlorine activation and ozone loss. Similarly, *Feng et al.* (2005) applied the same CTM and analyses to examine the evolution of the ozone hole during the highly disturbed Antarctic winter of 2002; although runs driven by both analyses reproduced the anomalous conditions in that winter, differences in the structure and magnitude of simulated total ozone were seen.

Lawrence et al. (2015) revisited the use of polar processing diagnostics to evaluate reanalyses, performing a comprehensive intercomparison of NASA's Modern Era Retrospective-analysis for Research and Applications (MERRA) and ERA-Interim over the period 1979 to 2013. Agreement between the two meteorological data sets changed substantially during this interval, with many stratospheric temperature and vortex characteristics converging to greater consistency over time as more high-quality observations were assimilated. *Lawrence et al.* (2015) concluded that for the years since 2002 the MERRA and ERA-Interim reanalyses are equally appropriate choices and either can be used with confidence in polar processing studies in both hemispheres. In a follow-up study, *Lawrence et al.* (2018) extended the application of polar processing diagnostics to encompass other current full-input reanalyses, including MERRA-2, the Japanese 55-year Reanalysis (JRA-55), and the NCEP Climate Forecast System Reanalysis / Climate Forecast System, version 2 (CFSR/CFSv2). Results from this later study are described in detail in *Section 10.4*.

In summary, several previous studies have found considerable discrepancies between meteorological analyses/reanalyses in various parameters of relevance for polar processing, revealing that significant quantitative and qualitative differences may arise from the choice of which meteorological products are used in a given study. The recent work of *Lawrence et al.* (2015, 2018) indicates that agreement among various modern reanalyses improved substantially for some polar processing diagnostics in the post-2001 timeframe, following the introduction of new data streams. Nevertheless, previous studies have not examined all reanalyses of interest for S-RIP; moreover, a comparison of metrics not explored in earlier papers would be informative. Thus a comprehensive reassessment is warranted.

In this chapter we intercompare recent full-input reanalyses using an extensive set of polar processing diagnostics. The specific reanalyses considered here are: MERRA, MERRA-2, ERA-Interim, JRA-55, and CFSR/CFSv2. *Fujiwara et al.* (2017) provide an overview of these reanalysis systems, and they are also described in detail in *Chapter 2* of this Report. We note that ECMWF stopped producing ERA-Interim in August 2019 and replaced it with ERA5. Because the bulk of the analysis for this chapter had already been completed by the time ERA5 became available, its performance has not been assessed here. Although we expect that ERA5 will prove to be at least as reliable for polar processing studies as other modern reanalyses, we can make no conclusive judgments about its suitability for such studies at this time.

The primary focus here is on reanalyses; however, in some cases where comparisons with atmospheric observations are made we also examine the ECMWF operational analysis (OA) and the NASA GMAO Goddard Earth Observing System Version 5.9.1 (GEOS-591) assimilation product. The ECMWF OA evaluated by *Hoffmann et al.* (2017a), whose results are summarized in *Section 10.6*, is characterized by 3-hr temporal resolution, 0.125°×0.125° horizontal resolution, and 91 vertical levels with an upper lid at 0.01 hPa. The GEOS-591 near-real-time analysis, which was produced by the GEOS-5 data assimilation system (*Molod et al.*, 2015; *Rienecker et al.*, 2011), was characterized by 3-hr temporal resolution, 0.625°×0.5° horizontal resolution, and 72 vertical levels with an upper lid at 0.01 hPa. This stable system, used by NASA Earth Observing System satellite instrument teams in their data processing, provided consistent meteorological fields over much of the Aura record and was thus somewhat akin to a reanalysis. It was assessed by *Lambert and Santee* (2018), whose results are summarized in *Section 10.7*.

Much of this chapter focuses on process-oriented and case studies. For many diagnostics, the degree of agreement between reanalyses is an important direct indicator of the systems' inherent uncertainties, for which comparisons to independent measurements are not required. In addition, some diagnostics are based on PV or other dynamical quantities that cannot be provided directly by any measurement system. These situations pertain to many of the polar temperature and vortex diagnostics presented in *Section 10.4*, including minimum lower stratospheric temperature, area and volume of stratospheric air with temperatures below PSC existence thresholds, maximum latitudinal gradients in PV (a measure of the strength of the winter polar vortex), area of the vortex exposed to sunlight each day, and vortex breakup dates, as well as the polar cap average diabatic heating rates discussed in *Section 10.5*. On the other hand, comparisons with atmospheric measurements can be made for some diagnostics, especially the more derived ones. Such comparisons typically demand fairly broad spatial coverage on a daily basis, which is best afforded by satellite measurements. For the most part, comparisons between reanalysis fields and independent observations are left to *Chapter 3*; *Long et al.* (2017) also presented comparisons of reanalysis temperatures against satellite observations. However, analyses/reanalyses are evaluated through comparisons with long-duration superpressure balloon temperature and wind measurements in *Section 10.6*. In addition, the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) global navigation satellite system (GNSS) radio occultation (RO) temperatures are examined in connection with PSC thermodynamic-consistency diagnostics in *Section 10.7*. The latter section also relies on vertical profiles of gas-phase HNO_3 and H_2O measured by the Aura Microwave Limb Sounder (MLS), as well as PSC characteristics determined from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar aerosol and cloud backscatter. Finally, because results from chemical models synthesize the interplay among the spatially and temporally varying differences between reanalyses and exemplify how their net effects impact the bottom-line conclusions of typical real-life studies, in *Section 10.8* we compare simulated sequestration of HNO_3 and H_2O in PSCs, chlorine activation, and ozone fields with those observed by Aura MLS and the Envisat Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) for one winter in each hemisphere. Model-based estimates of Antarctic chemical ozone loss in the stratospheric partial column are also compared with those derived from MLS data. Other commonly used ozone loss metrics such as ozone hole area, ozone mass deficit, *etc.*, are not included here, nor are other processes that affect polar ozone but that are covered extensively elsewhere in this Report (*e.g.*, sudden stratospheric warmings are discussed in *Chapter 6*). Further direct comparisons between observations and the ozone fields from the reanalyses can be found in *Chapter 4*.

10.2 Description of atmospheric measurements

10.2.1 Aura Microwave Limb Sounder

MLS measures millimeter- and submillimeter-wavelength thermal emission from the limb of Earth's atmosphere (*Waters et al.*, 2006). The Aura MLS field-of-view (FOV) points in the direction of orbital motion and vertically scans the limb in the orbit plane, providing data coverage from 82°S to 82°N latitude on every orbit. Because the Aura orbit is sun-synchronous (with a 13:45 local time ascending equator-crossing time), MLS observations at a given latitude on either the ascending (mainly day) or descending (mainly night) portions of the orbit have the same local solar time. Northern high latitudes are sampled by ascending measurements near midday local time, whereas southern high latitudes are sampled by ascending measurements in the late afternoon. Vertical profiles are measured every ~ 165 km along the suborbital track, yielding a total of ~ 3500 profiles per day.

Here, we use the MLS version 4.2 (v4.2) data (*Livesey et al.*, 2020). Detailed information on the quality of a previous version of MLS data, v2.2, can be found in dedicated validation papers by *Lambert et al.* (2007) for stratospheric H_2O, *Santee et al.* (2007) for HNO_3, *Santee et al.* (2008) for ClO, *Froidevaux et al.* (2008a) for HCl, *Froidevaux et al.* (2008b) for stratospheric O_3, and *Schwartz et al.* (2008) for temperature. The precision, resolution, and useful vertical range of the v4.2 measurements, as well as assessments of their accuracy through systematic error quantification (and, in some cases, validation comparisons with correlative data sets), are reported for each species by *Livesey et al.* (2020). Briefly, MLS measurements have single-profile precisions (accuracies) of 4 - 15 % (4 - 20 %) for H_2O, 0.6 ppbv (1 - 2 ppbv) for HNO_3, 0.1 ppbv (0.05 - 0.25 ppbv) for ClO, 0.2 - 0.3 ppbv (0.2 ppbv) for HCl, 0.05 - 0.1 ppmv (0.1 - 0.25 ppmv) for O_3, and 0.6 - 1.2 K (0 - 5 K) for temperature in the stratosphere.

We note that MERRA-2 assimilates MLS temperatures, but only at pressures less than 5 hPa and not within the pressure range investigated here (*Gelaro et al.*, 2017).

Errors in the MLS H_2O contribute a few tenths of a kelvin to the error in calculated frost point temperatures and are substantially smaller than the errors in the temperature limb sounding retrievals obtained from MLS. From August 2004 until December 2013, mean differences between NOAA frost point hygrometer and MLS H_2O data showed no statistically significant differences (agreement to better than <1 %) from 68 - 26 hPa, although significant biases at 100 hPa and 83 hPa were found to be 10 % and 2 %, respectively (*Hurst et al.*, 2014). However, increasing the time frame to mid-2015 revealed a long-term drift in MLS H_2O of up to 1.5 % per year starting around 2010 (*Hurst et al.,* 2016). Although changes to the MLS data processing system have substantially mitigated this drift in the version 5 MLS H_2O measurements (*Livesey et al.*, 2021), for the v4 data used here the effect on the calculated supercooled ternary solution (STS) reference and frost point temperatures is less than 0.1 K per year.

To aid in the analysis of MLS measurements, particularly in *Section 10.7,* we make use of MLS Derived Meteorological Products (DMPs). These files contain meteorological data (*e.g.*, temperature) and derived parameters (*e.g.*, equivalent latitude) interpolated from gridded reanalysis fields to the along-track geolocations of the MLS measurements. The original version of the MLS DMPs was described in detail by *Manney et al.* (2007). Here we use updated files (version 2, the DMP version of record for the MLS v4.2 data; see the MLS web page, **http://mls.jpl.nasa.gov**, for more details); the v2 DMPs are from the software described by *Manney et al.* (2011a). DMP files containing associated meteorological information at the MLS measurement locations have been produced for all five full-input reanalyses considered here.

10.2.2 Envisat MIPAS

The MIPAS instrument (*Fischer et al.*, 2008) was launched in March 2002 on the ESA Environment Satellite (Envisat) and was operational until April 2012. MIPAS was an infrared Fourier transform spectrometer for measuring limb emission spectra between 685 cm^{-1} and 2410 cm^{-1} (14.6 - 4.15 µm). Through azimuth scanning it provided global coverage from 87.5 °S to 89.3 °N. The instrument FOV was 30 km across-track and 3 km in the vertical, and the horizontal along-track sampling distance for nominal-mode observations was ~530 km from 2002 to 2004 and ~400 km from 2005 onward. Several retrieval algorithms have been developed for the MIPAS spectra; here we use profiles of temperature and atmospheric constituents generated by the KIT-IMF-ASF (Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Trace Gases and Remote Sensing) group in cooperation with the Instituto de Astrofísica de Andalucía (*von Clarmann et al.*, 2009). For $ClONO_2$ below 40 km, precision is 8 - 14 %, with vertical resolution 2.5 - 9 km (*Höpfner et al.*, 2007; *von Clarmann et al.*, 2009). Retrieval of $ClONO_2$ is hindered by the presence of optically thick PSCs along the MIPAS line of sight.

10.2.3 CALIPSO CALIOP

The CALIOP dual-wavelength elastic backscatter lidar (*Winker et al.*, 2009) flies on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite launched in April 2006. We use the CALIOP Level 2 operational data set L2PSCMask (v1 Polar Stratospheric Cloud Mask Product) produced by the CALIPSO science team. The Level 2 operational data consist of nighttime-only data and contain profiles of PSC presence, composition, optical properties, and meteorological information along the CALIPSO orbit tracks at a horizontal resolution of 5 km and a vertical resolution of 180 m. We have applied post-processing to generate coarser horizontal/vertical bins for a better comparison at the scale of the MLS along-track and vertical resolution (see *Section 10.7* for details). Each averaging bin is the size of the MLS along-track vertical profile separation (165 km) and the height between the mid-points of the retrieval pressure levels (2.16 km) for the MLS HNO3 data product. This we refer to as the MLS geometric FOV. There are approximately four hundred 5 km × 180 m CALIOP "pixels" within the MLS geometric FOV.

The CALIOP PSC classification scheme used here is described by *Pitts et al.* (2009), with modifications discussed by *Pitts et al.* (2013), and consists of four main PSC types. MIX1 and MIX2 denote detections of nitric acid trihydrate (NAT) particles, with the MIX1/MIX2 boundary marking a transition between lower (MIX1) and higher (MIX2) NAT number/volume densities. The STS type indicates supercooled liquid ternary solution ($H_2SO_4/HNO_3/H_2O$) particles, and ICE indicates water-ice particles.

10.2.4 COSMIC GNSS-RO

We use the US/Taiwan Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) network data obtained from the Universities for Cooperative Atmospheric Research (UCAR) COSMIC Data Analysis and Archive Center (CDAAC). Global navigation satellite system radio occultation (GNSS-RO) data have provided high accuracy (bias < 0.2 K and precision > 0.7 K, *Gobiet et al.*, 2007), global (day and night) coverage, coupled with excellent long term stability, for nearly two decades (*Anthes*, 2011). The vertical resolution is better than about 0.6 km over the 15 - 30 km vertical range considered here. The introduction of GNSS-RO has been documented to improve numerical weather prediction (NWP) forecast skill in the ECMWF Integrated Forecast System (IFS) (*Bonavita*, 2014) and to reduce tropopause and lower stratospheric

temperature biases in ERA-Interim (*Poli et al.*, 2010). The direct assimilation of bending angles or refractivity is now the common practice for many global reanalyses; however, for many other purposes the production of vertical atmospheric temperature profiles from GNSS-RO data is required. The retrieval of vertical atmospheric geophysical profiles from RO requires a number of assumptions because of the long ray path through a non-uniform atmosphere (*Ho et al.*, 2012). Therefore, corrections are required for ionospheric effects, variations in water vapour, and gradients in temperature along the ray path (*Anthes*, 2011; *Poli and Joiner*, 2004). Many other studies have intercompared GNSS-RO with independent operational analyses, *e.g.,* with forecast versions that have not assimilated the GNSS-RO data. The near real time COSMIC data (in the form of bending angles or refractivity) are ingested by most of the data assimilation procedures considered here (except for MERRA), and therefore these reanalyses are not strictly independent of the postprocessed COSMIC temperatures. We have chosen to use the COSMIC temperatures as a common reference to evaluate the reanalysis departures, rather than using the reanalysis ensemble mean.

10.2.5 Concordiasi superpressure balloon measurements

Superpressure balloons are aerostatic balloons, which are filled with a fixed amount of lifting gas, and for which the maximum volume of the balloon is kept constant by means of a closed, inextensible, spherical envelope. After launch, the balloons ascend and expand until they reach a float level where the atmospheric density matches the balloon density. On this isopycnic surface a balloon is free to float horizontally with the motion of the wind. Hence, superpressure balloons behave as quasi-Lagrangian tracers in the atmosphere. The Concordiasi field campaign in Antarctica in September 2010 to January 2011 was aimed at making innovative atmospheric observations to study the circulation and chemical species in the polar lower stratosphere and to reduce uncertainties in diverse fields in Antarctic science (*Rabier et al.*, 2010). During the field campaign, 19 superpressure balloons with 12 m diameter were launched from McMurdo Station (78 ° S, 166 ° E), Antarctica, by the French space agency, Centre National d'Etudes Spatiales (CNES). Balloons of this size typically drift at pressure levels of ~ 60 hPa and altitudes of ~ 18 km. The balloons were launched between 8 September and 26 October 2010, and each balloon flew in the mid- and high-latitude lower stratosphere for a typical period of 2 to 3 months.

The positions of the balloons were tracked every 60 s by means of global positioning satellite (GPS) receivers. At each observation time the components of the horizontal wind are computed by finite differences between the GPS positions. The uncertainty is about 1 m for the GPS horizontal position and 0.1 m s^{-1} for the derived winds (*Podglajen et al.*, 2014). Each balloon launched during Concordiasi was equipped with a meteorological payload called the Thermodynamical SENsor (TSEN). TSEN makes in situ measurements of atmospheric pressure and temperature every 30 s during the whole flight. The pressure is measured with an accuracy of 1 Pa and a precision of 0.1 Pa. The air temperature is measured via two thermistors. During daytime, the thermistors are heated by the sun, leading to daytime temperature measurements being warmer than the real air temperature. An empirical correction has been used to correct for this effect, which is described in detail by *Hertzog et al.* (2004). The precision of the corrected temperature observations is about 0.25 K during daytime and 0.1 K during nighttime.

Most of the measurements (*i.e.*, more than 90 %) took place between 25 September and 22 December 2010, at an altitude range of 17.0 - 18.5 km, and within a latitude range of 59 ° - 84 ° S. The pressure measurements are mostly within a range of 58.2 - 69.1 hPa and the temperature measurements within 189 - 227 K. The density of air, calculated from pressure and temperature, varies between 0.099 kg m^{-3} and 0.120 kg m^{-3}. The zonal winds are predominately westerly and mostly within a range of 1 - 44 m s^{-1}. The meridional wind distributions are nearly symmetric, with meridional winds being in the range of ± 17 m s^{-1}. Horizontal wind speeds are mostly within 5 - 47 m s^{-1}.

The Concordiasi balloon observations have been assimilated into the ECMWF, MERRA, and MERRA-2 data sets, but they were not considered for NCEP-NCAR R1. The observations therefore provide an independent data source only for the validation of the NCEP-NCAR R1 data set. However, as meteorological analyses are a result of combining various satellite and in situ observations, a forecast model, and a data assimilation procedure, a comparison of the meteorological data with the Concordiasi observations still provides information on the performance of the overall system, even for the reanalyses that assimilate those observations. As the observational data have been subject to downsampling and data thinning before they were assimilated, an assessment of the representation of small-scale structures due to gravity waves also remains meaningful.

Trajectory calculations for the Concordiasi balloon observations have been analyzed using the Lagrangian particle dispersion model Massive-Parallel Trajectory Calculations (MPTRAC) (*Hoffmann et al.*, 2016). Transport is simulated by calculating trajectories for large numbers of air parcels based on given wind fields from global meteorological reanalyses. The numerical accuracy and efficiency of trajectory calculations with MPTRAC was assessed by *Rößler et al.* (2018). Turbulent diffusion and subgrid-scale wind fluctuations are simulated based on the Langevin equation, closely following the approach implemented in the Flexible Particle

(FLEXPART) model (*Stohl et al.*, 2005).

10.3 Overview of reanalysis polar temperature differences

To provide context for later results derived from more complex analysis techniques, we show in **Figure 10.1** a basic overview of reanalysis temperatures in the polar lower stratosphere. We have chosen as a suitable metric the daily (12 UT) mean 60° polar cap temperature differences at 46 hPa. Time series of the temperature differences calculated for MERRA, MERRA-2, JRA-55, and CFSR/CFSv2 relative to ERA-Interim over 2008 - 2013 have been smoothed using a 10-day boxcar average. The daily mean standard error of the temperature differences is less than 0.1 K. In both hemispheres, temperature differences display annual cycles, with positive deviations mainly in summer and negative deviations mainly in winter. In the Antarctic the largest deviations are ~ 1 K in MERRA – ERA-Interim, whereas in the Arctic the largest deviations are in JRA-55 – ERA-Interim, but they only reach ~ 0.5 K.

The grey-shaded regions in **Figure 10.1** mark the useful wintertime measurement periods, chosen to capture the bulk of the PSC activity needed for the evaluation of the thermodynamic temperature comparisons that are discussed in *Section 10.7*. These periods also happen to largely coincide with times of smaller variability in the temperature differences, when biases of the other reanalyses with respect to ERA-Interim are predominantly negative. Therefore, we caution that intercomparisons of reanalyses undertaken using other time periods, especially for summertime, could even obtain temperature deviations of opposite sign whilst maintaining about the same magnitude. Indeed, whereas *Hoffmann et al.* (2017a) find MERRA to be the warmest and ERA-Interim the coldest compared to superpressure balloon temperature measurements made during the Antarctic Concordiasi campaign in September 2010 to January 2011, *Lambert and Santee* (2018) find the opposite order compared to the COSMIC and thermodynamic temperature references for May to August during 2008 to 2013. To reconcile this apparent discrepancy, in **Figure 10.2** daily mean temperature differences (at 12 UT) for MERRA and MERRA-2 relative to ERA-Interim are used to highlight the non-overlapping intervals of the PSC analysis window (green line) and the balloon flights (red line). Measurements in the later time period of the Concordiasi balloon flights (September - December) clearly sample different atmospheric conditions than those prevailing in the earlier time period (May - August). Moreover, differences between reanalysis temperatures along individual balloon trajectories are likely to be amplified compared to the differences in mean polar cap temperatures. We note that MERRA does not assimilate COSMIC data, whereas MERRA-2 and the other reanalyses investigated here do; hence some of the reduction in the bias of MERRA-2 compared to MERRA seen in **Figure 10.2** is likely attributable to the former's use of GNSS-RO data.

10.4 Polar temperature and vortex diagnostics

Lawrence et al. (2018) expanded on the diagnostics in *Lawrence et al.* (2015) and applied them to CFSR/CFSv2,

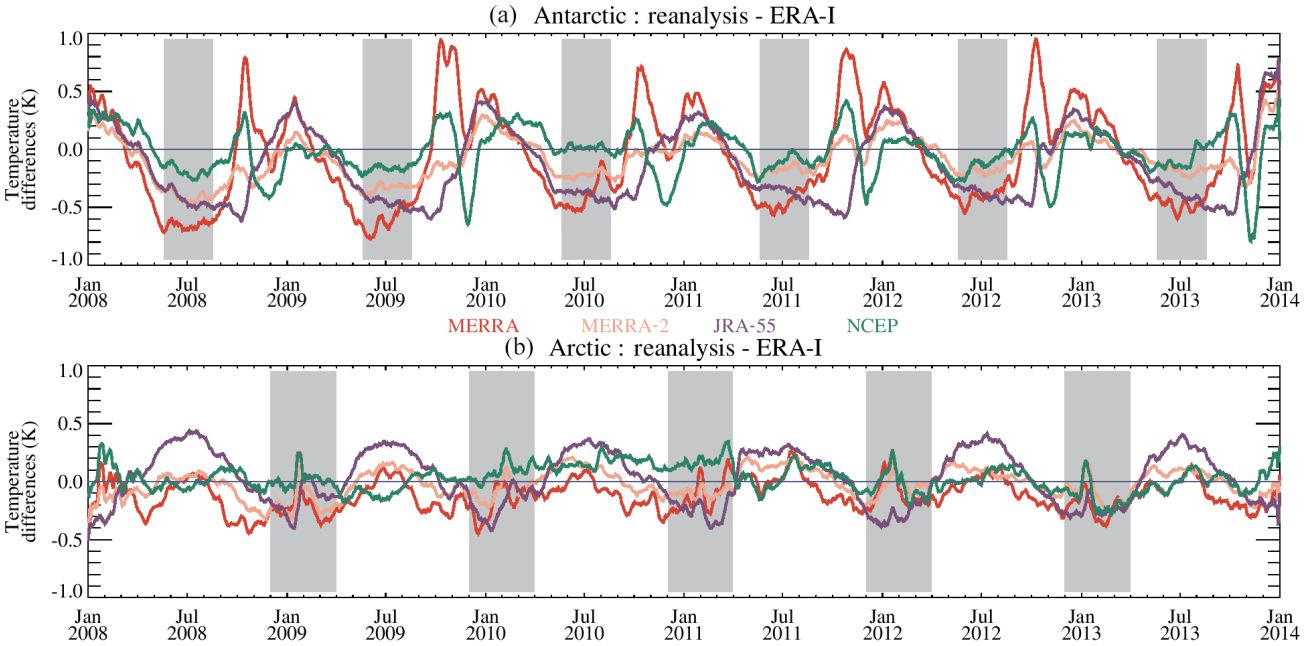

Figure 10.1: *(a) Time series (for 12 UT) from 2008 to 2013 of 10-day boxcar-smoothed temperature differences for MERRA, MERRA-2, JRA-55, and CFSR/CFSv2 relative to ERA-Interim at 46 hPa, averaged over the 60° Antarctic polar cap. The four reanalyses being differenced against ERA-Interim are shown in the colors indicated in the legend between the two panels. Grey regions indicate the periods defined for the analysis of PSC-related metrics (see Section 10.7). (b) Same, but for the Arctic. From Lambert and Santee (2018).*

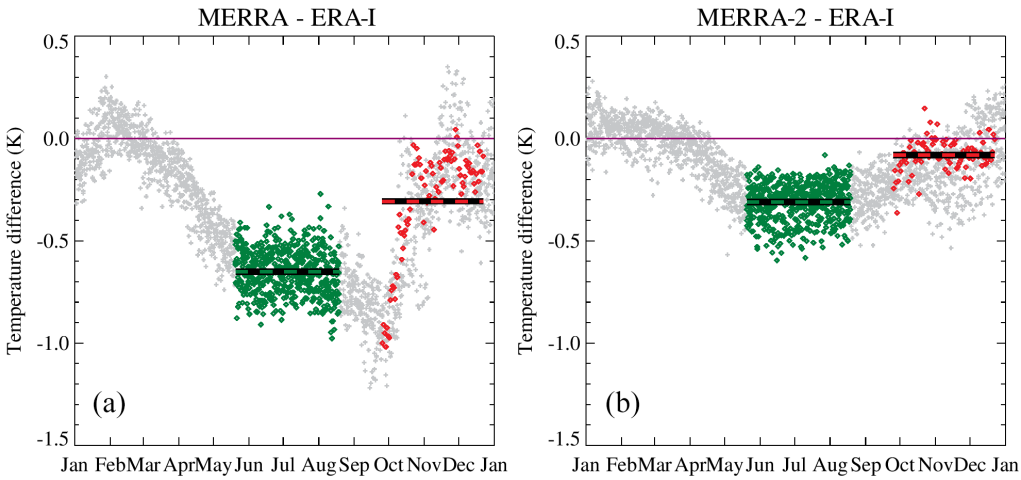

Figure 10.2: *Daily mean temperature differences (grey) of (a) MERRA and (b) MERRA-2, relative to ERA-Interim at 62 hPa (representative of the Concordiasi balloon float heights (see Section 10.6) in the 60° Antarctic polar cap for 2008 - 2013. The green-black dashed line indicates the mean of the 2008 - 2013 differences (green symbols) during the PSC analysis window. The red-black dashed line indicates the mean of the differences (red symbols) over the time span of 90 % of the Concordiasi balloon measurements in 2010 (Hoffmann et al., 2017a). From Lambert and Santee (2018).*

ERA-Interim, JRA-55, and MERRA-2 (evaluations were also done for MERRA but were not included in the paper). The full suite of diagnostics examined by *Lawrence et al.* (2018) includes minimum temperatures poleward of ± 40° latitude, the area of temperatures (poleward of ± 30° latitude) below PSC existence thresholds, the winter mean volume of lower stratospheric air with temperatures below PSC existence thresholds, maximum gradients in scaled PV as a function of equivalent latitude, the area of the vortex exposed to sunlight, and approximate dates of the breakup of the polar vortices. As noted by *Lawrence et al.* (2018), reanalysis comparisons are particularly critical to assess the uncertainties in these types of diagnostics because they are not quantities that can be compared directly with observations. The key results of *Lawrence et al.* (2018) are summarized below.

Figure 10.3 compares Southern Hemisphere (SH) extended winter season (MJJASO) minimum temperatures poleward of 40° S in the lower stratosphere from MERRA-2, ERA-Interim, JRA-55, and CFSR/CFSv2 with those from the reanalysis ensemble mean (REM). Since the REM values that go into each season's mean vary from day to day and are expected to change with any large change in any of the reanalyses, the differences from the REM quantify only how far each reanalysis is from that mean during each season, thus giving an idea of the range of values (which could be interpreted as an uncertainty in the diagnostic) but not of the absolute changes in those values. The standard deviations shown on the right of **Figure 10.3** help further quantify the spread among the reanalyses. The reanalyses converge towards much better agreement in the later years at all levels. There are step-like changes in agreement among the reanalyses around 1998 (especially ERA-Interim and MERRA-2), when the reanalyses (albeit not all at exactly the same time) changed from assimilating Tiros Operational Vertical Sounder (TOVS) to advanced TOVS (ATOVS) radiances; the latter provide higher-resolution constraints on stratospheric temperatures. On average, the individual reanalyses agree with the REM to within about 0.5 K in the most recent decade, though differences in the early years commonly exceed 3 K. The standard deviations of the differences increase with altitude (indicating larger maximum differences between the reanalyses) and show a modest decrease over the time period, with an abrupt decrease seen around 1998 in most reanalyses. CFSR/CFSv2 generally shows larger variance than the other reanalyses in the period since about 2002, and less of a change around 1998. These features are consistent with those from comparisons of other polar processing diagnostics. In the Northern Hemisphere (NH), the reanalyses agree much better throughout the 36 years (to within about 1.5 K before 1999), though convergence toward better agreement is also seen in most of the reanalyses (excepting CFSR/CFSv2) (*Lawrence et al.*, 2018); this is not unexpected since the input data density (from, *e.g.*, sondes and ground-based measurements) in the years before the TOVS/ATOVS transition was much greater in the NH, and thus the reanalyses' temperatures before that transition were much better constrained than those in the SH. *Lawrence et al.* (2018) also evaluated the area of temperatures below PSC thresholds, with results consistent with those shown here for minimum temperatures.

Figure 10.4 shows differences from the REM of daily maximum PV gradients (a measure of vortex strength) with respect to equivalent latitude averaged over the DJFM season in the NH, as well as the standard deviations of the daily differences. The climatological maximum values of this diagnostic increase with height from around $1 - 6 \times 10^{-6}$ s^{-1} deg^{-1} at about 430 K to over 20×10^{-6} s^{-1} deg^{-1} above about 600 K (*Lawrence et al.*, 2018). There is no obvious systematic decrease in the differences from the REM, and a very slight apparent decrease in the standard deviations for most reanalyses. Not shown is a small

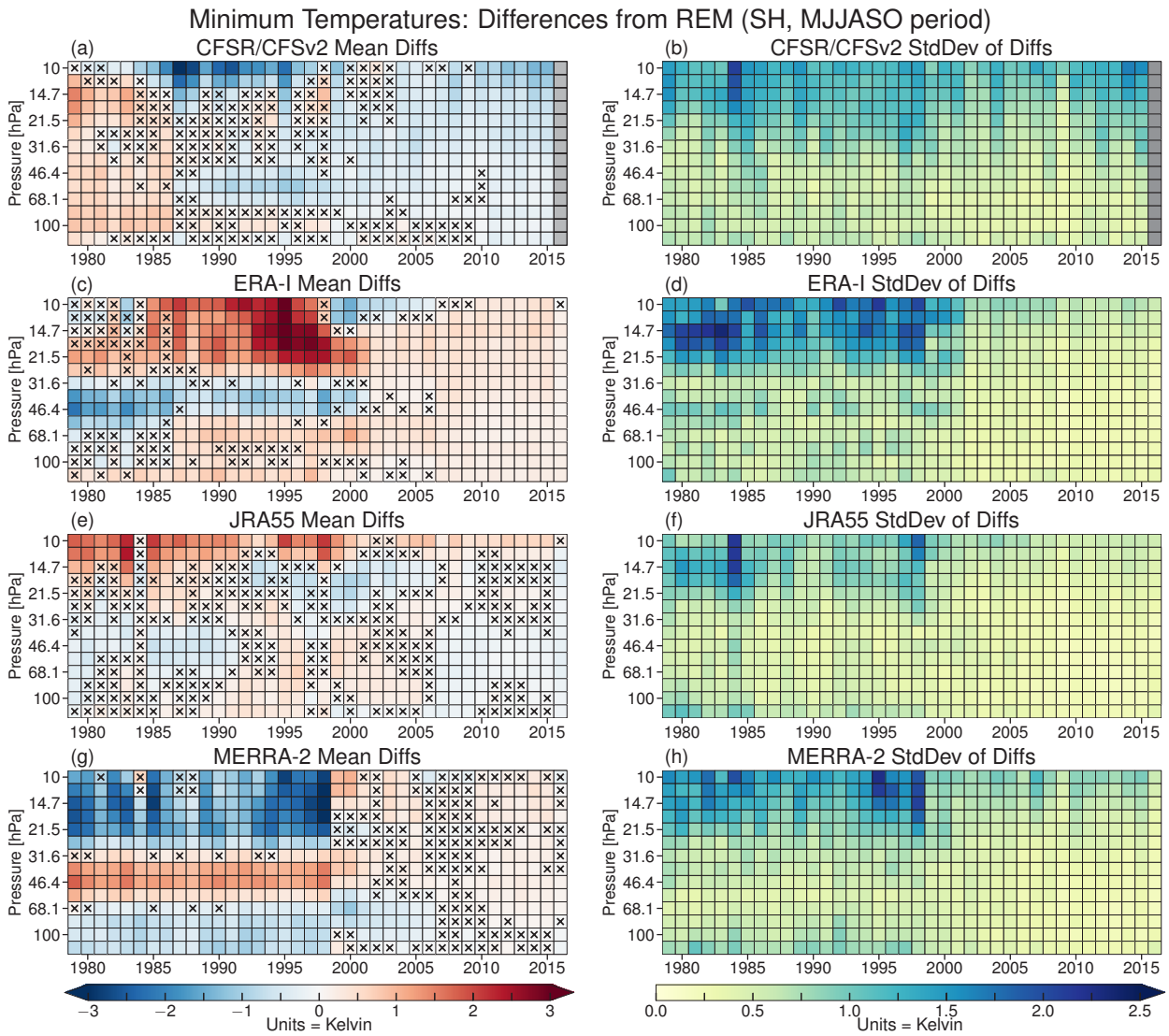

Figure 10.3: *SH extended winter season (MJJASO) (a, c, e, g) averages and (b, d, f, h) standard deviations of minimum daily temperature differences for each reanalysis from the reanalysis ensemble mean (REM, see text) as a function of year and pressure for the 1979 through 2017 winters, concatenated into pixel plots as described by Lawrence et al. (2018). Columns of grey pixels indicate years with no data. Pixels with x symbols inside indicate years and levels where the differences from the REM are insignificant according to our bootstrapping analysis (see Lawrence et al., 2018). In the average difference panels, negative values (reanalysis less than REM) are shown in blue and positive values (reanalysis greater than REM) are shown in red; in the standard deviation panels, yellows/deep blues represent low/high standard deviations of the reanalysis differences, respectively. From Lawrence et al. (2018).*

systematic decrease in the difference between MERRA and MERRA-2 for a narrow range of levels from 580 K to 700 K. JRA-55 shows generally stronger maximum PV gradients than the REM up through about 750 K, whereas CFSR/CFSv2 shows weaker gradients through most of that range. ERA-Interim and MERRA-2 are generally closer to the REM, with regions of stronger and weaker gradients alternating with height. As can be seen from the standard deviations, the variability in the daily differences increases strongly above about 520 K. The standard deviations also highlight a period of large variance in the differences between about 1995 and 2002 in the highest levels shown (above about 580 K) in ERA-Interim, JRA-55, and MERRA-2; this is a period when many data inputs were changing (including the TOVS to ATOVS transition) and during which the timing of those changes and the way they were handled (*e.g.*, whether a simple switch or using both older and newer inputs for a time) vary among the reanalyses. The maximum PV gradient diagnostic for the SH shows an improvement in agreement with the REM after about 1999, albeit not as strong as that for temperature diagnostics (*Lawrence et al.*, 2018).

Figure 10.5 shows the winter-mean (DJFM) volume of air with temperature below the NAT PSC threshold for the NH for each of the reanalyses, expressed as a fraction of the volume of air in the vortex. (The altitude for the depth in the volume calculation is obtained using the theta to altitude conversion approximation of *Knox* (1998).) The

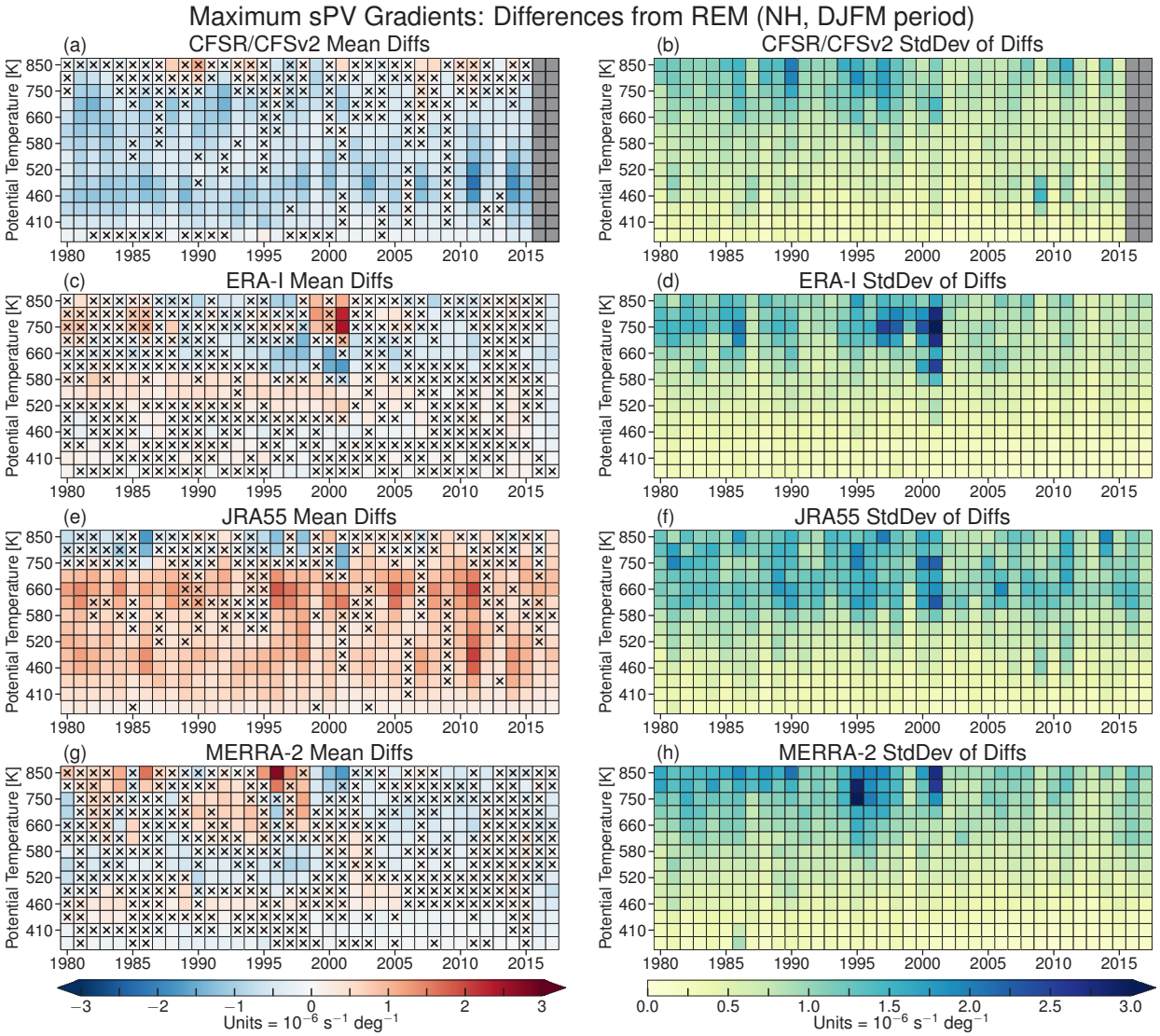

Figure 10.4: *As in* ***Figure 10.3****, but for NH maximum PV gradients for DJFM. Units are in 10^{-6} s^{-1} deg^{-1}. From Lawrence et al. (2018).*

extent of the bars shows the sensitivity of the diagnostic to using ± 1 K offsets in the threshold temperatures. Those threshold temperatures are calculated according to *Hanson and Mauersberger* (1988) for standard pressure surfaces (12 levels per decade in pressure, as have been used for several NASA satellite instrument data sets) using climatological HNO_3 and H_2O profiles (from all January Cryogenic Limb Array Etalon Spectrometer (CLAES) data and all January UARS MLS data, respectively); these values are most accurate for conditions that are neither denitrified nor dehydrated (see *Lawrence et al.*, 2018, and references therein, for further details). The threshold temperatures are then assigned to the "corresponding" standard theta level; the bar ranges for the ± 1 K offsets thus help to estimate uncertainties from the approximations for the HNO_3 and H_2O profiles and for the conversion to theta levels. The large interannual variability in the NH PSC potential is reflected accurately in all of the reanalyses throughout the period shown. Differences between the reanalyses are small throughout the period and do not show an obvious convergence to closer agreement. JRA-55 often shows a slightly larger volume than the other reanalyses. The sensitivity to the ± 1 K offsets is, as expected, largest when the volume itself is large. The corresponding diagnostics for the SH are shown in *Lawrence et al.* (2018); they indicate generally smaller differences, consistent with less interannual variability and much larger volumes with PSC potential in the Antarctic.

Lawrence et al. (2018) calculate vortex area and sunlit vortex area based on a vortex boundary defined by a climatological winter profile of the PV at the location of maximum PV gradients; that paper provides a detailed discussion of why this choice of vortex-edge definition is appropriate. Vortex area and sunlit vortex area show small but persistent biases in vortex size, with JRA-55 usually having a smaller vortex than ERA-Interim and CFSR/CFSv2 and the MERRA-2 vortex size difference from the REM varying with altitude. Consistent with a generally smaller vortex, JRA-55 tends to have earlier vortex decay dates than the other reanalyses evaluated; MERRA-2 has the latest vortex decay dates below about 550 K, and CFSR/CFSv2 the latest above that.

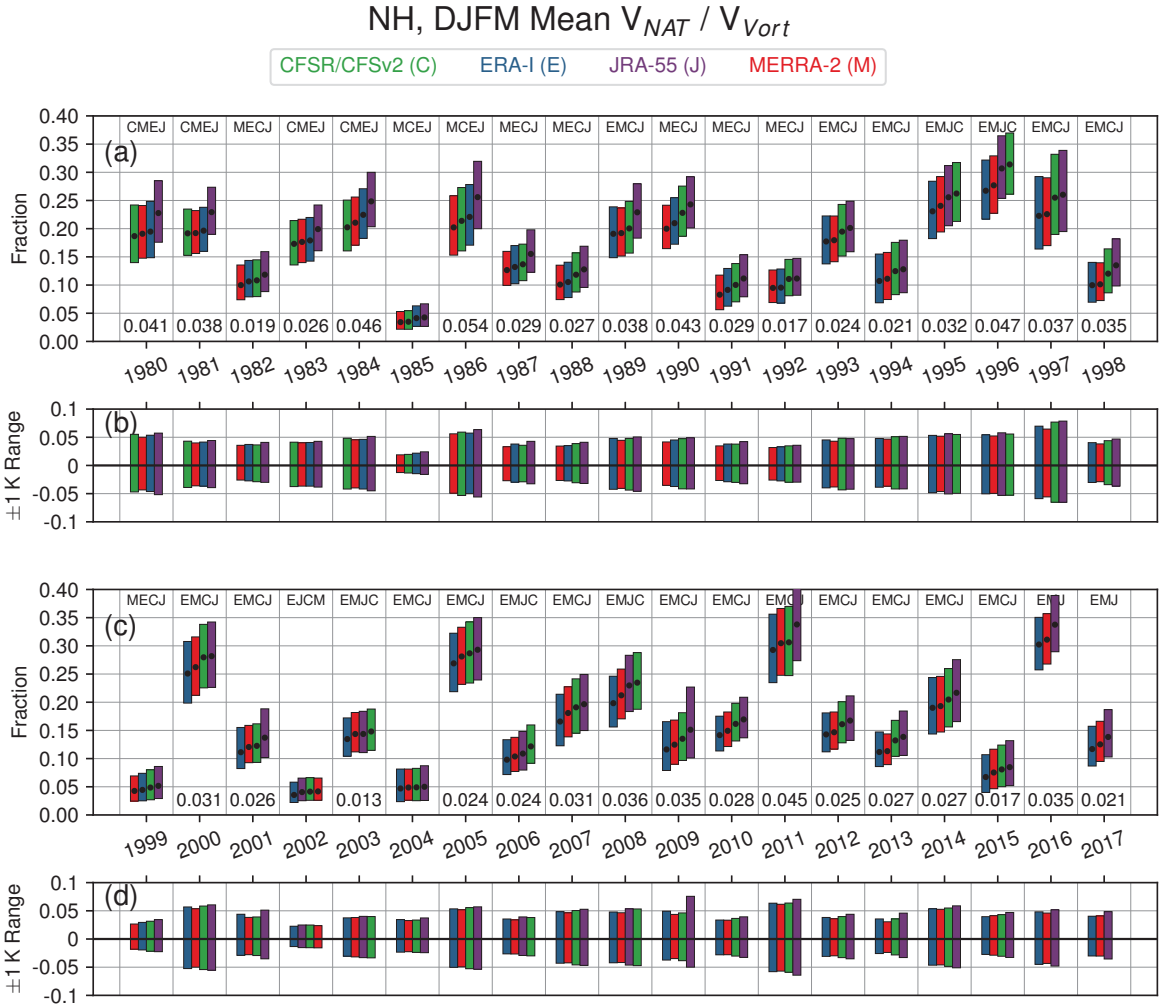

Figure 10.5: *Winter-mean (DJFM) fraction of vortex volume between the 390 K and 580 K isentropic surfaces with temperatures below T_{NAT} in the NH (a and c), and range of values obtained for the ± 1 K NAT threshold sensitivity tests (b and d). The bars are ordered from lowest to highest central values. The numbers at the bottom of (a) and (c) show the range of central values (that is, rightmost minus leftmost central value). Green, blue, purple, and red indicate CFSR/CFSv2, ERA-Interim, JRA-55, and MERRA-2, respectively. From Lawrence et al. (2018).*

Differences in vortex area and decay dates are small and do not obviously improve over the period studied.

10.5 Polar diabatic heating rates

Diabatic heating rates in the polar vortex regions are important to polar processing studies: Diabatic descent is the main driver of the downward transport of trace gases, including inorganic chlorine (Cl_y), and the replenishment of ozone in the lower stratospheric vortex, and it is thus critical to account for it in studies aimed at using observations to quantify chemical ozone loss (see, *e.g.*, *Manney et al.*, 2003a; *WMO*, 2007, and references therein, for reviews of methods for observational ozone loss studies). Although descent in the polar vortices can be estimated using quasi-conserved tracers such as N_2O or CO (*Ryan et al.*, 2018, and references therein), such estimates can be biased by mixing processes (*Livesey et al.*, 2015, and references therein), and in models and reanalyses the vertical motions of stratospheric air parcels are often constrained thermodynamically based on net diabatic heating, as initially laid out by *Murgatroyd and Singleton* (1961) and *Dunkerton* (1978).

Diabatic heating in the polar lower stratosphere is generally controlled by radiative effects – the sum of longwave cooling from thermal emission, and shortwave heating from solar absorption (when sunlight is present). For longwave cooling, the most important factors are temperature and the concentrations of major greenhouse gases carbon dioxide, water vapor, and ozone. Shortwave heating is determined primarily by the solar zenith angle and ozone amount. The physics of radiative transfer and the basic spectroscopic parameters needed to calculate net diabatic heating have been known for some time, and the foundation for understanding radiative heating and cooling in the stratosphere is well established (*e.g.*, *Mertens et al.*, 1999; *Olaguer et al.*, 1992; *Kiehl and Solomon*, 1986; *Ramanathan*, 1976). However, there are a number of important issues that remain for improving the accuracy of radiative heating and cooling calculations in the stratosphere, such as corrections to

broadband schemes used in models, variations in the representation of water vapor longwave radiative effects, and uncertainties in solar near-infrared spectral irradiances (*e.g.*, *Menang*, 2018; *Maycock and Shine*, 2012; *Forster et al.*, 2001).

We use the zonal mean model-generated diabatic heating rates from the MERRA-2, ERA-Interim, JRA-55, and CFSR reanalyses described by *Martineau et al.* (2018; Section 3.6.1 of that paper, also see *Wright*, 2017), which are provided on a common grid to facilitate comparisons. These are on a standard pressure grid and given as the potential temperature tendency due to all physics; further details, including the processes included and differences in products provided, are given by *Martineau et al.* (2018) [1]. For the figures shown here we construct daily averages from the six-hourly fields provided and average those over the polar cap (60°-90°) in each hemisphere. The climatological comparisons are done for 1980 through 2010 so that CFSR can be included (diabatic heating rates for CFSRv2 were not archived).

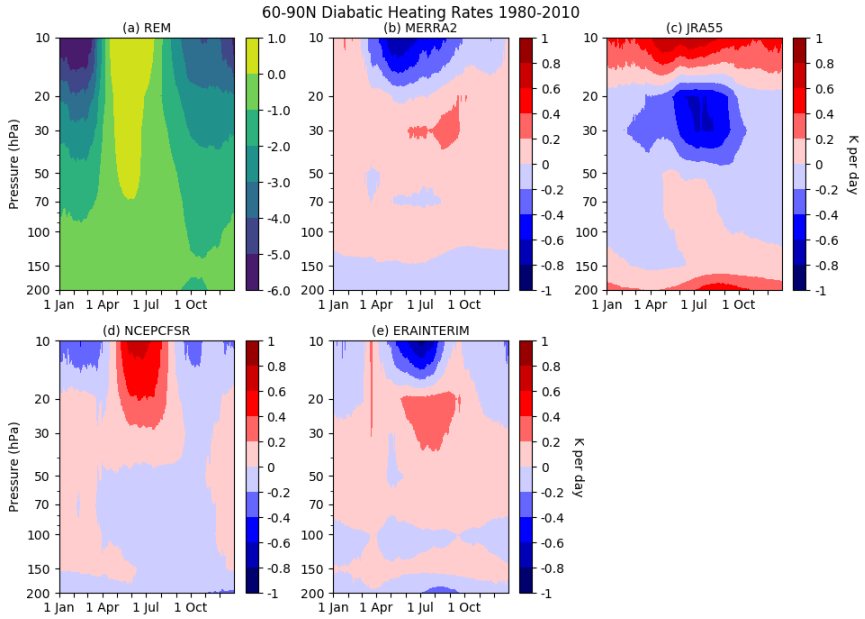

Figure 10.6: *Climatological 1980-2010 north polar cap (60°-90°N) diabatic heating rates in the lower stratosphere from four reanalyses: (a) REM for MERRA-2, ERA-Interim, JRA-55, and CFSR; (b through e) differences (reanalysis – REM, so positive values indicate that reanalysis values are higher than those for the REM) of each reanalysis from the REM climatology.*

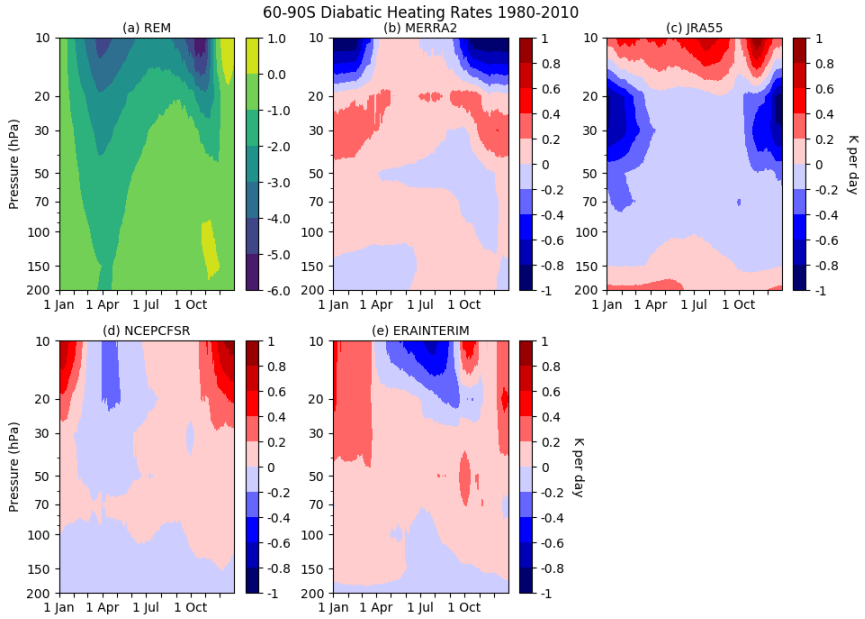

Figure 10.7: *As in* ***Figure 10.6*** *but for the south polar cap.*

Figure 10.6a shows NH polar cap averaged heating rates in the lower stratosphere for the REM (as defined in *Section 10.4*) constructed from MERRA-2, ERA-Interim, JRA-55, and CFSR. As expected, the strongest diabatic cooling (descent) is seen in late December through March, and stronger descent is seen at higher altitudes. Increased negative values in late January and February are coincident with the most common timing of strong sudden stratospheric warmings (SSWs), during which the temperatures far above radiative equilibrium give rise to stronger diabatic cooling (see, *e.g.*, *Manney et al.*, 2008, 2009b, and references therein). Differences of each reanalysis from the REM are shown in **Figure 10.6b** through **10.6e**. The largest differences between the reanalyses and the REM are at the lowest pressures, with differences up to about ±1 K/day near and above (at lower pressures than) 30 hPa. During the cold season (November through April) that we are most interested in here, the magnitude of the differences from the REM is typically no more than about 0.4 K/day except above (at pressures below) 20 hPa; cold-season differences are thus typically within about 10 %. While there is much interannual variability (in line with the large interannual variability in temperatures), examination of individual years indicates generally consistent biases between the reanalyses, with JRA-55 generally showing lower values (stronger diabatic cooling) than the other reanalyses in

[1] also see the footnote on diabatic heating rates in reanalyses in *Section 12.1.3*.

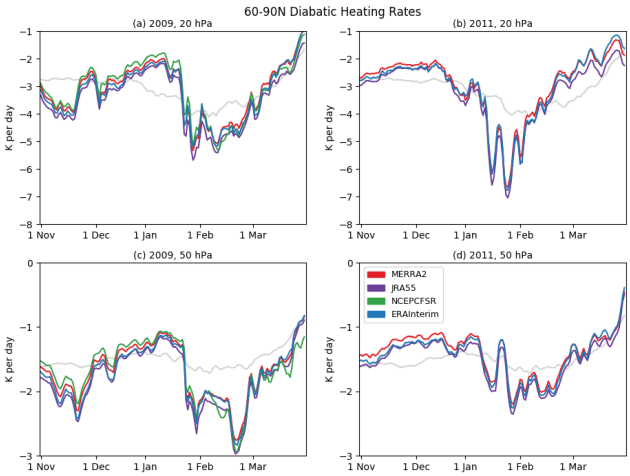

Figure 10.8: *North polar cap (60-90°N) average diabatic heating rates in the lower stratosphere from four reanalyses for 2009 (left panels) and 2011 (right panels) in the NH at 20 hPa (top panels) and 50 hPa (bottom panels) during the cold seasons. Coloured lines show the reanalyses as given in the legend. The grey line shows the REM 1980-2010 climatology for the same time of year as a reference. Values from CFSR are not available in 2011.*

a layer between approximately 50 hPa and 15 hPa in the warm season. The seasonality of the differences from the REM at the lowest pressures shown is strongly influenced by JRA-55, which has a weaker seasonal cycle in the diabatic heating rates than the other reanalyses.

Figure 10.7 shows a similar REM climatology and reanalysis differences for the SH polar cap. The strongest radiative cooling (largest negative values) is seen in late spring during the final warming (October and November), with strong cooling also seen in fall. As in the NH, the largest differences among the reanalyses occur during the warm season and in late spring, and at levels above (pressures below) about 20 hPa except for JRA-55 warm-season differences near 50 hPa to 15 hPa. Outside of those times and regions, the magnitude of differences is typically less than 0.4 K/day, and less than 0.2 K/day below (at pressures higher than) about 30 hPa, thus within about 15 %. As was the case in the NH, JRA-55 has a weaker seasonal cycle in the diabatic heating in the SH than the other reanalyses.

In light of the large changes in temperature agreement before and after the transition from assimilating TOVS to ATOVS data (*e.g.*, *Lawrence et al.*, 2018; *Long et al.*, 2017), we examined separately climatologies between 1980 and 1998, and between 2000 and 2010, for each hemisphere. Overall, the agreement among reanalysis diabatic heating rates does not differ substantially during the two periods, suggesting that the differences in the reanalyses' radiative transfer models and their inputs (*e.g.*, differences in handling ozone and water vapor) are likely a larger factor than changes in the satellite radiances used to define temperatures for the radiation calculations. Maximum values of differences at pressures below 30 hPa are slightly smaller for most of the reanalyses in the ATOVS period, but the differences are small and qualitative differences appear negligible in the climatology.

To examine the day-to-day differences among the reanalyses, we show time series at 20 hPa and 50 hPa for individual winters, choosing a relatively quiescent and a very disturbed winter in each hemisphere. **Figure 10.8** shows the NH cold seasons of 2008/2009 (a year with a strong, prolonged vortex-split SSW, *e.g.*, *Kuttippurath and Nikulin*, 2012; *Labitzke and Kunze*, 2009; *Manney et al.*, 2009a) and 2010/2011 (at the time, the year with the largest Arctic ozone loss on record, *e.g.*, *Balis et al.*, 2011; *Manney et al.*, 2011b; *Sinnhuber et al.*, 2011). At the onset time of the SSW in late January 2009, there is an abrupt increase in diabatic cooling (stronger negative values) at both levels shown, as temperatures rose far above radiative equilibrium. A previous period of unusually strong diabatic cooling occurred at both levels in November 2009, and a period of unusually weak diabatic cooling occurred immediately before the SSW onset. In contrast, in 2011, a sharp increase in cooling in late January at 20 hPa was associated with a minor SSW that did not strongly affect the lower stratospheric vortex; except during this period at the higher levels, diabatic cooling in 2011 was close to, but often somewhat weaker than, climatological values, consistent with lower-than-average temperatures. While the reanalyses show small biases (about 0.5 K at 20 hPa and 0.3 K at 50 hPa), they all follow the day-to-day variations quite well, such that any of these would give a representative picture of the daily evolution of diabatic heating. CFSR often shows slightly less cooling than the other reanalyses and occasionally shows qualitatively different behavior for a few days to a week or so (*e.g.*, at 50 hPa in mid-February and late March 2009).

Figure 10.9 shows 2002 (the winter with the only major SSW on record in the SH, *e.g.*, *Shepherd et al.*, 2005, and references therein) and 2006 (a year with one of the deepest ozone holes on record, *e.g.*, *WMO*, 2018). In contrast to the NH, the diabatic heating rates in the SH are usually near the climatological values. In 2006, diabatic cooling at 20 hPa was slightly

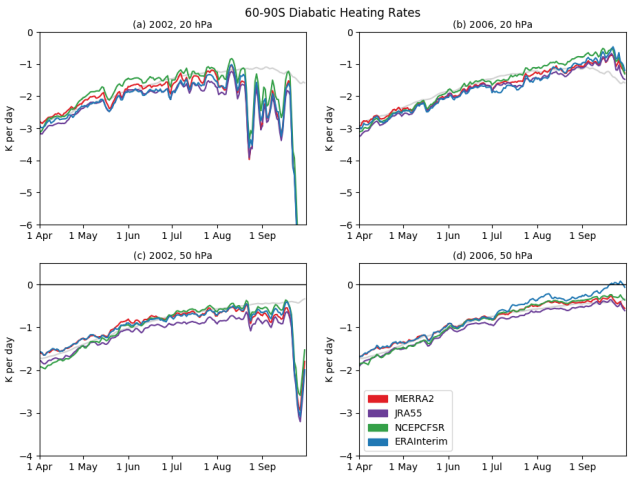

Figure 10.9: *As in* ***Figure 10.8*** *but for the south polar cap average in 2002 (left panels) and 2006 (right panels).*

weaker than the climatology at the end of the season, consistent with lower-than-usual temperatures (closer to radiative equilibrium) at that time. The contrast with 2002 is dramatic – the SSW occurred in late September, when there was an abrupt increase in diabatic cooling at all levels examined. Unusually large wave activity was seen for about a month prior to the SSW (*e.g.*, *Krüger et al.*, 2005; *Newman and Nash*, 2005), leading to modest temperature increases, consistent with the strengthening of the diabatic cooling over that in the climatology during the month prior to the SSW. During both winters, similar to the NH results, all of the reanalyses capture the day-to-day variability well, with some small overall biases apparent. The spread among the reanalyses is slightly larger than that in the NH, especially at 50 hPa, with a range among the reanalyses near 0.5 K at both levels shown. Again, CFSR often shows slightly less cooling than the other reanalyses, but ERA-Interim shows the weakest diabatic cooling among the reanalyses from July through September 2006. JRA-55 shows stronger diabatic cooling than the other reanalyses at 50 hPa during most of both winters shown.

This evaluation of diabatic heating rates in the lower stratosphere indicates generally good agreement among the reanalyses and that each of the reanalyses compared can reasonably be used in studies of polar processing, with some limitations; in particular, the weaker seasonal cycle in JRA-55 than in any of the other reanalyses suggests some caution in using heating rates from that reanalysis in the middle stratosphere and higher.

10.6 Concordiasi superpressure balloon comparisons

In this section we review the work presented by *Hoffmann et al.* (2017a) in the S-RIP special issue on the superpressure balloon measurements made during the Antarctic Concordiasi campaign (*Rabier et al.*, 2010) in September 2010 to January 2011 to evaluate meteorological analyses and reanalyses. The study covers the ERA-Interim, MERRA, MERRA-2, and NCEP-NCAR R1 reanalyses, as well as the higher-resolution ECMWF operational analysis (OA, see *Section 10.1* for details). The comparison was performed with respect to long-duration observations from 19 superpressure balloon flights, with most of the balloon measurements conducted at altitudes of 17 - 18.5 km and latitudes of 60° - 85° S. The five meteorological data sets considered by *Hoffmann et al.* (2017a) differ significantly in spatial and temporal resolution, with the truncation of the models limiting the capability of representing small-scale fluctuations.

10.6.1 Temperatures and winds

Hoffmann et al. (2017a) (see their Table 3) found positive temperature biases in the range of 0.4 K to 2.1 K, zonal wind biases in the range of – 0.3 m s^{-1} to 0.5 m s^{-1}, and meridional wind biases below 0.1 m s^{-1} for all data sets compared to the balloon measurements. The largest biases and standard deviations were typically found for NCEP-NCAR R1, which may be attributed to the fact that this data set is independent, whereas the Concordiasi balloon observations were assimilated into the other reanalyses. However, significant differences between ECMWF (OA and Interim), MERRA, and MERRA-2 are manifest, despite the balloon data being assimilated into each of these products. Hoffmann et al. suggest that this is a result of inter-model dependencies, such as the relative weighting of observations, types of forecast models, and the particular assimilation procedures of each model.

The superpressure balloon observations are a valuable source of data with which to study small-scale fluctuations in the atmosphere that are mostly attributed to gravity waves, as demonstrated in several studies (*e.g.*, *Jewtoukoff et al.*, 2015; *Vincent and Hertzog*, 2014; *Plougonven et al.*, 2013; *Hertzog et al.*, 2008, 2012). Statistics of high-pass filtered balloon data were used by *Hoffmann et al.* (2017a) (see their Table 4) to assess the representation of small-scale fluctuations in the meteorological analyses. ECMWF OA reproduces about 60 % of the standard deviations of the temperature and wind fluctuations of the balloons, and the temperature results are consistent with the recent studies of *Jewtoukoff et al.* (2015) and *Hoffmann et al.* (2017b). Therefore, the ECMWF operational model resolves the atmospheric gravity wave spectrum with higher fidelity than the reanalyses since ERA-Interim, MERRA, and MERRA-2 reproduce only about 30 % of the standard deviations of the temperature and wind fluctuations seen in the high-pass filtered balloon data, and NCEP-NCAR R1 reproduces only about 15 % for temperature and 30 % for the winds. These results are correlated with the spatial resolutions of the analyses, which determine the ability of the forecast models to simulate realistic gravity wave patterns.

Large-scale biases and standard deviations of temperatures and horizontal winds at different latitudes were averaged over the entire time period of the campaign by *Hoffmann et al.* (2017a) and are reproduced in **Figure 10.10**. An increasing temperature bias from middle to high latitudes is seen in all reanalyses. NCEP-NCAR R1 shows the largest temperature warm bias (3.1 K) at high latitudes of 80 - 85 ° S, but a bias is also present in reanalyses that assimilated the Concordiasi balloon observations, although to a lesser extent: MERRA (1.4 K), MERRA-2 (1.3 K), ERA-Interim (1.1 K), and ECMWF OA (0.5 K). Southern Hemisphere winter pole temperature biases were also reported for earlier winters in other studies (*Boccara et al.*, 2008; *Parrondo et al.*, 2007; *Gobiet et al.*, 2005), and the Concordiasi study indicates that such biases were still present in 2010/2011. *Gobiet et al.* (2005) speculate that the assimilation of microwave radiances from satellite measurements into ECMWF analyses may be a reason for the temperature bias. The NCEP-NCAR R1 warm bias found by *Hoffmann et al.* (2017a) is comparable to those found in earlier studies. Other analyses have smaller biases, which *Hoffmann et al.* attribute to the assimilation of the Concordiasi data.

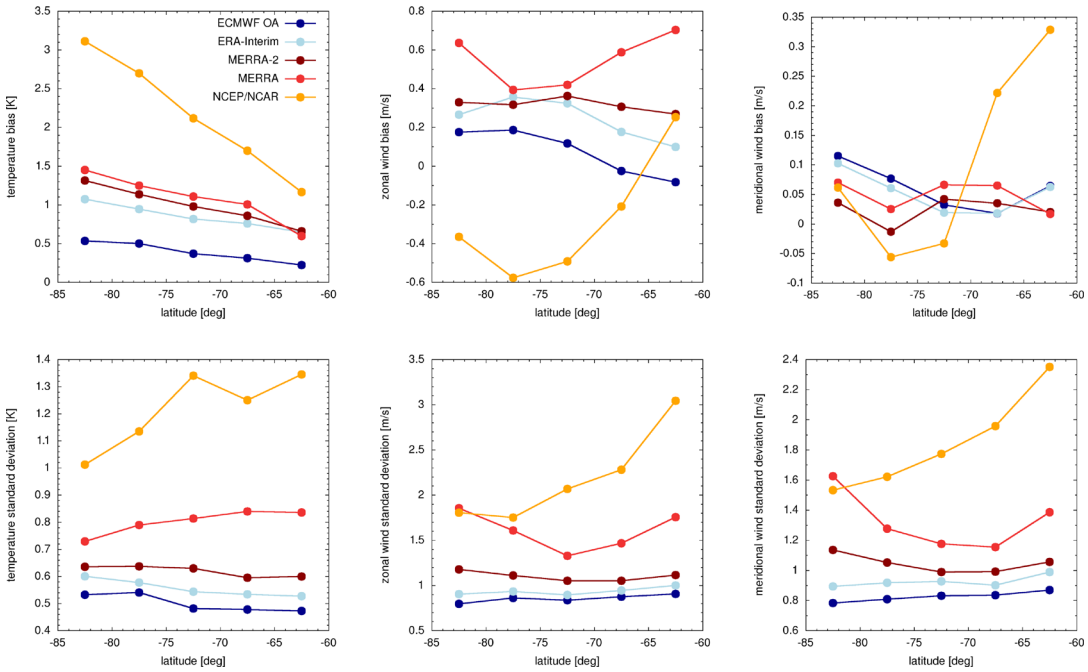

Figure 10.10: *Bias and standard deviations of temperature and horizontal winds of meteorological analyses minus Concordiasi balloon data at different latitudes. From Hoffmann et al. (2017a).*

Moreover, the wind biases as well as temperature and wind standard deviations shown in **Figure 10.10** are also generally largest for NCEP-NCAR R1, whereas they are smaller and more similar to each other for both ECMWF data sets and MERRA-2.

10.6.2 Trajectories

Hoffmann et al. (2017a) used the balloon tracking observations to evaluate reanalysis-driven trajectory calculations with the Lagrangian transport model MPTRAC, where the vertical motions of simulated trajectories were nudged to pressure measurements of the balloons. Absolute horizontal transport deviations (AHTDs) and relative horizontal transport deviations (RHTDs) are commonly used to compare trajectory calculations with observations or to evaluate results for different model configurations (*Stohl*, 1998; *Stohl et al.*, 1995; *Rolph and Draxler,* 1990; *Kuo et al.*, 1985). AHTD at a particular time is the average Euclidean distance between the observed balloon and modeled air parcel positions projected to the Earth's surface. RHTDs are calculated by dividing the AHTD of individual air parcels by the length of the corresponding reference trajectory. Of course, the accuracy of calculated trajectories depends strongly on the fidelity of the particle dispersion model being used as well as the details of its configuration (*e.g.*, the time step, which is 30 seconds in these runs, consistent with the sampling rate of the balloon data). Our purpose here, however, is not to evaluate the trajectories themselves, but rather to compare the air motions calculated using the same model setup but different reanalyses to drive the model with long-duration balloon observations as a tool for assessing transport in the reanalyses.

Figure 10.11 shows transport deviations for the different meteorological data sets calculated from over a hundred samples of 15-day trajectories, and at the end of this time the AHTDs are in the range 1400 km to 2200 km. Estimates of the growth rates of the AHTDs are typically within 60 km day^{-1} to 170 km day^{-1}. The RHTDs are in the range 4 - 12 % after 2 days, but converge to a smaller range of 6 - 9 % after 15 days. Although the transport deviations grow rather steadily, the relative differences between the data sets tend to get smaller over time. The largest transport deviations and growth rates were found for NCEP-NCAR R1, which did not assimilate the Concordiasi balloon observations. The results agree well with those reported by *Boccara et al.* (2008) for the Vorcore campaign

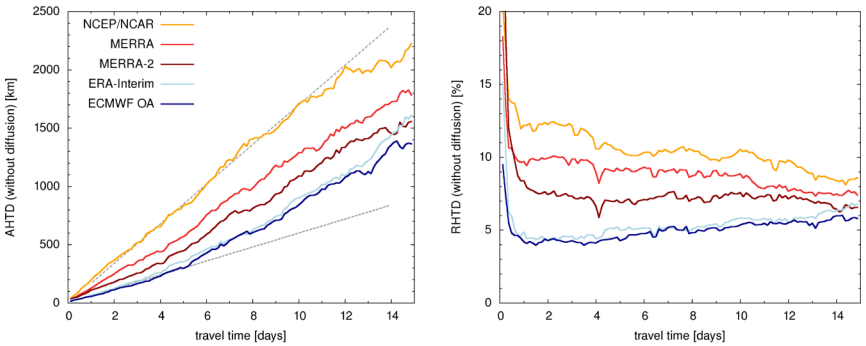

Figure 10.11: *Absolute (left, AHTD) and relative (right, RHTD) horizontal transport deviations of simulated and observed balloon trajectories for different meteorological analyses. Dotted grey lines represent AHTD growth rates of 60 km day^{-1} and 170 km day^{-1}. From Hoffmann et al. (2017a).*

in 2005, where mean spherical distances of about 1650 km (with an interquartile range of 800 - 3600 km) for ECMWF analyses and 2350 km (1400 - 3800 km) for NCEP-NCAR R1 data were found for 15-day trajectories.

10.7 PSC thermodynamic-consistency diagnostics

In this section we review the work presented by *Lambert and Santee* (2018) in the S-RIP special issue on comparisons of reanalysis temperatures with COSMIC GNSS-RO temperatures and with independent absolute temperature references derived from theoretical considerations of PSC formation defined by the equilibrium thermodynamics of STS and water-ice clouds. The PSC thermodynamic-consistency diagnostics rely on the near-simultaneous and colocated measurements of nitric acid, water vapour and cloud phases provided from the long-term precise formation flying of the CALIOP and Aura MLS instruments within the afternoon "A-Train" satellite constellation. The initial A-Train configuration of the CALIPSO and Aura spacecraft from April 2006 to April 2008 resulted in an across-track orbit offset of ~200 km, with the MLS tangent point leading the CALIOP nadir view by about 7.5 minutes. From April 2008 until September 2018, Aura and CALIPSO were operated to maintain positioning within tightly constrained control boxes, such that the MLS tangent point and the CALIOP nadir view were colocated to better than about 10 - 20 km and about 30 seconds.

The analysis refines and extends the methodology used originally by *Lambert et al.* (2012) to investigate the temperature existence regimes of different types of PSCs. The CALIOP lidar PSC classification is used to identify the presence of STS and ice PSCs in the MLS geometric FOV at the along-track resolution (165 km × 2.16 km). MLS provides ambient gas-phase H_2O and HNO_3 volume mixing ratios, which are required to calculate the theoretical equilibrium temperature dependence of the STS (T_{eq}) and ice (T_{ice}) PSCs, based on the expressions of *Carslaw et al.* (1995) and *Murphy and Koop* (2005), respectively. Observed and calculated temperature distributions are compared for both the uptake of HNO_3 in STS and the ice frost point for each reanalysis data set and for MLS temperature. Viewing scenes having a distinct dominant PSC classification in a sample volume similar in size to the MLS gas-species resolution are selected, with the requirement that 75 % or more of the CALIOP pixels in the MLS geometric FOV have the same PSC classification. Scenes satisfying this requirement for CALIOP STS detections are denoted as LIQ; those for CALIOP ice detections are denoted as ICE. Reanalysis temperature biases are then quantified relative to the respective LIQ and ICE reference temperatures. The calculated standard deviations of the temperature differences are used to estimate the measurement precisions. *Lambert et al.* (2012) show that the resulting root mean square (RMS) temperature uncertainties for the LIQ and ICE references are smaller than those derived for the MLS retrieved temperatures and comparable to the measurement capabilities of the GNSS RO technique (bias < 0.2 K, precision > 0.7 K) in the lower stratosphere. The estimated measurement precisions for the STS equilibrium and ice frost points are 0.4 K and 0.3 K, respectively, in the 68 - 21 hPa pressure range, with the corresponding estimated measurement accuracies in the range of 0.7 - 1.6 K for STS and 0.4 - 0.7 K for ice.

The approach for PSC thermodynamic-consistency diagnostics is summarized as follows:

- Identify LIQ and ICE PSCs using the CALIOP lidar measurements.
- Accumulate the CALIOP PSC types (LIQ and ICE) at the MLS along-track resolution (165 km × 2.16 km), ensuring that the same PSC type is detected in at least 75 % of the MLS FOV.
- Calculate the theoretical temperature dependence of STS (T_{eq}) and ice (T_{ice}) PSCs under equilibrium conditions using the spatially and temporally colocated MLS gas-phase HNO_3 and H_2O measurements.
- Compare reanalyses and MLS data with (a) calculated and observed HNO_3 uptake in STS and (b) ice temperature distribution *vs*. the frost point temperature.
- Create LIQ and ICE temperature distributions for each reanalysis (all interpolated to the MLS measurement times and locations) as well as MLS.
- Calculate the median and mean temperature deviations from T_{eq} and T_{ice} and their standard deviations for LIQ and ICE classifications, respectively.

In the Antarctic, six PSC seasons from 20 May (d140) to 18 August (d230) were investigated from 2008 to 2013 for latitudes poleward of 60 °S in the lower stratosphere (100 - 10 hPa). In the Arctic, five PSC seasons from 2 December (d336) to 31 March (d090) were investigated from 2008/2009 to 2012/2013 for latitudes poleward of 60 °N.

10.7.1 Mountain waves

Small-scale temperature fluctuations are not captured accurately by the reanalyses because of their limited spatial resolution (*e.g.*, *Hoffmann et al.*, 2017a). An orographic gravity wave case study over the Palmer Peninsula was used by *Lambert and Santee* (2018) to show that, at 50 hPa, the temperature amplitudes resolved by the reanalyses can vary by up to a factor of two in extreme circumstances. The differences in the wave amplitudes are well correlated with the spatial resolutions of the reanalyses. It is thus important to identify regions characterized by small-scale temperature fluctuations and remove them from further consideration in this study. The high vertical resolution of the COSMIC temperatures allows examination of the spectrum of temperature variance over the height region of PSC activity. An estimate of the potential energy density of the wave disturbance, E_p,

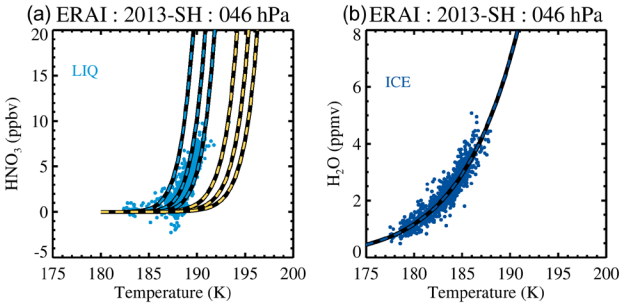

Figure 10.12: *Scatterplots of coincident MLS HNO_3 and H_2O vs ERA-Interim temperature for PSCs classified by CALIOP on the 46-hPa pressure level in the 2013 Antarctic winter. (a) MLS HNO_3 vs ERA-Interim temperature for liquid (LIQ) PSCs (light blue dots). Note that measurement noise can lead the MLS data processing algorithms to retrieve negative HNO_3 mixing ratios when abundances are low (e.g., under highly denitrified conditions). Though unphysical, such negative values must be retained to avoid introducing a positive bias into any averages calculated from the measurements. The theoretical equilibrium uptake of HNO_3 by STS is shown for representative ambient H_2O values by the light blue-black dashed lines. The yellow-black dashed lines show the corresponding NAT equilibrium curves. (b) MLS H_2O vs ERA-Interim temperature for ice (ICE) PSCs (dark blue dots). The dark blue-black dashed lines indicate the theoretical equilibrium for the frost point temperatures. Adapted from Lambert and Santee (2018).*

provides an effective means of applying low-pass filtering to reveal the large-scale temperature structure. A low-pass filter was applied by excluding profile matches with COSMIC temperature variances $>0.5\,K^2$ ($E_p > 1.5\,J\,kg^{-1}$). COSMIC temperatures were restricted to below 200 K to focus on regions of potential PSC existence.

10.7.2 Thermodynamic equilibrium

PSC theoretical existence temperatures, calculated based on equilibrium thermodynamics, are dependent on the ambient partial pressures of H_2O in the case of the ice frost point, T_{ice} (*Murphy and Koop*, 2005), and also of HNO_3 for STS (*Carslaw et al.*, 1995).

Figure 10.12 shows scatterplots against ERA-Interim temperature of coincident MLS HNO_3 (left panel) and H_2O (right panel) for Antarctic PSCs classified by CALIOP at 46 hPa in 2013, along with the theoretical HNO_3 gas-phase uptake curves for STS and NAT. It is clear that LIQ PSCs are closely associated with the equilibrium STS curve. Similarly, the scatter of gas-phase H_2O is closely associated with the frost point temperature in the presence of ICE PSCs.

In **Figure 10.13** we show the variation in MLS gas-phase HNO_3 with ERA-Interim temperature for CALIOP PSC classifications at 31 hPa for one Southern Hemisphere winter. MLS HNO_3 data are separated into corresponding CALIOP PSC categories, allowing comparison of

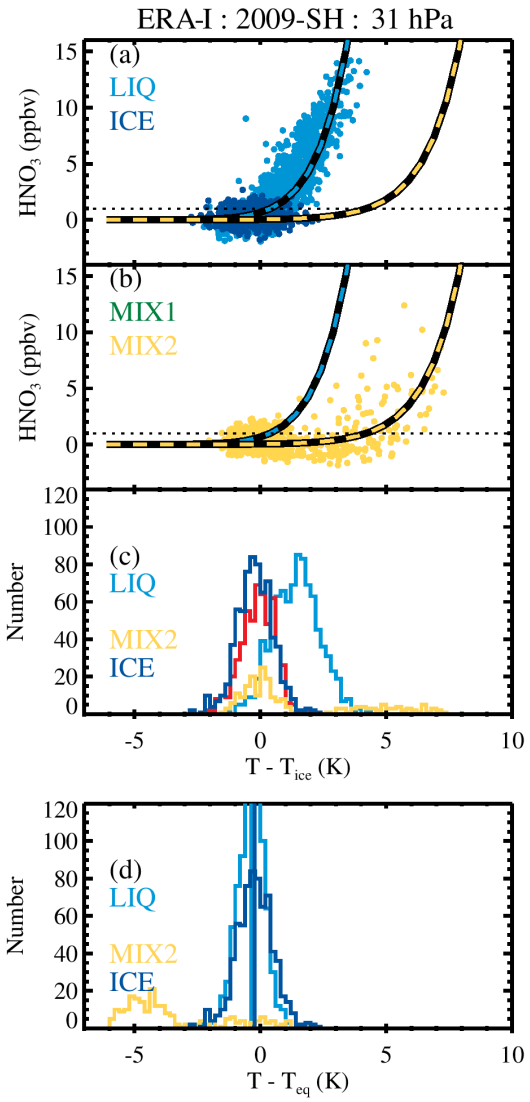

Figure 10.13: *Composite statistics for d140 - d230 (the bulk of the PSC existence period) of a representative year (2009) of the MLS gas-phase HNO_3 variation with ERA-Interim temperature corresponding to CALIOP PSC classifications at 31 hPa, with the added constraint that at least 75 % of the MLS FOV is filled with the same classification. (a) Scatter plot of HNO_3 vs temperature deviation from the frost point ($T - T_{ice}$) for PSCs classified as LIQ (light blue) and ICE (dark blue). (b) As for (a), but for MIX2 PSCs (yellow). Equilibrium STS (light blue-black dashed) and NAT (yellow-black dashed) curves show the theoretical uptake of HNO_3. Because of non-equilibrium effects, which cause larger temperature scatter, the CALIOP NAT classifications MIX1 and MIX2 are not used in this analysis, but this panel is shown to indicate the good discrimination between the solid and liquid uptake curve branches. Note that there are no MIX1 PSC detections for the case shown here. (c) Temperature histograms for HNO_3 mixing ratios > 1 ppbv for the LIQ PSC type; data in the ICE classification are not subject to this constraint. The red histogram indicates the distribution of LIQ PSCs that have HNO_3 below the 1 ppbv threshold. (d) Temperatures transformed according to the STS equilibrium curve for the LIQ classification and NAT equilibrium curve for the MIX2 classification; the ICE classification remains the same as in (c) for comparison. From Lambert and Santee (2018).*

observed and modeled uptake of HNO_3 in different types of PSCs. The scatter of MLS HNO_3 against the temperature deviation from the frost point (calculated using MLS H_2O) is shown in **Figure 10.13a** for LIQ and ICE PSCs. Uptake of gas-phase HNO_3 in the presence of liquid-phase LIQ PSCs follows the STS equilibrium curve. In contrast, HNO_3 abundances are very low (typically ~2 ppbv or less) in the presence of ICE PSCs. In **Figure 10.13b**, uptake of gas-phase HNO_3 in clouds identified by CALIOP as being solid NAT MIX2 shows significant non-equilibrium variation, lying between the STS and NAT equilibrium curves. For this reason, we do not use either of the CALIOP NAT classifications, MIX1 or MIX2, in this analysis. Histograms of the temperature distributions of **Figure 10.13a** and **b** are shown in **Figure 10.13c**. Light blue (LIQ) and yellow (MIX2) represent distributions for HNO_3 mixing ratios > 1 ppbv, whereas red (LIQ) represents distributions for HNO_3 mixing ratios < 1 ppbv. The tails of the temperature distributions for LIQ (light blue and red histograms) do not reach temperatures lower than those in the distribution for ICE (dark blue), and no peaks indicating the existence of PSCs at a frost-point depression near $T_{ice} - 3$ K are observed. In **Figure 10.13d**, the temperatures are transformed with respect to the corresponding equilibrium curves: STS for the LIQ classification and NAT for the MIX2 classification, with the ICE classification remaining the same as in **Figure 10.13c** for comparison. As a result, the LIQ histogram narrows and is shifted (*Lambert et al.*, 2012). This is just an illustrative case; since there are seven temperature data sources, collected over six years on six pressure levels, the total number of histograms is 252 for each hemisphere. Statistics of the temperature difference distributions were generated for each reanalysis data set for each year and pressure level.

10.7.3 Temperature difference profiles

Intercomparisons of the reanalysis temperature statistics are displayed in **Figure 10.14** as vertical profiles over 100 - 10 hPa averaged for the Antarctic winters 2008 - 2013. For ICE PSCs, median temperature bias values fall in the range −1.0 to +0.1 K with standard deviations of ~0.7 K, both of which are larger than those of their LIQ PSC counterparts. Median values for the ICE reference are more uniform throughout the profile than those for the LIQ reference, which become increasingly negative with decreasing pressure. Median values for the LIQ reference are consistently biased low relative to the corresponding values for ICE by ~0.5 K, although standard deviations for the LIQ reference are smaller than those for ICE. The largest bias is found for MLS. In addition to differences with respect to the LIQ and ICE reference points, we also show comparisons with the COSMIC temperatures (90° - 60° S, mean variance < 0.5 K^2, and COSMIC temperatures < 200 K). Biases for the COSMIC reference are generally smaller than those for either LIQ or ICE. MERRA does not assimilate COSMIC GNSS-RO data, and it tends to exhibit the largest bias with respect to the COSMIC reference. The comparisons with thermodynamic equilibrium temperatures are completely independent of any data assimilated by the reanalyses. Comparing the standard deviation of the temperature differences (SD), it is apparent that, except at 100 hPa, the (PSC temperature − Reanalysis) SDs (**Figure 10.14d, e**) are smaller than the (COSMIC − Reanalysis) SDs (**Figure**

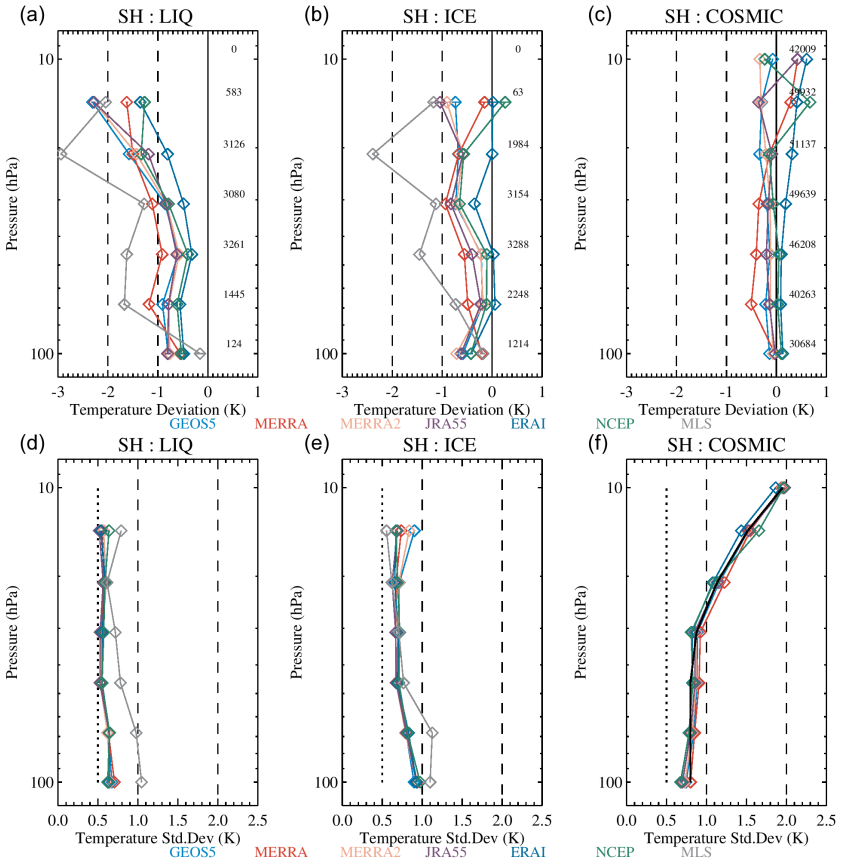

Figure 10.14: *Vertical profiles of median temperature deviations from T_{eq} for (a) LIQ PSCs and (b) ICE PSCs for the temperature distributions accumulated over Antarctic PSC seasons (20 May to 18 August, d140 - d230) from 2008 to 2013. Lines for the different reanalyses (MERRA, MERRA-2, JRA-55, ERA-Interim, and NCEP (CFSR/CFSv2)), GEOS-5.9.1, and MLS are colour coded (see legend); the numerical values on the right-hand side of the panel indicate the total number of observations in the distribution at the corresponding pressure level. Error bars for the median temperatures are displayed but are generally smaller than the symbol sizes. (c) as for (a,b) except for deviations with respect to COSMIC GNSS-RO. (d,e,f) The standard deviations of the corresponding temperature distributions shown in (a,b,c). Dotted lines indicate a standard deviation of 0.5 K. From Lambert and Santee (2018).*

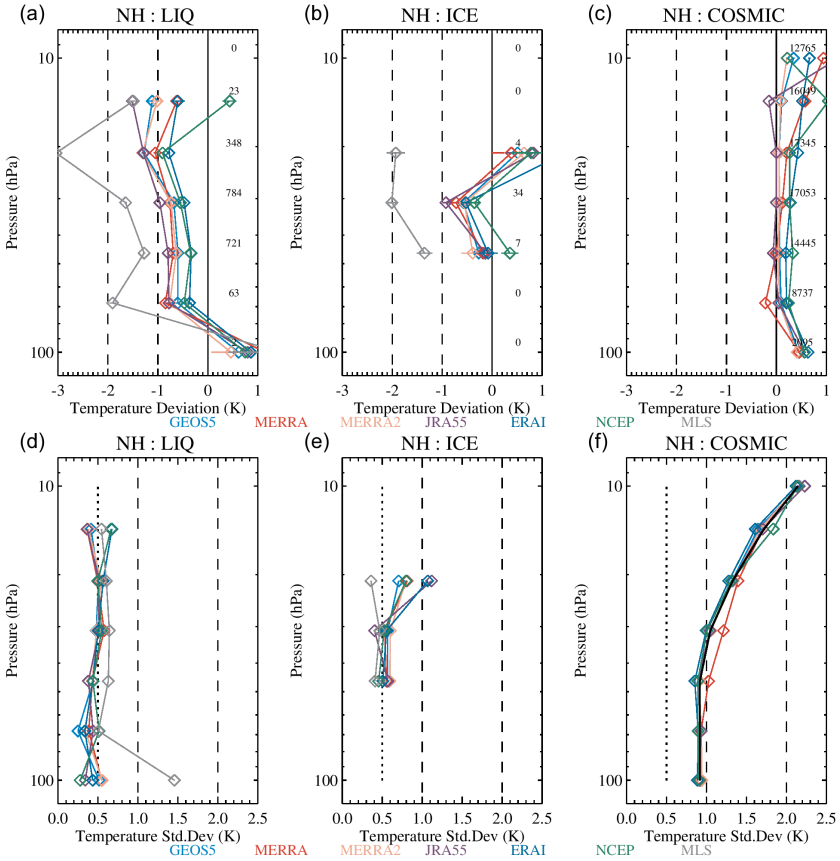

Figure 10.15: *As for* ***Figure 10.14****, but for Arctic PSC seasons (2 December to 31 March, d336 - d090) from 2008/2009 to 2012/2013. From Lambert and Santee (2018).*

(a) COSMIC GNSS-RO temperatures and (b) absolute temperature references obtained from the equilibrium properties of certain types of PSCs. A concise summary of the ranges of the mean temperature biases of the reanalyses relative to the LIQ (−1.6 K to −0.3 K) and ICE (−0.9 K to +0.1 K) equilibrium references as well as COSMIC (−0.5 K to +0.2 K) is given in **Figure 10.16**, which depicts bias ranges for the pressure domain 68 - 21 hPa. For all reference points, the coldest reanalysis biases tend to be found in the Antarctic and the warmest in the Arctic. The fact that GNSS-RO data are not assimilated in MERRA is evident in its higher biases with respect to COSMIC temperatures in the Antarctic (where there is a relative paucity of conventional measurements) compared to other reanalyses.

Reanalysis temperatures are found to be lower than the absolute reference points by 0.3 K to 1 K for LIQ and 0 K to 1 K for ICE at the altitudes of peak PSC occurrence (68 - 32 hPa). For LIQ, the negative biases are larger above 32 hPa than below that level. Median deviations for LIQ are consistently biased lower than those for ICE by ~ 0.5 K. At 46 hPa, median differences of the reanalyses with respect to the reference temperatures all depart from zero, and their scatter falls within the range of about 0.6 K for LIQ and 0.5 K for ICE. Although the biases are larger for LIQ, their standard deviations (~ 0.6 K) are smaller than those for ICE (~ 0.7 K). To put these comparisons with LIQ and ICE reference temperatures into context, temperature measurements from long-duration balloon flights have typical nighttime biases of 0.5 K, precisions of 0.4 K (*Pommereau et al.*, 2002), and measured standard deviations of 1.0 K to 1.3 K for temperature differences with respect to ECMWF operational analyses (*Knudsen et al.*, 2002).

Finally, the polar temperatures from recent full-input reanalyses are in much better agreement than were the reanalyses from previous decades. As a consequence, strong justification is warranted in modeling studies that seek to match the simulated chlorine activation and/or ozone loss by making arbitrary systematic adjustments to the reanalysis temperatures of 1 - 2 K or higher.

10.14f). Therefore, in terms of random errors, the PSC-based temperature references perform better than the COSMIC temperature measurements. **Figure 10.15** shows the corresponding observations for the Arctic, where limited ice cloud production and hence few data points preclude a robust conclusion for ICE PSCs. Again, MLS shows the largest bias.

Although we have not directly matched A-Train locations with COSMIC occultations, we can estimate the differences between the COSMIC temperatures and the thermodynamic reference temperatures by elimination of the reanalysis temperature biases, $\overline{\Delta T_{re}}$. The temperature bias difference profiles for all reanalyses are tightly clustered (not shown), especially over the pressure range 68 - 21 hPa, justifying the assumption that $\overline{\Delta T_{re}}$ can be eliminated. The biases for $(\overline{\Delta T_{LIQ}} - \overline{\Delta T_{COSMIC}})$ have similar magnitude and profile shape over the pressure range 68 - 21 hPa in both hemispheres, with $\overline{\Delta T_{LIQ}}$ being smaller than $\overline{\Delta T_{COSMIC}}$ by about 0.5 K to 1.0 K. Likewise, the biases for $\overline{\Delta T_{ICE}}$ are about 0 K to 0.5 K smaller than $\overline{\Delta T_{COSMIC}}$ in the Antarctic, but there are too few data points to make a meaningful comparison in the Arctic.

10.7.4 Summary of the temperature differences

The accuracy and precision of several contemporary reanalysis data sets were evaluated through comparisons with

10.8 Chemical modeling diagnostics

In this section, we compare simulated fields of nitric acid, water vapour, several chlorine species, nitrous oxide, and

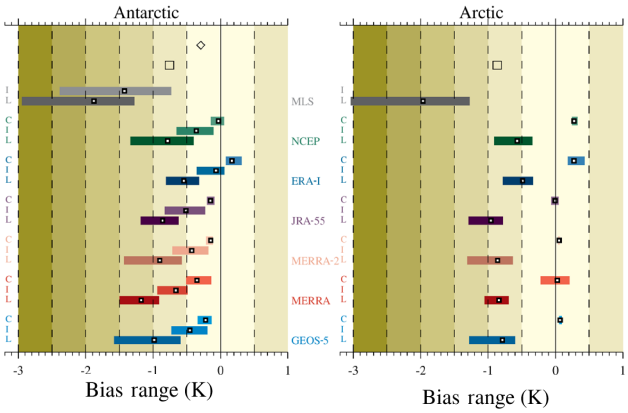

Figure 10.16: *Temperature bias ranges of the reanalyses (MERRA, MERRA-2, JRA-55, ERA-Interim, and NCEP (CFSR/CFSv2)), and MLS, relative to the LIQ (labelled "L" on the y-axis) and ICE ("I" on the y-axis) equilibrium references, and COSMIC ("C" on the y-axis), for Antarctic (left) and Arctic (right) winters 2008-2013, poleward of 60°, and for pressure levels from 68 hPa to 21 hPa. The bias ranges are obtained from the extrema of the yearly mean bias values over 68-21 hPa weighted by the yearly standard deviations. The horizontal coloured bars indicate the ranges of the minimum to maximum bias for MLS and each of the reanalyses as indicated in the legend. White squares with black border indicate the mean bias over 68-21 hPa. Open square (diamond) symbols indicate the mean values of* $\overline{\Delta T_{LIQ}} - \overline{\Delta T_{COSMIC}}$ *(*$\overline{\Delta T_{ICE}} - \overline{\Delta T_{COSMIC}}$*) over 68-21 hPa. There are insufficient statistics for a reliable comparison with the ICE reference in the Arctic. Note that MLS has not been compared directly to COSMIC. Background shading indicates 0.5 K increments in the bias range. From Lambert and Santee (2018).*

ozone with those observed by Aura MLS and Envisat MIPAS. As noted in *Section 10.1*, chemical model simulations provide a means of assessing the net effects of multiple (in some cases competing) spatially and temporally varying differences between reanalyses. Thus results from a chemical model driven by different reanalyses may provide further insights into which reanalyses are well suited for polar process studies beyond those obtained through simpler, more direct diagnostics.

It must be borne in mind that some of the differences seen between observed and simulated fields may arise through inherent shortcomings in the formulation or setup of the chemical model itself and not through the meteorological fields being used to drive it. Ideally the same set of simulations should be performed by a range of chemical models and the full suite of runs examined to elucidate how underlying differences in their chemistry and physics manifest when different models are forced by different reanalyses. It was not practical to carry out such an extensive investigation in the context of this project, however, and only a single model is applied here. Nevertheless, because our purpose is intercomparison of the reanalyses (not model evaluation or scientific study of specific polar processing events), using multiple reanalyses to drive the same (albeit imperfect) chemical model over the same time period can illuminate how various differences between reanalyses combine to affect simulated trace gas fields.

We have chosen to use the chemistry transport model developed for the Belgian Assimilation System for Chemical ObsErvations (BASCOE). The BASCOE CTM has been validated against several satellite data sets and shown to successfully reproduce observed stratospheric composition in the polar regions (*Huijnen et al.*, 2016; *Lindenmaier et al.*, 2011; *Daerden et al.*, 2007); its performance is thus adequate for this reanalysis intercomparison study. Experiments were performed for five reanalyses: MERRA, MERRA-2, JRA-55, ERA-Interim, and CFSR/CFSv2.

10.8.1 Details of the BASCOE system and experimental setup

We use here the BASCOE CTM (version 6.2). The transport module and its setup are described in *Chapter 5* for the mean age of air simulations (for full details see *Chabrillat et al.*, 2018). It is based on the Flux-Form Semi-Lagrangian (FFSL) scheme (*Lin and Rood*, 1996). We use a low-resolution latitude-longitude grid (2° × 2.5°) that is common to all five experiments, but the vertical grids differ between experiments because each of them retains the native sigma-pressure vertical grid of its input reanalysis. All five reanalyses have been expressed as spherical harmonics and identically truncated to wavelength number 47 before derivation of the wind fields on the common horizontal grid. On the basis of mass conservation, vertical velocities (expressed as $\omega = dp/dt$) are derived afterwards within the FFSL module. Hence, in contrast to diabatic transport models that use the isentropic coordinate, such as SLIMCAT (*Chipperfield*, 2006) or CLaMS (*Grooß et al.*, 2011), this transport scheme does not compute heating rates, nor does it read them from the input reanalysis. Diabatic transport models have distinct advantages in modeling the stratosphere, because the use of isentropic levels reduces spurious vertical mixing by providing a true separation between horizontal and vertical motion (*Chipperfield et al.*, 1997). These issues were especially important with the previous generation of reanalyses because they often suffered from spurious surface pressure increments caused by data assimilation (*Pawson et al.*, 2007; *Meijer et al.*, 2004). In modern kinematic transport models, these issues are mitigated by pre-processing the input wind fields in order to correct for the small inconsistencies in the pressure tendency compared with the divergence fields (*Chabrillat et al.*, 2018; *Chipperfield*, 2006). While the simpler kinematic approach used by the BASCOE CTM may provide less realistic simulations of transport in and around the polar vortex (*Hoppe et al.*, 2016), it should ease the interpretation of results in the context of S-RIP because the temperature fields impact only the chemical and microphysical processes, while the surface pressure and wind fields determine entirely the transport processes. This approach also allows us to keep the sigma-pressure vertical grid of each input reanalysis, providing an intercomparison that takes their different vertical resolutions into account.

The photochemical scheme and parameterization of PSCs are described by *Huijnen et al.* (2016). The kinetic rates for heterogeneous chemistry are determined by the parameterization of *Fonteyn and Larsen* (1996), using classical expressions for the uptake coefficients on sulphate aerosols (*Hanson and Ravishankara*, 1994) and on PSCs (*Sander et al.*, 2000). The surface area density of stratospheric aerosols uses an aerosol number density climatology based on SAGE II observations (*Hitchman et al.*, 1994). Ice PSCs are presumed to exist at any grid point in the winter/spring polar regions where water vapour partial pressure exceeds the vapour pressure of water ice (*Murphy and Koop*, 2005). NAT PSCs are assumed when the HNO_3 partial pressure exceeds the vapour pressure of condensed HNO_3 at the surface of NAT PSC particles (*Hanson and Mauersberger*, 1988). The surface area density is set to $2\times10^{-6}cm^2cm^{-3}$ for ice PSCs and $2\times10^{-7}cm^2cm^{-3}$ for NAT PSCs. The sedimentation of PSC particles causes denitrification and dehydration. This process is approximated by an exponential decay of HNO_3 with a characteristic timescale of 20 days for grid points where NAT particles are supposed to exist, and an exponential decay of HNO_3 and H_2O with a characteristic timescale of 9 days for grid points where ice particles are supposed to exist (*Huijnen et al.*, 2016).

Experiments were performed for five reanalyses: MERRA, MERRA-2, JRA-55, ERA-Interim, and CFSR/CFSv2. The simulations span 13 months, from March 2009 to April 2010, to allow examination of chemical processing during one Antarctic and one Arctic winter. Each experiment starts on 1 March 2009 from the CTM simulation described in *Huijnen et al.* (2016), *i.e.*, a higher-resolution ($1.125°\times1.125°$) experiment that started one year earlier from a combined analysis of Envisat MIPAS and Aura MLS chemical observations (*Errera et al.*, 2008, 2016, 2019) and was driven by ECMWF analyses similar to ERA-Interim. We first investigate the period from May 2009 through October 2009 over the Antarctic (*Section 10.8.3*), followed by the period from November 2009 through April 2010 over the Arctic (*Section 10.8.4*). Hence, the five experiments were spun up with their respective input reanalyses during the two months of March and April 2009.

10.8.2 Analysis approach

The BASCOE model output is interpolated to the geolocations and times of the satellite observations to facilitate comparisons, and both observations and model results are vertically interpolated onto a potential temperature grid with 10 K spacing. Daily averages of observed and modeled quantities are calculated over two equivalent latitude ranges representing the inner vortex core (75°-90°) and outer vortex collar region (60°-75°) in both hemispheres. Temperature, PV, equivalent latitude, and other parameters at the MLS geolocations are obtained from the MLS DMP files; as described in *Section 10.2.1*, DMP files containing associated meteorological information at the MLS measurement locations are available for all five reanalyses considered here. Thus results from the model simulations driven by each of the five reanalyses are compared to corresponding MLS data analyzed using the respective set of DMPs. Since no DMP files are available for MIPAS data, they are analyzed in terms of geographic latitude. In the following two subsections, we examine two case studies: the 2009 Antarctic winter (May through October) and the 2009/2010 Arctic winter (November through April).

10.8.3 Case study: 2009 Antarctic winter

Model/measurement comparisons of the evolution of various chemical constituents are shown as a function of potential temperature (300 K to 800 K) for the inner and outer vortex in **Figures 10.17** and **10.18**, respectively. Observations (from MLS and MIPAS) are presented in the left-hand column, with corresponding model results in the middle column. The simulation driven by MERRA-2 is selected as a representative example. The right-hand column shows the reanalysis ranges, that is, the differences between the maximum and the minimum from the full set of five simulations, to illustrate the spread between the runs forced by the different reanalyses.

Overall, the model performs well in reproducing the observed stratospheric conditions inside the vortex during Antarctic winter, as has been demonstrated previously (*Huijnen et al.*, 2016). There are, however, areas where agreement is less satisfactory (**Figure 10.17**). Again, it should be emphasized that some of the model/measurement discrepancies may be attributable to fundamental deficiencies in the formulation of the model itself, independent of the reanalysis being used to drive it. The observations indicate higher peak HNO_3 abundances than in the model (forced by MERRA-2), so although sequestration of HNO_3 in PSCs is fairly well modeled, the simulated values start off from a lower maximum. Renitrification, or the redistribution of HNO_3 to lower altitudes via evaporation of sedimenting PSC particles, is visible in the observations but not represented in the model, likely because the simulated HNO_3 is permanently removed rather than being transported downward by sedimentation. The decline in HCl, indicative of chlorine activation, begins about two weeks earlier, progresses more rapidly, and is more complete in the model, such that nearly all of the available HCl in the Antarctic inner vortex is consumed by late May. This overestimation of HCl depletion at the beginning of Antarctic winter is in contrast to results from other chemical models, which typically see a delay in chlorine activation relative to observations (*e.g.*, *Wohltmann et al.*, 2017; *Grooß et al.*, 2018, and references therein). Because the deep vortex core is essentially in darkness until mid-winter, active chlorine remains tied up in forms other than ClO (primarily the ClO dimer, ClOOCl, as well as Cl_2

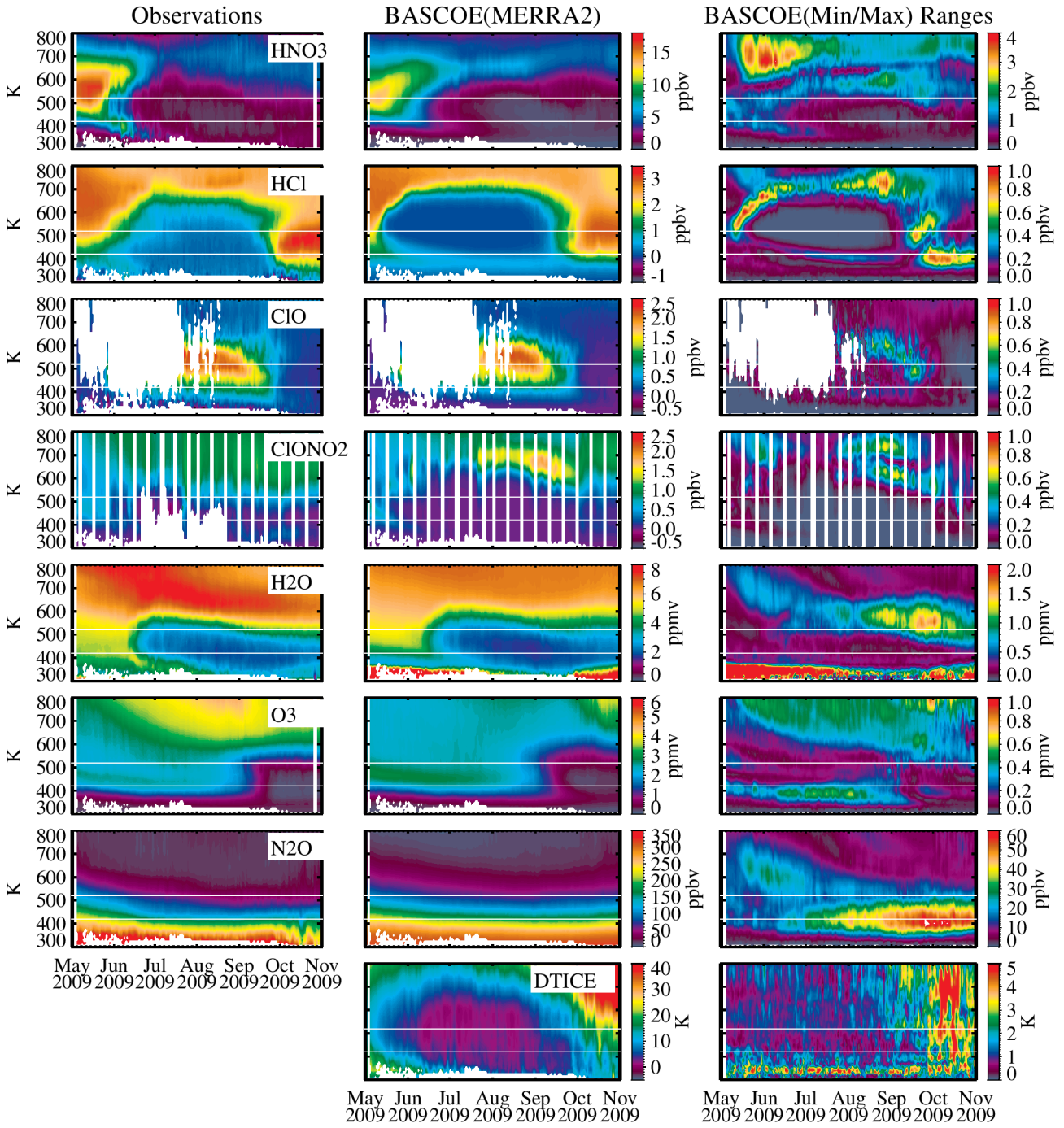

Figure 10.17: *Time/height cross sections of daily averages calculated over the region of the inner vortex (75 °- 90 °S) for the 2009 Antarctic winter (May through October). The first column shows measurements of HNO_3, HCl, ClO, $ClONO_2$, H_2O, O_3, and N_2O. The $ClONO_2$ data are from MIPAS, averaged over geographic latitude; all other species were measured by Aura MLS and averaged over equivalent latitude. The middle column shows corresponding results from the BASCOE simulation driven by MERRA-2. The third column shows the range (maximum – minimum) of model simulations driven by all five reanalyses. The bottom row of the middle column shows the time/height cross section of the MERRA-2 temperature deviation from the ice frost point calculated using MLS H_2O (DTICE), and the bottom row of the third column shows the range in reanalysis temperatures. Blank spaces arise for several reasons: (1) MIPAS is observing in another mode or data are otherwise missing, e.g., MIPAS measurements are affected by the presence of PSCs, (2) the potential temperature surface falls below the lowest recommended MLS retrieval pressure level, or (3) no measurements in sunlight are available in polar night. The latter situation pertains to the ClO panels; because active chlorine is converted to ClO in daylight, only data from the ascending portions of the orbit are used for ClO (see Section 10.2.1). Horizontal white lines mark the 420 K and 520 K potential temperature surfaces, which are examined in detail in subsequent figures.*

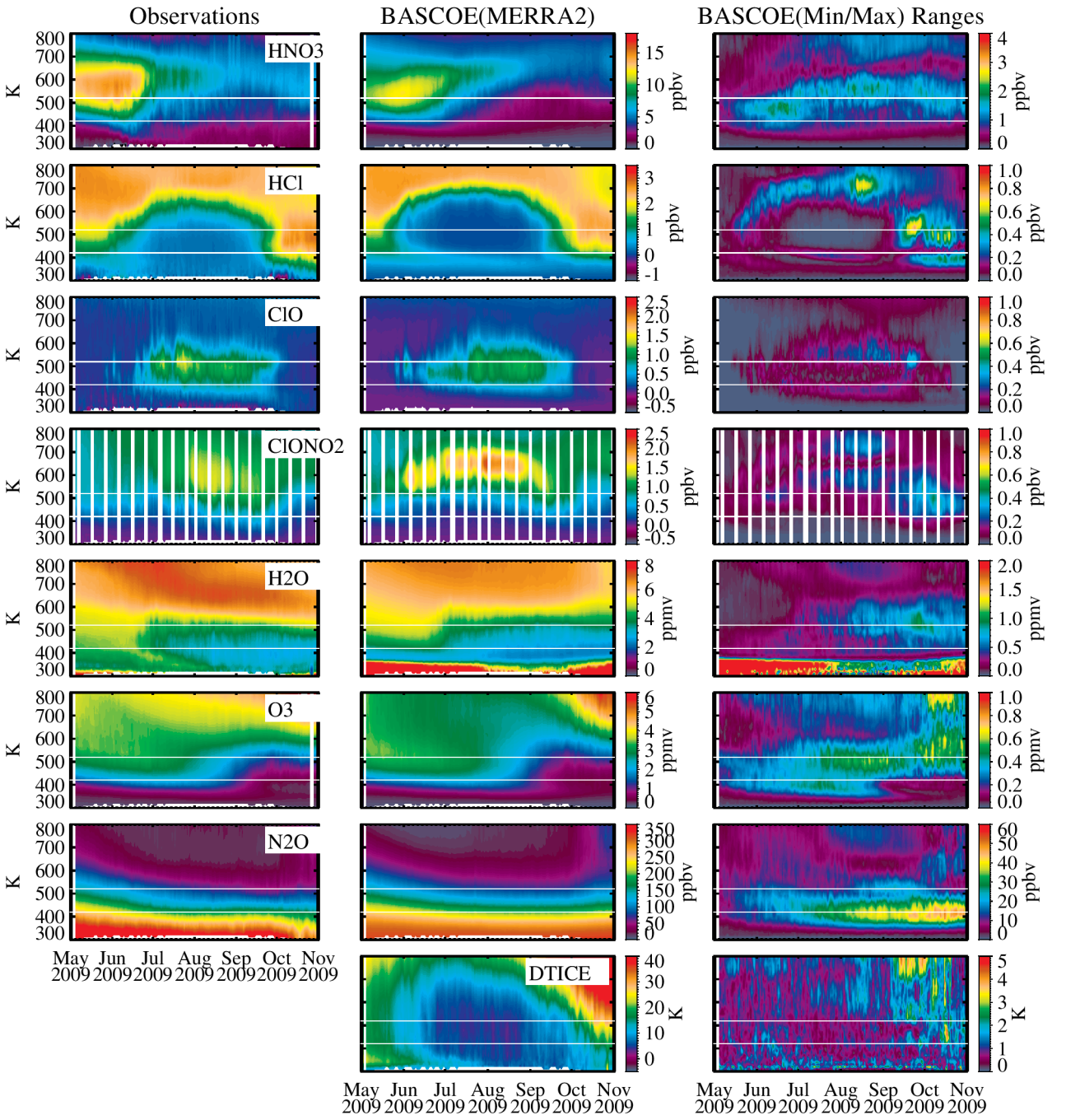

Figure 10.18: *As in* ***Figure 10.17*** *but for the region of the outer vortex (60 °- 75 °S).*

and HOCl); the resulting patchiness complicates interpretation of the ClO panel. Nevertheless, **Figure 10.17** shows that the modeled and measured morphology and magnitude match fairly closely. As was the case for HCl, modeled $ClONO_2$ depletion is spatially and temporally more extensive than observed. The model/measurement discrepancies in PSC processes and chlorine activation may be attributable to the relatively simple PSC parameterizations implemented in the BASCOE model (*e.g., Errera et al.*, 2019).

Examination of N_2O as a tracer of stratospheric air motions, along with H_2O and O_3 in the non-dehydrated, non-ozone-depleted regions where those species act as conserved tracers (*e.g.*, above 600 K), suggests that either confined diabatic descent inside the winter polar vortex is weaker in the simulations than indicated by the observations, or mixing across the vortex edge is greater. The latter possibility would be consistent with the findings of *Hoppe et al.* (2014), who reported that model simulations based on the FFSL transport scheme (see *Section 10.8.1*) tend to underestimate the strength of the transport barrier at the edge of the polar vortex compared to observations. On the other hand, some models nudged to reanalyses have been shown to have difficulty in accurately reproducing the strength of the diabatic descent inside the winter polar vortices (*e.g.*, *Froidevaux et al.*, 2019; *Khosrawi et al.*, 2018). Disentangling the contributions of advection and mixing in the model representation of

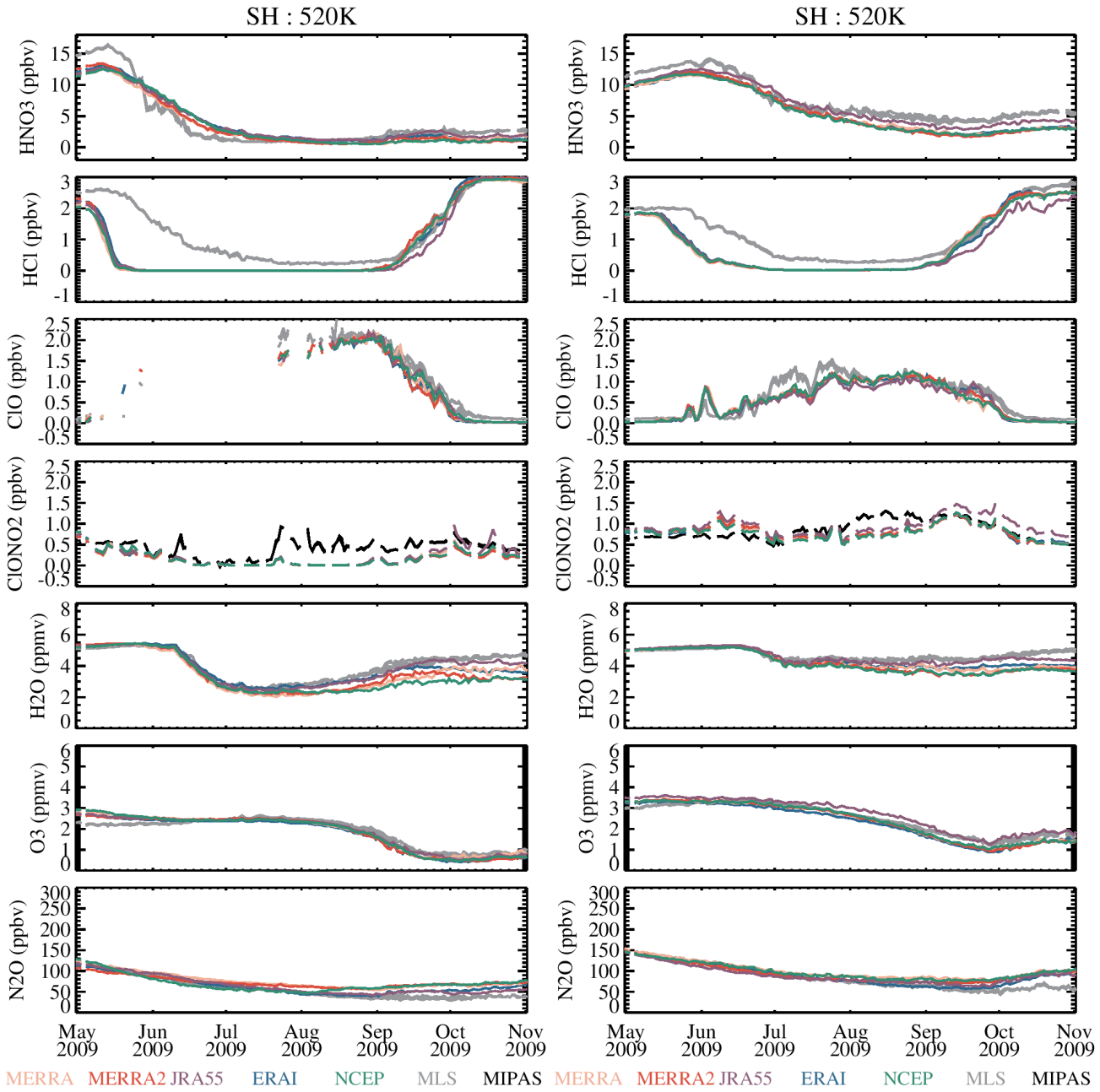

Figure 10.19: *Time series of daily averages calculated over the region of the inner vortex (75°-90°S, left) and outer vortex (60°-75°S, right) at 520 K. MIPAS $ClONO_2$ data (black lines) are averaged in geographic latitude bands. For all other species, averages of MLS data are calculated over equivalent latitude based on all five reanalyses (grey lines, largely indistinguishable). Corresponding BASCOE results for each of the simulations driven by the reanalyses (MERRA, MERRA-2, JRA-55, ERA-Interim, and NCEP (CFSR/CFSv2)) are colour-coded as indicated in the legend.*

transport is challenging, because both processes suffer from large and competing uncertainties that depend not only on the input reanalysis but also on the offline (or nudged) transport model (*Minganti et al.*, 2020, and references therein).

Differences between the individual realizations of the model are captured in the right-hand column of **Figure 10.17**. For the most part, the model runs forced by the different reanalyses produce very similar results. However, substantial disparities are evident in a few places, in particular in those regions where the gradients are largest.

Figure 10.18 paints a generally similar picture for the outer vortex. One notable difference is that the collar region is more consistently in sunlight, affording greater coverage of daytime ClO abundances. Here it is again clear in the HCl and $ClONO_2$ that the modeled chlorine activation is stronger and more extensive than that observed; the ClO also indicates slightly earlier activation in the model, although the peak ClO abundances reached in midwinter are slightly smaller than those recorded by MLS. In terms of the reanalysis ranges, again the largest spread between simulations driven by different reanalyses is found where

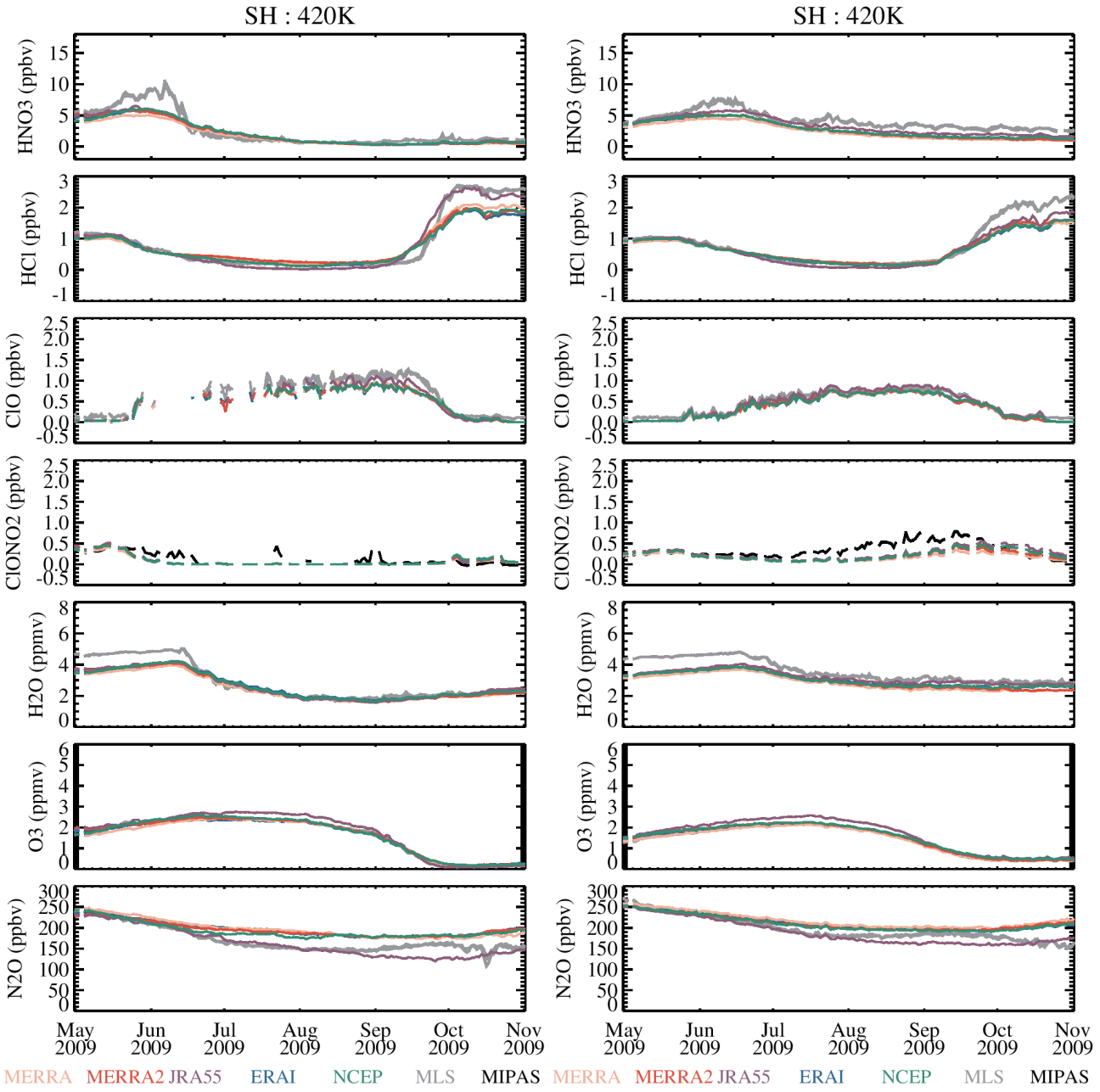

Figure 10.20: *As for* ***Figure 10.19*** *but at 420 K.*

the gradients are largest, but in general the discrepancies are smaller than those in the vortex core.

Figures 10.19 and **10.20** show slices through the data at the 520 K and 420 K potential temperature levels, respectively, marked by the horizontal white lines in the cross section plots (**Figures 10.17** and **10.18**). Averages for the inner vortex core are shown on the left and for the outer vortex collar region on the right. Conclusions similar to those discussed above can be drawn from **Figure 10.19**. Initially HNO_3 is ~20 % lower in the model than observed, but thereafter its sequestration in PSCs is reasonably well reproduced in all runs. As before, the most dramatic differences between the observed and simulated behavior are seen in HCl in early and mid-winter. However, the simulation lines are fairly tightly clustered, with the exception of JRA-55 at the end of the season, implying that the discrepancies arise from issues in the underlying model and not the reanalyses being used to force it. Although dehydration is simulated well by the model, at the end of winter H_2O recovers to larger values in the observations than in the model runs, again except for JRA-55, which agrees well with MLS. These results for water vapour suggest that the model does not overestimate the mixing across the polar vortex edge, because such erroneous mixing would lead to overestimations of H_2O inside the vortex (unless ongoing dehydration processes promptly removed the excess water vapor being mixed in throughout mid-winter), whereas the model delivers in all cases an underestimation.

After the month of July, no simulation reproduces the observed low values in N_2O in either the core or the

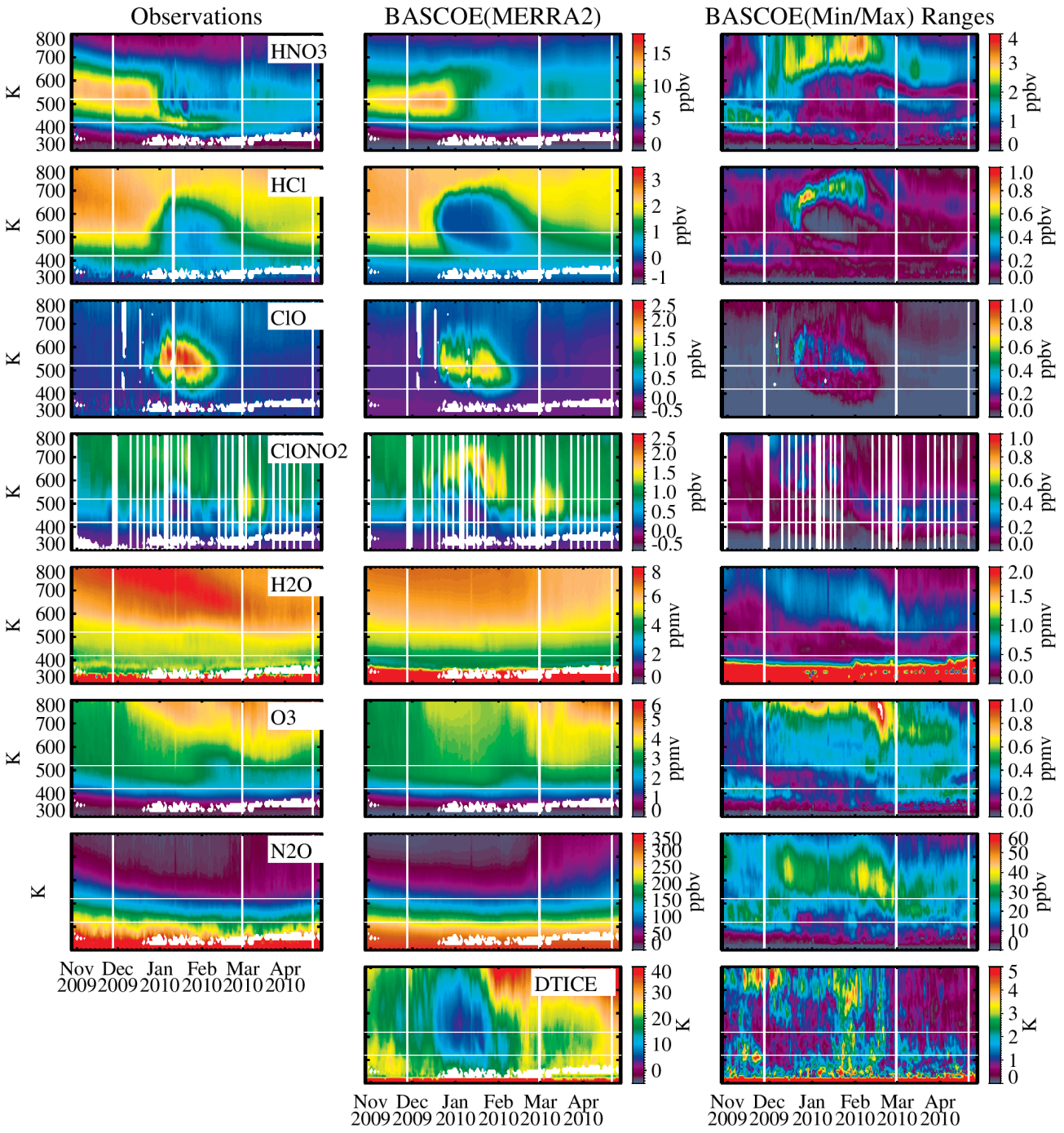

Figure 10.21: *As in* ***Figure 10.17*** *but for the 2009/2010 Arctic winter (November through April).*

collar of the vortex, although in the vortex core ERA-Interim and JRA-55 N_2O abundances are closer to those measured by MLS than the other reanalyses. Erroneous mixing across the vortex edge is probably not responsible for the overestimation in the simulations using MERRA, MERRA-2 and CFSR/CFSv2 because this issue is expected to have a similar impact in all simulations. Rather, the N_2O results at 520 K suggest that the model underestimates the strength of confined diabatic descent in the vortex interior, except for JRA-55 and ERA-Interim during the first half of the season. After the month of August, the rate of N_2O decrease is underestimated in all five simulations. As noted above, inaccurate depiction of downward transport is a known model problem. The underestimation of N_2O decrease may be interpreted as an underestimation of the diabatic descent in the vortex, but recent calculations of the Transformed Eulerian Mean budget of N_2O, using the same reanalyses and CTM, suggest that it may instead be due to underestimation of the impact of horizontal mixing with the BASCOE CTM (*Minganti et al.*, 2020). The representation of descent is also a serious issue at 420 K (**Figure 10.20**), where the contrast between JRA-55 and the other reanalyses

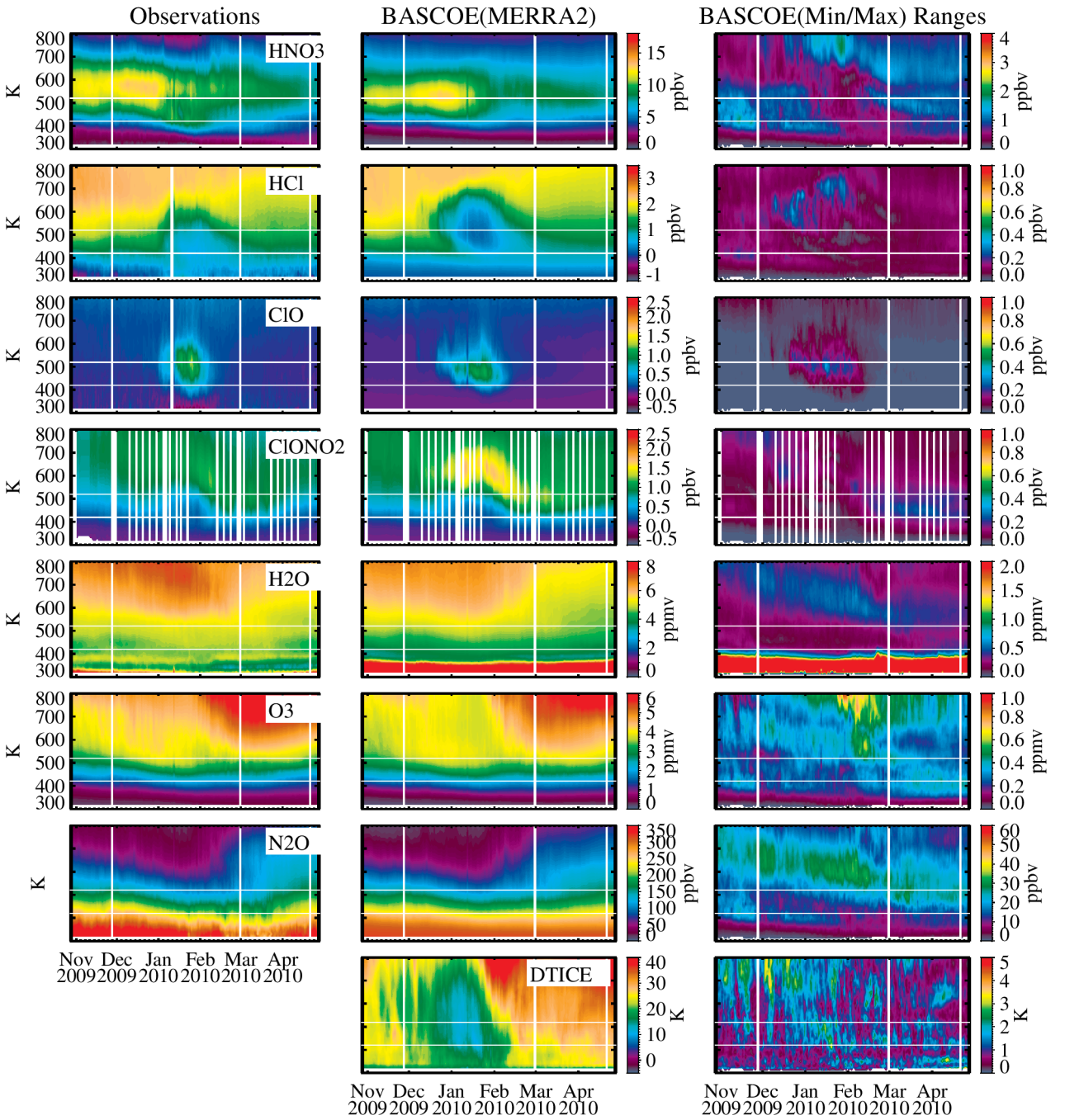

Figure 10.22: *As in* ***Figure 10.18*** *but for the Arctic.*

is even greater. At this level, however, JRA-55 seems to suggest slightly stronger descent than the MLS observations throughout the vortex, whereas the other reanalyses indicate weaker descent, as at 520 K. Other noteworthy results at 420 K include the stronger renitrification seen by MLS than any of the simulations. Faithful reproduction of observed denitrification and renitrification features has often been a challenge for models (*e.g.*, *Braun et al.*, 2019; *Khosrawi et al.*, 2018); as suggested earlier, the lack of renitrification in these BASCOE runs is likely attributable to the HNO_3 being permanently removed rather than transported downward by sedimentation. In addition, the HCl increase in October is much larger (and closer to that observed) with JRA-55 than with any other reanalysis.

The areas of largest disagreement between measured and modeled behavior are generally consistent across all five simulations, suggesting the presence of underlying problems with BASCOE not associated with the particular reanalysis being used to force the model. However, considerable spread between the simulations becomes evident in middle and late winter, particularly in the H_2O and N_2O fields. Apparently deficiencies in modeled diabatic descent within the vortex, together with other shortcomings in model physics

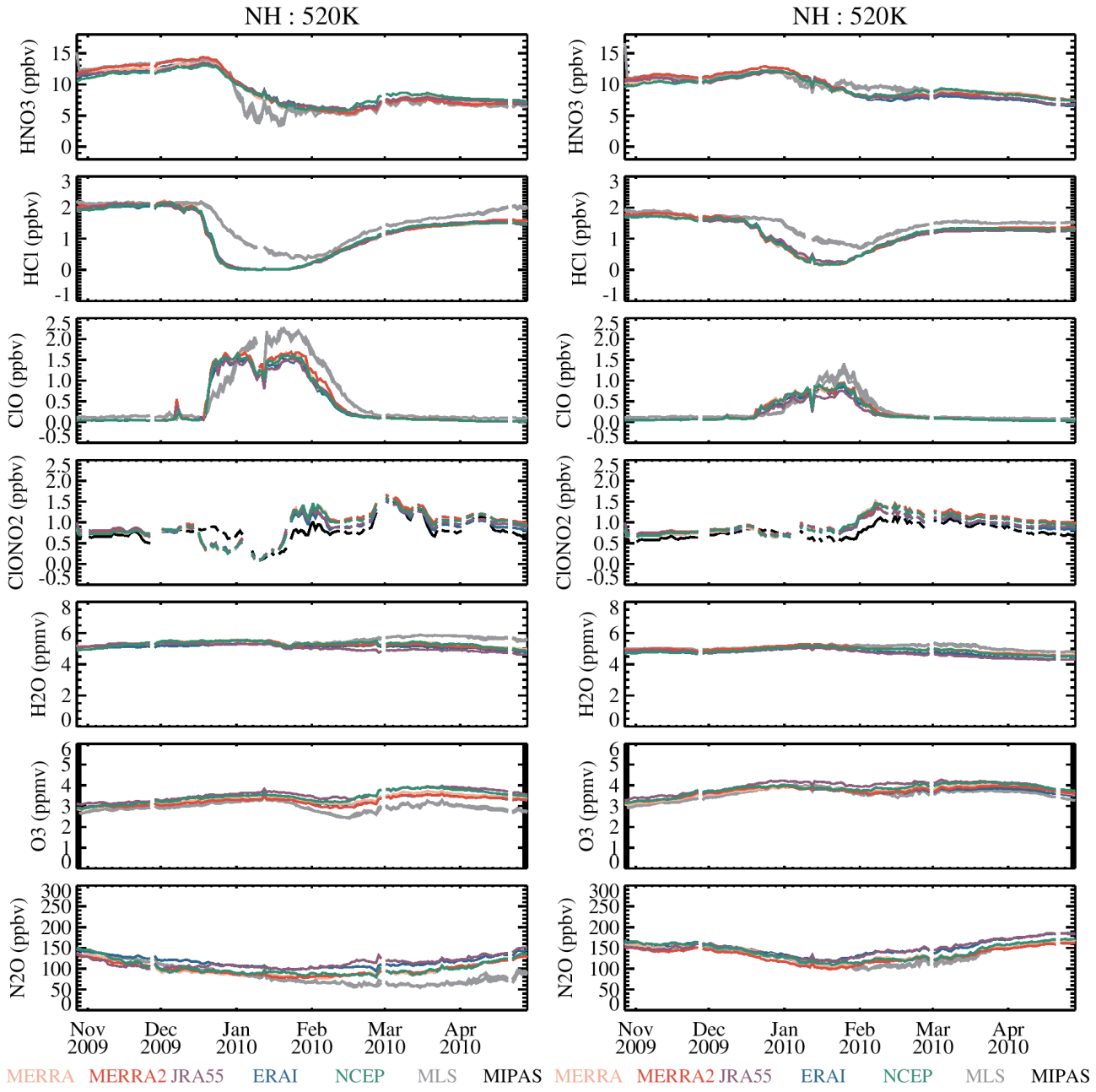

Figure 10.23: *As for* ***Figure 10.19*** *but for the Arctic.*

or chemistry schemes, manifest to varying degrees depending on the driving reanalysis, and these issues give rise to the divergence in simulated behavior seen in **Figures 10.19** and **10.20**.

10.8.4 Case study: 2009/2010 Arctic winter

Similar to the analysis for the Antarctic, cross section comparisons of the evolution over the 2009/2010 Arctic winter (November through April) are shown in **Figures 10.21** and **10.22** for the inner and outer vortex, respectively, and slices through the data at 520 K and 420 K are shown in **Figures 10.23** and **10.24**, respectively.

Again, the model performs well overall, although many of the same issues arise as for the Southern Hemisphere. At the beginning of the season, observed HNO_3 values are slightly larger than those modeled, although the agreement with MERRA-2 in particular is fairly good at 520 K. Here too MLS measurements indicate a clear signature of renitrification at 420 K that is lacking in any of the simulations (**Figure 10.24**). As before, HCl depletion at 520 K is substantially overestimated in all model realizations (**Figure 10.23**). In contrast to the situation in the Antarctic, however, peak ClO enhancement in the Arctic vortex core is considerably smaller and less persistent in the simulations than observed. Similar indications of weaker modeled diabatic descent inside the vortex as those discussed

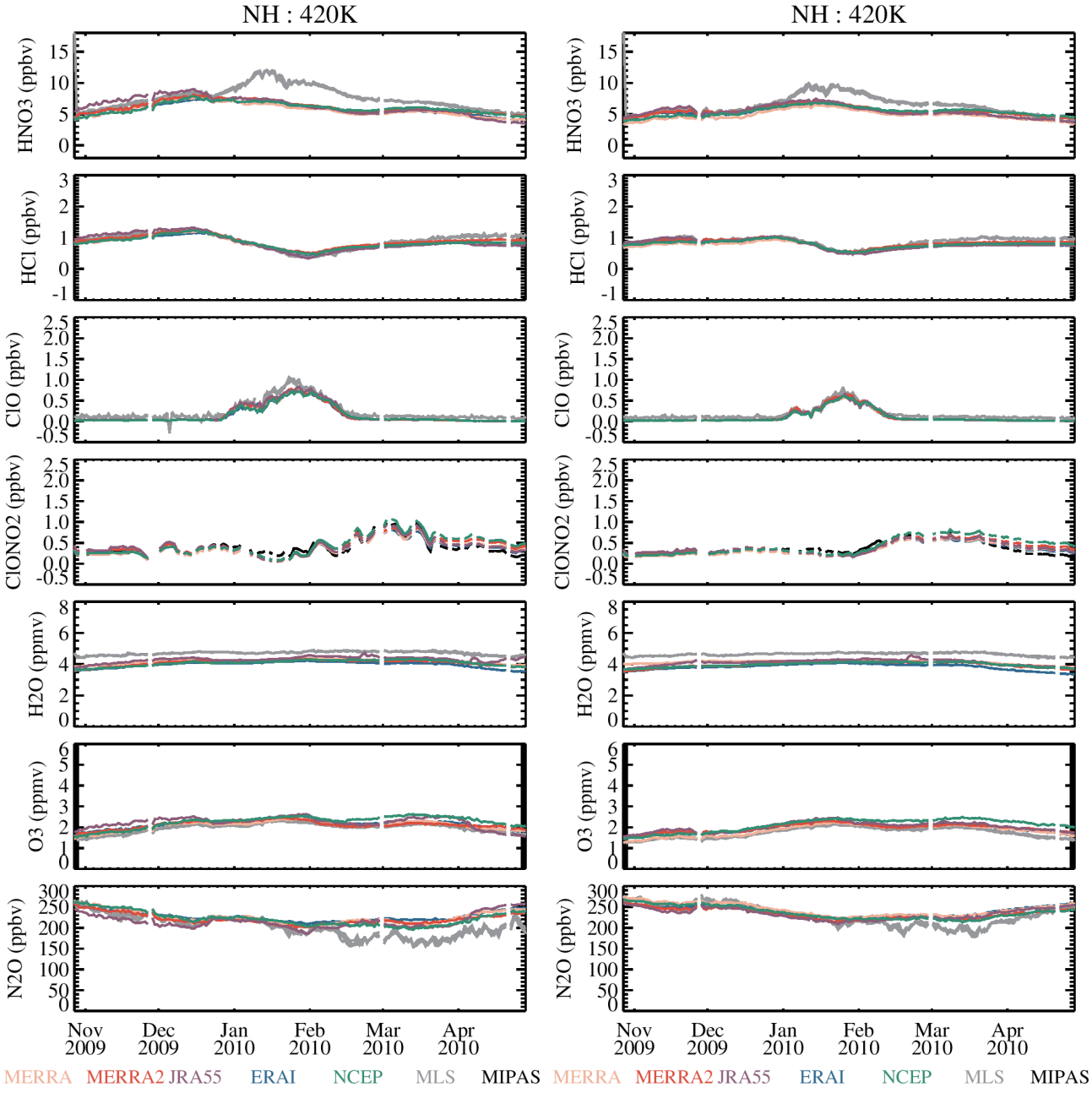

Figure 10.24: *As for* ***Figure 10.20*** *but for the Arctic.*

above are also seen at both levels in the Arctic. However, unlike in the Antarctic, where the simulations driven by JRA-55 and ERA-Interim more closely match observed behavior, in the Arctic they track the MLS measurements even less well than the other reanalyses. Overall, the maximum – minimum range of the various reanalyses is smaller for most species than that seen in the Antarctic.

10.8.5 Comparisons of chemical ozone loss

The ultimate goal of many stratospheric polar processing studies over the last several decades has been to quantify the degree of chemical ozone loss, which requires accounting for the effects of dynamics on the distribution of stratospheric ozone. Several different techniques have been developed to remove transport-induced changes in ozone and thus isolate the signature of chemical destruction (*e.g.*, *WMO*, 2007). Here we adopt the approach recently employed by *Strahan and Douglass* (2018). Ozone partial columns are calculated over the range of MLS retrieval pressure levels encompassing the majority of depletion, 261 - 12 hPa. Averages over 10-day periods in early winter (1 - 10 July) and late winter (11 - 20 September) are calculated in an attempt to reduce the effects of short-term dynamical fluctuations on ozone. The early-winter averaged partial column is then subtracted from the late-winter value to permit an estimate of chemical loss

while mitigating dynamical variability to some extent. Although it is arguable whether transport variations are fully accounted for by this method, our purpose here is not to provide the most accurate quantification of chemical ozone loss, but rather to apply a convenient methodology for doing so consistently to modeled and measured fields to facilitate intercomparison of the reanalyses.

Results are shown in **Figure 10.25** for all five BASCOE simulations as well as MLS data (analyzed separately using meteorological information from the respective reanalyses), for averages over both the inner (dark shading) and outer (light shading) vortex regions. Estimates of ozone loss based on MLS are relatively insensitive to the choice of reanalysis used for interpolation of the measurements to isentropic surfaces and identification of vortex data points, although slightly larger differences between the MLS estimates for the inner and outer vortex are seen using JRA-55 and CFSR/CFSv2. In contrast, loss estimates based on the modeled O_3 fields do differ substantially, with MERRA indicating much smaller loss, closest to that calculated from MLS data. Estimates derived from the MERRA and MERRA-2 runs disagree by ~10 DU. ERA-Interim also provides relatively weak depletion, while JRA-55 indicates much stronger depletion than the other reanalyses. Thus, from a purely BASCOE-based perspective, using JRA-55 to drive the model instead of MERRA could yield ~25 DU (30 %) more chemical ozone loss in the Antarctic vortex core.

10.8.6 Discussion and implications

In summary, five recent full-input reanalyses were used to drive the BASCOE CTM, and the results were compared to satellite observations from MLS and MIPAS. The simulations spanned a full year, from May 2009 to May 2010, allowing chemical processing during winter to be examined in both hemispheres. Overall, the model reproduced the observed seasonal evolution of stratospheric constituents well, although some discrepancies between measured and modeled behavior were noted. In terms of reanalysis intercomparison, agreement between the individual model realizations was generally very good, and in most cases only small deviations between results from the different simulations were discernible. Thus the inter-simulation spread was generally smaller than the disparities between the model (regardless of how it was forced) and the observations.

Since the main areas of disagreement between measured and modeled behavior were typically replicated across all five simulations, as discussed above in connection with the case studies, they likely arise from underlying deficiencies in model physics or chemistry not associated with the particular reanalysis being used to drive the model. One notable exception was found in long-lived tracer abundances (N_2O in particular; also H_2O and O_3 to a lesser extent) in middle and late winter in both hemispheres. Although none of the model runs faithfully reproduced the observed tracer evolution in either the core or edge regions of the vortex, considerable spread between simulations developed at this time, especially in the core of the Antarctic vortex. Results from most simulations indicated weaker confined diabatic descent inside the vortex compared to observations, with the exception of the JRA-55 run in the Arctic at 420 K, which pointed to stronger descent throughout the vortex than implied by the MLS N_2O measurements. The divergence in simulated behavior suggests that issues with the modeled depiction of transport are worse for some reanalyses than others. It must be noted that these results were obtained with a kinematic transport model wherein diabatic descent is derived from the wind fields contained in the reanalyses. As noted in *Chapter 5* for age of air calculations, a different outcome may be expected with diabatic transport models that read the heating/cooling rates computed by the radiative transfer models of the reanalysis systems (see also *Martineau et al.*, 2018, and *Section 10.5*).

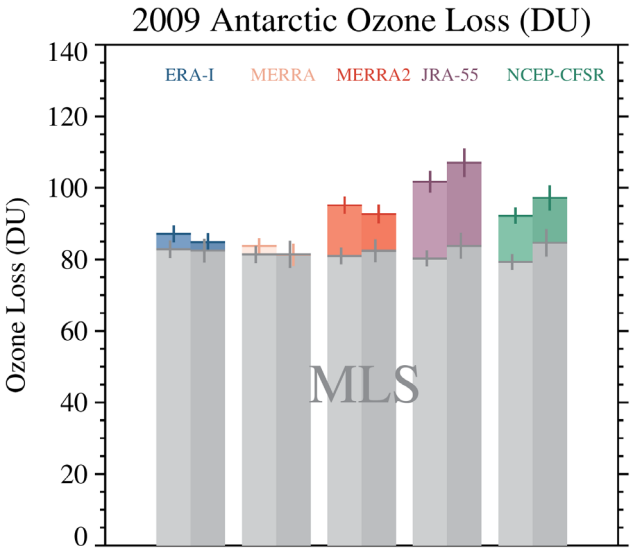

Figure 10.25: *Estimates of ozone loss for the 2009 Antarctic season (see text for details about the approach) for equivalent latitude ranges representing the outer vortex collar region (light shading: 60 °- 75 °S) and the inner vortex core (dark shading: 75 °- 90 °S), calculated from the BASCOE model output driven by each reanalysis (ERA-Interim, MERRA, MERRA-2, JRA-55, and NCEP (CFSR/CFSv2); coloured shading) and from MLS observations (grey shading). MLS data are analyzed separately for comparison to each simulation using meteorological information from the respective reanalysis; thus the five grey bars differ slightly. Note that in almost all cases, modeled ozone loss estimates exceed those derived from MLS measurements and are thus plotted as increments above the MLS values. Error bars represent the standard errors of the mean ozone losses calculated from the model data (coloured) and from the MLS observations (grey).*

10.9 Summary, key findings, and recommendations

This chapter employs an extensive set of diagnostics of relevance to polar chemical processing and dynamics to evaluate and intercompare recent full-input reanalyses, including MERRA, MERRA-2, ERA-Interim, JRA-55, and CFSR/CFSv2. The GMAO GEOS-591 operational analysis, a stable system providing consistent meteorological fields used by NASA satellite instrument teams, is also examined for PSC thermodynamic-consistency diagnostics.

To provide an overview and set the stage for more detailed intercomparisons, time series (2008 - 2013) of daily mean polar cap temperature differences at 46 hPa relative to ERA-Interim are presented first to examine seasonal (and interannual) variations. The rest of the chapter focuses on winter conditions. The comprehensive suite of polar temperature and vortex diagnostics considered in both hemispheres includes daily minimum lower stratospheric temperature poleward of 40 °, area and volume of stratospheric air with temperatures below PSC existence thresholds, maximum latitudinal gradients in PV (a measure of vortex strength), area of the vortex exposed to sunlight each day, vortex breakup dates, and polar cap average diabatic heating rates.

Comparisons with superpressure balloon measurements made during the Antarctic Concordiasi campaign in September 2010 to January 2011 are used to quantify biases in reanalysis temperatures and horizontal winds and the growth of errors along 15-day trajectories calculated using them, as well as to assess the capacity of the reanalyses to represent small-scale atmospheric fluctuations. The accuracy and precision of the reanalysis temperatures are evaluated through comparisons with both COSMIC GNSS-RO temperatures and independent absolute temperature references derived from theoretical considerations of PSC formation defined by the equilibrium thermodynamics of STS and water ice clouds. The thermodynamic-consistency diagnostics rely on near-simultaneous and colocated measurements of PSC classifications and gas-phase nitric acid and water vapour from CALIPSO CALIOP and Aura MLS satellite measurements, respectively.

Finally, because chemical model simulations synthesize the interplay among the spatially and temporally varying differences between reanalyses and exemplify how their net effects impact the bottom-line conclusions of typical real-life studies, simulated fields of nitric acid, water vapour, several chlorine species, nitrous oxide, and ozone are compared with those observed by Aura MLS and Envisat MIPAS. The key findings of this work, along with recommendations that follow from them, are summarized below.

Key findings

- In both hemispheres, differences in lower stratospheric daily polar cap averaged temperatures between ERA-Interim and other recent full-input reanalyses display annual cycles, with mainly positive deviations from ERA-Interim in summer and mainly negative deviations in winter. Thus intercomparisons of the same reanalyses undertaken for different time periods could indicate temperature deviations of roughly the same magnitude but opposite sign. Largest differences reach ~ 1 K in the Antarctic and ~ 0.5 K in the Arctic.

- Polar winter temperatures from recent full-input reanalyses are in much better agreement in the lower and middle stratosphere than were those from older reanalysis systems in common use in prior decades.

- In the Southern Hemisphere especially, a dramatic convergence toward better agreement between the reanalyses is seen after 1999 (when a major change in the input data occurred; see *Section 10.4* for more details). Average absolute differences in wintertime daily minimum temperatures poleward of 40°S from the reanalysis ensemble mean (REM) have been reduced from over 3 K prior to 1999 to generally less than 0.5 K in the most recent decade, and average differences from the REM in the area with temperatures below PSC thresholds have been reduced from over 1.5 % of a hemisphere to less than about 0.5 %. Other polar temperature and vortex diagnostics suggest a more complex picture, showing similar improvements for some reanalyses but persistent differences for others. Although convergence toward better agreement is also apparent in the Northern Hemisphere, the changes are less dramatic there because reanalysis temperatures are more consistent throughout the whole comparison period (1979 - 2015) than they are in the Southern Hemisphere.

- For many polar temperature and vortex diagnostics, the reanalyses generally agree better in the Antarctic than in the Arctic. The extremely cold conditions and relatively small interannual variability in the Antarctic mean that winter seasons tend to have similar polar processing potential and duration every year, and thus the sensitivity to differences in meteorological conditions among the reanalyses is low. In contrast, the generally warmer and more disturbed vortex and large interannual variability of Arctic winters lead to conditions that are frequently marginal, with temperatures hovering around PSC existence thresholds, and thus the sensitivity to reanalysis differences is high.

- Comparisons of polar-cap averaged diabatic heating rates in the lower stratosphere of both hemispheres show that MERRA-2, ERA-Interim, JRA-55, and CFSR give consistent results for the climatology and day-to-day evolution at pressures greater than about 20 hPa and should generally be suitable for polar processing studies. In most cases, the range of differences among the reanalyses is less than 0.5 K/day in the cold season, representing no more than about 10 - 15 % differences. These differences appear primarily as biases between the time series rather than as a failure of any reanalysis to capture day-to-day variations.

- Reanalyses differ in spatial and temporal resolution, and truncation of the models is an important factor in how well they represent small-scale fluctuations such as gravity waves. Comparisons of a subset of full-input reanalyses (ERA-Interim, MERRA, MERRA-2) with long-duration balloon observations in the Antarctic show that they reproduce the temperature and horizontal wind fluctuations of the balloons at about the 30 % level; thus a significant portion of the atmospheric gravity wave spectrum is not captured by the reanalyses. A case study of a mountain wave event shows that the temperature amplitudes resolved by the reanalyses can vary by up to a factor of two in extreme circumstances, with the differences being well correlated with the spatial resolutions of the reanalyses.

- Trajectory calculations from a Lagrangian transport model based on a subset of full-input reanalyses were evaluated using long-duration balloon observations in the Antarctic over the September 2010 to January 2011 period; relative horizontal transport deviations of 4 - 12 % and error growth rates of 60 - 170 km day^{-1} over 15-day trajectories were found for the different reanalyses.

- GNSS-RO data are not assimilated in MERRA, which consequently shows larger biases (of as much as ~ 0.5 K at 68 hPa) with respect to COSMIC temperatures compared to the other reanalyses, which do assimilate the GNSS-RO data. Differences between MERRA and COSMIC temperatures are particularly large in the Antarctic, where there is a relative paucity of conventional measurements.

- Absolute temperature references are derived from the thermodynamic equilibrium properties of certain types of PSCs and are completely independent of any data that is assimilated by the reanalyses. The reanalysis temperatures are found to be lower than these absolute reference points by up to 1 K.

- In poth polar regions, winter-long simulations from a chemistry transport model driven by different full-input reanalyses generally produce very similar results through most of the season for most species. However, substantial disparities between model runs are seen where composition gradients are largest. In particular, comparisons with satellite long-lived tracer measurements indicate that for most of the reanalyses the chemistry transport model underestimates the strength of confined diabatic descent inside the winter polar vortex. This underestimation of descent, together with other shortcomings in the model chemistry and physics schemes, manifests to varying degrees depending on the particular reanalysis used to force the model; as a consequence, considerable spread between the different simulations becomes evident by late winter.

- Results from a case study of a representative Antarctic winter (2009) show that estimates of chemical ozone loss based on satellite observations are relatively insensitive to the choice of reanalysis used for interpolation of the measurements to isentropic surfaces and identification of the vortex boundary. In contrast, chemical loss estimates based on simulated ozone fields from a chemistry transport model can differ substantially; in the case study, forcing the same model with different reanalyses yielded differences in the estimates of chemical ozone loss in the Antarctic vortex core as large as ~ 25 DU (20% - 30 %).

Recommendations

- Any of the recent full-input reanalyses (MERRA, MERRA-2, ERA-Interim, JRA-55, and CFSR/CFSv2) can be suitable for studies of lower stratospheric polar processing. However, substantial differences between the various reanalyses are found in some instances; therefore, the choice of which reanalysis to use in a given study may depend on the specific science questions being addressed.

- Temperature biases and other artifacts in older meteorological reanalyses often rendered them unsuitable for accurate modeling of the interannual variability in PSC formation and consequent denitrification, chlorine activation, and chemical ozone loss; in particular, ERA-40, NCEP-NCAR R1, and NCEP-DOE R2 are obsolete and should no longer be used for stratospheric studies of any kind.

- Because of the limitations of earlier reanalyses, it was not uncommon for modeling studies to try to match simulated chlorine activation and/or ozone loss by imposing arbitrary systematic adjustments of 1 - 2 K or more on the reanalysis temperatures. Increased confidence in the accuracy of current polar reanalysis temperatures provides tighter constraints on model parameterizations of microphysics/chemistry used to represent polar chemical processing. As a consequence, strong justification is warranted in modeling studies that seek to ascribe deficiencies in modeled chlorine activation and/or ozone losses to reanalysis temperature biases.

- Despite the overall good agreement between the polar temperatures from current full-input reanalyses, whenever feasible it is best to employ multiple reanalyses, even for studies involving recent winters for which it can reasonably be expected that differences between the reanalyses are small; using more than one reanalysis allows estimation of uncertainties and the potential impact of those uncertainties on the results, especially for quantities that cannot be directly compared with observations.

- If multiple reanalyses are used, they should be treated in the same manner to the extent possible to reduce the uncertainty in sources of differences; *e.g.*, data from one reanalysis on native model levels (*i.e.*, sigma coordinates) should not be used in conjunction with data from another reanalysis on pressure levels because of errors when interpolating from model levels to a coarser standard pressure grid.

- It would be helpful for the reanalysis centers to provide standard sets of products on standardized isobaric and isentropic levels. In particular, having PV available on model levels in future reanalyses would be extremely valuable. Vertical sampling of pressure and potential temperature levels comparable to that of model levels is also desirable.

- Reanalysis temperatures are generally unsuitable for assessment of trends in temperature-based diagnostics. One issue is that major changes in the input observational data in the assimilation systems are often made at approximately the same time in all of the reanalyses, hindering determination of the impact of such changes through reanalysis intercomparisons. Caution is especially advised for the estimation of trends in diagnostics that aggregate low temperatures over months and/or vertical levels in the Northern Hemisphere, such as the winter-mean fraction of the vortex volume with air cold enough for PSC existence; these types of diagnostics are particularly sensitive to the specific PSC threshold chosen, which is subject to non-negligible interannual variability.

Evaluation table

Figure 10.26 provides a summary evaluation of selected diagnostics examined in *Chapter 10*, as a quick reference to help users identify which reanalyses may be most suitable for a given issue related to stratospheric polar chemical processing. Only those diagnostics specifically examined either in this Chapter or in previously published papers are assigned a "score" in the table; otherwise they are marked "unevaluated" (tan shading). In particular, many of the polar processing diagnostics have not been formally assessed for the earlier ERA-40, NCEP-NCAR R1, and NCEP-DOE R2 reanalyses. However, given the considerable shortcomings in their representation of stratospheric temperatures and/or winds identified in prior studies, it is extremely unlikely that those reanalyses would perform well for the remaining diagnostics, and further evaluation of them was not undertaken here. Note that the results for some diagnostics used in this chapter are too complex to be amenable to such a rating system and have therefore been omitted. Most notably, the assessment of transport processes in the chemistry transport model yielded results that vary substantially by hemisphere, time of year, altitude, location in the polar vortex (core *vs.* edge region), and species, precluding simple categorization. Similarly, the PSC thermodynamic-consistency diagnostics have also been omitted from **Figure 10.26**.

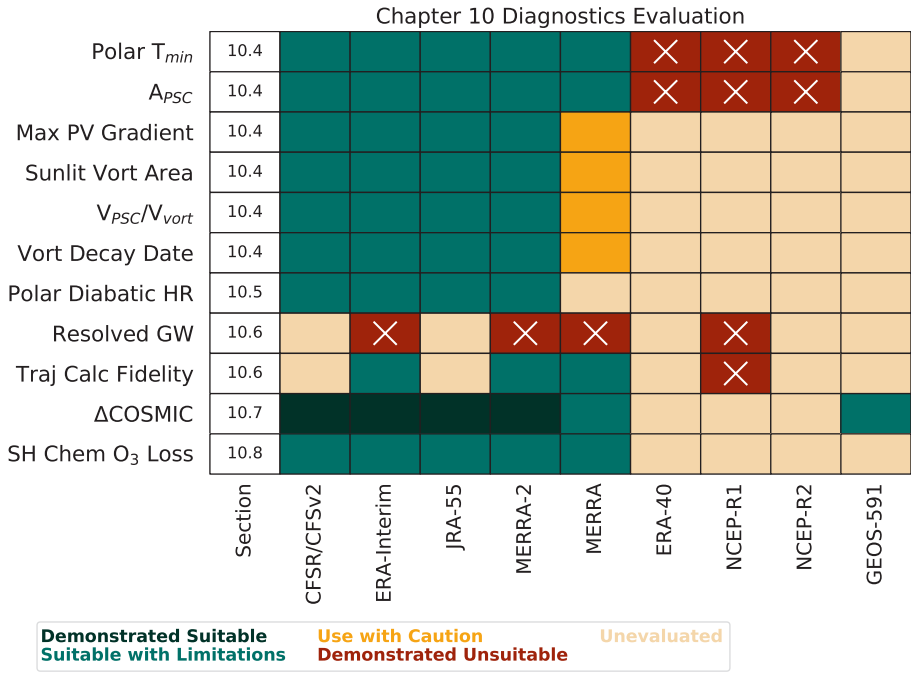

Figure 10.26: *Summary evaluation table for most of the Chapter 10 diagnostics. Along with a "score" capturing the overall performance of each reanalysis considered for each diagnostic (see table legend), the first column contains a pointer to the specific Chapter 10 section in which the diagnostic is discussed in detail.*

Data availability

Data availability for the reanalysis products is described in *Chapter 2* of this Report.

CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center (*CALIPSO Science Team*, 2015a; *CALIPSO Science Team*, 2015b).

The Concordiasi Thermodynamical SENsor (TSEN) dataset is provided by the Laboratoire de Météorologie Dynamique (*Concordiasi Science Team*, 2010).

COSMIC data were obtained from the University Corporation for Atmospheric Research COSMIC Data Analysis and Archive Center (*COSMIC Science Team*, 2013).

GEOS-5.9.1 data were obtained from the Goddard Earth Sciences Data and Information Services Center (*Global Modeling and Assimilation Office*, 2013).

MLS data are archived at the NASA Goddard Earth Sciences Data Information and Services Center (*Froidevaux et al.*, 2015; *Lambert et al.*, 2015; *Manney et al.*, 2015; *Santee et al.*, 2015; *Schwartz et al.*, 2015a,b).

Acknowledgements

We thank the Microwave Limb Sounder team at JPL, especially Brian W. Knosp, Luis F. Millán, and Ryan A. Fuller, for computational, data processing, data management, and analysis support. We gratefully acknowledge members of the teams associated with the CALIOP, COSMIC, and MIPAS satellite missions and the Concordiasi campaign, as well as the reanalysis data centers. We thank Jonathon Wright for producing zonal mean diabatic heating rates files for the reanalyses. Helpful discussions with G. Stiller and T. von Clarmann were greatly appreciated, as were the comments and suggestions from authors of other S-RIP Report chapters. Work at the Jet Propulsion Laboratory, California Institute of Technology, was carried out under a contract with the National Aeronautics and Space Administration.

References

Anthes, R.A., 2011: Exploring Earth's atmosphere with radio occultation: contributions to weather, climate and space weather, *Atmos. Meas. Tech.*, **4**, 1077 - 1103, doi: 10.5194/amt-4-1077-2011.

Balis, D., *et al.*, 2011: Observed and modelled record ozone decline over the Arctic during winter/spring 2011. *Geophys. Res. Lett.*, **38**, L23801, doi: 10.1029/2011GL049259.

Boccara, G., A.Hertzog, C. Basdevant, and F. Vial, 2008: Accuracy of NCEP/NCAR reanalyses and ECMWF analyses in the lower stratosphere over Antarctica in 2005. *J. Geophys. Res. Atmos.*, **113**, doi: 10.1029/2008JD010116.

Bonavita, M., 2014: On some aspects of the impact of GPSRO observations in global numerical weather prediction. *Q. J. Roy. Meteor. Soc.*, **140**, 2546 - 2562, doi: 10.1002/qj.2320.

Braun, M. *et al., 2019*: Nitrification of the lowermost stratosphere during the exceptionally cold Arctic winter 2015 - 2016. *Atmos. Chem. Phys.*, **19**, 13 681 - 13 699, doi: 10.5194/acp-19-13681-2019.

CALIPSO Science Team, 2015a: CALIPSO Science Team: CALIPSO/CALIOP Level 1B, Lidar Profile Data, versions 3.01, 3.02 and 3.30. Hampton, VA, USA: *NASA Atmospheric Science Data Center* (ASDC), doi: 10.5067/CALIOP/CALIPSO/CAL_LID_L1-ValStage1-V301_L1B-003.01; 10.5067/CALIOP/CALIPSO/CAL_LID_L1-ValStage1-V3-01_L1B-003.02; 10.5067/CALIOP/CALIPSO/CAL_LID_L1-ValStage1-V3-01_L1B-003.30.

CALIPSO Science Team, 2015b: CALIPSO Science Team: CALIPSO/CALIOP Level 2, Polar Stratospheric Cloud Data, versions 1.00. Hampton, VA, USA: *NASA Atmospheric Science Data Center* (ASDC), doi: 10.5067/CALIOP/CALIPSO/CAL_LID_L2_PSCMASK-PROV-V1-00_L2-001.00.

Carslaw, K.S., B.P. Luo, and T. Peter, 1995: An analytic expression for the composition of aqueous HNO_3 - H_2SO_4 stratospheric aerosols including gas phase removal of HNO_3. *Geophys. Res. Lett.*, **22**, 1877 - 1880, doi: 10.1029/95GL01668.

Chabrillat, S., *et al.*, 2018: Comparison of mean age of air in five reanalyses using the BASCOE transport model. *Atmos. Chem. Phys.*, **18**, 14 715 - 14 735, doi: 10.5194/acp-18-14715-2018.

Chipperfield, M.P., 2006: New version of the TOMCAT/SLIMCAT off-line chemical transport model: Intercomparison of stratospheric tracer experiments. *Q. J. Roy. Meteor. Soc.*, **132**, 1179 - 1203, doi: 10.1256/qj.05.51.

Chipperfield, M.P., *et al.*, 1997: Model studies of chlorine deactivation and formation of $ClONO_2$ collar in the Arctic polar vortex. *J. Geophys. Res.*, **102**, 1467 - 1478, doi: 10.1029/96JD00442.

Concordiasi Science Team, 2010: Concordiasi Thermodynamical SENsor (TSEN) dataset. Available at: **http://www.umr-cnrm.fr/concordiasi-dataset/index.php**, last access on 14 July 2021.

COSMIC Science Team, 2013: COSMIC Data Analysis and Archive Center, Constellation Observing System for Meteorology, Ionosphere and Climate, University Corporation for Atmospheric Research, Atmospheric Profiles from COSMIC Occultation Data, *Research Data Archive at the National Center for Atmospheric Research*, Computational and Information Systems Laboratory, Boulder CO, Available at: **http://rda.ucar.edu/datasets/ds723.0/**, last access on 14 July 2021.

Daerden, F., *et al.*, 2007: A 3D-CTM with detailed online PSC-microphysics: analysis of the Antarctic winter 2003 by comparison with satellite observations. *Atmos. Chem. Phys.*, **7**, 1755 - 1772, doi: 10.5194/acp-7-1755-2007.

Davies, S. *et al.*, 2003: Modeling the effect of denitrification on Arctic ozone depletion during winter 1999/2000. *J. Geophys. Res.*, **108**, 8322, doi: 10.1029/2001JD000445.

Dunkerton, T., 1978: On the mean meridional mass motions of the stratosphere and mesosphere. *J. Atmos. Sci.*, **35**, 2325 - 2333, doi: 10.1175/1520-0469(1978)035<2325:OTMMMM>2.0.CO;2.

Errera, Q., *et al.*, 2008: 4D-Var assimilation of MIPAS chemical observations: ozone and nitrogen dioxide analyses. *Atmos. Chem. Phys.*, **8**, 6169 - 6187, doi: 10.5194/acp-8-6169-2008.

Errera, Q., *et al.*, 2016: Harmonisation and diagnostics of MIPAS ESA CH_4 and N_2O profiles using data assimilation. *Atmos. Meas. Tech.*, **9**, 5895 - 5909, doi: 10.5194/amt-9-5895-2016.

Errera, Q. *et al.*, 2019: Technical note: Reanalysis of Aura MLS chemical observations. *Atmos. Chem. Phys.*, **19**, 13 647 - 13 679, doi: 10.5194/acp-19-13647-2019.

Feng, W., *et al.*, 2005: Three-dimensional model study of the Antarctic ozone hole in 2002 and comparison with 2000. *J. Atmos. Sci.*, **62**, 822 - 837, doi: 10.1175/JAS3335.1.

Fischer, H., *et al.*, 2008: MIPAS: an instrument for atmospheric and climate research. *Atmos. Chem. Phys.*, **8**, 2151 - 2188, doi: 10.5194/acp-8-2151-2008.

Fonteyn, D. and N. Larsen, 1996: Detailed PSC formation in a two-dimensional chemical transport model of the stratosphere. *Ann. Geophys.*, **14**, 315 - 328, doi: 10.1007/s00585-996-0315-0.

Forster, P.M., M. Ponater, and W.Y. Zhong, 2001: Testing broadband radiation schemes for their ability to calculate the radiative forcing and temperature response to stratospheric water vapour and ozone changes. *Meteorol. Z.*, **10**, 387 - 393, doi: 10.1127/0941-2948/2001/0010-0387.

Froidevaux, L. *et al.*, 2008a: Validation of Aura Microwave Limb Sounder HCl measurements. *J. Geophys. Res.*, **113**, D15S25, doi: 10.1029/2007JD009025.

Froidevaux, L. *et al.*, 2008b: Validation of Aura Microwave Limb Sounder stratospheric ozone measurements. *J. Geophys. Res.*, **113**, D15S20, doi: 10.1029/2007JD008771.

Froidevaux, L., N. Livesey, and W. Read, 2015: MLS/Aura Level 2 Hydrogen Chloride (HCl) Mixing Ratio V004. Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Center* (GES DISC). doi 10.5067/Aura/MLS/DATA2010.

Froidevaux, L., *et al.*, 2019: Evaluation of CESM1 (WACCM) free-running and specified dynamics atmospheric composition simulations using global multispecies satellite data records. *Atmos. Chem. Phys.*, **19**, 4783 - 4821, doi: 10.5194/acp-19-4783-2019.

Fujiwara, M., *et al.*, 2017: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems. *Atmos. Chem. Phys.*, **17**, 1417 - 1452, doi: 10.5194/acp17-1417-2017.

Gelaro, R., *et al.*, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, **30**, 5419 - 5454, doi: 10.1175/JCLI-D-16-0758.1.

Global Modeling and Assimilation Office (GMAO), 2013: GEOS5.9.1 v1.2 NRT Assimilation Products FP-IT (Forward Processing for Instrument Teams), available by subscription at: **https://gmao.gsfc.nasa.gov/products/index.php**.

Gobiet, A., *et al.*, 2005: Climatological validation of stratospheric temperatures in ECMWF operational analyses with CHAMP radio occultation data. *Geophys. Res. Lett.*, **32**, doi: 10.1029/2005GL022617.

Gobiet, A., *et al.,* 2007: Retrieval of temperature profiles from CHAMP for climate monitoring: Intercomparison with Envisat MIPAS and GOMOS and different atmospheric analyses. *Atmos. Chem. Phys.*, **7**, 3519 - 3536, doi: 10.5194/acp-7-3519-2007.

Grooß, J.-U., *et al.*, 2011: Stratospheric ozone chemistry in the Antarctic: what determines the lowest ozone values reached and their recovery? *Atmos. Chem. Phys.*, **11**, 12 217 - 12 226, doi: 10.5194/acp-11-12217-2011.

Grooß, J.-U., *et al.*, 2018: On the discrepancy of HCl processing in the core of the wintertime polar vortices. *Atmos. Chem. Phys.*, **18**, 8647 - 8666, doi: 10.5194/acp-18-8647-2018.

Hanson, D. and K. Mauersberger, 1988: Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere. *Geophys. Res. Lett.*, **15**, 855 - 858, doi: 10.1029/GL015i008p00855.

Hanson, D.R. and A.R. Ravishankara, 1994: Reactive Uptake of $ClONO_2$ onto Sulfuric Acid Due to Reaction with HCl and H_2O. *J. Phys. Chem.*, **98**, 5728 - 5735, doi: 10.1021/j100073a026.

Hertzog, A., C. Basdevant, F. Vial, and C.R. Mechoso, 2004: The accuracy of stratospheric analyses in the northern hemisphere inferred from long-duration balloon flights. *Q. J. Roy. Meteor. Soc.*, **130**, 607 - 626, doi: 10.1256/qj.03.76.

Hertzog, A., *et al.*, 2008: Estimation of Gravity Wave Momentum Flux and Phase Speeds from Quasi-Lagrangian Stratospheric Balloon Flights. Part II: Results from the Vorcore Campaign in Antarctica. *J. Atmos. Sci.*, **65**, 3056 - 3070, doi: 10.1175/2008JAS2710.1.

Hertzog, A., M.J. Alexander, and R. Plougonven, 2012: On the Intermittency of Gravity Wave Momentum Flux in the Stratosphere. *J. Atmos. Sci.*, **69**, 3433 - 3448, doi: 10.1175/JAS-D-12-09.1.

Hitchman, M.H., M. McKay, and C.R. Trepte, 1994: A climatology of stratospheric aerosol. *J. Geophys. Res.*, **99**, 20 689 - 20 700, doi: 10.1029/94JD01525.

Ho, S.-P., *et al.*, 2012: Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile inter-comparison of CHAMP climate records 2002 to 2008 from six data centers. *J. Geophys. Res.*, **117**, doi: 10.1029/2012JD017665.

Hoffmann, L., *et al.*, 2016: Lagrangian transport simulations of volcanic sulfur dioxide emissions: Impact of meteorological data products. *J. Geophys. Res.*, **121**, 4651 - 4673, doi: 10.1002/2015JD023749.

Hoffmann, L., *et al.*, 2017a: Intercomparison of meteorological analyses and trajectories in the Antarctic lower stratosphere using Concordiasi superpressure balloon observations. *Atmos. Chem. Phys.*, **17**, 8045 - 8061, doi: 10.5194/acp-17-8045-2017.

Hoffmann, L., *et al.*, 2017b: A decadal satellite record of gravity wave activity in the lower stratosphere to study polar stratospheric cloud formation. *Atmos. Chem. Phys.*, **17**, 2901 - 2920, doi: 10.5194/acp-17-2901-2017.

Höpfner, M. *et al.*, 2007: Validation of MIPAS $ClONO_2$ measurements. *Atmos. Chem. Phys.*, **7**, 257 - 281, doi: 10.5194/acp-7-257-2007.

Hoppe, C.M., *et al.*, 2014: The implementation of the CLaMS Lagrangian transport core into the chemistry climate model EMAC 2.40.1: application on age of air and transport of long-lived trace species. *Geosci. Model Dev.*, **7**, 2639 - 2651, doi: 10.5194/gmd-7-2639-2014.

Hoppe, C.M., F. Ploeger, P. Konopka, and R. Müller, 2016: Kinematic and diabatic vertical velocity climatologies from a chemistry climate model. *Atmos. Chem. Phys*., **16**, 6223 - 6239, doi: 10.5194/acp-16-6223-2016.

Huijnen, V., *et al.*, 2016: C-IFS-CB05-BASCOE: stratospheric chemistry in the Integrated Forecasting System of ECMWF. *Geosci. Model Dev.*, **9**, 3071 - 3091, doi: 10.5194/gmd-9-3071-2016.

Hurst, D.F., *et al.*, 2014: Validation of Aura Microwave Limb Sounder stratospheric water vapor measurements by the NOAA frost point hygrometer. *J. Geophys. Res.*, **119**, 1612 - 1625, doi: 10.1002/2013JD020757.

Hurst, D.F., *et al.*, 2016: Recent divergences in stratospheric water vapor measurements by frost point hygrometers and the Aura Microwave Limb Sounder. *Atmos. Meas. Tech*., **9**, 4447 - 4457, doi: 10.5194/amt-9-4447-2016.

Jewtoukoff, V., *et al.*, 2015: Comparison of Gravity Waves in the Southern Hemisphere Derived from Balloon Observations and the ECMWF Analyses. *J. Atmos. Sci*., **72**, 3449 - 3468, doi: 10.1175/JAS-D-14-0324.1.

Khosrawi, F., *et al.*, 2018: Comparison of ECHAM5/MESSy Atmospheric Chemistry (EMAC) simulations of the Arctic winter 2009/2010 and 2010/2011 with Envisat/MIPAS and Aura/MLS observations. *Atmos. Chem. Phys.*, **18**, 8873 - 8892, doi: 10.5194/acp-18-8873-2018.

Kiehl, J.T. and S. Solomon, 1986: On the radiative balance of the stratosphere. *J. Atmos. Sci.*, **43**, 1525 - 1534, doi: 10.1175/1520-0469(1986)043<1525:OTRBOT>2.0.CO;2.

Knox, J.A., 1998: On converting potential temperature to altitude in the middle atmosphere. *Eos Trans. AGU*, **79**, 376, doi: 10.1029/98EO00290.

Knudsen, B.M., 1996: Accuracy of arctic stratospheric temperature analyses and the implications for the prediction of polar stratospheric clouds. *Geophys. Res. Lett.*, **23**, 3747 - 3750, doi: 10.1029/96GL03541.

Knudsen, B.M., *et al*., 2001: Comparison of stratospheric air parcel trajectories based on different meteorological analyses. *J. Geophys. Res*., **106**, 3415 - 3424, doi: 10.1029/2000JD900608.

Knudsen, B.M., *et al*., 2002: Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights 2. Results. *J. Geophys. Res.*, **107**, doi: 10.1029/2001JD001329.

Krüger, K., B. Naujokat, and K. Labitzke, 2005: The unusual midwinter warming in the southern hemisphere stratosphere 2002: A comparison to northern hemisphere phenomena. *J. Atmos. Sci.*, **62**, 603 - 613, doi: 10.1175/JAS-3316.1.

Kuo, Y.-H., M. Skumanich, P.L. Haagenson, and J.S. Chang, 1985: The Accuracy of Trajectory Models as Revealed by the Observing System Simulation Experiments. *Mon. Weather Rev*., **113**, 1852 - 1867, doi: 10.1175/1520-0493(1985)113<1852:TAOTMA>2.0.CO;2.

Kuttippurath, J. and G. Nikulin, 2012: A comparative study of the major sudden stratospheric warmings in the Arctic winters 2003/2004 - 2009/2010. *Atmos. Chem. Phys*., **12**, 8115 - 8129, doi: 10.5194/acp-12-8115-2012.

Labitzke, K. and M. Kunze, 2005: Stratospheric temperatures over the Arctic: Comparison of three data sets. *Meteorol. Z.*, **14**, 65 - 74, doi: 10.1127/0941-2948/2005/0014-0065.

Labitzke, K. and M. Kunze, 2009: On the remarkable Arctic winter in 2008/2009. *J. Geophys. Res*., **114**, doi: 10.1029/2009JD012273.

Lambert, A. and M.L. Santee, 2018: Accuracy and precision of polar lower stratospheric temperatures from reanalyses evaluated from A-Train CALIOP and MLS, COSMIC GPS RO, and the equilibrium thermodynamics of supercooled ternary solutions and ice clouds. *Atmos. Chem. Phys*., **18**, 1945 - 1975, doi: 10.5194/acp-18-1945-2018.

Lambert, A., M.L. Santee, D.L. Wu, and J.H. Chae, 2012: A-train CALIOP and MLS observations of early winter Antarctic polar stratospheric clouds and nitric acid in 2008. *Atmos. Chem. Phys.*, **12**, 2899 - 2931, doi: 10.5194/acp-12-2899-2012.

Lambert, A., W. Read, and N. Livesey, 2015: MLS/Aura Level 2 Water Vapor (H_2O) Mixing Ratio V004, Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Center* (GES DISC). doi: 10.5067/AURA/MLS/DATA2009.

Lambert, A. *et al*., 2007: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements. *J. Geophys. Res*., **112**, D24S36, doi: 10.1029/2007JD008724.

Lawrence, Z.D., *et al*., 2015: Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses. *Atmos. Chem. Phys*., **15**, 3873 - 3892, doi: 10.5194/acp-15-3873-2015.

Lawrence, Z.D., G.L. Manney, and K. Wargan, 2018: Reanalysis intercomparisons of stratospheric polar processing diagnostics. *Atmos. Chem. Phys*., **18**, 13 547 - 13 579, doi: 10.5194/acp-18-13547-2018.

Lin, S.-J. and R.B. Rood, 1996: Multidimensional Flux-Form Semi-Lagrangian Transport Schemes. *Mon. Weather Rev.*, **124**, 2046 - 2070, doi: 10.1175/1520-0493(1996)124<2046:mffslt>2.0.co;2.

Lindenmaier, R., *et al*., 2011: A study of the Arctic NO_y budget above Eureka, Canada. *J. Geophys. Res*., **116**, doi: 10.1029/2011JD016207, 2011.

Livesey, N.J., M.L. Santee, and G.L. Manney, 2015: A Match-based approach to the estimation of polar stratospheric ozone loss using Aura Microwave Limb Sounder observations. *Atmos. Chem. Phys*., **15**, 9945 - 9963, doi: 10.5194/acp-15-9945-2015.

Livesey, N.J. *et al.*, 2020: Version 4.2x Level 2 and 3 data quality and description document. *Tech. Rep. JPL* **D-33509** Rev. E, Jet Propulsion Laboratory, available at **http://mls.jpl.nasa.gov**.

Livesey, N.J., *et al.*, 2021: Investigation and amelioration of long-term instrumental drifts in water vapor and nitrous oxide measurements from the Aura Microwave Limb Sounder (MLS) and their implications for studies of variability and trends, *Atmos. Chem. Phys*., **21**, 15409 - 15430, doi:10.5194/acp-21-15409-2021.

Long, C.S., *et al*., 2017: Climatology and Interannual Variability of Dynamic Variables in Multiple Reanalyses Evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP). *Atmos. Chem. Phys*., **17**, 14,593 - 14,629, doi: 10.5194/acp-17-14593-2017.

Manney, G., *et al*., 2015: MLS/Aura Level 2 Nitric Acid (HNO_3) Mixing Ratio V004, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), doi: 10.5067/AURA/MLS/DATA2012.

Manney, G.L., *et al*., 1996: Comparison of U.K. Meteorological Office and U.S. National Meteorological Center stratospheric analyses during northern and southern winter. *J. Geophys. Res*., **101**, 10,311 - 10,334, doi: 10.1029/95JD03350.

Manney, G.L., *et al.*, 2003a: Variability of ozone loss during Arctic winter (1991 to 2000) estimated from UARS Microwave Limb Sounder measurements. *J. Geophys. Res*., **108**, 4149, doi: 10.1029/2002JD002634.

Manney, G.L., *et al*., 2003b: Lower stratospheric temperature differences between meteorological analyses in two cold Arctic winters and their impact on polar processing studies. *J. Geophys. Res*., **108**, 8328, doi: 10.1029/2001JD001149.

Manney, G.L. *et al.* 2005: Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter. *Mon. Weather Rev.*, **133**, 1261 - 1278, doi: 10.1175/MWR2926.1.

Manney, G.L. *et al.*, 2007: Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura Microwave Limb Sounder. *J. Geophys. Res.*, **112**, D24S50, doi: 10.1029/2007JD008709.

Manney, G.L. *et al.*, 2008: The evolution of the stratopause during the 2006 major warming: Satellite data and assimilated meteorological analyses. *J. Geophys. Res*., **113**, D11115, doi: 10.1029/2007JD009097.

Manney, G.L., *et al.*, 2009a: Aura Microwave Limb Sounder observations of dynamics and transport during the record-breaking 2009 Arctic stratospheric major warming. *Geophys. Res. Lett*., **36**, L12815, doi: 10.1029/2009GL038586.

Manney, G.L. *et al*., 2009b: Satellite observations and modelling of transport in the upper troposphere through the lower mesosphere during the 2006 major stratospheric sudden warming. *Atmos. Chem. Phys*., **9**, 4775 - 4795, doi: 10.5194/acp-9-4775-2009.

Manney, G.L., *et al.*, 2011a: Jet characterization in the upper troposphere/lower stratosphere (UTLS): applications to climatology and transport studies. *Atmos. Chem. Phys*., **11**, 6115 - 6137, doi: 10.5194/acp-11-6115-2011.

Manney, G.L., *et al.*, 2011b: Unprecedented Arctic Ozone Loss in 2011. *Nature*, **478**, 469 - 475, doi: 10.1038/nature10556.

Martineau, P., J.S. Wright, N. Zhu, and M. Fujiwara, M., 2018: Zonal-mean data set of global atmospheric reanalyses on pressure levels. *Earth Syst. Sci. Data*, **10**, 1925 - 1941, doi: 10.5194/essd-10-1925-2018.

Maycock, A.C. and K.P. Shine, 2012: Stratospheric water vapor and climate: Sensitivity to the representation in radiation codes. *J. Geophys. Res*., **117**, D13102, doi: 10.1029/2012JD017484.

Meijer, E.W., B. Bregman, A. Segers, and P.F.J. van Velthoven, 2004: The influence of data assimilation on the age of air calculated with a global chemistry-transport model using ECMWF wind fields. *Geophys. Res. Lett*., **31**, doi: 10.1029/2004GL021158.

Menang, K.P., 2018: Assessment of the impact of solar spectral irradiance on near-infrared clear-sky atmospheric absorption and heating rates. *J. Geophys. Res.*, **123**, 6460 - 6468, doi: 10.1029/2018JD02834.

Mertens, C.J., M.G. Mlynczak, R.R. Garcia, and R.W. Portmann, 1999: A detailed evaluation of the stratospheric heat budget: 1. Radiation transfer. *J. Geophys. Res. Atmos.*, **104**, 6021 - 6038, doi: 10.1029/1998JD200100.

Minganti, D., *et al.*, 2020: Climatological impact of the Brewer–Dobson Circulation on the N_2O budget in WACCM, a chemical reanalysis and a CTM driven by four dynamical reanalyses. *Atmos. Chem. Phys.,* **20***,* 12609 - 12631, doi: 10.5194/acp-20-12609-2020.

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339 - 1356, doi: 10.5194/gmd-8-1339-2015.

Murgatroyd, R.J. and F. Singleton, 1961: Possible meridional circulations in the stratosphere and mesosphere. *Q. J. Roy. Meteor. Soc*., **87**, 125 - 135, doi: 10.1002/qj.49708737202.

Murphy, D.M. and T. Koop, 2005: Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Q. J. Roy. Meteor. Soc.*, **131**, 1539 - 1565, doi: 10.1256/qj.04.94.

Newman, P.A. and E.R. Nash, 2005: The unusual southern hemisphere stratosphere winter of 2002. *J. Atmos. Sci.*, **62**, 614 - 628, doi: 10.1175/JAS-3323.1.

Olaguer, E.P., H. Yang, and K.K. Tung, 1992: A reexamination of the radiative balance of the stratosphere. *J. Atmos. Sci.*, **49**, 1242 - 1261, doi: 10.1175/1520-0469(1992)049<1242:AROTRB>2.0.CO;2.

Parrondo, M.C., *et al.*, 2007: Mid-winter lower stratosphere temperatures in the Antarctic vortex: comparison between ob-

servations and ECMWF and NCEP operational models. *Atmos. Chem. Phys*., **7**, 435 - 441, doi: 10.5194/acp-7-435-2007.

Pawson, S., *et al.*, 2007: Stratospheric transport using 6-h-averaged winds from a data assimilation system. *J. Geophys. Res.*, **112**, doi: 10.1029/2006JD007673, 2007.

Pitts, M.C., L.R. Poole, and L.W. Thomason, 2009: CALIPSO polar stratospheric cloud observations: second-generation detection algorithm and composition discrimination. *Atmos. Chem. Phys*., **9**, 7577 - 7589, doi: 10.5194/acp-9-7577-2009.

Pitts, M.C., L.R. Poole, A. Lambert, and L.W. Thomason, 2013: An assessment of CALIOP polar stratospheric cloud composition classification. *Atmos. Chem. Phys*., **13**, 2975 - 2988, doi: 10.5194/acp-13-2975-2013.

Plougonven, R., A. Hertzog, A. and L. Guez, 2013: Gravity waves over Antarctica and the Southern Ocean: consistent momentum fluxes in mesoscale simulations and stratospheric balloon observations. *Q. J. Roy. Meteor. Soc*., **139**, 101 - 118, doi: 10.1002/qj.1965.

Podglajen, A., A. Hertzog, R. Plougonven, and N. Žagar, 2014: Assessment of the accuracy of (re)analyses in the equatorial lower stratosphere. *J. Geophys. Res.*, **119**, 11,166 - 11,188, doi: 10.1002/2014JD021849.

Poli, P. and J. Joiner, 2004: Effects of horizontal gradients on GPS radio occultation observation operators. I: Ray tracing. *Q. J. Roy. Meteor. Soc.*, **130**, 2787 - 2805, doi: 10.1256/qj.03.228.

Poli, P., S.B. Healy, and D.P. Dee, 2010: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA–Interim reanalysis. *Q. J. Roy. Meteor. Soc*., **136**, 1972 - 1990, doi: 10.1002/qj.722.

Pommereau, J.-P., *et al.*, 2002: Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights 1. Measurements. *J. Geophys. Res*., **107**, doi: 10.1029/2001JD001379.

Rabier, F., *et al.*, 2010: The Concordiasi Project in Antarctica. *Bull. Am. Meteor. Soc.*, **91**, 69 - 86, doi: 10.1175/2009BAMS2764.1.

Ramanathan, V., 1976: Radiative transfer within the Earth's troposphere and stratosphere: a simplified radiative-convective model. *J. Atmos. Sci.*, **33**, 1330 - 1346, doi: 10.1175/1520-0469(1976)033<1330:RTWTET>2.0.CO;2.

Rieder, H.E. and L.M. Polvani, 2013: Are recent Arctic ozone losses caused by increasing greenhouse gases? *Geophys. Res. Lett.*, **40**, 4437 - 4441, doi: 10.1002/grl.50835.

Rienecker, M.M. *et al*., 2011: MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624 - 3648, doi: 10.1175/JCLI-D-11-00015.1.

Rolph, G.D. and R.R. Draxler, 1990: Sensitivity of Three-Dimensional Trajectories to the Spatial and Temporal Densities of the Wind Field. *J. App. Meteorol.*, **29**, 1043 - 1054, doi: 10.1175/1520-0450(1990)029<1043:SOTDTT>2.0.CO;2.

Rößler, T., *et al.*, 2018: Trajectory errors of different numerical integration schemes diagnosed with the MPTRAC advection module driven by ECMWF operational analyses. *Geosci. Model Dev.*, **11**, 575 - 592, doi: 10.5194/gmd-11-575-2018.

Ryan, N.J., *et al*., 2018: Assessing the ability to derive rates of polar middle-atmospheric descent using trace gas measurements from remote sensors. *Atmos. Chem. Phys.*, **18**, 1457 - 1474, doi: 10.5194/acp-18-1457-2018.

Sander, S.P., *et al.*, 2000: Chemical Kinetics and Photochemical Data for Use in Stratospheric Modeling. Supplement to evaluation 12: Update of key reactions. *Evaluation Number 13*, Publication **00–3**, JPL, available at: **http://jpldataeval.jpl.nasa.gov/**.

Santee, M.L. *et al.*, 2007: Validation of the Aura Microwave Limb Sounder HNO_3 measurements. *J. Geophys. Res.*, **112**, D24S40, doi: 10.1029/2007JD008721.

Santee, M.L. *et al.*, 2008: Validation of the Aura Microwave Limb Sounder ClO measurements. *J. Geophys. Res.*, **113**, D15S22, doi: 10.1029/2007JD008762.

Santee, M., N. Livesey, and W. Read, 2015: MLS/Aura Level 2 Chlorine Monoxide (ClO) Mixing Ratio V004. Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Center* (GES DISC). doi:10.5067/Aura/MLS/DATA2004.

Schwartz, M., N. Livesey, and W. Read, 2015a: MLS/Aura Level 2 Temperature V004. Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Center* (GES DISC), doi: 10.5067/AURA/MLS/DATA2021.

Schwartz, M., L. Froidevaux, N. Livesey, and W. Read, 2015b: MLS/Aura Level 2 Ozone (O3) Mixing Ratio V004. Greenbelt, MD, USA, *Goddard Earth Sciences Data and Information Services Center* (GES DISC). doi: 10.5067/Aura/MLS/DATA2017.

Schwartz, M.J., et al., 2008: Validation of the Aura Microwave Limb Sounder Temperature and Geopotential Height Measurements. *J. Geophys. Res*., **113**, D15S02, doi: 10.1029/2007JD008805.

Shepherd, T.G., R.A. Plumb, and S.C. Wofsy, 2005: PREFACE (Special Issue on 2002 Antarctic Major Sudden Stratospheric Warming). *J. Atmos. Sci.*, **62**, 565 - 566, doi: 10.1175/JAS-9999.1.

Sinnhuber, B.-M., *et al.*, 2011: Arctic winter 2010/2011 at the brink of an ozone hole. *Geophys. Res. Lett*., **38**, L24814, doi: 10.1029/2011GL049784.

Solomon, S., D. Kinnison, J. Bandoro,and R. Garcia, 2015: Simulation of polar ozone depletion: An update. *J. Geophys. Res. Atmos*., **120**, 7958 - 7974, doi: 10.1002/2015JD023365.

Stohl, A., 1998: Computation, accuracy and applications of trajectories–A review and bibliography. *Atmos. Env.*, **32**, 947 - 966, doi: 10.1016/S1352-2310(97)00457-3.

Stohl, A., G. Wotawa, P. Seibert, and H. Kromp-Kolb, 1995: Interpolation Errors in Wind Fields as a Function of Spatial and Temporal Resolution and Their Impact on Different Types of Kinematic Trajectories. *J. App. Meteorol.*, **34**, 2149 - 2165, doi: 10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2.

Stohl, A., *et al.*, 2005: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.*, **5**, 2461 - 2474, doi: 10.5194/acp-5-2461-2005.

Strahan, S.E. and A.R. Douglass, 2018: Decline in Antarctic Ozone Depletion and Lower Stratospheric Chlorine Determined From Aura Microwave Limb Sounder Observations. *Geophys. Res. Lett.*, **45**, 382 - 390, doi: 10.1002/2017GL074830.

Tilmes, S., et al., 2006: Chemical ozone loss in the Arctic and Antarctic stratosphere between 1992 and 2005. Geophys. Res. Lett., 33, L20812, doi: 10.1029/2006GL026925, 2006.

Vincent, R.A. and A. Hertzog, 2014: The response of superpressure balloons to gravity wave motions. *Atmos. Meas. Tech.*, **7**, 1043 - 1055, doi: 10.5194/amt-7-1043-2014.

von Clarmann, T., *et al.*, 2009: Retrieval of temperature, H_2O, O_3, HNO_3, CH_4, N_2O, $ClONO_2$ and ClO from MIPAS reduced resolution nominal mode limb emission measurements. *Atmos. Meas. Tech.*, **2**, 159 - 175, doi: 10.5194/amt-2-159-2009.

Waters, J.W. *et al.*, 2006: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite. *IEEE Trans. Geosci. Remote Sens.*, **44**, 1075 - 1092, doi: 10.1109/TGRS.2006.873771.

Wegner, T. *et al.*, 2012: Heterogeneous chlorine activation on stratospheric aerosols and clouds in the Arctic polar vortex. *Atmos. Chem. Phys.*, **12**, 11,095 - 11,106, doi: 10.5194/acp-12-11095-2012.

Winker, D.M., *et al.*, 2009: Overview of the CALIPSO Mission and CALIOP data processing algorithms. *J. Atmos. Oceanic Technol.*, **26**, 2310 - 2323, doi: 10.1175/2009JTECHA1281.1.

WMO, 2007: Scientific assessment of ozone depletion: 2006. *Global Ozone Res. and Monit. Proj.* — Report No. **50**, Geneva, Switzerland.

WMO, 2018: Scientific assessment of ozone depletion: 2018. *Global Ozone Res. and Monit. Proj.* — Report No. **58**, Geneva, Switzerland.

Wohltmann, I., R. Lehmann, and M. Rex, 2017: A quantitative analysis of the reactions involved in stratospheric ozone depletion in the polar vortex core. *Atmos. Chem. Phys.*, **17**, 10,535 - 10,563, doi: 10.5194/acp-17-10535-2017.

Wright, J., 2017: S-RIP: Zonal-mean heating rates of global atmospheric reanalyses on pressure levels. *Centre for Environmental Data Analysis*, doi: 10.5285/70146c789eda4296a3c3ab6706931d56, last access on 14 July 2021.

Major abbreviations and terms

4D-Var	Four-Dimensional Variational data assimilation scheme
AHTD	Absolute Horizontal Transport Deviation
ATOVS	Advanced TIROS Operational Vertical Sounder
BASCOE	Belgian Assimilation System for Chemical ObsErvations
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CDAAC	COSMIC Data Analysis and Archive Center
CFSR/CFSv2	NCEP Climate Forecast System Reanalysis/Climate Forecast System, version 2
CLAES	Cryogenic Limb Array Etalon Spectrometer
CNES	Centre National d'Etudes Spatiales
COSMIC	Constellation Observing System for Meteorology, Ionosphere and Climate
CPC	Climate Prediction Center
CTM	Chemical Transport Model
DAO	Goddard Space Flight Center Data Assimilation Office
DOE	Department of Energy
ECMWF	European Centre for Medium-range Weather Forecasts
ECMWF OA	ECMWF Operational Analysis
ERA-40	ECMWF 40-year reanalysis

ERA-Interim	ECMWF interim reanalysis
ERA5	the fifth major global reanalysis produced by ECMWF
ESA	European Space Agency
FLEXPART	Flexible Particle model
FOV	Field Of View
GEOS	Goddard Earth Observing System
GMAO	NASA Global Modeling and Assimilation Office
GNSS	global navigation satellite system
GNSS-RO	global navigation satellite system - radio occultation
GPS	Global Positioning System
ICE	used to indicate water-ice particle type in PSC classification schemes
IFS	Integrated Forecast System
JRA-55	Japanese 55-year Reanalysis
MERRA	Modern Era Retrospective-Analysis for Research and Applications
MERRA-2	Modern Era Retrospective-Analysis for Research and Applications, Version 2
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
MJJASO	May-June-July-August-September-October
MLS	Microwave Limb Sounder
MPTRAC	Massive-Parallel Trajectory Calculations
NAT	Nitric Acid Trihydrate
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction of the NOAA
NCEP-DOE R2	Reanalysis 2 of the NCEP and DOE
NCEP-NCAR R1	Reanalysis 1 of the NCEP and NCAR
NH	Northern Hemisphere
NMC	US National Meteorological Center (now NCEP)
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
ppmv / ppbv	parts per million by volume / parts per billion by volume
PSC	Polar Stratospheric Cloud
PV	Potential Vorticity
REM	Reanalysis Ensemble Mean
RHTD	relative horizontal transport deviation
RMS	Root Mean Square
SD	Standard Deviation (used here for standard deviation of temperature differences)
SH	Southern Hemisphere
S-RIP	SPARC Reanalysis Intercomparison Project
SSW	Sudden Stratospheric Warming
STS	Supercooled Ternary Solutions
TOVS	TIROS Operational Vertical Sounder
TSEN	Thermodynamical SENsor
UCAR	Universities for Cooperative Atmospheric Research
UKMO	UK Meteorological Office
V_{PSC}	Volume of air with temperatures below the PSC threshold
WMO	World Meteorological Organization