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

Alyn Lambert[1] and Michelle L. Santee[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

Correspondence to: A. Lambert, (Alyn.Lambert@jpl.nasa.gov)

Abstract.

We investigate the accuracy and precision of polar lower stratospheric temperatures (100–10 hPa during 2008–2013) reported in several contemporary reanalysis data sets comprising two versions of the Modern-Era Retrospective analysis for Research and Applications (MERRA and MERRA-2), the Japanese 55-year Reanalysis (JRA-55), the European Centre for Medium-range Weather Forecasts (ECMWF) interim reanalysis (ERA-I), and the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (NCEP-CFSR). We also include the Goddard Earth Observing System Model version 5.9.1 near real-time analysis (GEOS-5.9.1). Comparisons of these datasets are made with respect to retrieved temperatures from the Aura Microwave Limb Sounder (MLS), Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) Global Positioning System (GPS) Radio Occultation (RO) temperatures, and independent absolute temperature references defined by the equilibrium thermodynamics of supercooled ternary solutions (STS) and ice clouds. Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) observations of polar stratospheric clouds are used to determine the cloud particle types within the Aura MLS geometric field of view. The thermodynamic calculations for STS and the ice frost point use the colocated MLS gas-phase measurements of HNO_3 and H_2O. The estimated accuracy and precision for the STS temperature reference, over the 68 to 21 hPa pressure range, is 0.6–1.5 K and 0.3–0.6 K, respectively; for the ice temperature reference they are 0.4 K and 0.3 K, respectively. These uncertainties are smaller than those estimated for the retrieved MLS temperatures and also comparable to GPS RO uncertainties (accuracy$<$0.2 K, precision $>$0.7 K) in the same pressure range.

We examine a case study of the time-varying temperature structure associated with layered ice clouds formed by orographic gravity waves forced by flow over the Palmer Peninsula, and compare how the wave amplitudes are reproduced by each reanalysis data set. We find that the spatial and temporal distribution of temperatures below the ice frost point, and hence the potential to form ice PSCs in model studies driven by the reanalyses, varies significantly because of the underlying differences in the representation of mountain wave activity.

High-accuracy COSMIC temperatures are used as a common reference to intercompare the reanalysis temperatures. Over the 68–21 hPa pressure range, the biases of the reanalyses with respect to COSMIC temperatures for both polar regions fall

within the narrow range of −0.6 K to +0.5 K. GEOS-5.9.1, MERRA, MERRA-2 and JRA-55 have predominantly cold biases, whereas ERA-I has a predominantly warm bias. NCEP-CFSR has a warm bias in the Arctic, but becomes substantially colder in the Antarctic.

Reanalysis temperatures are also compared with the PSC reference temperatures. Over the 68–21 hPa pressure range, the reanalysis temperature biases are in the range −1.6 K to −0.3 K with standard deviations ∼0.6 K for the CALIOP STS reference, and in the range −0.9 K to +0.1 K with standard deviations ∼0.7 K for the CALIOP ice reference. Comparisons of MLS temperatures with the PSC reference temperatures reveal vertical oscillations in the MLS temperatures, and a significant low bias in MLS temperatures of up to 3 K.

1 Introduction

Over the last couple of decades, global reanalysis datasets have become one of the workhorse tools of the climate research community for understanding atmospheric processes and variability (Fujiwara et al., 2017) and more recently for potentially investigating climate changes (Thorne and Vose, 2010; Dee et al., 2014; Simmons et al., 2014). A reanalysis system combines observations with predictions from a global forecast model that propagates information forward in time and space using an assimilation scheme to produce a control-weighted blend of the observations and the new forecast. Unlike operational analysis schemes, which are updated as needed to improve numerical weather prediction (NWP) capabilities, reanalysis systems are designed to be conservative and retain the same code-base, with the aim of producing consistent and low-artifact output content over their entire multi-decade timeseries for a given product version generation.

The state-of-the-art of NWP and reanalysis data has improved greatly over the years as computational technology has evolved and new observation systems, such as global navigation satellite system (GNSS) radio occultation (RO) and data from other advanced satellites (including one-of-a-kind research satellites), have also been brought within the realm of data assimilation. Currently, there are numerous NWP centers that produce reanalysis products. Understandably, there are differences in the technical implementation of the highly complex reanalysis schemes between the NWP centers, and these can lead to differences in the reanalysis products. Accordingly, there is a need for the research community to know how the deficiencies of a particular reanalysis may impact their investigations. The SPARC (Stratosphere-troposphere Processes And their Role in Climate) Reanalysis Intercomparison Project (S-RIP) (Fujiwara et al., 2017) is a coordinated activity with the aim of understanding the underlying causes of differences among global reanalysis data products to help support the needs of the research community. Comparisons of reanalysis data with independent observations form an indispensable part of this assessment, and a core component of the evaluation process is to use satellite observations as a reference.

The focus of this paper is on the "Polar Processes" theme of the upcoming associated S-RIP report outlined by Fujiwara et al. (2017), in particular the intercomparison of reanalysis temperatures, and we seek to answer the question *"Can the thermodynamics of polar stratospheric clouds be used to provide an absolute temperature reference in the polar regions?"*. Thermodynamic properties dictate that only two types of polar stratospheric cloud particles are conducive for this application: supercooled ternary solutions and water ice. As discussed in detail in Section 3, PSCs composed of nitric acid trihydrate particles are not suitable because they are often out of equilibrium with the ambient HNO_3. Polar processes, such as the potential for ozone loss via heterogeneous reactions (Solomon, 1999; Solomon et al., 2015), depend critically on temperature. The potential ozone loss is related to the volume of stratospheric air that is below certain critical temperature thresholds associated with formation of PSCs (Rex et al., 2004; Orsolini et al., 2009; Harris et al., 2010). Mean winter temperatures in the Arctic stratosphere are significantly higher than those in the Antarctic (e.g. Waugh and Polvani, 2010), such that the propensity for the formation of persistent Arctic PSCs is much lower, and the development of widespread synoptic ice PSCs is quite rare (Engel et al., 2013). Since the results of polar modeling studies are especially susceptible to errors in the underlying temperature fields, Chemistry-Climate Models (CCMs) (Butchart et al., 2011) are preferably used in a specified-dynamics (SD) mode for studying polar processes. In the SD-mode, CCMs are constrained by using Newtonian relaxation to nudge the model output at each time step towards the reanalysis temperatures and wind fields (e.g. Whole Atmosphere Community Climate Model (WACCM), Wegner et al., 2013)), and are therefore adapted better to represent polar winters with unusual characteristics such as split vortices, exceptionally cold conditions, sudden warmings, etc. Accurate temperatures are critical, not only to process studies, but also to assessment of trends in stratospheric winter conditions, especially considering the prospects of an increased frequency of future episodes of severe Arctic ozone depletion (Rex et al., 2006; Sinnhuber et al., 2011; Langematz et al., 2014), following the record low seen in 2011 (Manney et al., 2011). A temperature difference of less than 1 K can have a significant effect on the outcome of model runs that are driven by reanalysis data. Errors in the extent of heterogeneous processing, arising from temperature biases, result in under/over estimate of activated chlorine, thus leading to too little/much chemical ozone loss. Temperature adjustments have often been applied in order to better match model predictions with observations (e.g. Danilin et al., 2000). Wohltmann et al. (2013) found a discrepancy in modeled and observed HCl and ClO which could be improved by adjusting the reanalysis temperatures by -1.0 K for ERA-I in the Arctic 2009/2010 winter. Brakebusch et al. (2013) found an improvement in modeled ozone loss with a -1.5 K adjustment for an unspecified version of GEOS5 in the Arctic 2004/2005 winter. Wegner et al. (2012) showed a comparison of ERA-I Arctic temperatures in March 2005 that indicates a 1.5 K warm bias in ERA-I below 205 K compared to the ambient temperatures measured with the Geophysica high-altitude research aircraft. Solomon et al. (2015) demonstrated that a -2.0 K adjustment to the MERRA temperatures in the 2011 Arctic winter provides a better fit of the zonally averaged modeled total column ozone to the observations. Improved knowledge of the temperature biases in reanalysis data would enhance confidence in the attribution of model errors to the underlying physico-chemical properties of PSCs and heterogeneous reactions.

Several previous intercomparisons of analyses and reanalyses generated by various national centers have been carried out (e.g. Manney et al., 1996; Pawson et al., 1999; Manney et al., 2003, 2005) to assess their accuracy and ultimate suitability specifically for stratospheric polar studies. Independent datasets such as radiosondes, satellite observations, Global Positioning

System (GPS) Radio Occultation (RO) (Nedoluha et al., 2007) and temperature sensors on long-duration balloons (Hertzog et al., 2004; McDonald and Hertzog, 2008) have also been used to assess reanalysis temperatures. Data denial experiments, in which the analyses are compared with and without ingestion of a particular dataset, are often used to test assimilation systems (e.g. Bauer et al., 2014).

Historically, the uncertainties in polar reanalysis temperatures, especially for the southern hemisphere, have been higher than those in other regions of the globe, because of sparse coverage from conventional temperature measurements such as radiosondes (e.g. Gobiet et al., 2005; de la Torre Juárez et al., 2009). In recent decades, the augmentation of the coverage of polar regions by an increase in satellite missions, including assimilation of research satellite data, and notably the advent of GNSS RO, has dramatically improved the situation (Wang and Lin, 2007). Lawrence et al. (2015) examined a 34-year record of polar processing diagnostics for MERRA and ERA-I reanalyses. They documented the introduction of new data streams and noted better agreement in the post-2001 timeframe following the assimilation of Aqua Atmospheric Infrared Sounder (AIRS) and Geostationary Operational Environmental Satellite (GOES) radiances into both schemes.

In section 2 we review the satellite instruments, reanalysis datasets, and methodology. In section 3 we review the equilibrium thermodynamics associated with the formation of stratospheric ternary solutions and ice clouds. In section 4 we present and discuss the results of the comparisons. Finally, in section 5 we present the conclusions.

2 Datasets and methodology

We confine our investigations to the temperature data over the past decade from six contemporary reanalysis datasets and introduce a novel analysis based on the thermodynamics of supercooled ternary solutions and the ice frost point to provide an absolute temperature reference. Near-simultaneous and colocated measurements of nitric acid, water vapor and cloud phases are currently only afforded by the precise formation flying of two satellite instruments in the A-train. These are the Aura Microwave Limb Sounder (MLS) instrument, which measures the gas-phase species, and the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar, which is used to classify PSC types. The analysis presented refines and extends the methodology used originally by Lambert et al. (2012) to investigate the temperature existence regimes of different types of PSCs. Supercooled Ternary Solution (STS) and ice PSCs are identified by the CALIOP lidar PSC classification. We accumulate statistics on the existence regimes for STS and ice PSCs by using CALIOP to identify the presence of PSCs in the MLS geometric field of view at the along track resolution (165 km by 2.16 km). MLS is used to obtain the ambient gas-phase H_2O and HNO_3 volume mixing ratios. These are required to calculate the theoretical equilibrium temperature dependence of the STS (T_{eq}) and ice (T_{ice}) PSCs. We compare the observed and calculated temperature distributions of (a) the uptake of HNO_3 in STS, and (b) the ice frost point, for each reanalysis data set and for MLS temperature.

We select viewing scenes in which there is a distinct dominant PSC classification in a sample volume that is similar in size to the MLS gas-species resolution. The requirement is that 75% or more of the CALIOP pixels in the MLS geometric field of view (FOV) have the same PSC classification. Scenes satisfying this requirement for CALIOP STS detections we denote as `LIQ`, and for CALIOP ice detections we denote as `ICE`. This is discussed in detail in Section 4.4.

To summarize, our approach includes the following steps :

- 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 $\times$ 2.16 km), ensuring that the same PSC type is detected in at least 75% of the MLS field of view.
- 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 (a) calculated and observed HNO_3 uptake in STS and (b) ice temperature distribution vs the frost point with reanalysis data and MLS temperatures.
- Create LIQ and ICE temperature distributions for each reanalysis dataset (all sampled by interpolating to the MLS measurement times and locations) and additionally for MLS temperature.
- Calculate the median and mean temperature deviations from T_{eq} and T_{ice} and their standard deviations for LIQ and ICE classifications, respectively.

The accuracy (bias) of the reanalysis temperatures is obtained relative to the reference temperatures (LIQ, ICE, and COSMIC). We use the calculated standard deviations of the temperature differences to estimate the measurement precisions, i.e. high precision is determined by the narrowness of repeated measurements with respect to a measure of their central value such as the mean or median.

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

2.1 Reanalysis temperature data

The reanalysis data used are from four analysis centers (NASA, NOAA/NCEP, ECMWF, JMA). Details pertaining to the S-RIP intercomparisons can be found in Fujiwara et al. (2017) and are summarized in Table 1.

The synoptic gridded reanalysis temperatures are interpolated in log pressure to a common vertical pressure grid (100–10 hPa), with 6 levels per decade ($p = 100 \times 10^{-i/6}; i = 0, \ldots, 6$), and to the MLS measurement times and geophysical locations. A separate dataset of colocated reanalysis temperature profiles is also created for the COSMIC measurement times and locations, with the COSMIC temperatures linearly interpolated in log pressure to the same 6 level per decade grid.

2.2 CALIOP PSC data

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 dataset

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 5 km horizontal by 180 m vertical resolution. 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. 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 HNO_3 data product. This we refer to as the MLS geometric field of view. There are approximately four hundred CALIOP 5 km x 0.18 km "pixels" within the MLS geometric field of view.

2.3 MLS gas-phase constituents and temperature

MLS is onboard the Aura spacecraft launched in July 2004 and measures thermal emission at millimeter and sub-millimeter wavelengths from the Earth's limb (Waters et al., 2006) along the forward direction of the Aura spacecraft flight track, with a vertical scan from the surface to 90 km every 24.7 s. Each orbit consists of 240 scans spaced at 1.5° (165 km) along-track, with a total of almost 3500 profiles per day and latitudinal coverage from 82° S to 82° N. The Level-1 limb radiance measurements are inverted using 2-D optimal estimation (Livesey et al., 2006) to produce Level-2 profiles of atmospheric temperature and composition. Validation of a previous version of the MLS H_2O and HNO_3 data products and error estimations are discussed in detail by Read et al. (2007), Lambert et al. (2007) and Santee et al. (2007). MLS temperature validation and error analysis is discussed by Schwartz et al. (2008). Here we use the MLS version 4 (v4) data (Livesey et al., 2017), which have single-profile precisions (accuracies) of 4–15 % (4–7 %) for H_2O, 0.6 ppbv (1–2 ppbv) for HNO_3 and for temperature a precision of 0.7 K and a bias in the range -2–0 K. 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 frost point temperature 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 (Hurst et al., 2014) showed no statistically significant differences (agreement to better than $<1\%$) from 68–26 hPa, although significant biases at 100 and 83 hPa were found to be 10% and 2%, respectively. Increasing the time frame to mid-2015 (Hurst et al., 2016) suggests a long-term drift in MLS H_2O of up to 1.5% per year starting around 2010. This is still under investigation by the MLS science team, but the effect on the calculated STS reference and frost point temperatures would be less than 0.1 K per year. Estimated uncertainties in the reference point temperatures are discussed in Sections 3.1 and 3.2.

2.4 CALIPSO and Aura configuration within the A-train

CALIPSO and Aura are part of the afternoon "A-train" satellite constellation at 705 km nominal altitude and 98° inclination, with daily near-global coverage attained in 14.5 orbits. 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 $\sim$200 km, with the MLS tangent point leading the CALIOP nadir view by about 7.5 minutes. Since April 2008, Aura and CALIPSO have been operated to maintain positioning

within tightly constrained control boxes, such that the MLS tangent point and the CALIOP nadir view are colocated to better than about 10-20 km and about 30 seconds.

2.5 COSMIC GPS RO temperatures

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). GPS RO data have provided high accuracy (< 0.2 K, Gobiet et al., 2007), global (day and night) coverage, coupled with excellent long term stability for nearly two decades (Anthes, 2011). The introduction of GPS RO has been documented to improve NWP forecast skill in the ECMWF Integrated Forecast System (IFS) (Bonavita, 2014) and to reduce tropopause and lower stratospheric temperature biases in ERA-I (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 GPS RO data is required. Even though the RO measurements of signal phase delays are traceable to the SI second (Anthes, 2011), in common with all satellite limb sounding techniques, 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 vapor, and gradients in temperature along the ray path (Poli and Joiner, 2004; Anthes, 2011). Many other studies have intercompared GPS RO with independent operational analyses, e.g. with forecast versions that have not assimilated the GPS 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.

Schreiner et al. (2007) investigated the precision of the COSMIC data by leveraging the close configuration of the six satellites during the early deployment phase of the mission. Further work by Staten and Reichler (2008) examined the tropopause temperatures obtained from various GPS RO datasets and determined a global mean bias of <0.1 K between the different RO instruments and of <0.5 K between RO instruments and radiosondes. A later study by the same authors (Staten and Reichler, 2009) reported a more comprehensive analysis of the RO temperature precision for the COSMIC data that examined variations with height, latitude and season. The concept of apparent precision was introduced, whereby the temperature differences between colocated neighboring RO measurements are organized by their temporal and spatial differences. A quadratic polynomial surface was fit to the RMS temperature differences, acting as a noise filter for the atmospheric variability and providing the apparent temperature precision of perfectly colocated data via extrapolation of the fitted RMS values to the time and distance origin. For the 15 km to 26 km height range, the RMS differences increase monotonically from about 0.5 K to 1.5 K. The results presented graphically by Staten and Reichler (2009) appear to be applicable to the RMS of colocated measurement pairs and should therefore be divided by $\sqrt{2}$ to represent the precision on a single GPS measurement.

Alexander et al. (2014) estimated the COSMIC GPS RO temperature precision by examining the RMS difference between a large ensemble of pairs of independent COSMIC observations (both wet and dry temperature profiles). Temperature accuracy

was assessed by comparing to colocated ECMWF analyses, yielding an average bias of 0.1 K (RO−ECMWF) for 10 to 30 km and −0.5 K above 35 km. Comparisons of COSMIC GPS RO with ERA-I (Dee et al., 2011) showed deviations of ±1 K at 100 hPa and ±2 K at 10 hPa for the polar regions. Ladstädter et al. (2015) made COSMIC GPS RO comparisons to Vaisala RS90/92 radiosondes and found mean temperature differences for 100–30 hPa of <0.2 K since 2006 and of <0.5 K for 30–10 hPa. Scherllin-Pirscher et al. (2011b) modeled the observed vertical error structure of GPS RO measurements, including the dry temperature error variation. Differences between RO instruments from different missions were found to have global mean standard deviations differing by at most 0.2 K at all levels between 4 and 35 km. The GPS RO model error, s_{model}, is a constant value, $s_0 = 0.7$ K around the tropopause, increasing exponentially with height into the stratosphere and given by

$$s_{model} = \begin{cases} s_0 & \text{for } z_{Ttop} < z < z_{S_{bot}}, \\ s_0\, exp\left[\frac{z - z_{S_{bot}}}{H_{S_{bot}}}\right] & \text{for } z_{Sbot} \leq z \end{cases} \tag{1}$$

where z_{Ttop} is the top level of the tropopause, z_{Sbot} is the bottom level of the stratopause. Latitudinal and seasonal variations are governed by an adjustable scale height parameter, H_S, which has its lowest values at high latitudes and during the winter months. The GPS model error has a different functional form for the troposphere (Scherllin-Pirscher et al., 2011b), which we do not require here.

The GPS RO data provide a measurement of the atmospheric refractivity, N, which is dependent on the temperature, T (in K), atmospheric pressure, p (in hPa), and the water vapor partial pressure, p_w (in hPa). For heights above 4 km, and following correction for ionospheric electron density, N, is given by (Smith and Weintraub, 1953)

$$N = c_1 \frac{p}{T} + c_2 \frac{p_w}{T^2} \tag{2}$$

where $c_1 = 77.60$ K hPa^{-1} and $c_2 = 3.73 \times 10^5$ K^2 hPa^{-1}. The first term in the equation represents the contribution to refractivity from the total molecular polarizability of the dry atmosphere, whereas the second term corresponds to the effects of the permanent dipole moment of water vapor molecules. In regions of low atmospheric humidity, above about 14 km at low latitudes and above about 9 km at high latitudes, dry temperatures provide an adequate representation of the physical atmosphere (Scherllin-Pirscher et al., 2011a). We used the latest available COSMIC GPS version 2013.3520 "atmPrf" dry temperatures processed at the UCAR CDAAC. These have been compared recently to Vaisala RS92 radiosonde temperatures (Ho et al., 2017), and the mean radiosonde − RO global daytime temperature difference over 200–20 hPa was found to be 0.20 K, with a mean standard deviation of 1.5 K. For completeness, we also evaluated the "wetPrf" wet temperatures for the same data version from 2008 to 2013 in the polar regions. We obtained mean temperature differences (dry − wet) in the range −0.1 to +0.05 K at 100 hPa and −0.4 to +0.15 K at 10 hPa, with standard deviations of 0.15 K at 100 hPa increasing to 1.1 K at 10 hPa. The wet temperatures yield results that are consistent with the dry temperatures and therefore are not considered further in this paper.

We have not implemented the line of sight corrections (Feltz et al., 2014b, a) that are possible when comparing against gridded temperature fields. On a case-by-case basis the correction can be >1 K, depending on the orientation of the GPS raypath with respect to temperature gradients, but the raypath averaging method has been shown to reduce the bias (Feltz et al.,

2014b, a). For larger sample sizes the biases are reduced because the positive/negative temperature gradients are averaged away even for a simple closest matching pair method (Feltz et al., 2014b, a).

In Section 4.3 we compare each reanalysis data set to the COSMIC data and calculate the bias and standard deviation of the temperature differences.

3 Thermodynamics of PSC formation

The growth of solid nitric acid trihydrate (NAT) crystals is kinetically limited (e.g. Voigt et al., 2005) and is frequently out of equilibrium with the gas-phase HNO_3 abundance. This makes NAT PSCs unsuitable for use as a local temperature reference, since their growth by HNO_3 uptake is dependent on the temperature history (Larsen et al., 1997). On the other hand, liquid STS reacts more quickly to ambient temperature changes, effectively acting as a thermometer (Hoyle et al., 2013), provided that very rapid temperature changes are excluded (Voigt et al., 2005). Ice PSCs also react quickly to temperature changes, because of the high ambient H_2O. Critical supercooling is the temperature depression below the frost point where homogeneous freezing of the STS can take place to form ice (Larsen, 2000). A depression of about 3 K is typically required (Carslaw et al., 1998a; Koop et al., 1998). However, heterogeneous nucleation, such as the formation of ice on pre-existing NAT or freezing of STS mediated by active sites on foreign nuclei immersed within the liquid drop (Engel et al., 2013), does not require a temperature depression below the frost point.

Theoretical existence temperatures of the PSC types are calculated using equilibrium thermodynamics and 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 HNO_3 for STS (Carslaw et al., 1995). Errors in the calculations of these reference temperatures arising from uncertainties in the MLS H_2O and HNO_3 data are estimated to be $\leq$0.5 K for T_{ice} and $\leq$0.7 K for T_{eq} in the pressure range 68–21 hPa.

It is expected that theoretical equilibrium calculations are more appropriate for 'steady state' conditions, e.g. PSCs produced by slow synoptic cooling, than rapidly changing temperatures associated with the cold phases of mountain waves (Carslaw et al., 1998b; Dörnbrack et al., 2001). Therefore, mountain wave production of ice that is detected by the CALIOP wave-ice classification is excluded from this study. However, mountain wave activity also influences the production of STS (Carslaw et al., 1998a). This can be potentially excluded to some extent by analysis of the wave activity in the reanalysis datasets (Knudsen, 2003) and rejection of data in the affected locations. Alternatively, since the CALIOP data have resolution superior to that of the reanalysis grids, we can address the challenge posed by the disparity between the CALIOP, MLS, and reanalysis horizontal and vertical sampling scales by selecting only those scenes with a consistent CALIOP PSC classification that fills a substantial proportion of the synoptic-scale MLS field of view (see Section 4.4). Therefore, we can effectively mitigate the visual speckle effect (Pitts et al., 2009) of disparate PSCs occurring within the MLS field of view generated by sub-grid temperature fluctuations that are not resolved by the reanalysis data or that arise because of random noise in the CALIOP signals. Scenes satisfying this requirement for CALIOP STS detections we denote as `LIQ`, and for CALIOP ice detections we denote as `ICE`.

3.1 Ice cloud equilibrium

The frost point temperature (FPT, T_{ice}) variation over the lower stratosphere is shown in Figure 1(a) for 5 ppmv H_2O. The FPT derivative, shown in Figure 1(b), indicates that a change in H_2O of 1 ppmv results in a change in FPT of $\sim$1.05–1.2 K. The FPT derivative is used to convert the estimated MLS H_2O single-profile retrieval errors (Livesey et al., 2017) into an equivalent error in FPT, and the result is shown in Figure 1(c). The accuracy and precision are estimated to be 0.6–0.3 K and 0.3 K from 68–21 hPa, respectively. The root mean square error (RMS) is the quadrature addition of the accuracy and precision. Hence the estimated RMS errors in FPT arising from the uncertainties in the MLS H_2O measurements from 68–21 hPa are 0.7–0.4 K. As noted this is the error associated with a single measurement; although the statistical precision can be assumed to diminish as the square root of the number of measurements used for averaging, the accuracy remains as the overriding error source. In any case, the error in FPT is substantially lower than the estimated RMS errors for the MLS lower stratosphere temperature measurements (Section 2.3).

3.2 Supercooled ternary solution (STS) equilibrium

Sample STS equilibrium curves are calculated using the analytic formula of Carslaw et al. (1995) and shown in Figure 2(a) for a range of lower stratospheric conditions; pressures 68, 46, and 32 hPa; H_2O mixing ratios 4, 5, and 6 ppmv; and H_2SO_4 mixing ratios 0.1 and 0.5 ppbv. Figure 2(b) shows the effect of removing the water vapor partial pressure variation through a coordinate transformation of the temperature scale relative to T_{ice}. In Figure 2(c) we show that the estimated accuracy and precision in T_{eq} arising from the uncertainties in the MLS HNO_3 measurements from 68–21 hPa are 0.6–1.5 K and 0.4 K, respectively. The corresponding RMS errors are 0.7–1.6 K. This is greater than the FPT error, but still lower than the estimated RMS errors for the MLS lower stratosphere temperature measurements (Section 2.3).

4 Results

In order to provide context for the later results that use more complex analysis techniques, we first provide a basic overview of the reanalysis temperatures in the polar lower stratosphere, and we choose as a suitable metric the daily (12UT) mean 60° polar cap temperature differences at 46 hPa for the 2008–2013 time frame. Figure 3 shows the timeseries of these temperature differences calculated for MERRA, MERRA-2, JRA55 and NCEP-CFSR relative to ERA-I, and smoothed using a 10-day boxcar average. Both hemispheres reveal underlying annual cycles with positive (negative) deviations in the summer (winter). In the Antarctic, MERRA displays the largest and most rapid excursions of up to 0.9 K from ERA-I, recurring in September through October every year, whereas in the Arctic the temperature excursions for all reanalyses are confined to within 0.5 K from ERA-I. The gray shaded regions in Figure 3 indicate the measurement periods we used for the main statistical analysis, and although these were chosen in advance, based on capturing the bulk of the PSC formation, by chance they are synchronous with more stable plateau regions in the temperature differences for the Antarctic. Therefore to achieve valid comparisons with results from other investigations, one must first acertain their compatibility based on their time overlap. We return to this theme

in Section 4.5 concerning comparisons of the reanalyses with Antarctic measurements by superpressure balloons (Hoffmann et al., 2017a).

4.1 Temperature fluctuations

Mesoscale temperature fluctuations have a significant effect on the formation of PSCs (Murphy and Gary, 1995; Dörnbrack and Leutbecher, 2001; Gary, 2006), and therefore the capability of models to accurately mimic PSC formation requires consideration of these sub-grid processes (Engel et al., 2014; Hoyle et al., 2013). The high vertical resolution of the COSMIC temperatures allows an examination of the spectrum of the temperature variance over the height region of PSC formation. For this aspect of the analysis, we interpolated (in log pressure) the COSMIC temperature profiles to a set of fixed pressure levels with 24 levels per decade ($p = 1000 \times 10^{-i/24}; i = 0, \ldots, 49$, approximately 2/3 km vertical spacing). We have analysed the atmospheric temperature variability in the 80 to 20 hPa range by fitting cubic polynomials to the COSMIC profiles and calculating the mean variance, $\overline{T'^2}$, over this range. The cubic fit provides data smoothing to estimate the unperturbed background temperature state (Whiteway, 1999; Knudsen, 2003), and therefore the mean variance is a statistic that serves as a simple measure of the vertical wave disturbances. The orientation of the COSMIC line of sight to the planes of the gravity wave fronts (Preusse et al., 2002) has an influence on the detected wave amplitudes and hence the variances. No correction for this effect is attempted.

The mean temperature variance relative to the background unperturbed temperature state, i.e. the mean fractional variance $\overline{\left(\frac{T'}{\bar{T}}\right)^2}$, is related to the potential energy density of the wave disturbance, E_p in J kg^{-1}, by the relation (Whiteway, 1999; Tsuda et al., 2000):

$$E_p = \frac{1}{2}\left(\frac{g}{N}\right)^2 \overline{\left(\frac{T'}{\bar{T}}\right)^2} \tag{3}$$

where g is the gravitational acceleration, T' is the perturbation amplitude, $\bar{T}$ is the mean temperature, and N is the bouyancy frequency. Wave activity was categorized by Whiteway (1999) and Knudsen (2003) as inhibited for $E_p < 1$ J kg^{-1} and enhanced for $E_p > 2$ J kg^{-1}. Assuming typical values for the polar lower stratosphere, $N \sim 0.02$ s^{-1}, $\bar{T} = 200$ K, we find that an energy density of E_p =1.5 J kg^{-1} corresponds to $\overline{T'^2}$ =0.5 K^2, and we have used this value as a convenient variance threshold to delineate the transition between quiescent and enhanced wave activity.

The mean (2008–2013) COSMIC vertical temperature variances for the southern and northern hemisphere winters are shown in Figure 4 for temperatures below 200 K. The Arctic temperature variance spectrum in Figure 4(d) shows a longer tail, extending well beyond 1 K^2, than for the Antarctic in Figure 4(a). The geographic distributions for the profiles with variances exceeding the 0.5 K^2 threshold are shown in Figures 4(b,e). The signature of orographically forced gravity waves over the Palmer Peninsula is captured by the variance statistic, even though it is not specifically tuned to detect orographic waves. Similarly, in the Arctic a slight increase in variance can be seen along the East coast of Greenland. These spatial imprints of gravity wave activity are largely consistent with the patterns identified by the AIRS 15 µm brightness temperature variations (Hoffmann et al., 2017b) and the Antarctic winter monthly-mean E_p patterns also derived from COSMIC GPS-RO (Hindley et al., 2015). However, since any kind of wave activity giving rise to vertical temperature fluctuations will increase the cal-

culated variance, non-orographic features are also visible in these COSMIC temperature variance maps. Figures 4(a) and (d) indicate that around 4% and 12% of the COSMIC profiles exceed the 0.5 K^2 (E_p =1.5 J kg^{-1}) threshold in the southern and northern hemispheres, respectively.

Small-scale temperature fluctuations cannot be captured accurately by the reanalysis data because of their limited spatial resolution. For example, a study of superpressure balloon measurements by Hoffmann et al. (2017a) found that ERA-I, MERRA and MERRA-2 reproduced about 30% of the standard deviation of the balloon temperature fluctuations, whereas the lower spatial resolution NCEP/NCAR reanalysis reproduced only 15% and the higher resolution ECMWF operational analysis reproduced 60%. ECMWF analyses also underestimate both gravity wave momentum fluxes derived from the balloon measurements (Jewtoukoff et al., 2015) and wave amplitudes derived from Aqua AIRS (Hoffmann et al., 2017b). Similarly, the COSMIC data should perform better than the reanalyses because of their higher resolution. The standard deviation of the differences between the reanalysis and COSMIC temperatures is the result of the quadrature addition of their individual standard deviations. The effect of removing the colocated profiles that show gravity wave activity should be to decrease the spread in the temperature differences of reanalysis data compared to COSMIC. This is seen to be the case in Figures 4(c) and (f), where the standard deviations of the temperature differences between ERA-I and COSMIC improve successively as the more highly fluctuating profiles (as determined by the COSMIC mean vertical temperature variances) that are less well captured by the ERA-I reanalysis scheme are excluded. The mean differences are also seen to improve substantially for the Arctic, but not for the Antarctic. Another effect to consider is that this method also reduces the variability of the atmospheric path along the COSMIC line of sight, and consequently the fidelity of the COSMIC temperatures to the true atmosphere will also be improved. Therefore, we cannot attribute the entirety of the observed reduction in the standard deviation to the ERA-I reanalysis data alone.

We have analyzed the other reanalyses with respect to the mean and standard deviations from the COSMIC temperatures, also using the COSMIC estimate of E_p as an effective means of applying low-pass filtering to reveal the large-scale temperature structure as described above for ERA-I. Similar to the results in Figures 4(c) and (f) for ERA-I, we also find smaller standard deviations for the other reanalyses for the low-pass filtering cases (not shown).

In the main statistical analysis of the large-scale reanalysis temperature comparisons with COSMIC we have applied a low-pass filter by excluding profile matches with COSMIC temperature variance >0.5 K^2 (E_p >1.5 J kg^{-1}). We have also restricted the COSMIC temperatures to below 200 K to better match the regions of potential PSC formation.

4.2 Orographic gravity wave case study

In the next section we report on the statistical analysis of the temperature differences, but first we examine a case study taken from a synoptic cooling period coupled with a gravity wave event on 28 August 2008 (2008d241) over the Palmer Peninsula region as shown in Figure 5. The case study serves to illustrate the quite large differences in the representation of wave structure between the reanalyses and the dependence on the spatial resolution. The 50 hPa maps (Figure 5(a)) show the reanalysis temperatures at 18UT on 28 August 2008, and a large variation in the wave activity is seen, with JRA-55, ERA-I and NCEP having much larger amplitudes than GEOS-5.9.1, MERRA and MERRA-2. A longitude section from -70° to -50° at

latitude 72° S is selected for further study. The mountain wave is forced by flow over the topographic ridge along the Palmer Peninsula, and the wind direction is almost orthogonal to the ridge line, with wind speeds $u \sim$ 25–35 ms^{-1} corresponding to the observation of ice clouds in the pressure range 80–40 hPa. The vertical wavelength, λ_z, is about 7–8 km, with a temperature amplitude of $\pm$12 K determined by MLS. PSC ice formation detected by CALIOP (not shown) correlates with the cold phases of the gravity wave, but not with all regions colder than the frost point. The horizontal wavelength, $\lambda_h \sim$ 400–500 km, is resolved by AIRS (not shown) using the techniques of Eckermann et al. (2009). From the linear gravity wave dispersion relation, we obtain the vertical wavelength, $\lambda_z = 2\pi u/N$, in the range 7-11 km, in agreement with that observed.

Figure 5b shows the temperature variation at five pressure levels (100, 70, 50, 30 and 20 hPa) along the longitude section. Temperatures below the ice frost point, calculated using MLS H_2O, are highlighted by square symbols, and the mountain terrain is indicated by the black shading. At 50 hPa the peak to trough temperature differences are 7.4, 6.2. 8.0, 10.0, 6.1 and 14.9 K for GEOS-5.9.1, MERRA, MERRA-2, JRA-55, ERA-I and NCEP-CFSR, respectively; i.e., over a factor of two difference is seen in the temperature amplitudes between the reanalyses. These differences in the wave amplitudes are rather well correlated with the spatial resolutions of the reanalyses (see Table 1). Figures 5(c) and (d) show the time variation of the gravity wave induced temperature disturbance during 28-29 August, along the longitude section at 50 hPa and 30 hPa, respectively. It is clear from these figures that the height variation, timing of the onset and extent of temperatures cold enough to potentially form ice PSCs are quite different in each of the reanalyses. The lowest temperatures are below the ice frost point at every time stamp for the 30-hPa NCEP-CFSR data, whereas for ERA-I the temperatures are only sporadically below the ice frost point. In this particular case, a PSC model run based on the NCEP-CFSR temperatures would clearly produce much larger, and longer lasting, ice volumes than one using the ERA-I temperatures.

In the rest of this paper we concentrate on the large-scale temperatures, facilitated by the effective low-pass filtering using the E_p threshold to remove gravity-wave fluctuations. However, further work on issues related to the representation of gravity waves in the reanalyses would be worthwhile, and the high-pass side of the E_p threshold may provide useful insights.

4.3 Reanalysis temperatures compared to COSMIC GPS RO

Distributions of the temperature differences between the reanalyses and COSMIC were compiled for the polar regions for the 100–10 hPa pressure range. The structure of the temperature differences has been obtained from summary statistics consisting of zonal medians and standard deviations calculated over the 2008–2013 period (Figure 6). Both JRA-55 and NCEP-CFSR show abrupt vertical transitions in the median differences above 21 hPa in both hemispheres. This is likely to be an interpolation error in the case of NCEP-CFSR since the 14.7 hPa common vertical grid level is distant from the closest native vertical grid points at 20 and 10 hPa. However, for JRA-55 the native vertical grid has a higher vertical resolution and the interpolation error is small. The standard deviations for the Arctic are similar in zonal structure for all the reanalyses, but the values are systematically smaller in the Antarctic.

Figure 7 is a summary of the temperature bias ranges, extracted from the data shown in Figure 6, that readily captures the salient information over the pressure range from 68 to 21 hPa, which encompasses the bulk of the PSC vertical formation region (Pitts et al., 2009). The main attributes of the reanalyses are revealed by the relative size of the boxes and their location

relative to the zero line. The box width is set to the largest difference in minimum to maximum bias at any height (68–21 hPa), and the box height is set to the largest difference in minimum to maximum meridional bias. The MERRA boxes are larger than those for any of the other reanalyses, with widths (heights) of 0.54 K (0.54 K) in the Antarctic and 0.86 K (0.50 K) in the Arctic. The smallest box widths are for Arctic NCEP-CFSR (0.26 K) and Antarctic JRA-55 (0.26 K). The smallest box heights are for Arctic NCEP-CFSR (0.18 K) and Antarctic MERRA-2 (0.14 K).

For both polar regions the reanalysis temperature differences from COSMIC are encompassed within a quite narrow range of -0.61 to $+0.48$ K. GEOS-5.9.1, MERRA, MERRA-2 and JRA-55 have cold biases (-0.61 to -0.05 K) in the Antarctic, whereas ERA-I has a warm bias ($+0.04$ to $+0.38$ K). The NCEP-CFSR bias (-0.30 to $+0.16$ K) changes with decreasing pressure from warm to cold. Overall, the average Arctic temperature biases for individual reanalyses with respect to COSMIC are shifted by $\sim$0.1–0.4 K from the Antarctic values.

4.4 PSC types from CALIOP and their representation in MLS observations

Figure 8 is an example of the probability density distribution, presented as a ternary graph, of the relative proportion of the three groups of CALIOP PSC types, `STS`, `MIX` and `ICE` occurring within the MLS geometric field of view. Here we define the congener names `MIX`, for the sum of the `MIX1`, `MIX2`, and `MIX2-enh` types, and `ICE`, which also includes wave-ice. This figure is compiled by accumulating the co-existence frequency of these three groups of PSCs for the 2013 Antarctic winter on the 46 hPa pressure level for temperatures below the ice frost point (as determined by ERA-I). The three axes lie along the sides of an equilateral triangle, and the three components are constrained to sum to unity. At the center of any side the percentage of the two adjoining PSC types is 50:50, and the third type at the opposite vertex is zero. At the center of the triangle the proportion of all three PSC types is equal to 1/3. The most populated areas consist of two branches close to the `MIX/STS` and `MIX/ICE` axes. Away from these two branches, the `STS/ICE` region with less than 10% `MIX` is essentially devoid of data. This graph is a striking demonstration that the co-existence of STS and ice PSCs on a spatial scale of the size of the MLS geometric field of view is relatively rare. The implication is that the rapidity of freezing of STS to form ice and the resulting re-equilibration of the water vapor and cloud ice is faster than can be captured by the satellite instrument integration times. The A-train orbital speed is $\sim$7.5 km s^{-1}, and hence the 165 km along-track MLS FOV is traversed in about 22 s.

We illustrate the concept of the field of view fill-fraction in Figures 9 and 10. Figure 9 shows CALIOP and MLS data recorded over a 7000 km transect over Antarctica on 23 June 2008 (2008d175). The CALIOP PSC types show large contiguous areas of STS and ice, but with other embedded PSC types causing visual speckle as mentioned above (Figure 9(a)). The MLS temperatures are up to 4 K below the frost point (Figure 9(b)). Large regions of HNO_3 depletion are seen where abundances have been reduced to practically zero values by sequestration and denitrification by PSCs compared to background values of over 12 ppbv (Figure 9(c)). A smaller dehydrated region has H_2O values 1-2 ppmv lower than the $\sim$5 ppmv background (Figure 9(d)). The second and third rows in the figure show expanded views of a selected region at 32 hPa centered on a particular MLS geometric field of view. At each magnification step the CALIOP data reveal increasingly finer detail down to the 5-km by 0.180 km pixel size, whereas the MLS data become over-sampled. Within the highlighted MLS geometric field of view, all CALIOP PSC types except wave ice are present, consisting of STS:68.9%, ICE:1.6%, MIX1:0.6%, MIX2:12.9%,

and MIX2E:13.2%. Clear sky cases make up the remaining 2.8%. We impose an upper limit on the occupancy fraction of cloud-free pixels of 0.25 and require the fraction of the dominant LIQ or ICE PSCs to be greater than 0.75. The example scene has too great a variation of PSC types within the MLS field of view and fails this acceptance test since it only has a 0.69 LIQ fill-fraction. Orbit transects of the MLS FOV PSC fill-fraction for the different PSC types are shown in Figure 10. The LIQ and ICE filled FOV threshold criteria are mutually exclusive. Despite the PSCs extending over a few thousand km, the strict acceptance criteria, albeit essential to ensure a more uniform single PSC type across the MLS FOV, severely reduces the total number of scenes available for analysis. Only about 15% of the scenes containing STS or ice PSCs are accepted.

Figure 11 shows scatter plots of the coincident MLS HNO_3 and H_2O for Antarctic PSCs classified by CALIOP vs ERA-I temperature at 46 hPa in 2013. Figures 11(a) and (c) show the scatter of HNO_3 vs temperature and inverse temperature, respectively, along with the theoretical HNO_3 gas-phase uptake curves for STS and NAT, indicating clearly that LIQ PSCs are closely associated with the STS curve. Similarly, Figures 11(b) and (d) show the scatter of gas-phase H_2O and its close association with the frost point temperature in the presence of ICE PSCs.

In Figure 12 we show the temperature variation of the estimated partial pressure of water vapor over ice, p_{ice}, for atmospheric pressure levels in the range 100–10 hPa. The partial pressure is determined simply from the product of the MLS H_2O volume mixing ratio and the atmospheric pressure. Since CALIOP indicates the presence of coincident ice PSCs, this is an estimate of the partial pressure of water vapor over ice. The bias in the MLS temperatures is evident in the departure of the p_{ice} observations from the theoretical equilibrium curve in Figure 12(a), whereas the ERA-I temperatures place the p_{ice} observations in much better agreement.

4.5 Reanalysis temperatures compared to LIQ and ICE reference points

In Figure 13 we show the variation of the MLS gas-phase HNO_3 with ERA-I temperature corresponding to CALIOP PSC classifications at 32 hPa for one Southern Hemisphere winter. The MLS HNO_3 data are separated into corresponding CALIOP PSC categories, allowing comparison of 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 13(a) for LIQ and ICE PSCs. The HNO_3 gas-phase uptake in the presence of liquid-phase LIQ PSCs follows the STS equilibrium curve. In contrast, the HNO_3 abundance is very low in the presence of ICE PSCs. In Figure 13(b), the HNO_3 gas-phase uptake in clouds associated with the CALIOP solid NAT type MIX2 shows significant non-equilibrium variation and lies between the STS and NAT equilibrium curves. Histograms of the temperature distributions of (a) and (b) are shown in Figure 13(c). The light blue (LIQ) and yellow (MIX2) colored histograms are for HNO_3 mixing ratios >1 ppbv, whereas the red (LIQ) colored histogram is for HNO_3 mixing ratios <1 ppbv. The tails of the temperature distributions for LIQ (light blue and red histograms) do not extend to temperatures lower than those in the distribution for ICE (dark blue), and no peaks are observed that would indicate the existence of PSCs at a frost-point depression near $T_{ice} - 3$ K. In Figure 13(d), the temperatures are 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. Note that the LIQ histogram is shifted and becomes much

narrower (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 for each reanalysis data set of the temperature difference distributions were generated from histograms analogous to those in Figure 13(d) for each year and pressure level. Statistics of the combined data for the 2008–2013 Antarctic winters at 46 hPa are shown in Tables 2 and 3. The distributions for LIQ PSCs (Table 2) all show negative temperature biases compared to the STS equilibrium reference, with mean values in the range -0.9 to -0.3 K and standard deviations $\sim$0.6 K. The standard error of the mean is $\sim$0.01 K, and the medians are very similar to the mean values, indicating a small skewness of the distribution. The scatter follows a normal distribution reasonably well in the tails of the distribution, but with a mismatch near the peak of the distribution (not shown). The normalized χ^2 values are in the range 1.4–2.2, which for around 30 degrees of freedom (i.e. the difference between the number of histogram bins and the number of fitted parameters) indicates a poor goodness-of-fit. We note that improved fits can be obtained by introducing a secondary normal distribution with a smaller standard deviation (not shown), which seems to imply the underlying presence of two distinct modes of temperature precision in the reanalyses. We have not investigated this finding any further, but one pertinent avenue of exploration, given that the density of observations ingested into the system is sparse in the polar regions, would be to look for evidence that the reanalysis temperatures near grid points that are located closer to the geographic source of assimilated observations have better precision than those at grid points remote from observations.

The distributions for ICE PSCs (Table 3) show mainly negative temperature biases compared to the equilibrium ice frost point reference. Mean temperature bias values for ICE PSCs are in the range -0.6 to 0 K with standard deviations of $\sim$0.7 K, both of which are larger than their LIQ PSC counterparts. The ICE distribution fits were also improved in some cases by introducing a secondary normal distribution.

Intercomparisons of the reanalysis temperature statistics are displayed in Figure 14 for the Antarctic in 2013 (a representative year). In addition to the differences relative to the LIQ and ICE reference points, we also show the comparison with the COSMIC temperatures (latitude range 90° S to 60° S, mean variance $<$0.5 K^2 and COSMIC temperatures $<$200 K). The largest bias is for MLS, and biases for the LIQ reference are consistently about 0.5 K more negative than for ICE. Biases for the COSMIC reference are smaller than those for LIQ or ICE. MERRA does not assimilate the COSMIC GPS RO data, and it has the largest bias with respect to the COSMIC reference. Standard deviations for LIQ are consistently smaller than those for ICE, and standard deviations for COSMIC are between those of LIQ and ICE. Interannual variability for the 46 hPa level is shown in Figure 15, where median values for the COSMIC reference display less variability than those for the LIQ or ICE references. Two years (2010 and 2012) stand out in the LIQ median reference as anomalously high, and we note that those Antarctic winters were warmer than in the other years (Kuttippurath et al., 2015). Standard deviations show less interannual variability than median values, especially for the COSMIC reference.

The vertical temperature differences over the 100–10 hPa range are shown in Figure 16 for the Antarctic winters 2008–2013. Median values for the ICE reference are more constant with height than those for the LIQ reference, which become more negative with increasing height. Median values for the LIQ reference are biased lower than those for ICE by $\sim$0.5 K, although standard deviations for the LIQ reference are smaller than those for ICE. The corresponding observations for the Arctic are

shown in Figure 17, but there the number of data points in the ICE reference is much lower, and the lack of sufficient ice cloud production in the Arctic precludes a robust conclusion for ICE PSCs. Again, MLS shows the largest bias.

Das and Pan (2014) intercompared COSMIC GPS RO temperatures with satellite measurements from the Sounding of the Atmosphere using Broadband Emission Radiometry (SABER) instrument and Aura MLS and determined that the structure of the observed median temperature differences was due in large part to inherent retrieval biases of SABER and MLS (December 2010 to November 2011). Seasonal and meridional bins were used to compare with SABER and MLS. At high latitudes, in the 100–10 hPa range, median profiles for SABER showed an obvious positive bias (2–3 K higher than COSMIC) in the lower stratosphere, increasing with decreasing atmospheric pressure. Median profiles for MLS showed negative biases of 0 to 2 K, but with larger oscillations than seen with SABER. These results indicate the difficulty of obtaining sub-kelvin temperature accuracy with microwave and infrared limb sounders.

The standard deviations result from the combination of the quadrature addition of the precisions of the reanalyses and the reference points. In Figures 1 and 2, we estimated the minimum precision of the ICE and LIQ references to be $\sim$0.3 K. For the LIQ reference comparison (Figures 16(d) and 17(d)), the minimum combined precision is 0.5 K (in both NH and SH); therefore, subtracting the LIQ reference precision results in an estimated precision of 0.4 K for the reanalysis data. The theoretical vertical structure of the COSMIC GPS RO standard deviation values follows an exponential increase with height as shown by Scherllin-Pirscher et al. (2011b) and is defined by Equation 1. We have determined the best fit parameters for each hemisphere from Equation 1 by minimizing the RMS deviations of the ensemble means of the (reanalysis $-$ COSMIC) standard deviations. The solid black lines in Figures 16(f) and 17(f) are the quadrature addition of the estimated reanalysis temperature error, $\sigma_{T_{RE}} = 0.4$ K, and the GPS RO model error given in Equation 1 with best fit parameters $s_0 = (0.83, 0.69)$ K, $H_s = (10.1, 8.8)$ km, and $Z_S = (22.5, 23.1)$ km for the Arctic and Antarctic, respectively; i.e. our values for s_0 are close to the 0.7 K value given by Scherllin-Pirscher et al. (2011b).

The MERRA temperatures should be uncorrelated with the COSMIC data, and in the main they do show a higher combined standard deviation than the other reanalyses. Some correlation with the other reanalyses that assimilate COSMIC data may be expected to reduce the standard deviation, and this appears to be the case except for the top two pressure levels.

For the southern hemisphere, the LIQ reanalysis standard deviation (Figure 16(d)) appears to be approximately constant from 46 to 14 hPa, rising at the lower stratospheric levels, whereas the northern hemisphere (Figure 17(d)) shows no increase at the lower levels. Overall the combined standard deviations for the reanalyses and COSMIC data are higher in the Arctic (Figure 17(f)) than in the Antarctic (Figure 16(f)) by $\sim$0.1 K.

Recently, Hoffmann et al. (2017a) used superpressure balloon measurements made during the Antarctic Concordiasi campaign in September 2010 to January 2011 to validate meteorological analyses and reanalyses. Over the flight paths of the balloons (17–18.5 km (58.2–69.1 hPa) and 60° S – 85° S), Hoffmann et al. found warm biases ranging from 0.6 to 1.5 K for ERA-I, MERRA and MERRA-2, increasing towards the pole, and standard deviations of 0.5–0.8 K, with ERA-I showing smaller standard deviations than MERRA. Whereas Hoffmann et al. find MERRA to be the warmest and ERA-I the coldest, we find the opposite order compared to the COSMIC and thermodynamic temperature references (see Figure 16). Also we find negative temperature biases for MERRA and MERRA-2 and only a slight positive bias for ERA-I, with no significant latitu-

dinal gradients (Figure 6). To reconcile this apparent discrepancy we show daily mean temperatures (at 12UT) for MERRA and MERRA-2 relative to ERA-I in Figure 18 and highlight the non-overlapping time periods of the PSC analysis window (green line) and the balloon flights (red line). Clearly, measurements in the later time period of the balloon flights (September-December) sample different atmospheric conditions than experienced in the earlier time period (May-August). Differences between reanalysis temperatures along the individual balloon trajectories are likely to be amplified compared to the differences in mean polar cap temperatures.

As we noted previously MERRA does not assimilate COSMIC data, whereas MERRA-2 and the other reanalyses we investigated do assimilate COSMIC data, and hence some of the reduction in the bias of MERRA-2 compared to MERRA seen in Figure 18 is likely to be attributable to the use of GPS RO data in the former.

Large vertical oscillations of up to $\pm$2–3 K were reported in the differences between ECMWF operational analyses and the CHallenging Minisatellite Payload (CHAMP) RO data in the 2002 to 2006 June-August Antarctic polar vortices (Gobiet et al., 2005, 2007). Similar height-dependent features were also found (Parrondo et al., 2007) in comparisons of ECMWF analyses with radiosondes. The value of the RO dataset as a reference was ably demonstrated by comparisons of CHAMP RO with ECMWF and other meteorological data (GEOS-4, MetO, NCEP/CPC, NCEP Rean) for June-August 2003 (Gobiet et al., 2007). None of the other dataset differences showed features consistent with ECMWF, nor were any two consistent with each other, e.g. GEOS-4 displayed an Antarctic cold bias relative to CHAMP at 25–30 km, whereas NCEP Rean displayed a warm bias in roughly the same range. The absence of spurious vertical features in the comparisons we have presented here underlines the general much closer agreement amongst the more modern reanalysis systems.

Although we have not directly matched A-train locations with the COSMIC occultations, we can estimate the differences between the COSMIC temperatures and the thermodynamic reference temperatures by elimination of the reanalysis temperature biases as follows. First, we decompose the reanalysis, radio occultation and equilibrium temperatures into $\mathrm{T_x} = \mathrm{T'} + \Delta\mathrm{T_x}$, where $\mathrm{T'}$ is the true temperature, and the temperature bias is $\Delta\mathrm{T_x}$, such that $\mathrm{x} = \mathrm{re, ro, eq}$ for the reanalysis, radio occultation and equilibria, respectively. Then, we have $\overline{(\mathrm{T_{re}} - \mathrm{T_{ro}})} = \overline{\Delta\mathrm{T_{re}}} - \overline{\Delta\mathrm{T_{ro}}}$ and $\overline{(\mathrm{T_{re}} - \mathrm{T_{eq}})} = \overline{\Delta\mathrm{T_{re}}} - \overline{\Delta\mathrm{T_{eq}}}$. With the tacit implication that the $\overline{\Delta\mathrm{T_{re}}}$ are unchanging and can be eliminated from both equations, we derive $\overline{(\mathrm{T_{re}} - \mathrm{T_{eq}})} - \overline{(\mathrm{T_{re}} - \mathrm{T_{ro}})} = \overline{\Delta\mathrm{T_{ro}}} - \overline{\Delta\mathrm{T_{eq}}}$. The results of these operations are shown in Figure 19 for `LIQ` and COSMIC, $\overline{\Delta\mathrm{T_{LIQ}}} - \overline{\Delta\mathrm{T_{COSMIC}}}$, and `ICE` and COSMIC, $\overline{\Delta\mathrm{T_{ICE}}} - \overline{\Delta\mathrm{T_{COSMIC}}}$. The temperature bias difference profiles for the different reanalyses are very closely grouped, especially over the pressure range 68–21 hPa, justifying the assumption that $\overline{\Delta\mathrm{T_{re}}}$ can be eliminated. The biases for ($\overline{\Delta\mathrm{T_{LIQ}}} - \overline{\Delta\mathrm{T_{COSMIC}}}$) are similar in their shape and value for both hemispheres over the pressure range 68–21 hPa, with $\overline{\Delta\mathrm{T_{LIQ}}}$ around 0.5 to 1.0 K lower than $\overline{\Delta\mathrm{T_{COSMIC}}}$. The biases for $\overline{\Delta\mathrm{T_{ICE}}}$ are around 0 to 0.5 K lower than $\overline{\Delta\mathrm{T_{COSMIC}}}$ for the Antarctic, but the Arctic has too few data points to make a meaningful comparison. The `LIQ` and `ICE` profiles for the Antarctic have similar values at 32 hPa with $\overline{\Delta\mathrm{T_{ICE}}} - \overline{\Delta\mathrm{T_{COSMIC}}} \sim -0.5$ K, whereas the `LIQ` values diverge from those of `ICE` by up to -1.5 K above and below this level. The reasons for these `LIQ` and `ICE` systematic discrepancies are not known; however, the differences are within the expected uncertainty limits in the 68–21 hPa pressure range.

Finally, a summary of the mean temperature bias ranges of the reanalyses relative to the `LIQ` (-1.6 to -0.3 K) and `ICE` (-0.9 to $+0.1$ K) equilibrium references and COSMIC (-0.5 to $+0.2$ K) is shown in Figure 20. The ranges quoted are for

the pressure range 68–21 hPa. For all reference points, the coldest reanalysis biases tend to be found in the Antarctic and the warmest in the Arctic.

5 Conclusions

We have evaluated the accuracy and precision of several contemporary reanalysis data sets compared to (a) the COSMIC GPS RO temperatures and (b) the absolute temperature references obtained from the equilibrium properties of certain types of PSC. In the polar regions, for temperatures below 200 K, the range in the mean biases of the reanalyses with respect to COSMIC temperatures is only -0.61 to $+0.48$ K. The lack of assimilated GPS RO data in MERRA is more evident in the higher biases for this data set with respect to COSMIC temperatures in the Antarctic (i.e. where there is a relative paucity of conventional measurements) compared to the other reanalyses. Significantly larger negative biases, except at 100 hPa, and a vertical oscillation are seen in the MLS temperatures. Standard deviations are similar for the reanalyses, but the standard deviations for MLS temperature are substantially larger.

The extent to which the equilibrium thermodynamics of STS and ice PSCs can be used as absolute temperature references has been explored in detail. 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. The corresponding estimated measurement accuracies are in the range 0.7–1.6 K for STS and 0.4–0.7 K for ice. The resulting RMS uncertainites are smaller than those derived for the MLS retrieved temperatures and comparable to the measurement capabilities of the GPS RO technique (accuracy$<$0.2 K, precision $>$0.7 K) in the lower stratosphere. The reanalysis temperatures were found to be lower than the absolute reference points by 0.3 to 1 K for `LIQ` and 0 to 1 K for `ICE` at the peak heights of PSC occurrence (68–32 hPa). Vertical profiles for `LIQ` show larger negative deviations above 32 hPa than below that level, and also compared to `ICE`. Medians for `LIQ` are consistently biased lower than those for `ICE` by $\sim$0.5 K. On the 46 hPa pressure level, medians of the reanalyses all depart from zero, and their scatter falls within the range of about 0.6 K for `LIQ` and 0.5 K for `ICE`. Although the biases are larger for `LIQ` than for `ICE`, the standard deviations for `LIQ` ($\sim$0.6 K) are smaller than those for `ICE` ($\sim$0.7 K).

To put these `LIQ` and `ICE` reference temperatures into context with other independent polar temperature measurements, it is instructive to compare them to the temperature measurement errors from long-duration balloon flights, which have typical nighttime biases of 0.5 K, precisions of 0.4 K (Pommereau et al., 2002) and measured standard deviations of 1.0 to 1.3 K for temperature differences with respect to ECMWF operational analyses (Knudsen et al., 2002).

The polar temperatures from several contemporary reanalyses are in much better agreement than were the reanalyses from previous decades. Hence, in explaining any systematic deficiencies in modeled chlorine activation and/or modeled ozone losses based on these reanalyses, the burden shifts to finding alternative explanations other than the arbitrary adjustment of the reanalysis temperatures by as much as 1–2 K to offset such discrepancies. However, even though the reanalyis temperature differences are in general in very good agreement, under certain conditions such as wave-driven events, individual temperature comparisons may vary by several kelvin and are therefore important for specific case studies.

6 Data availability

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

CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center (CALIOP L1; CALIOP L2).

COSMIC data were obtained from the University Corporation for Atmospheric Research COSMIC Data Analysis and Archive Center (COSMIC)

GEOS5.9.1 data were obtained from the Goddard Earth Sciences Data and Information Services Center (GEOS5.9.1).

MERRA data were obtained from the Goddard Earth Sciences Data and Information Services Center (MERRA).

MERRA-2 data were obtained from the Goddard Earth Sciences Data and Information Services Center (MERRA2).

JRA-55 data were obtained from the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory JRA-55.

ERA-I data were obtained from the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory (ECMWF)

NCEP-CFSR data were obtained from the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory (NCEP-CFSR).

IDL software for calculation of PSC thermodynamic properties provided by M. E. Hervig was obtained from the GATS Scientific Software 95 website (http://gwest.gats-inc.com/software/software_page.html).

Acknowledgements. We gratefully acknowledge members of the teams associated with the CALIOP and MLS instruments, and the reanalysis data centers. MLS data are archived at the NASA Goddard Earth Sciences Data Information and Services Center. CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. IDL software for calculation of PSC thermodynamic properties provided by M. E. Hervig was obtained from the GATS Scientific Software website (http://gwest.gats-inc.com/software/software_page.html). Work at the Jet Propulsion Laboratory, California Institute of Technology, was carried out under a contract with the National Aeronautics and Space Administration. This research was conducted as a component of the SPARC (Stratosphere-troposphere Processes And their Role in Climate) S-RIP activity, under the guidance and sponsorship of the World Climate Research Programme.

References

Alexander, P., de la Torre, A., Llamedo, P., and Hierro, R.: Precision estimation in temperature and refractivity profiles retrieved by GPS radio occultations, J. Geophys. Res., 119, 8624–8638, https://doi.org/10.1002/2013JD021016, 2014.

Anthes, R. A.: Exploring Earth's atmosphere with radio occultation: contributions to weather, climate and space weather, Atmospheric Measurement Techniques, 4, 1077–1103, https://doi.org/10.5194/amt-4-1077-2011, 2011.

Bauer, P., Radnóti, G., Healy, S., and Cardinali, C.: GNSS Radio Occultation Constellation Observing System Experiments, Mon. Weather Rev, 142, 555–572, https://doi.org/10.1175/MWR-D-13-00130.1, 2014.

Bonavita, M.: On some aspects of the impact of GPSRO observations in global numerical weather prediction, Quarterly Journal of the Royal Meteorological Society, 140, 2546–2562, https://doi.org/10.1002/qj.2320, 2014.

Brakebusch, M., Randall, C. E., Kinnison, D. E., Tilmes, S., Santee, M. L., and Manney, G. L.: Evaluation of Whole Atmosphere Community Climate Model simulations of ozone during Arctic winter 2004–2005, J. Geophys. Res., 118, 2673–2688, https://doi.org/10.1002/jgrd.50226, 2013.

Butchart, N., Charlton-Perez, A. J., Cionni, I., Hardiman, S. C., Haynes, P. H., Krüger, K., Kushner, P. J., Newman, P. A., Osprey, S. M., Perlwitz, J., Sigmond, M., Wang, L., Akiyoshi, H., Austin, J., Bekki, S., Baumgaertner, A., Braesicke, P., Brühl, C., Chipperfield, M., Dameris, M., Dhomse, S., Eyring, V., Garcia, R., Garny, H., Jöckel, P., Lamarque, J.-F., Marchand, M., Michou, M., Morgenstern, O., Nakamura, T., Pawson, S., Plummer, D., Pyle, J., Rozanov, E., Scinocca, J., Shepherd, T. G., Shibata, K., Smale, D., Teyssèdre, H., Tian, W., Waugh, D., and Yamashita, Y.: Multimodel climate and variability of the stratosphere, J. Geophys. Res., 116, https://doi.org/10.1029/2010JD014995, 2011.

CALIOP L1: 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) (last access: 6 July 2017), https://doi.org/10.5067/CALIOP/CALIPSO/CAL_LID_L1-ValStage1-V3-01_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, 2015.

CALIOP L2: CALIPSO Science Team: CALIPSO/CALIOP Level 2, Polar Stratospheric Cloud Data, versions 1.00, Hampton, VA, USA: NASA Atmospheric Science Data Center (ASDC) (last access: 6 July 2017), https://doi.org/10.5067/CALIOP/CALIPSO/CAL_LID_L2_PSCMask-Prov-V1-00_L2-001.00, 2015.

Carslaw, K. S., Luo, B. P., and Peter, T.: 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, 1995.

Carslaw, K. S., Wirth, M., Tsias, A., Luo, B. P., Dörnbrack, A., Leutbecher, M., Volkert, H., Renger, W., Bacmeister, J. T., and Peter, T.: Particle microphysics and chemistry in remotely observed mountain polar stratospheric clouds, J. Geophys. Res., 103, 5785–5796, 1998a.

Carslaw, K. S., Wirth, M., Tsias, A., Luo, B. P., Dörnbrack, A., Leutbecher, M., Volkert, H., Renger, W., Bacmeister, J. T., Reimer, E., and Peter, T.: Increased stratospheric ozone depletion due to mountain-induced atmospheric waves, Nature, 391, 675–678, 1998b.

COSMIC: 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, (last access: 6 July 2017), http://rda.ucar.edu/datasets/ds723.0/, 2013.

Danilin, M. Y., Santee, M. L., Rodriguez, J. M., Ko, M. K. W., Mergenthaler, J. L., Kumer, J. B., Tabazadeh, A., and Livesey, N.: Trajectory hunting: A case study of rapid chlorine activation in December 1992 as seen by UARS, J. Geophys. Res., 105, 4003–4018, 2000.

Das, U. and Pan, C. J.: Validation of FORMOSAT-3/COSMIC level 2 "atmPrf" global temperature data in the stratosphere, Atmos. Meas. Tech., 7, 731–742, https://doi.org/10.5194/amt-7-731-2014, 2014.

de la Torre Juárez, M., Marcus, S., Dörnbrack, A., Schroder, T. M., Kivi, R., Iijima, B. A., Hajj, G. A., and Mannucci, A. J.: Detection of temperatures conducive to Arctic polar stratospheric clouds using CHAMP and SAC-C radio occultation data, J. Geophys. Res., 114, https://doi.org/10.1029/2008JD011261, 2009.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Dee, D. P., Balmaseda, M., Balsamo, G., Engelen, G., Simmons, A. J., and Thépaut, J.-N.: Toward a Consistent Reanalysis of the Climate System, Bull. Am. Meteor. Soc., 95, 1235–1248, https://doi.org/10.1175/BAMS-D-13-00043.1, 2014.

Dörnbrack, A. and Leutbecher, M.: Relevance of mountain waves for the formation of polar stratospheric clouds over Scandinavia: A 20 year climatology, J. Geophys. Res., 106, 1583–1593, 2001.

Dörnbrack, A., Leutbecher, M., Reichardt, J., Behrendt, A., Müller, K.-P., and Baumgarten, G.: Relevance of mountain waves for the formation of polar stratospheric clouds over Scandinavia: Mesoscale dynamics and observations for January 1997, J. Geophys. Res., 106, 1569–1581, 2001.

Eckermann, S. D., Hoffmann, L., Hoepfner, M., Wu, D. L., and Alexander, M. J.: Antarctic NAT PSC belt of June 2003: Observational validation of the mountain wave seeding hypothesis, Geophys. Res. Lett., 36, L02 807, doi:10.1029/2008GL036 629, 2009.

ECMWF: European Centre for Medium-Range Weather Forecasts, ERA-Interim Project, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder CO (last access: 6 July 2017), https://doi.org/10.5065/D6CR5RD9, 2009.

Engel, I., Luo, B. P., Pitts, M. C., Poole, L. R., Hoyle, C. R., Grooß, J.-U., Dörnbrack, A., and Peter, T.: Heterogeneous formation of polar stratospheric clouds? Part 2: Nucleation of ice on synoptic scales, Atmos. Chem. Phys. Discuss., 13, 8831–8872, https://doi.org/10.5194/acpd-13-8831-2013, 2013.

Engel, I., Luo, B. P., Khaykin, S. M., Wienhold, F. G., Vömel, H., Kivi, R., Hoyle, C. R., Grooß, J.-U., Pitts, M. C., and Peter, T.: Arctic stratospheric dehydration – Part 2: Microphysical modeling, Atmos. Chem. Phys., 14, 3231–3246, https://doi.org/10.5194/acp-14-3231-2014, 2014.

Feltz, M., Knuteson, R., Ackerman, S., and Revercomb, H.: Application of GPS radio occultation to the assessment of temperature profile retrievals from microwave and infrared sounders, Atmos. Meas. Tech., 7, 3751–3762, https://doi.org/10.5194/amt-7-3751-2014, 2014a.

Feltz, M. L., Knuteson, R. O., Revercomb, H. E., and Tobin, D. C.: A methodology for the validation of temperature profiles from hyperspectral infrared sounders using GPS radio occultation: Experience with AIRS and COSMIC, J. Geophys. Res., 119, 1680–1691, https://doi.org/10.1002/2013JD020853, 2014b.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis In-

tercomparison Project (S-RIP) and overview of the reanalysis systems, Atmos. Chem. Phys., 17, 1417–1452, https://doi.org/10.5194/acp-17-1417-2017, 2017.

Gary, B. L.: Mesoscale temperature fluctuations in the stratosphere, Astroparticle Phys., 6, 4577–4589, https://doi.org/10.5194/acp-6-4577-2006, 2006.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

GEOS5.9.1: Global Modeling and Assimilation Office (GMAO): 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 (last access: 6 July 2017), 2013.

Gobiet, A., Foelsche, U., Steiner, A. K., Borsche, M., Kirchengast, G., and Wicker, J.: Climatological validation of stratospheric temperatures in ECMWF operational analyses with CHAMP radio occultation data, Geophys. Res. Lett., 32, L12 806, doi:10.1029/2005GL022 617, 2005.

Gobiet, A., Kirchengast, G., Manney, G. L., Borsche, M., Retscher, C., and Stiller, G.: 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, https://doi.org/10.5194/acp-7-3519-2007, 2007.

Harris, N. R. P., Lehmann, R., Rex, M., and von der Gathen, P.: A closer look at Arctic ozone loss and polar stratospheric clouds, Atmos. Chem. Phys., 10, 8499–8510, https://doi.org/10.5194/acp-10-8499-2010, 2010.

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

Hindley, N. P., Wright, C. J., Smith, N. D., and Mitchell, N. J.: The southern stratospheric gravity wave hot spot: individual waves and their momentum fluxes measured by COSMIC GPS-RO, Atmos. Chem. Phys., 15, 7797–7818, https://doi.org/10.5194/acp-15-7797-2015, https://www.atmos-chem-phys.net/15/7797/2015/, 2015.

Ho, S.-P., Hunt, D., Steiner, A. K., Mannucci, A. J., Kirchengast, G., Gleisner, H., Heise, S., von Engeln, A., Marquardt, C., Sokolovskiy, S., Schreiner, W., Scherllin-Pirscher, B., Ao, C., Wickert, J., Syndergaard, S., Lauritsen, K. B., Leroy, S., Kursinski, E. R., Kuo, Y.-H., Foelsche, U., Schmidt, T., and Gorbunov, M.: Reproducibility of GPS radio occultation data for climate monitoring: Profile-to-profile intercomparison of CHAMP climate records 2002 to 2008 from six data centers, J. Geophys. Res., 117, https://doi.org/10.1029/2012JD017665, 2012.

Ho, S.-P., Peng, L., and Vömel, H.: Characterization of the long-term radiosonde temperature biases in the upper troposphere and lower stratosphere using COSMIC and Metop-A/GRAS data from 2006 to 2014, Atmos. Chem. Phys., 17, 4493–4511, https://doi.org/10.5194/acp-17-4493-2017, 2017.

Hoffmann, L., Hertzog, A., Rößler, T., Stein, O., and Wu, X.: Validation of meteorological analyses and trajectories in the Antarctic lower stratosphere using Concordiasi superpressure balloon observations, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-71, 2017a.

Hoffmann, L., Spang, R., Orr, A., Alexander, M. J., Holt, L. A., and Stein, O.: A decadal satellite record of gravity wave activity in the lower stratosphere to study polar stratospheric cloud formation, Atmos. Chem. Phys., 17, 2901–2920, https://doi.org/10.5194/acp-17-2901-2017, 2017b.

Hoyle, C. R., Engel, I., Luo, B. P., Pitts, M. C., Poole, L. R., Grooß, J.-U., and Peter, T.: Heterogeneous formation of polar stratospheric clouds – Part 1: Nucleation of nitric acid trihydrate (NAT), Atmos. Chem. Phys., 13, 9577–9595, https://doi.org/10.5194/acp-13-9577-2013, 2013.

Hurst, D. F., Lambert, A., Read, W. G., Davis, S. M., Rosenlof, K. H., Hall, E. G., Jordan, A. F., and Oltmans, S. J.: Validation of Aura Microwave Limb Sounder stratospheric water vapor measurements by the NOAA frost point hygrometer, J. Geophys. Res., 119, 1612–1625, https://doi.org/10.1002/2013JD020757, 2014.

Hurst, D. F., Read, W. G., Vömel, H., Selkirk, H. B., Rosenlof, K. H., Davis, S. M., Hall, E. G., Jordan, A. F., and Oltmans, S. J.: Recent divergences in stratospheric water vapor measurements by frost point hygrometers and the Aura Microwave Limb Sounder, Atmos. Meas. Tech., 9, 4447–4457, https://doi.org/10.5194/amt-9-4447-2016, 2016.

Jewtoukoff, V., Hertzog, A., Plougonven, R., de la Cámara, A., and Lott, F.: Comparison of Gravity Waves in the Southern Hemisphere Derived from Balloon Observations and the ECMWF Analyses, J. Atmos. Sci., 72, 3449–3468, https://doi.org/10.1175/JAS-D-14-0324.1, 2015.

JRA-55: JRA-55: Japanese 55-year Reanalysis, Daily 3-Hourly and 6-Hourly Data, https://doi.org/10.5065/D6HH6H41, 2013.

Knudsen, B. M.: On the accuracy of analysed low temperatures in the stratosphere, Atmos. Chem. Phys., 3, 1759–1768, https://doi.org/10.5194/acp-3-1759-2003, 2003.

Knudsen, B. M., Pommereau, J.-P., Garnier, A., Nunes-Pinharanda, M., Denis, L., Newman, P., Letrenne, G., and Durand, M.: Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights 2. Results, J. Geophys. Res., 107, https://doi.org/10.1029/2001JD001329, 2002.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, J. Meteorol. Soc. Japan. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Koop, T., Ng, H. P., Molina, L. T., and Molina, M. J.: A new optical technique to study aerosol phase transitions: the nucleation of ice from H_2SO_4 aerosols, J. Phys. Chem. A, 102, 8924–8931, 1998.

Kuttippurath, J., Godin-Beekmann, S., Lefèvre, F., Santee, M. L., Froidevaux, L., and Hauchecorne, A.: Variability in Antarctic ozone loss in the last decade (2004–2013): high-resolution simulations compared to Aura MLS observations, Atmos. Chem. Phys., 15, 10 385–10 397, https://doi.org/10.5194/acp-15-10385-2015, 2015.

Ladstädter, F., Steiner, A. K., Schwärz, M., and Kirchengast, G.: Climate intercomparison of GPS radio occultation, RS90/92 radiosondes and GRUAN from 2002 to 2013, Atmos. Meas. Tech., 8, 1819–1834, https://doi.org/10.5194/amt-8-1819-2015, 2015.

Lambert, A., Read, W. G., Livesey, N. J., Santee, M. L., Manney, G. L., Froidevaux, L., Wu, D. L., Schwartz, M. J., Pumphrey, H. C., Jimenez, C., Nedoluha, G. E., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Pickett, H. M., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., Murtagh, D., Elkins, J. W., and Atlas, E.: Validation of the Aura Microwave Limb Sounder middle atmosphere water vapor and nitrous oxide measurements, J. Geophys. Res., 112, D24S36, doi:10.1029/2007JD008 724, 2007.

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

Lambert, A., Read, W., and Livesey, N.: MLS/Aura Level 2 Water Vapor (H_2O) Mixing Ratio V004, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (last access: 6 July 2017), https://doi.org/10.5067/AURA/MLS/DATA2009, 2015.

Langematz, U., Meul, S., Grunow, K., Romanowsky, E., Oberländer, S., Abalichin, J., and Kubin, A.: Future Arctic temperature and ozone: The role of stratospheric composition changes, J. Geophys. Res., 119, 2092–2112, https://doi.org/10.1002/2013JD021100, 2014.

Larsen, N.: Polar Stratospheric Clouds Microphysical and Optical Methods, Scientific Report 00-06, Danish Meteorological Institute, 2000.

Larsen, N., Knudsen, B. M., Rosen, J. M., Kjome, N. T., Neuber, R., and Kyrö, E.: Temperature histories in liquid and solid polar stratospheric cloud formation, J. Geophys. Res., 102, 23,505–23,517, 1997.

Lawrence, Z. D., Manney, G. L., Minschwaner, K., Santee, M. L., and Lambert, A.: Comparisons of polar processing diagnostics from 34 years of the ERA-Interim and MERRA reanalyses, Atmos. Chem. Phys., 15, 3873–3892, https://doi.org/10.5194/acp-15-3873-2015, 2015.

Livesey, N. J., Snyder, W. V., Read, W. G., and Wagner, P. A.: Retrieval algorithms for the EOS Microwave Limb Sounder (MLS), IEEE Trans. Geosci. Remote Sens., 44, 1144–1155, 2006.

Livesey, N. J., Read, W. G., Wagner, P. A., Froidevaux, L., Lambert, A., Manney, G. L., Valle, L. F. M., Pumphrey, H. C., Santee, M. L., Schwartz, M. J., Wang, S., Fuller, R. A., Jarnot, R. F., Knosp, B. W., and Martinez, E.: Version 4.2x Level 2 data quality and description document, Tech. Rep. JPL D-33509 Rev. C, Jet Propulsion Laboratory, available from http://mls.jpl.nasa.gov., 2017.

Manney, G., Santee, M., Froidevaux, L., Livesey, N., and Read, W.: MLS/Aura Level 2 Nitric Acid (HNO_3) Mixing Ratio V004, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (last access: 6 July 2017), https://doi.org/10.5067/AURA/MLS/DATA2012, 2015.

Manney, G. L., Swinbank, R., Massie, S. T., Gelman, M. E., Miller, A. J., Nagatani, R., O'Neill, A., and Zurek, R. W.: 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, 1996.

Manney, G. L., Sabutis, J. L., Pawson, S., Santee, M. L., Naujokat, B., Swinbank, R., Gelman, M. E., and Ebisuzaki, W.: 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:10.1029/2001JD001 149, 2003.

Manney, G. L. et al.: Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter, Mon. Weather Rev., 133, 1261–1278, 2005.

Manney, G. L., Santee, M. L., Rex, M., Livesey, N. J., Pitts, M. C., Veefkind, P., Nash, E. R., Wohltmann, I., Lehmann, R., Froidevaux, L., Poole, L. R., Schoeberl, M. R., Haffner, D. P., Davies, J., Dorokhov, V., Gernandt, H., Johnson, B., Kivi, R., Kyro, E., Larsen, N., Levelt, P. F., Makshtas, A., McElroy, C. T., Nakajima, H., Concepcion Parrondo, M., Tarasick, D. W., von der Gathen, P., Walker, K. A., and Zinoviev, N. S.: Unprecedented Arctic ozone loss in 2011, Nature, 478, 469–U65, 2011.

McDonald, A. J. and Hertzog, A.: Comparison of stratospheric measurements made by CHAMP radio occultation and Stratéole/Vorcore in situ data, Geophys. Res. Lett., 35, https://doi.org/10.1029/2008GL033338, 2008.

MERRA: Global Modeling and Assimilation Office (GMAO), inst6_3d_ana_Nv: MERRA 3D Analyzed State, Meteorology Instantaneous 6-hourly V5.2.0, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (last access: 6 July 2017), https://doi.org/10.5067/WGY2HAX25374, 2008.

MERRA2: Global Modeling and Assimilation Office (GMAO): MERRA-2 inst3_3d_asm_Nv: 3d, 3-Hourly, Instantaneous, Model-Level, Assimilation, Assimilated Meteorological Fields V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (last access: 6 July 2017), https://doi.org/10.5067/WWQSXQ8IVFW8, 2015.

Murphy, D. M. and Gary, B. L.: Mesoscale temperature fluctuations and polar stratospheric clouds, J. Atmos. Sci., 52, 1753–1760, 1995.

Murphy, D. M. and Koop, T.: Review of the vapour pressures of ice and supercooled water for atmospheric applications, Q. J. Roy. Meteor. Soc., 131, 1539–1565, 2005.

NCEP-CFSR: NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010, https://doi.org/10.5065/D69K487J, 2010.

Nedoluha, G. E., Benson, C. M., Hoppel, K. W., Alfred, J., Bevilacqua, R. M., and Drdla, K.: Antarctic dehydration 1998–2003: Polar Ozone and Aerosol Measurement III (POAM) measurements and Integrated Microphysics and Aerosol Chemistry on Trajectories (IMPACT) results with four meteorological models, J. Geophys. Res., 112, D07 305, doi:10.1029/2006JD007 414, 2007.

Orsolini, Y. J., Karpechko, A. Y., and Nikulin, G.: Variability of the Northern Hemisphere polar stratospheric cloud potential: the role of North Pacific disturbances, Q. J. Roy. Meteor. Soc., 135, 1020–1029, https://doi.org/10.1002/qj.409, 2009.

Parrondo, M. C., Yela, M., Gil, M., von der Gathen, P., and Ochoa, H.: Mid-winter lower stratosphere temperatures in the Antarctic vortex: comparison between observations and ECMWF and NCEP operational models, Atmos. Chem. Phys., 7, 435–441, 2007.

Pawson, S., Krüger, K., Swinbank, R., Bailey, M., and O'Neill, A.: Intercomparison of two stratospheric analyses: Temperatures relevant to polar stratospheric cloud formation, J. Geophys. Res., 104, 2041–2050, 1999.

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

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

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

Pommereau, J.-P., Garnier, A., Knudsen, B. M., Letrenne, G., Durand, M., Nunes-Pinharanda, M., Denis, L., Vial, F., Hertzog, A., and Cairo, F.: Accuracy of analyzed stratospheric temperatures in the winter Arctic vortex from infrared Montgolfier long-duration balloon flights 1. Measurements, J. Geophys. Res., 107, https://doi.org/10.1029/2001JD001379, 2002.

Preusse, P., Dörnbrack, A., Eckermann, S. D., Riese, M., Schaeler, B., Bacmeister, J. T., Broutman, D., and Grossmann, K. U.: Space-based measurements of stratospheric mountain waves by CRISTA 1. Sensitivity, analysis method, and a case study, J. Geophys. Res., 107, https://doi.org/10.1029/2001JD000699, 2002.

Read, W. G., Lambert, A., Bacmeister, J., Cofield, R. E., Christensen, L. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Fetzer, E., Froidevaux, L., Fuller, R., Herman, R., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Kelly, K., Knosp, B. W., Kovalenko, L. J., Livesey, N. J., Liu, H. C., Manney, G. L., Pickett, H. M., Pumphrey, H. C., Rosenlof, K. H., Sabounchi, X., Santee, M. L., Schwartz, M. J., Snyder, W. V., Stek, P. C., Su, H., Takacs, L. L., Thurstans, R. P., Vomel, H., Wagner, P. A., Waters, J. W., Webster, C. R., Weinstock, E. M., and Wu, D. L.: Aura Microwave Limb Sounder upper tropospheric and lower stratospheric H_2O and relative humidity with respect to ice validation, J. Geophys. Res., 112, D24S35, doi:10.1029/2007JD008 752, 2007.

Rex, M., Salawitch, R. J., von der Gathen, P., Harris, N. R. P., Chipperfield, M. P., and Naujokat, B.: Arctic ozone loss and climate change, Geophys. Res. Lett., 31, L04 116, doi:10.1029/2003GL018 844, 2004.

Rex, M. et al.: Arctic winter 2005: Implications for stratospheric ozone loss and climate change, Geophys. Res. Lett., 33, L23 808, doi:10.1029/2006GL026 731, 2006.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S.,

Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H. V. D., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP Climate Forecast System Reanalysis, Bull. Am. Meteor. Soc., 91, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1, 2010.

Santee, M. L., Lambert, A., Read, W. G., Livesey, N. J., Cofield, R. E., Cuddy, D. T., Daffer, W. H., Drouin, B. J., Froidevaux, L., Fuller, R. A., Jarnot, R. F., Knosp, B. W., Manney, G. L., Perun, V. S., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Waters, J. W., Muscari, G., de Zafra, R. L., Dibb, J. E., Fahey, D. W., Popp, P. J., Marcy, T. P., Jucks, K. W., Toon, G. C., Stachnik, R. A., Bernath, P. F., Boone, C. D., Walker, K. A., Urban, J., and Murtagh, D.: Validation of the Aura Microwave Limb Sounder HNO_3 measurements, J. Geophys. Res., 112, D24S40, doi:10.1029/2007JD008 721, 2007.

Scherllin-Pirscher, B., Kirchengast, G., Steiner, A. K., Kuo, Y.-H., and Foelsche, U.: Quantifying uncertainty in climatological fields from GPS radio occultation: an empirical-analytical error model, Atmos. Meas. Tech., 4, 2019–2034, https://doi.org/10.5194/amt-4-2019-2011, 2011a.

Scherllin-Pirscher, B., Steiner, A. K., Kirchengast, G., Kuo, Y.-H., and Foelsche, U.: Empirical analysis and modeling of errors of atmospheric profiles from GPS radio occultation, Atmos. Meas. Tech., 4, 1875–1890, https://doi.org/10.5194/amt-4-1875-2011, 2011b.

Schreiner, W., Rocken, C., Sokolovskiy, S., Syndergaard, S., and Hunt, D.: Estimates of the precision of GPS radio occultations from the COSMIC/FORMOSAT-3 mission, Geophys. Res. Lett., 34, https://doi.org/10.1029/2006GL027557, l04808, 2007.

Schwartz, M., Livesey, N., and Read, W.: MLS/Aura Level 2 Temperature V004, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC) (last access: 6 July 2017), https://doi.org/10.5067/AURA/MLS/DATA2021, 2015.

Schwartz, M. J., Manney, G. L., Lambert, A., Read, W. G., Livesey, N. J., Waters, J. W., Wu, D. L., Froidevaux, L., Ao, C. O., Cofield, R. E., Daffer, W. H., Drouin, B. J., Fetzer, E. J., Fuller, R. A., Jarnot, R. F., Jiang, J. H., Jiang, Y. B., Knosp, B. W., Li, J.-L. F., Santee, M. L., Snyder, W. V., Stek, P. C., Thurstans, R. P., Wagner, P. A., Pawson, S., Bernath, P. F., Walker, K. A., Boone, C. D., III, J. M. R., Mlynczak, M. G., Li, J.-L. F., Krüger, K., and Tompkins, A. M.: Validation of the Aura Microwave Limb Sounder Temperature and Geopotential Height Measurements, J. Geophys. Res., 113, https://doi.org/doi:10.1029/2007JD008783, 2008.

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S., and Peubey, C.: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim, Q. J. Roy. Meteor. Soc., 140, 329–353, https://doi.org/10.1002/qj.2317, 2014.

Sinnhuber, B.-M., Stiller, G., Ruhnke, R., von Clarmann, T., Kellmann, S., and Aschmann, J.: Arctic winter 2010/2011 at the brink of an ozone hole, Geophys. Res. Lett., 38, https://doi.org/10.1029/2011GL049784, 2011.

Smith, E. and Weintraub, S.: The constants in the equation for atmospheric refractive index at radio frequencies. in Proceedings of Proc. IRE 41, 1953.

Solomon, S.: Stratospheric ozone depletion: A review of concepts and history, Rev. Geophys., 37, 275–316, 1999.

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

Staten, P. W. and Reichler, T.: Use of radio occultation for long-term tropopause studies: Uncertainties, biases, and instabilities, J. Geophys. Res., 113, https://doi.org/10.1029/2008JD009886, 2008.

Staten, P. W. and Reichler, T.: Apparent precision of GPS radio occultation temperatures, Geophys. Res. Lett., 36, https://doi.org/10.1029/2009GL041046, 2009.

Thorne, P. W. and Vose, R. S.: Reanalyses Suitable for Characterizing Long-Term Trends, Bull. Am. Meteor. Soc., 91, 353–361, https://doi.org/10.1175/2009BAMS2858.1, 2010.

Tsuda, T., Nishida, M., Rocken, C., and Ware, R. H.: A Global Morphology of Gravity Wave Activity in the Stratosphere Revealed by the GPS Occultation Data (GPS/MET), Journal of Geophysical Research: Atmospheres, 105, 7257–7273, https://doi.org/10.1029/1999JD901005, 2000.

Voigt, C., Schlager, H., Luo, B. P., Dornbrack, A. D., Roiger, A., Stock, P., Curtius, J., Vossing, H., Borrmann, S., Davies, S., Konopka, P., Schiller, C., Shur, G., and Peter, T.: Nitric Acid Trihydrate (NAT) formation at low NAT supersaturation in Polar Stratospheric Clouds (PSCs), Atmos. Chem. Phys., 5, 1371–1380, 2005.

Wang, K.-Y. and Lin, S.-C.: First continuous GPS soundings of temperature structure over Antarctic winter from FORMOSAT-3/COSMIC constellation, Geophysical Research Letters, 34, https://doi.org/10.1029/2007GL030159, l12805, 2007.

Waters, J. W., Froidevaux, L., Harwood, R. S., Jarnot, R. F., Pickett, H. M., Read, W. G., Siegel, P. H., Cofield, R. E., Filipiak, M. J., Flower, D. A., Holden, J. R., Lau, G. K. K., Livesey, N. J., Manney, G. L., Pumphrey, H. C., Santee, M. L., Wu, D. L., Cuddy, D. T., Lay, R. R., Loo, M. S., Perun, V. S., Schwartz, M. J., Stek, P. C., Thurstans, R. P., Boyles, M. A., Chandra, K. M., Chavez, M. C., Chen, G. S., Chudasama, B. V., Dodge, R., Fuller, R. A., Girard, M. A., Jiang, J. H., Jiang, Y. B., Knosp, B. W., LaBelle, R. C., Lam, J. C., Lee, K. A., Miller, D., Oswald, J. E., Patel, N. C., Pukala, D. M., Quintero, O., Scaff, D., Van Snyder, W., Tope, M. C., Wagner, P. A., and Walch, M. J.: The Earth Observing System Microwave Limb Sounder (EOS MLS) on the Aura satellite, IEEE Trans. Geosci. Remote Sens., 44, 1075–1092, 2006.

Waugh, D. W. and Polvani, L. M.: Stratospheric Polar Vortices, pp. 43–57, The Stratosphere: Dynamics, Transport, and Chemistry (eds L. M. Polvani, A. H. Sobel and D. W. Waugh), American Geophysical Union, Washington, D. C. https://doi.org/10.1002/9781118666630.ch3, 2010.

Wegner, T., Grooß, J.-U., von Hobe, M., Stroh, F., Sumińska-Ebersoldt, O., Volk, C. M., Hösen, E., Mitev, V., Shur, G., and Müller, R.: Heterogeneous chlorine activation on stratospheric aerosols and clouds in the Arctic polar vortex, Atmos. Chem. Phys., 13, 11 095–11 106, https://doi.org/10.5194/acp-12-11095-2012, 2012.

Wegner, T., Kinnison, D. E., Garcia, R. R., and Solomon, S.: Simulation of polar stratospheric clouds in the specified dynamics version of the whole atmosphere community climate model, J. Geophys. Res., 118, 4991—5002, https://doi.org/10.1002/jgrd.50415, 2013.

Whiteway, J. A.: Enhanced and Inhibited Gravity Wave Spectra, J. Atmos. Sci., 56, 1344–1352, https://doi.org/10.1175/1520-0469(1999)056<1344:EAIGWS>2.0.CO;2, 1999.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP data processing algorithms, J. Atmos. Oceanic Technol., 26, 2310–2323,doi:10.1175/2009JTECHA1281.1, 2009.

Wohltmann, I., Wegner, T., Müller, R., Lehmann, R., Rex, M., Manney, G. L., Santee, M. L., Bernath, P., Sumińska-Ebersoldt, O., Stroh, F., von Hobe, M., Volk, C. M., Hösen, E., Ravegnani, F., Ulanovsky, A., and Yushkov, V.: Uncertainties in modelling heterogeneous chemistry and Arctic ozone depletion in the winter 2009/2010, Atmos. Chem. Phys., 13, 3909–3929, https://doi.org/10.5194/acp-13-3909-2013, 2013.

Table 1. Reanalysis temperature data. Details are valid for the 100–10 hPa region and 2008–2013.

Dataset	Horizontal resolution	Time resolution	Vertical resolution	
GEOS-5.9.1*	0.625° by 0.5°	3 hr	1.1–1.4 km	Rienecker et al. (2011)
	Goddard Earth Observing System Model, Version 5			
	Near real time assimilation system			
	Global Modeling and Assimilation Office			
	NASA GMAO			
MERRA**	0.666° by 0.5°	6 hr	1.1–1.4 km	Rienecker et al. (2011)
	Modern-Era Retrospective analysis for Research and Applications			
	NASA GMAO			
MERRA2	0.626° by 0.5°	3 hr	1.1–1.4 km	Gelaro et al. (2017)
	Modern-Era Retrospective analysis for Research and Applications Version 2			
	NASA GMAO			
JRA55	0.582° by 0.56°	6 hr	1.1–1.5 km	Kobayashi et al. (2015)
	Japanese 55-year Reanalysis			
	Japanese Meteorological Agency (JMA)			
ERA-I	0.75° by 0.75°	6 hr	1.2–1.5 km	Dee et al. (2011)
	European Centre for Medium-range Weather Forecasts (ECMWF) interim reanalysis			
	ECMWF			
NCEP-CFSR	0.5° by 0.5°	6 hr	2.3–4.8 km	Saha et al. (2010)
	National Centers for Environmental Prediction (NCEP)			
	Climate Forecast System Reanalysis (CFSR)			
	National Oceanic and Atmospheric Administration (NOAA)			

*GEOS-5.9.1 is a near real-time assimilation product provided to the NASA Earth Observing System (EOS) instrument science teams.

**All except MERRA include GPS RO data during the time period studied here.

Table 2. Statistics of the reanalysis temperature distributions for LIQ PSCs relative to the theoretical STS equilibrium temperature (T_{eq}) at 46 hPa for the combined 2008-2013 Antarctic winters (3261 profiles).

Dataset	Median (K)	Mean (K)	Standard Deviation (K)	Standard Error (K)	χ^2
GEOS5.9.1	-0.56	-0.57	0.54	0.009	1.62
MERRA	-0.90	-0.91	0.58	0.010	1.40
MERRA2	-0.54	-0.56	0.54	0.009	1.62
JRA55	-0.59	-0.62	0.53	0.009	2.21
ERA-I	-0.29	-0.32	0.57	0.010	2.08
NCEP-CFSR	-0.37	-0.38	0.59	0.010	1.82

Table 3. As for Table 2, except for ICE PSCs relative to the theoretical ice frost point temperature (3288 profiles)

Dataset	Median (K)	Mean (K)	Standard Deviation (K)	Standard Error (K)	χ^2
GEOS5.9.1	-0.21	-0.18	0.71	0.012	1.94
MERRA	-0.58	-0.55	0.70	0.012	1.95
MERRA2	-0.23	-0.21	0.70	0.012	1.74
JRA55	-0.41	-0.38	0.70	0.012	1.56
ERA-I	+0.01	+0.03	0.72	0.013	1.54
NCEP-CFSR	-0.12	-0.11	0.73	0.013	1.64

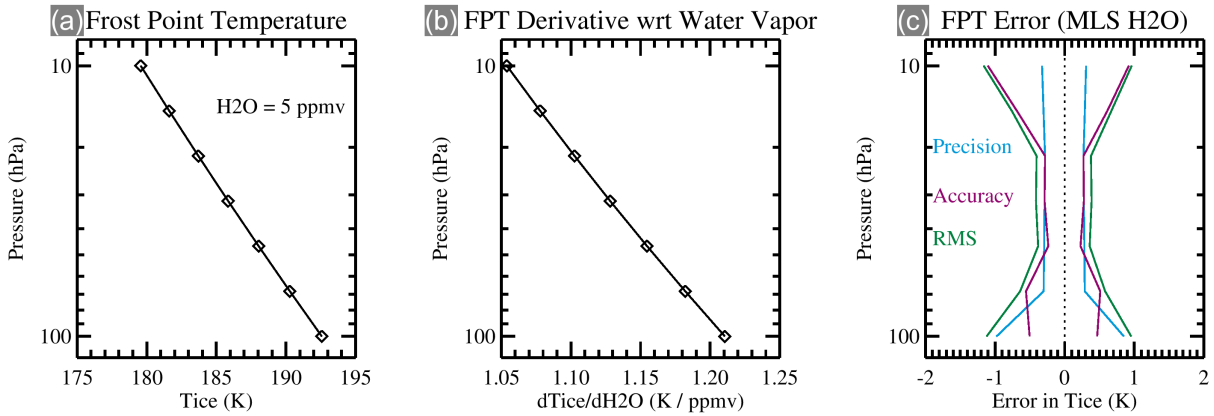

Figure 1. (a) Variation of frost point temperature with pressure in the lower stratosphere for a fixed H_2O volume mixing ratio (5 ppmv). (b) Derivative of the frost point temperature with respect to water vapor. (c) Error in determination of frost point temperature (T_{ice}) arising from the uncertainties in the MLS H_2O measurement.

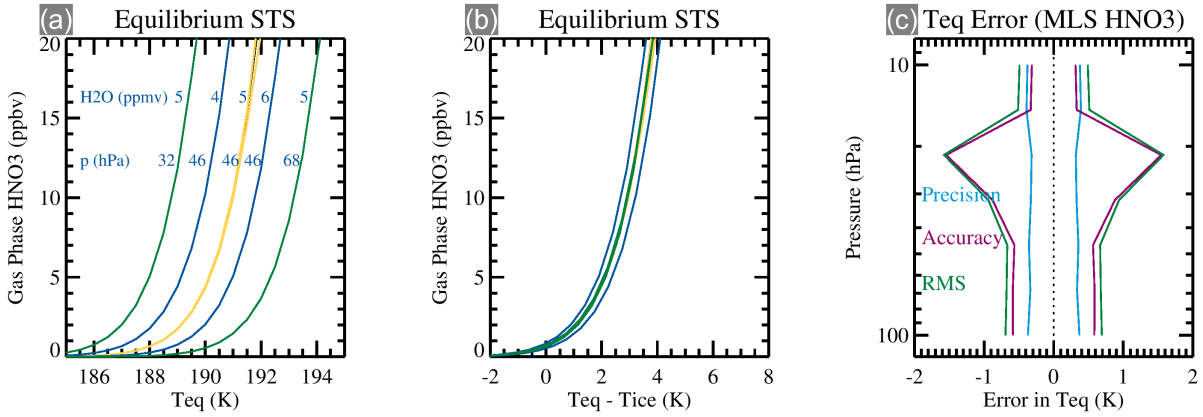

Figure 2. (a) Equilibrium HNO_3 uptake curves vs temperature for a variety of H_2O mixing ratios (4, 5, 6 ppmv), H_2SO_4 mixing ratios (0.1, 0.5 ppbv), and pressures (68, 46, 32 hPa). (b) As for (a), except for a coordinate transformation of the temperature scale relative to the ice frost point (T_{ice}). (c) Error in the equilibrium STS temperature arising from the errors in the MLS HNO_3.

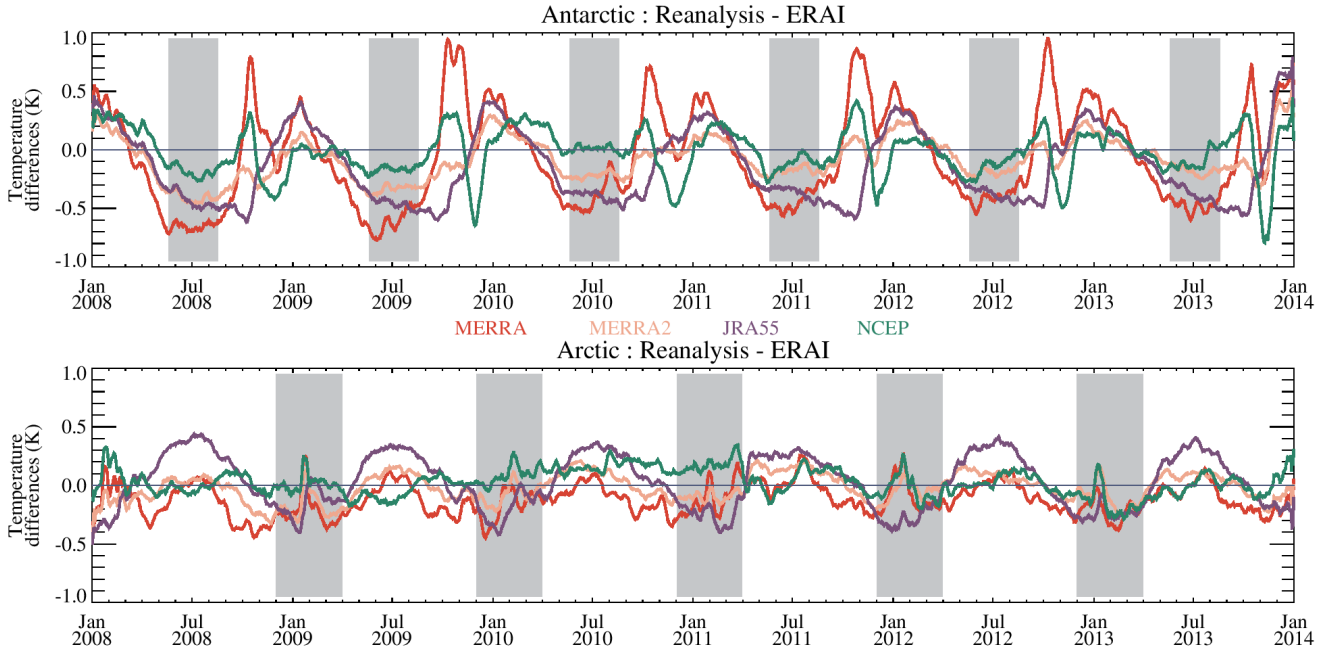

Figure 3. (Upper) Timeseries (for 12UT) from 2008 to 2013 of 10-day boxcar smoothed temperature differences for MERRA, MERRA2 JRA55 and NCEP-CFSR, relative to ERA-I at 46 hPa and averaged over the 60°Antarctic polar cap. The gray regions indicate the analysis periods defined for the analysis of PSC related metrics. (Lower) Similar, except for the Arctic.

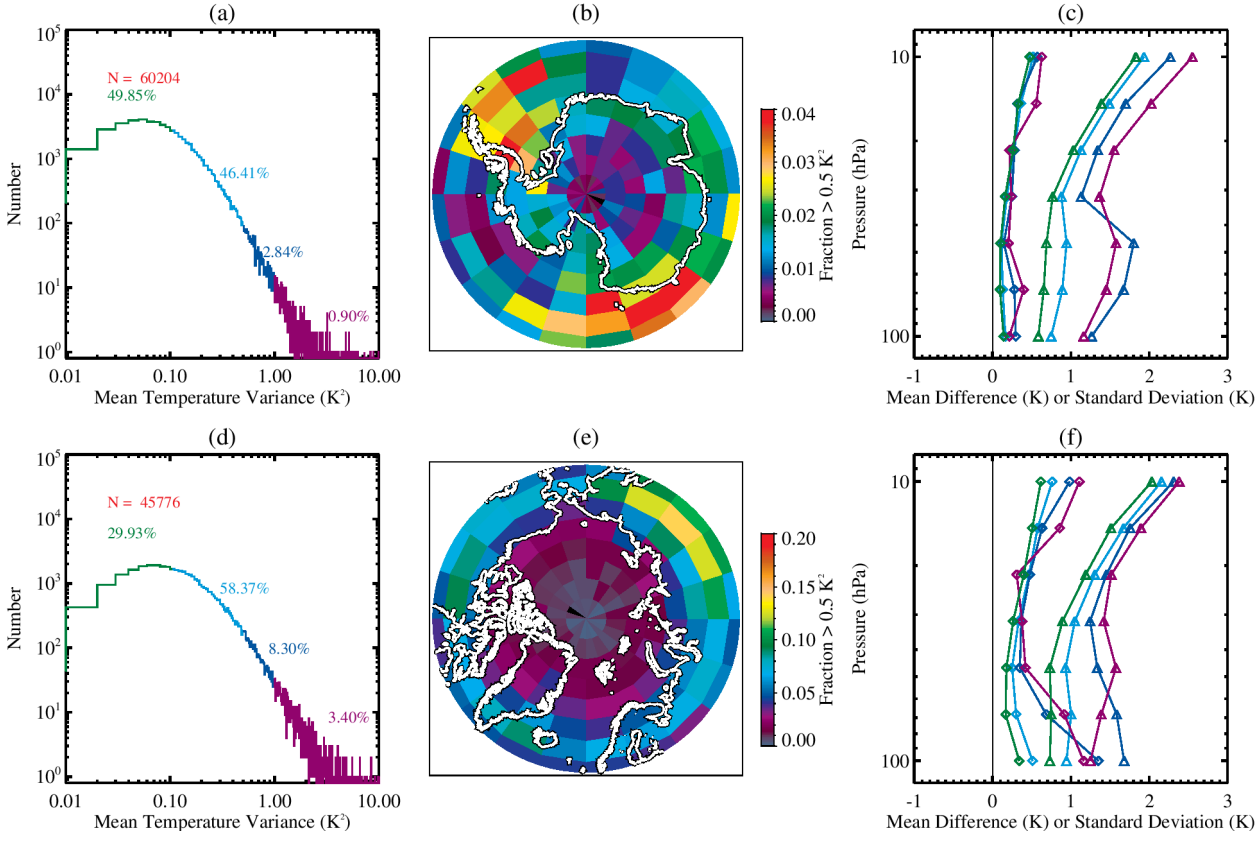

Figure 4. Temperature variance, $\sigma^2 = \overline{T'^2}$, in K^2, of the COSMIC profiles averaged over the pressure range 80–20 hPa for 2008–2013 for the Antarctic and Arctic. (a,d) Histogram of the mean COSMIC temperature variance with color coding for $\sigma^2 \leq 0.1$ K^2 (green), 0.1 $K^2 > \sigma^2 \leq 0.5$ K^2 (cyan), 0.5 $K^2 > \sigma^2 \leq 1.0$ K^2 (blue) and $\sigma^2 \geq 1.0$ K^2 (purple). N is the total number of profiles, and the percentages falling in each of the four colored σ^2 regions are also given. (b,e) Geographic distribution of the fraction of COSMIC profiles with $\sigma^2 > 0.5$ K^2. (c,f) Mean values (diamonds) and standard deviations (triangles) of the temperature differences between ERA-I and COSMIC, color coded for the COSMIC temperature variance as given in (a,d). The estimated noise variance is about 0.03 K^2.

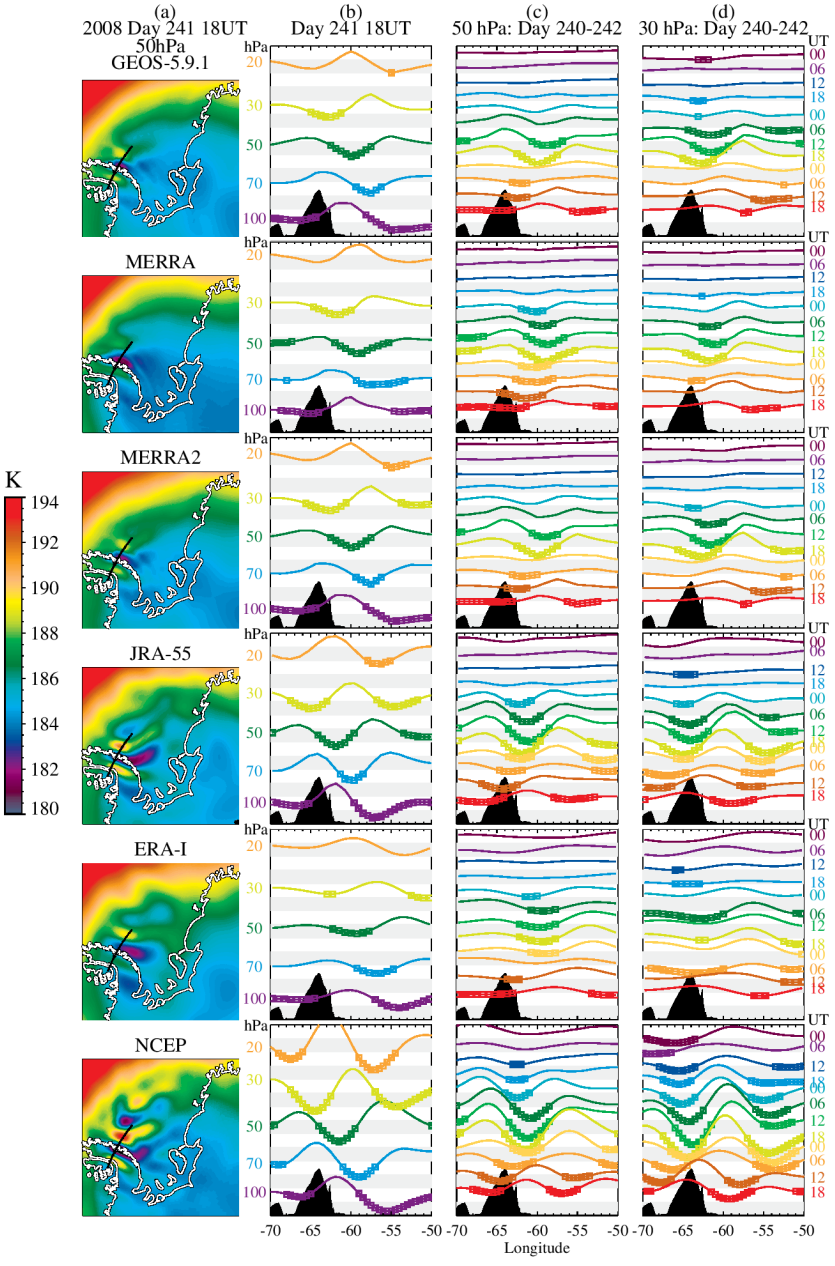

Figure 5. (a) Geographic distribution of 50 hPa temperatures over a region of the Palmer Peninsula on 28 August 2008 (day 241) at 18UT for different reanalyses. (b) Longitudinal reanalysis temperature structure from -70° to -50° on the 100, 70, 50, 30, and 20 hPa pressure levels along the 72° S latitude circle (shown as a black line in (a)). The ordinate is an arbitrary temperature scale with offset colored lines representing the longitudinal temperature variation for each of the labeled pressure levels. (c,d) Time series of the 50 hPa (30 hPa) reanalysis temperatures starting from 27 August (d240) at 0UT and ending on 29 August (d242) at 18UT. The ordinate is an arbitrary temperature scale with offset colored lines representing the temperature variation at the 6-hourly intervals labeled on the right. In (b,c,d), temperatures below the ice frost point are highlighted by square symbols, and mountain terrain is indicated by the black shading. Gray/white shading denotes both 5 K temperature bands and 500 m terrain increments.

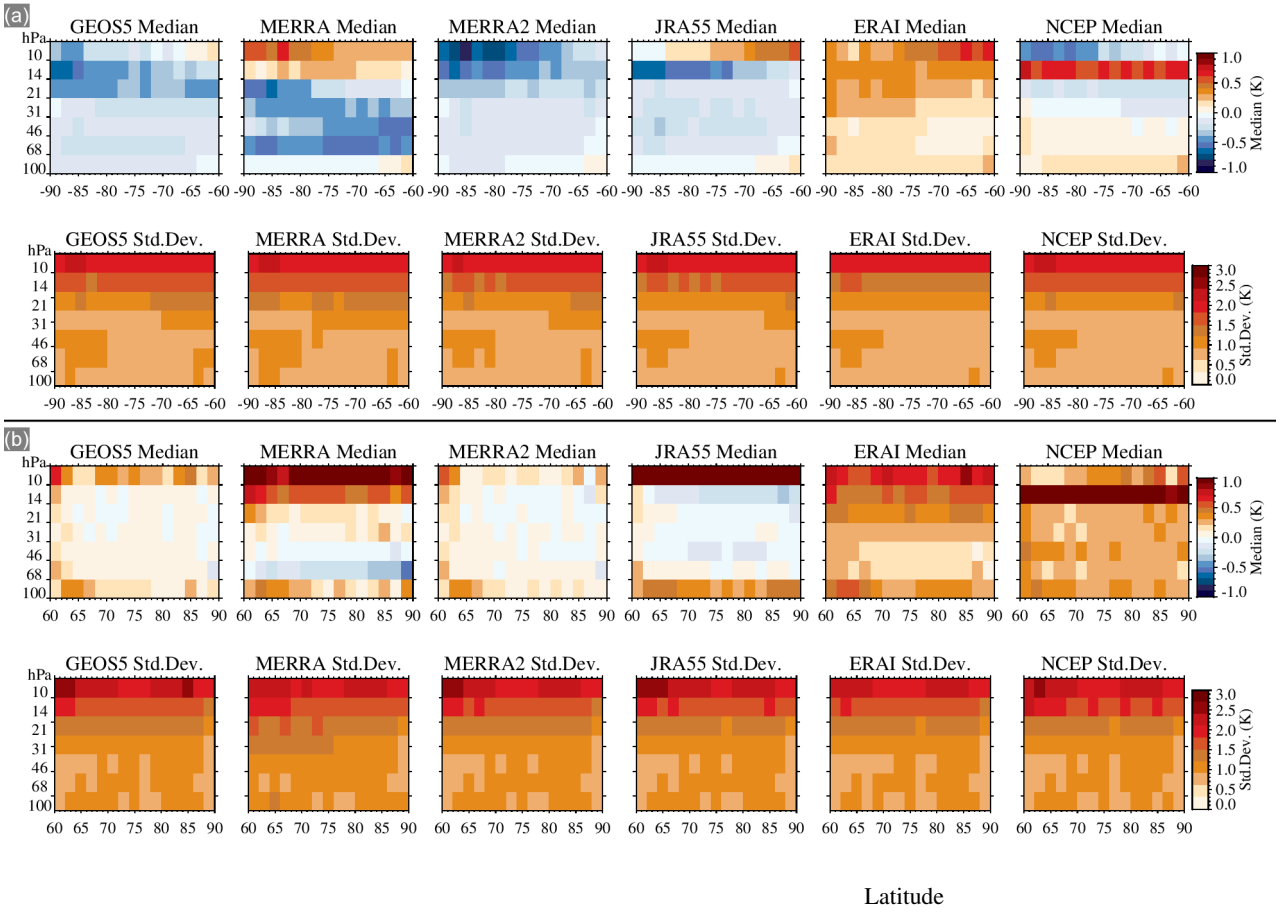

Figure 6. Zonal distributions of the median and standard deviations of the differences between the reanalysis temperatures and COSMIC temperatures ($T_{re} - T_{COSMIC\ RO}$) averaged over 2008–2013 for (a) the Antarctic (90° S–60° S) and (b) the Arctic (60° N–90° N).

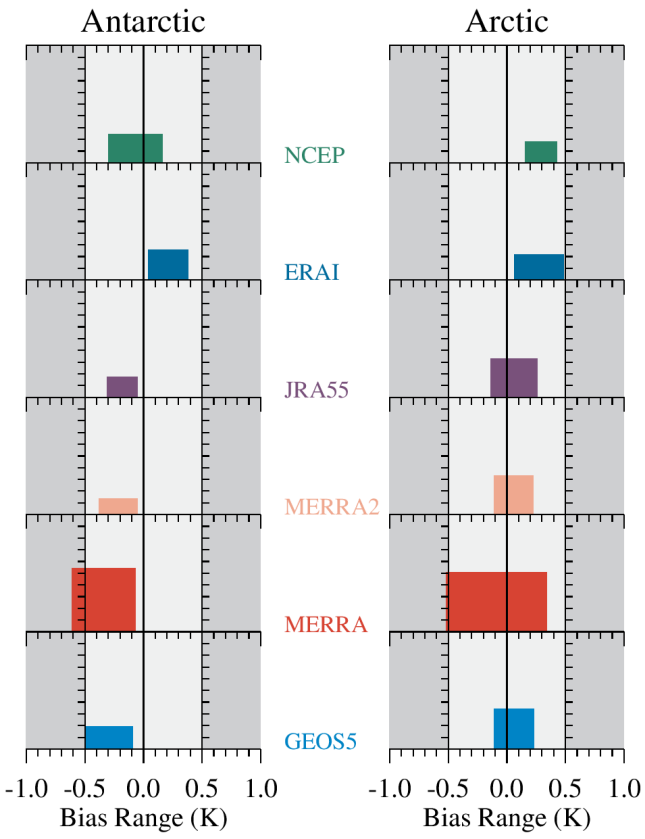

Figure 7. Temperature bias ranges of the reanalyses relative to COSMIC GPS RO for Antarctic and Arctic winters 2008–2013, poleward of 60°, and for pressure levels 68–21 hPa, derived from the data in Figure 6. The bias ranges are obtained from the extrema of the individual yearly mean bias values over 68–21 hPa. Each colored box indicates the bias ranges for the particular reanalysis scheme given in the legend; the abscissa gives the range of minimum to maximum bias; the ordinate gives the largest meridional bias range. Light grey shading indicates a ±0.5 K scale in the bias range and a 0–1 K scale for the meridional bias range.

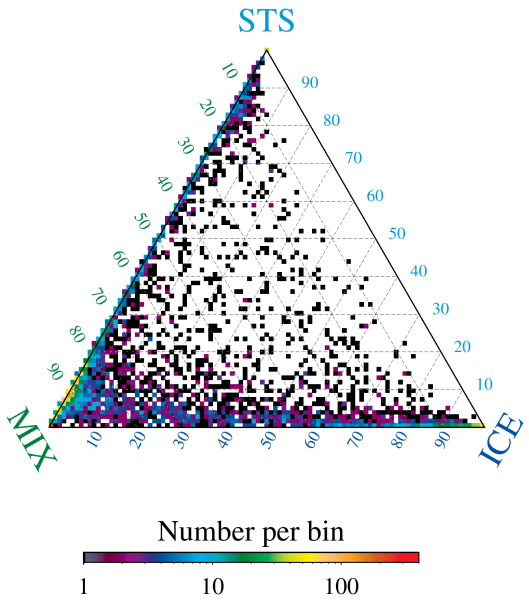

Figure 8. Ternary density plot of the relative proportions of `STS/MIX/ICE` occurring in the MLS geometric FOV at temperatures below the ice frost point for the 2013 Antarctic winter at 46 hPa.

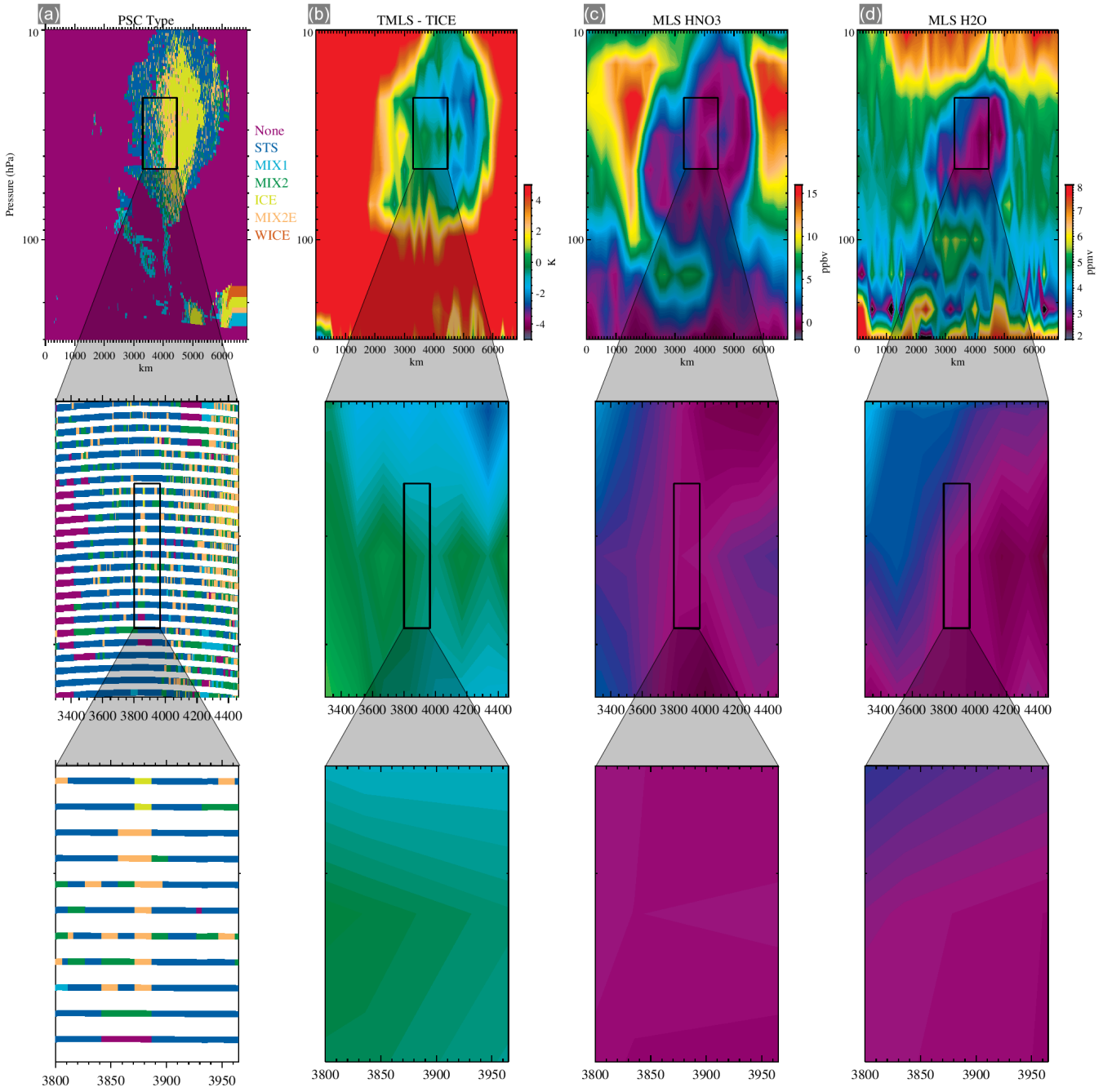

Figure 9. Along-track vertical cross-sections of PSCs, temperature, HNO_3 and H_2O over a 7000 km transect crossing Antarctica on 23 June 2008 (2008d175). Column (a) CALIOP PSC types. Column (b) MLS temperature relative to the ice frost point. Column (c) MLS HNO_3. Column (d) MLS H_2O. Selected regions (black rectangular boxes), centered at 32 hPa in the top row, are shown enlarged in the middle row (1165 km along-track by 4.32 km high), and similarly enlarged again in the bottom row (165 km by 2.16 km). In the bottom row, the resolution is that of the MLS geometric FOV, in which MLS reports only a single measurement value, whereas CALIOP reports ∼400 individual 5 km by 0.180 km pixels.

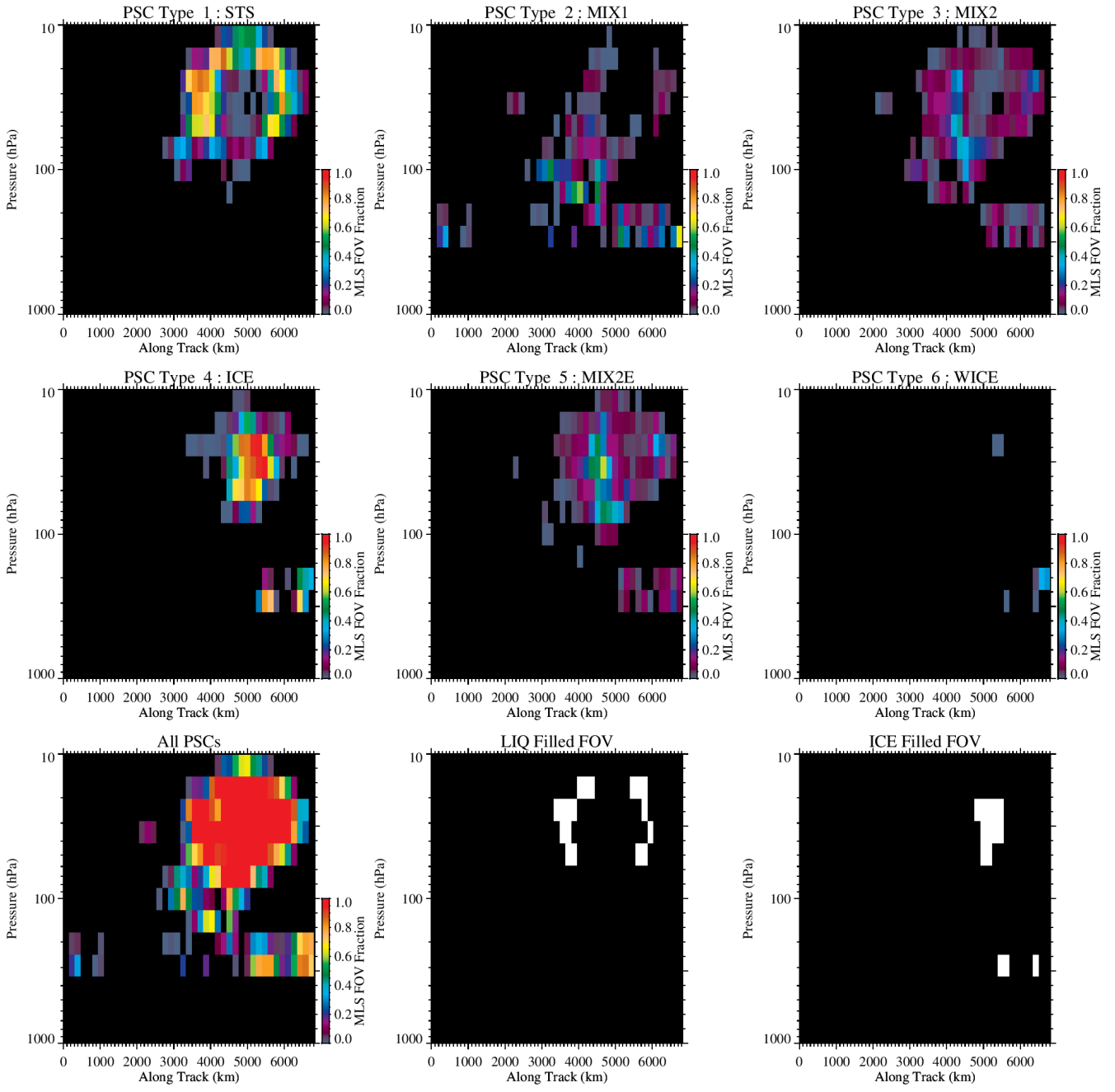

Figure 10. Orbit transect from Figure 9 indicating the fraction of the MLS geometric FOV that is filled by each CALIOP classified PSC type and all PSCs. The white areas in the `LIQ` and `ICE` Filled FOV plots indicate the locations that pass the strict FOV filling criteria given in the text.

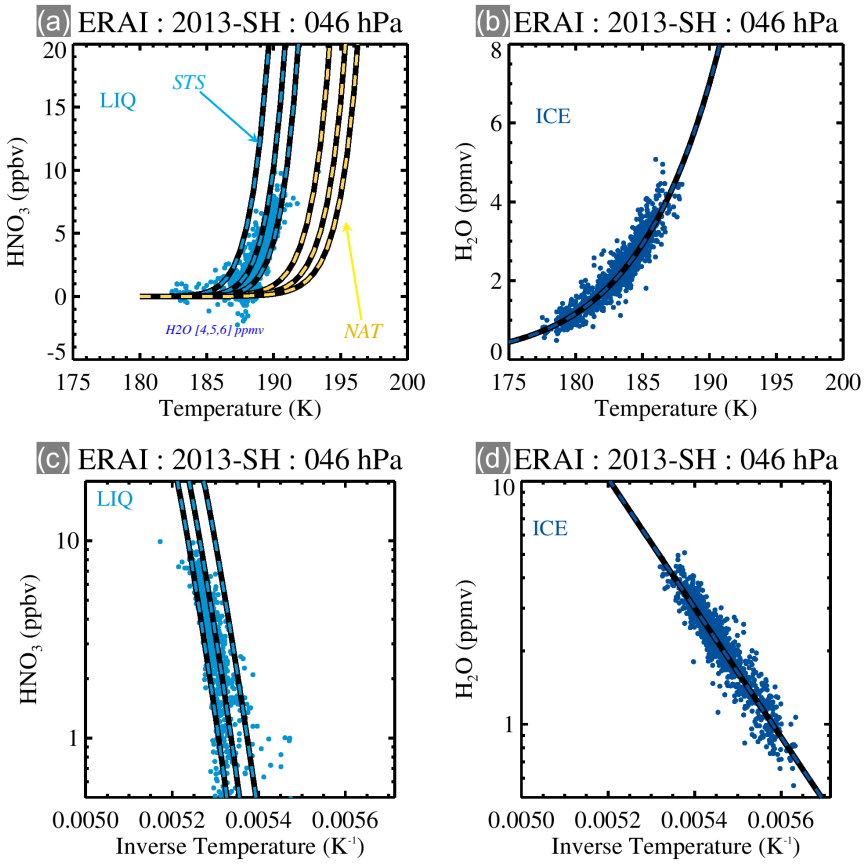

Figure 11. Scatter plots of coincident MLS HNO_3 and H_2O vs ERA-I temperature for PSCs classified by CALIOP on the 46 hPa pressure level in the 2013 Antarctic winter. (a) MLS HNO_3 vs ERA-I temperature for liquid (LIQ) PSCs. The theoretical equilibrium uptake of HNO_3 by STS is shown for representative ambient H_2O values by the cyan-black dashed lines. The yellow-black dashed lines show the corresponding NAT equilibrium curves. (b) MLS H2O vs ERA-I temperature for ice (ICE) PSCs. The blue-black dashed lines indicate the theoretical equilibrium for the frost point temperatures. (c,d) As for (a,b), except plotted as log mixing ratio vs inverse temperature.

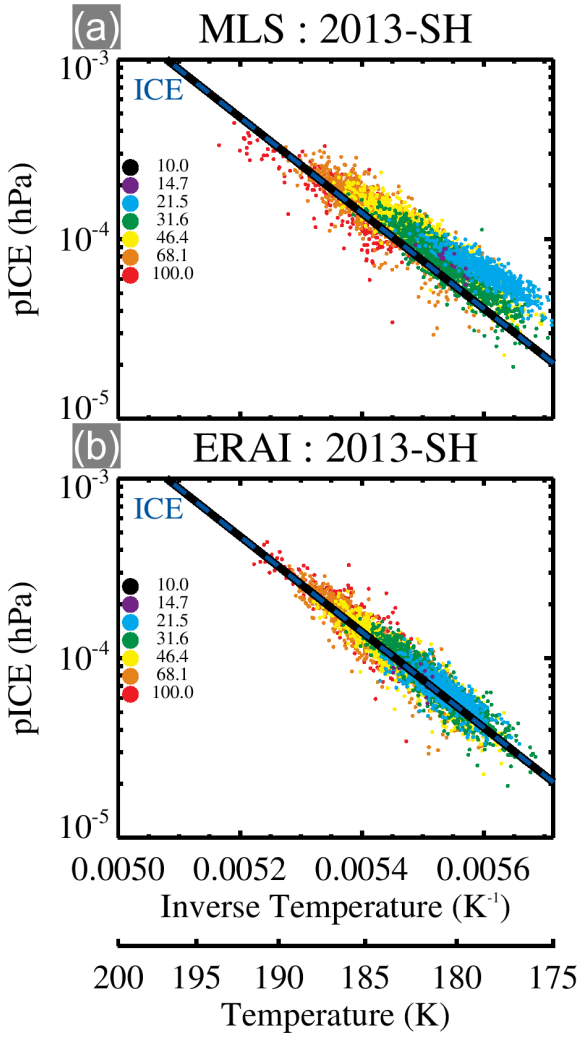

Figure 12. Temperature variation of the estimated partial pressure of water vapor over ice, p_{ice}, determined from the product of the MLS H_2O volume mixing ratio and the total atmospheric pressure, for Antarctic ice PSCs in 2013. The pressure levels (hPa) are color coded and given in the legend. (a) MLS temperature. (b) ERA-I temperature. The blue-black dashed lines indicate the theoretical equilibrium p_{ice} as a function of the frost point temperature.

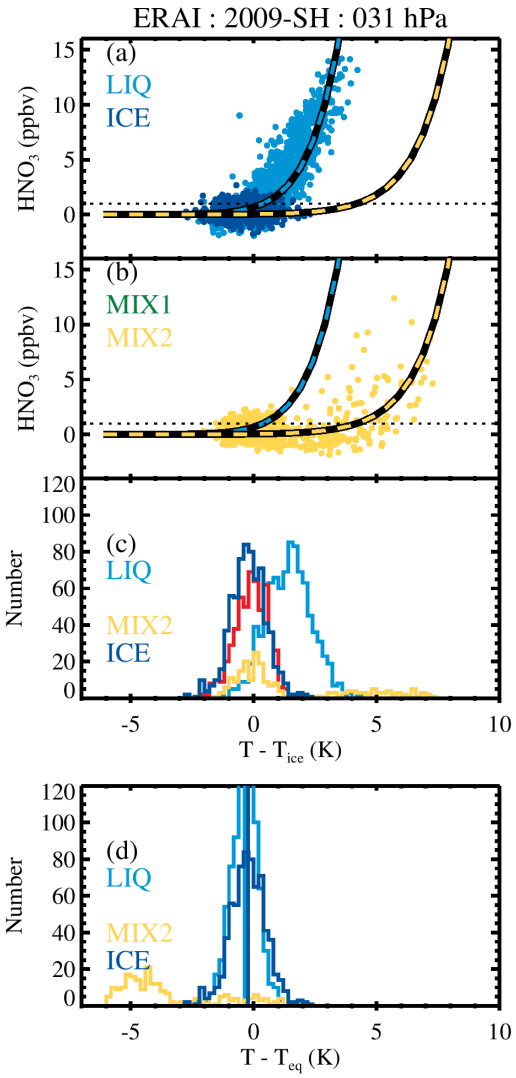

Figure 13. Composite statistics for d140–d230 (the bulk of the PSC formation period) of a representative year (2009) of the MLS gas-phase HNO_3 variation with ERA-I temperature corresponding to CALIOP PSC classifications at 32 hPa, with the added constraint that at least 75% of the MLS FOV is filled 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` (cyan) and `ICE` (blue). (b) As for (a), but for `MIX2` PSCs (yellow; because of non-equilibrium effects, which cause larger temperature scatter, `MIX2` PSCs are not used in this analysis, but this panel is shown to indicate the good discrimination between the solid and liquid uptake curve branches.) Equilibrium STS (black/cyan dashed) and NAT (black/yellow dashed) curves show the theoretical uptake of HNO_3. (c) Temperature histograms for HNO_3 mixing ratio >1 ppbv for `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.

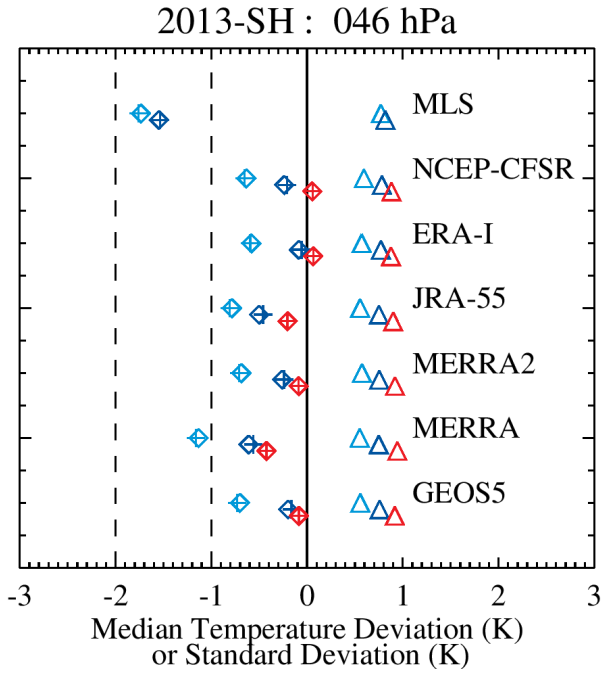

Figure 14. Median deviations (diamonds) and standard deviations (triangles) of the temperature distributions for `LIQ` (cyan), `ICE` (blue) and COSMIC (red) for each reanalysis, GEOS-5.9.1, and MLS, for a representative year (2013) at 46 hPa. The plus symbols denote the means of the distributions. Their similarity to the median values indicates that the distributions have small skewness.

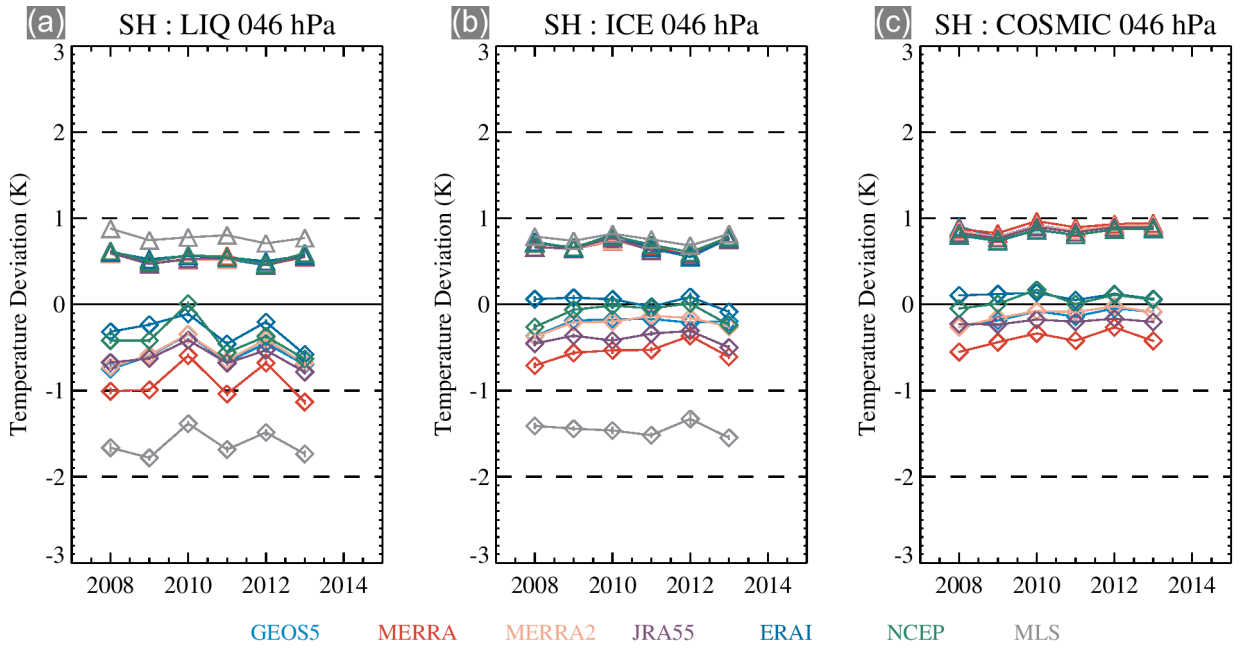

Figure 15. Median temperature deviations (diamonds) from T_{eq} for (a) `LIQ` PSCs and from T_{ice} for (b) `ICE` PSCs for the temperature distributions in each year 2008–2013 at 46 hPa. The standard deviations of the temperature deviations are also shown (triangles). Lines for the different reanalyses, GEOS-5.9.1, and MLS are color coded (see legend). (c) as for (a,b) except for deviations with respect to COSMIC GPS RO.

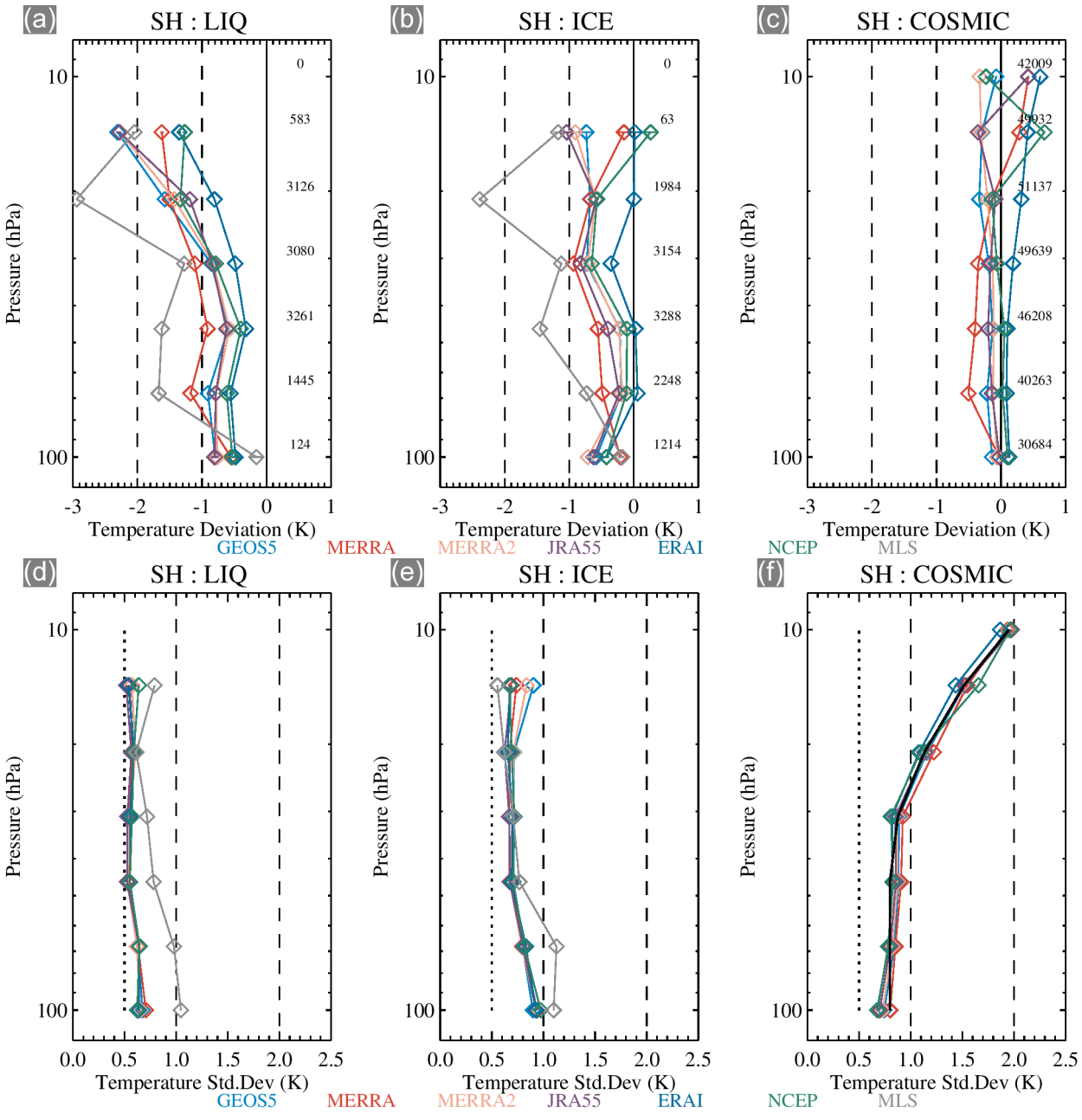

Figure 16. Vertical profiles of median temperature deviations from T_{eq} for (a) `LIQ` PSCs and for (b) `ICE` PSCs for the temperature distributions accumulated over the Antarctic PSC seasons (20 May (d140) to 18 August (d230) from 2008 to 2013). Lines for the different reanalyses, GEOS-5.9.1, and MLS are color coded (see legend); the numerical values on the right-hand side of each panel indicate the total number of observations in the distribution at the corresponding pressure level. (c) as for (a,b) except for deviations with respect to COSMIC GPS 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

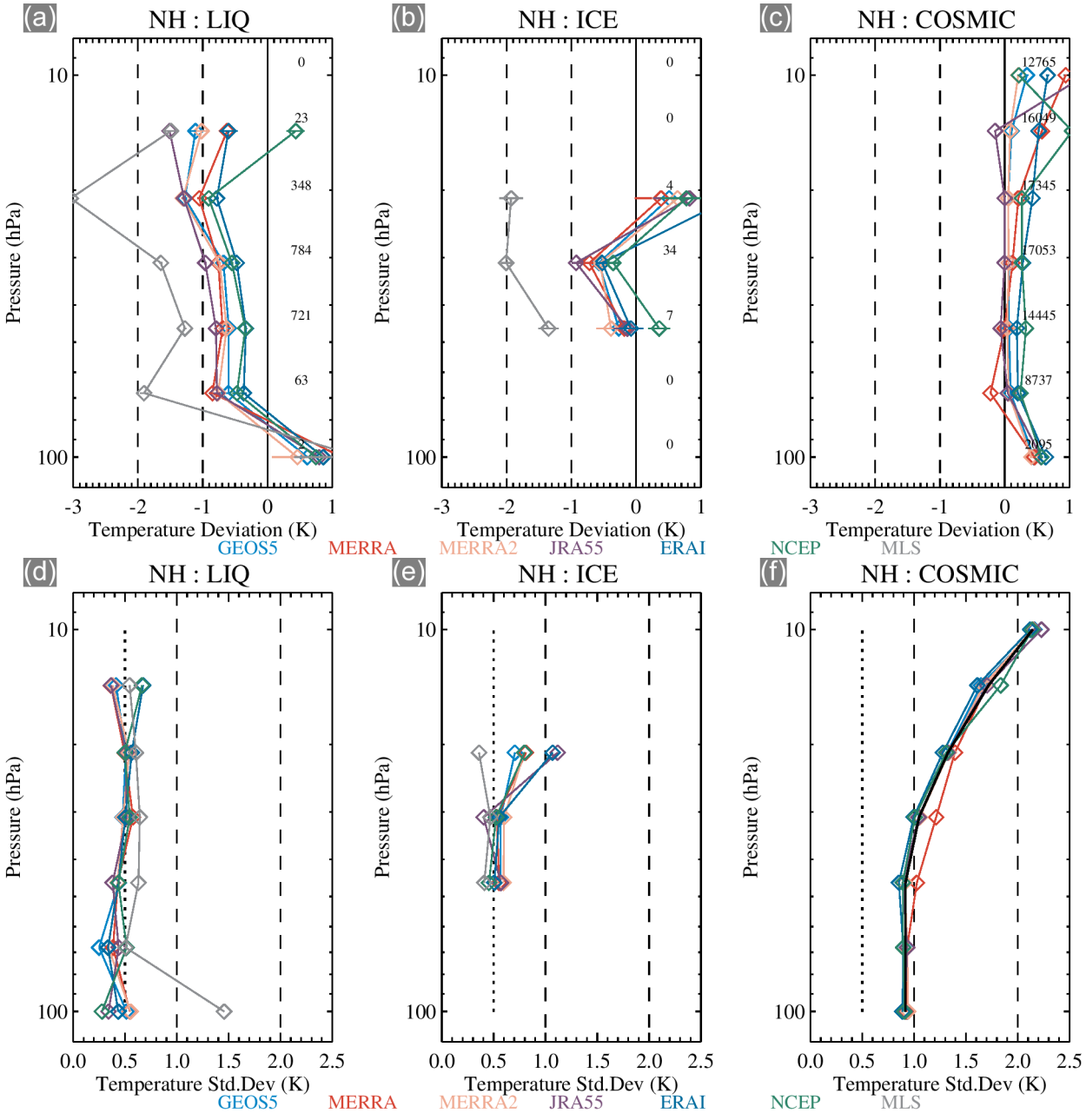

Figure 17. As for Figure 16, except for the Arctic PSC seasons (2 December (d336) to 31 March (d090) from 2008/2009 to 2012/2013).

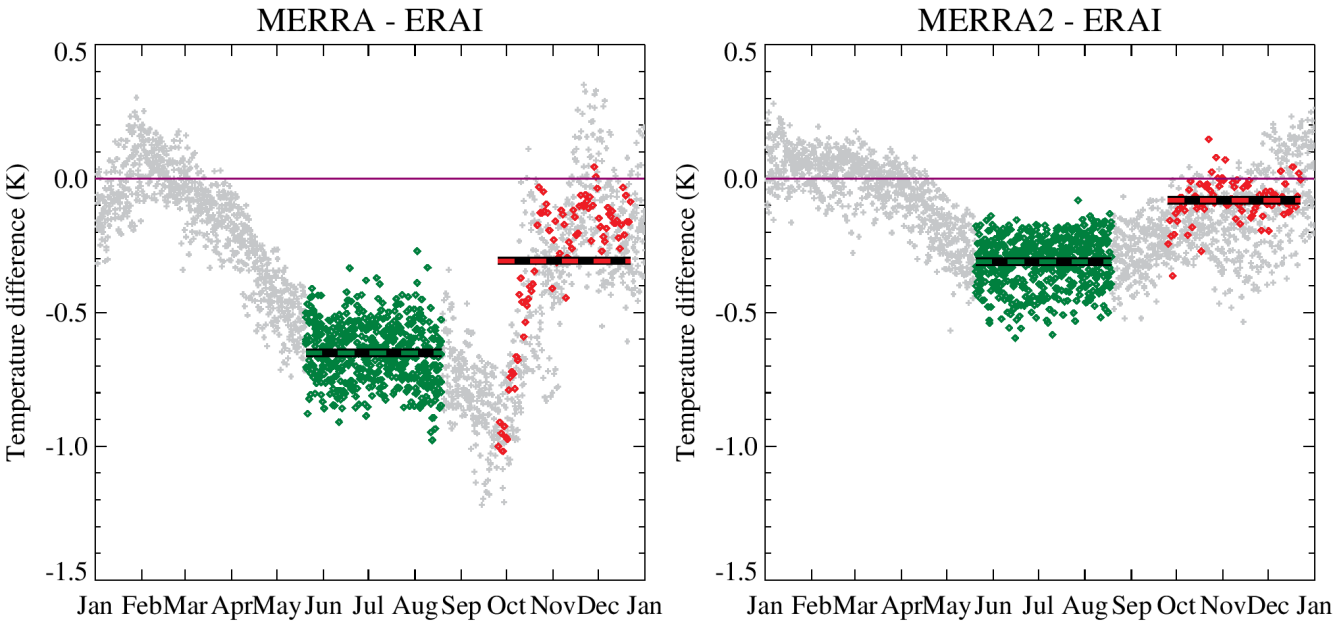

Figure 18. Daily mean temperature differences (not smoothed) of MERRA (left) and MERRA-2 (right) relative to ERA-I at 62 hPa (representative of the balloon float heights) in the 60° Antarctic polar cap for 2008–2013. Green-black dashed line indicates the mean of the 2008-2013 differences (green symbols) in the PSC analysis window. Red-black dashed line indicates the mean of the differences (red symbols) over the time-span of 90% of the balloon measurements.

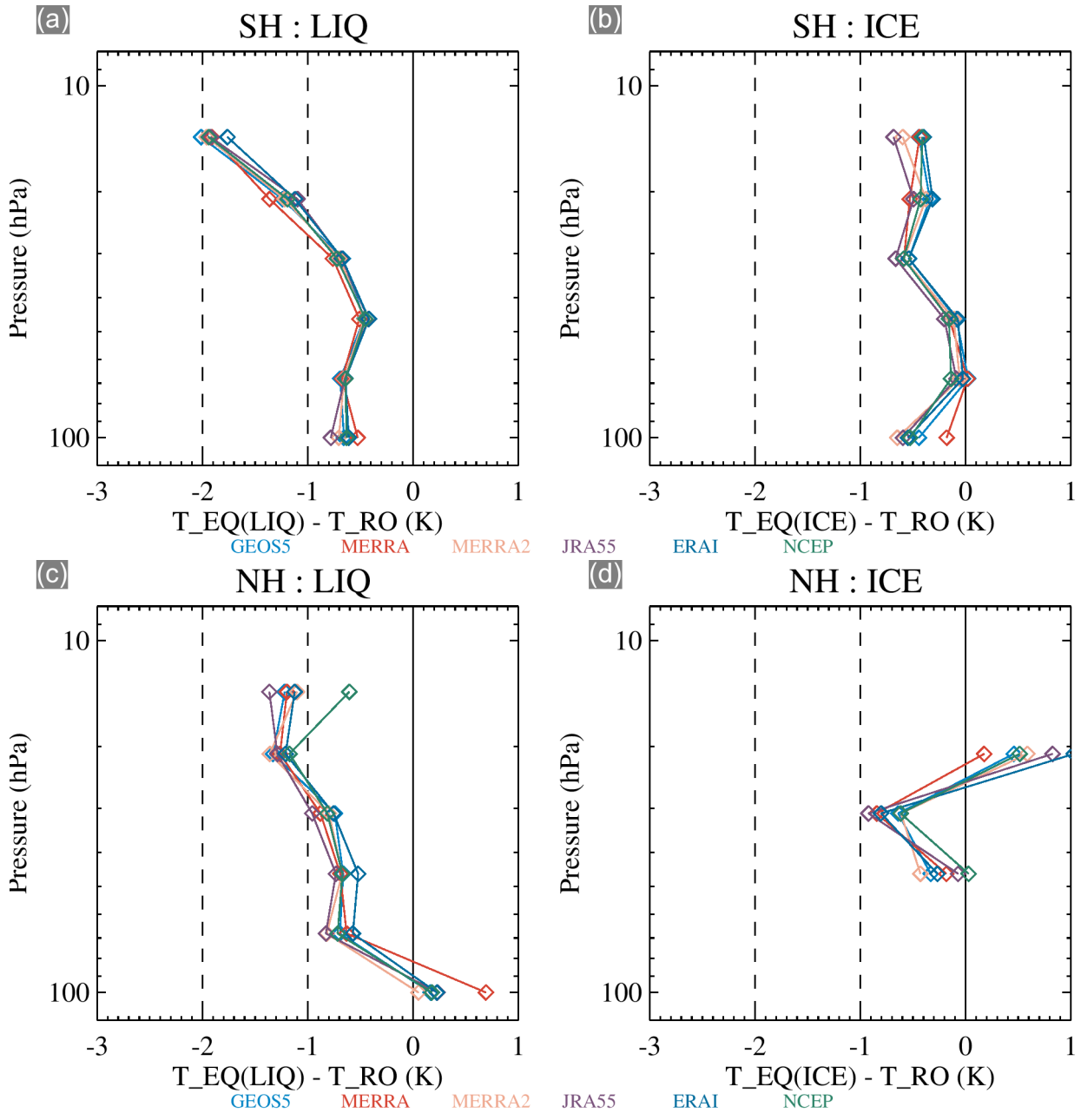

Figure 19. Vertical profiles of the temperature differences between (a) LIQ equilibrium and COSMIC GPS RO, $\overline{\Delta T_{LIQ}} - \overline{\Delta T_{COSMIC}}$, and (b) ICE equilibrium and COSMIC RO, $\overline{\Delta T_{ICE}}$ - $\overline{\Delta T_{COSMIC}}$, for the Antarctic. (c,d) As for (a,b) except for the Arctic. Lines for the different reanalyses are color coded according to the legend.

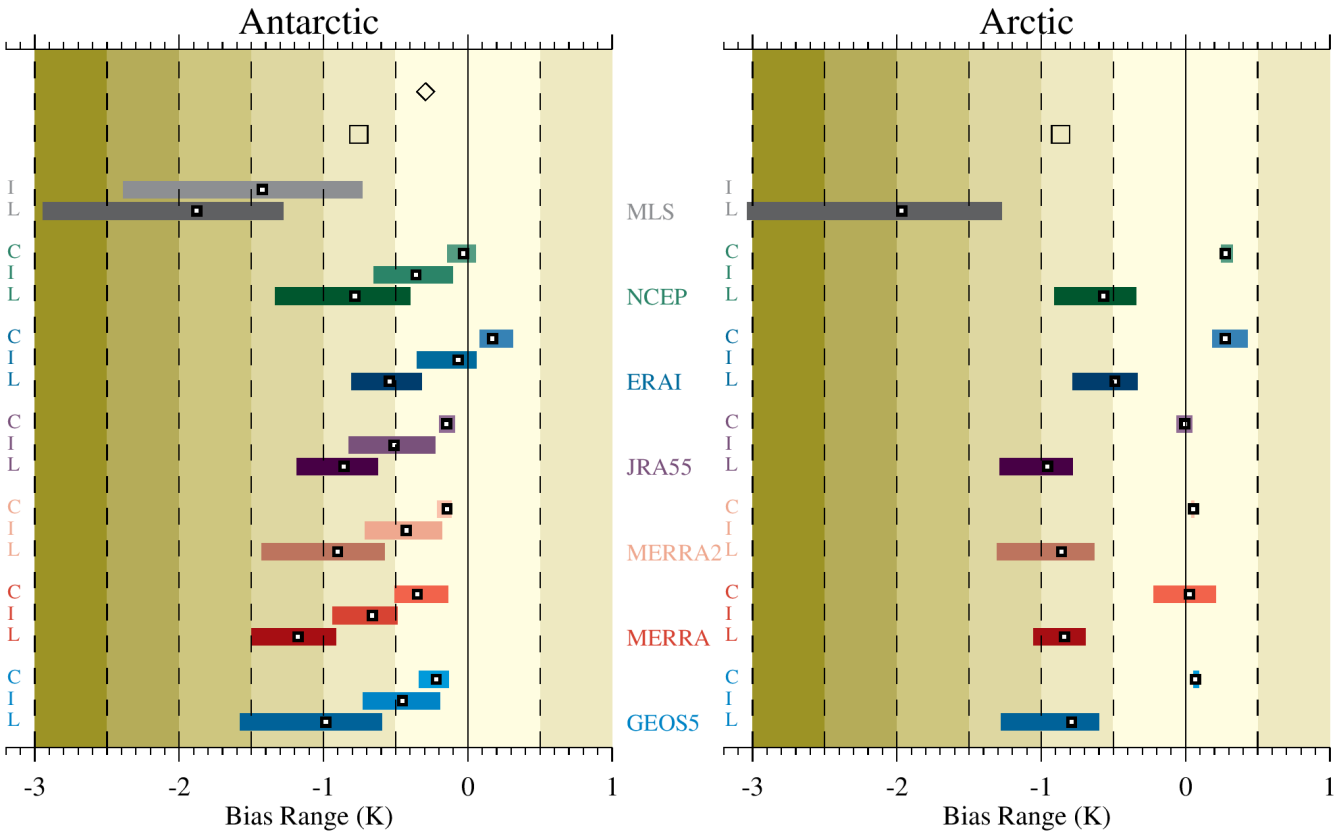

Figure 20. Temperature bias ranges of the reanalyses, and MLS, relative to the LIQ (L) and ICE (I) equilibrium references, and COSMIC (C), for Antarctic and Arctic winters 2008–2013, poleward of 60°, and for pressure levels from 68–21 hPa, derived from the data in Figures 16 and 17. The bias ranges are obtained from the extrema of the yearly mean bias values over 68–21 hPa weighted by the yearly standard deviations. Each horizontal colored bar indicates the range of the minimum to maximum bias for MLS and the particular reanalysis scheme given 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, obtained from Figure 19. 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.