The evolution of the stratopause during the 2006 major warming: Satellite data and assimilated meteorological analyses

Gloria L. Manney,1,2 Kirstin Krüger,3 Steven Pawson,4 Ken Minschwaner,2 Michael J. Schwartz,1 William H. Daffer,1 Nathaniel J. Livesey,1 Martin G. Mlynczak,5 Ellis E. Remsberg,5 James M. Russell III,6 and Joe W. Waters1

Received 25 June 2007; revised 25 January 2008; accepted 15 February 2008; published 12 June 2008.

[1] Microwave Limb Sounder and Sounding of the Atmosphere with Broadband Emission Radiometry data provide the first opportunity to characterize the four-dimensional stratopause evolution throughout the life-cycle of a major stratospheric sudden warming (SSW). The polar stratopause, usually higher than that at midlatitudes, dropped by ~30 km and warmed during development of a major “wave 1” SSW in January 2006, with accompanying mesospheric cooling. When the polar vortex broke down, the stratopause cooled and became ill-defined, with a nearly isothermal stratosphere. After the polar vortex started to recover in the upper stratosphere/lower mesosphere (USLM), a cool stratopause reformed above 75 km, then dropped and warmed; both the mesosphere above and the stratosphere below cooled at this time. The polar stratopause remained separated from that at midlatitudes across the core of the polar night jet. In the early stages of the SSW, the strongly tilted (westward with increasing altitude) polar vortex extended into the mesosphere, and enclosed a secondary temperature maximum extending westward and slightly equatorward from the highest altitude part of the polar stratopause near the vortex edge. The temperature evolution in the USLM resulted in strongly enhanced radiative cooling in the mesosphere during the recovery from the SSW, but significantly reduced radiative cooling in the upper stratosphere. Assimilated meteorological analyses from the European Centre for Medium-Range weather Forecasts (ECMWF) and Goddard Earth Observing System Version 5.0.1 (GEOS-5), which are not constrained by data at polar stratopause altitudes and have model tops near 80 km, could not capture the secondary temperature maximum or the high stratopause after the SSW; they also misrepresent polar temperature structure during and after the stratopause breakdown, leading to large biases in their radiative heating rates. ECMWF analyses represent the stratospheric temperature structure more accurately, suggesting a better representation of vertical motion; GEOS-5 analyses more faithfully describe stratopause level wind and wave amplitudes. The high-quality satellite temperature data used here provide the first daily, global, multiannual data sets suitable for assessing and, eventually, improving representation of the USLM in models and assimilation systems.


1. Introduction

[2] Understanding the dynamics of the upper stratosphere/lower mesosphere (USLM) is of interest for studies related to climate change and ozone recovery: USLM temperatures are a sensitive indicator of climate change [e.g., Rind et al., 1998; Manzini et al., 2003; WMO, 2003]. Via changes in the Brewer-Dobson circulation [e.g., Andrews et al., 1987], USLM dynamical changes influence the entire stratospheric circulation [e.g., Hitchman et al., 1989; Garcia and Boville, 1994; Shepherd, 2007], which can in turn affect tropospheric weather and climate (Baldwin et al. [2007] and references therein). Stratospheric sudden warmings (SSWs) provide perhaps the most dramatic dem-
onstracion of coupling of the USLM with the lower atmosphere, and some of the most dramatic changes during SSWs occur in the USLM ([Labitzke [1972], Siskind et al. [2005, 2007], and references therein]). Until recently, there have been few global, vertically resolved data available on a daily basis suitable for detailed study of USLM dynamics, especially during SSWs.

[3] Grided meteorological analyses from data assimilation systems have long been an important tool for studying day-to-day and long term variations in the stratosphere. Some operational assimilated analyses, including the Goddard Earth Observing System (GEOS) analyses from NASA’s Global Modeling and Assimilation Office (GMAO) and the European Centre for Medium-Range Weather Forecasts (ECMWF), have extended their products through the USLM. However, their highest altitude inputs are currently from nadir sounding satellite radiances representing a very broad layer (~20 km) in the upper stratosphere, and the models have tops in the mesosphere. Analyzed USLM fields are thus from very near the models’ tops and largely unconstrained by data, relying primarily on the underlying general circulation models (GCMS), which have differing treatments of the mesosphere, model top effects, and gravity wave (GW) and other parameterizations. Since data for comparisons have been sparse, the quality of the analyses in the USLM is largely unknown. Previous studies showed increasingly large discrepancies between different analyses above the middle stratosphere [e.g., Manney et al., 2005c] and since the USLM is very near the models’ tops, decreasing ability in this region is not unexpected. Since long term reanalyses that might be used for trend or interannual variability studies are planned from these assimilation systems, since winds and temperatures from these analyses are widely used for dynamical and transport calculations, and since improving GCMS and hence climate and forecast models hinges upon understanding current limitations, it is valuable to determine to what degree and which features of USLM dynamics are represented well in these products.

[4] Labitzke [1972] used rocket profiles and radiation calculations to schematically describe the evolution of the stratosphere and mesosphere during SSWs. From an initial stratospheric altitude near 60 km in early winter, she showed the stratosphere drooping by ~20 km and warming, accompanied by mesospheric cooling; when a complete breakdown of the stratospheric vortex took place (during “major” SSWs), the warm stratosphere layer was entirely destroyed. Local temperature maps up to 60 km derived from rocket observations over a week, and Selective Chopper Radiometer (SCR) satellite radiances in a ~20 km layer centered near 42 km, allowed a first glimpse of some features of USLM synoptic structure. Dropping and warming of the stratosphere has also been reported during “minor” (without a complete vortex breakdown) SSWs using single-station lidar data [e.g., von Zahn et al., 1998]. A number of other studies have used mesospheric (typically above ~70 km) station data in conjunction with meteorological analyses to study relationships between stratospheric and mesospheric winds and temperatures, showing mesospheric wind reversals preceding those in the stratosphere (e.g., Hoffmann et al. [2002, 2007], and references therein) and confirming mesospheric cooling associated with SSWs (e.g., Walterscheid et al. [2000], Hernandez [2003], Cho et al. [2004], and references therein).

[5] Hitchman et al. [1989] used Limb Infrared Monitor of the Stratosphere (LIMS) and Nimbus-6 Pressure Modulator Radiometer (PMR) data to show the seasonal zonal mean evolution of the polar winter stratosphere, documenting the elevated “separated” polar winter stratosphere, and showing that its formation is consistent with the results of GW processes.

[6] The data available for these, and other, previous observational studies were limited by combinations of poor resolution (e.g., PMR with ~15 km resolution), limited vertical extent (e.g., LIMS data, with good resolution (~3 km), but extending only to 65–70 km, lidar and rocket data typically to ~65–70 km, radar data only above ~70 km), lack of coverage (e.g., rocket, lidar or other single-station data), and/or short-term availability (e.g., LIMS data for ~7 months in 1978–1979, daily or near daily lidar data for sporadic, short periods). Other data sets extending through the mesosphere (e.g., temperatures from the Upper Atmosphere Research Satellite (UARS) Halogen Occultation Experiment (HALOE), which measured two latitudes each day and did not cover the polar winter) suffer similar limitations. Because of these limitations, it has not previously been possible to study USLM evolution in detail throughout an entire SSW lifecycle.

[7] The Sounding of the Atmosphere with Broadband Emission Radiometry (SABER) instrument [Mylczak and Russell, 1995] has been providing daily hemispheric temperatures with good quality and resolution since January 2002; mesospheric cooling associated with stratospheric sudden warmings has been reported using SABER data [e.g., Coy et al., 2005; Siskind et al., 2005]. However, the ~60-day “yaw” cycle whereby SABER observes high latitudes in one hemisphere at a time limits studies of high-latitude synoptic to seasonal evolution; thus, SABER alone did not view the whole lifecycle of recent (see below) major SSWs. Since August 2004, the Earth Observing System Aura Microwave Limb Sounder (MLS) [Waters et al., 2006] has provided high-quality daily near-global temperature profiles from the upper troposphere through the mesosphere, with coarser vertical resolution than SABER, but denser horizontal measurement spacing. These two data sets together provide data for much more detailed study of USLM evolution and structure than ever before possible.

[8] One of the strongest, most prolonged SSWs on record began in January 2006; Figure 1 shows a reversal of 10 hPa zonal mean winds north of 60°N to easterly (the criterion, with accompanying reversal of high-latitude temperature gradients, for a major SSW) on 21 January. An accompanying strong wind reversal extending below 60°N at 1 hPa preceded the 10-hPa reversal by about 6–8 d, with an earlier, brief reversal about 1 January. The scaled potential vorticity (sPV, in “vorticity units” as described by Dunkerton and Delisi [1986] and Manney et al. [1994b]) maps in Figure 2 show the synoptic evolution of this event in the middle and upper stratosphere (850 and 1500 K) and near the typical stratospheric altitude (2500 K). The evolution of the stratospheric circulation during the 2006 SSW was very similar to that during the 2004 major SSW described by Manney et al. [2005a], with easterlies
throughout the upper stratosphere at 10 hPa for over a month (Figure 1), a rapid recovery in early February to a very strong vortex in the upper stratosphere (Figure 1, 1 hPa, and Figure 2, right panels), weak recovery in the middle stratosphere (Figure 1, 10 hPa, and Figure 2, right panels) and below, and a very late final warming in May. Hoffmann et al. [2007] show the onset of mesospheric cooling and reversal of mesospheric (70–90 km) winds at stations in the high Arctic about 6 d before the 10 hPa wind reversal; during late January to late February (when mesospheric winds were easterly), they showed a large increase in gravity wave activity between 70 and 90 km. As shown below, and apparent in Figure 2 at 850 K on 16 January, the 2006 SSW was primarily a “wave one” event, with the vortex shifted far off the pole before fragmenting in a highly nonlinear manner (Figure 2, 30 January). Hoffmann et al. [2007] showed that the enhanced GW activity in the mesosphere began as stratospheric wave one amplitudes decreased in late January, and persisted until late February. Siskind et al. [2007] used SABER data to show an extremely high-altitude polar stratopause in mid-February 2006 after the January SSW; simulations with a high-altitude GCM suggested that this formation resulted from filtering by the disturbed stratospheric flow of GWs (generated by the model’s orographic

Figure 1. Latitude-time series of GEOS-5 (described below) zonal mean zonal winds at (top) 1 hPa and (bottom) 10 hPa, from 1 December 2005 through 15 May 2006.

Figure 2. Maps of GEOS-5 sPV at (top to bottom) 2500, 1500, and 850 K on (left to right) 16 and 30 January and 14 February 2006. Overlaid contours are static stability values of 3.6 x 10^-4 s^-2 (dashed) and 4.0 x 10^-4 s^-2 (solid). Projection is orthographic, with 0° longitude at the bottom and 90°E to the right; domain is 0 to 90°N.
GW scheme) that would ordinarily break near 50 km. As was first predicted by Holton [1983] and successfully demonstrated in an assimilation model study by Ren et al. [2008] only eastward propagating GW can propagate upward into the mesosphere during and after the onset of a SSW, producing an eastward net GW anomaly leading to an anomalous upwelling and cooling in the polar mesosphere.

[9] The evolution of the stratopause throughout the 2005–2006 Arctic winter and its relationship to polar vortex evolution is detailed here using MLS and SABER data, and compared with that seen in GMAO GEOS-5 (Goddard Earth Observing System Version 5.0.1) and ECMWF meteorological analyses. Comparisons of the high-altitude satellite data with the analyses in the USLM are used to assess the ability of those systems to represent the dynamics of the stratopause region during this extreme event.

2. Satellite Data and Assimilated Analyses

[10] Version 2.2 (v2.2) Aura MLS temperature and geopotential height data are used here, and have been validated by Schwartz et al. [2008]. MLS v2.2 temperatures show a slight high bias with respect to many correlative measurements in the middle and lower stratosphere, and alternating high and low biases in the upper stratosphere. The vertical resolution of MLS v2.2 temperature is ∼4 km in the stratosphere, and ∼8–9 km near the stratopause and in the lower mesosphere; precision is better than ∼1 K through the upper stratosphere, degrading to 2–2.5 K above that [Schwartz et al., 2008]. MLS profiles are available each day every ∼160 km along the orbit tracks, to 82° in each hemisphere. MLS orbit-track data are gridded at 2° latitude by 5° longitude using a weighted average around each grid point of all measurements in a given day. Horizontal winds are calculated from gridded MLS geopotential heights using a form of the primitive equations that neglects the vertical advection and time tendency terms [Randel, 1987; Newman et al., 1989]. PV is calculated from those winds and the gridded MLS temperatures as described by Newman et al. [1989] and Manney et al. [1996], as noted by Manney et al. [1996], winds calculated in this manner agree quite well with those from the assimilation systems except at low latitudes (differences in the upper stratosphere typically within ±2 ms⁻¹ and slightly more “noise” than corresponding assimilated winds). Winds, and especially PV (more highly derived) calculated this way are also somewhat noisy because of the coarse and irregular spacing of orbit-track measurements and the crude gridding used to convert to a regular latitude/longitude grid. However, at times and levels where assimilated analyses are expected to be accurate, MLS-derived winds agree well with those from the analyses, and PV fields agree well qualitatively (morphology and PV gradients).

[11] The SABER temperature data used are version 1.06 (v1.06). The effective vertical resolution is ∼2 km [Remsberg et al., 2003]; precision is better than ∼1 K, becoming somewhat larger (∼1.5 K) by the middle mesosphere. Test days of v1.06 SABER temperatures compare well with profiles from the UARS HALOE [Remsberg et al., 2002] and other correlative data. SABER samples every ∼300 km along the orbit track, to ∼84° in the hemisphere favored by the yaw cycle; it observes high northern latitudes from mid-January through mid-March in Arctic winter. SABER data are gridded in the same manner as for MLS, but with 4° latitude spacing. Winds and PV have been calculated from gridded SABER data in the same manner as for MLS, on a 4° latitude by 5° longitude grid.

[12] GEOS Version 5.0.1 (GEOS-5) analyses [Reinecker et al., 2007] have been produced for the period of the Aura mission, from August 2004 through the present, and have replaced GEOS-4 [Bloom et al., 2005] as the operational system. GEOS-5 uses the Grid point Statistical Analysis method of Wu et al. [2002], a 3D-Var system, and a 6-h analysis window. The interface between the observations and the GCM is performed using the incremental analysis update (IAU) approach [Bloom et al., 1996], which avoids shocking the model, thus producing smoother analyses. GEOS-5 analyses are provided on 72 model levels from the surface to 0.01 hPa, on a 0.5° latitude by 2/3° longitude grid. Orogenic GWs are parameterized in GEOS-5 as described by McFarlane [1987] and Kiehl et al. [1998], with directional variance of subgrid orography and strength of the resolved flow used to determine the GW source strength; propagation and breaking follow McFarlane [1987].

[13] A simple nonorographic GW parameterization [Garcia and Boville, 1994] is used to represent waves with nonzero phase speed that are important in the USLM; while this GW scheme conserves angular momentum in formulation, it is nonconservative in implementation in that waves are allowed to propagate out of the model top. The finite-volume dynamical core used in GEOS-5 [Lin, 2004] includes a monotonicity constraint that infers nonlinear diffusion that is local in space and time; the impacts of this are not straightforward to quantify. The model also includes a polar filter to damp the smallest spatial scales in the polar cap. The GEOS-5 output fields include PV calculated within the model, as well as model radiative heating rates.

[14] While GEOS-4 analyses are also available up to 0.01 hPa, comparisons with satellite and lidar data indicate a persistent high bias in GEOS-4 temperatures near the stratopause in high latitude winter [e.g., Manney et al., 2008] that is to a large degree alleviated in GEOS-5 (at least partly as a result of changes in the bias correction used in radiance assimilation in the upper stratosphere); because of this, and GEOS-4 having been discontinued after April 2007, we show only GEOS-5 results below.

[15] ECMWF uses a 4D-Var system based on a spectral GCM [e.g., Simmons et al., 2005]. ECMWF data shown here are from the T799/91-level system with a top at 0.01 hPa that became operational in February 2006 (e.g., Untch et al. [2006], available at http://www.ecmwf.int/publications/newsletters/, and references therein). Model level data from the T799/91-level system are used at levels up to 0.01 hPa; the data were extracted on a 2.5° × 2.5° horizontal grid. The T799/91-level system was running starting in October 2005, and pre-operational data were made available to users. The ECMWF orographic GW drag code, including flow blocking, is described by Lott and Miller [1997]; mountain wave propagation and drag follows the model of Miller et al. [1989]. In lieu of an explicit nonorographic GW parameterization, ECMWF uses Rayleigh friction at altitudes above 5 hPa to slow down the
otherwise too strong polar night jet. A $\nabla^4$ horizontal diffusion is applied, with large coefficients near the model top, to prevent reflection of waves at the top boundary, and to remove numerical aliasing at the end of the spectrum. This high-resolution version of the ECMWF model generates some gravity waves, e.g., Lee waves in versions with resolution of T512 or greater [e.g., Maturilli and Dörnbrack, 2006; Smith et al., 2006]. PV is calculated from ECMWF winds and temperatures as described by Manney et al. [1996].

[16] Both GEO-5 and ECMWF are operational systems, and they include very nearly the same data inputs. Neither assimilates products from research satellites such as MLS or SABER. In the upper stratosphere, Advanced TIROS Operational Vertical Sounder (ATOVs) radiances are the only data input; these constrain the analyses up to ~0.2 hPa, but the constraint is very weak above the peak of the uppermost channel (AMSU-A Channel 14) near 1 hPa.

[17] In section 3.3, total radiative heating rates calculated from each of the data sets are shown. A longwave band model adapted from the Community Climate Model, Version 2 (CCM2) radiation code [Briegleb, 1992a] is used. This is a 100 cm$^{-1}$ band model that considers infrared opacity by H$_2$O, CO$_2$, O$_3$, CH$_4$, N$_2$O, CFC-11, and CFC-12. The shortwave radiative heating is calculated using the 5-Eddington approximation with 18 spectral bands [Briegleb, 1992b]. The CCM2 radiation code has been validated against line-by-line longwave calculations and in comparisons with Earth Radiation Budget Experiment longwave fluxes and shortwave albedos [Briegleb, 1992a, 1992b]. H$_2$O and O$_3$ are derived from MLS profiles; other trace gases, tropospheric temperatures, and tropospheric clouds are based on a combined UARS climatology appropriate for high-latitude winter [Minschwaner et al., 1998]. To quantify the sensitivity of diabatic heating to temperature differences, all inputs are fixed for a given day, except for temperature: zonal mean profiles of 12-UT temperature (daily averages centered around 12 UT for the satellite data sets) from MLS, SABER, GEO-5, and ECMWF are used.

3. Stratopause Evolution in Satellite Data and Comparisons With Analyses

[18] Figures 3 through 5 provide an overview of stratopause structure as represented in satellite data and meteorological analyses. Figure 3 shows zonal mean 70°N temperatures, static stability, and zonal winds during the 2005–2006 Arctic winter, along with 70°N geopotential height wave one and wave two amplitudes. Static stability is calculated as given by Andrews et al. [1987] in log-pressure coordinates. To define the stratopause, we use the “warm-point” in temperature; this coincides closely with the $4 \times 10^{-4}$ s$^{-2}$ static stability contour shown in Figure 3, with (away from the period of the major SSW) the stratopause having generally higher values and the mesosphere lower values. Figure 4 shows the latitude dependence of stratopause altitude and temperature. Figure 5 shows snapshots of stratopause evolution in cross-sections of temperature around the 70°N latitude circle. In the following, we first describe stratopause evolution in the satellite data (section 3.1), then compare with the representation in the analyses (section 3.2), and finally discuss some implications of the evolution and comparisons (section 3.3).

3.1. Stratopause Evolution in MLS and SABER Data

[19] Before early January, the polar stratopause altitude was fairly constant near 55–60 km (Figure 3, left panels, and Figure 4). Through late December, the relatively sharp increase in zonal mean stratopause altitude and temperature near ~60°N (Figure 4) reflected the climatological pattern noted by Hitchman et al. [1989] during winter; zonal mean easterlies appeared in the USLM in early January (Figure 3), and descended through the stratosphere to fulfill the conditions for a major SSW by 21 January (Figure 1); the easterlies descended in two pulses, with the second associated with the major warming (see also Figure 1). Easterlies appeared nearly concurrently at all levels above ~1 hPa, and later at progressively lower levels below that. Easterlies extended down to 100 hPa by early February, indicating a complete breakdown of the stratospheric vortex; by this time the USLM vortex had already begun to recover and westerlies reappeared above ~0.1 hPa. Stratospheric wave one amplitudes were large in December through ~10 January, with significantly enhanced wave two amplitudes from ~5 to 10 January; a rapid further wave one amplification at the beginning of early January coincided with the first wind reversal (Figure 3), and even larger amplification accompanied by wave two amplification (due to a strong distortion of the vortex, e.g., Figure 2, 16 January) coincided with the start of the prolonged wind reversal.

[20] Concurrently with the onset of USLM easterlies, the stratopause began dropping and warming rapidly. The stratopause began to cool in mid- to late-January, when high-latitude USLM easterlies were at a maximum but the underlying stratosphere was still warming; the westerly USLM jet was redeveloping at lower latitudes (Figure 4), and the cooling started just poleward of that jet core. By late January, the stratopause was very low, but ill-defined: There was a cool, nearly isothermal region from ~30 to 0.1 hPa, and static stability (Figure 3) was nearly constant at about $4 \times 10^{-4}$ s$^{-2}$. This pattern reflected a complete breakdown of the warm stratopause layer, similar to that reported by Labitzke [1972]. During the period when the polar stratopause was low and ill-defined, there were still strong gradients in stratopause altitude near 60°N (Figure 4), indicating that the polar stratopause (in as far as defining it was meaningful) remained separated from that at midlatitudes. The midlatitude stratopause (~20–50°N in Figure 4) remained near the same altitude and warmed gradually throughout the winter.

[21] During the recovery period after SSW, MLS and SABER data show that the stratopause reappeared as a well-defined structure near the end of January near 0.01 hPa (~75 km) with much lower temperatures than before the SSW (Figures 3 and 4). The strong latitudinal gradients in stratopause altitude poleward of the jet core after its reformation (Figure 4) show the continuing separation of the polar from the midlatitude stratopause. During the recovery period, wave 1 amplified at very high altitude, near the level of the reforming stratopause, and gradually descended in parallel with the descent of the stratopause. Model studies have shown planetary waves to be forced in
situ in the mesosphere during SSWs by filtered GWs [e.g., Liu and Roble, 2002].

Above the reforming stratopause, Hoffmann et al. [2007] reported enhanced gravity wave activity and turbulent energy dissipation rates indicative of GW breaking near ~70–85 km. during the recovery period in February and March, consistent with reduced stratospheric planetary wave amplitudes (Figure 3) having allowed propagation of eastward-propagating GWs to higher altitudes than usual. This GW drag contributes to the cooling above the reforming stratopause. Siskind et al. [2007] used GCM simulations to show that modeling the high stratopause after the SSW in mid-February 2006 depended critically on the GW parameterization used. After mid-February, the stratopause dropped and warmed; enhanced radiative cooling in the USLM (modeled by Siskind et al. [2007], also see below), the descent of the altitude of strong planetary waves, and observations of enhanced mesospheric GW activity are consistent with the descent and warming of the stratopause being related to dissipation of mesospheric planetary waves and associated GW activity [e.g., Holton, 1983; Garcia and Boville, 1994; Siskind et al., 2007]. The underlying stratosphere continued to cool as the stratopause warmed and descended (Figure 3). By mid-March, the stratopause reached an altitude and temperature similar to those in early winter (and in more quiescent winters, e.g., Manney et al. [2008]).

Longitude-height temperature sections at 70°N (Figure 5) show the synoptic patterns of stratopause structure and differences in how it is represented in the satellite data and the analyses. Overlaid contours of negative eddy geopotential height indicate the structure and position of the polar vortex (white contours are more negative than black, so mark the vortex interior). Prior to, and in the early stages of, the SSW (Figure 5, 1 and 16 January), the stratopause was separated in longitude as well as latitude, with high temperatures at the minimum altitude tilting downward and eastward near 240–260°E, high temperatures at the maximum altitude tilting poleward and westward near 260–280°E and low temperatures in between; the higher portion of the stratopause corresponds to the portion within the polar vortex, which tilts strongly westward with height, as reported in previous studies of SSWs [e.g., Fairlie et al., 1990; Manney et al., 1994a, 2005b]. On both 1 and 16 January, the satellite data show a secondary temperature maximum extending up from the high altitude side of the
stratopause, overlying the stratopause from \( \sim 120 \) to \( 270^\circ \text{E} \) (Figure 5). On 16 January large wave one amplitudes extend to very high altitudes (also apparent in Figure 3), well into the mesosphere, with the strongly westward-tilting vortex enclosing the region of the secondary temperature maximum. The “double-stratopause” this appears to be an upward extension of the strong “baroclinic zones” previously reported before the peak of SSWs [e.g., Fairlie et al., 1990; Manney et al., 1994a, 2005b]. The model study of Fairlie et al. [1990] suggested that such baroclinic zones could be a source of upward-propagating inertia-gravity waves; such features would not likely be resolved by MLS or SABER. Examination of MLS and SABER data for other periods and for the Antarctic shows the double-stratopause to be a common pattern in winter, and its structure is discussed further below (section 4.1).

The 30 January plots in Figure 5 emphasize how nearly isothermal the polar stratosphere was and how ill-defined the high-latitude stratopause was after the SSW. Both MLS and SABER show weak local temperature maxima near 0.1 and 0.03 hPa. On 25 February, when the stratopause had dropped and warmed toward more typical values, the satellite data show the altitude varying smoothly around the latitude circle from about 0.01 to 0.68 hPa (Figures 5, with lowest temperatures in the transition regions between low and high stratopauses, and the higher stratopause in the polar vortex region. The overlaid eddy geopotential height contours show the reemergence of a strong vortex above \( \sim 40 \) km, shifted off the pole by a strong wave1 (e.g., Figure 3), but upright rather than tilted as before the SSW, consistent with typical stratospheric vortex evolution at this stage of SSWs [e.g., Fairlie et al., 1990; Manney et al., 1994a, 2005b].

[24] The evolution of temperatures, wave amplitudes, and synoptic vortex and temperature structure in the USLM agrees well between SABER and MLS data throughout the winter. The stratopause is typically slightly cooler in MLS than in SABER, consistent with the global-mean low bias in MLS temperatures between \( \sim 0.3 \) and 0.01 hPa reported by Schwartz et al. [2008]. The broader stratopause region in MLS when the stratopause is at higher altitudes is consistent with the coarsening vertical resolution there. MLS and SABER wave amplitudes and the locations (in time and altitude) of the maxima agree very well throughout the period of overlap, with slightly lower maxima from SABER; this may be related to the coarser horizontal coverage of SABER. Small quantitative MLS/SABER differences in the secondary stratopause feature seen on 16 January (Figure 5) are also likely related to the poorer vertical resolution of MLS in the mesosphere and the coarser horizontal sampling of SABER. The satellite data thus provide a consistent, detailed picture of stratopause structure and evolution that can be used to evaluate how it is represented in the assimilated analyses.

3.2. Comparisons of Stratopause Evolution With Analyses

[25] Stratopause evolution in the satellite data and the GEOS-5 and ECMWF analyses agrees well qualitatively until late January when the stratopause became ill-defined, and again after early March when the stratopause had dropped and warmed (Figures 3 through 5). However, the analyses do not capture the reformation of the stratopause at high altitude; in fact, the level where it reforms is very near or above the models’ tops. GEOS-5 shows the polar stratopause reforming near 0.08 hPa with much higher
temperatures than observed, whereas the ECMWF polar
stratopause remains cool and somewhat ill-defined until
early March, and warms later than that in the satellite data.
When the stratopause was cool in late January and February,
GEOS-5 temperatures were higher than those in the satellite
data and the cold region was confined nearer the pole
(Figure 4); ECMWF showed lower temperatures than sat-
ellite data, and they remained low longer, during this period.
The ECMWF temperature bias is likely related to the use of
Rayleigh friction instead of a parameterized nonorographic
GW scheme, since the temperature evolution described
above (section 3.1) is consistent with a residual meridional
circulation forced by GW drag producing enhanced descent
over the pole that warms the stratopause region [e.g., Garcia
and Boville, 1994; Siskind et al., 2007].

GEOS-5 and ECMWF zonal mean winds (Figure 3)
show larger peak values (both easterly and westerly for
winds) than in the satellite data; this may be related to the
high horizontal resolution of the analyses compared to the
satellite data sets, and/or to the balanced wind calculations
used to get winds from satellite data. Zonal mean winds
and waves in the analyses, especially the ECMWF, have strong
values extending considerably higher than those derived
from the satellite data; this suggests deficiencies in the
models related to the GW parameterizations (in conjunction
with treatment of the model top), with the Rayleigh friction
used in ECMWF being even less effective than the simple
nonorographic scheme used in GEOS-5 in “closing off” the
jets and planetary scale waves above the stratopause.

The secondary temperature maximum shown in
Figure 5 extends beyond the analyses’ model tops, and they
do not even capture the lower portion of it very accurately,
showing a lower, less extended secondary maximum with a
shallower upward slope. The analyses do, however, capture
the primary stratopause structure fairly accurately before and
near the beginning of the SSW (1 and 16 January). However,
the eddy geopotential heights once again reflect the analyses' failure to sufficiently reduce the vortex strength above the stratopause (e.g., 1 January, more pronounced in ECMWF).

When the stratopause was ill-defined after the warming (Figure 5, 30 January), the GEOS-5 analysis shows a vertically compressed version of the same pattern, overestimating them just below that, whereas the ECMWF analysis shows a better representation of the lower part of the pattern, but no overlying second maximum. Consistent with this, the structure of the weak vortex in ECMWF corresponds more closely in the altitudes within the model domain to that in the satellite data sets than does the structure in the GEOS-5 analysis.

On 25 February (Figure 5), though the stratopause had dropped closer to typical levels, the analyses capture only the lower portion of the temperature structure, below about 10 hPa (where the analyses begin to be quite well constrained by data). They place the higher-altitude maxima at lower levels than seen in the satellite data, thus having a much smaller range in stratopause altitude. As was the case before the SSW, the analyses once again have too strong a vortex at the higher levels. The analyses' temperature biases are particularly apparent on this day, with GEOS-5 showing much higher, and ECMWF much lower, temperatures than the satellite data in the higher altitude portion of the stratopause. GEOS-5 shows a sharp low bias in temperature just below the higher part of the stratopause; this is near the top level where direct temperature information from ATOVS is a significant constraint, and thus suggests that, in addition to the lack of assimilated information above this level, the procedures for assimilating ATOVS radiances could be important in the analyses.

There are some, albeit more modest than in the polar regions, biases between the satellite data and the analyses at midlatitudes (Figure 4). GEOS-5 and ECMWF stratopause altitudes are typically ~1–3 km higher than those in MLS and SABER data, with GEOS-5 biases reduced after the SSW. The ECMWF stratopause is up to ~5 K warmer than that in the satellite data, and the GEOS-5 stratopause slightly cooler (~0–5 K, larger differences in March). The midlatitude stratopause, at ~40–47 km, is well below the model tops and at an altitude where some temperature information is assimilated. Previous work [e.g., Manney et al., 2005c] has shown that at these altitudes (where the temperature information is from ATOVS soundings comprising data from three channels, each of which gives information on the temperature of a thick layer of the atmosphere) differences in temperatures between analyses increase, and are of similar magnitude to those seen between the analyses shown here and the satellite data.

The qualitative differences between GEOS-5 and ECMWF analyses highlight the dependence on the underlying models and parameterizations; in particular, the differences between GEOS-5 and ECMWF indicate that, in addition to the (equally) low model tops that limit both models’ ability to represent the stratopause, the different GW parameterizations are an important factor. Shaw and Shepherd [2007, and references therein] showed that the parameterized GW response to an imposed radiative perturbation is highly sensitive to model lid height and details of the GW formulation if momentum is not conserved; since the Rayleigh friction used by ECMWF is nonconservative in formulation and the scheme used by GEOS-5 is nonconservative in implementation, both systems are expected to be sensitive to these details. The fact that GEOS-4, using an identical nonorographic GW scheme to that in GEOS-5, had a larger high bias in stratopause temperatures than GEOS-5 [Manney et al., 2008] indicates that other factors are important as well; this particular difference is thought to be related to changes in the bias correction scheme for ATOVS radiances between GEOS-4 and GEOS-5.

GW drag is still expected to be important at lower altitudes via downward control [Garcia and Boville, 1994] and/or spurious downward influence of GW parameterizations [e.g., Shepherd et al., 1996; Shepherd and Shaw, 2004; Shaw and Shepherd, 2007]. Other factors that may play a significant role include radiation calculations (e.g., ozone climatologies used [Forster and Shine, 1999; Tegtmeier, 2006]), bias correction in data assimilation (e.g., Polavarapu et al. [2005b], especially in assimilating ATOVS radiances), vertical information propagation (also affected by GW drag feedbacks, e.g., Ren et al. [2008]), and horizontal diffusion and its possible effects on resolved GWs.

3.3. Implications of Stratopause Evolution and Representation in Analyses

The dramatic variations in the stratopause (and temperature evolution throughout the stratosphere and mesosphere) during the SSW compared to more quiescent winters have significant implications for transport. In addition, the implications of the misrepresentation of the stratopause structure and evolution in the analyses are significant for many studies using them, not limited to those focusing on the USLM. Radiative heating rates calculated from each of the data sets are shown here to elucidate some of these factors.

Figure 6 shows heating rates calculated using the CCM2 code (section 2) with daily 70°N zonal mean temperature profiles from each data set on selected days following the SSW (we use the same code, and vary only the input temperatures to highlight the dependence on temperature). Prior to the SSW (not shown), the qualitative structure of the temperature profiles, and hence the calculated radiative heating, agrees well (though modest GEOS-4 and ECMWF temperature biases even then lead to some quantitative differences). MLS and SABER calculations agree well throughout the period shown, with SABER showing slightly stronger diabatic cooling in the lower mesosphere consistent with the low bias in MLS temperatures there (the “kink” in MLS temperatures just above 1 hPa is an artifact related to instrumental and/or retrieval effects that are under investigation [Schwartz et al., 2008]). Comparison of MLS radiative heating rates for 2005 (a cold, relatively undisturbed Arctic winter) and 2006 (dotted and solid black lines, respectively) demonstrates strongly enhanced the radiative cooling above ~0.04 hPa following the 2006 SSW, consistent with the model calculations of Siskind et al. [2007]. However, radiative cooling in 2006 was substantially reduced through mid-February compared to that in 2005 at levels below ~0.1 hPa. Consistent with these calculations, heating rates provided with the GEOS-5 analyses in the two years show nearly twice the diabatic
cooling above ~0.4 hPa, with reduced cooling below, in February 2006 compared to February 2005. Enhanced mesospheric cooling (i.e., diabatic descent) in 2006 versus 2005 is consistent with the observational results of Randall et al. [2006] and Manney et al. [2008] and model results of Siskind et al. [2007] showing stronger descent of trace gases from the mesosphere into the strong vortex that reformed after the SSW.

The qualitative difference in GEOS-5 temperatures from those in the satellite data sets first shown in Figure 5 is evident here, notably the higher temperatures at the stratopause and sharp temperature minimum just below that near 0.6 hPa; this pattern leads to too much diabatic cooling between ~0.4 and 0.03 hPa, and unrealistic near-zero cooling near 1 hPa (this pattern is consistent with the diabatic heating rates provided in the GEOS-5 analyses). During the beginning of the recovery from the SSW, ECMWF temperatures (see also Figure 5) showed qualitative structure much closer to the satellite data sets than that in GEOS-5 below ~0.03 hPa, with lower temperatures than the satellite data above that. Consistent with this, heating rates calculated from ECMWF temperatures agree with those from MLS and SABER up to ~0.1 hPa; above ~0.03 hPa, ECMWF temperatures have a strong low bias from late January through most of February, reflected in much less diabatic cooling. Since enhanced radiative cooling contributes to the descent and warming of the stratopause during the recovery, insufficient enhancement in ECMWF and excessive enhancement in GEOS-5 (consistent with the opposite stratopause temperature bias seen in those data sets, GEOS-5 too warm and ECMWF too cool) acts to prolong the period of strong biases.

The dramatic differences in heating rates between the analyses and the satellite data imply that descent rates calculated from these heating rates would also be inaccurate during the recovery from the SSW. Since the analyses’ biases are strongly altitude dependent, the net impact on calculated descent is difficult to quantify; however, our expectations and rough calculations suggest that ECMWF would show a moderate low bias in the amount of descent from levels above ~0.1 hPa, and GEOS-5 a stronger high bias in a broader altitude range. Since these operational analyses are among those most commonly used for transport modeling (along with the Met Office stratosphere-troposphere assimilation products, which do not extend into the lower mesosphere and which more egregiously misrepresent stratopause structure and evolution at all times), it is important to be aware of these deficiencies, especially in modeling extreme events such as SSWs that are often of particular interest for transport studies. In any but the shortest calculations, significant biases extending down through the midstratosphere may affect even simulations focused on the lower stratosphere.

4. Stratopause and Polar Vortex Synoptic Structure and Evolution

In section 3.1 (e.g., Figure 5), it was apparent that the high polar stratopause before and during the recovery from the SSW was confined to the polar vortex region, with the vortex extending into the mesosphere before the SSW in
in (or above) the polar vortex. The secondary maximum seems to be associated with situations where the vortex tilts westward with height in the stratosphere (typical behavior during major and minor SSWs before their peaks [e.g., Manney et al., 1994a, 1994c, 2005a, 2005b]) on 16 January, when the secondary maximum was particularly strong and well defined. Figure 5 shows that the tilted vortex structure extended well into the mesosphere. We use MLS data and derived winds and PV fields here to examine stratopause and polar vortex structure and evolution in more detail during the SSW.

4.1. Synoptic Stratopause/Polar Vortex Relationship

Latitude/pressure sections of MLS temperature, with overlaid eddy geopotential heights, on 1 and 16 January 2006 and 25 February 2006 (Figure 7) at 270°E confirm, as indicated in Figure 5, that separation of the midlatitude and polar stratopause occurs across the edge of the polar vortex; on 25 February, well into the recovery and when the vortex is relatively pole-centered, the latitudinal structure appears similar to the zonal mean winter stratopause structure previously reported [e.g., Hitchman et al., 1989], with the high polar stratopause separated from that at midlatitudes across the edge of the polar vortex; a similar pattern appeared on 1 January, in the period leading to the SSW, when the polar stratopause had only just begun to drop (reflected in the zonal means in Figure 4). The 25 February polar stratopause was still considerably higher than during typical, less disturbed, winters. At longitudes where the polar vortex is strong (e.g., 270°E on both 1 January and 25 February) polar stratopause temperatures are near or slightly cooler than those at the midlatitude stratopause, but separated by a much cooler region. The secondary temperature maximum seen on 1 and 16 January is located directly in (or above) the polar vortex.

Maps of MLS stratopause altitude and temperature are shown in Figure 8; overlaid 2500 K GEOS-5 sPV contours are typically near the average high-latitude stratopause altitude before the warming and after the recovery (on 1 January, 14 and 25 February). As shown above, highest stratopause altitudes are in the vortex region; on 16 January, the upper stratospheric vortex had already broken down and was ill-defined (e.g., Figures 1 and 2), and on 30 January both vortex and stratopause were ill-defined. Highest stratopause temperatures are usually well outside the vortex and localized in the region between the vortex and the anticyclone (e.g., 1 January and 25 February); a pattern with highest temperatures between the vortex and anticyclone has previously been reported in the stratosphere during SSWs [e.g., Manney et al., 1994a, 2005a, 2005b]. In the vortex region, there tend to be both low and moderately high temperatures, with the higher temperatures in the vortex core and low temperatures near the vortex edge (e.g., 1 January, 14 and 25 February). This temperature pattern is persistent before, during, and after the SSW, and indicates that the higher zonal mean temperatures at the polar stratopause previously reported [e.g., Hitchman et al., 1989] and seen in Figure 4 result from averaging a more localized region of higher temperatures with lower temperatures in midlatitudes, with a broader (in longitude coverage) region of moderately high temperatures near the pole.

The 30 January plots in Figure 8 emphasize how nearly isothermal the polar stratosphere was immediately following the SSW, rendering the stratopause very ill-defined. Cold, nearly isothermal conditions covered the entire polar region north of ~60°N on 30 January. By 14 February, when the stratopause had dropped and warmed toward more typical values, both stratopause and vortex were again well-defined and relatively pole-centered.

Figure 2 showed the structure of the polar vortex in the middle through upper stratosphere in GEOS-5 analyses. The overlaid static stability contours of 4.0 (solid) and 3.6 × 10⁻⁴ s⁻² (dashed) indicate what portions of the maps are in the stratosphere (static stability greater than 4.0 × 10⁻⁴ s⁻²) and mesosphere. However, the large inaccuracies (qualitative as well as quantitative) in GEOS-5 temperatures at higher levels from mid-January through early February result in misrepresentation of the vortex/stratopause relation and also affect the sPV fields; these discrepancies become more acute at higher altitudes. Thus to examine the vortex structure across the stratopause, Figure 9 shows sPV maps and static stability contours calculated from MLS geopotential heights and temperatures from near stratopause level.
through the lower mesosphere. Although MLS-derived sPV fields are somewhat noisy, comparing the 2500 K panels of Figure 9 (bottom row) and Figure 2 (top row) shows fair agreement in the morphology of the sPV fields. The large difference between GEOS-5 and MLS static stability contours reflects the deficiencies in GEOS-5 temperatures above \( \sim 10 \) hPa during the SSW, and, consistent with the effect of this on the sPV fields, sPV agrees much more closely, especially in morphology, on days when the GEOS-5 static stability is more accurate (not shown). Disagreement in the magnitudes stems partly from the different calculations (the GEOS-5 PV fields being those provided by GMAO directly from the model output) and from unrealistically strong winds at stratopause level in GEOS-5 throughout the winter (section 3.2 and Figure 3). However, magnitude differences, as well as the differences in morphology, are also related to the biases and incorrect structure in the GEOS-5 temperature fields.

Before the SSW, the vortex region is in the stratosphere at 2500 K, with the extravortex region in the mesosphere (not shown). On 16 January, static stability values are characteristic of the mesosphere at all levels, except in a localized region at 3200 and 4000 K, where values characteristic of the stratosphere extend upward along the secondary temperature maximum shown in Figure 5. On 30 January, when the stratosphere was nearly isothermal, a pole-centered vortex was already starting to reform at 4000 and 3200 K, and static stability values (also see Figure 3) oscillate with altitude, with values typical of the mesosphere everywhere at 3200 K, and characteristic stratospheric values in the vortex regions at 2500 and 4000 K; these reflect the multiple weak temperature maxima seen on that day in Figure 5. By 14 February, a strong circum-polar vortex had reformed in the USLM, to below 1500 K (Figure 2), and only the vortex region was in the stratosphere above 2500 K (not shown).

### 4.2. Stratopause/Polar Vortex Relationship Summary

Figure 10 shows the structure of the stratopause in relation to the USLM jet as a function of equivalent-latitude (EqL, the latitude that would enclose the same area as a given potential vorticity (PV) contour, e.g., Butchart and Remsberg [1986]) and potential temperature. This summarizes in an average sense the vortex/stratopause relationships detailed above, but with a more vortex-centered perspective than that provided by zonal means shown in previous studies. PV fields and winds calculated from MLS are used here; MLS temperatures mapped with GEOS-5 high-resolution PV fields are qualitatively similar, while GEOS-5 winds show unrealistic values at higher levels (section 3.2). The 1 December panel shows a typical early winter structure (reflected in Figure 4), with an elevated

![Maps of MLS stratopause altitude (km, left) and stratopause temperature (K, right) on (top to bottom) 1, 16, and 30 January and 14 and 25 February 2006. Overlaid contours are GEOS-5 sPV values of 1.6 and 2.8 \( \times 10^{-8} \) s\(^{-1}\) at 2500 K. Projection is orthographic, with 0° longitude at the bottom and 90°E to the right; domain is 0 to 90°N. Glitches in contouring on 30 January occur when stratopause is ill-defined and algorithm to locate it fails.](image)
polar stratopause separated from the midlatitude stratopause across the polar night jet core (elevation does not appear as large in isentropic coordinates because vortex core temperatures are lower than those at mid- to low-EqLs). Similar structure persists through 30 December. The polar stratopause during this period is slightly cooler than that at mid-EqLs, reflecting the synoptic pattern seen in Figures 7 and 8 more clearly than the zonal means in Figure 4.

By 16 January, the vortex had broken down in the USLM (weak winds) and was rapidly weakening in the midstratosphere; the stratopause near 70°EqL had dropped to just above the altitude of the weak jet, while that at higher EqLs remained level with the mid-EqL stratopause and warmed. By 5 February, the polar stratopause had virtually disappeared; the polar vortex started to reform in the USLM, but with a lower latitude jet (larger vortex area). By 25 February, the vortex was very strong in the USLM, but still weak in the middle stratosphere. The mid-EqL stratopause tilted upward along the jet axis, transitioning to being very high and cool at highest EqLs. In the following weeks, the polar stratopause dropped further and warmed, but temperatures remained slightly lower than those of the mid-EqL stratopause (not shown).

5. Discussion and Conclusions

[46] MLS and SABER temperature data during the 2005–2006 Arctic winter provide the first opportunity to thoroughly characterize the temporal and spatial evolution of the polar upper stratosphere/lower mesosphere (USLM) throughout the lifecycle of a major stratospheric sudden warming (SSW). These data are used to provide a detailed four-dimensional picture of the stratopause and its relationship to polar vortex evolution during the SSW, and to assess the ability of assimilated meteorological analyses to capture aspects of that evolution.

[47] Stratopause evolution during the 2006 major SSW was characterized by the following:

[48] A drop of 20–30 km in polar stratopause altitude during the SSW, with warming during and cooling after the peak of the SSW. Mesospheric cooling accompanied the SSW. Similar behavior has been previously documented using localized measurements [e.g., Labitzke, 1972; von Zahn et al., 1998].

Figure 9. Maps of sPV calculated from MLS data at (top to bottom) 4000, 3200, and 2500 K on (left) 16 and (right) 30 January 2006. Overlaid contours are static stability values (calculated from MLS temperatures) of $3.6 \times 10^{-4}$ s$^{-2}$ (dashed) and $4.0 \times 10^{-4}$ s$^{-2}$ (solid). Projection is orthographic, with 0° longitude at the bottom and 90°E to the right; domain is 0 to 90°N.

Figure 10. EqL/potential temperature sections of MLS temperature (K) on (right to left, top to bottom) 1 December 2005, 16 January 2006, and 5 and 25 February 2006. EqL for mapping is from PV calculated from MLS data. Overlaid contours are wind speeds calculated from MLS data from 30 to 90 m/s by 10 m/s.
[49] Complete disappearance of the warm stratopause layer accompanying the vortex breakdown (consistent with schematic of Labitzke [1972]). The vortex and stratopause breakdown followed strong planetary wave amplification from ~30 to 0.1 hPa.

[50] Reformation of a cool stratopause at very high altitude, followed by a rapid drop and gradual warming (with concurrent cooling of the mesosphere above and the stratosphere below). This evolution during the recovery phase has not previously been documented. The stratopause reformation above ~75 km marks a range of stratopause altitudes of ~40 km over the lifecycle of the SSW.

[51] Planetary wave amplification in the mesosphere as the stratopause began to reform, with large-amplitude waves descending in parallel with the stratopause descent and warming.

[52] A previously undocumented secondary temperature maximum extending westward and equatorward from the separated polar stratopause over the midlatitude stratopause before the peak of the SSW. The “double-stratopause” extended upward and westward in the region of the strongly westward tilted stratospheric vortex (which extended well into the mesosphere as conventionally defined) characteristic of the early stages of SSWs.

[53] Separation of the polar from the midlatitude stratopause (as documented by Hitchman et al. [1989]) across the polar night jet core (vortex edge) throughout the winter including the period of the SSW, with only gradual changes in the midlatitude stratopause comparable to those in more quiescent winters.

[54] Radiative cooling rates in the lower mesosphere (above ~0.04 hPa) more than double those in an undisturbed winter (consistent with enhanced descent of trace gases [Randall et al., 2006; Manney et al., 2008]), with substantially reduced radiative cooling in the upper stratosphere (below ~0.1 hPa).

[55] GEOS-5 and ECMWF assimilated analyses show deficiencies in representing many aspects of USLM evolution during this extreme event:

[56] Failure to capture the secondary temperature maximum and the polar stratopause reformation at high altitude, features that extended above the model tops.

[57] Misrepresentation of details of the nearly isothermal region after the vortex breakdown, and underestimation of polar stratopause altitude variations during and following the SSW. While ECMWF showed realistic structure up to ~0.1 hPa during the stratopause breakdown and reformation, GEOS-5 shows unrealistic structure above ~10 hPa, with vertically compressed temperature variations and an unrealistic temperature minimum just below the stratopause.

[58] The analyses’ winds are too strong (both easterly and westerly) and planetary wave amplitudes too high in the region surrounding the stratopause, especially in the ECMWF analyses, in which Rayleigh friction is used as a surrogate for nonorographic GW drag.

[59] As a result of the unrealistic temperature structure and evolution, the analyses’ radiative heating rates are badly biased, with GEOS-5 strongly overestimating them above ~0.2 hPa and underestimating them near 0.4 hPa, and ECMWF strongly underestimating them above ~0.04 hPa.

[60] The inability of current analyses to capture many features of stratopause structure during these extreme conditions is no surprise given the limitations of the GCMs and assimilated data. The model tops at 0.01 hPa are in themselves a severe limitation, as demonstrated clearly in the failure to capture significant features that extend beyond the models’ tops. Along with the temperature differences, not only at the highest levels, but also extending through at least the middle stratosphere, differences in wind and wave amplitudes indicate that behavior of the analysis systems is very sensitive to the parameterizations of GWs and their treatment at the model top. Failure to accurately represent some features away from the model boundaries also raises questions about the adequacy of aspects of radiation calculations, procedures (especially bias correction) for assimilating ATOVS radiances (and dwindling of information from ATOVS above ~1 hPa), and the remote impact of data insertion at lower levels. Several groups are beginning to address such issues, for example, using assimilation systems with very high tops [e.g., Polavarapu et al., 2005a] and more sophisticated GW codes (S. Polavarapu, personal communication, 2007; Hoppel et al., 2008), testing upward propagation of effects of data inserted at lower levels [Ren et al., 2008], or assimilating SABER and MLS temperatures to provide direct data constraints in the mesosphere [Hoppel et al., 2008]; early results from both an assimilation using a higher model top and more comprehensive GW scheme and one assimilating MLS and SABER temperatures show promise of significant improvements [Hoppel et al., 2008; Polavarapu et al., 2008]. Comprehensive data sets such as those from MLS and SABER are critical to provide the high-quality, global, multiannual daily temperature data through the mesosphere that are needed to understand the USLM and to assess the representation of that region in models and assimilation systems. This assessment is in turn a critical prerequisite for studies aimed at improving the performance of the assimilation systems.

[61] Despite the above deficiencies, the assimilated analyses did, away from the extremely disturbed period of the SSW, qualitatively capture most features of the USLM, representing well the time evolution of stratopause altitude and structure until the vortex breakdown and after the recovery. There may thus be aspects of stratopause evolution and variability about which we can extract useful information from long-term reanalyses using these assimilation systems, or similar ones with modest improvements guided by detailed comparisons such as those provided here.

[62] Which analysis system is most appropriate to use in detailed scientific studies may depend on the focus of those studies: ECMWF analyses shown here represent the upper stratospheric temperature structure better than GEOS-5 under the extreme conditions during and after the SSW. GEOS-5 does, however, represent wind and planetary wave amplitudes better near the stratopause. For large-scale transport studies driven with GEOS-5 or ECMWF winds, the large biases in radiative heating rates in the USLM in both analyses must be considered in interpretation of the results, but rough calculations suggest that ECMWF represents descent better below ~0.1 hPa.

[63] The prolonged SSW in the 2003–2004 winter showed very similar evolution to the one in 2006 described herein [Manney et al., 2005a]. [Manney et al., 2008] used GEOS-4 with MLS and SABER data to compare polar
stratopause evolution in these two years and the contrasting cold, undisturbed 2004–2005 Arctic winter. GEOS-4 temperatures during the 2003–2004 SSW show a very similar pattern of stratopause evolution to the representation in GEOS-4 of that in 2005–2006; although SABER was viewing high southern latitudes prior to and during the development of the 2004 SSW, those data also show a nearly isothermal stratopause during vortex breakdown and reformation of a very high-altitude stratopause afterward. In contrast, the stratopause altitude in 2004–2005 remained near 55–60 km throughout the winter, and its altitude and time evolution were captured by GEOS-4 (albeit with significant temperature biases) and GEOS-5. \textit{Labitzke [1972]}, using rocket profiles, and other studies [e.g., \textit{von Zahn et al., 1998}] using local data, showed that strong stratopause warming and altitude changes can occur even when the stratospheric vortex does not break down in a major SSW; some GCM studies [e.g., \textit{Braesicke and Langematz, 2000}] suggest that strong stratopause warmings occur in most Arctic winters. Further study of stratopause evolution in existing and future satellite data, and in the historical record using the satellite data/analysis comparisons shown here and “spot checks” with previous sparser or short-term data sets as guidance for the types of behavior captured accurately by the analyses, will help advance our understanding of USLM dynamics, and allow development of a climatology of and evaluation of interannual variability in global stratopause structure and evolution. These studies are important for improving our ability to model the USLM and allowing exploration of possible implications of USLM variability/changes for the lower atmosphere and climate change.


References


