HOKKAIDO UNIVERSITY

Title	Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP)
Author(s)	Long, Craig S.; Fujiwara, Masatomo; Davis, Sean et al.
Citation	Atmospheric Chemistry and Physics, 17(23), 14593-14629 https://doi.org/10.5194/acp-17-14593-2017
Issue Date	2017-12-07
Doc URL	https://hdl.handle.net/2115/71108
Rights(URL)	https://creativecommons.org/licenses/by/3.0/
Type	journal article
File Information	acp-17-14593-2017.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Atmos. Chem. Phys., 17, 14593–14629, 2017
https://doi.org/10.5194/acp-17-14593-2017
© Author(s) 2017. This work is distributed under
the Creative Commons Attribution 3.0 License.

Atmospheric Chemistry and Physics
Open Access
EGU

Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP)

Craig S. Long[1], Masatomo Fujiwara[2], Sean Davis[3,4], Daniel M. Mitchell[5], and Corwin J. Wright[6]

[1]Climate Prediction Center, National Centers for Environmental Prediction, National Oceanic and Atmospheric Administration, College Park, MD 20740, USA
[2]Faculty of Environmental Earth Science, Hokkaido University, Sapporo 060-0810, Japan
[3]Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA
[4]Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, CO 80309, USA
[5]School of Geographical Sciences, University of Bristol, Bristol, UK
[6]Centre for Space, Atmosphere and Ocean Science, University of Bath, Claverton Down, Bath, UK

Correspondence to: Craig S. Long (craig.long@noaa.gov) and Masatomo Fujiwara (fuji@ees.hokudai.ac.jp)

Received: 28 March 2017 – Discussion started: 7 April 2017
Revised: 21 August 2017 – Accepted: 25 September 2017 – Published: 7 December 2017

Abstract. Two of the most basic parameters generated from a reanalysis are temperature and winds. Temperatures in the reanalyses are derived from conventional (surface and balloon), aircraft, and satellite observations. Winds are observed by conventional systems, cloud tracked, and derived from height fields, which are in turn derived from the vertical temperature structure. In this paper we evaluate as part of the SPARC Reanalysis Intercomparison Project (S-RIP) the temperature and wind structure of all the recent and past reanalyses. This evaluation is mainly among the reanalyses themselves, but comparisons against independent observations, such as HIRDLS and COSMIC temperatures, are also presented. This evaluation uses monthly mean and 2.5° zonal mean data sets and spans the satellite era from 1979–2014. There is very good agreement in temperature seasonally and latitudinally among the more recent reanalyses (CFSR, MERRA, ERA-Interim, JRA-55, and MERRA-2) between the surface and 10 hPa. At lower pressures there is increased variance among these reanalyses that changes with season and latitude. This variance also changes during the time span of these reanalyses with greater variance during the TOVS period (1979–1998) and less variance afterward in the ATOVS period (1999–2014). There is a distinct change in the temperature structure in the middle and upper stratosphere during this transition from TOVS to ATOVS systems. Zonal winds are in greater agreement than temperatures and this agreement extends to lower pressures than the temperatures. Older reanalyses (NCEP/NCAR, NCEP/DOE, ERA-40, JRA-25) have larger temperature and zonal wind disagreement from the more recent reanalyses. All reanalyses to date have issues analysing the quasi-biennial oscillation (QBO) winds. Comparisons with Singapore QBO winds show disagreement in the amplitude of the westerly and easterly anomalies. The disagreement with Singapore winds improves with the transition from TOVS to ATOVS observations. Temperature bias characteristics determined via comparisons with a reanalysis ensemble mean (MERRA, ERA-Interim, JRA-55) are similarly observed when compared with Aura HIRDLS and Aura MLS observations. There is good agreement among the NOAA TLS, SSU1, and SSU2 Climate Data Records and layer mean temperatures from the more recent reanalyses. Caution is advised for using reanalysis temperatures for trend detection and anomalies from a long climatology period as the quality and character of reanalyses may have changed over time.

Published by Copernicus Publications on behalf of the European Geosciences Union.

1 Introduction

Reanalyses are used in many ways, including as initial conditions for historical model runs, developing climatologies, comparison with experimental models, and the examination of atmospheric features or conditions over long periods of time. This paper mainly evaluates eight reanalysis data sets: NCEP/NCAR Reanalysis 1 (Kalnay et al., 1996; Kistler et al., 2001; referred to hereafter as "R1"; see Appendix A for abbreviations), ERA-40 (Uppala et al., 2005), JRA-25 (Onogi et al., 2007), NCEP/CFSR (Saha et al., 2010), ERA-Interim (Dee et al., 2011; referred to hereafter as ERA-I), MERRA (Rienecker et al., 2011), JRA-55 (Kobayashi et al., 2015), and MERRA-2 (Gelaro et al., 2017; GMAO, 2015), with some notes on NCEP/DOE Reanalysis 2 (Kanamitsu et al., 2002) (referred to hereafter as R2) and 20CR (Compo et al., 2011). See Fujiwara et al. (2017) for more information about these reanalyses. The ERA-15 (Gibson et al., 1997) is not included in this intercomparison due to its short period and subsequent replacement by ERA-40. When a reanalysis product is chosen for use in a study or comparison, the choice is made based upon several factors such as newness of the reanalysis systems, span of time evaluated, horizontal and vertical resolution, top layer, and observational data assimilated. In this paper, we present an intercomparison of these 10 reanalyses focusing mainly upon their temperature and zonal wind fields. The five more recent reanalyses (CFSR, MERRA, ERA-I, JRA-55, and MERRA-2) are the primary focus and we concentrate on how these reanalyses intercompare in the upper troposphere and entire stratosphere.

Intercomparisons of middle atmosphere winds and temperatures using reanalyses have been performed since the very first reanalyses were generated in the late 1990s. Pawson and Fiorino (1998a, b, 1999) were the first to evaluate reanalyses winds and temperatures by comparing R1 and ERA-15 analysis of the tropics before and after satellite data were used in the reanalyses. Randel et al. (2004) intercompared wind and temperature climatologies from R1, ERA-15, and ERA-40 along with meteorological centre analyses. R1 and the ERA-40 have been used by thousands of researchers for tropospheric studies. Notable middle atmosphere studies evaluating R1 and ERA-40 winds and temperatures include the following. Manney et al. (2005) used these two reanalyses along with other analyses to evaluate their ability to capture the unique 2002 Antarctic winter, while Charlton and Polvani (2007) intercompared the two for detecting Northern Hemispheric sudden stratospheric warmings (SSWs). Martineau and Son (2010) used temperature and wind fields from R1, R2, JRA-25, ERA-I, and MERRA to compare their depiction of stratospheric vortex weakening and intensification events against GPSRO temperature data. Simmons et al. (2014) intercompared the ERA-I, MERRA, and JRA-55 stratospheric temperature analyses over the 1979–2012 period showing where and when they agreed and disagreed and the reasons why they did so. They also pointed out the difficulties of the transition from the TOVS to ATOVS observations, most notably in the upper stratosphere and lower mesosphere. Lawrence et al. (2015) used polar processing diagnostics to compare the ERA-I and MERRA. They noted good agreement in the diagnostics after 2002, but cautioned that the choice of one over the other could influence the results of polar processing studies. Miyazaki et al. (2016) intercompared six reanalyses (R1, ERA-40, JRA-25, CFSR, ERA-I, JRA-55) to study the mean meridional circulation in the stratosphere and eddy mixing and their implications upon the strength of the Brewer–Dobson circulation. Fujiwara et al. (2015) used nine reanalyses (JRA-55, MERRA, ERA-I, CFSR, JRA-25, ERA-40, R1, R2, and 20CR) to examine their stratospheric temperature response to the eruptions of Mount Agung (1963), El Chichón (1982), and Mount Pinatubo (1991). Mitchell et al. (2015) performed a multiple linear regression analysis on the same nine reanalyses to test the robustness of their variability. Martineau et al. (2016) intercompare eight reanalyses (ERA-40, ERA-I, R1, R2, CFSR, JRA-25, JRA-55, and MERRA) for dynamical consistency of wintertime stratospheric polar vortex variability. Kawatani et al. (2016) compare the representation of the monthly mean zonal wind in the equatorial stratosphere with a focus on the quasi-biennial oscillation (QBO; Baldwin et al., 2001) among nine reanalyses (R1, R2, CFSR, ERA-40, ERA-I, JRA-25, JRA-55, MERRA, and MERRA-2).

The report by the SPARC Reference Climatology Group (SPARC, 2002) and the subsequent journal article by Randel et al. (2004) were in response to the need to compare and evaluate the then-existing middle atmosphere climatologies that were housed and made readily available to the research community at the SPARC Data Center. Both reports provide an intercomparison of eight middle atmosphere climatologies: UK Met Office data assimilation, NOAA Climate Prediction Center objective analysis, UK Met Office objective analysis using TOVS data, the Free University of Berlin Northern Hemisphere subjective analysis, CIRA86 (COSPAR International Reference Atmosphere, 1986), R1, ERA-15, and ERA-40. This intercomparison was mostly based upon analyses rather than reanalyses, as only the R1, ERA-15, and ERA-40 reanalyses were available at that time. Notable differences were found among analyses for temperatures near the tropical tropopause and polar lower stratosphere and zonal winds throughout the tropics. Comparisons of historical reference atmosphere and rocketsonde temperature observations with the more recent global analyses showed the influence of decadal-scale cooling of the stratosphere. Detailed comparisons of the tropical semiannual oscillation (SAO) and QBO showed large differences in amplitude among analyses; the more recent data assimilation schemes showed better agreement with equatorial radiosonde, rocket, and satellite data (e.g. Baldwin and Gray, 2005).

About 10 years after the SPARC climatology report (SPARC, 2002), SPARC started a new project, the SPARC

Table 1. Information about NCEP, JMA, ECMWF, and GMAO earlier and latter reanalyses pertinent to the stratosphere. Information includes the model version, horizontal and vertical resolution, model top pressure, and radiative transfer model (RTM) used.

	Model version		Horizontal resolution	Vertical resolution	Model top level	RTM
R1/R2	NCEP MRF	(1995/98)	T62: 1.875°	28 (σ)	3 hPa	Temperature retrievals
CFSR	NCEP CFS	(2007)	T382: 0.3125°	64 (hybrid $\sigma-p$)	$\sim$ 0.266 hPa	CRTM
ERA-40	IFS Cycle 23r4	(2001)	T_L159: $\sim$ 125 km	60 (hybrid $\sigma-p$)	0.1 hPa	RTTOVS-5
ERA-I	IFS Cycle 31r1	(2007)	T_L255: $\sim$ 79 km	60 (hybrid $\sigma-p$)	0.1 hPa	RTTOVS-7
JRA-25	JMA GSM	(2004)	T106: 1.125°	40 (hybrid $\sigma-p$)	0.4 hPa	RTTOVS-6: TOVS RTTOVS-7: ATOVS
JRA-55	JMA GSM	(2009)	T_L319: $\sim$ 55 km	60 (hybrid $\sigma-p$)	0.1 hPa	RTTOVS-9
MERRA	GEOS 5.0.2	(2008)	0.5° lat × 0.667° lon	72 (hybrid $\sigma-p$)	0.01 hPa	CRTM
MERRA-2	GEOS 5.12.4	(2015)	0.5° lat × 0.625° lon	72 (hybrid $\sigma-p$)	0.01 hPa	CRTM

Reanalysis Intercomparison Project (S-RIP; Fujiwara et al., 2017). The goals of this project are (1) to better understand the differences among current reanalysis products and their underlying causes, (2) to provide guidance to reanalysis data users by documenting the results of this reanalysis intercomparison, and (3) to create a communication platform between the SPARC community and the reanalysis centres that helps to facilitate future reanalysis improvements. This paper will present the key findings from the S-RIP Chapter 3 team on "Climatology and Interannual Variability of Dynamical Variables." In this paper we show the results from the eight "full-input" (Fujiwara et al., 2017) reanalyses, which are systems that assimilate surface and upper-air conventional and satellite data (i.e. MERRA-2, MERRA, ERA-I, JRA-55, CFSR, JRA-25, ERA-40, R1), though we will show one figure for 20CR, which is one of the "surface-input" reanalyses. We will concentrate only on the satellite era period of 1979 to 2014. Several of the reanalyses do not cover the entire span of the later period (e.g. ERA-40 ends in August 2002, 20CR ends in December 2012 (for its version 2), and JRA-25 ends in January 2014). The R2 is an updated version of the R1. Almost all of the changes and enhancements incorporated into the R2 were surface or boundary layer oriented. The only possible change to the stratosphere would be due to a change to a newer ozone climatology (Fujiwara et al., 2017). As a result, preliminary comparisons of R1 and R2 show very minor differences in temperatures and winds above the boundary layer. Therefore, we will not show R2 comparisons, but one can expect all R2 qualities to be nearly exactly the same as R1. All of the reanalyses except for the CFSR used the same forecast model and assimilation scheme throughout their time span. In 2010 the CFSR had an undocumented update to its GSI assimilation scheme. Another change to the CFSR occurred in 2011 with the implementation of the version 2 Climate Forecast System (CFSv2; Saha et al., 2014), in which the resolution, forecast model, and assimilation scheme were all upgraded. Fujiwara et al. (2017) distinguish this latter analysis as the CDAS-T574.

The rest of this paper will be organized as follows: Sect. 2 presents a summary of changes and improvements from each reanalysis centre's earlier to later versions. Section 3 presents and discusses temperature variability with time of the reanalyses. Section 4 presents the methodology used to compare the various reanalyses, the creation of a reanalysis ensemble mean (REM), and the ensemble mean attributes and variability with time. Section 5 presents the differences in the temperatures and winds in individual reanalyses from the REM. Section 6 examines the seasonal temperature amplitude of the reanalyses in the polar latitudes. Section 7 discusses the results of comparisons with observations that are not assimilated in the reanalyses by showing specific data analyses. Section 8 provides summaries and main conclusions.

We characterize the stratosphere into altitude ranges using the following generalizations: "upper" for 1 to 5 hPa, "middle" for 7 to 30 hPa, and "lower" for 50 to 100 hPa.

2 Improvements from older reanalyses to newer versions

Fujiwara et al. (2017) provide many details of each reanalysis, such as model characteristics, physical parameterizations used, observations assimilated, execution stream characteristics, and assimilation strategies. The most recent reanalyses are later generations of earlier versions (MERRA and MERRA-2, ERA40 and ERA-Interim, JRA-55 and JRA-25, CFSR, and R1 and R2). Using information contained in Fujiwara et al. (2017), we will highlight what we consider the major improvements and changes from the earlier version to the more recent version. Pertinent to the stratosphere, we present in Table 1 a summary of the earlier and latter reanalysis model used, model resolution, top pressure level, and radiative transfer model (RTM) used. Several reanalyses

improved their model horizontal and vertical resolution between versions. All of the later versions used a more recent version of the RTM. Explanations for the various labelling of horizontal resolution can be found in Fujiwara et al. (2017).

3 Reanalysis global mean temperature anomaly variability

The 1979–2014 period includes the assimilation of satellite observations in addition to the assimilation of conventional (surface, aircraft, and balloon) observations (see Fujiwara et al., 2017, for details). During this period, there are multiple transitions, additions, and removals of satellites and instruments observing the atmosphere. The calibration and quality control of the observations from these satellite instruments in many instances have improved over time from the earlier reanalysis systems to the more current reanalysis systems. The radiative transfer models used in the forecast models have also improved over time. Reanalysis centres devote major efforts to minimizing the transition from one satellite or observing system to the next (e.g. TOVS to ATOVS in 1998; see Fujiwara et al., 2017). However, the forecast models used by the reanalysis centres have their own biases throughout the atmosphere. If and how well the bias correction is performed will also dictate how the reanalysis uses these observations. Additionally, most reanalyses are not run as one stream, but rather it is more efficient timewise and computationally for the reanalysis to be broken up into multiple streams with overlap periods of at least 1 or more years. These overlap periods are intended to allow the new stream to spin up sufficiently to ensure minimal discontinuity when the older stream ends. Because of these factors, it will be shown that the more recent reanalyses have fewer discontinuities at different times throughout this data record than older reanalyses.

To illustrate how well the various reanalyses were able to transition between satellites and other data sources, Fig. 1 presents time series for each reanalysis of the global mean temperature anomalies from their own long-term (1979–2014) monthly means. In all of the time series plots, several climatic features are evident: the tropospheric warming during the 1998 and 2010 El Niño events (located on the time axis with an "e") and the lower and middle stratospheric warming associated with the El Chichón (1982) and Mount Pinatubo (1991) volcanic eruptions (located on the time axis with a "v"). However, the older reanalyses (ERA-40 and JRA-25) show several distinct discontinuities in the stratosphere. The ERA-40, which was the first reanalysis to assimilate SSU radiances, shows discontinuities during several changes in the NOAA polar satellites with the SSU instrument in the early 1980s. The ERA-40 assimilated both SSU and AMSU-A radiances from the end of 1998 through 2002 (Uppala et al., 2005). The JRA-25 shows smaller discontinuities in the 1980s but has an abrupt change in 1998 coincident with the immediate transition from TOVS (SSU, MSU) to the ATOVS (AMSU) observing systems. The bias correction schemes for the TOVS and ATOVS radiances were also different. The combination of both resulted in large discontinuity in the stratosphere (Onagi et al., 2007). Of the five more recent reanalyses, the CFSR shows multiple discontinuities in the upper and middle stratosphere. This is because the CFSR is made up of six streams (end years 1986, 1989, 1994, 1999, 2005, 2009) and also because it corrects the biases in the SSU channel 3 observations with a forecast model that has a noted warm bias in the upper stratosphere. After 1998 the CFSR only used the AMSU-A radiances (it did not assimilate channel 14) and just monitored the SSU channels (Saha et al., 2010). The ERA-I shows two distinct discontinuities: in 1985 from the transition from NOAA-7 SSU to NOAA-9 SSU and in August 1998 when ATOVS observing systems began to be assimilated. ERA-I assimilated both SSU and AMSU-A radiances until 2005. Channel 3 of the SSU prior to August 1998 and AMSU-A channel 14 were not bias corrected. After August 1998 the SSU channel 3 radiances were bias corrected (Simmons et al., 2014). MERRA merged the SSU and AMSU observations over a period of time. The version of the CRTM (Han et al., 2006) that MERRA used for other satellite radiances was not able to work with the SSU radiances, and as an alternative the GLATOVS (Susskind et al., 1983) was used. The latter was not updated with the necessary adjustments to the channels due to pressure cell leaks and changes in the stratospheric CO_2 concentration (Gelaro et al., 2017). MERRA immediately stopped using the SSU channel 3 in October 1998 but continued to assimilate channels 1 and 2 through 2005. JRA-55 also merged the SSU and AMSU observations, but for a shorter overlap period of 1 year, and bias corrected all the SSU and AMSU-A channels (Kobayashi et al., 2015). MERRA-2 shows a discontinuity in 1995 from the transition from NOAA-11 to NOAA-14 SSU channel 3 radiances. A second discontinuity occurs when MERRA-2 immediately transitions from SSU and MSU to the AMSU in October 1998. A third discontinuity occurs when it begins using Aura MLS observations in August 2004. Just as with MERRA, MERRA-2 did not bias correct SSU channel 3 and AMSU-A channel 14. MLS temperatures were used to remove a bias in the upper stratosphere and to sharpen the stratopause (Gelaro et al., 2017). R1, R2, and the 20CR reanalyses only extend up to 10 hPa due to their fewer model layers, so the upper stratosphere is not analysed. R1 and R2 use NESDIS-derived temperature retrievals, which minimized satellite transitions. The 20CR is shown as an example that assimilated only surface-based observations. Therefore, it shows no discontinuities, but its forecast model included the volcanic aerosols and the historical changes in carbon dioxide to produce interannual variations in the stratosphere (see Fujiwara et al., 2017, for more details).

The timing and degree of these discontinuities will play a role in how well the various reanalyses compare with

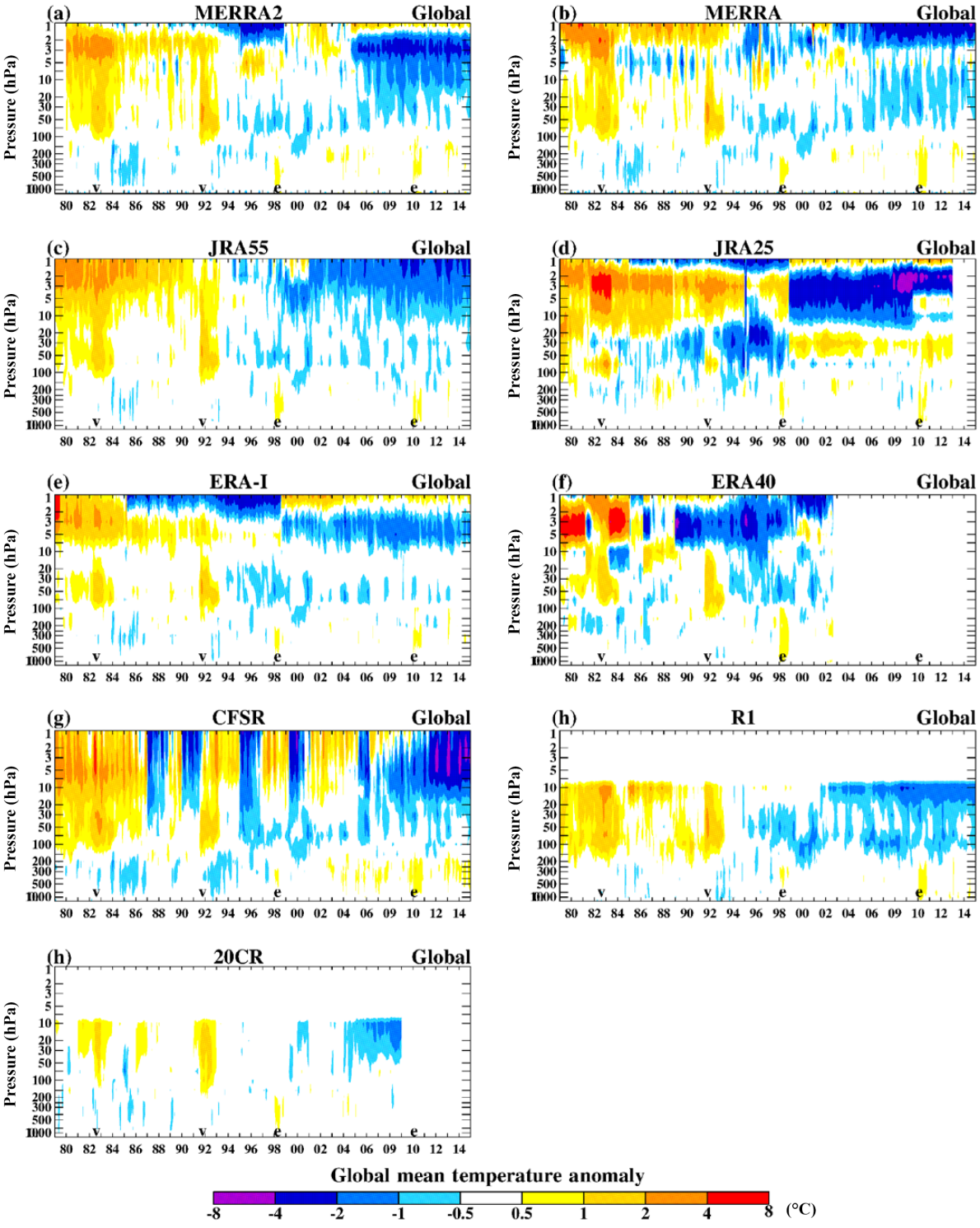


Figure 1. Pressure vs. time plots of the global mean temperature anomalies (K) of reanalyses. The anomalies are from the monthly climatology of each reanalysis. The reanalyses shown are **(a)** MERRA-2, **(b)** MERRA, **(c)** JRA-55, **(d)** JRA-25, **(e)** ERA-Interim, **(f)** ERA-40, **(g)** CFSR, **(h)** R1, and **(i)** 20CR. Note that R1 and 20CR do not provide analyses above 10 hPa; "v" and "e" denote the occurrence of volcanos and El Niños.

each other over time. Difficulties associated with assimilating the SSU observations due to their CO_2 pressure-modulated cells slowly leaking, and the changing of atmospheric CO_2 impaired the earlier reanalyses (ERA-40, JRA-25, MERRA). The more recent reanalyses should agree more closely with each other after 1998 because there are fewer issues assimilating the ATOVS, AIRS, and GPSRO observations (MERRA did not assimilate GPSRO data.)

Because of these discontinuities and transitions discussed above, reanalyses should be viewed very carefully for use in trend analysis and trend detection, especially in the middle and upper stratosphere.

4 Reanalysis ensemble mean (REM)

4.1 Methodology

No one reanalysis is the de facto standard for all variables and processes. Consequently a reanalysis ensemble mean (REM) of three of the more recent reanalyses (MERRA, ERA-I, and JRA-55) will be used as the reference from which differences and anomalies will be determined. The CFSR is excluded from the REM primarily because of the stream-change impacts upon the temperature structure in the middle and upper stratosphere. MERRA-2 is not included in the REM because it had just become available at the time of the preparation of this paper and does not include 1979. The data sets used to perform the intercomparisons are monthly mean zonal means at a 2.5° resolution. Standard post-processed pressure levels are used (1000, 850, 700, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10, 7, 5, 3, 2, 1 hPa). The focus time period of this intercomparison is 1979 through 2014. The current WMO 30-year climatology period (1981–2010) will be the base period of the climatology used. It should be noted that most reanalyses, with the exception of MERRA and MERRA-2, provide data below the surface for some regions (e.g. at 1000 hPa under Antarctica and the Tibetan Plateau). These data are calculated via vertical extrapolation. When the REM is created, re-gridded zonal means are first calculated for each reanalysis, and then the three data sets are averaged where valid data exist. Since most of the latitude zones poleward of 60° S are part of the Antarctic land mass with surface elevations reaching 3 km, pressures higher than 700 hPa have invalid data and hence are not analysed.

4.2 Climatology of the REM

4.2.1 Temperature

The seasonal variation in the REM temperature monthly means and their interannual variability in three different zonal regions (60–90° N, 10° S–10° N, and 90–60° S) are shown in Fig. 2. It is of note that at polar latitudes the lowest temperatures occur in the upper stratosphere in November (for the Northern Hemisphere, NH) and May (for the Southern Hemisphere, SH) and descend with time such that the lowest temperatures in the lower stratosphere do not occur until January in the NH and September in the SH. Thus, when lower stratospheric temperatures are reaching a minimum, upper stratospheric temperatures are already increasing. The upper stratosphere polar circulation is well defined prior to solstice shutting down any meridional advection of heat into the polar region. Consequently, radiative cooling drives the temperatures to their lowest values prior to solstice. The lowest temperatures occur at about 30 hPa in both polar regions. However, the lowest SH polar temperatures are more than 15 K colder than the lowest NH polar temperature. The interannual variability graphs show that the greatest variability in the NH temperatures is in the upper stratosphere in February when wave activity is most pronounced. In the SH the greatest variability occurs in October and November, associated with the winter to spring transition from low to high temperatures when wave activity becomes significant in that hemisphere. This variability is associated with how quickly that transition occurs. In some years the circulation over Antarctica is very zonal and stable, which prolongs the period of low temperatures in the polar latitudes. In other years there may be greater wave activity transporting heat from the extratropics into the polar latitudes, thus shortening the period of low temperatures. In the tropics, the variability is much smaller than in the polar regions but is associated with the phase of the SAO and the QBO in the upper and middle stratosphere, respectively.

4.2.2 Zonal wind

The seasonal variation in the REM zonal wind monthly means and their interannual variability in three different zonal regions (40–80° N, 10° S–10° N, and 80–40° S) are shown in Fig. 3. In the NH polar jet region (40–80° N) the maximum winds occur in the upper stratosphere in November and December, and the greatest variability occurs from December through March. In the SH polar jet region (80–40° S) wintertime westerlies are about 30 m s^{-1} stronger than the wintertime NH westerlies. These stronger westerlies are due to the much weaker disruption of the polar vortex by the vertically propagating planetary-scale waves and the stronger temperature gradients. Similar to the temperature variability, the variability in the SH polar night jet between May and August is not as great as in the NH polar jet. The SH zonal wind variability increases during the final warming and transition from westerlies to easterlies as wave activity increases from August through November.

In the tropical upper stratosphere, there is a strong semi-annual oscillation (SAO; Ray et al., 1998) with maximum westerlies of up to 20 m s^{-1} at equinox and intervening easterlies during the solstice periods. There is a marked asymmetry in the amplitude of the easterly SAO phase, with amplitudes of -40 to -50 m s^{-1} in the easterlies in December to February but only -20 to -30 m s^{-1} in July–September.

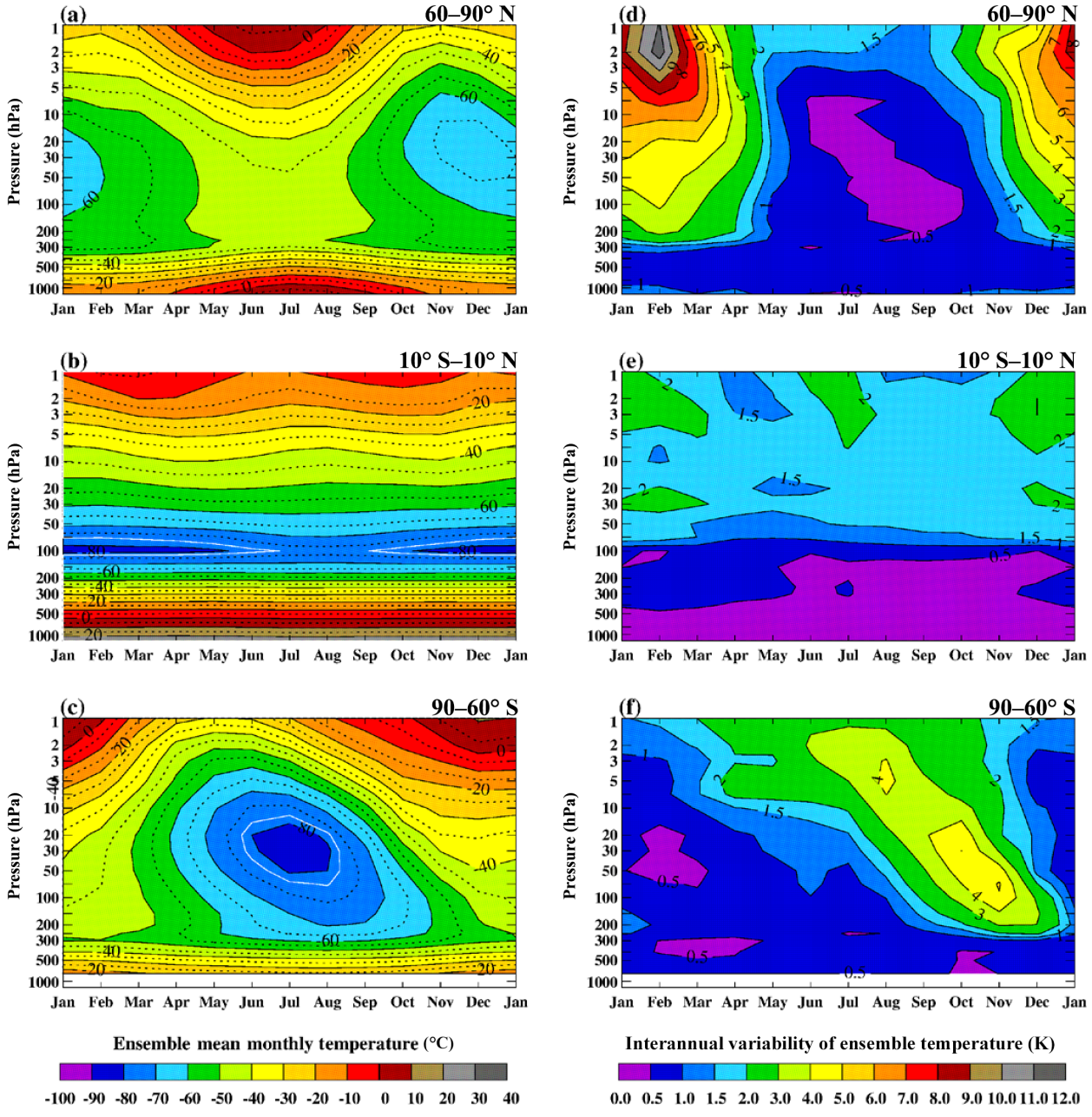


Figure 2. Annual variation in the REM temperature monthly means (°C) and their SD (K) in three different zonal regions: 60–90° N **(a, b)**, 10° S–10° N **(c, d)**, and 90–60° S **(e, f)**.

The easterly SAO phase is believed to result from the advection of easterlies from the summer hemisphere by the Brewer–Dobson circulation (Gray and Pyle, 1987), and this asymmetry is consistent with the much stronger circulation in December to February associated with greater wave activity in the NH winter. In the equatorial mid-stratosphere where the QBO dominates, the climatological winds in the tropical middle stratosphere have mean easterlies of -5 to $-10\,\mathrm{m\,s^{-1}}$. Because of the quasi-biennial nature of the winds, the interannual variability is very large, peaking between 10 and 20 hPa. The SAO wind transition in the upper stratosphere also shows a high amount of interannual variability.

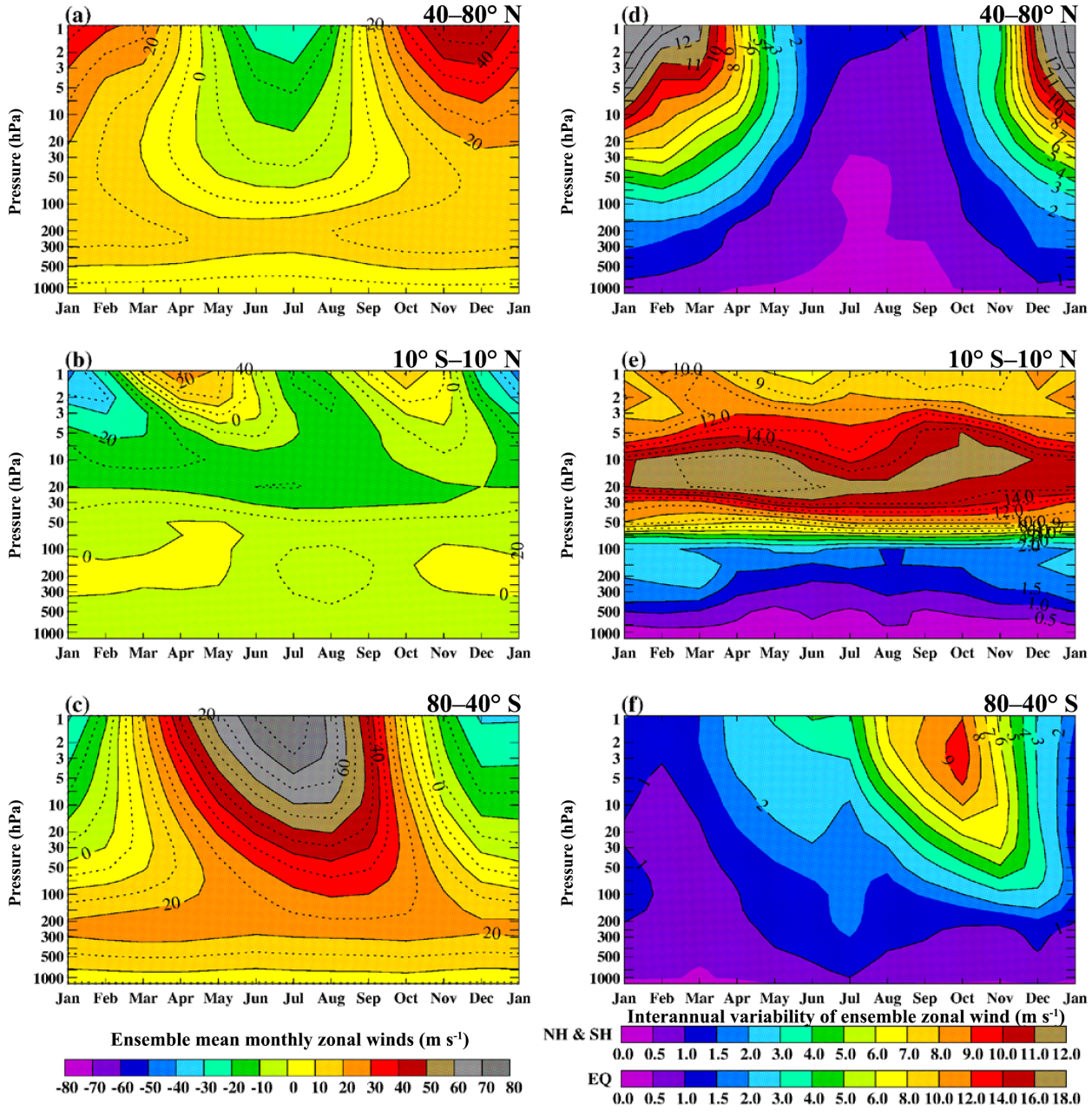


Figure 3. Annual variation in the REM zonal wind monthly means ($m s^{-1}$) and their SD ($m s^{-1}$) in three different zonal regions: 40–80° N **(a, b)**, 10° S–10° N **(c, d)**, and 80–40° S **(e, f)**.

4.3 Agreement among the REM members

4.3.1 Temperature

The previous section dealt with the mean of three of the more recent reanalyses (MERRA, ERA-I, and JRA-55). Now we examine their variability or "degree of disagreement" over time. We define the degree of disagreement as the SD of the three reanalyses for each month, for each latitude zone, and for each pressure level for the 1979–2014 period. Latitude zones (e.g. 60–90° N) are the cosine-weighted summations of the 2.5° zonal SDs. We must note that agreement of the three reanalyses does not imply correctness because the three reanalyses could possibly have similar erroneous analyses. For some months in the upper stratosphere, the temperature disagreement can be greater than 5 K. Figure 4 presents pressure vs. time series plots of the temperature SD (K) of the three members of the ensemble. Figure 4 shows how the monthly temperature disagreement varies in three latitude zones (60–90° N, 10° S–10° N, and 90–60° S) in a time vs.

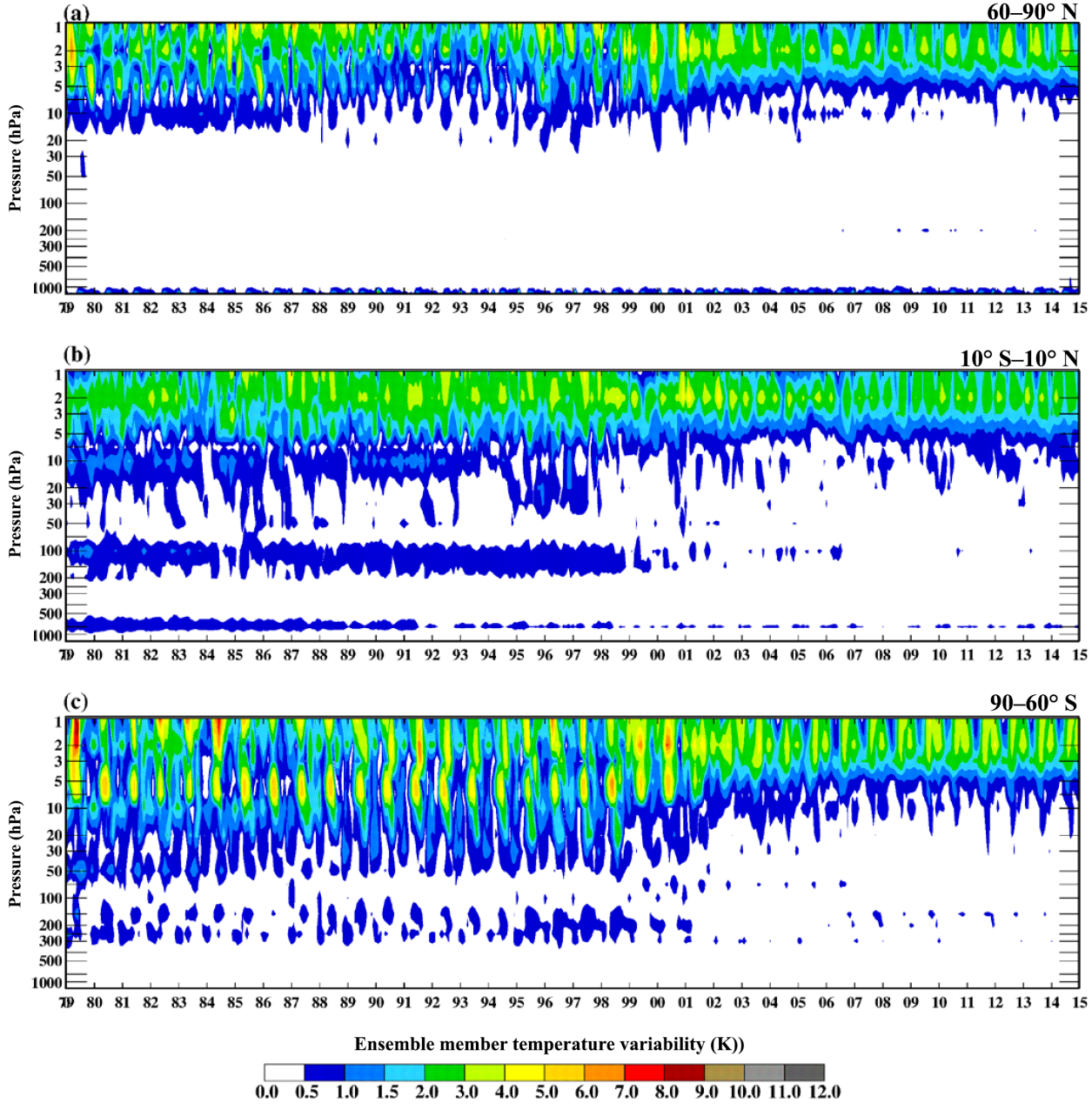


Figure 4. Pressure vs. time plots of the temperature SD (K) for each month of the three reanalyses making up the REM for three zonal regions: 60–90° N **(a)**, 10° S–10° N **(b)**, and 90–60° S **(c)**.

pressure plot. The mid-latitude plots are not shown but evaluations will be presented below. In all three latitude bands the disagreements are greatest at pressures lower than 10 hPa at which there are fewer conventional observations available for assimilation and the satellite observations generally have very broad weighting functions in the vertical. The 60–90° N plot shows that at pressures greater than the 20 hPa level, all three reanalyses agree with each other very well, with an SD smaller than 0.5 K. Generally, from 1979 to 2001 the pressure at which the 0.5 K difference contour occurs stays constant between 20 and 10 hPa. Interrupting this period during the 1990s, the NH polar activity was unusually quiet and cold (Pawson and Naujokat, 1999; Charlton and Polvani, 2007). Then from 2001 to 2014 the pressure at which the 0.5 K contour occurs moves upward to between 7 and 5 hPa. The increased agreement between 20 and 7 hPa is most likely due to the assimilation of AMSU and AIRS observations. The disagreement among the three reanalyses is greater in June–

August than in other months due to the ERA-I having warmer temperatures at this level than MERRA and JRA-55.

In the tropics, the disagreement maximizes in two separate layers: between 150 and 70 hPa during the TOVS period (1979–1998) and above 20 hPa throughout the entire 1979–2014 period. The former disagreement is at the vertical location of the cold point temperature. Apparently, there is greater disagreement among the three reanalyses in determining this temperature during the TOVS period than during the ATOVS period. During the TOVS period there are only four MSU channels sounding the troposphere and lower stratosphere. The AMSU-A instrument has five channels (5 through 9; 1 through 4 are water vapour channels) sounding the same layer. These additional channels provide information about the temperature structure near the tropopause, thus allowing the reanalyses to better analyse and agree upon the temperature structure there. The pressure at which the greatest differences (3–4 K) occur is 2 hPa and has a seasonally varying pattern.

In the 90–60° S zone, the disagreement among the three reanalyses extends lower into the stratosphere than the NH polar zone. This region encompasses all of Antarctica and the ocean surrounding it. There are very few observation sites in this latitude zone. Manney et al. (2005) and Lawrence et al. (2015) have shown that reanalyses of temperatures in the polar stratosphere can differ significantly depending on what observations are available. Differences greater than 0.5 K during the TOVS period extend to 70 hPa. There are two layers of greatest disagreement in the TOVS period: between 7 and 5 hPa and above 3 hPa. The disagreement between 7 and 5 hPa terminates after 2001, which may be due to the assimilation of AIRS radiances.

The northern mid-latitude (30–60° N) disagreement does not change significantly throughout the entire 1979–2014 period. Values larger than 0.5 K begin at pressures lower than 7 hPa and have summertime peaks of 2 K at pressures lower than 3 hPa.

The southern mid-latitude (60–30° S) disagreement is similar to the SH polar disagreement in that the disagreement during the TOVS period extends to higher pressures (between 20 and 30 hPa) than during the ATOVS period (between 7 and 10 hPa). Also similar to the SH polar region, there are two layers of greatest disagreement in the TOVS period between 7 and 5 hPa and above 3 hPa. The 7 to 5 layer disagreements also terminate after 2001, just as in the SH polar region.

4.3.2 Zonal winds

Figure 5 presents pressure vs. time series plots of the zonal wind SD ($m\,s^{-1}$) of the three members of the ensemble mean. Figure 5 shows the disagreement of the monthly ensemble members' zonal wind in the polar jet regions (40–80° N and 80–40° S) and in the tropics (10° S–10° N). As with the temperatures, the zonal wind disagreement in the mid-latitudes are not shown but are described in the text below. There is very good agreement of the zonal winds among the three reanalyses in the NH and SH polar jet regions with SDs smaller than $0.5\,m\,s^{-1}$. In the NH polar jet region, significant disagreement ($> 0.5\,m\,s^{-1}$) among the three reanalyses is consistently confined to pressures lower than 5 hPa. Disagreements greater than $0.5\,m\,s^{-1}$ are nearly eliminated after the transition to ATOVS observations occurs at the end of 1998.

The altitude range of disagreement greater than $0.5\,m\,s^{-1}$ in the SH polar jet region extends from the upper stratosphere down into the middle stratosphere (10–20 hPa) during the TOVS time period, but improves considerably in the ATOVS time period.

The tropical zonal wind disagreement shows much larger values of the order of $10\,m\,s^{-1}$ in the upper stratosphere than the polar jet values, resulting from disagreement in SAO and QBO winds and winds near the surface at 850 hPa. There is improvement with time in the agreement of the QBO winds and 850 winds, but this improvement does not extend to the SAO height region. The greater improvement in the NH and SH polar jet winds after 1998 vs. minor improvement in the equatorial winds illustrates the differences between the mechanisms controlling these winds. The polar jet winds are largely dictated by the latitudinal thermal gradient and resulting thermal wind. However, in the tropics the thermal wind relation breaks down and the wind fields are not well constrained by the assimilated satellite radiances. In addition, the tropical winds are primarily determined by the transfer of momentum from upward-propagating waves with spatial scales that are too small to be adequately resolved by the forecast models used in these reanalyses (Baldwin et al., 2001). The tropical winds are therefore highly dependent upon radiosonde observations for speed and direction (and these only extend to ~ 10 hPa). In general the amplitude of the reanalysis tropical winds are smaller than observations. Following the change to ATOVS data, the differences among the reanalyses decrease slightly. No single forecast model included in the REM is capable of generating a QBO on its own. To date, only the forecast model used in MERRA-2 is capable of doing so, and Coy et al. (2016) show that after 2000 the MERRA-2 QBO winds are greatly improved vs. those in MERRA.

The characteristics of the NH and SH mid-latitude regions (20–40° N and 40–20° S, respectively) are very similar to their respective polar jet regions. The NH mid-latitude disagreements during the TOVS period occur at pressures lower than 7 hPa and do not exceed $1.5\,m\,s^{-1}$. During the ATOVS period the disagreements are more sporadic and occur at pressures lower than 3 hPa.

The SH mid-latitude disagreement occurs at pressures lower than 20 hPa during the TOVS period with values not exceeding $4\,m\,s^{-1}$. During the ATOVS period the disagreements become more sporadic, smaller in value, and occur at pressures lower than 7 hPa.

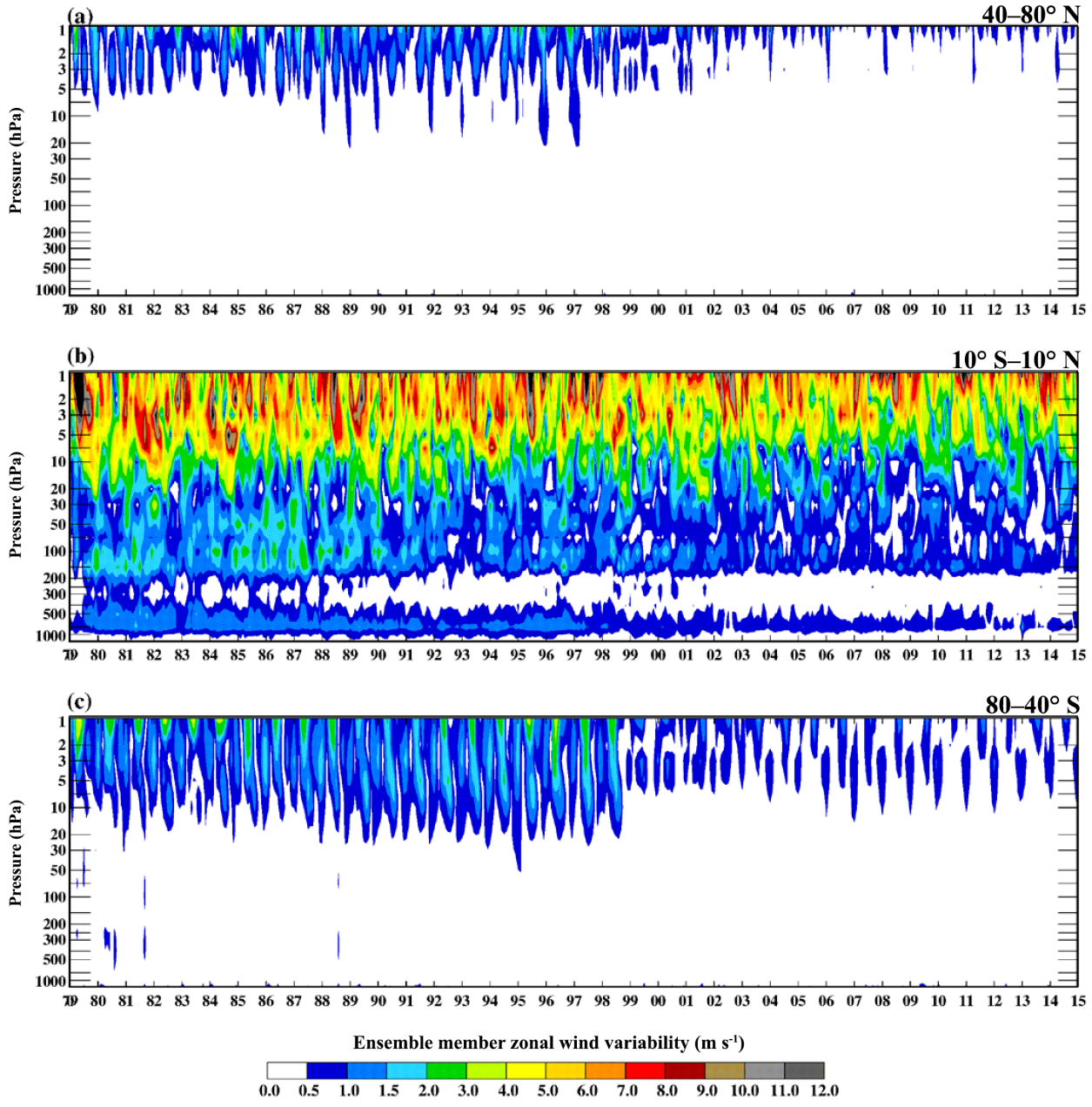


Figure 5. Pressure vs. time plots of the zonal wind SD ($m\,s^{-1}$) for each month of the three reanalyses making up the REM for three zonal regions: 40–80° N **(a)**, 10° S–10° N **(b)**, and 80–40° S **(c)**.

5 Intercomparisons of the reanalyses

In this section we extend our evaluation to the individual reanalyses and examine how each of eight reanalyses (CFSR, MERRA, ERA-I, JRA-55, MERRA-2, JRA-25, ERA-40, and R1) differs from the REM for both temperatures and winds. We do not show comparisons of R2, but one can expect all R2 qualities to be nearly the same as those of R1. We also do not show comparisons with the 20CR as that reanalysis assimilated no upper-air observations. As a result the 20CR does not show any QBO features in the tropical winds or temperatures, does not observe the occurrences of sudden stratospheric warmings making NH winters 5 K colder and polar zonal winds stronger than they should, and is 3–4 K warmer at 100 hPa in the tropics, which may be due to its coarse model vertical resolution.

5.1 Temperature

Figures 6–8 present the time-mean zonal mean temperature difference from the REM (reanalysis – REM) for each month (left columns). The right columns show the time series of the zonal-mean monthly mean differences from 1979 through 2014. The left columns show the gross monthly mean differences, while the right columns show the monthly differences over time. Both are useful to illustrate where in the vertical and when in the annual cycle the differences occur and whether these improve over time. Differences in the right column typically do not extend throughout the entire 1979–2014 period. Rather, much like the other differences discussed earlier, large improvements are seen going from the TOVS to ATOVS time periods, with the TOVS time period having the larger differences extending down further from the upper stratosphere into the middle stratosphere. Except where specifically mentioned, temperature differences between the individual reanalyses and the REM are within ±0.5 K. In general the earlier reanalyses (JRA-25, ERA-40, and R1) show greater differences from the REM than the more recent reanalyses (MERRA-2, MERRA, ERA-I, JRA-55, and CFSR). Also, the NH and SH polar latitudes generally show similar difference patterns, with much greater differences in the SH. Thus, in the following, we start with the description on the SH polar latitudes, then mention the NH polar latitudes relatively briefly, and finally describe the equatorial latitudes where the patterns are quite different from those at higher latitudes.

5.1.1 SH polar latitudes

MERRA-2 has a year-round cold bias of −1 to −2 K compared to the REM from 1 to 2 hPa, a year-round warm bias from 3 to 5 hPa, and a cold bias at 10 hPa from March through June. The time series shows that these biases are largest during the TOVS period, with much smaller differences during the ATOVS period, and that any bias is greatly reduced after August 2004 when Aura MLS temperatures at pressures less than 5 hPa are assimilated.

MERRA shows a warm bias of 1 to 2 K in the time-mean plot compared to the REM between 2 and 3 hPa from July through February. Below this, between 5 and 20 hPa, there is a cold bias of −1 to −2 K from April through August. The time series plot shows that this cold bias only exists during the TOVS period, while the warm bias at higher altitudes persists throughout the entire period.

The ERA-I has a mixture of cold (−1 K, March through August) and warm (2 K, November through February) biases compared to the REM between 1 and 3 hPa. An opposite set of biases exist slightly below, between 5 and 10 hPa, during roughly the same time periods. The time series plot shows that the upper stratosphere cold bias exists during the 1990s. The upper stratosphere warm bias occurs after 1998, while the warm bias between 10 and 5 hPa persists throughout the entire TOVS period.

The JRA-55 shows a cold bias (−2 to −4 K) compared to the REM between 1 and 5 hPa from July through March, which then descends to 7 hPa as a warm bias forms between 1 and 2 hPa from March through June. The time series plot shows that temperature differences transitioned from the TOVS to ATOVS period with the cold bias of −4 to −6 K becoming the dominant feature during this later period.

The CFSR temperatures are 6–8 K warmer than the REM in the upper stratosphere, peaking during the period of minimum temperatures in that region between March and July. Just below this warm region, there is a small altitude region with colder temperatures than the REM of −1 and −2 K. The time series plot shows that the CFSR upper stratospheric warm bias occurs throughout the entire 1979–2014 time span with similar seasonal variability.

The JRA-25 time-mean plot shows greater differences from the REM than the above five reanalyses, with a year-long warm bias (8 to 10 K) compared to the REM from 1 to 3 hPa and a very cold bias (−4 to −6 K) during the SH winter period between 5 and 10 hPa. In the middle stratosphere there are periods of persistent cool bias with a maximum (−2 to −4 K) occurring in the August–November months. The time series plot shows that the upper stratosphere warm bias (8 to 12 K) persists throughout the entire time period, with greater values (> 12 K) in the TOVS period. The cold bias (ranging between −2 and −10 K) just below the warm bias occurs mostly during the ATOVS time period. The middle stratosphere cold bias (−2 to −6 K) occurs during the TOVS period (see Sect. 5.2 of Fujiwara et al., 2017, for the reason).

The ERA-40 time-mean plot shows a strong cold bias (−2 to −6 K) compared to the REM persisting year-long between 2 and 10 hPa. Just below this is a warm bias (2 to 4 K) between 10 and 30 hPa. The annual cycle of both the cold bias and warm bias show a slight rising in summer and a lowering in winter months. In the lower stratosphere and upper troposphere, there are layers and monthly periods of slight cold (> −2 K) and slight warm (< 2 K) bias. The time series plot shows that these biases occur throughout most of the ERA-40 time period, which ends in 2002.

R1 does not analyse at pressures lower than 10 hPa, so there is no evaluation in the upper stratosphere. However, there is a nearly year-round warm bias (1 to 2 K) compared to the REM between 10 and 50 hPa peaking between June and September. Another shallow layer of warm bias (1 to 2 K) exists between 100 and 400 hPa. The time series plot shows that the middle stratospheric warm bias is most pronounced in the TOVS period.

5.1.2 NH polar latitudes

Many features in the upper stratosphere are seasonally common between the NH and SH polar latitudes (Fig. 7). However, differences with the REM in the middle and lower

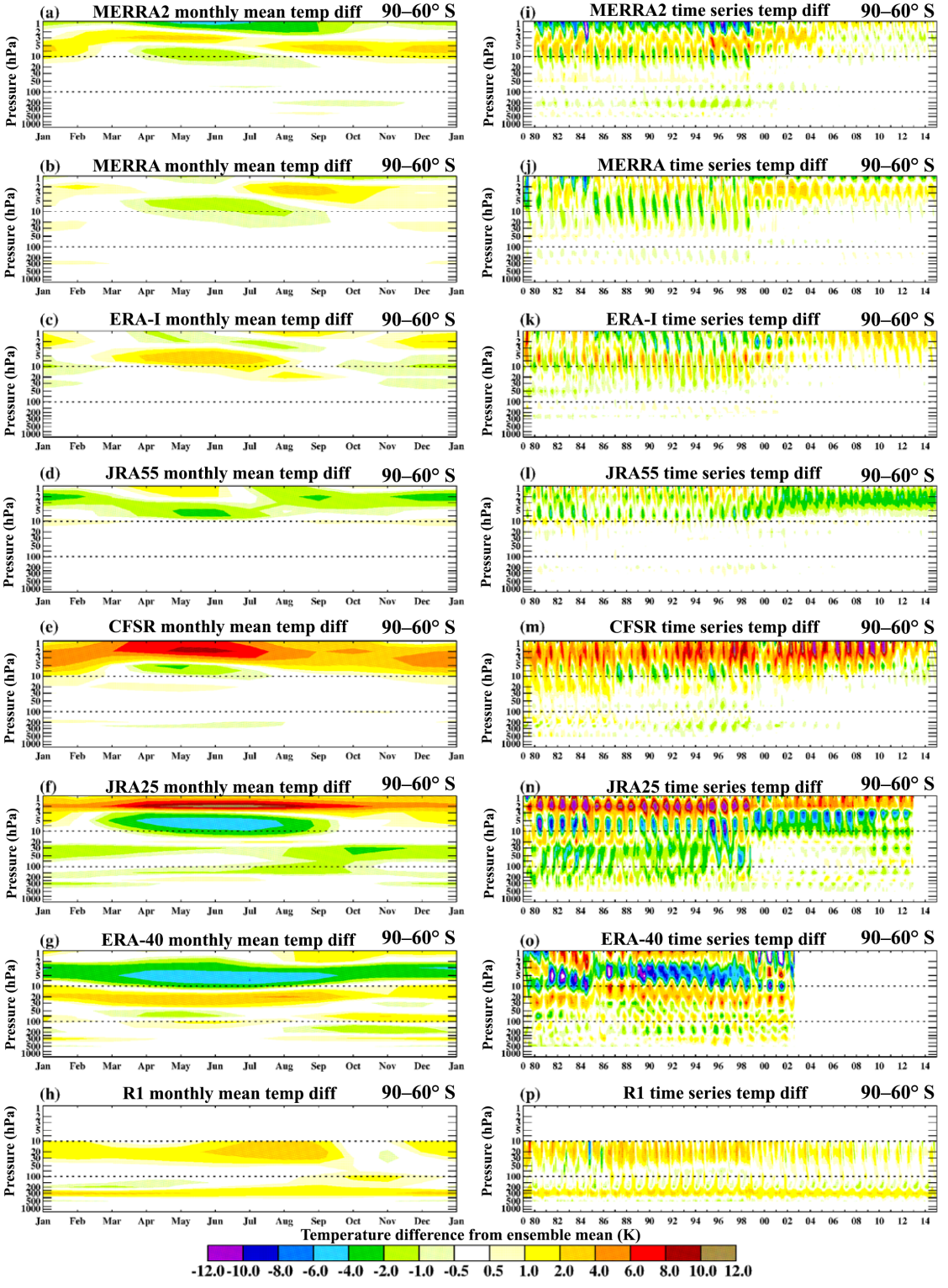


Figure 6. Pressure vs. month plots **(a–h)** and pressure vs. time plots **(i–p)** of the temperature difference (K) in individual reanalyses from the REM for the zonal region 90–60° S. The reanalyses are **(a, i)** MERRA-2, **(b, j)** MERRA, **(c, k)** ERA-I, **(d, l)** JRA-55, **(e, m)** CFSR, **(f, n)** JRA-25, **(g, o)** ERA-40, and **(h, p)** R1. The left column plots are the monthly mean differences for the entire 1979–2014 period. The right column plots are each month's difference from the REM for that same month.

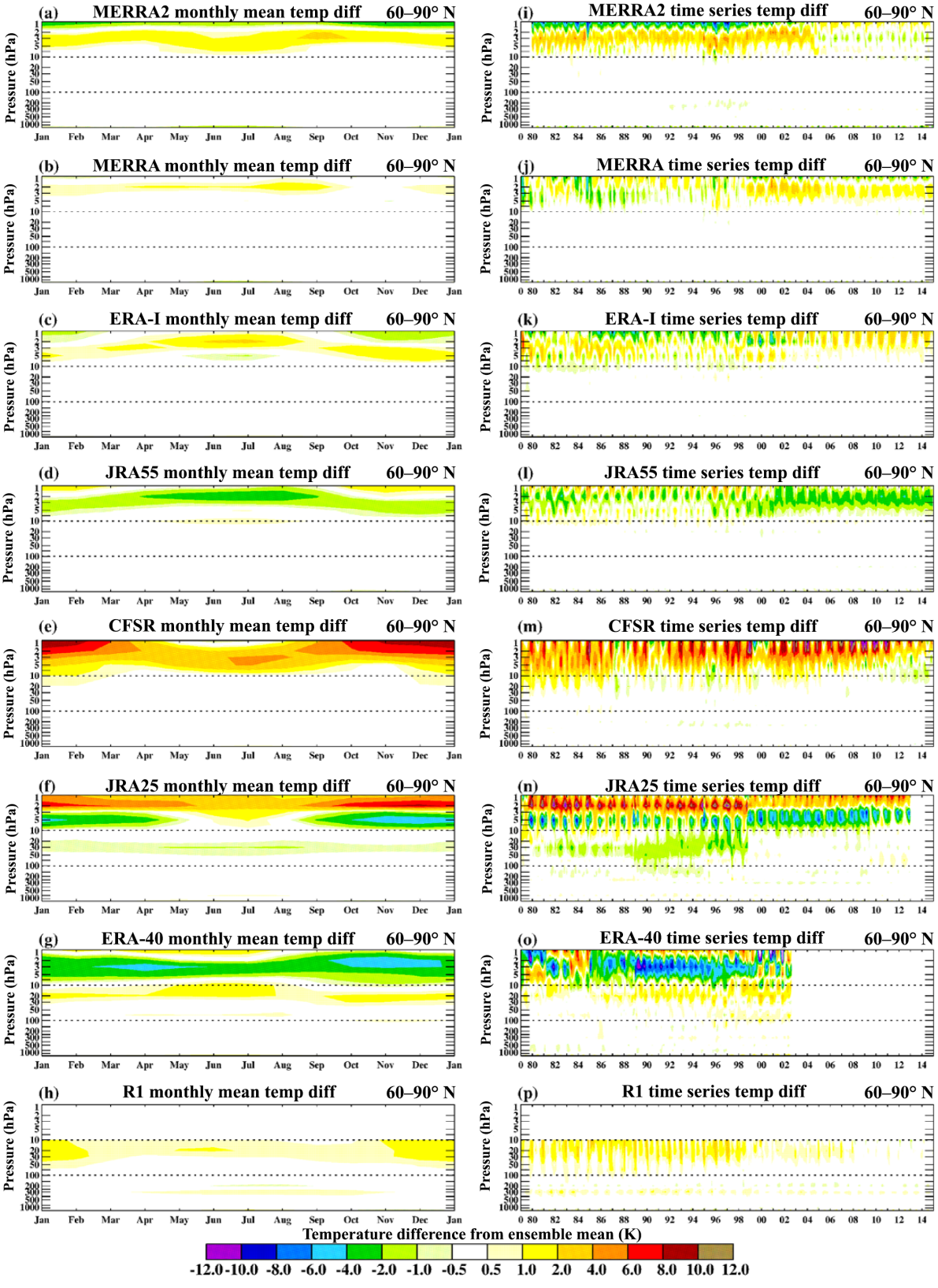


Figure 7. Same as Fig. 6 but for the 60–90° N latitude zone.

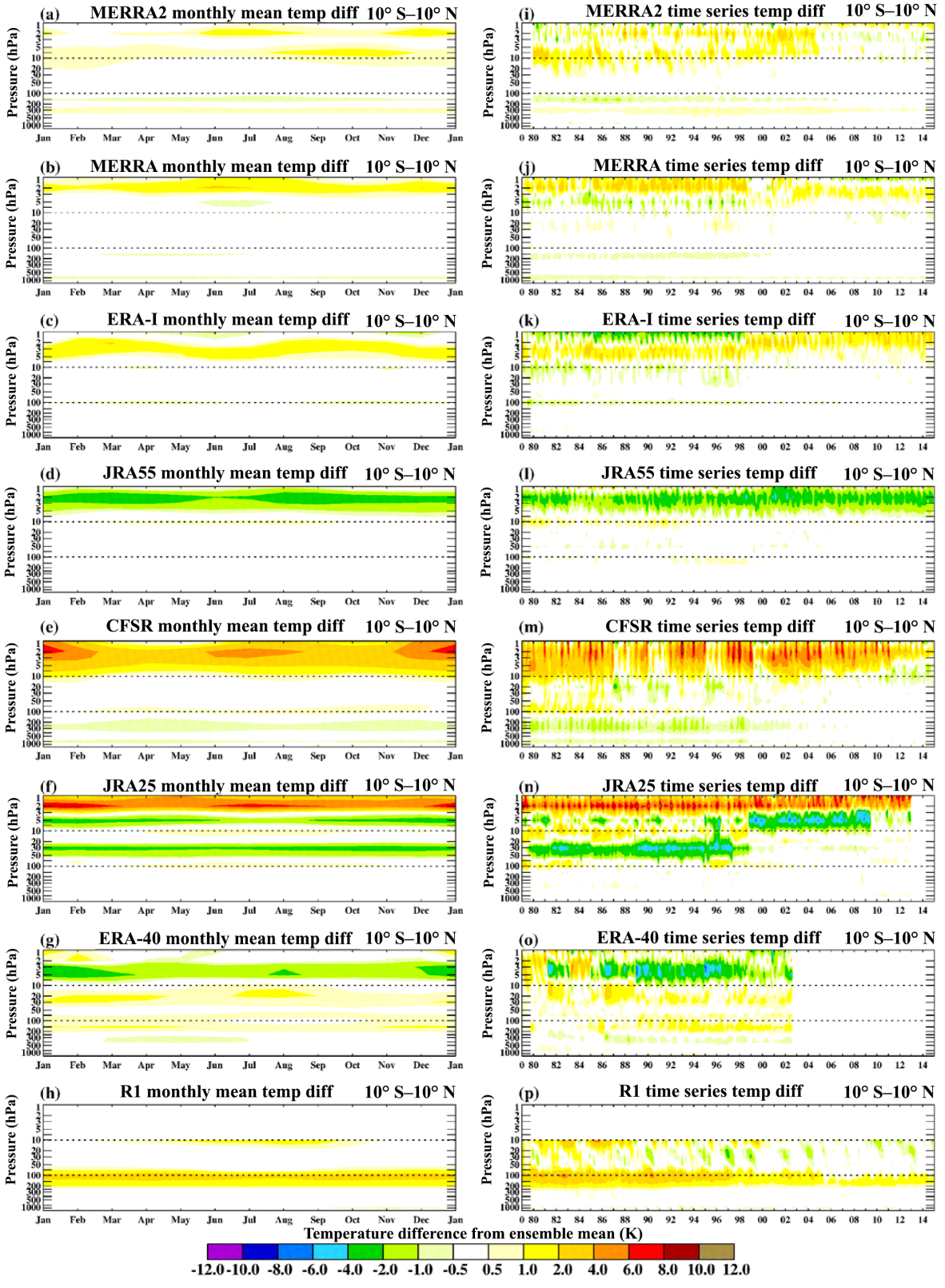


Figure 8. Same as Fig. 6 but for the 10° S–10° N latitude zone.

stratosphere in the SH are reduced or eliminated in the NH. The cold bias that occurred between 10 and 5 hPa in the MERRA-2 differences during the SH winter season is not present in the NH winter differences. MERRA differences from the REM in the NH are much smaller in the monthly means, with just a thin warm bias layer between 3 and 5 hPa. The time series shows only slight differences in the middle and lower stratosphere during the TOVS period compared to the same altitude region in the SH. The ERA-I and JRA-55 have very similar seasonal biases as those that occurred in the SH. Similar to MERRA, the time series of differences for the ERA-I during the TOVS period in the middle and lower stratosphere are nearly eliminated. The JRA-55 time series does not have noticeable differences from what was observed in the SH. The CFSR wintertime warm bias that occurs at pressures lower than 7 hPa extends from October through March. There is no evidence of a cold bias underneath this warm bias in the monthly means as occurs in the SH. The time series of differences shows that the differences that occur in the middle and lower stratosphere in the SH do not exist in the NH. The JRA-25, ERA-40, and R1 all show similar seasonal biases from the REM in the upper stratosphere. Their time series show reduced differences in the middle and lower stratosphere.

5.1.3 Equatorial latitudes

Differences in reanalysis temperatures from the REM in the equatorial regions (10° S–10° N) vary more on a semi-annual basis. Figure 8 shows that such is the case for the CFSR upper stratosphere warm bias of 2 to 4 K and for the JRA-55 upper stratosphere cold bias of −2 to −4 K. MERRA-2 shows relatively small differences (< 1 K) at all altitudes compared to the REM and the near elimination of any bias after August 2004 when MLS temperatures at pressures less than 5 hPa were assimilated. The MERRA and ERA-I exhibit a slight warm bias at pressures lower than 5 hPa. The time series plots for the CFSR show the jumps associated with the different streams and the gradually increasing warm bias in the upper stratosphere during each of these streams. A warm bias centred at 100 hPa and a cold bias below persist though the TOVS period. The MERRA and ERA-I have temperature biases that are greater during the TOVS period than the ATOVS period. In the ATOVS period the bias in both reanalyses is confined to the upper stratosphere at pressures less than 3 hPa with a warm bias of 0.5 to 2 K. The JRA-55 reanalyses show that the cold biases are nearly constant throughout the entire time series. The JRA-25 has a consistent warm bias of 4 to 6 K in the upper stratosphere at pressures less than 3 hPa. Immediately below this at 5 hPa is a cold bias of −2 to −8 K that is largest during the ATOVS period. Between 30 and 50 hPa, there is another layer of cold bias of −2 to −6 K that is present only during the TOVS period. ERA-40 has a persistent cold bias of −2 to −6 K in the upper stratosphere between 2 and 7 hPa and two layers of warm bias of 0.5 to 1 K in the middle stratosphere and tropopause regions. R1 in the middle stratosphere has slight warm and cold biases associated with the QBO (seen in the time series plot). There is also a persistent warm bias of 2 to 4 K in the upper troposphere to tropopause layer between 70 and 200 hPa. This warm bias persists from the TOVS period to the ATOVS period when its magnitude decreases to a warm bias of 1 to 2 K. Randel et al. (2004) pointed this out in their comparison of analyses and attributed the inability to capture lower tropopause temperatures to the coarse vertical resolution and the assimilation of retrieved temperatures (as opposed to radiances).

As discussed in Sect. 4.3 the three members of the ensemble mean have their greatest disagreement in the upper stratosphere. From the above differences compared to the REM temperatures, the upper stratospheric warm bias of MERRA and ERA-I at all latitudes is nearly counterbalanced by the cold bias of the JRA-55. The ERA-I warm bias between 5 and 7 hPa in the SH polar latitudes is counterbalanced somewhat equally by the MERRA and JRA-55 reanalyses.

5.1.4 NH and SH mid-latitudes

The NH and SH mid-latitude zone (30–60° N and 60–30° S, respectively) monthly mean temperature differences and time series temperature differences are nearly exactly the same in character, altitude, and value as the respective polar region differences.

5.2 Zonal wind

5.2.1 SH polar latitudes

The time-mean SH polar jet differences (see the Supplement) of the individual reanalyses from the REM are relatively small, ranging from −2 to 1 $m s^{-1}$, with most differences smaller in magnitude than that. As presented in Sect. 4.3.2, the REM members agree quite well in the polar jet region in both hemispheres. Some notable features are as follows. For all reanalyses except R1, the upper stratosphere is the region where the greatest differences from the REM are seen, but shows much improvement from the TOVS to ATOVS periods. MERRA-2 shows further improvements after 2004 when the MLS temperatures started to be assimilated at pressures less than 5 hPa. JRA-25 and ERA-40 show greater differences compared to more recent reanalyses. Finally, R1 shows an easterly bias to the westerlies during the transition months from westerlies to easterlies in the middle and lower stratosphere for most of the entire time series.

5.2.2 NH polar latitudes

Just as with the NH temperature differences in Sect. 4.1.2, the NH polar jet wind differences from the REM (see the Supplement) are smaller in magnitude than the SH differences and are restricted mainly to the upper stratosphere.

5.2.3 Equatorial latitudes

In Fig. 9, differences in the stratosphere at pressures less than 7 hPa show how the reanalyses differ from each other in the strength of the westerly and the easterly phases in the SAO region. CFSR and JRA-55 have weaker westerlies and thus have negative biases of greater than $-5\,\mathrm{m\,s^{-1}}$ during the March–April and September–November westerly periods. They also have positive biases greater than $3\,\mathrm{m\,s^{-1}}$ during the December–February easterly period. MERRA and ERA-I have stronger westerlies and show positive biases of greater than $3\,\mathrm{m\,s^{-1}}$ during the March–April and September–November westerly periods. They also have stronger easterlies during the December–February period but differ slightly during the July–August easterly period. This results in the MERRA and ERA-I having negative biases of less than $-3\,\mathrm{m\,s^{-1}}$ during the former period. The SAO westerlies in MERRA-2 are more than $10\,\mathrm{m\,s^{-1}}$ stronger than those in the REM. The time series shows that the stronger westerlies occur primarily during the TOVS period. Kawatani et al. (2016) and Molod et al. (2015) note that the downward-propagating westerly phase of the SAO is enhanced during the 1980s and could be caused by strong gravity wave forcing.

MERRA-2 also transitions from QBO westerlies to easterlies more rapidly than the REM during the TOVS period. The time series plots also show where each reanalysis has a slight easterly or westerly bias associated with the phase of the descending QBO winds. The JRA-25 and R1 show greater differences from the REM than the other reanalyses. R1 shows a westerly bias of $> 4\,\mathrm{m\,s^{-1}}$ during the easterly phase of the QBO from 10 hPa down to 100 hPa. This was also discussed by Pawson and Fiorino (1998b). The JRA-25 has an easterly bias of $> 4\,\mathrm{m\,s^{-1}}$ during the easterly phase of the QBO from 10 hPa down to 30 hPa. It should be noted that the CFSR used ERA-40 zonal winds as substitute observations between 30° S and 30° N and from 1 to 30 hPa from 1 July 1981 to 31 December 1998 (Saha et al., 2010); hence their differences from the REM during that time period and in that pressure range are very similar.

Interestingly, in Fig. 9 there are also sizable differences in the troposphere. The CFSR zonal winds in the tropical upper troposphere during the TOVS years have an easterly bias. This may be associated with the CFSR having a cold bias of about 1 K in the upper troposphere during this time period. The JRA-55 zonal winds have a westerly bias during this time period. The MERRA and ERA-I zonal wind differences in the upper troposphere are no larger than $0.5\,\mathrm{m\,s^{-1}}$. Hence, the differences from the REM show that the CFSR has a consistent layer of negative biases of -1 to $-2.5\,\mathrm{m\,s^{-1}}$ from 50 to 300 hPa. The JRA-55 shows the other extreme of a consistent positive bias of 1 to $2\,\mathrm{m\,s^{-1}}$ from 30 to 200 hPa. The time series plots confirm that these upper troposphere zonal wind biases are persistent during the TOVS time period and are reduced in the ATOVS period. MERRA-2 shows large positive differences of $> 6\,\mathrm{m\,s^{-1}}$ from the REM in the upper stratosphere (SAO region). The time series show that these large differences occur mostly during the 1980s and periodically extend to 20 hPa. These large differences continue throughout the time series but are confined to the upper stratosphere after the 1990s.

5.2.4 NH and SH mid-latitudes

Characteristically, the zonal winds in the NH and SH mid-latitudes (20–40° N and 40–20° S, respectively) are different in depending upon the altitude. In the troposphere there is the tropical jet with maximum winds near 200 hPa. In the lower stratosphere there is a lull between the equatorial winds and the polar jet. The upper stratosphere is seasonally transitioning from the SAO to the winter polar jet. The differences from the REM show that all the reanalyses are in very good agreement with the tropospheric tropical jet. In the lower stratosphere R1 has a westerly bias of 0.5 to $1\,\mathrm{m\,s^{-1}}$, which is greatest in the early 1980s and diminishes to nil by the 2000s. The CFSR, interestingly, has an easterly bias of -0.5 to $-1\,\mathrm{m\,s^{-1}}$ during the TOVS period and is eliminated in the ATOVS period. All the other reanalyses are in good agreement (differences within $\pm 0.5\,\mathrm{m\,s^{-1}}$) with the REM in the lower stratosphere. In the middle stratosphere the JRA-25 has differences between -0.5 and $-1\,\mathrm{m\,s^{-1}}$ from the REM in both the NH and SH mid-latitudes. In the upper stratosphere the more recent reanalyses have differences between -1 and $1\,\mathrm{m\,s^{-1}}$ from the REM, which diminish further during the ATOVS period. The JRA-25 and ERA-40 have slightly larger differences, which also diminish appreciably in the ATOVS period.

5.2.5 Comparisons with Singapore QBO winds

Kawatani et al. (2016) provides a thorough evaluation of the RMS differences in QBO (70–10 hPa) zonal winds among the more recent reanalyses and observations from all the radiosonde sites in the equatorial-latitude zone. Kawatani et al. (2016) also show that of the nearly 220 radiosonde stations in the 20° S–20° N zone, Singapore (1° N, 104° E) is the only station that reports 10 hPa observations 80–100% of the time between 1979 and 2001. For this reason, we will focus just upon comparisons between the reanalyses and zonal winds at Singapore. This is not to imply that Singapore is representative of the entire tropical zone, which it is not because there is longitudinal variability in the zonal-mean zonal winds (Kawatani et al., 2016). Correlations among the monthly mean MERRA-2, MERRA, ERA-I, JRA-55, and CFSR QBO zonal winds (interpolated to Singapore) and the monthly mean radiosonde wind observations at Singapore (obtained from the Free University of Berlin) are mostly above 0.9. More information about how the reanalyses differ from the Singapore winds can be obtained by evaluating the linear regression line between the observed and analysed QBO winds and their scatter. Figure 10a–c shows the

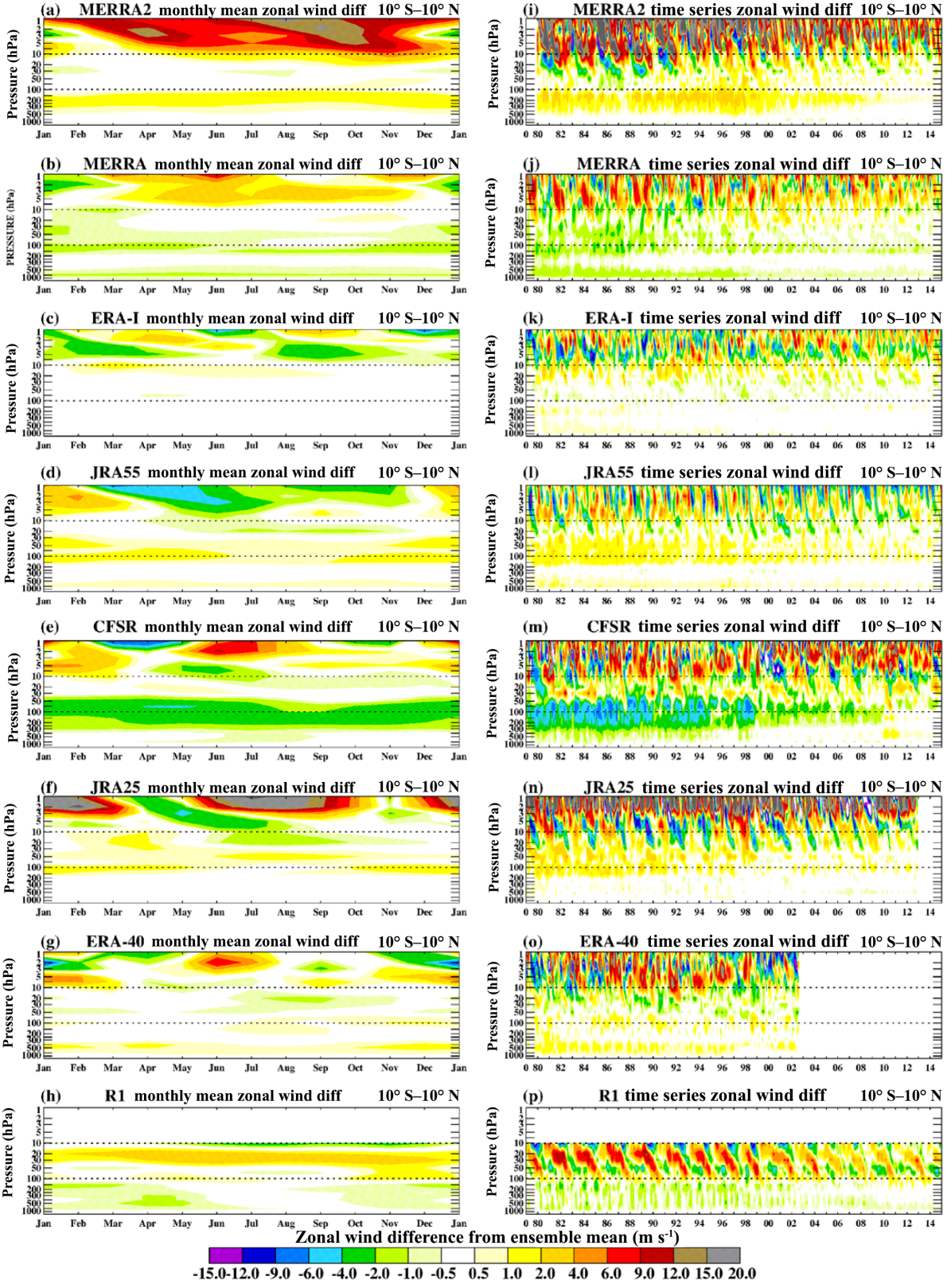


Figure 9. Pressure vs. month plots **(a–h)** and pressure vs. time plots **(i–p)** of the zonal wind difference ($m s^{-1}$) in individual reanalyses from the REM for the zonal region 10° S–10° N. The reanalyses are **(a, i)** MERRA-2, **(b, j)** MERRA, **(c, k)** ERA-I, **(d, l)** JRA-55, **(e, m)** CFSR, **(f, n)** JRA-25, **(g, o)** ERA-40, and **(h, p)** R1.

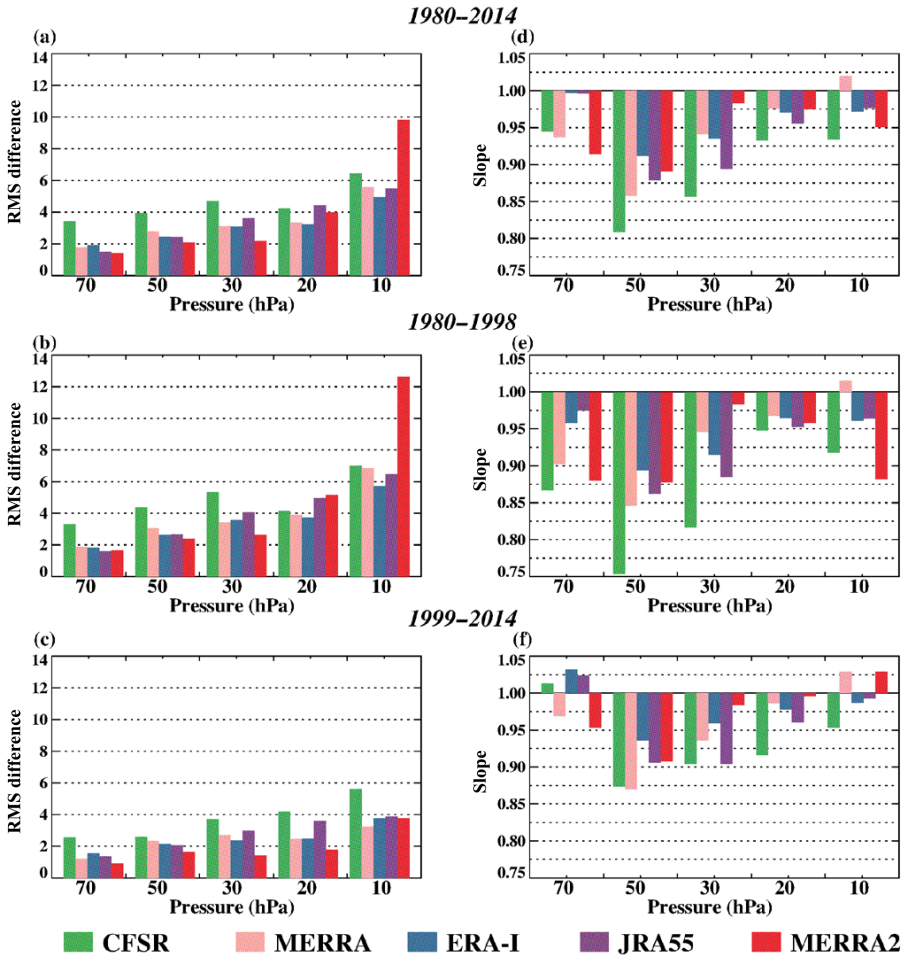


Figure 10. RMS differences ($m\,s^{-1}$) **(a–c)** and linear slopes **(d–f)** of the matched QBO zonal wind anomalies at 70, 50, 30, 20, and 10 hPa for the CFSR, MERRA, ERA-I, JRA-55, and MERRA-2 reanalyses interpolated to Singapore (1° N, 104° E) vs. the observed Singapore monthly mean zonal winds from the FUB. RMS differences and slopes are computed for the 1980–2014 time period **(a, d)**, the 1980–1998 period **(b, e)**, and the 1999–2014 period **(c, f)**. Slopes less than 1.0 indicate that the reanalysis zonal winds are weaker than the Singapore zonal winds.

RMS differences in the reanalyses QBO winds and those at Singapore. Comparisons are shown for the entire 1980–2014 period and then divided into the TOVS (1980–1998) and ATOVS (1999–2014) periods. All of the reanalyses RMS differences are smaller during the ATOVS period. All of the RMS differences increase from 70 to 10 hPa as does the amplitude of the winds at these levels. The RMS differences decrease by one-half to one-third from the TOVS to the ATOVS period. Of these five reanalyses, the CFSR performs the poorest with higher RMS differences at nearly all pressure levels during all periods. MERRA-2 has the largest RMS differences at 10 hPa during the TOVS period, but improves during the ATOVS period. As seen in Fig. 9, MERRA-2 has large irregularities in the 1980s and in 1993. As mentioned earlier, these irregularities are a result of overly strong SAO westerlies that propagate down to the middle stratosphere. Coy et al. (2016) explain that during the 1980s and early 1990s MERRA-2 overemphasized the annual signal. Figure 10d–f show the slope of the regression line between the individual reanalysis QBO winds and the Singapore QBO winds. The maximum underestimation (slope smaller than 1) at 50 hPa is present in all of the reanalyses. The reanalysis winds and Singapore winds become more similar in strength at lower pressure levels and are closer in strength during the ATOVS period than the TOVS period. The CFSR has consistently weaker winds at all pressure levels during both the TOVS and ATOVS periods. No one reanalysis is better than the others at all QBO levels in either the TOVS or ATOVS period.

6 Amplitude of polar annual temperature cycle

Another way to examine the differences among the reanalyses is to compare their annual temperature amplitude (warmest summer month minus coldest winter month) in the

polar latitudes. If a reanalysis has a wintertime warm bias or a summertime cold bias, then its annual temperature amplitude will be smaller compared to the other reanalyses. Generally, as Fig. 2 shows, the summertime temperatures do not vary much from year to year, while the wintertime temperatures have greater interannual variability. The mean polar temperatures in Fig. 2 indicate which months would likely be used as the warmest and coldest at the various pressure levels. For these differences we use the coldest (warmest) month from November through March and the warmest (coldest) month from May through September for the Northern Hemisphere (Southern Hemisphere). The lower variability in the SH temperatures ensures that the same months are used for the 1979 to 2014 period. However, in the NH the coldest month at a particular pressure level depends upon whether an SSW occurs. In the upper stratosphere, after an SSW the low temperatures following the warming are usually the lowest of the year. Without a warming the lowest temperatures may well have occurred in November or December. In the middle stratosphere the lowest temperatures will usually occur in December. In the lower stratosphere the lowest temperatures will usually occur in December or January. In Fig. 11 a time series of the SH and NH polar zone annual temperature amplitudes is presented. In general the SH annual amplitudes in the middle and upper stratosphere are up to 25 K larger than at the same level in the NH, largely because of the persistent and colder SH winters. At pressures greater than 300 hPa, temperature amplitudes in the SH are smaller than those in the NH. SH temperature amplitudes increase from 5–15 K in the troposphere to 45–60 K in the middle stratosphere. Maximum amplitudes (60–70 K) in the SH occur above 10 hPa. In the NH polar latitudes, the minimal amplitude of 5–15 K occurs at the polar tropopause. Between the surface and the tropopause, the temperature amplitude is larger at 15–25 K. Above the tropopause the temperature amplitude increases up to about 2–3 hPa where the temperature amplitude lies in the 55–60 K range, although the depth of this layer is not nearly as extensive as in the SH polar regions. There is good agreement among these five more recent reanalyses on the years of peak amplitude in the NH polar region upper stratosphere. The peak SH amplitudes of the five reanalyses are in lesser agreement in year and pressure range.

Individually, the five more recent reanalyses agree well with each other from the surface through the lower stratosphere in both hemispheres. However, the ERA-I shows an annual temperature amplitude in the middle stratosphere that is 5–15 K smaller than the other four reanalyses in the SH and about 5 K smaller in the NH polar regions from 1979–2002. The JRA-55 has smaller maximum amplitudes in the SH than the other four reanalyses, which is associated with its seasonally low temperature bias in the upper stratosphere, whereas the CFSR tends to have consistently large maximum amplitudes which are associated with its seasonally warm bias. However, the CFSR temperature amplitudes peak at greater pressures in the upper stratosphere and then decrease rapidly between 3 and 1 hPa in both hemispheres, particularly in the ATOVS period. This is most likely due to the fact that the CFSR did not bias correct the SSU channel 3 observations and did not assimilate the top AMSU-A channel 14.

As a group the NH plots show that the greatest amplitudes occur at 2 hPa. The years with this large amplitude are years in which an SSW occurred. This is a result of the very cold air that immediately follows the warming in the upper stratosphere. The years in which an SSW did not occur (e.g. the 1990s) have smaller temperature amplitudes in the upper stratosphere. The SH years in which there was a great amount of wave activity during the winter months had warmer winters and consequently smaller annual amplitudes. This is particularly noticeably in 2002 and 2010. These two years exhibited a very early transition from winter circulation to summer circulation, similar to a final warming in the NH. Final warmings are not followed by very cold air in the upper stratosphere. The ERA-I stands out as having smaller annual amplitudes in the SH middle stratosphere compared to the other four reanalyses during the TOVS period.

7 Comparisons with satellite temperature observations

7.1 HIRDLS and MLS temperatures

The NASA Earth Observing System (EOS) Aura spacecraft was launched in July 2004 and has several onboard instruments that measure multiple atmospheric constituents. The High Resolution Dynamics Limb Sounder (HIRDLS; Gille et al., 2008) instrument on the Aura spacecraft made measurements from the upper troposphere through the mesosphere until it prematurely ceased functioning in mid-2008. Quality temperature measurements extend from January 2005 through March 2008. The HIRDLS measurements were not assimilated by any of the reanalyses and thus are independent measurements. Monthly mean temperature differences in reanalyses from the HIRDLS (reanalysis – HIRDLS) temperatures at NH high latitudes (60–80° N), the tropics (10° S–10° N), and SH high latitudes (60° S) were generated for the 2005 through 2008 period. Figures 12–14 present the differences in MERRA-2, MERRA, ERA-I, JRA-55, and CFSR from the HIRDLS monthly means for these latitude zones, respectively. The time, location, and amplitude of the SH differences are generally similar to those of the reanalyses from the REM (Fig. 6). MERRA-2 has a warm bias all year long at 1 hPa and a −1 to −2 K cold bias from November through March. MERRA has a cold bias of 2–4 K from August through April from 1–3 hPa and a 2 K warm bias from May through July. ERA-I has a −2 K cold bias at 2 hPa from February through May. JRA-55 has a −4 to −6 K cold bias from July through April between 2 and 3 hPa that becomes thinner in altitude from April to July as a warm bias occurs from 1 to 2 hPa. The CFSR has a very warm bias of over 14 K in the April to July period at pressures lower than

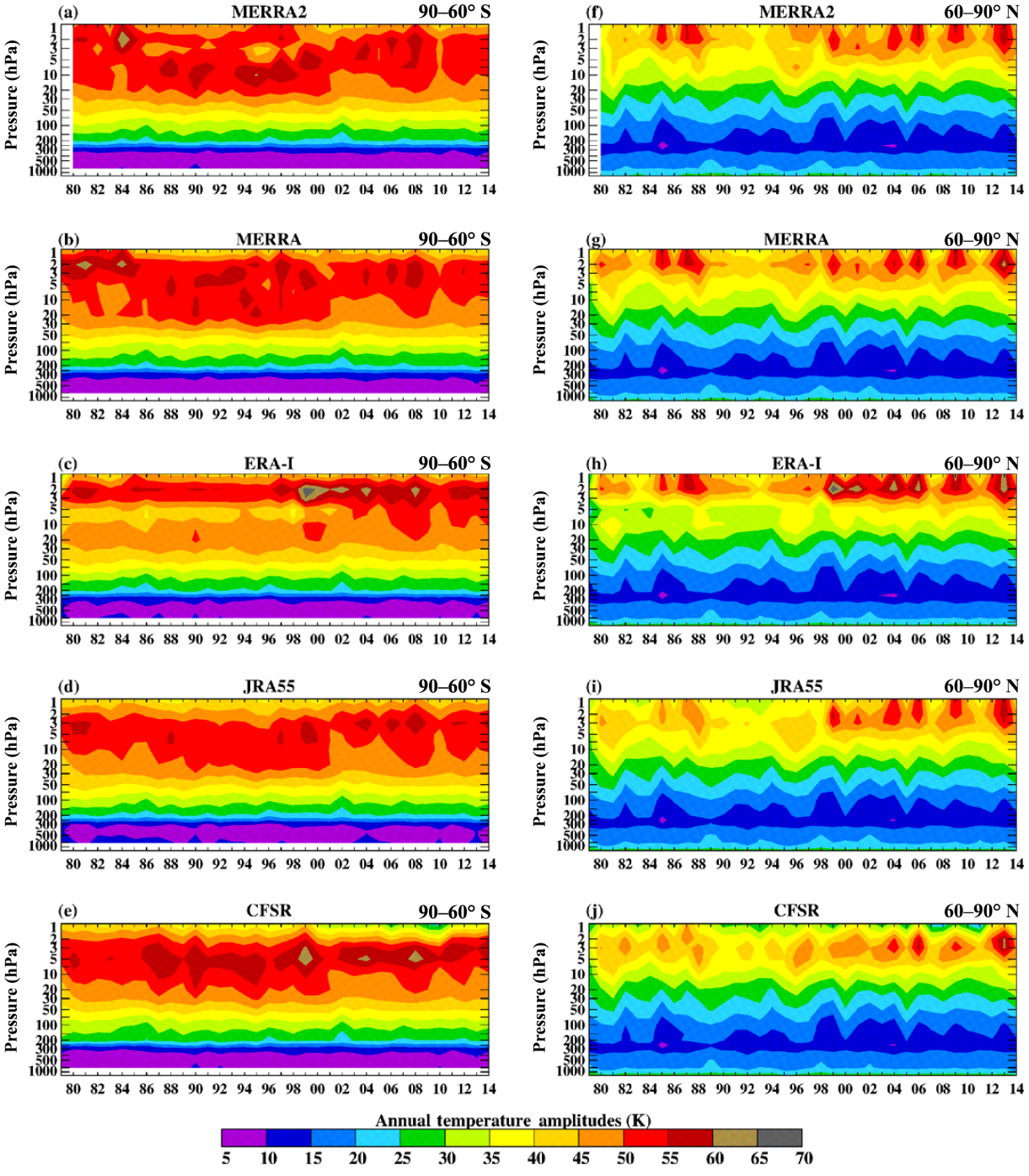


Figure 11. Yearly annual temperature amplitude (K) for 90–60° S **(a–e)** and 60–90° N **(f–j)** from the **(a, f)** MERRA-2, **(b, g)** MERRA, **(c, h)** ERA-I, **(d, i)** JRA55, and **(e, j)** CFSR reanalyses. Note that the SH annual amplitude is much larger than the NH amplitude. No analysis is performed between 1000 and 700 hPa for the SH plots as this is below the Antarctic surface.

5 hPa with a cold bias at 7 hPa during this same time period. All of the reanalyses show a slight (< 1 K) warm bias in the middle stratosphere during the November through March period.

In the NH, the cold bias of MERRA-2 in the summer period is smaller in the NH, while the year-long warm bias exists at 1 hPa. The cold bias that MERRA has in the SH does not exist in the NH. The midwinter warm bias that was in the SH is about 1 °C warmer in the NH. Similarly, the ERA-I

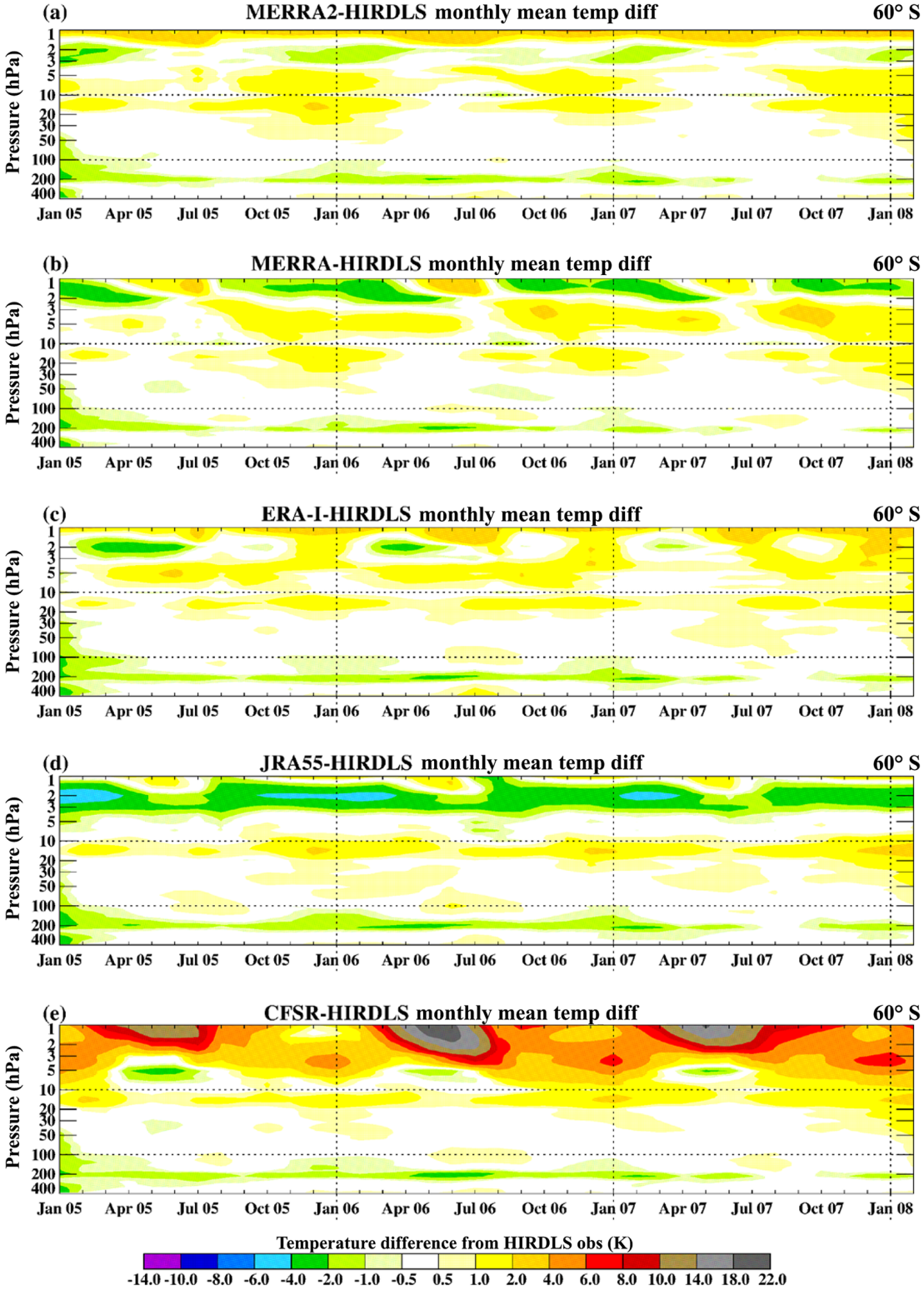


Figure 12. Pressure vs. time plots of differences in reanalyses minus HIRDLS temperatures (K) from January 2005 through January 2008 for the Southern Hemisphere high-latitude zone (60° S). The reanalyses are **(a)** MERRA-2, **(b)** MERRA, **(c)** ERA-I, **(d)** JRA-55, and **(e)** CFSR.

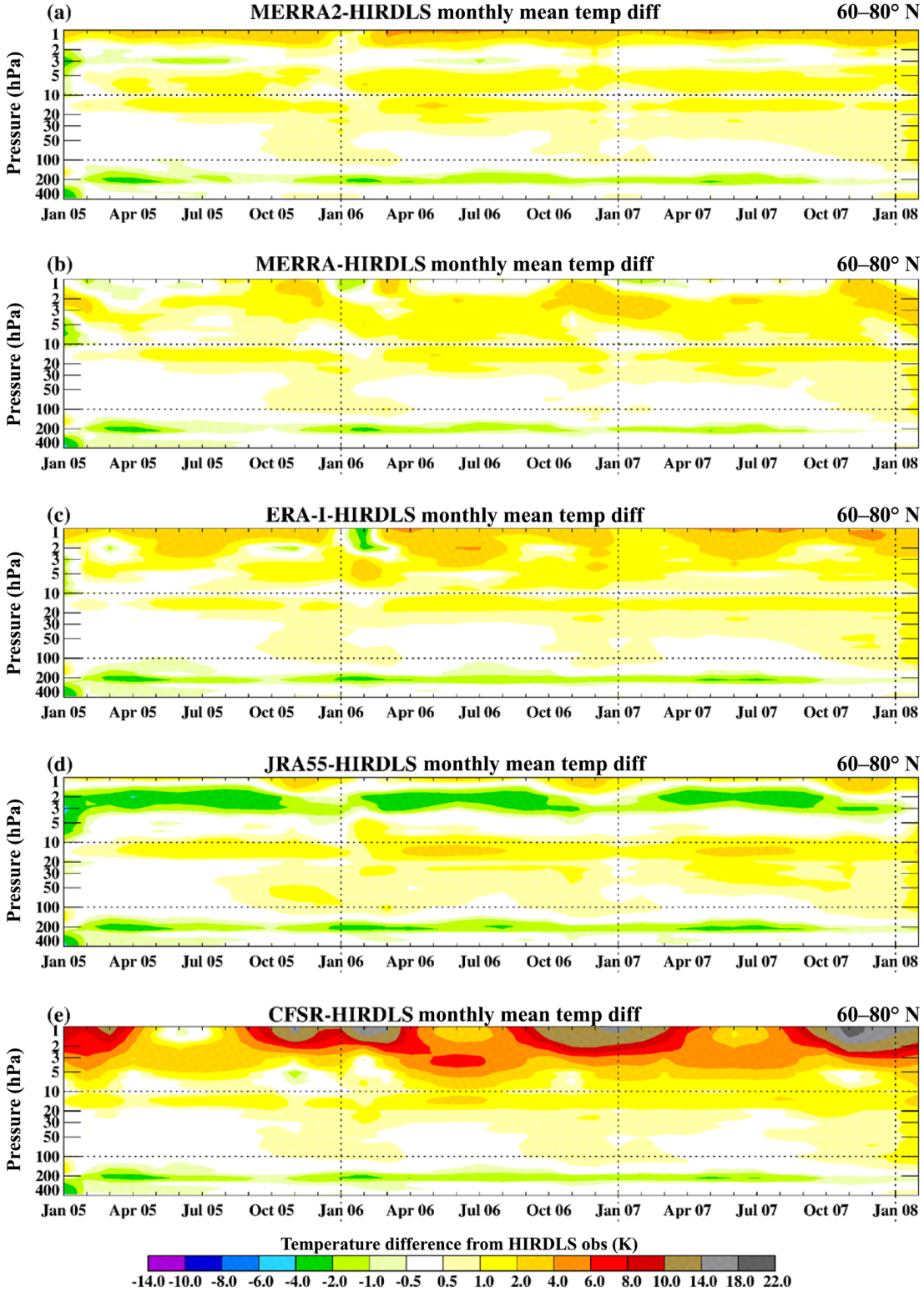


Figure 13. Same as Fig. 12 except for the Northern Hemisphere high-latitude zone (60–80° N).

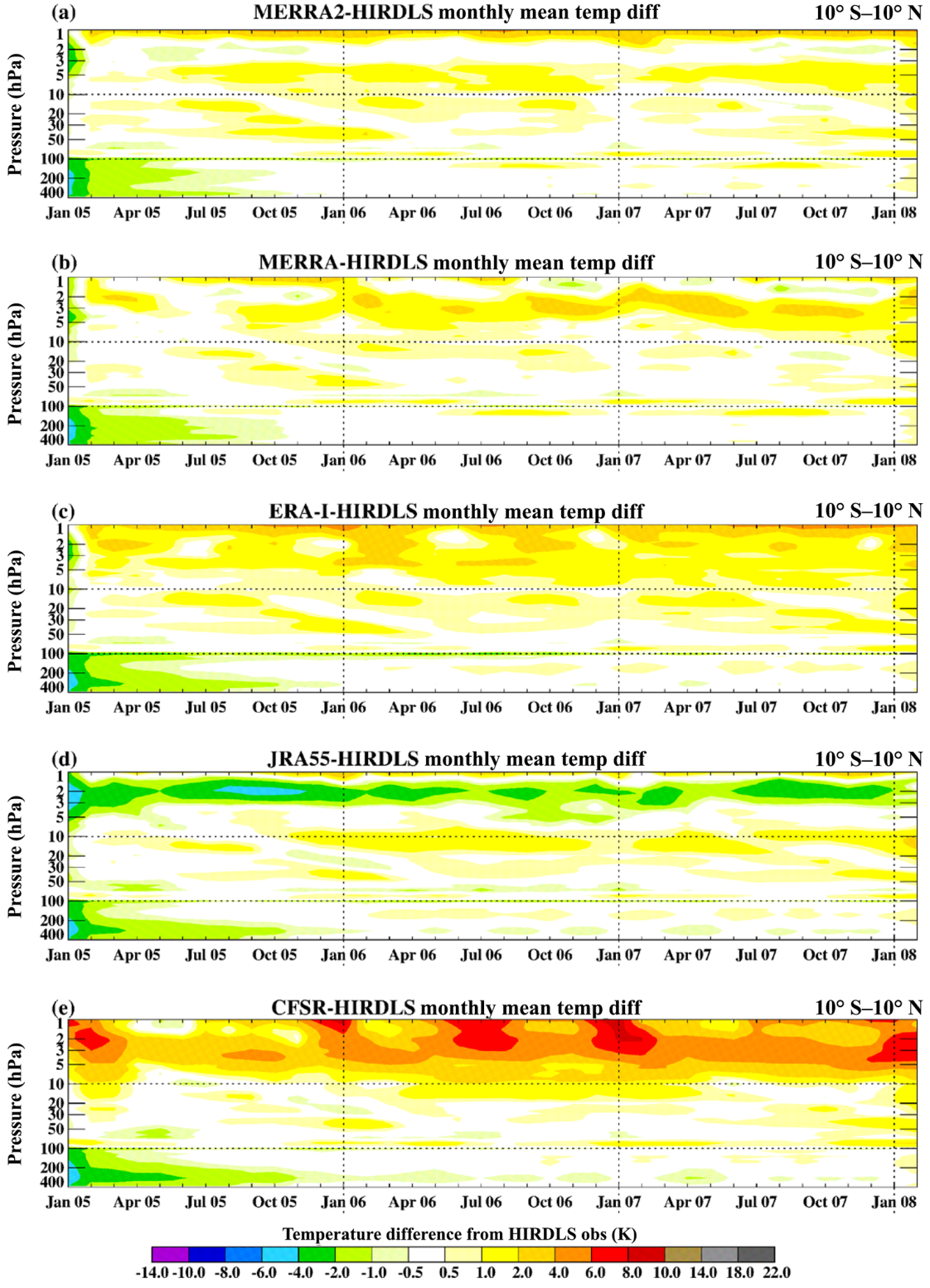


Figure 14. Same as Fig. 12 except for the equatorial-latitude zone (10° S–10° N).

does not have a cold bias in the late winter–spring period, but there is a warm bias in midsummer in the upper stratosphere. The CFSR and JRA-55 differences with HIRDLS occur in the same seasons as in the SH with little change in amplitude. Of interest is that all the reanalyses show a similar warm bias as in the SH during the November through March period.

In the tropics, MERRA-2 continues to have a year-long warm bias at 1 hPa and a slight warm bias near 10 hPa. In 2006–2007 MERRA has a warm bias between 2 and 3 hPa during January and February and moves lower to 5 to 10 hPa during the other months of the year. ERA-I seems to have a year-long 0.5 to 1 K warm bias at pressures lower than 10 hPa. JRA-55 has a year-long −1 to −2 K cold bias between 5 and 2 hPa. The CFSR has a similar warm bias on a semi-annual basis in the upper stratosphere.

The Microwave Limb Sounder (MLS) is also on the EOS Aura spacecraft. Monthly zonal means of temperatures from the version 4 retrievals were provided by the MLS team for comparisons with reanalyses for the 2005–2014 period. The characteristics of the MLS temperatures are described by Schwartz et al. (2008) and Livesey et al. (2015). Note again that among the reanalyses, MERRA-2 is the only one that assimilated MLS temperatures but only at pressures less than 5 hPa. HIRDLS temperatures have been noted to be colder than the Aura MLS temperatures (Gille et al., 2008) in the upper stratosphere. Evidently, differences in MERRA-2, MERRA, ERA-I, JRA-55, and CFSR temperatures from the MLS temperatures (not shown) are very similar to those with the HIRDLS but less positive. Differences greater than ±2 K only occur above 10 hPa. Bands of differences of the order of 1 K are present below 10 hPa; however, the MLS documentation notes that there are known oscillations of this magnitude in comparison with other satellite temperature sensors, so these latter differences are not considered significant. Overall differences from the MLS observations are in agreement with the characteristics already described for each of these reanalyses.

7.2 Comparisons with COSMIC temperature observations

COSMIC GPSRO monthly zonal mean dry temperatures from January 2007 through December 2014 (level 3, version 1.3) were obtained from the JPL GENESIS data portal. Leroy et al. (2012) explain the technique through which the RO observations were turned into temperatures and transposed from altitude to pressure surfaces. We use these data to compare against the MERRA-2, MERRA, ERA-I, JRA-55, and CFSR monthly zonal mean temperature for the same period. The COSMIC data set provides temperature from 400 to 10 hPa. We will not perform comparisons with data at pressures higher than 200 hPa as atmospheric water vapour causes deviations in the actual temperatures from the dry temperatures. Figure 15 shows the 8-year time series of differences (reanalysis – COSMIC) between the reanalysis temperatures and the COSMIC temperature in the SH polar latitudes (90–60° S). Most obvious is a recurring −1 K difference between the reanalyses and COSMIC from January through July from 10 hPa down to 100 hPa. This is during the transition from SH summer to winter. During the transition from SH winter to summer, there is a 0.5 to 1 K difference also extending from 10 to 100 hPa. The source of these two biases could be in how the COSMIC zonal mean temperatures are generated as there is a 3–5-day time averaging in which temporal transitions may be smoothed out. All of the reanalyses differed (except MERRA) in assimilating either the GPSRO bending angle or refractivity (Curcurull et al., 2007; Poli et al., 2010).

Figure 16 shows the reanalysis minus COSMIC differences for the NH polar region (60–90° N). Similar negative differences occur during the transition from NH summer to winter. The depth and time length of the −1 K differences are smaller than the SH differences. There are also short-term negative differences that extend from 10 to 100 hPa during the years in which an SSW occurred (2009, 2010, and 2013). In 2009 this is preceded by a short-term (1-month) positive difference also extending from 10 to 100 hPa. The positive differences occur during the months when the SSW produced very warm temperatures in the NH polar region. The negative spikes occurred in the month(s) following the warming when very cold temperatures followed the warming in the upper and middle stratosphere. These differences imply that the dry temperature data set does not capture the maximum warming during the SSW or the cooling which follows. This may be due to the fewer COSMIC observations in the polar region vs. the number of observations peaking between 50 and 60° in both hemispheres.

Differences between the reanalyses and COSMIC dry temperatures in the tropics (10° S–10° N) (Fig. 17) show much smaller negative differences. MERRA-2, JRA-55, and especially ERA-I show very few occurrences of differences larger than −0.5 K. The few differences with the JRA-55 have a seasonal occurrence from December through February. MERRA, which did not assimilate the GPSRO data, has negative differences fairly consistent between 10 and 30 hPa. CFSR, which did assimilate GPSRO observations, has more occurrences of negative differences than MERRA-2, JRA-55, and ERA-I.

The NH and SH mid-latitudes (not shown) have seasonal differences similar to their respective polar regions but to a smaller time extent and shallower from 10 hPa down into the middle atmosphere. We conclude that between 60° S and 60° N, the lower stratosphere temperatures in the more recent reanalyses and the COSMIC dry temperatures are within ±0.5 K of each other consistently throughout the year.

7.3 Atmospheric layer temperature anomalies

Long-term satellite observations from NOAA polar orbiting satellites of temperatures in the lower stratosphere

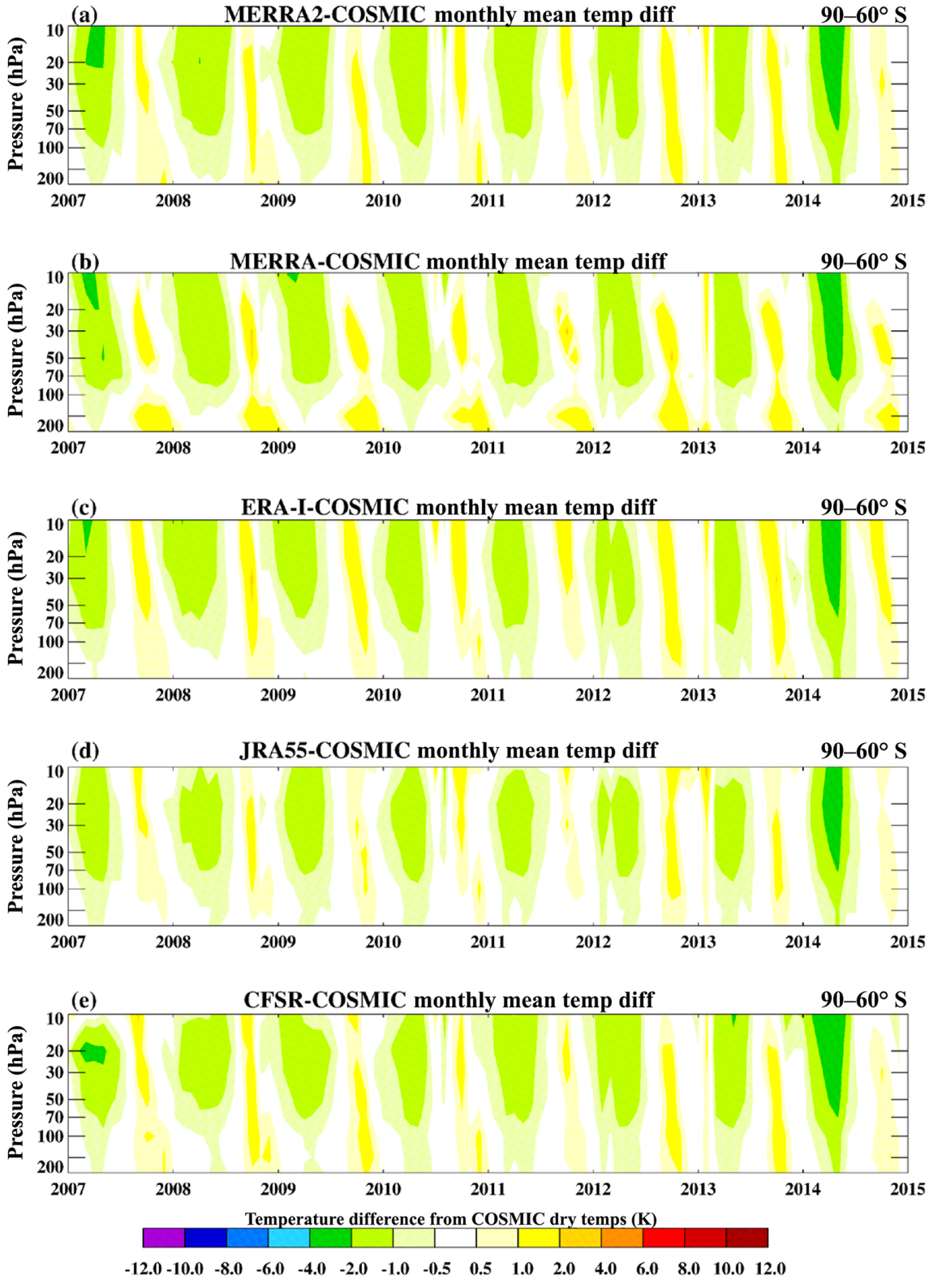


Figure 15. Pressure vs. time plots of differences in reanalyses minus COSMIC dry temperatures (K) from January 2007 through December 2014 for the Southern Hemisphere high-latitude zone (90–60° S). The reanalyses are **(a)** MERRA-2, **(b)** MERRA, **(c)** ERA-I, **(d)** JRA-55, and **(e)** CFSR.

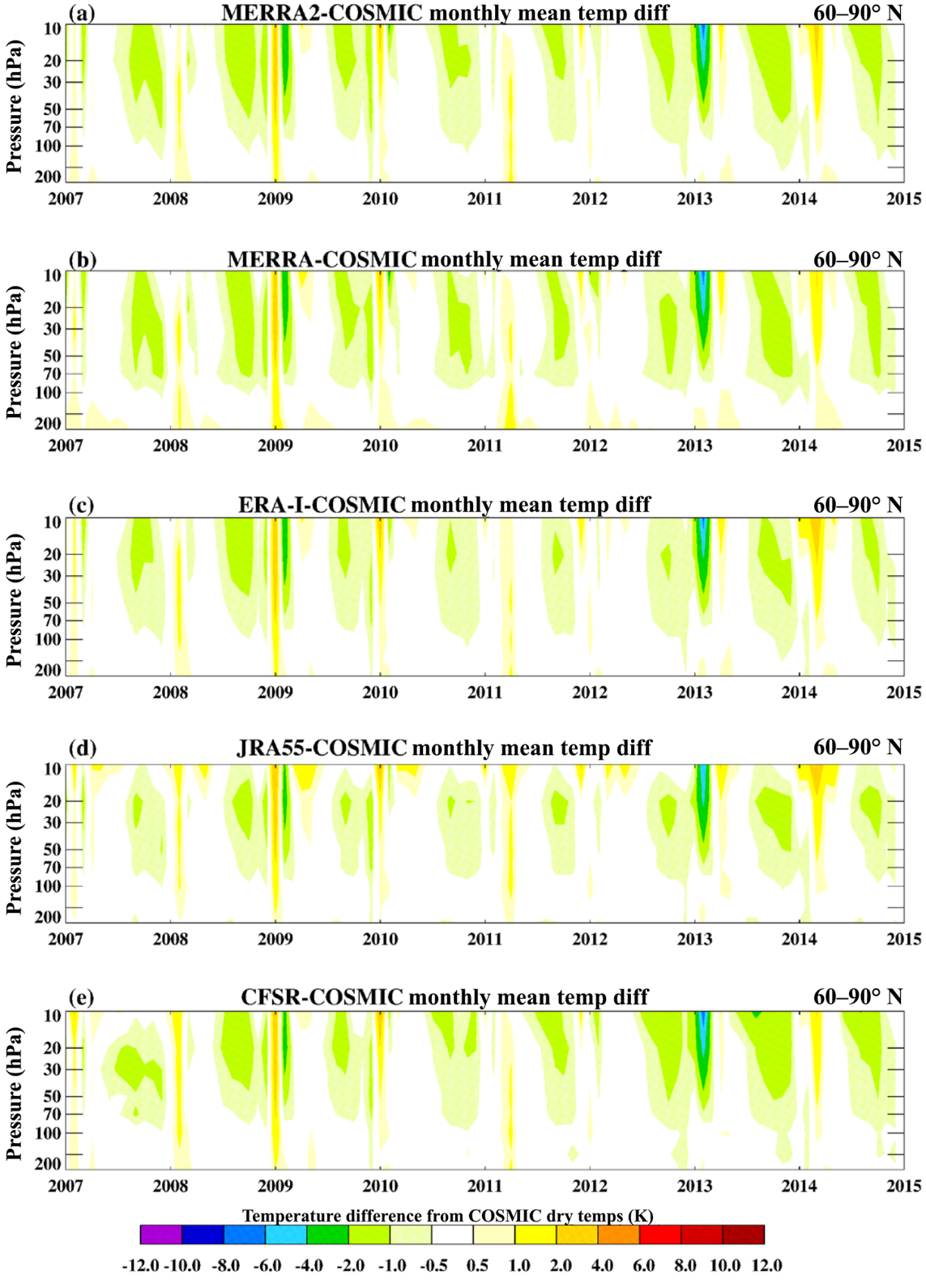


Figure 16. Same as Fig. 15 except for the Northern Hemisphere high-latitude zone (60–90° N).

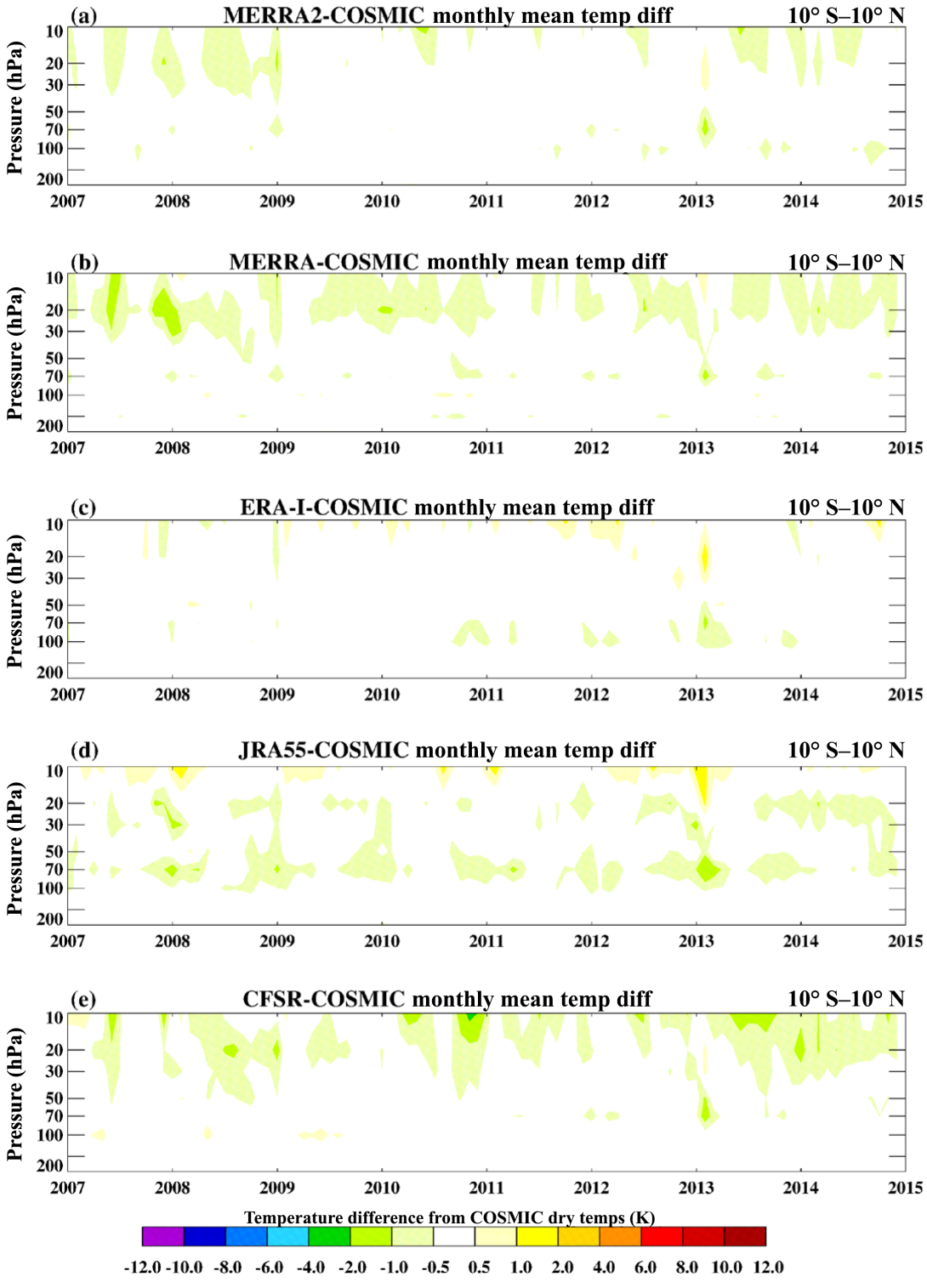


Figure 17. Same as Fig. 15 except for the equatorial-latitude zone (10° S–10° N).

(TLS) are available from the MSU-4 and AMSU-A9 microwave channels, while the Stratospheric Sounding Unit channel 1 (SSU1) and channel 2 (SSU2) provide temperature observations of the middle and upper stratosphere, respectively. Zou and Qian (2016) explain the process of merging and extending the infrared-based SSU observations

with the microwave-based AMSU-A and ATMS observations. The satellite weighting functions for these three channels can be found in Fujiwara et al. (2017, their Fig. 7) and Seidel et al. (2016, their Fig. 1) and on the NOAA STAR SSU website (http://www.star.nesdis.noaa.gov/smcd/emb/mscat/index.php). These satellite-observation climate data records have been used to compare with climate model runs to determine whether the model accurately captures the atmospheric vertical temperature changes since 1979 (Zhao et al., 2016). Other studies use these temperature data records to monitor changes in the Brewer–Dobson circulation (Young et al., 2011, 2012). Randel et al. (2016) compared global and latitudinal trends from SSU with Aura MLS and SABER temperatures. Simmons et al. (2014) discuss the impacts of the MSU, SSU, AMSU-A, HIRS, and AIRS channels assimilated in the ERA-I. Seidel et al. (2016) intercompared the TLS trends from three satellite centres for the entire (1979–2015) period and separate trends for pre-1997 and post-1997. Mitchell et al. (2015) generated TLS and SSU channel-weighted temperatures from reanalyses to see how well they compare with the satellite observations. We perform a similar exercise by applying the TLS, SSU1, and SSU2 weights to the reanalyses temperatures at their standard pressure-level temperatures. Table 2 provides weighting function information about each of the SSU and MSU-4 channels. SSU3 layer temperatures were not generated because there were insufficient pressure levels from the majority of the reanalyses to adequately represent this layer in the lower mesosphere. Global mean TLS, SSU1, and SSU2 temperatures are generated for each month from 1979 through 2014. Anomalies from the 30-year period (1981–2010) for the TLS, SSU1, and SSU2 are generated. These anomalies are compared against the NOAA STAR SSU v2.0 data set (Zou et al., 2014) and MSU/AMSU mean layer atmospheric temperature v3.0 (Zou and Wang, 2012). The left column of Fig. 18 shows the monthly TLS, SSU1, and SSU2 temperature anomalies from the CFSR, ERA-I, JRA-55, MERRA, and MERRA-2 from 1979 through 2014 with the NOAA STAR anomalies overplotted in black. In general, the anomalies show that the layer temperatures were higher in the 1980s than at present. The El Chichón and Mt. Pinatubo volcanic eruptions increased the layer mean temperature by over 1 K from 1982–1984 and 1991–1993, respectively. Smaller impacts occurred in the SSU1 and SSU2 layer temperatures, as the volcanic influence was mostly in the lower stratosphere. The TLS temperature anomalies show a flat trend between the two volcanos and after Mt. Pinatubo. The SSU1 and SSU2 temperature anomalies have a persistent cooling trend from 1979 to 2010 and have become flatter since then.

To better assess how each reanalysis differs from the NOAA STAR anomalies, the right column shows the differences in the anomalies of each reanalysis from the NOAA STAR anomalies. The reanalyses TLS anomalies differ from the NOAA STAR anomalies by less than ±0.5 K for most of the time series. Most noticeable is that the ERA-I has smaller anomalies than NOAA STAR in the early 1980s and then has larger anomalies after 2006. Aside from the ERA-I, the other reanalyses seem to agree with the NOAA STAR anomalies during the El Chichón volcanic period (1982–1984), with the exception of MERRA and MERRA-2, which have smaller anomalies during the Mt. Pinatubo volcanic period (1991–1993). There is a noticeable decrease in the reanalyses anomalies with respect to the NOAA STAR anomalies in 1999 followed by a gradual increase in time until 2006, after which the reanalyses begin to disagree more with each other. GPSRO observations from the COSMIC constellation became available for assimilation in 2006.

Table 2. Pressure (hPa) of SSU channels 1, 2, and 3 and MSU channel 4 weighting function peaks, 50 % of peak weight above, 50 % of peak weight below, 10 % of peak weight above, and 10 % of peak weight below the peak.

	Peak	50 % above	50 % below	10 % above	10 % below
SSU Ch 3	1.5	0.5	5	0.15	45
SSU Ch 2	3.5	1.5	20	0.30	100
SSU Ch 1	15.0	4.5	60	1.10	150
MSU Ch 4	85.0	35.0	150	15.0	175

The SSU1 temperature anomalies from the CFSR show large temperature jumps associated with the six streams, preventing any useful evaluation. The other four reanalyses differ from the NOAA STAR by less than ±0.5 K for most of the time series. The ERA-I, MERRA, MERRA-2, and JRA-55 all show smaller anomalies than the NOAA STAR in the early 1980s. There is minor disagreement among the four reanalyses with the NOAA STAR between the late 1980s and the early 2000s. MERRA exhibits two spikes in the SSU1 and SSU2 differences from NOAA STAR. The first spike is a result of missing SSU data from 8 April–21 May 1996. The second is from a lack of AMSU-A channel 14 data on NOAA-15 from 30 October–31 December 2000 (W. McCarthy, personal communication, 2017). When there are no observations to constrain the model in the upper stratosphere, analyses migrate to the model climatology, which is warmer than the observations. MERRA-2 found the missing SSU observations in 1996 and began using NOAA-16 AMSU-A observations earlier than in MERRA to shrink the gap to just several days. Beginning in 2006, just as with the TLS anomalies, the disagreement among the four reanalyses increases.

Just as with the SSU1 anomalies, the large temperature jumps associated with the CFSR stream transitions prevent a proper evaluation of its SSU2 time series. Aside from the CFSR, the other four reanalyses are within ±0.5 K of the NOAA STAR anomalies. The JRA-55 matches the NOAA STAR SSU2 observations very well throughout the entire time series with the exception of a period in the late 1990s and early 2000s when its anomalies are smaller than the

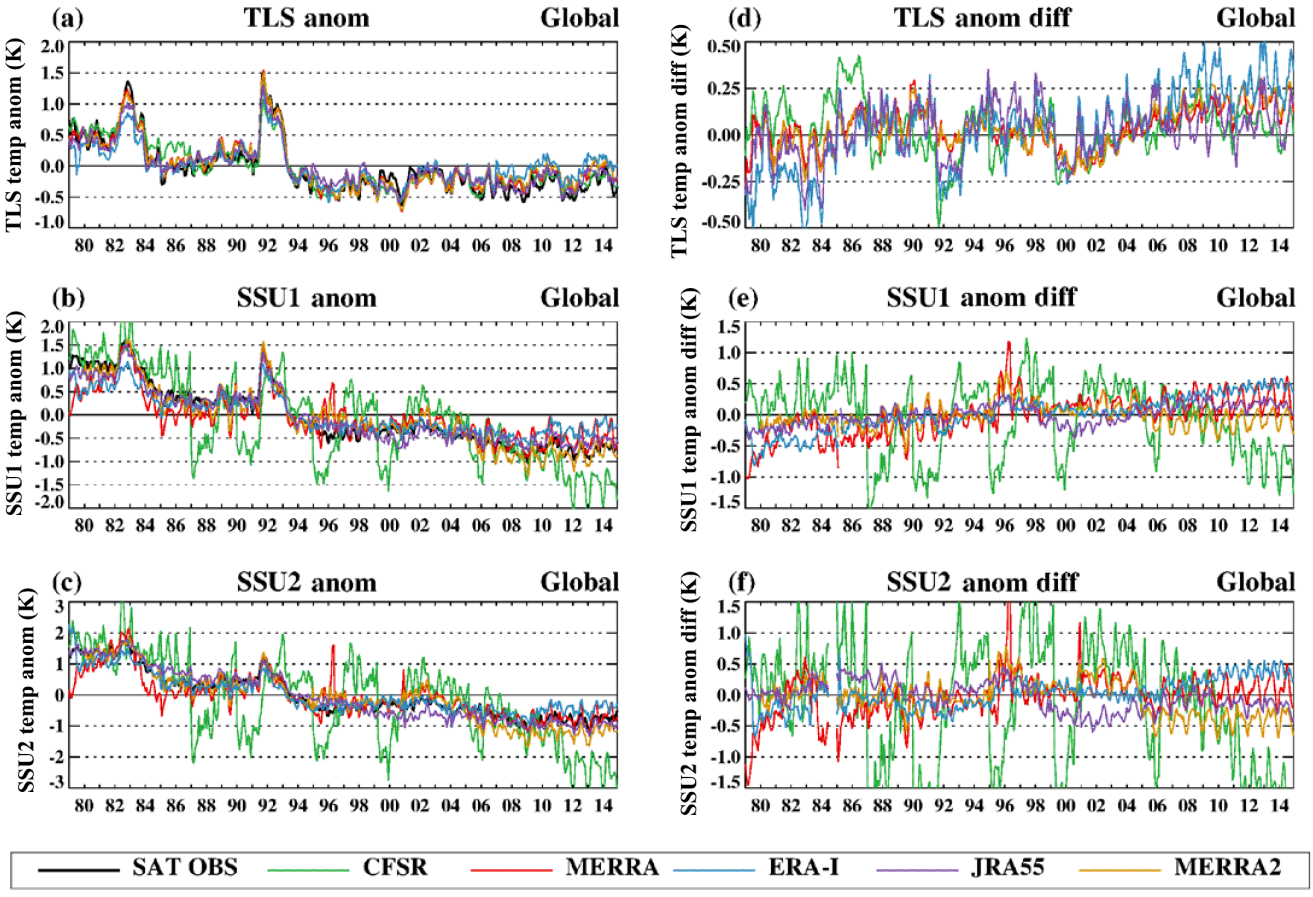


Figure 18. Time series plots of the global layer mean temperature anomalies (K) from the 1981–2010 climatology **(a–c)** and reanalyses anomaly differences from the NOAA STAR anomalies **(d–f)** for **(a, d)** the lower stratosphere (TLS) equivalent to the MSU 4 observations, **(b, e)** the middle stratosphere (SSU1) equivalent to the SSU channel 1 observations, and **(c, f)** the upper stratosphere (SSU2) equivalent to the SSU channel 2 observations. TLS, SSU1, and SSU2 weights are applied to the MERRA-2, MERRA, ERA-I, JRA-55, and CFSR pressure-level data to produce layer mean temperatures and anomalies. NOAA STAR TLS, SSU1, and SSU2 anomalies are plotted along with the reanalyses in the left column.

NOAA STAR anomalies. The ERA-I matches the NOAA STAR SSU2 observations very well except after 2006 when it exhibits a positive trend. Simmons et al. (2014) state that the use of radiosonde data that are not bias adjusted is the likely cause of this trend. MERRA initially begins with lower SSU2 anomalies than NOAA STAR, whereas MERRA-2 anomalies are much closer to the NOAA STAR anomalies. MERRA-2 separates from MERRA after 2005 with more negative anomalies. This is most likely due to the assimilation of MLS temperatures at pressures less than 5 hPa, which have been shown to produce lower temperatures than before 2005.

The CFSR, JRA-55, ERA-I, and MERRA-2 all use GPSRO observations after 2006, yet the later years in Fig. 18 show that their anomalies increasingly disagree with each other after 2006. Figure 19a presents the actual TLS temperatures for these four reanalyses over time from 1980–2014. There is a large spread in the TLS temperatures of 0.8 K between the coldest TLS temperature (ERA-I) and the warmest TLS temperature (CFSR). Over time this large spread decreases until the difference is less than 0.1 K. This illustrates how the various reanalyses actually approach agreement of the TLS values as more observations are assimilated. Figure 19b presents the SD of the four reanalyses TLS temperatures over time. There is a large decrease from 1986 to 1987, which is attributed to the CFSR TLS values cooling during the transition from its initial stream to its second. Another drop in 1999 follows the availability of ATOVS in Fig. 18; the quality and character of the temperature values between 1981 and 2010 changed. This makes generating long-term climatology and anomalies misleading.

Similar comparisons of the SSU1 and SSU2 temperatures are not presented as the temperature biases of each reanalysis above 10 hPa prevents agreement in the layer mean temperature. This shows the value of the GPSRO data to anchor the

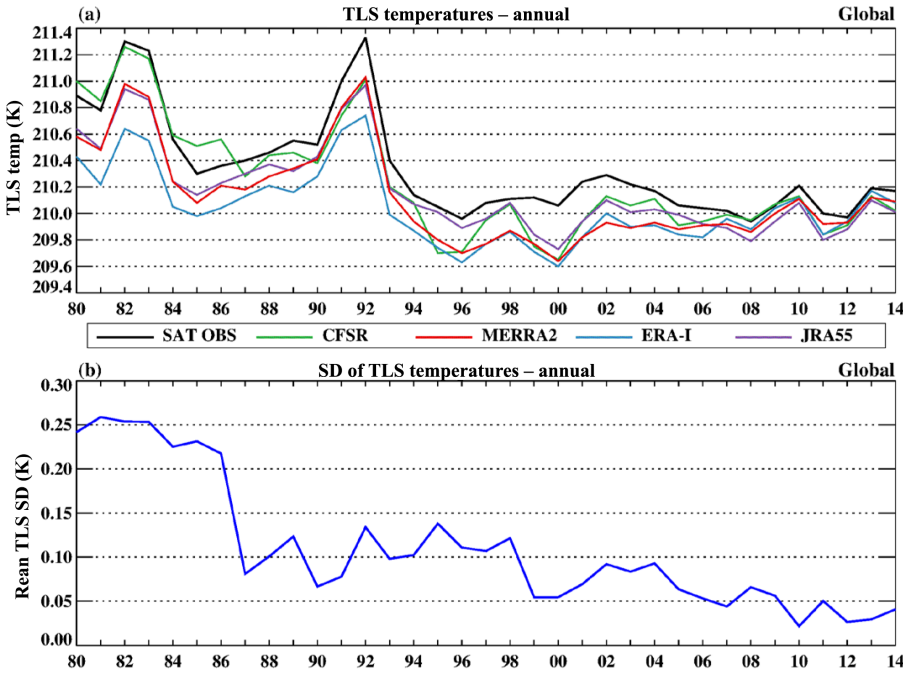


Figure 19. Time series plot of the **(a)** global annual average of the lower stratospheric temperature layer (TLS) temperatures (°C) for MERRA-2, ERA-Interim, JRA-55, CFSR, and the NOAA STAR TLS CDR. **(b)** The TLS temperature SD (K) of the four reanalyses for each year. The climatological period spanned from 1981–2010. COSMIC GPSRO observations began to be assimilated in 2006.

temperatures in the middle and lower stratosphere, which is where most of the TLS weighting function occurs.

8 Summary and conclusions

In this paper a comparison of monthly zonal mean temperatures and zonal winds from the five more recent reanalyses and several older reanalyses were evaluated and intercompared. Our initial evaluation was to look for temperature discontinuities in the time series of each of the reanalyses. This showed that the earlier reanalyses (ERA-40 and JRA-25) had multiple temporal discontinuities in the 1980s in the stratosphere associated with changes in the biases of the data from the NOAA TOVS and SSU instruments. The R1 and R2 did not show such discontinuities because they used NESDIS-generated temperature profiles, not the original radiance data. NESDIS most likely strived to minimize such discontinuities in the profile temperatures. Almost all the reanalyses have a temporal discontinuity in 1998 when the ATOVS observations became available and the reanalyses either switched immediately or transitioned from the TOVS to the ATOVS over several years. The CFSR has temporal discontinuities at the time of switching from one stream to the next. The CFSR bias corrected the top SSU channel 3. The model used by the CFSR had a warm bias in the upper stratosphere and slowly warmed about 5 K during the course of each stream. Because of the presence of the discontinuities and transitions discussed above, great caution should be exercised in using reanalyses for trend analysis and/or trend detection, especially in the middle and upper stratosphere.

So as not to favour any one particular reanalysis, a reanalysis ensemble mean (REM) of three of the more recent reanalyses (MERRA, ERA-I, and JRA-55) was generated. We presented the climatological mean (1981–2010) of the temperature and zonal wind REM and showed the altitudes and seasons with the largest variance in the REM. The temperature and zonal winds have the greatest interannual variability in the NH polar region from January through March because of the large variability in wave activity, including the frequent occurrence of strong stratospheric warming events. This variability is greatest in the upper stratosphere as planetary-scale wave amplitudes and the associated temperature and zonal wind changes during strong stratospheric warming events are largest in the upper stratosphere. In the SH polar region the interannual variability is not as large in magnitude and is prevalent throughout the stratosphere. Because midwinter wave activity is much smaller in the SH, most of the interannual variability in the SH polar region is associated with the springtime transition to summer circulation patterns and polar vortex breakdown when wave activity shows larger interannual variability in timing and magnitude.

Time series of the temperature variance in the three REM members showed that the greatest disagreement occurs during the TOVS time period (1979–1998) in all latitude zones, and agreement improves during the ATOVS time period

(1999 to present). The disagreement in the SH polar latitudes extended lower into the stratosphere than in the NH polar latitudes. The zonal wind variance was smaller than the temperature variance in the polar latitudes, but had a similar temporal difference between the TOVS and ATOVS time periods. In the tropics, the zonal wind variance was much larger than in the polar regions as the disagreement of the SAO and QBO zonal winds was quite large. Thus, improving equatorial winds in future reanalyses is an important goal.

The characteristics of each reanalysis were identified as differences from the temperature and zonal wind REM. The CFSR had a seasonal warm bias compared to the REM in the upper stratosphere that persisted during both the TOVS and ATOVS time periods. The JRA-55, on the other hand, had a seasonal cold bias that persisted during both the TOVS and ATOVS time periods. ERA-I and MERRA had smaller differences from the temperature REM except that the ERA-I had a warm bias in the SH polar latitudes between 7 and 5 hPa that occurred only during the austral winter and only during the TOVS time period. MERRA-2 had very small differences from the REM except in the upper stratosphere in the polar regions where it had a year-long cool bias at 1 hPa and a warm bias between 2 and 3 hPa. These biases greatly diminished during the ATOVS period. Temperature differences from the REM in the earlier reanalyses (JRA-25, ERA-40, and R1) extended throughout the stratosphere and the upper troposphere. These differences occurred through both the TOVS and ATOVS time periods. This illustrates the progress made by the reanalysis centres to improve the analyses from the earlier versions to the later versions. This results in better agreement among the more recent reanalyses.

In the tropics, the individual reanalyses exhibited smaller temperature differences than in the polar latitudes. However, the characteristic biases in the upper stratosphere observed in the polar latitudes were maintained in the tropics. The zonal wind differences from the REM of the individual reanalyses are very large in the SAO region. In the QBO region the differences frequently show dissimilarities in the timing of the descending westerlies and easterlies as well as the amplitude of these winds. Zonal wind differences from the REM were not confined to the stratosphere as several reanalyses also had sizable differences in the troposphere.

Specifically comparing the more recent reanalyses QBO zonal winds (70–10 hPa) against the zonal winds observed at Singapore using the FUB data set showed that the CFSR had the largest RMS differences from the Singapore winds than the other reanalyses at most levels and during both the TOVS and ATOVS periods However, MERRA-2 10 hPa zonal winds were nearly twice as large as the other reanalyses during the TOVS period, mostly due to an overly aggressive gravity wave parameterization. The RMS differences from the Singapore zonal winds were smaller during the ATOVS period for all the reanalyses. The CFSR had the largest amplitude biases from the Singapore winds as shown by the linear slope of their matched monthly values. The linear slopes of all the reanalyses were furthest from unity at 50 and 30 hPa during the TOVS period.

There are several reasons why the ATOVS period is an improvement over the TOVS period. The primary reason is that the AMSU-A instrument has five narrower channels in the stratosphere instead of the broader three SSU channels. (The MSU channel 4 and AMSU-A channel 9 weighting functions are almost identical.) Another reason is that the SSU was the only instrument monitoring the thermal structure of the stratosphere from 1978 through 1998. From 1999 onward there are additional satellite instruments monitoring the stratosphere: AIRS, IASI, MLS, and GPSRO. Hence the quantity and quality of data monitoring in the stratosphere increases from 1999 to the present.

The amplitude of the annual temperature cycle (warmest summer month minus the coldest winter month) in the SH polar latitudes is larger than the NH polar latitude temperature amplitude by 5–15 K. The region of large amplitude extends throughout the middle and upper stratosphere in the SH polar latitudes. In the NH polar latitudes the vertical region of large temperature amplitudes is confined to the upper stratosphere and occurs during the years with an SSW. The ERA-I has a noticeably smaller annual temperature amplitude in the SH polar latitudes than the other ensemble members from 3 to 30 hPa. This is due to its warm bias during the SH winter months in this latitude region. The CFSR temperature amplitude decreases rapidly above 3 hPa due to its warm bias in the upper stratosphere in both SH and NH polar latitudes.

Comparisons against HIRDLS (January 2005–March 2008) and Aura MLS (2005–2014) temperatures concur with the previous characteristics of the various reanalyses in the upper stratosphere. The CFSR has a definite warm bias compared to HIRDLS temperatures, while the JRA-55 has a definite cold bias. Both MERRA and ERA-I have a slight warm bias during the summer months between 3 and 7 hPa. MERRA has a slight cold bias above this between 1 and 2 hPa nearly all year long. MERRA-2 assimilates Aura MLS temperatures at pressures less than 5 hPa and consequently differences are very small.

The NOAA STAR TLS, SSU1, and SSU2 data sets (Zou et al., 2014; Zou and Qian, 2016) are a much improved CDR than the version used in Thompson et al. (2012), which pointed out the dissimilarities between the NOAA and Met Office SSU data records. The comparison between the version used in this paper and the appropriately weighted reanalyses is much better than previous papers using the older version and the Met Office CDR. All of the more recent reanalyses capture the characteristics of the NOAA STAR TLS anomalies. Excluding the CFSR, the other reanalyses (MERRA-2, MERRA, ERA-I, and JRA-55) capture the basic features of the SSU1 and SSU anomalies. We learn from this intercomparison that the GPSRO observations provide an anchor that drives the reanalyses to closer agreement in the middle and lower stratosphere. We also learn that using a long period climatology may not be the best practice to gen-

erate anomalies in parts of the atmosphere which are more sensitive to the changes in data sources, which impacts their quality and accuracy over time.

In this paper we have examined the thermal and dynamical characteristics of the older and the more recent reanalyses. We find that the more recent reanalyses have fewer discontinuities in their temperature and wind time series due to better data assimilation techniques and transition among different sets of observations. We also find that the larger temperature and wind differences among the older reanalyses have become smaller among the more recent reanalyses. However, the transition from the TOVS to ATOVS satellite periods continues to be problematic. The reanalysis QBO winds during the ATOVS period also agree much better with the Singapore radiosonde observations than during the TOVS period. We expect that future reanalyses will have better QBO winds as their forecast models improve to produce a spontaneous QBO in the tropics. We have shown that the more recent reanalyses agree quite well with each other in the lower and middle stratosphere, but greater differences exist in the upper stratosphere and lower mesosphere. This latter disagreement is a result of differences in model top and vertical resolution and what data is assimilated. Due to these disagreements we caution data users from using any one reanalysis for comparisons and even more so for the detection of trends and/or changes in climate.

Improving the TOVS time period would be highly beneficial to future reanalyses. However, the TOVS time period may never be as good as the ATOVS period due to the sparsity of data. Model improvements, improvements to the variational bias corrections to handle the broad SSU weighting functions, and non-orographic gravity wave parameterization improvements (so the forecast models can generate a QBO on their own) are some of the ways this period can be improved upon.

Additional literature will be generated from other aspects of the S-RIP initiative. An evaluation of ozone and water vapour in the reanalyses has recently been published (Davis et al., 2017). Future work will evaluate the following: the Brewer–Dobson circulation; stratosphere–troposphere dynamical coupling; upper tropospheric–lower stratospheric processes,; stratospheric–tropospheric exchange in the extratropics and tropics; the QBO, SAO, and tropical variability; stratospheric polar dynamic and chemical processes that lead to ozone depletion; and dynamics and transport in the upper stratosphere–lower mesosphere.

Data availability. Reanalysis data were obtained from these public sites: MERRA-2: https://disc.sci.gsfc.nasa.gov/uui/datasets?keywords=%22MERRA-2%22, MERRA: https://disc.sci.gsfc.nasa.gov/uui/datasets?keywords=%22MERRA%22, ERA-I: http://apps.ecmwf.int/datasets/, ERA-40: http://apps.ecmwf.int/datasets/, JRA-25: through NCAR RDA at https://rda.ucar.edu/datasets/ds625.1, JRA-25: through NCAR RDA at https://rda.ucar.edu/datasets/ds628.1, CFSR, pressure-level data: through NCAR RDA at https://rda.ucar.edu/datasets/ds094.2, R1, pressure-level data: through NCAR RDA at https://rda.ucar.edu/datasets/ds090.0, R2, pressure-level data: through NCAR RDA at https://rda.ucar.edu/datasets/ds091.0 and 20CR, pressure-level data: through NCAR RDA at https://rda.ucar.edu/datasets/ds131.1.

Appendix A: Major abbreviations and terms

20CR	20th Century Reanalysis of NOAA and CIRES
ATOVS	Advanced TIROS Operational Vertical Sounder
AMSU	Advanced Microwave Sounding Unit (AMSU-A for Unit A)
Aura	A satellite in the EOS A-Train satellite constellation
CIRA86	COSPAR International Reference Atmosphere, 1986
CIRES	Cooperative Institute for Research in Environmental Sciences (NOAA and the University of Colorado Boulder)
CFSR	Climate Forecast System Reanalysis of NCEP
COSMIC	Constellation Observing System for Meteorology, Ionosphere, and Climate
COSPAR	Committee on Space Research
CRTM	Community Radiative Transfer Model
DOE	Department of Energy
ECMWF	European Centre for Medium-Range Weather Forecasts
EOS	NASA Earth Observing System
ERA-15	ECMWF 15-year reanalysis
ERA-40	ECMWF 40-year reanalysis
ERA-I or ERA-Interim	ECMWF interim reanalysis
GLATOVS	Goddard Laboratory for Atmospheres TOVS forward model
GPSRO	Global Positioning System radio occultation
GENESIS	Global Environmental and Earth Science Information System
HIRDLS	High Resolution Dynamics Limb Sounder
JRA-25	Japanese 25-year reanalysis
JRA-55	Japanese 55-year reanalysis
MERRA	Modern Era Retrospective-Analysis for Research (MERRA-2 for its version 2)
MLS	Microwave Limb Sounder
MSU	Microwave Sounding Unit
NASA	National Aeronautics and Space Administration
NCAR	National Center for Atmospheric Research
NCEP	National Centers for Environmental Prediction of NOAA
NESDIS	National Environmental Satellite, Data, and Information Service of NOAA
NH	Northern Hemisphere
NOAA	National Oceanic and Atmospheric Administration
NOAA-*	NOAA polar-orbiting operational meteorological satellite (* indicates the satellite number)
QBO	Quasi-biennial oscillation
R1	NCEP-NCAR Reanalysis 1
R2	NCEP-DOE Reanalysis 2
REM	Reanalysis ensemble mean
RMS	Root mean square
S-RIP	SPARC Reanalysis Intercomparison Project
SAO	Semi-annual oscillation
SH	Southern Hemisphere
SPARC	Stratosphere–troposphere Processes and their Role in Climate
SSU	Stratospheric Sounding Unit (SSU1 and SSU2 for SSU channel 1 and 2, respectively)
SSW	Sudden stratospheric warming
STAR	Center for Satellite Applications and Research of NESDIS
TIROS	Television Infrared Observation Satellite
TLS	Temperature of the lower stratosphere (MSU channel 4 and AMSU channel 9)
TOVS	TIROS Operational Vertical Sounder
WMO	World Meteorological Organization

The Supplement related to this article is available online at https://doi.org/10.5194/acp-17-14593-2017-supplement.

Competing interests. The authors declare that they have no conflict of interest.

Special issue statement. This article is part of the special issue "The SPARC Reanalysis Intercomparison Project (S-RIP) (ACP/ESSD inter-journal SI)". It is not associated with a conference.

Acknowledgements. We acknowledge the scientific guidance and sponsorship of the World Climate Research Programme coordinated in the framework of SPARC. We are grateful for the internal reviews, alterations, and improvements by two of the S-RIP co-leads Gloria L. Manney and Lesley J. Gray. We thank the reanalysis centres for providing their support and data products. We thank Michael Schwartz for providing the MLS monthly temperature data. Masatomo Fujiwara's contribution was financially supported in part by the Japan Society for the Promotion of Science (JSPS) through Grants-in-Aid for Scientific Research (26287117 and 16K05548), and Corwin Wright's was supported by NERC grant NE/K015117/1.

Edited by: Peter Haynes
Reviewed by: three anonymous referees

References

Baldwin, M. P. and Gray, L. J.: Tropical stratospheric zonal winds in ECMWF ERA-40 reanalysis, rocketsonde data and rawinsonde data, Geophys. Res. Lett., 32, L09806, https://doi.org/10.1029/2004GL022328, 2005.

Baldwin, M. P., Gray, L. J., Dunkerton, T. J., Hamilton, K., Haynes, P. H., Randel, W. J., Holton, J. R., Alexander, M. J., Hirota, I., Horinouchi, T., Jones, D. B. A., Kinnersley, J. S., Marquardt, C., Sato, K., and Takahashi, M.: The quasi-biennial oscillation, Rev. Geophys., 39, 179–229, https://doi.org/10.1029/1999RG000073, 2001.

Charlton, A. J. and Polvani, L. M.: A new look at stratospheric sudden warmings, Part I: Climatology and modelling benchmarks, J. Climate, 20, 449–469, https://doi.org/10.1175/JCLI3996.1, 2007.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteorol. Soc., 137, 1–28, https://doi.org/10.1002/qj.776, 2011.

Coy, L., Wargan, K., Molod, A. M., McCarty, W. R., and Pawson, S.: Structure and dynamics of the quasi-biennial oscillation in MERRA-2, J. Climate, 29, 5339–5354, https://doi.org/10.1175/JCLI-D-15-0809.1, 2016.

Cucurull, L,. Derber, J. C., Treadon, R., and Purser, R.: Assimilation of global positioning system radio occultation observations into NCEP's global data assimilation system, Mon. Weather Rev., 35, 3174–3193, 2007.

Davis, S. M., Hegglin, M. I., Fujiwara, M., Dragani, R., Harada, Y., Kobayashi, C., Long, C., Manney, G. L., Nash, E. R., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapor and ozone in reanalyses as part of S-RIP, Atmos. Chem. Phys., 17, 12743–12778, https://doi.org/10.5194/acp-17-12743-2017, 2017.

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. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Fujiwara, M., Hibino, T., Mehta, S. K., Gray, L., Mitchell, D., and Anstey, J.: Global temperature response to the major volcanic eruptions in multiple reanalysis data sets, Atmos. Chem. Phys., 15, 13507–13518, https://doi.org/10.5194/acp-15-13507-2015, 2015.

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 Intercomparison 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.

Gelaro, R., McCarty, W., Suarez, M., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., Da Silva, A., Gu, W., Kim, G., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S., Sienkiewicz, M., and Zhao, B.: The Modern-Efa 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.

Gille, J., Barnett, J., Arter, P., Barker, M., Bernath, P., Boone, C., Cavanaugh, C., Chow, J., Coffey, M., Craft, J., Craig, C., Dials, M., Dean, V., Eden, T., Edwards, D. P., Francis, G., Halvorson, C, Harvey, L., Hepplewhite, C., Khosravi, R., Kinnison, D., Krinsky, C., Lambert, A., Lee, H., Lyjak, L., Loh, J., Mankin, W., Massie, S., McInerney, J., Moorhouse, J., Nardi, B., Packman, D., Randall, C., Reburn, J., Rudolf, W., Schwartz, M., Serafin, J., Stone, K., Torpy, B., Walker, K., Waterfall, A.,

Watkins, R., Whitney, J., Woodard, D., and Young, G.: High resolution dynamics limb sounder: experiment overview, recovery, and validation of initial temperature data, J. Geophys. Res., 113, D16S43, https://doi.org/10.1029/2007JD008824, 2008.

Gibson, J. K., Kållberg, P., Uppala, S. M., Nomura, A., Hernandez, A., and Serrano, E.: ECMWF Re-analysis Final Project Report, Series 1, ECMWF, Reading, 72 pp., 1997.

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

Gray, L. and Pyle, J.: Two-dimensional model studies of equatorial dynamics and tracer distributions, Q. J. Roy. Meteorol. Soc., 113, 635–651, https://doi.org/10.1002/qj.49711347611, 1987.

Han, Y., van Delst, P., Liu, Q., Weng, F., Yan, B., Treadon, R., and Derber, J.,: JCSDA Community Radiative Transfer Model (CRTM) – Version 1, NOAA Tech. Rep. 122, NOAA/NESDIS, USA, 33 pp. 2006.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40 year reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, 1996.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II reanalysis (R-2), B. Am. Meteorol. Soc., 83, 1631–1643, 2002.

Kawatani, Y., Hamilton, K., Miyazaki, K., Fujiwara, M., and Anstey, J. A.: Representation of the tropical stratospheric zonal wind in global atmospheric reanalyses, Atmos. Chem. Phys., 16, 6681–6699, https://doi.org/10.5194/acp-16-6681-2016, 2016.

Kistler, R., Collins, W., Saha, S., White, G., Woollen, J., Kalnay, E., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP–NCAR 50 year reanalysis: monthly means CD-ROM and documentation, B. Am. Meteorol. Soc., 82, 247–267, 2001.

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. Jpn., 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

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.

Leroy, S., Ao, C., and Verkhoglyadova, O. P.: Mapping GPS Radio Occultation Data by Bayesian Interpolation, O. P., J. Atmos. Ocean. Tech., 29, 1062, https://doi.org/10.1175/JTECH-D-11-00179.1, 2012.

Livesey, N., Read, W., Wagner, P., Froidevaux, L., Lambert, A., Manney, G., Milan Valle, L., Pumphrey, H., Santee, M., Schwartz, M., Wang, S., Fuller, R., Jarnot, R., Knosp, B., and Martinez, E.: EOS MLS version 4.2x level 2 data quality and description document, Tech. Rep. D-33509 Rev. A, Jet Propul. Lab., Calif. Inst. of Technol., Pasadena, Calif., 2015.

Manney, G. L., Allen, D. R., Kruger, K., Naujokat, B., Santee, M. L., Sabutis, J. L., Pawson, S., Swinbank, R., Randall, C. E., Simmons, A. J., and Long, C.: Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter, Mon. Weather Rev., 133, 1261–1278, https://doi.org/10.1175/MWR2926.1, 2005.

Martineau, P. and Son, S.-W.: Quality of reanalysis data during stratospheric vortex weakening and intensification events, Geophys. Res. Lett., 37, L22801, https://doi.org/10.1029/2010GL045237, 2010.

Martineau, P., Son, S.-W., and Taguchi, M.: Dynamical consistency of reanalysis datasets in the extratropical stratosphere, J. Climate, 29, 3057–3074, https://doi.org/10.1175/JCLI-D-15-0469.1, 2016.

Mitchell, D. M., Gray, L. J., Fujiwara, M., Hibino, T., Anstey, J. A., Ebisuzaki, W., Harada, Y., Long, C., Misios, S., Stott, P. A., and Tan, D.: Signatures of naturally induced variability in the atmosphere using multiple reanalysis datasets, Q. J. Roy. Meteorol. Soc., 141, 2011–2031, https://doi.org/10.1002/qj.2492, 2015.

Miyazaki, K., Iwasaki, T., Kawatani, Y., Kobayashi, C., Sugawara, S., and Hegglin, M.: Inter-comparison of stratospheric mean-meridional circulation and eddy mixing among siz reanalysis datasets, Atmos. Chem. Phys., 16, 6131–6152, https://doi.org/10.5194/acp-16-6131-2016, 2016.

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

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., Matsumoto, T., Yamazaki, N., Kamahori, H., Takahashi, K., Kadokura, S., Wada, K., Kato, K., Oyama, R., Ose, T., Mannoji, N., and Taira, R.: The JRA-25 reanalysis, J. Meteorol. Soc. Jpn., 85, 369–432, https://doi.org/10.2151/jmsj.85.369, 2007.

Pawson, S. and Fiorino, M.: A comparison of reanalyses in the tropical stratosphere: Part 1: Thermal structure and the annual cycle, Clim. Dynam., 14, 631–644, 1998a.

Pawson, S. and Fiorino, M.: A comparison of reanalyses in the tropical stratosphere: Part 2: The quasi-biennial oscillation, Clim. Dynam., 14, 645–658, 1998b.

Pawson, S. and Fiorino, M.: A comparison of reanalyses in the tropical stratosphere: Part 3: Inclusion of the pre-satellite data era, Clim. Dynam., 15, 241–250, 1999.

Pawson, S. and Naujokat, B.: The cold winters of the middle 1990s in the northern lower stratosphere, J. Geophys. Res., 104, 14209–14222, https://doi.org/10.1029/1999JD900211, 1999.

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. Meteorol. Soc., 136, 1972–1990, https://doi.org/10.1002/qj.722, 2010.

Randel, W., Udelhofen, P., Fleming, E., Geller, M., Gelman, M., Hamilton, K., Karoly, D., Ortland, D., Pawson, S., Swinbank, R., Wu, F., Baldwin, M., Chanin, M.-L., Kechhut, P., Labitzke, K., Remsberg, E., Simmons, A., and Wu, D.: The SPARC intercomparison of middle-atmosphere climatologies, J. Climate, 17, 986–1003, 2004.

Randel, W., Smith, A. K., Wu, F., Zou, C.-Z., and Qian, H.: Stratospheric temperature trends over 1979–2015 derived from combined SSU, MLS, and SABER satellite observations,

J. Climate, 29, 4843–4858, https://doi.org/10.1175/JCLI-D-15-0629.1, 2016.

Ray, E. A., Alexander, M. J., and Holton, J. R.: An analysis of the structure and forcing of the equatorial semiannual oscillation in zonal wind, J. Geophys. Res., 103, 1759–1774, https://doi.org/10.1029/97JD02679, 1998.

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., van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., 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, B. Am. Meteorol. Soc., 91, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1, 2010.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: The NCEP climate forecast system version 2, J. Climate, 27, 2185–2208, https://doi.org/10.1175/JCLI-D-12-00823.1, 2014.

Schwartz, M. J., Lambert, A., Manney, G., Read, W., Livesey, N., Froidevaux, L., Ao, C., Bernath, P., Boone, C., Cofield, R., Daffer, W., Drouin, B., Fetzer, D., Fuller, R., Jarnot, R., Jiang, J., Jiang, Y., Knosp, B., Krüger, K., Li, J.-L., Mlynczak, M., Pawson, S., Russell, J., Santee, M., Snyder, W., Stek, P., Thurstans, R., Tompkins, A., Wagner, P., Walker, K., Waters, J., and Wu, D.: Validation of the aura microwave limb sounder temperature and geopotential height measurements, J. Geophys. Res., 113, D15S11, https://doi.org/10.1029/2007JD008783, 2008.

Seidel, D. J., Li, J., Mears, C., Moradi, I., Nash, J., Randel, W. J., Saunders, R., Thompson, D. W. J., and Zou, C.-Z.: Stratospheric temperature changes during the satellite era, J. Geophys. Res.-Atmos., 121, 664–681, https://doi.org/10.1002/2015JD024039, 2016.

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. Meteorol. Soc., 140, 329–353, https://doi.org/10.1002/qj.2317, 2014.

SPARC: SPARC intercomparison of middle atmosphere climatologies, edited by: Randel, W., Chanin, M.-L., and Michaut, G., WCRP-116, WMO/TD-No. 1142, SPARC Rep. 3, Verrières-le-Buisson CEDEX, France, 96 pp., 2002.

Susskind, J., Rosenfield, J., and Reuter, D.: An accurate radiative transfer model for use in the direct physical inversion of HIRS and MSU temperature sounding data, J. Geophys. Res., 88, 8550–8568, 1983.

Thompson, D., Seidel, D., Randel, W., Zou, C.-Z., Butler, A., Lin, R., Long, C., Mears, C., and Osso, A.: The mystery of recent stratospheric temperature trends, Nature, 491, 692–697, https://doi.org/10.1038/nature11579, 2012.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 reanalysis, Q. J. Roy. Meteorol. Soc., 131, 2961–3012, https://doi.org/10.1256/qj.04.176, 2005.

Young, P. J., Thompson, D. W. J., Rosenlof, K. H., Solomon, S., and Lamarque, J.: The seasonal cycle and interannual variability in stratospheric temperatures and links to the Brewer–Dobson circulation: an analysis of MSU and SSU Data, J. Climate, 24, 6243–6258, 2011.

Young, P. J., Rosenlof, K. H., Solomon, S., Sherwood, S. C., Fu, Q., Lamarque, J.-F.: Changes in stratospheric temperatures and their implications for changes in the Brewer–Dobson circulation, 1979–2005, J. Climate, 25, 1759–1772, 2012.

Zhao, L., Xu, J., Powell, A. M., Jiang, Z., and Wang, D: Use of SSU/MSU satellite observations to validate upper atmospheric temperature trends in CMIP5 simulations, Remote Sens., 8, https://doi.org/10.3390/rs8010013, 2016.

Zou, C.-Z. and Qian, H.: Stratospheric temperature climate data record from merged SSU and AMSU-A observations, J. Atmos. Ocean. Tech., 33, 1967–1984, https://doi.org/10.1175/JTECH-D-16-0018.1, 2016.

Zou, C.-Z. and Wang, W.: MSU/AMSU radiance fundamental climate data record calibrated using simultaneous nadir overpasses Climate Algorithm Theoretical Basis Document (C-ATBD), NOAA/NESDIS, USA, 2012.

Zou, C.-Z., Qian, H., Wang, W., Wang, L., and Long, C.: Recalibration and merging of SSU observations for stratospheric temperature trend studies, J. Geophys. Res.-Atmos.,119, 13180–13205, https://doi.org/10.1002/2014JD021603, 2014.