Response: We thank Referee #1 for these valuable comments and suggestions.

Reviewer_1: This study provides a comprehensive overview of the temperature and zonal wind biases in eight reanalysis data products, with a focus on stratospheric levels. The study identifies biases in each reanalysis from the "reanalysis mean" (defined as the mean of the MERRA, ERA-Interim, and JRA-55 reanalyses). It then examines reanalysis temperature biases with respect to HIRDLS (an independent satellite measurement) and MSU/AMSU/SSU satellite data products. The authors identify systematic biases and notable change points in the reanalyses associated with discontinuities in data sources, such as the transition from TOVS to ATOVS around 1998-1999. One of the key conclusions of the study is the pervasive uncertainty in zonal winds in the tropical stratosphere, largely because of the inability of reanalyses to resolve the waves that drive zonal wind variability in this region.

Reviewer_1: This paper is not likely one that most readers will read from beginning to end, as it contains a highly technical description of reanalysis biases. While many of the is- sues discussed have been discussed in previous literature, this document serves a centralized review by the SPARC S-RIP Project of these issues, providing a guidance document to reanalysis users (to understand biases) and to reanalysis data centers (to improve upon existing reanalysis products). For these reasons, I recommend publication of this manuscript. However, I think the paper would be more useful if it provided more detailed guidance and suggestions as to the improvements necessary in future reanalysis products. Comments and suggested revisions are detailed below.

Response: In the summary we will include comments about the TOVS period and the transition to ATOVS. 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 (VBC) 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.

Changes in Manuscript: P 28 L 1-5 added response to summary section

Minor Revisions

1. Reviewer_1: The authors could do more to provide guidance to improve future reanalysis products, particularly focusing on what improvements were already made from ERA-40 to ERA-Interim, JRA-25 to JRA-55, and MERRA to MERRA-2 to reduce biases. This knowledge would be particularly helpful in interpreting the results in Figs. 6-9, where the authors compare the biases among these reanalysis products. For example, if ERA-Interim has smaller biases than ERA-40 in a certain region, it would be useful to more clearly emphasize what improvements might have reduced these biases.

Response: We will add a section briefly highlighting the improvements from the older version to newer reanalysis. The common improvements are to the Radiative Transfer Model (RTM) in both the forecast model and that used in the assimilation step, model horizontal and vertical resolution, and bias correction.

Changes in Manuscript: P 4 L 29, P 5 L 1-9 Added a section and Table 1 discussing the important changes from the older reanalysis version to the newer version for each reanalysis center.

2. Reviewer_1: I'm curious as to why the authors did not directly evaluate the reanalysis temperatures against GPSRO data. GPSRO provides high vertical resolution satellite-derived temperature measurements up to 40 km altitude. It is clear from Fig. 15 that the inclusion of GPSRO data in some reanalysis products had a substantial impact after 2006.

Response: A section with supporting figures will be added showing the comparisons of the more recent reanalyses (CFSR, ERA-I, JRA55, MERRA, and MERRA2 vs COSMIC monthly zonal mean temperature from 400-10 hPa for the years 2007-2014.

Changes in Manuscript: P 21 L 11-31, P 22 L 1-12 Added comparisons of the reanalyses with COSMIC dry temperatures from 2007-2014. Presented reanalysis-COSMIC temperature differences in Figure 15, 16, and 17.

3. Reviewer_1: I'm also curious about why the authors focus on the polar regions and tropics and do not discuss biases at midlatitudes. Is there a reason why midlatitudes are not discussed in this paper?

Response: A section (without supporting figures) will be added discussing the mid-latitudes.
Changes in Manuscript: Mid-latitude paragraphs are added to:

Section 4.3.1 P 10 L 11-17
Section 4.3.2 P 11 L 13-19
Section 5.1.4 P 15 L 19-23
Section 5.2.4 P 17 L 5-16

4. The paper deserves a thorough and careful proofreading. I caught a number of inconsistencies between the manuscript text and the figures, which need to be corrected prior to publication. I've listed some examples below, but I'm sure there are others that I may have missed.

a. p. 8, Line 27: In Fig. 4c, the disagreement between 7 and 5 hPa appears to terminate in 2002, not in 1998 (TOVS/ATOVS transition).

Response: Will check and correct text as needed.

Changes in Manuscript: P 10 L 9 reviewer is correct and text corrected.

b. p. 11, Line 11: persistent cool bias from August to November

Response: Will do

Changes in Manuscript: P 13 L 19 Corrected

c. p. 11, Line 12: upper stratosphere warm bias

Response: Will do

Changes in Manuscript: P 13 L 20 Corrected

d. p. 12, Line 14: In Fig. 8i, the CFSR biases near 100 hPa appear to stop at the TOVS/ATOVS transition, not continue through it as the text states.

Response: Will do

Changes in Manuscript: P 14 L 30 Corrected

e. p. 12, Line 16: 0.5 to 2 K

Response: Will do

Changes in Manuscript: P 15 L 2 Corrected

f. p. 12, Lines 19-28: Please double-check the magnitudes in this paragraph, as they seem inconsistent with Fig. 8f.

Response: Will do

Changes in Manuscript: P 15 L5-13 Checked and Corrected

g. p. 13, Line 25: It does not appear from Fig. 9m that the westerlies are stronger during the TOVS period. They look stronger throughout the entire data record.

Response: Will check and correct text as needed

Changes in Manuscript: P 16 L 15 The MERRA-2 SAO westerlies are much stronger during the TOVS period. Text is not changed..

h. p. 15: The color ranges in Fig. 11 do not match those discussed in the text in section 5.

Response: Will check and correct text as needed.

Changes in Manuscript: P 19 & 20 Ranges are changed to be consistent with the color bar.

i. p. 17, Lines 4-6: In Fig. 14b, the MERRA warm bias only occurs in November through February during the first year (Nov. 2005-Feb. 2006). After that, the warm bias is primarily confined to the 5-10 hPa pressure range.

Response: Will check and correct text as needed.

Changes in Manuscript: P 20 L 26 Text is corrected.

j. p. 20, Line 14: cool bias at 1 hPa and warm bias between 2-3 hPa

Response: Will check and correct text as needed.

Changes in Manuscript: P 26 L 3-4 Corrected

k. Line-by-line comments

p. 1, Line 19: among the reanalyses themselves
Response: Will do
Changes in Manuscript: P 1 L19 Changed

p. 2, Line 19: I didn't see any mention of the v and w wind fields in the text.

Response: Focused this paper upon temps and zonal winds.

Changes in Manuscript: No changes to text

p. 5, Line 13: The volcanic warming is primarily confined to the lower stratosphere.
Response: Yes
Changes in Manuscript: P 5 L 28 Figure 1 shows a response as well in the middle stratosphere. Text changed.

p. 6, Line 27: Why do the minimum temperatures occur before the winter solstice?

Response: The polar circulation forms in austral fall at the top of the stratosphere shutting out horizontal advection. Radiative cooling takes over and progressively moves towards the surface.

Changes in Manuscript: P 8 L 3-5 Explanation added to text.

p. 7, Line 12: This sentence seems out of place. The QBO and SAO are not introduced until the following paragraph.

Response: Will check and correct text as needed.

Changes in Manuscript: P 8 L 13-14 the text is discussing the temperature response to the SAO and QBO. No changes are made.

p. 9, Line 27: How large are the 20CR biases in the stratosphere? It might be useful to warn readers against using 20CR data, as large biases in stratospheric dynamics might also have a substantial impact at tropospheric levels.

Response: Will add text to summarize deficiencies of the 20CR:

- 20CR does not have a QBO, hence no time variability of temps and winds in the lower stratospheric tropics,
- 20CR does not capture SSW, hence NH winter temps are > 5C colder, and polar jets are stronger
- 20CR is 3-4C warmer at 100 hPa in the tropics (possible result of coarse model vertical resolution)
- 20CR has larger annual temp oscillation from 200-850hPa in the tropics.

Changes in Manuscript: P 11 L 24-27 The 20CR deficiencies are added to the text.

p. 14, Lines 5-11: MERRA-2 is not discussed in this paragraph, but it looks as if it also has sizeable wind biases in the tropical troposphere.

Response: Will check and correct text as needed

Changes in Manuscript: P 17 L 2-4 MERRA-2 discussion is added.

p. 14, Line 20: 1980-2014 period

Response: Will check and correct text as needed

Changes in Manuscript: P 17 L 29 Corrected

p. 15, Lines 26-29: I'm not sure that I understand how a year-round temperature bias (+ for CFSR and – for JRA-55) impacts the amplitude of the annual cycle. Perhaps this could be clarified.

Response: Clarification will be added.

Changes in Manuscript: P 19 L 9-11 The CFSR warm bias and the JRA-55 cold bias is seasonal not year long as the original text states. The timing of these biases dictate the amplitude response.

p. 16, Lines 3-8: Why would a sudden stratospheric warming increase the amplitude of the annual cycle in the Northern Hemisphere but decrease it in the Southern Hemi- sphere (2002)?

Response: The cold temp anomalies following the warming in the upper stratosphere enlarge the winter_min/summer_max difference. The 2002 SH warming (which occurred in SH spring) did not have cold air following the warming. In fact the winter time temps were warmer than normal thus decreasing the annual amplitude for that year. This will be added to the text.

Changes in Manuscript: P 19 L 16-21 above explanation is added to the text.

p. 18, Line 13: 0.5 K

Response: Will check and correct text as needed

Changes in Manuscript: P 23 L 15 Corrected

Figs. 4-5: The authors need to more clearly describe what they are plotting in these figures. The standard deviation of 3 data sets seems somewhat of an unusual metric, as standard deviation is typically used for larger sample sizes than 3. It might be clearer to simply show the difference between the maximum value of the 3 reanalyses and minimum value of the 3 reanalyses at each month/latitude/pressure.

Response: Keeping St Dev just as an index of the degree of disagreement.

Changes in Manuscript: P 9 L 11-12, P 10 L 19-20 Added text explaining the figures. St Dev are retained.

Fig. 9: It might be helpful to mark the QBO phases somehow on these figures. Other- wise, it is extremely difficult to see what the authors are discussing in section 4.2.3.

Response: Understood. But difficult to add that to these plots.

Changes in Manuscript: Figure 9 is already very busy. It would be good to see during which phase these differences are occurring. The figure is left as is.

Fig. 10: Pressure axis needs to be labeled.

Response: Will do

Changes in Manuscript: Figure 10 'Pressure (hPa)' is added to each plots x axis.

Fig. 15: It would useful to give the approximate altitude/pressure ranges for the TLS, SSU1, and SSU2 weighting functions, as some readers may not be familiar with them.

Response: Agreed, these will be included in text.

Changes in Manuscript: P 22 L 29-30 Table 2 is added to describe the TLS, SSU1, SSU2, and SSU3 weighting function characteristics.

Response: We thank Referee #2 for their valuable comments and suggestions.

Reviewer_2: The manuscript describes the results of intercomparisons of the zonal mean temperature and zonal winds obtained by assimilation at different centers as a part of the SPARC-Reanalysis Intercomparison Project (S-RIP). The focus of the comparison is the middle atmosphere below 1 hPa during the period of satellite observations (1979– 2014). The comparison is mainly of the reanalyses produced by the different centers. A large discontinuity is found corresponding to the change in the NOAA observing system from TOVS to ATOVS. During the ATOVS period (1999–2014), the agreement among recent reanalyses largely improves. Comparisons against independent satellite observations such as HIRDLS temperature have also been conducted as supplementary analysis. Reanalysis zonal winds and temperature are widely used in different areas of atmospheric science research. Therefore, the paper merits publication in Atmospheric Chemistry and Physics once the general remarks below have been addressed.

Reviewer_2: General remarks. The introduction of the present paper states that the goal of this project is to better understand the differences between current reanalysis products and their underlying causes. However, the present paper mainly describes the differences but there is little discussion of their causes. To understand the differences in temperature and zonal winds, it is also necessary to analyze the difference in forcings such as radiative heating rate, and resolved and unresolved momentum forcings. For instance, MERRA-2 is a unique assimilation system making use of non-orographic gravity wave parameterization. Such an effect on the equatorial zonal winds should be detectable by analyzing momentum forcing.

Response: Investigating the difference in radiative and momentum forcings is beyond the scope of this paper. This has been and will be addressed in additional S-RIP papers more focused on these forcings.

Equatorial zonal winds in MERRA-2 and other reanalyses have been discussed in Kawatani et al. (2016).

Heat budgets of the tropical upper troposphere and lower stratosphere have been extensively discussed in Wright, J. S. and Fueglistaler, S.: Large differences in reanalyses of diabatic heating in the tropical upper troposphere and lower stratosphere, Atmos. Chem. Phys., 13, 9565–9576, doi: 10.5194/acp-13-9565-2013, 2013.

Changes in Manuscript: P 28 L 6-10 No additional text added discussing forcings. Additional S-RIP foci description is added to the Summary section. No additional text is added discussing heat budget.

Reviewer_2: Also, in the comparison using the reanalysis ensemble of just three members, it is not possible to detect errors or deficiencies that are common to the three members. Temperatures are compared with HIRDLS, but this is not sufficient. The assimilation increment is an important measure of the quality of assimilated products, and should therefore be analyzed and compared.

Response: Agreed. Assimilation increments are not readily available for all reanalyses and will need help from the reanalysis centers to acquire. Thus examination of the increments will have to be addressed in a future paper/document.

Changes in Manuscript: No additional text is added concerning the examination of increments.

Reviewer_2: The large differences between the periods of observation by TOVS and ATOVS are repeatedly mentioned, but the reason for the improvement is not discussed. Is this due to increased vertical resolution resulting from the increased number of channels for ATOVS, or something else?

Response: Additional text will be devoted to explaining why the ATOVS period is superior to the TOVS period.

Changes in Manuscript: P 26 L 23-28 Text is added explaining why the ATOVS period is better than the TOVS period.

Reviewer_2: It would more convenient to display and discuss the results such as in Fig. 1 grouped by family similarly to the companion paper by Fujiwara et al. (2017): "Introduction to the SPARC Reanalysis Intercomparison Project". ECMWF reanalyses: (a) ERA- I, (b) ERA40, JMA reanalyses: (c) JRA55, (d) JRA25, NASA GMAO reanalyses: (e) MERRA-2, (f) MERRA, NOAA/NCEP and related reanalyses: (g) CFSR, (h) R1, (i) 20CR.

Response: We will add a section briefly highlighting the improvements from the older version to newer reanalysis. The common improvements are to the radiative transfer model (RTM) to both the forecast model and the assimilation step, model horizontal and vertical resolution, and bias correction. Fig 1 is rearranged to compare older with newer reanalyses.

Changes in Manuscript: Figure 1 is grouped such that that each center's newer reanalysis is on the left column and the older version is in the right column.

P 5 L 1-9 Text and Table 1 is added discussing the changes made from the older reanalyses to the newer version grouped by center.

Reviewer_2: Minor comments. 1) Page 2, line 23 Typo: Fiorina -> Fiorino

Response: Will do

Changes in Manuscript: P 2 L 24 Corrected

2) Page 7, lines 25–27: "The easterly SAO phase is believed to result from .... " Please add the reference.

Response: Will do

Changes in Manuscript: P 8 L 29 Added reference

3) Page 8, line 1 Agreement does not necessarily mean good performance, if they have common errors.

Response: True and will be noted.

Changes in Manuscript: P 9 L 9-11 Text is added stating this point

4) Page 8, line 13: "the 0.5 K contour occurs moves upward from between 20 and 10 hPa to between 7 and 5 hPa." After 2001, there is no reduction of differences in the upper stratosphere. Is there some explanation for this?

Response: Most likely due to the assimilation of AIRS data. And will be noted in the text.

Changes in Manuscript: P 9 L 18-25 Additional text is added giving more explanation of this figure. We believe the addition of AIRS and IASI accounts for the change of the 0.5 K contour to rise from 10-20 hPa to 5-7 hPa.

5) Page 8, line 15: "The disagreement between the three reanalyses is greater in June- August" Why? Is this due to the dynamical heating difference?

Response: This is when the cold biased JRA55 disagrees the greatest with MERRA and ERA-I. See Fig 7. And will be noted in the text.

Changes in Manuscript: P 9 L 24-25 We found that the ERA-I is warmer at 5 hPa during the winter months than MERRA and JRA-55 during the TOVS period.

6) Page 8, line 20: "disagreement between the three reanalyses in determining this temperature during the TOVS period than during the ATOVS period" Is this because ATOVS has better vertical resolution around the tropopause region?

Response: Correct and will be noted in the text.

Changes in Manuscript: P 9 L 29-30, P 10 L 1-2 Correct. We add text explaining that there are more channels in ATOVS with sharper weighting functions, thus able to isolate the tropopause better.

7) Page 10, lines 4–5: "The left hand column shows the gross monthly mean differences ...." Because there are large differences between the TOVS to ATOVS periods, it would be better to plot the monthly difference for two different periods before and after 1998.

Response: That would be nice, but some of the reanalyses do not have a transition in 1998.

Changes in Manuscript: P 12 L 2-3 the text is only changed to address other reviewer's comments.

8) Page 10, lines 10–12: "In general, the earlier reanalyses (JRA-25, ERA-40, and R-1) show greater differences from the REM than the more recent reanalyses (CFSR, ERA-I, JRA-55, MERRA, and MERRA-2)" This may be true, but because the reanalysis ensemble mean (REM) is calculated from ERA-I, JRA55 and MERRA, the difference from REM should be smaller for these three reanalysis products.

Response: The REM members have great agreement from 10-1000hPa. Older reanalyses differ from the REM (and hence each of its members) in this pressure range because their analyses are inferior. One purpose of these figures is to show how improvements in the newer reanalysis systems improve their analysis making them more consistent with each other and hence smaller differences.

Changes in Manuscript: P 12 L 11-12 The text was not changed.

9) Page 13, line 24: "westerlies in MERRA-2 are more than 10 m/s stronger than those in the REM westerlies." Is this because of the non-orographic gravity wave parametrization introduced in MERRA-2? Comparison of the momentum forcing may clarify the cause.

Response: Momentum forcing has been addressed in Kawatani et al. (2016) and will be addressed in future S-RIP related papers.

Changes in Manuscript: P 16 L 15-16 Added text that Kawatani et al. (2016) and Molod et al. (2015) found the MERRA-2 non-orographic gravity wave parameterization to be too aggressive thus making the SAO winds too strong.

10) Page 14, line 18: "linear slope of their matched monthly winds" This phrase is not easy to understand. Is this the linear regression line between observed and assimilated QBO winds?

Response: Yes. That would be a clearer way of explaining what that plot represents.

Changes in Manuscript: P 17 L 27-28 Changed text to the reviewers suggestion.

11) Page 16, line 1 : " the fact that the CFSR did not bias correct the SSU Channel 3" Please give information about the bias corrections introduced in ERA, JRA and MERRA. JRA55 has an apparently smaller gap between the TOVS and ATOVS periods. Is this due to bias corrections applied for JRA55?

Response: More text will be added to discuss what each center did to transition from the TOVS to ATOVS observations and which channels were not bias corrected and why.

Changes in Manuscript: P 6 L 1-30 Added text describing what was done for each reanalysis during the transition from TOVS to ATOVS observations.

12) Page 16, lines 19–20: " differences are generally similar to those of the reanalyses from the REM (Figure 6)." Because the comparison so far has been made against REM, it would be useful first to show the difference between REM and HIRDLS.

Response: Agreed. This would show how well the REM represents the mean reanalysis conditions.

Changes in Manuscript: P 20 L 5 Due to time and space constraints, we decided that generating the REM-HIRDLS plot was not necessary.

Response: We thank Referee #3 for valuable comments and suggestions.

Reviewer_3: This paper provides a comparison of climatological aspects of the stratospheric temperature and wind fields produced by multiple reanalyses, covering the current and previous generations of product from the major providers. It is in general quite clearly written, and the discussion of results is even-handed in its treatment of the various reanalyses. It and companion papers from the S-RIP are likely to provide useful points of reference for users of reanalyses and some new information for producers of reanalyses, even though the latter may be quite well versed in the characteristics of their own products and more focused on the performance of new systems that are either in production or being prepared for production. The paper merits publication, although it should at least be amended to address several issues noted below that can be easily fixed. The authors should also consider whether it is feasible to make any more substantial change in response to the comments below.

Reviewer_3: The paper lacks emphasis on the lower stratosphere, and related to this is a lack of demonstration of the benefits of assimilating GPSRO data. Many of the figures are dominated by the large differences in the upper stratosphere, but the smaller differences found lower down may be of importance for those undertaking studies of the UTLS region. Here differences of a few tenths of a Kelvin may be significant, and it has been found that assimilation of GPSRO data brings the reanalyses that employ it into much closer agreement. Of the three contributors to the authors' REM, ERA-Interim and JRA-55 do assimilate GPSRO data and MERRA does not. This has proved instructive, for example, in identifying reanalysis bias in tropical tropopause temperatures (as discussed in GCOS Publication 195, page 235). Further comments are given below in the discussion of section 6.2; note in particular the final comment on this section. Perhaps contour intervals that provide more detail where differences are small could be used, or the authors could consider using times series plots for a few selected levels as well as (or instead of) the pressure/time cross-sections.

Response: Comparisons with COSMIC GPSRO temperatures throughout the upper troposphere, lower and middle stratosphere will be generated and a section will be added with supporting figures.

Changes in Manuscript: P 21 L 11-31, P 22 L 1-13 A new section (7.2) was added. COSMIC dry temperatures were compared with reanalyses. Plots of differences were generated (Figures 15, 16, and 17) and evaluated.

Reviewer_3: What happens during the transition from the TOVS to the ATOVS systems needs to be properly explained. The existing text is wrong. It is incorrect to state that only some reanalyses "merged the observations . . . over a period of time [Page 5]" and that others "switched immediately from the TOVS to the ATOVS observations [Page 19]". ERA-Interim is one of the reanalyses with a sharp discontinuity during the changeover from TOVS to ATOVS, but it is due (as documented in several of the peer-reviewed papers on this reanalysis) to a change in the anchoring of the bias correction; ATOVS observations were assimilated from the time they first became available and TOVS observations were assimilated for as long as they were available in ERA-Interim. The discontinuity in ERA-Interim is indeed related to the shift from TOVS to ATOVS, but arises simply from changing the anchoring high-sounding channel for which bias-correction was not applied from SSU-3 to AMSU-A14. More generally, the assimilating models and the data from the high-sounding SSU and AMSU channels all have biases in the upper stratosphere, and the differences in structure functions between SSU and AMSU instruments complicate the picture. It is a continuing challenge to deal with these biases in the absence of independent data that reach above the levels reached by radiosondes. GPSRO data help a little bit, but do not reach much higher, and are not available in substantial numbers prior to late 2006.

Response: Thank you for pointing out this egregious mistake. More text will be added discussing how each reanalysis center transitioned from the TOVS period to the ATOVS period. Additional text will be added discussing which channels (SSU and AMSU) have and have not been bias corrected. The new section devoted to GPSRO comparisons should show how the temperature disagreement of the reanalyses that assimilate the GPSRO data decreases after 2006 at pressures higher than 10 hPa. But at pressures less than 10 hPa the disagreement in temperatures does not change.

Changes in Manuscript: P 6 L 1-30 Text was added describing how each reanalysis dealt with the transition from TOVS to ATOVS.

P 24 L 12-25 TLS temperatures are compared and shown to merge in agreement after 2006 when COSMIC observations were assimilated. SSU1 and SSU2 temperatures were examined (not shown) and found that the temperature biases above 10 hPa prevents merging of the reanalyses SSU1 and SSU2 layer temperatures.

Reviewer_3: The paper does not discuss synoptic characteristics of the temperature and wind reanalyses. I assume that these are being taken care of in other S-RIP papers, and that this will be obvious to the reader of this paper. Some cross-referencing might be helpful, however.

Response: In the summary we will add a note that forthcoming S-RIP related papers will discuss in greater detail the circulation patterns, waves, fluxes, processes, and interactions pertaining to this part of the atmosphere.

Changes in Manuscript: P 28 L 6-10 Text is added discussing the additional S-RIP foci and what stratospheric processes they will be addressing.

Reviewer_3: Page 1, lines 27 and 28. It is noted that the zonal winds are in greater agreement than temperatures, and that this agreement extends to lower pressures. This could be a consequence firstly of the temperature differences being predominantly in the global average, or at least of large horizontal scale, which is what one would expect if they are predominantly due to differences in radiance bias correction. Thermal wind balance would then suggest only small differences in wind fields. Moreover, if the temperature differences are of broad vertical scale, the associated thermal wind differences would be expected to be larger higher up, as the thermal wind is the vertical integral of the temperature gradient. Is this too simplistic an explanation of the authors' finding?

Response: Or that it is not the actual temperatures but their gradients that reflect in the smaller wind differences.

Changes in Manuscript: P 1 L 27-28 The reviewer's comment is well noted and appreciated. These ideas would have to be evaluated to be sure what the reason is that the zonal winds were not as affected as the temperatures. Consequently, we did not add any text to make such a suggestion.

Reviewer_3: Page 2, line 23. "Fiorina" should be "Fiorino".

Response: Will do

Changes in Manuscript: P 2 L 24 Corrected

Reviewe_3r: Page 4, lines 13-16. Is it appropriate to consider CFSR to be a single reanalysis? The fundamental idea of a reanalysis is that it uses a fixed data assimilation system to analyse past observations. The upgrading of the resolution of the assimilating model and of the data assimilation in 2011 seems contrary to this fundamental idea. What is the justification for treating CFSR as a single reanalysis? Would it be better to show CFSR results only until 2010? The CFSR results in Fig.1 and (for SSU) Fig. 10 suggest a change in behaviour of CFSR in the upper stratosphere around the end of 2010.

Response: Agreed. It will be noted in the text that 'CFSR' does include the CDAS_574 analyses and that the change does impact anomalies post 2010.

Changes in Manuscript: P 4 L 15-18 The original text pointed out this change. We add that Fujiwara et al. (2017) makes a distinction between the CFSR and the CDAS-574.

Reviewer_3: Page 6, line 6. "Sceptically" is not the word I would use. Reanalysis data have a part to play along with other sources of information for the study of climate trends, but that use has to be careful, selective, and guided by studies such as the one under review. For some purposes, discontinuities can be corrected for, as has been done for ERA-Interim in published work.

Response: The purpose of this paper and other works by the S-RIP group is to inform the users of the good and bad of the various reanalyses in the middle atmosphere. We hope that from this information that inappropriate usage of reanalyses will be prevented.

Changes in Manuscript: P 7 L 6-7 Replaced 'skeptically' with 'carefully'. As in our response we want users to be aware of potential issues using any reanalysis and how those issues can impact their research.

Reviewer_3: Page 6, line 19. I would write that the 1000hPa level is "under" rather than "over" the Antarctic and Tibetan plateau.

Response: Will do

Changes in Manuscript: P 7 L 19 Replace 'over' with 'under'.

Reviewer_3: Page 9, lines 21 and 22. The MERRA-2 model is stated to be the only of the assimilating models capable of generating a QBO on its own, but MERRA-2 has the poorest agreement with the Singapore radiosonde data at 10hPa for the 1980-1998 period. Why? Also, the remark that MERRA-2 winds are greatly improved versus those of MERRA after 2000 does not hold for the Singapore winds at 10hPa, as these are fitted better by MERRA than MERRA-2. Furthermore, it is stated earlier in the paragraph that zonal winds in the tropics are not well constrained by the assimilated radiances, yet the MERRA-2 winds are apparently much better in the ATOVS period than the TOVS period. This needs further discussion.

Response: Kawatani et al. (2016) point out that in MERRA-2 the zonal winds at 10 hPa "exhibit spurious variations…during the easterly phase of the QBO. The downward propagation of the westerly SAO phases is enhanced during these periods, which could be caused by overly strong gravity wave forcing (Molod et al., 2015)." We will make note of this observation and add it in our discussion.

Changes in Manuscript: P 16 L 15-16, P 18 L 4-5, We interject our response into the text appropriately to explain why the MERRA-2 10 hPa winds are much stronger than observed.

Reviewer_3: Page 10. Section 4.1 is introductory, but refers in the beginning to a figure relating to SH polar latitudes. This text would seem more appropriately located at the start of section 4.1.1. Results are shown only for polar and tropical latitudes. What about middle latitudes? Extra figures may not be necessary, but the situation in middle latitudes should at least be summarized in the text.

Response: Mid-latitude sections will be added to the text without supporting figures.

Changes in Manuscript: P 12 L 1-2 Text is modified to introduce the characteristics of the figures not to get into specifics.

Changes in Manuscript: Mid-latitude paragraphs are added to:

Section 4.3.1 P 10 L 11-17

Section 4.3.2 P 11 L 13-19

Section 5.1.4 P 15 L 19-23
Section 5.2.4 P 17 L 5-16

Reviewer_3: Page 14. Is it sufficient to compare the QBO winds only with the radiosonde data from Singapore? Are the authors confident that this gives a representative picture of the equatorial region as a whole? Did Kawatani et al. (2016) find much variation from one radiosonde station to another?

Response: Although there are 14 radiosonde stations in the tropics with good observation records, Singapore is the only site with more than 80% available at 10 hPa. They found that during any period when radiosonde data was not available at a site that the St.Dev. of the reanalyses increased. They also found that the reanalyses have greater disagreement in the tropical regions void of radiosonde. Because of the void of observations in extensive longitude zones the equatorial zonal mean zonal wind is not as good a comparison as individual radiosonde sites.

Changes in Manuscript: P 17 L 18-31 Text is added explaining how evaluations performed by Kawatani et al. (2016) guided us to use only Singapore for comparisons. We make note that Kawatani et al. (2016) also show that there is longitudinal variations in the zonal winds.

Reviewer_3: Page 16. The comparison with HIRDLS data is good to see, but again a summary should be given of what is found for middle latitudes.

Response: Will do

Changes in Manuscript: The previous mid-latitude evaluations did not have significant features to remark about. Hence, we felt that mid-latitude evaluations with HIRDLS would not provide any additional information.

Reviewer_3: Page 17. Several comments relate to section 6.2:

(i) Reviewer: The start of the section could be clearer. The opening sentence should include reference to the AMSU-A and ATMS channels that provide temperature observations in the middle and upper stratosphere, and give links to where their weighting functions can be found. SSU observations cease in 2006, so some of what is shown in Figure 15 must not be SSU observations, but rather the observational equivalents from AMSU-A and ATMS data. This should be clarified.

Response: Additional text addressing this point will be added. Yes, NESDIS/STAR extends the MSU-4 record with the channel 9 of AMSU-A and channel 10 of the ATMS and extends the SSU channels 1,2, and 3 records with combinations of the AMSU-A and ATMS. See: Zou, C.-Z., H. Qian, W. Wang, L. Wang, and C. Long (2014), Recalibration and merging of SSU observations for stratospheric temperature trend studies, J. Geophys. Res. Atmos., 119, 13,180–13,205, doi:10.1002/2014JD021603.

Changes in Manuscript: P 22 L 16-18 The original text had a sentence pointing out that Zou and Qian (2016) explain how the AMSU channels were used to replicate the SSU channels. This sentence was move to follow the first sentence. We state that the weighting functions can be found in Fujiwara et al. (2017) and Seidel et al. (2016) and NOAA/STAR web site. We add Table 2 to provide weighting function information for the TLS, and SSU channels.

(ii) Reviewer_3: Why are SSU3 data and their equivalents not included in Figure 15? They are assimilated by the reanalyses.

Response: Not all reanalysis centers have model levels high (altitude) enough (and post process to pressures) to adequately represent the SSU Channel 3. Hence, it is not included. Barely enough post processed levels exist to represent SSU Channel 2. Text has been added providing the pressure levels at which 50% and 90% of the greatest weight occur.

Changes in Manuscript: P 23 L 1-3 We state that too much of the SSU3 weighting function is above 1 hPa, which is the lowest pressure most reanalyses output. Due to this, a proper recreation of the SSU3 cannot be generated and is not presented.

(iii) Reviewer_3: The ERA-Interim-focused paper by Simmons et al. (2014) is referred to earlier, but the reference could be repeated here. Simmons et al. included results on how well ERA-Interim fitted HIRS and AIRS data, two types of stratospheric temperature data that are not mentioned in section 6.6.

Response: Will Do and will note these results.

Changes in Manuscript: P 22 L 25-26 We add the reference to Simmons et al. (2014).

(iv) Reviewer_3: The reference to Seidel et al. (2016) should be checked, and amended if necessary. I believe it did not consider SSU trends for 1979-2015, as it did not use the AMSU-A and ATMS extensions of the SSU data record.

Response: Will check and will modify the text as needed.

Changes in Manuscript: P 22 L 26 Checked Seidel et al. (2016) and found that they just discuss the TLS layer for the 1979-2015 period.

(v) Reviewer_3: It is stated (lines 18 and 19 on page 18) that the reanalyses begin to disagree more with each other after GPSRO data become available in 2006. It must be made clear that this statement refers to anomalies not the actual temperatures. Assimilating GPSRO data brings the reanalyses closer to each other in the near-tropopause region and the lower stratosphere. The reference period used for the calculation of the climatologies on which the anomalies are based includes only a few years during which GPSRO data are assimilated, and the differences in anomalies shown after 2006 are mainly an indication of the differences in climatologies for those reanalyses that assimilate GPSRO data.

Response: Agreed, this text needs to be revised. A climatology is based upon a variable which is consistent throughout a time period (eg 30 years). This consistency is not true for any extensive period of any reanalysis especially in the middle atmosphere. An additional figure has been added showing the actual TLS temperatures which shows that after 2006 there is much greater agreement. And as mentioned earlier a section with figures has been added comparing GPSRO temperatures with the more recent reanalyses.

Changes in Manuscript: P 24 L 12-25 Figure 19 is added to show that the TLS temperatures are merging to agreement during the period in which GPSRO data are assimilated. Text is added discussing the pearls associated with using too long of a climatology with reanalyses as the quality and characteristics of reanalyses change over time thus making a climatology not really representative.

(vi) Reviewer_3: The caption to Fig. 15 states that the anomalies are with respect to 1980-2010 climatologies. If this is a typo, it should be corrected. Otherwise the figure should be changed to use the standard 30-year period 1981-2010.

Response: Will be corrected

Changes in Manuscript: P 38 L 16 Figure 18 caption is corrected.

(vii) Reviewer_3: The SSU curves for MERRA-2 in Figure 15 show a substantial spike around the beginning of 1996. This needs some discussion.

Response: The spike occurs due to fewer SSU observations assimilated during that short time period. This allows the model bias to impact the temperatures at 2 hPa. Text will be added discussing this spike.

Changes in Manuscript: P 23 L 27-30, P 24 L 1 Inquired about these spikes with folks at the GMAO. The reasons are added to the text.

(viii) Reviewer_3: The (probably erroneous) trend in ERA-Interim after 2006 noted in the last line of page 18 and the first line of page 19 was noted by Simmons et al. (2014), who identified the likely cause – an increasing use of radiosonde data that had not been bias-corrected.

Response: Text will be added citing the Simmons et al finding.

Changes in Manuscript: P 24 L8-9 Text added with the reviewers comment.

(ix) Reviewer_3: One of the reasons one does reanalysis is to produce height-resolved estimates of the meteorological fields, drawing on the better vertical resolution of radiosonde and GPSRO data than is provided by satellite radiances. A reanalysis whose bias does not change sign near the tropopause may well fit the layer average TLS dataset more poorly than a reanalysis whose bias changes sign close to the tropopause, but may be no worse or even better than the other reanalysis when it comes to comparisons at discrete pressure levels. This appears to be what is happening here: as noted in the paper under review, it is indeed most noticeable that ERA-Interim does not fit the TLS dataset well, but the evidence discussed at length in Simmons et al. (2014), supplemented by more recent results such as that on GPSRO discussed in the second paragraph of this review, does not point to a specific problem for ERA-Interim in its representation of lower stratospheric temperatures, but does indicate that the temperature differences between ERA-Interim and either JRA-55 or MERRA change sign between 70 and 100hPa. The discussion of the TLS data fits needs to acknowledge the above in order not to give a misleading impression of the capabilities of different reanalyses for representing temperature trends and variations at lower stratospheric levels such as 70 or 50hPa. The authors should also consider calculating and including in the paper time series of the fits at various levels of the reanalyses to a bias-adjusted radiosonde dataset, such as one of those produced by Haimberger.

Response: It will be noted that at the critical levels of 70 and 100 hPa (for the TLS) that the reanalyses differ from each other due to various causes. A figure showing the actual temps at 70 and 100 hPa and supporting text may be added to illustrate this. However, no additional comparisons with bias adjusted radiosondes will be performed.

Changes in Manuscript: P 24 L 13-25 The addition of Figure 19 addresses much of the reviewer's comment. The merging of the reanalyses toward a similar TLS temperature shows that improvements are being made to the reanalyses. An investigation of each reanalyses temperature near the tropopause will be the subject of another S-RIP report. No additional text is added addressing this comment.

Reviewer_3: Page 19, lines 9 and 10. It would probably be better to refer to "changes in the biases of the data from the TOVS/SSU instruments" rather than "changes in the TOVS/SSU instruments." The main differences in bias from one instrument to another stem from differences in cell pressure, and from the gaseous composition of the cell, linked to the behavior of the cell seals, rather than from changes in the basic design of the instrument.

Response: Text will be modified.

Changes in Manuscript: P 24 L 30 The text is changed to the reviewer’s suggestion.

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
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)

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 both observed by conventional systems, cloud tracked, and derived from height fields which in turn are 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 ~~between~~ 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 degree zonal mean data sets and spans the satellite era from 1979 – 2014. There is very good agreement in temperature seasonally and latitudinally between 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 between 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 between NOAA's 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 a. nd anomalies from a long climatology period as the quality and character of reanalyses may have changed over time.

1 Introduction

Reanalyses are used in many ways: as initial conditions for historical model runs, developing climatologies, comparison with experimental models, examination of atmospheric features or conditions over long periods of time, etc. This paper evaluates mainly eight reanalysis data sets: NCEP/NCAR Reanalysis 1(Kalnay et al.,1996, and Kistler et al.,2001), (referred to hereafter as "R-1"; see Appendix A for abbreviations), ERA-40 (Uppala et al.,2005), JRA-25 (Onoagi et al., 2007), NCEP/CFSR (Saha et al., 2010), ERA-Interim (Dee et al., 2011) (referred to hereafter as ERA-I), MERRA (Reinecher et al, 2011), JRA-55 (Kobayaoshi et al., 2015), and MERRA-2 (Bosilovich Gelaro et al., 2016)7, GMAO, 2015), with some notes on NCEP/DOE Reanalysis 2 (Kanamitsu et al.,2002) (referred to hereafter as "R-2") 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, observational data assimilated, etc. In this paper, we present an intercomparison of these ten reanalyses focusing mainly upon their temperature and zonal wind fields, though the mean meridional winds and vertical velocities are also briefly discussed. 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 1990's1990s. Pawson and Fiorinoa (1998a, 1998b, 1999) were the first to evaluate reanalyses winds and temperatures comparing R-1 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 R-1, ERA-15, ERA-40, along with meteorological centers' analyses. R-1 and the ERA-40 have been used by thousands of researchers for tropospheric studies. Notable middle-atmosphere studies evaluating R-1 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 (SSW). Martineau and Son (2010) used temperature and wind fields from R-1, R-2, 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, JRA-55 stratospheric temperature analyses over the 1979-2012 period showing where and when they agreed and disagreed, and 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-lower mesosphere. Lawrence et al. (2014) 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. (2015) intercompared 6 reanalyses (R-1, 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 9 reanalyses (JRA-55, MERRA, ERA-I, CFSR, JRA-25, ERA-40, R-1, R-2, 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) intercompares eight reanalyses (ERA-40, ERA-I, R-1, R-2, 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 the focus on the Quasi-Biennial Oscillation (QBO; Baldwin et al., 2001) among nine reanalyses (R-1, R-2, 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 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, Free University of Berlin's Northern Hemisphere subjective analysis, CIRA86 (COSPAR International Reference Atmosphere, 1986), R-1, ERA-15 and ERA-40. This intercomparison was mostly based upon analyses rather than reanalyses, as only the R-1, ERA-15 and ERA-40 reanalyses were available at that time. Notable differences were found between 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 between 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 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 that are systems that assimilate surface and upper-air conventional and satellite data (i.e., MERRA-2~~CFSR~~, MERRA, ERA-I, JRA-55, ~~MERRA-2~~CFSR, JRA-25, ERA-40, R-1), 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 R-2 is an updated version of the R-1. Almost all of the changes and enhancements incorporated into the R-2 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 R-1 and R-2 show very minor differences in temperatures and winds above the boundary layer. Therefore, we will not show R-2 comparisons, but one can expect all R-2 qualities to be nearly exactly the same as R-1. All of the reanalyses except for the CFSR used the same forecast model and assimilation scheme throughout their time span. In 2010 the CFSR made an undocumented update to the GSI assimilation scheme, then again in 2011, with the implementation of the version 2 Climate Forecast System (CFSv2, Saha et al., 2014), the resolution, forecast model and assimilation scheme all were upgraded in the CFSR. In Fujiwara et al. (2017) they distinguish this latter analysis as the CDAS-T574.

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

We characterize the stratosphere in 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 R-1/R-2). 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 latter versions used a more recent version of the RTM. Explanations of the various labeling of horizontal resolution can be found in Fujiwara et al., (2017).

3~~2~~ 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/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. Also the radiative transfer models used in the forecast models have improved over time. Reanalysis centers devote major efforts to minimize 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 centers 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 one 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, Figure 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 ~~Nino~~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 of the NOAA polar satellites with the SSU instrument in the early ~~1980's~~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 ~~1980's~~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 the 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 ~~from the immediate transition from TOVS~~when ~~to~~ ATOVS observing systems began to be assimilated. ERA-I assimilated both SSU and AMSU-A radiances until 2005. Channel 3 of the SSU previous 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, 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 leaks and changes in the stratospheric CO_2 concentration (Gelaro et al., 2017). ~~They~~ 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 one year and bias corrected all the SSU and AMSU-A channels (Kobayashi et al., 2014). 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 ~~and then shows another transition~~ 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). R-1, R-2, and the 20CR reanalyses only extend up to 10 hPa, due to their fewer number of model layers, so the upper stratosphere is not analyzed. R-1 and R-2 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 inter-annual 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 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 observations, AIRS, and GPSRO observations (MERRA did not assimilate GPSRO data.)~~.~~

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

4~~3~~ Reanalysis Ensemble Mean (REM)

4~~3~~.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 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 ~~over~~ under the Antarctica and 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 kilometers pressures higher than 700 hPa have invalid data and hence are not analyzed.

4~~3~~.2 Climatology of the REM

4~~3~~.2.1 Temperature

The seasonal variation of 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 Figure 2. Of note is that at polar latitudes the lowest temperatures occur in the upper stratosphere in November (for Northern Hemisphere, NH) and May (for 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 ~~more~~ greater wave activity transporting heat from the extra-tropics into the polar latitudes, thus shortening the period of low temperatures~~ in the polar latitudes~~. In the tropics, the variability is much smaller than in the polar regions, but is associated with the phase of the SAO and~~ with~~ the QBO in the upper ~~middle~~ and middle ~~upper~~ stratosphere, respectively.

4~~3~~.2.2 Zonal Wind

The seasonal variation of 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 Figure 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 stronger than the winter time 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 with the temperature variability, the variability of the SH polar night jet between May through August is not as great as in the NH polar jet. The SH zonal wind variability does increase 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 at equinox and intervening easterlies during the solstice periods. There is a marked asymmetry in amplitude of the easterly SAO phase, with amplitudes of -40 to -50 m/s easterlies in December-to-February but only -20 to -30 m/s in July-September. The easterly SAO phase is believed to result from 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 midstratosphere where the QBO dominates, the climatological winds in the tropical middle stratosphere have mean easterlies of -5 to -10 m/s. 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.

43.3 Agreement Among the REM Members

43.3.1 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 standard deviation 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 standard deviations. 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 standard deviation (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 versus pressure plot. The mid-latitudes 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 where 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 20 hPa level all three reanalyses agree with each other very well, with a standard deviation smaller than 0.5 K. Generally, as time moves forward 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 from between 20 and 10 hPa to between 7 and 5 hPa. Interrupting this upward trend is the period during the 1990's in which the NH polar activity was unusually quiet and cold. (Pawson and Naujokat, 1999, Charlton and Polvani, 2007). The increased agreement between 20 and 7 hPa is most likely due to the assimilation of AMSU and AIRS observations. The disagreement between 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 between 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 5 channels (5 through 9, 1 through

4 are water vapor channels) sounding the same layer. These additional channels provide information about the temperature structure near the tropopause. Thus allowing the reanalyses to better analyze and agree upon the temperature structure there. The pressure at which the greatest differences (3-4 K) occur is at 2 hPa and has a seasonally varying pattern.

In the 90°-60° S zone, the disagreement between 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, ~~with the change from TOVS to ATOVS periods. There is improvement in the differences after 2002~~ 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~~.3.2 Zonal Winds

Figure 5 presents pressure vs time series plots of the zonal wind standard devition (m/s) 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 between the three reanalyses in the NH and SH polar jet regions with standard deviations smaller than 0.5 m/s. In the NH polar jet region significant disagreement (>0.5 m/s) between the three reanalyses is consistently confined to pressures lower than 5 hPa. Disagreements greater than 0.5 m/s are nearly eliminated after the transition to ATOVS observations occurs at the end of~~in~~ 1998.

The altitude range of disagreement greater than 0.5 m/s 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 improve considerably in the ATOVS time period.

The tropical zonal wind disagreement shows much larger values of the order of 10 m/s 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 to the NH and SH polar jet winds after 1998 versus minor improvement of 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, whose spatial scales 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 between the reanalyses decrease slightly. None of the forecast models included in the REM are 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 versus 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. 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. During the ATOVS period the disagreements become more sporadic, smaller in value and occur at pressures lower than 7 hPa.

54 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 R-1) differs from the REM for both temperatures and winds. We do not show comparisons of R-2, but one can expect all R-2 qualities to be nearly the same as those of R-1. 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 it coarse model vertical resolution.

5~~4~~.1 Temperature

Figures 6, 7 and 8 present ~~In the left column of Figure 6~~ the time-mean ~~SH polar (90°-60° S)~~ zonal mean temperature difference from the REM (Reanalysis – REM) for each month (left columns) ~~is presented.~~, W~~w~~hile the right columns show~~s~~ the time series of the zonal mean monthly mean differences from 1979 through 2014. The left ~~hand~~ columns show~~s~~ the gross monthly mean differences while the right ~~hand~~ columns show~~s~~ 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 reanalysis and the REM are within ±0.5 K. In general the earlier reanalyses (JRA-25, ERA-40, and R-1) show greater differences from the REM than the more recent reanalyses (MERRA-2~~(CFSR~~, MERRA, ERA-I, JRA-55~~, MERRA~~, and CFSR~~MERRA-2~~). Also, in general, the NH and SH polar latitudes 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~~4~~.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.

~~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's upper stratospheric warm bias occurs throughout the entire 1979-2014 time span with similar seasonal variability~~

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 ~~1990's~~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 ~~year-round~~ 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 latter 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's upper stratospheric warm bias occurs throughout the entire 1979-2014 time span with similar seasonal variability.

~~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.~~

The JRA-25 time-mean plot shows greater differences from the REM than the above five reanalyses, with a yearlong 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 ~~warm~~ cool bias with maximum (-2 to -4 K) occurring in the August-November months. The time series plot shows that the upper stratosphere warm bias~~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 to -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 Section 5.2 of Fujiwara et al., 2017 for its reason).

The ERA-40 time-mean plot shows a strong cold bias (-2 to -6 K) compared to the REM persisting yearlong between 2 and 10 hPa. Just below this is a warm bias (2 to 4 K) between 10-30 hPa. The annual cycle of both the cold bias and warm bias show a slight rising in summer and 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.

R-1 does not analyze at pressures lower than 10 hPa, so there is no evaluation in the upper stratosphere. But 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.

54.1.2. NH Polar Latitudes

Many features in the upper stratosphere are seasonally common between the NH and SH polar latitudes (Figure 7). However, differences with the REM in the middle and lower 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. ~~The CFSR's winter time 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.~~ 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 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's winter time 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 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.~~ The JRA-25, ERA-40, and R-1 all show similar seasonal biases from the REM in the upper stratosphere. Their time series show reduced differences in the middle and lower stratosphere.

54.1.3. Equatorial Latitudes

Differences of 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's upper stratosphere warm bias of 2 to 4 K and for the JRA-55's 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 shows 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 centered at 100 hPa, and a cold bias below, persist though the ~~A~~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~~1~~ K. The JRA-55 reanalyses shows that the cold biases are nearly constant throughout the entire time series. ~~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 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~~6~~ to -8 K that is largest during the ATOVS period. Between 30 and 50 hPa, there is another layer of cold bias of -2~~4~~ to -6 K that is present only during the TOVS period. ERA-40 has a persistent cold bias of -2~~6~~ to -6~~8~~ 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. R-1 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 where 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 Section 4~~3~~.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 5-7 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~~4~~.2 Zonal Wind

5~~4~~.2.1. SH Polar Latitudes

The time-mean SH polar jet differences (see Supplement) of the individual reanalyses from the REM are relatively small, ranging from -2 to 1 m/s, with most differences smaller in magnitude than that. As presented in section 4~~3~~.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 R-1, the upper stratosphere is the region where the greatest differences from the REM is 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, R-1 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.

54.2.2. NH Polar Latitudes

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

54.2.3. Equatorial Latitudes

In Figure 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 m/s during the March-April and September-November westerly periods. They also have positive biases of greater than 3 m/s during the December-February easterly period. MERRA and ERA-I have stronger westerlies and show positive biases of greater than 3 m/s 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 m/s during the former period. The SAO westerlies in MERRA-2 are more than 10 m/s 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 R-1 show greater differences from the REM than the other reanalyses. R-1 shows a westerly bias of > 4 m/s during the easterly phase of the QBO from 10 hPa down to 100 hPa. This was discussed as well by Pawson and Fiorino (1998b). The JRA-25 has an easterly bias > 4 m/s 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-30° N and 1 to 30 hPa from July 1, 1981 to December 31, 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 Figure 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 m/s. Hence, the differences from the REM show the CFSR has a consistent layer of negative biases of -1 to -2.5 m/s from 50 to 300 hPa. The JRA-55 shows the other extreme of a consistent positive bias of 1 to 2 m/s 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 m/s from the REM in the upper stratosphere (SAO region). The time series differences 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 m/s 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 m/s during the TOVS period and is eliminated in the ATOVS period. All the other reanalyses are in good agreement (differences within ±0.5 m/s) with the REM in the lower stratosphere. In the middle stratosphere the JRA-25 has differences between -0.5 and -1 m/s 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 m/s 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~~4~~.2.5~~4~~. Comparisons with Singapore QBO Winds

Kawatani et al. (2016) provides a thorough evaluation of the RMS differences of QBO (70-10 hPa) zonal winds between 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, that 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, w~~W~~e will focus just upon comparisons between the reanalyses and zonal winds at Singapore ~~(1°N,104°E)~~. This is not to imply that Singapore is representative of the entire tropical zone, which it does not because there is longitudinal variability in the zonal mean zonal winds (Kawatani et al., 2016). Correlations between the monthly mean ~~CFSR,~~ MERRA-2, MERRA, ERA-I, JRA-55, and CFSR ~~MERRA-2~~ 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 ~~slope of their matched monthly winds~~regression line between the observed and analyzed QBO winds and their scatter. Figure 10 (a, b, and c) shows the RMS differences of the reanalyses QBO winds and that at Singapore. Comparisons are shown for the entire 1980~~79~~-2014 period and then divided into the TOVS (1980-1998) and ATOVS (1999-2014) periods. All of the reanalyses' RMS differences ~~improve~~ are smaller during the ATOVS period. All of the RMS differences increase from 70 hPa 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 having 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 Figure 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 1980’s and early 1990’s that MERRA-2 overemphasized the annual signal. Figure 10 (d, e, and f) shows 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 periods.

65 Amplitude of Polar Annual Temperature Cycle

Another way to examine the differences between 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 their annual temperature amplitude will be smaller compared to the other reanalyses. Generally, as Figure 2 shows, the summer time temperatures do not vary much from year to year, while the winter time temperatures have greater interannual variability. The mean polar temperatures in Figure 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 of 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 Figure 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 56-152 K in the troposphere to 458-6054 K in the middle stratosphere. Maximum amplitudes (6054-760 K) in the SH occur above 10 hPa. In the NH polar latitudes the minimal amplitude of 65-15-12 K occurs at the polar tropopause. Between the surface and the tropopause the temperature amplitude is larger at 15~~8~~-25~~4~~ K. Above the tropopause the temperature amplitude increases up to about 2-3~~5~~ hPa where the temperature amplitude lies in the 55~~4~~-60 K range, although the depth of this layer is not nearly as extensive as in the SH polar regions. There is good agreement between these five more recent reanalyses of 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 than is 5~~6~~-15~~2~~ K smaller than the other four reanalyses in the SH and about 5~~6~~ 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's temperature amplitudes peak at greater pressures in the upper stratosphere and then decreases 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 when a SSW did not occur (e.g. the ~~1990's~~1990s) have smaller temperature amplitudes in the upper stratosphere. In the SH years in which there was a great amount of wave activity during the winter months had warmer winters and consequently had 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~~6~~ Comparisons with Satellite Temperature Observations

7.1 HIRDLS and MLS Temperatures

~~HIRDLS and MLS Temperatures~~

The NASA Earth Observing System (EOS) Aura spacecraft was launched in July 2004 and has onboard several instruments that measure multiple atmospheric constituents. The ~~HIgh~~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 any of the reanalyses and thus are independent measurements. Monthly mean temperature differences of reanalyses from the HIRDLS (Reanalysis – HIRDLS) temperatures at NH high latitudes (60º - 80º N), tropics (10º S-10º N) and SH high latitudes (60º S) were generated for the 2005 through 2008 period. Figures 12, 13, and 14 present the differences of MERRA-2 ~~CFSR~~, MERRA, ERA-I~~, MERRA~~, JRA-55 and CFSR ~~MERRA-2~~ 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 (Figure 6). MERRA-2 has a warm bias all yearlong at 1 hPa and a -1 to -2 K cold bias from November through March. ~~The CFSR has a very warm bias of over 14 K in the April to July period at pressures lower than 5 hPa with a cold bias at 7 hPa during this same time period.~~ 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 become 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 5 hPa with a cold bias at 7 hPa during this same time period. ~~MERRA-2 has a warm bias all yearlong at 1 hPa and a -1 to -2 K cold bias from November through March.~~ All of the reanalyses show a slight (< 1 K) warm bias in the middle stratosphere during the November through March period.

In the NH, t~~T~~he cold bias of MERRA-2 in the summer period is smaller in the NH, while the yearlong warm bias exists at 1 hPa. ~~–, the CFSR and JRA-55 differences with HIRDLS occur in the same seasons as in the SH with little change in amplitude.~~ The cold bias that MERRA has in the SH does not exist in the NH. The mid-winter warm bias that was in the SH is about a degree warmer in the NH. Similarly, the ERA-I does not have a cold bias in the late winter-spring period, but there is a warm bias in mid-summer 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. ~~The cold bias of MERRA-2 in the summer period is smaller in the NH, while the yearlong warm bias exists at 1 hPa.~~ 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,~~ the CFSR has a similar warm bias in a semi-annual basis in the upper stratosphere~~ . ~~.~~ MERRA-2 continues to have a yearlong warm bias at 1 hPa and a slight warm bias near 10 hPa. In 2006-2007 MERRA has a warm bias between 2-3 hPa ~~that is highest in altitude~~ during ~~November through~~ January and February and move lower to 5 to 10 hPa during the other months of the year. ERA-I seems to have a yearlong 0.5 to 1 K warm bias at pressures lower than 10 hPa. JRA-55 has a yearlong -1 to -2 K cold bias between 5 and 2 hPa. The CFSR has a similar warm bias in a semi-annual basis in the upper stratosphere.~~MERRA-2 continues to have a yearlong warm bias at 1 hPa and a slight warm bias near 10 hPa.~~

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.

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 of MERRA-2 ~~CFSR~~, MERRA, ~~,~~ ERA-I, ~~,~~ JRA-55, and CFSR ~~MERRA-2~~ 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) was obtained from the JPL GENESIS data portal. Leroy et al. (2012) explains the technique in which the RO observations were turned into temperatures and transposed from altitude to pressure surfaces. We use this 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 higher pressures than 200 hPa as atmospheric water vapor deviates the actual temperatures from the dry temperatures. Figure 15 shows the eight year time series of differences (Reanalysis-COSMIC) between the reanalysis temperatures and the COSMIC temperature in the SH polar latitudes (90°S-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 hPa 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 differenced against (except MERRA) assimilate 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°N-90°N). Similar negative differences occur during the transition from NH summer to winter. The depth and time length of the -1 K differences is smaller than the SH differences. There are also short term negative differences that extend from 10 to 100 hPa during the years in which a SSW occurred (2009, 2010, and 2013). In 2009 this is preceded by a short term (one 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 is not capturing the maximum warming during the SSW as well as the cooling which follows. This may be due to the fewer number of COSMIC observations in the polar region versus the number of observations peaking between 50° and 60° in both hemispheres.

Differences between the reanalyses and COSMIC dry temperatures in the tropics (10S-10N) 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 do 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 at 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~~6~~.3~~2~~ Atmospheric Layer Temperature Anomalies

Long-term satellite observations from NOAA polar orbiting satellites of temperatures of the lower stratosphere (TLS) are available from the MSU 4 and AMSU-A 9 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 e.g. Fujiwara et al. (2017, their Fig. 7) and ~~Zou~~ Seidel et al. (2016, their Fig. 1~~4~~) and on the NOAA/STAR SSU website (http://www.star.nesdis.noaa.gov/smcd/emb/mscat/index.php). ~~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.~~ These satellite-observation climate data records have been used to compare with climate model runs to see if the model is accurately capturing 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, Young et al., 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 ~~and SSU channel~~ trends from three satellite centers for the entire (1979-2015) period and separate trends for pre-1997 and post-1997. Mitchell et al. (2014) generated TLS, and SSU channel weighted temperatures from reanalyses to see how well they compare with the satellite observations. We perform a similar exercise 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~~5~~. Anomalies from the 30 year period (1981-2010)~~ period~~ 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 Figure 18~~5~~ 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 of each reanalysis anomalies from the NOAA/STAR anomalies. The reanalyses TLS anomalies differ from the NOAA/STAR anomalies by less than ±0~~2~~.5 K for most of the time series. Most noticeable is that the ERA-I has smaller anomalies than NOAA/STAR in the early ~~1980's~~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) but with the exception of MERRA and MERRA-2 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. 2006 is when GPSRO observations from the COSMIC constellation became available for assimilation.

The SSU1 temperature anomalies from the CFSR shows 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 ~~1980's~~1980s. There is minor disagreement between the four reanalyses with the NOAA/STAR between the late ~~1980's~~1980s through the early 2000~~'~~s. 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 (McCarthy, personal communication). When there are no observations to constrain the model in the upper stratosphere, analyses migrate to the model's 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 between the four reanalyses increases.

Just as with the SSU1 anomalies, the large temperature jumps associated with the CFSR’s stream transitions prevents 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 ~~1990’s~~1990s and early 2000~~’~~s when its anomalies are smaller than the 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 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 ~~has~~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 latter years in Figure 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 of 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 are actually approaching agreement of the TLS values as more observations are assimilated. Figure 19b presents the standard deviation 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 Figure 18 is that the quality and character of the temperature values between 1981 and 2010 changed. This makes generating a 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 temperatures in the middle and lower stratosphere, which is where most of the TLS weighting function occurs.

8~~7~~ 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 ~~1980’s~~1980s in the stratosphere associated with changes in the biases of the data from the NOAA TOVS/SSU instruments. The R-1 and R-2 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 their profile temperatures. Almost all the reanalyses have a temporal discontinuity in 1998 when the ATOVS observations became available and the reanalyses either switched immediately ~~from the TOVS to the ATOVS observations or in the case of MERRA and JRA-55~~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 favor 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 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 mid-winter 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 show larger interannual variability in timing and magnitude.

Time series of the temperature variance of the three REM members showed that the greatest disagreement occurs during the TOVS time period (1979-1998) in all ~~three~~ 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 ~~occurred through all four seasons and~~ persisted during both the TOVS and ATOVS time periods. The JRA-55, on the other hand, had a seasonal cold bias that ~~occurred through all four seasons and~~ 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 yearlong ~~warm~~ cool bias at 1 hPa and ~~cool~~ warm bias between 2-3 hPa. These biases greatly diminished during the ATOVS period. Temperature differences from the REM in the earlier reanalyses (JRA-25, ERA-40, and R-1) 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 centers to improve the analyses from the earlier versions to the later version. This results in better agreement between the more recent reanalyses.

In the tropics, the individual reanalyses exhibited smaller temperature differences than in the polar latitudes. But 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 ~~differences~~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 parmaterization. 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 than the ATOVS period.

Several reasons exist why the ATOVS period is an improvement over the TOVS period. The primary reason is that the AMSU-A instrument has 5 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 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~~6~~-15~~2~~ K. The region of large amplitude extends throughout the middle and upper stratosphere in the SH polar latitude. 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's 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 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 – 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, and Zou and Qian, 2016) are a much improved CDR than the version used in the Thompson et al.,(2012) paper 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 generate 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 as well as 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 between different sets of observations. We also find that the larger temperature and wind differences between the older reanalyses have become smaller between 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 reanalysis 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 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 vapor in the reanalyses has recently been published (Davis et al., 2017). Future works will evaluate the Brewer-Dobson circulation, stratosphere-troposphere dynamical coupling, upper-tropospheric-lower stratospheric processes, stratospheric-tropospheric exchange in the extra-tropics 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

– 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 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’s 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 Laboratoty 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 the NOAA

NESDIS: National Environmental Satellite, Data, and Information Service of the NOAA

NH: Northern Hemisphere

NOAA: National Oceanic and Atmospheric Administration

NOAA-* (where * is number): NOAA’s polar-orbiting operational meteorological satellite, number *

QBO: quasi-biennial oscillation

R-1: NCEP-NCAR Reanalysis 1

R-2: 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 the NESDIS

TIROS: Television Infrared Observation Satellite

TLS: temperature of the lower stratosphere (MSU channel 4 / AMSU channel 9)

TOVS: TIROS Operational Vertical Sounder

WMO: World Meteorological Organization

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) ~~the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT)~~ through Grants-in-Aid for Scientific Research (26287117 and 16K05548), and Corwin Wright's by NERC grant NE/K015117/1.

References

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., Takahashi, M.: The quasi-biennial oscillation, Rev. Geophys., 39(2), 179–229, doi:10.1029/1999RG000073, 2001.

Baldwin, M. and Gray, L.J.: Tropical stratospheric zonal winds in ECMWF ERA-40 reanalysis, rocketsonde data and rawinsonde data. Geophys. Res. Lett., 32 (9), 2005

Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial evaluation of the climate, NASA Tech. Rep. Series on Global Modeling and Data Assimilation, NASA/TM–2015-104606, Vol. 43, 2015.

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, doi: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. Meteor. Soc., 137, 1–28, doi:10.1002/qj.776, 2011.

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

Cucurull L, Derber JC, Treadon R, 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., Potter, G. L., Tegtmeier, S., Wang, T., Wargan, K., and Wright, J. S.: Assessment of upper tropospheric and stratospheric water vapour and ozone in reanalyses as part of S-RIP, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-377, in review, 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. Meteor. Soc., 137, 553–597, doi: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, Atmospheric Chemistry and Physics, 15, 13507-13518, doi: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, doi: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. Clim., 30, 5419-5454, doi: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, doi:10.1029/2007JD008824, 2008.

Gibson, J. K., Kållberg, P., Uppala, S. M., Nomura, A., Hernandez, A., and Serrano, E.: ERA description. ERA-15 Rep. Series 1, 72 pp., 1997.

Global Modeling and Assimilation Office (GMAO): MERRA-2 inst3_3d_asm_Nv: 3d, 3-Hourly,Instantaneous, Model-Level, Assimilation, Assimilated Meteorological Fields V5.12.4, Greenbelt, MD, USA, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed 1 November 2015, doi:10.5067/WWQSXQ8IVFW8, 2015.

Gray, L, Pyle, J.:, Two-Dimensional Model Studies of Equatorial Dynamics and Tracer Distributions, QJRMS, 113, 635-651, doi: 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, 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.: Representation of the tropical stratospheric zonal wind in global atmospheric reanalyses, Atmospheric Chemistry and Physics, doi:10.5194/acp-2016-76, 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, doi: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, doi:10.5194/acp-15-3873-2015, 2015.

Leroy, S., C. Ao, O.P. Verkhoglyadova (2012), Mapping GPS Radio Occultation Data by Bayesian Interpolation, J. Atmosph. Oceanic Technology, 29, 1062, doi: 10.1175/JTECH-D-11-00179.1.

Livesey, N. J., et al.: 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.

Madhavi, G., Kishore, P., Rao, S., Velicogna, I., Basha, G.:Two-day wave observations over the middle and high latitudes in the NH and SH using COSMIC GPSRO measurements, Adv. Space Res.,55, 722-731, doi:10.1016/j.asr.2014.09.032, 2014.

Manney, G. L., Sabutis, J. L., Pawson, S., Santee, M. L., Naujokat, B., Swinbank, R., Gelman, M. E., and Ebisuzaki, W.: Lower stratospheric temperature differences between meteorological analyses in two cold Arctic winters and their impact on polar processing studies, J. Geophys. Res., 108(D5), 8328, doi:10.1029/2001JD001149, 2003.

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," Monthly Weather Review 133, 1261-1278, doi:10.1175/MWR2926.1, May 2005.

Manney, G.L., Hegglin, M.I., Daffer ,W.H., Schwartz, M.J., Santee, M.L., and Pawson, S., "Climatology of Upper Tropospheric/Lower Stratospheric UTLS Jets and Tropopauses in MERRA," J. Climate 27, 3248-3271, doi:10.1175/JCLI-D-13-00243.1, 2014.

Martineau, P., and Son, S.-W.: Quality of reanalysis data during stratospheric vortex weakening and intensification events, Geophys. Res. Lett., 37, L22801, doi: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, doi: 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. R. Meteorol. Soc., 141, 2011–2031, doi: 10.1002/qj.2492, 2015.

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, doi: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, doi: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., Labitzke, K., Lenschow, R., Naujokat, B., Rajewski, B., Wiesner, M. and Wohlfart, R.-C.: Climatology of the Northern Hemisphere Stratosphere Derived from Berlin Analyses. Part 1: Monthly Means. Meteorologische Abhandlung, Institut für Meteorologie der F. U. Berlin, Serie A, Band 7, Heft 3, 1993.

Pawson, S., and Naujokat, B.: The cold winters of the middle 1990s in the northern lower stratosphere, J. Geophys. Res., 104(D12), 14209–14222, doi:10.1029/1999JD900211, 1999.

Poli P, Healy SB, Dee DP.: Assimilation of Global Positioning System radio occultation data in the ECMWF ERA–Interim reanalysis. Q. J. R. Meteorol. Soc. 136: 1972–1990, doi: 10.1002/qj.722, 2010.

Randel, W., et al.: The SPARC intercomparison of middle-atmosphere climatologies, J. Climate, 17, 986–1003, 2004.

Randel, W., Smith, A.K., Wu, F., Zou, C-Z., Qian, H., Stratospheric temperature trends over 1979-2015 derived from combined SSU, MLS, and SABER satellite observations, J. Climate, 29, 4843-4858, DOI: 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(D2), 1759–1774, doi: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, doi: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, doi:10.1175/2010BAMS3001.1, 2010.

Schwartz, M. J., et al.: Validation of the Aura Microwave Limb Sounder temperature and geopotential height measurements, J. Geophys. Res., 113, D15S11, doi:10.1029/2007JD008783, 2008.

Seidel, D. J., J. Li, C. Mears, I. Moradi, J. Nash, W. J. Randel, R. Saunders, D. W. J. Thompson, and C.-Z. Zou: Stratospheric temperature changes during the satellite era, J. Geophys. Res. Atmos., 121, doi: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.R. Meteorol. Soc., 140: 329–353. doi:10.1002/qj.2317, 2014.

SPARC : SPARC intercomparison of middle atmosphere climatologies. SPARC Rep. 3, 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

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. Meteor. Soc., 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

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, doi:10.1038/nature11579, 2012

Young, P.J., Thompson, D.W.J., Rosenlof, K.H., Solomon, S., 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. Clim., 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. Journal of Climate 25:5, 1759-1772, 2012.

Zhao, L.; Xu, J.; Powell, A.M.; Jiang, Z.; Wang, D. Use of SSU/MSU Satellite Observations to Validate Upper Atmospheric Temperature Trends in CMIP5 Simulations. Remote Sens. 2016, 8, 13.

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, 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.,119,13,180-13,205, doi:10.1002/2014JD021603, 2014.

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

Figure Captions:

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

Figure 2. Annual variation of the REM temperature monthly means (°C) and their standard deviation (K) in three different zonal regions: 60° - 90° N (top); 10° S - 10° N (middle); and 90°- 60° S (bottom).

Figure 3. Annual variation of the REM zonal wind monthly means (m/s) and their standard deviation (m/s) in three different zonal regions: 40° - 80° N (top); 10° S - 10° N (middle); and 80° - 40° S (bottom).

Figure 4. Pressure vs time plots of the temperature standard deviation (K) for each month of the three reanalyses making up the REM for three zonal regions: 60°-90° N (top); 10° S-10° N (middle); and 90°-60°S (bottom).

Figure 5. Pressure vs time plots of the zonal wind standard deviation (m/s) for each month of the three reanalyses making up the REM for three zonal regions: 40° - 80° N (top); 10° S - 10° N (middle); and 80° - 40° S (bottom).

Figure 6. Pressure vs month plots (left~~(right~~) and pressure vs time plots (right~~left~~) of the temperature difference (K) of individual reanalyses from the REM for the zonal region 90° - 60° S. The reanalyses are a,i) MERRA-2~~CFSR~~, b,j) MERRA, c,k) ERA-I, d,l) JRA-55, e,m) CFSR~~MERRA-2~~, f,n) JRA-25, g,o) ERA-40, and h,p-) R-1. 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 Figure 6 but for the 60° - 90° N latitude zone.

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

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

Figure 10. RMS differences (m/s) (left~~ column~~) and linear slopes (right~~ column~~) 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 (1ºN, 104ºE) versus the observed Singapore monthly mean zonal winds from the FUB. RMS differences and slopes are computed for the 1980-2014 time period (~~top~~a,d), the 1980-1998 period (~~middle~~b,e), and the 1999-2014 period (~~bottom~~c,f). Slopes less than 1.0 indicate that the reanalysis zonal winds are weaker than the Singapore zonal winds.

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

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

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

Figure 14. Same as Figure 12 except for the Equatorial latitude (10º S – 10º N) zone.

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

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

Figure 17. Same as Figure 15 except for the Equatorial latitude (10º S – 10º N) zone.

Figure 18~~5~~. Time series plots of the global layer mean temperature anomalies (K) from the 1981~~0~~-2010 climatology (left) and reanalyses anomaly differences from the NOAA/STAR anomalies (right) for: a,d) the lower stratosphere (TLS) equivalent to the MSU 4 observations, ~~(top~~b,e)~~,~~ the middle stratosphere (SSU1) equivalent to the SSU channel 1 observations, and ~~(middle~~c,f) ~~, and~~ the upper stratosphere (SSU2) equivalent to the SSU channel 2 observations~~ (bottom)~~. TLS, SSU1 and SSU2 weights are applied to the ~~CFSR~~MERRA-2, MERRA, ERA-I, JRA-55, and CFSR~~MERRA-2~~ 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.

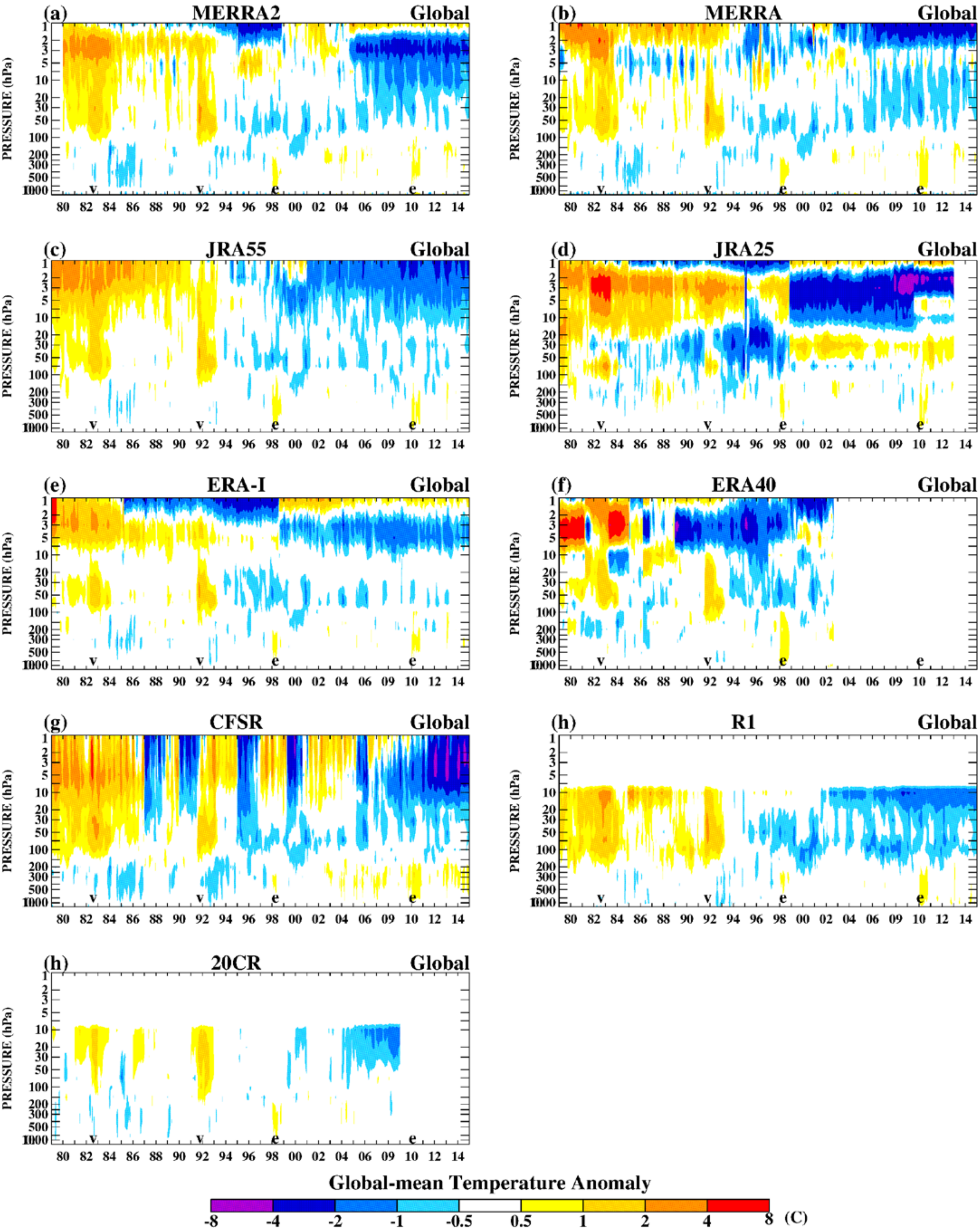
Figure 19. Time series plot of the a) global annual average of the lower stratospheric temperature layer (TLS) temperatures (°C) for MERRA-2, ERA-Interim, JRA-55, CFSR, and the NOAA/STAR TLS CDR, and b) the TLS temperature standard deviation (K) of the four reanalyses for each year. The climatological period spanned from 1981-2010. COSMIC GPSRO observations began to be assimilated in 2006. ~~Reanalyses anomaly differences from the NOAA/STAT anomalies are plotted in the right column.~~

	Model Version	Horizontal Resolution	Vertical Resolution	Model Top Level	RTM
R1/R2	NCEP MRF (1995/1998)	T62 : 1.875°	28 (σ)	3 hPa	Temperature retrievals
CFSR	NCEP CFS (2007)	T382 : 0.3125°	64 (hybrid σ-p)	~0.266 hPa	CRTM
ERA-40	IFS Cycle 23r4 (2001)	T_L159: ~125km	60 (hybrid σ-p)	0.1 hPa	RTTOVS-5
ERA-I	IFS Cycle 31r1 (2007)	T_L255 : ~79km	60 (hybrid σ-p)	0.1 hPa	RTTOVS-7
JRA-25	JMA GSM (2004)	T106: 1.125°	40 (hybrid σ-p)	0.4 hPa	RTTOVS-6 : TOVS RTTOVS-7 : ATOVS
JRA-55	JMA GSM (2009)	T_L319: ~55km	60 (hybrid σ-p)	0.1 hPa	RTTOVS-9
MERRA	GEOS 5.0.2 (2008)	0.5° lat x 0.667° lon	72 (hybrid σ-p)	0.01 hPa	CRTM
MERRA-2	GEOS 5.12.4 (2015)	0.5° lat x 0.625° lon	72 (hybrid σ-p)	0.01 hPa	CRTM

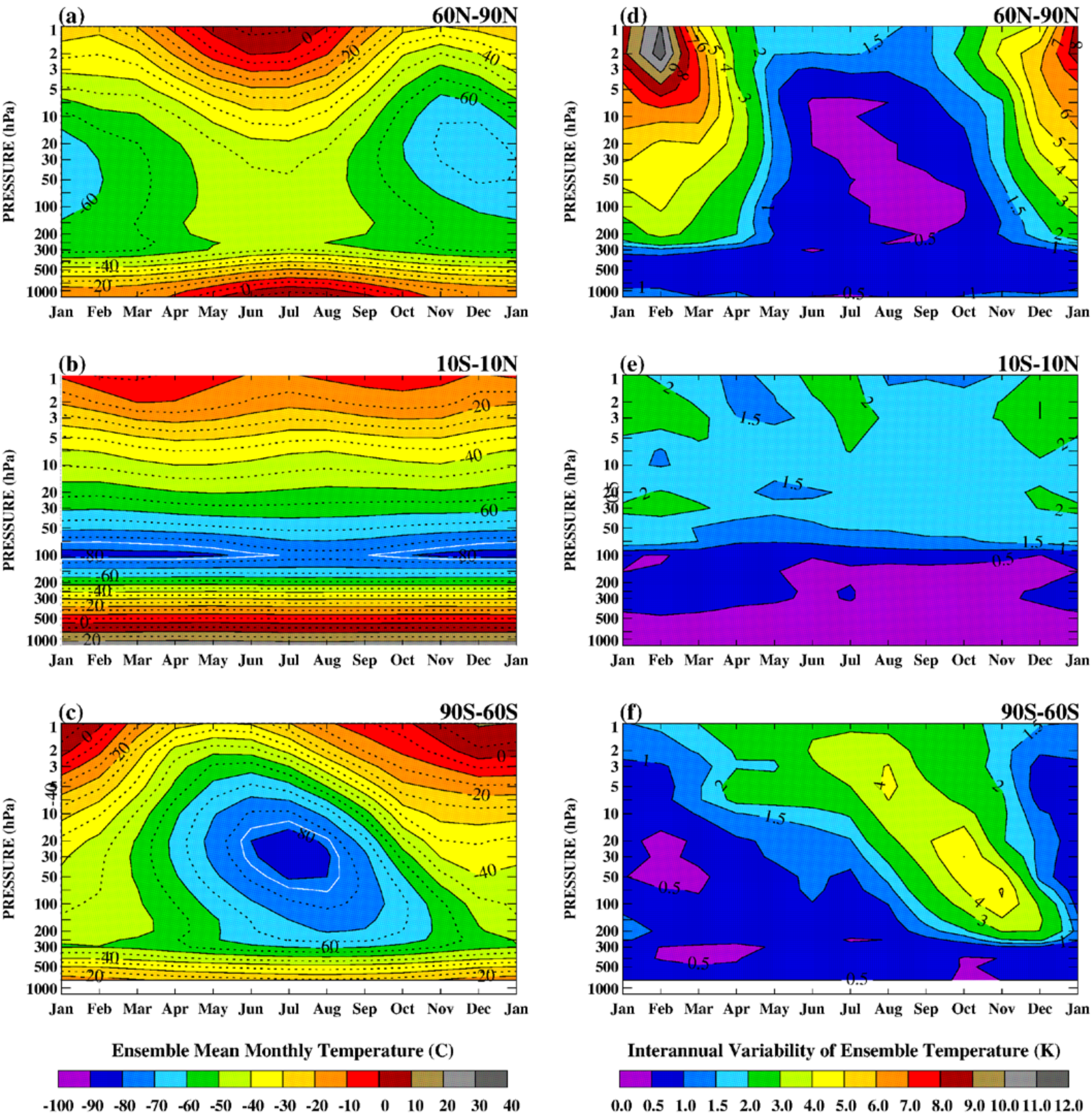
Table 1. Information about NCEP, JMA, ECMWF, and GMAO's earlier and latter reanalyses pertinent to the stratosphere. Information includes the model version, horizontal and vertical resolution, model top pressure, and radiative transfer model (RTM) used.

	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

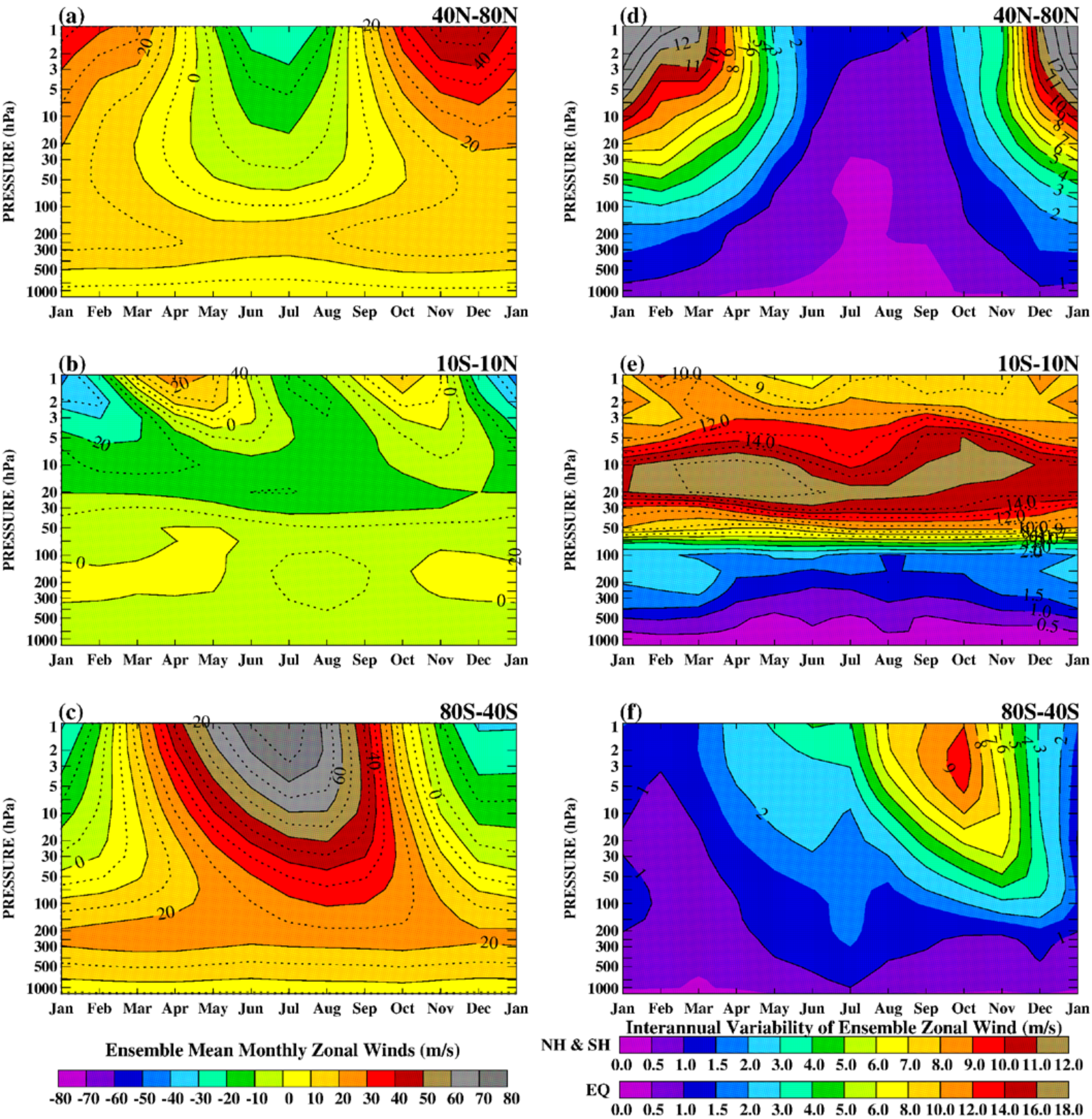
Table 21. Pressure (hPa) of SSU channels 1, 2, 3 and MSU channel 4 weighting function peaks, 50% of peak weight above, 50% of peak weight below, 10% of peak weight above, and 10% of peak weight below the peak.



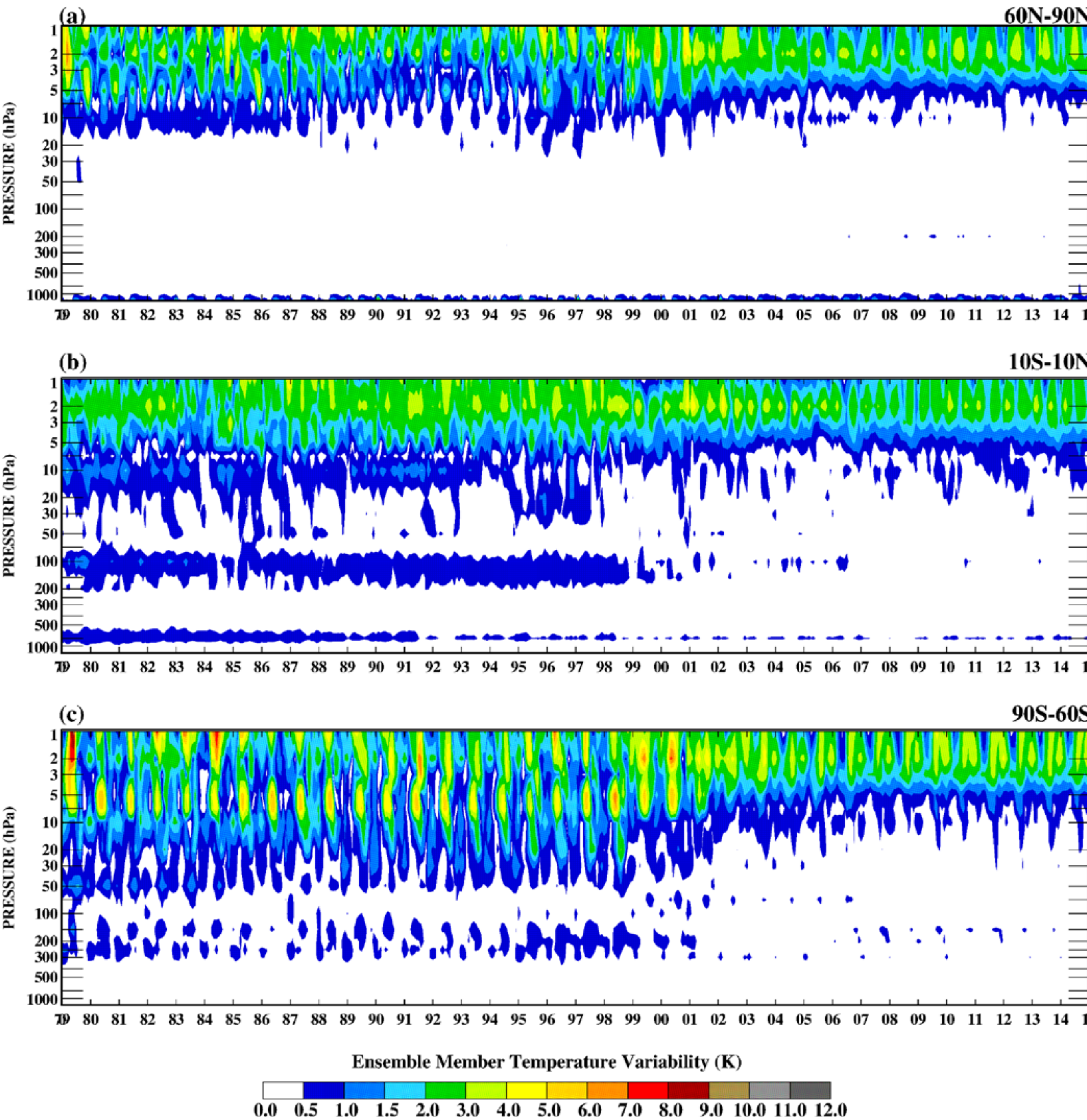
(a) MERRA2 Global
(b) MERRA Global
(c) JRA55 Global
(d) JRA25 Global
(e) ERA-I Global
(f) ERA40 Global
(g) CFSR Global
(h) R1 Global
(h) 20CR Global
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 300 500 1000
80 82 84 86 88 90 92 94 96 98 00 02 04 06 09 10 12 14
v v e e
Global-mean Temperature Anomaly
-8 -4 -2 -1 -0.5 0.5 1 2 4 8 (C)

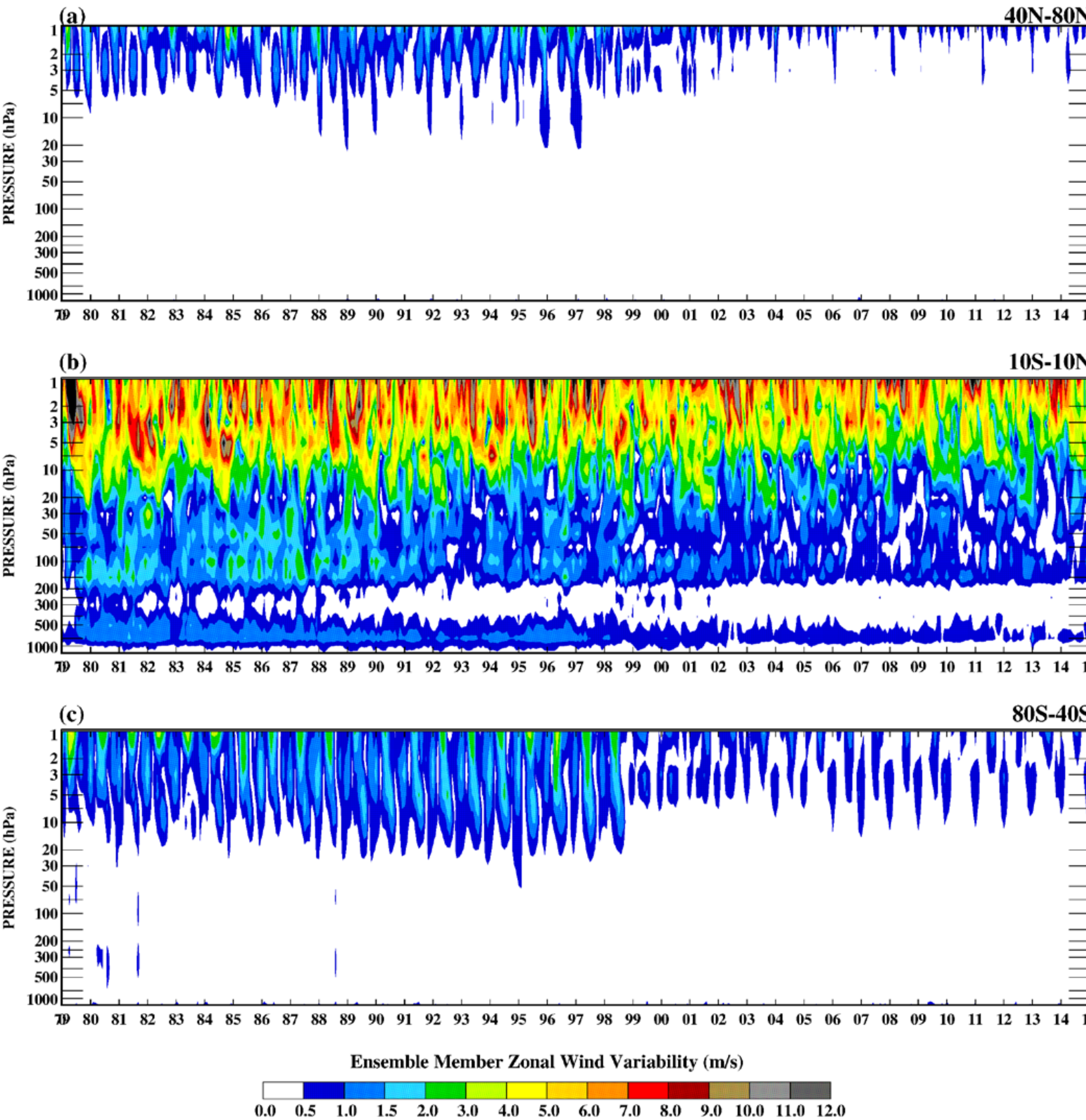


(a) 60N-90N
(b) 10S-10N
(c) 90S-60S
(d) 60N-90N
(e) 10S-10N
(f) 90S-60S
PRESSURE (hPa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
Ensemble Mean Monthly Temperature (C)
-100 -90 -80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40
Interannual Variability of Ensemble Temperature (K)
0.0 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0

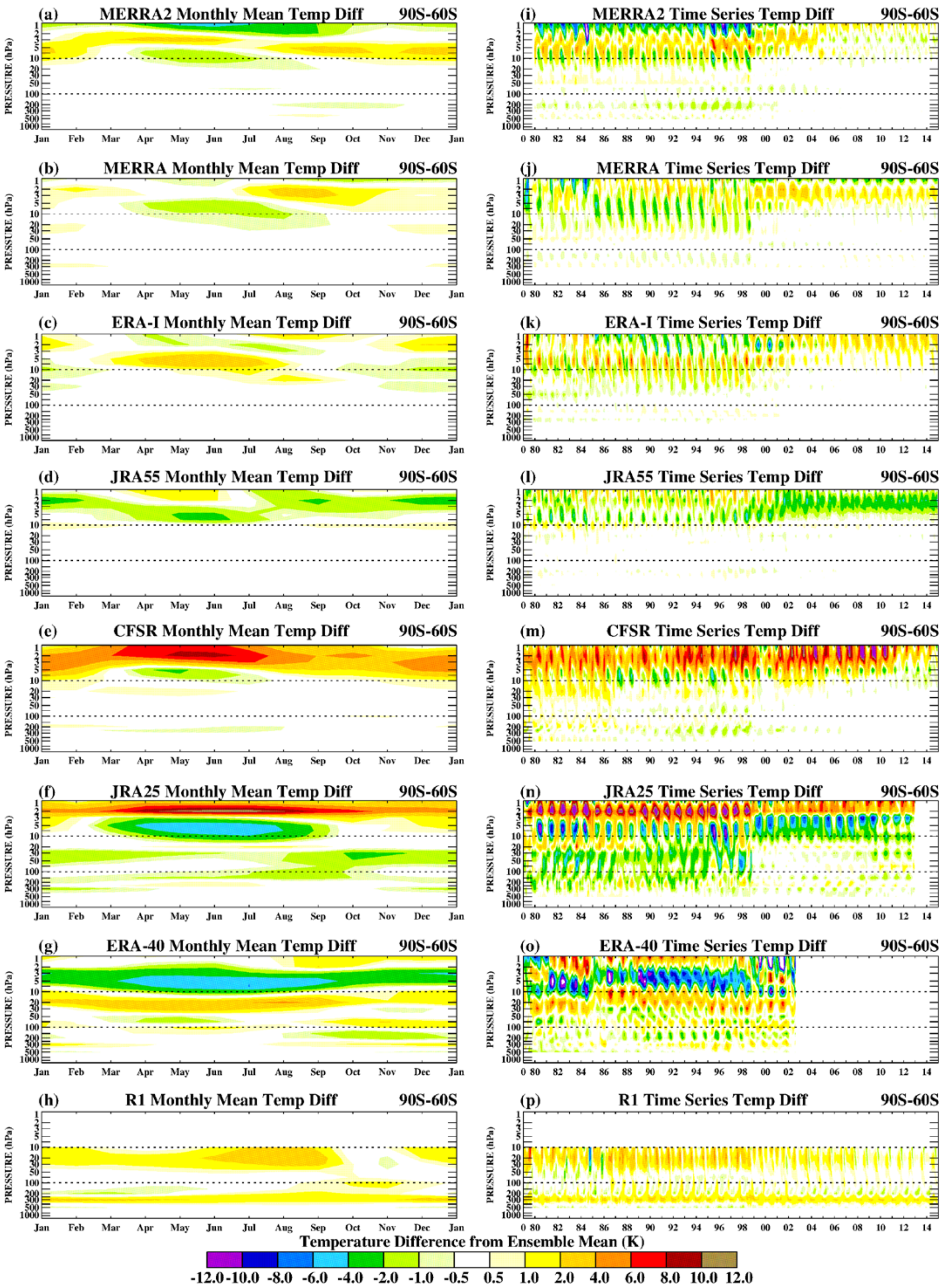


(a) 40N-80N
(b) 10S-10N
(c) 80S-40S
(d) 40N-80N
(e) 10S-10N
(f) 80S-40S
PRESSURE (hPa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
Ensemble Mean Monthly Zonal Winds (m/s)
-80 -70 -60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60 70 80
Interannual Variability of Ensemble Zonal Wind (m/s)
NH & SH
0.0 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0
EQ
0.0 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.0 8.0 10.0 12.0 14.0 16.0 18.0

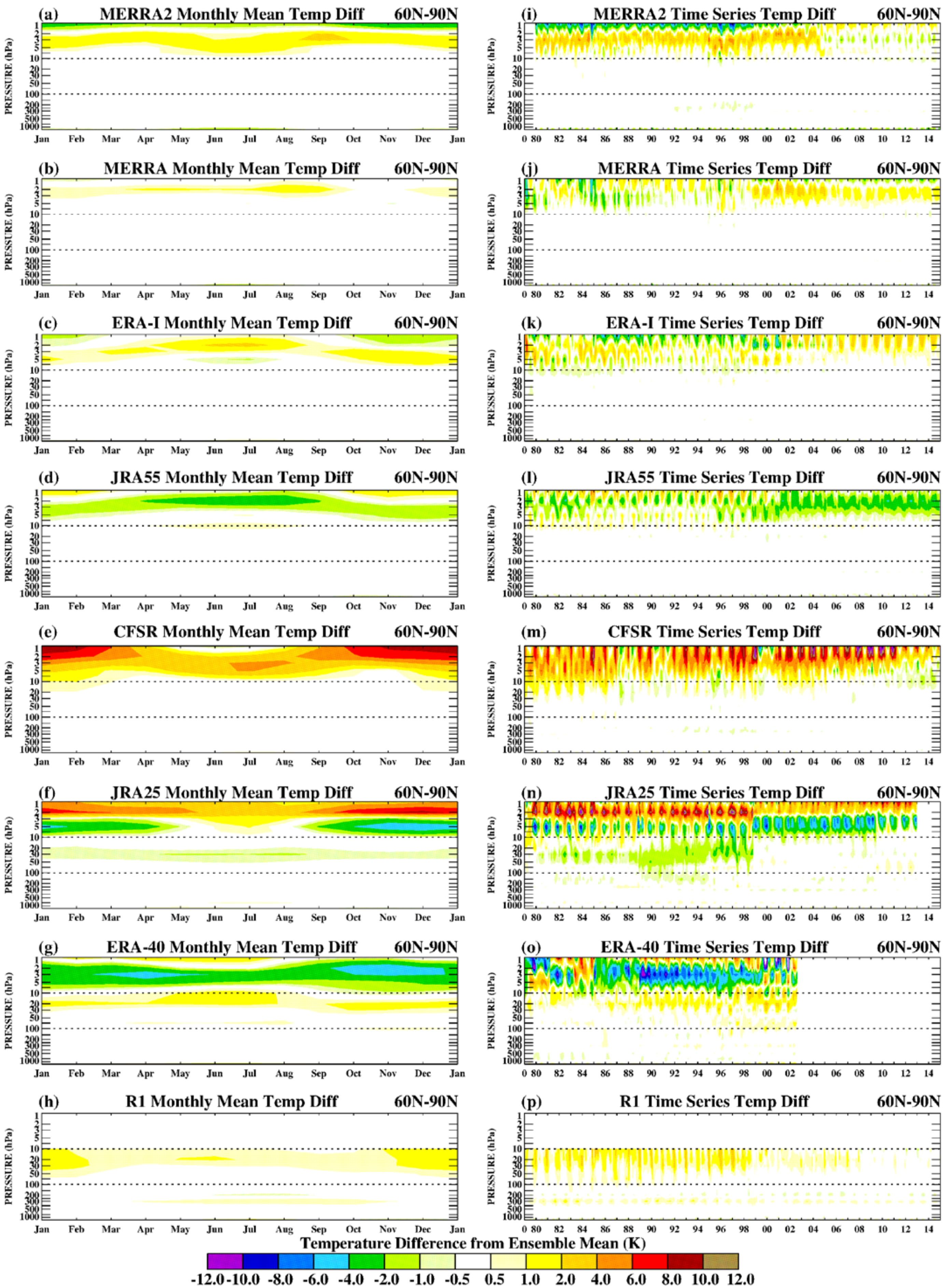

(a)
60N-90N
(b)
10S-10N
(c)
90S-60S
PRESSURE (hPa)
1
2
3
5
10
20
30
50
100
200
300
500
1000
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
00
01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
Ensemble Member Temperature Variability (K)
0.0
0.5
1.0
1.5
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0

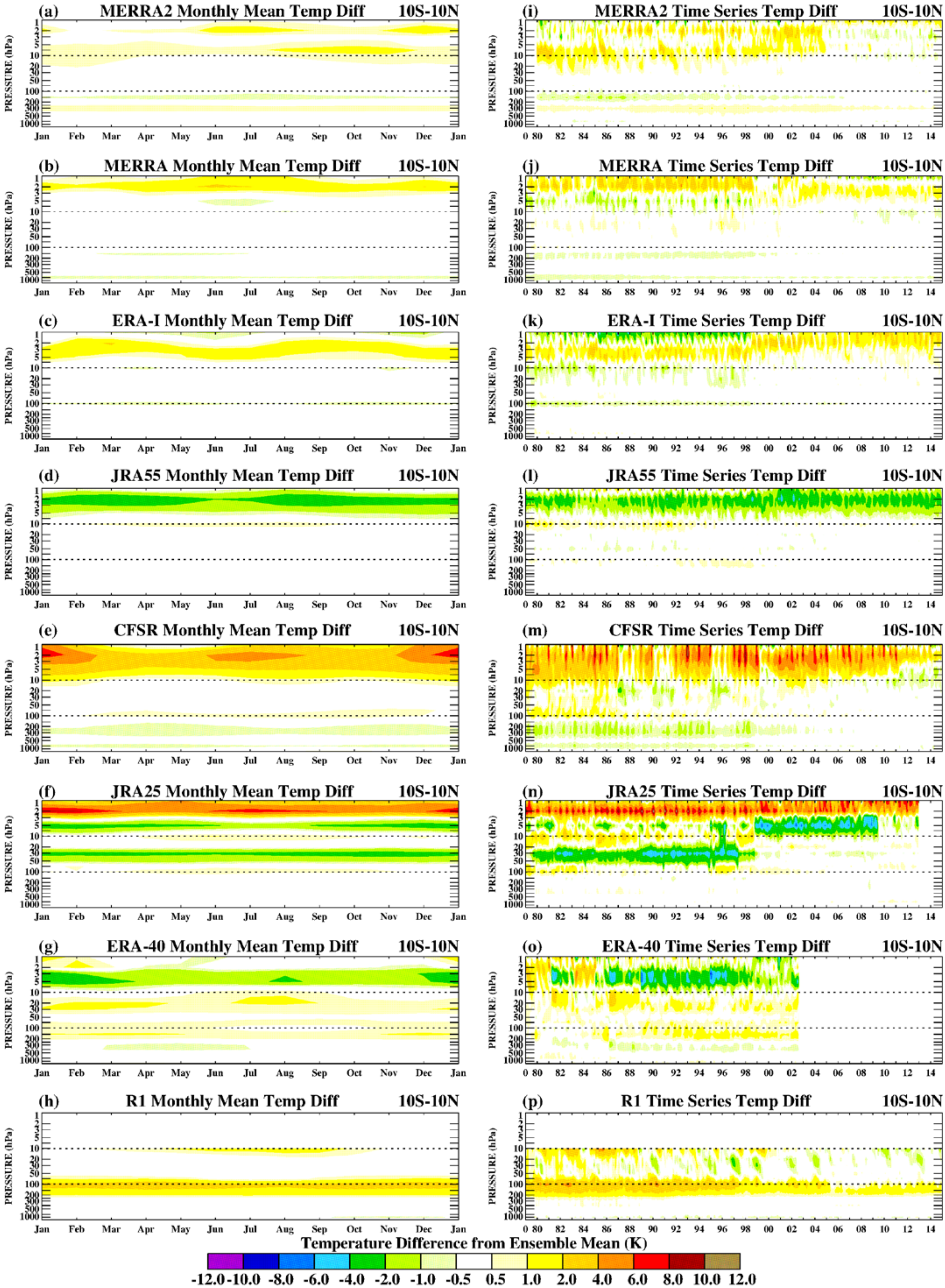


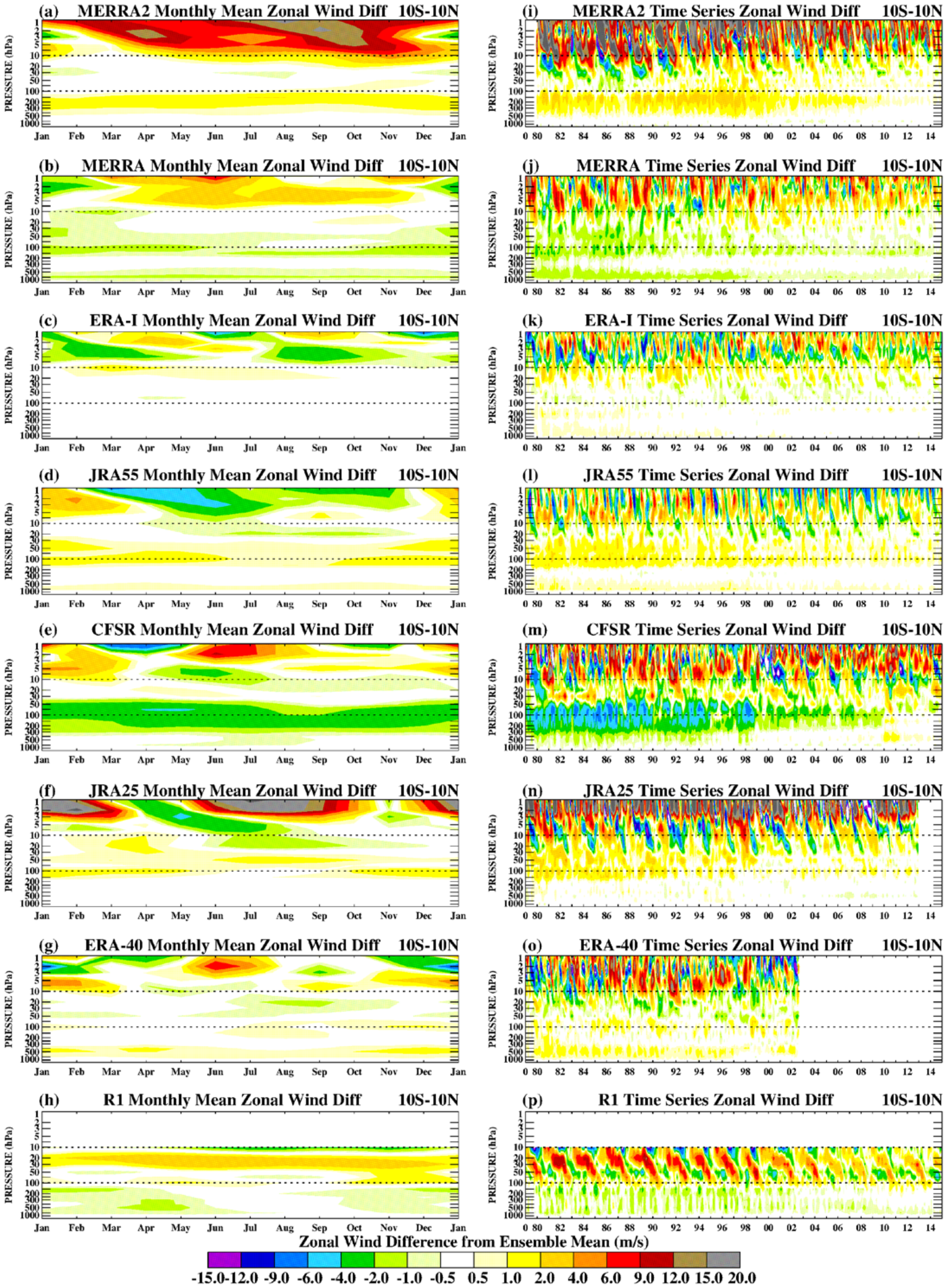
(a)
40N-80N
(b)
10S-10N
(c)
80S-40S
PRESSURE (hPa)
1
2
3
5
10
20
30
50
100
200
300
500
1000
79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15
Ensemble Member Zonal Wind Variability (m/s)
0.0 0.5 1.0 1.5 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0



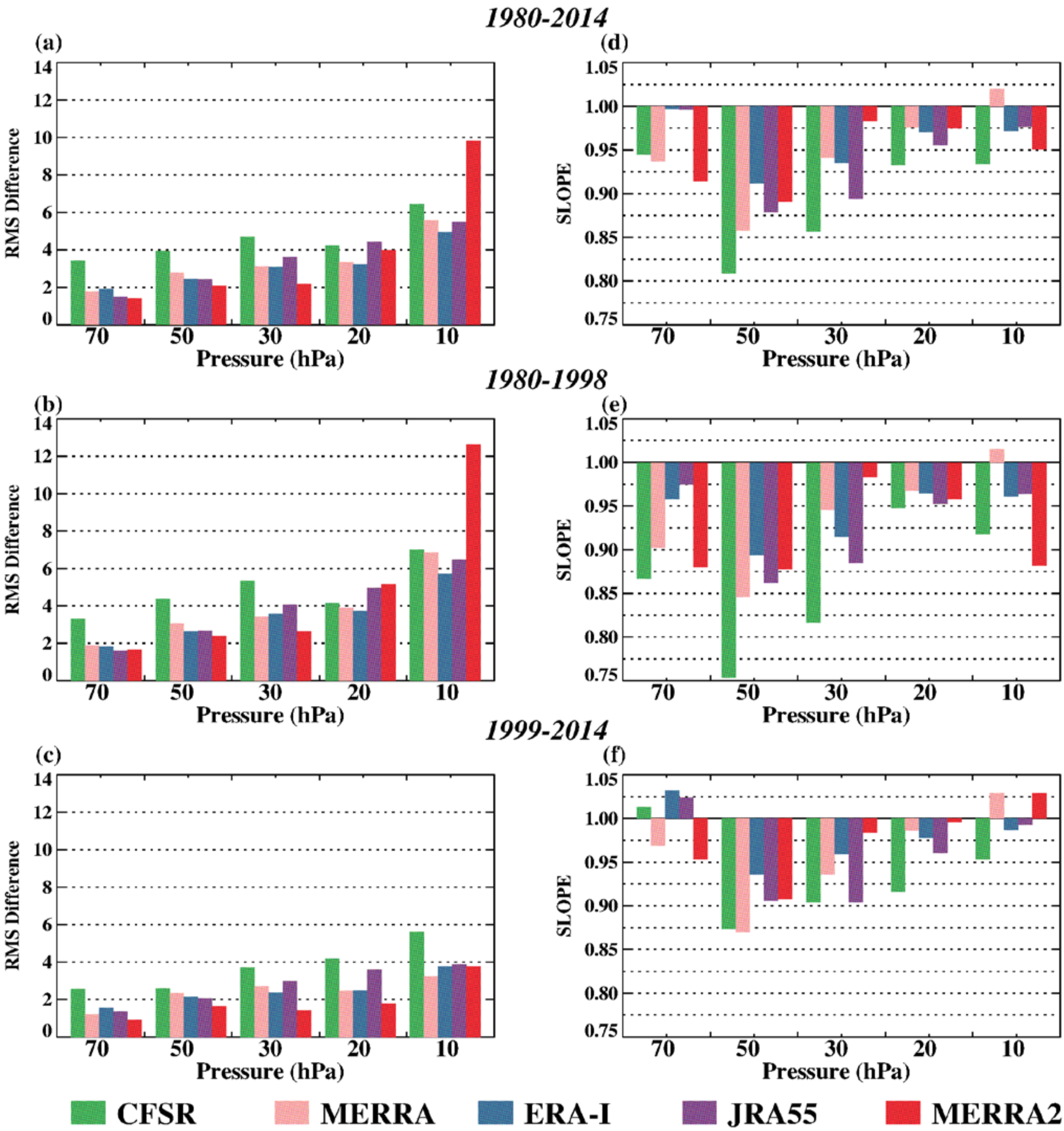
(a) MERRA2 Monthly Mean Temp Diff 90S-60S
(b) MERRA Monthly Mean Temp Diff 90S-60S
(c) ERA-I Monthly Mean Temp Diff 90S-60S
(d) JRA55 Monthly Mean Temp Diff 90S-60S
(e) CFSR Monthly Mean Temp Diff 90S-60S
(f) JRA25 Monthly Mean Temp Diff 90S-60S
(g) ERA-40 Monthly Mean Temp Diff 90S-60S
(h) R1 Monthly Mean Temp Diff 90S-60S
(i) MERRA2 Time Series Temp Diff 90S-60S
(j) MERRA Time Series Temp Diff 90S-60S
(k) ERA-I Time Series Temp Diff 90S-60S
(l) JRA55 Time Series Temp Diff 90S-60S
(m) CFSR Time Series Temp Diff 90S-60S
(n) JRA25 Time Series Temp Diff 90S-60S
(o) ERA-40 Time Series Temp Diff 90S-60S
(p) R1 Time Series Temp Diff 90S-60S
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 300 500 1000
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
0 80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
Temperature Difference from Ensemble Mean (K)
-12.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 12.0

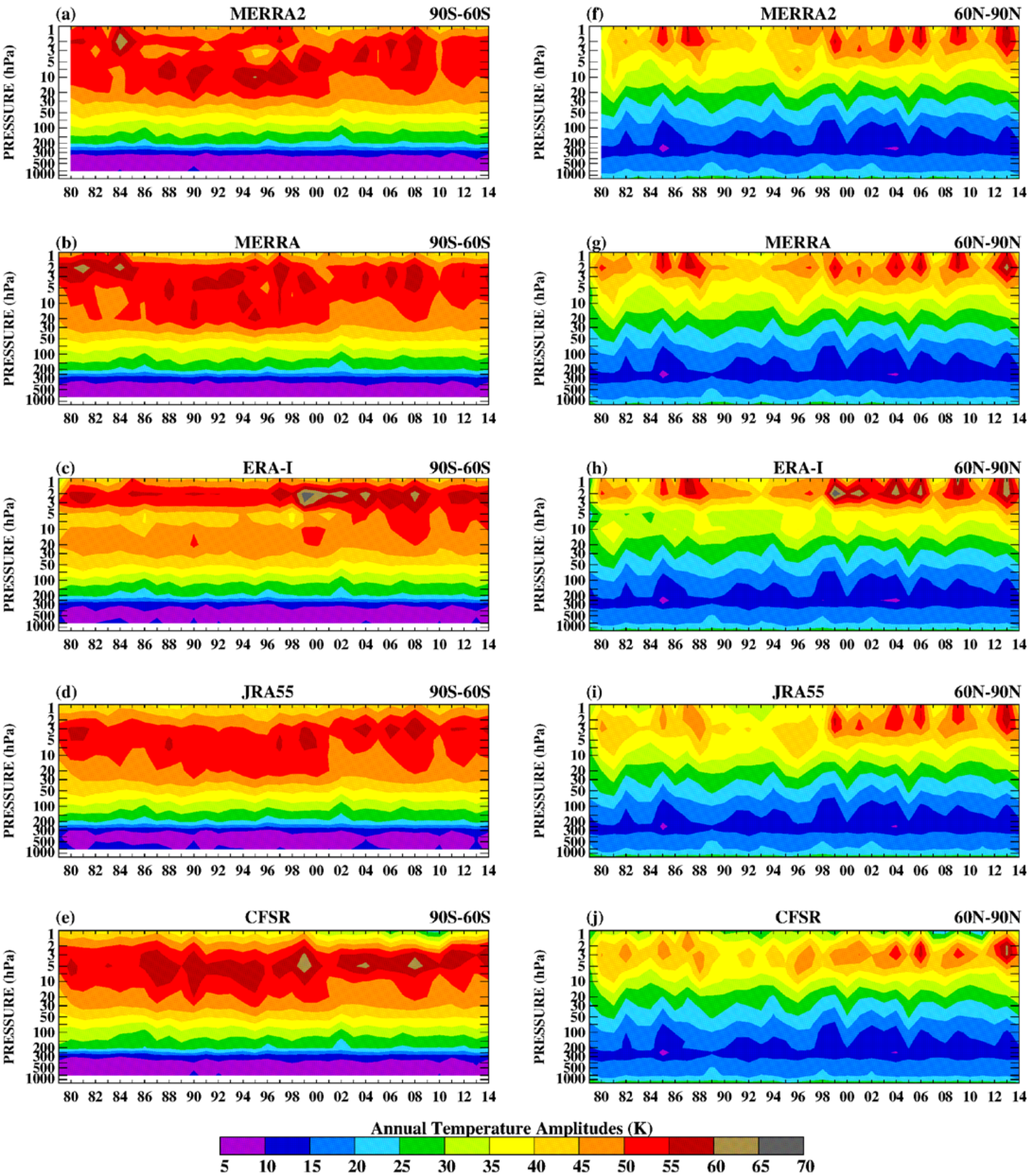

(a) MERRA2 Monthly Mean Temp Diff 60N-90N
(b) MERRA Monthly Mean Temp Diff 60N-90N
(c) ERA-I Monthly Mean Temp Diff 60N-90N
(d) JRA55 Monthly Mean Temp Diff 60N-90N
(e) CFSR Monthly Mean Temp Diff 60N-90N
(f) JRA25 Monthly Mean Temp Diff 60N-90N
(g) ERA-40 Monthly Mean Temp Diff 60N-90N
(h) R1 Monthly Mean Temp Diff 60N-90N
(i) MERRA2 Time Series Temp Diff 60N-90N
(j) MERRA Time Series Temp Diff 60N-90N
(k) ERA-I Time Series Temp Diff 60N-90N
(l) JRA55 Time Series Temp Diff 60N-90N
(m) CFSR Time Series Temp Diff 60N-90N
(n) JRA25 Time Series Temp Diff 60N-90N
(o) ERA-40 Time Series Temp Diff 60N-90N
(p) R1 Time Series Temp Diff 60N-90N
PRESSURE (hPa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
0 80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
Temperature Difference from Ensemble Mean (K)
-12.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 12.0


(a) MERRA2 Monthly Mean Temp Diff 10S-10N
(b) MERRA Monthly Mean Temp Diff 10S-10N
(c) ERA-I Monthly Mean Temp Diff 10S-10N
(d) JRA55 Monthly Mean Temp Diff 10S-10N
(e) CFSR Monthly Mean Temp Diff 10S-10N
(f) JRA25 Monthly Mean Temp Diff 10S-10N
(g) ERA-40 Monthly Mean Temp Diff 10S-10N
(h) R1 Monthly Mean Temp Diff 10S-10N
(i) MERRA2 Time Series Temp Diff 10S-10N
(j) MERRA Time Series Temp Diff 10S-10N
(k) ERA-I Time Series Temp Diff 10S-10N
(l) JRA55 Time Series Temp Diff 10S-10N
(m) CFSR Time Series Temp Diff 10S-10N
(n) JRA25 Time Series Temp Diff 10S-10N
(o) ERA-40 Time Series Temp Diff 10S-10N
(p) R1 Time Series Temp Diff 10S-10N
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 300 500 1000
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
0 80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
Temperature Difference from Ensemble Mean (K)
-12.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 12.0

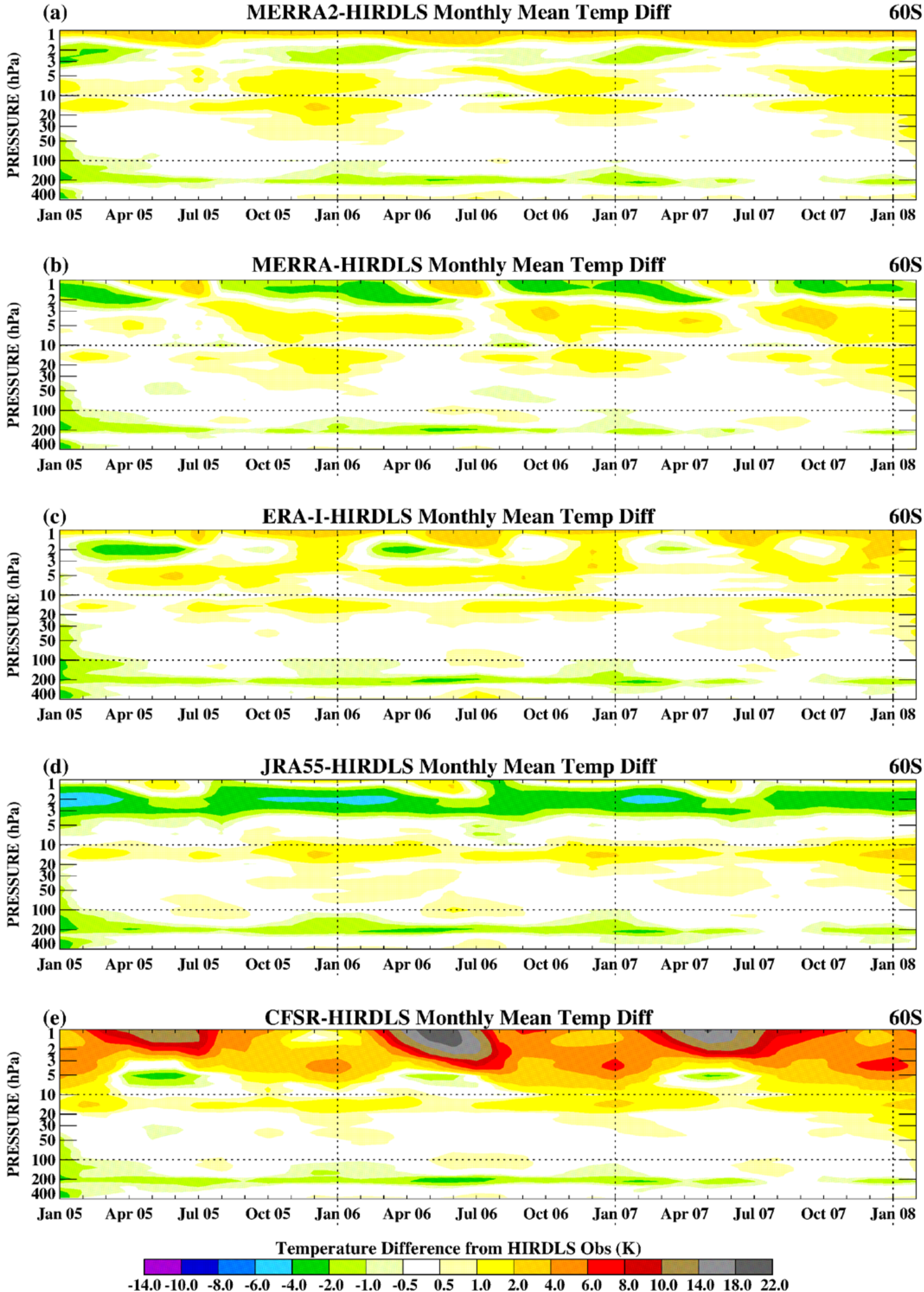


(a) MERRA2 Monthly Mean Zonal Wind Diff 10S-10N
(b) MERRA Monthly Mean Zonal Wind Diff 10S-10N
(c) ERA-I Monthly Mean Zonal Wind Diff 10S-10N
(d) JRA55 Monthly Mean Zonal Wind Diff 10S-10N
(e) CFSR Monthly Mean Zonal Wind Diff 10S-10N
(f) JRA25 Monthly Mean Zonal Wind Diff 10S-10N
(g) ERA-40 Monthly Mean Zonal Wind Diff 10S-10N
(h) R1 Monthly Mean Zonal Wind Diff 10S-10N
(i) MERRA2 Time Series Zonal Wind Diff 10S-10N
(j) MERRA Time Series Zonal Wind Diff 10S-10N
(k) ERA-I Time Series Zonal Wind Diff 10S-10N
(l) JRA55 Time Series Zonal Wind Diff 10S-10N
(m) CFSR Time Series Zonal Wind Diff 10S-10N
(n) JRA25 Time Series Zonal Wind Diff 10S-10N
(o) ERA-40 Time Series Zonal Wind Diff 10S-10N
(p) R1 Time Series Zonal Wind Diff 10S-10N
PRESSURE (hPa)
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
0 80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
Zonal Wind Difference from Ensemble Mean (m/s)
-15.0 -12.0 -9.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 9.0 12.0 15.0 20.0

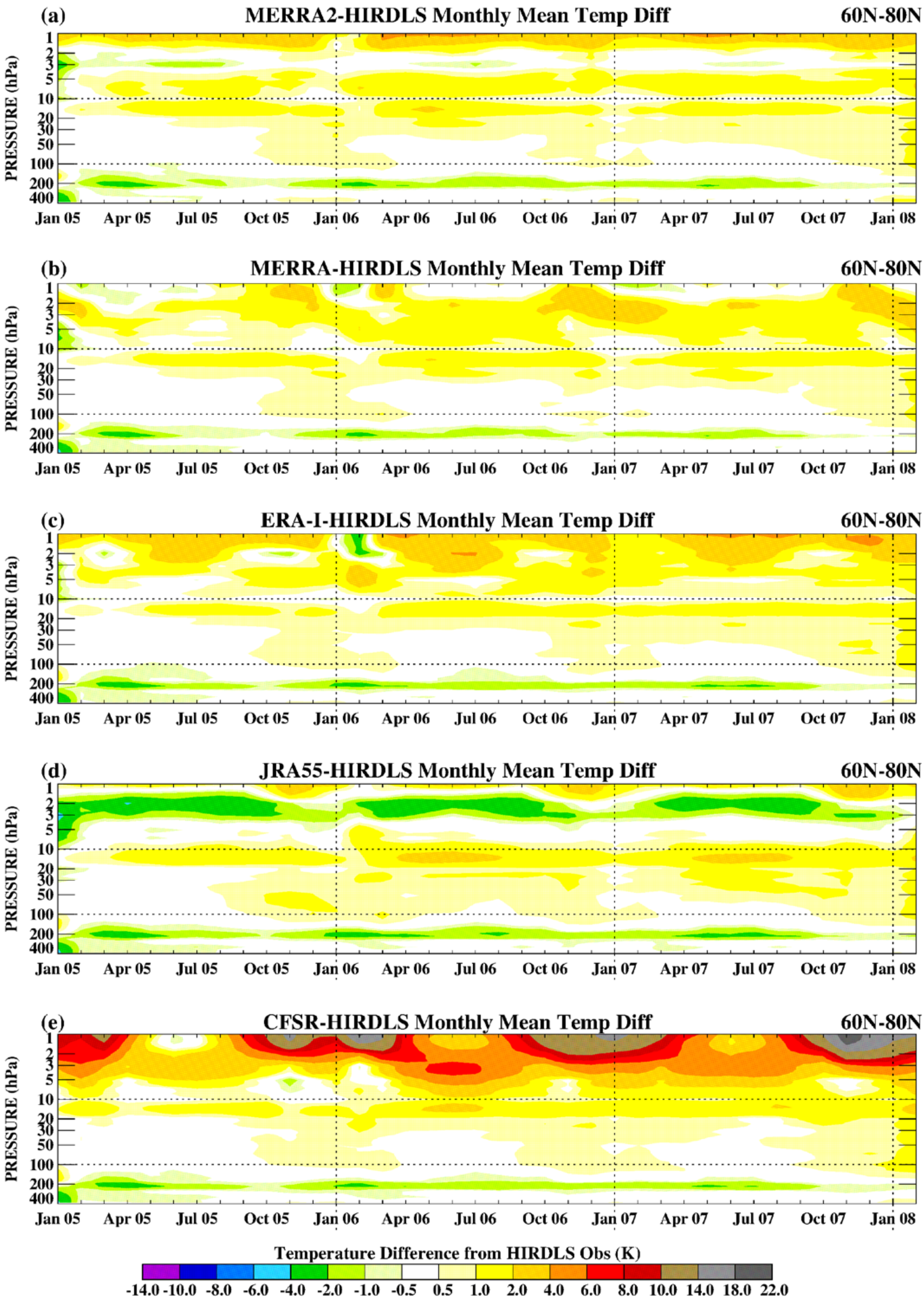

1980-2014
(a)
RMS Difference
Pressure (hPa)
(d)
SLOPE
Pressure (hPa)
1980-1998
(b)
RMS Difference
Pressure (hPa)
(e)
SLOPE
Pressure (hPa)
1999-2014
(c)
RMS Difference
Pressure (hPa)
(f)
SLOPE
Pressure (hPa)
CFSR
MERRA
ERA-I
JRA55
MERRA2



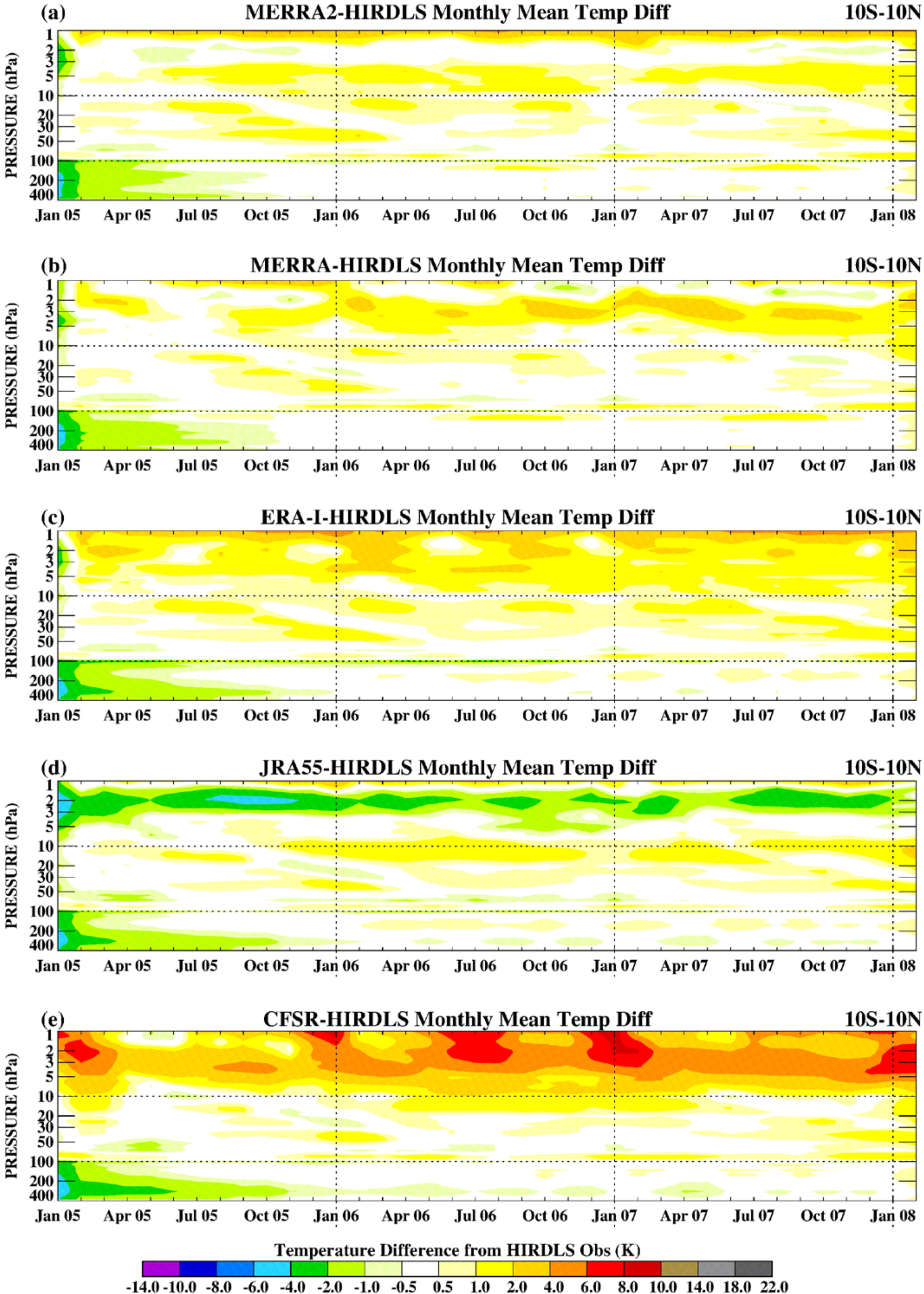
(a) MERRA2 90S-60S
(b) MERRA 90S-60S
(c) ERA-I 90S-60S
(d) JRA55 90S-60S
(e) CFSR 90S-60S
(f) MERRA2 60N-90N
(g) MERRA 60N-90N
(h) ERA-I 60N-90N
(i) JRA55 60N-90N
(j) CFSR 60N-90N
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 300 500 1000
80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
Annual Temperature Amplitudes (K)
5 10 15 20 25 30 35 40 45 50 55 60 65 70

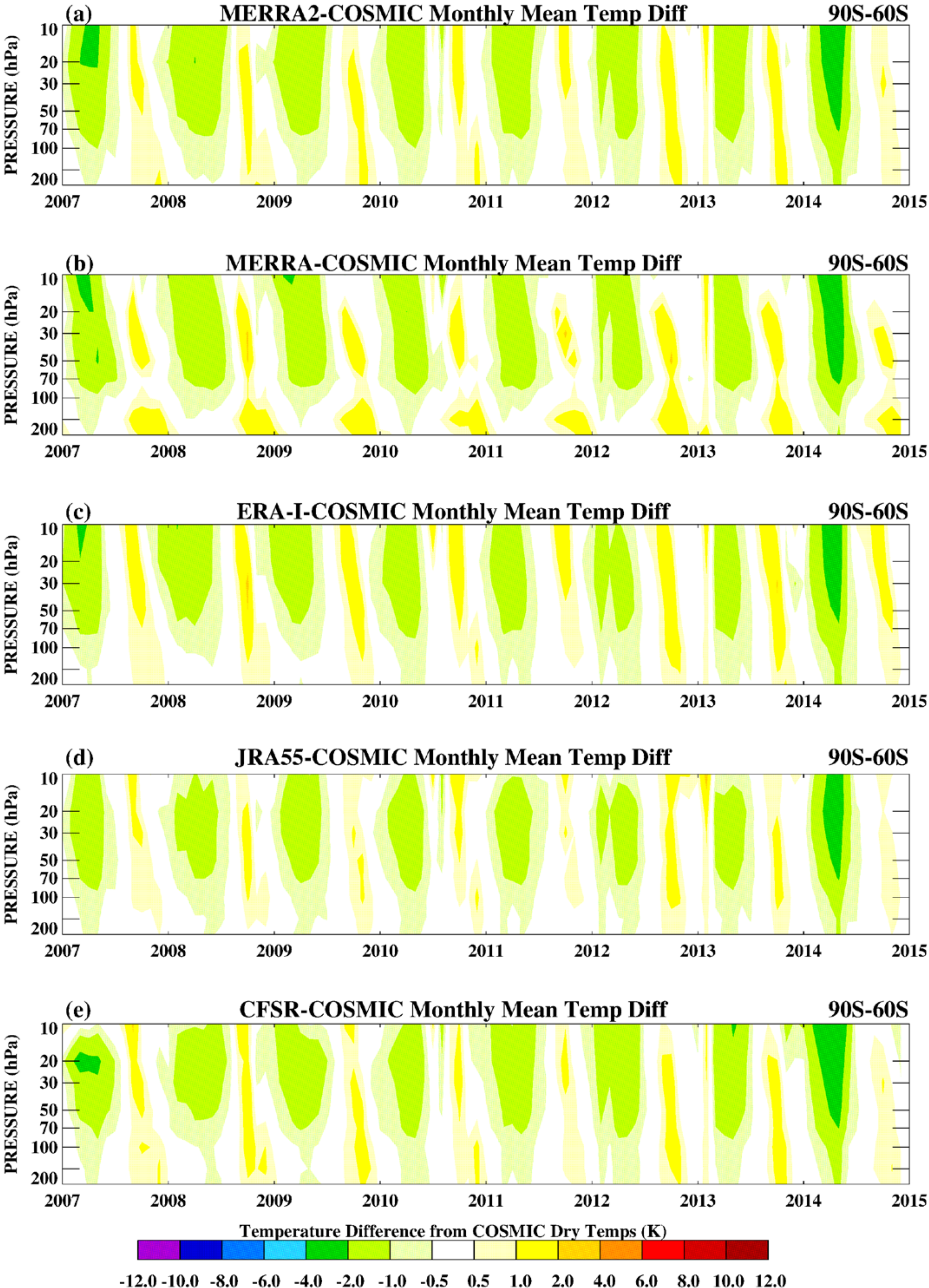


(a) MERRA2-HIRDLS Monthly Mean Temp Diff 60S
(b) MERRA-HIRDLS Monthly Mean Temp Diff 60S
(c) ERA-I-HIRDLS Monthly Mean Temp Diff 60S
(d) JRA55-HIRDLS Monthly Mean Temp Diff 60S
(e) CFSR-HIRDLS Monthly Mean Temp Diff 60S
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 400
Jan 05 Apr 05 Jul 05 Oct 05 Jan 06 Apr 06 Jul 06 Oct 06 Jan 07 Apr 07 Jul 07 Oct 07 Jan 08
Temperature Difference from HIRDLS Obs (K)
-14.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 14.0 18.0 22.0

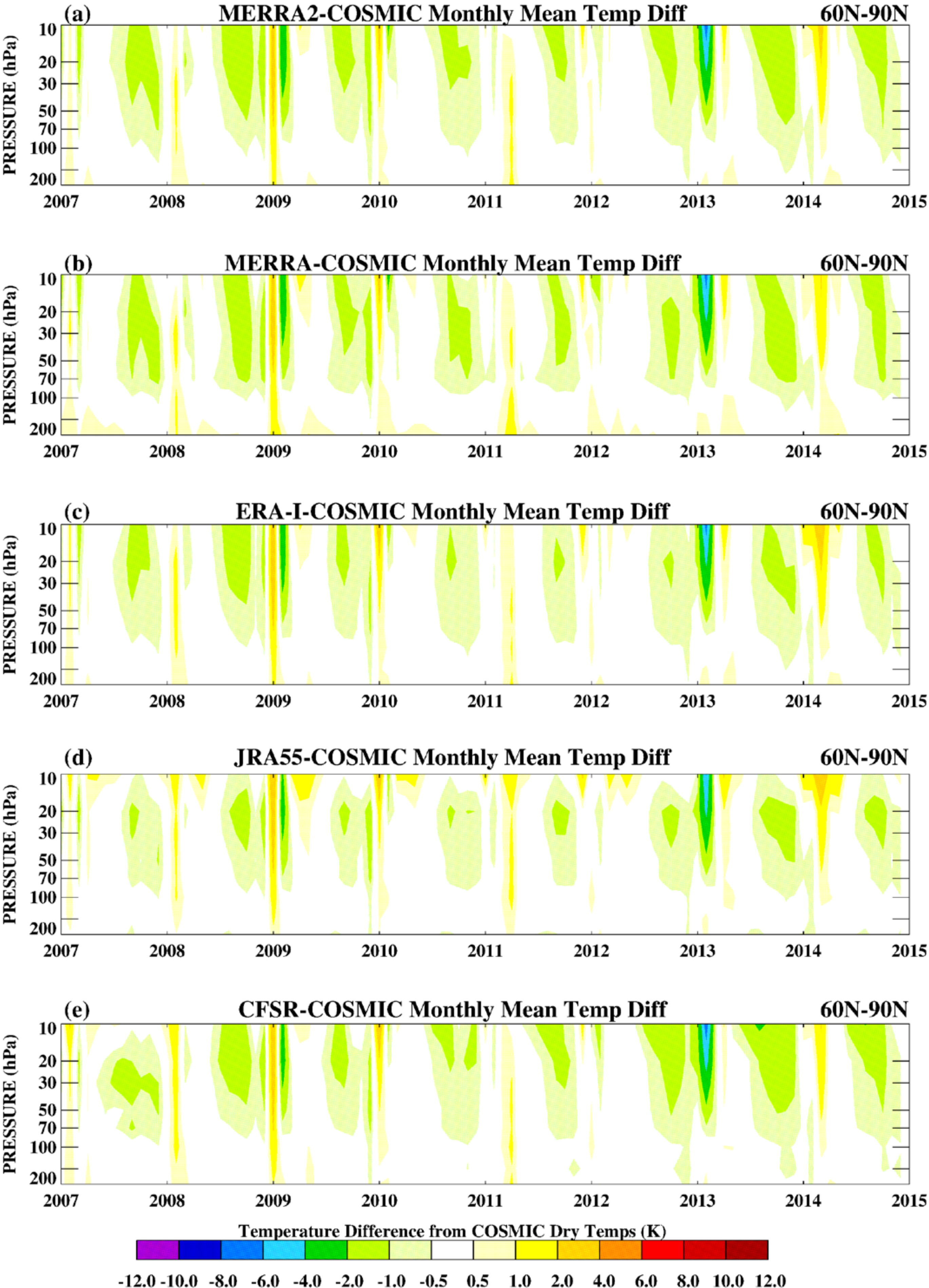


(a)
MERRA2-HIRDLS Monthly Mean Temp Diff
60N-80N
(b)
MERRA-HIRDLS Monthly Mean Temp Diff
60N-80N
(c)
ERA-I-HIRDLS Monthly Mean Temp Diff
60N-80N
(d)
JRA55-HIRDLS Monthly Mean Temp Diff
60N-80N
(e)
CFSR-HIRDLS Monthly Mean Temp Diff
60N-80N
PRESSURE (hPa)
1
2
3
5
10
20
30
50
100
200
400
Jan 05
Apr 05
Jul 05
Oct 05
Jan 06
Apr 06
Jul 06
Oct 06
Jan 07
Apr 07
Jul 07
Oct 07
Jan 08
Temperature Difference from HIRDLS Obs (K)
-14.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 14.0 18.0 22.0

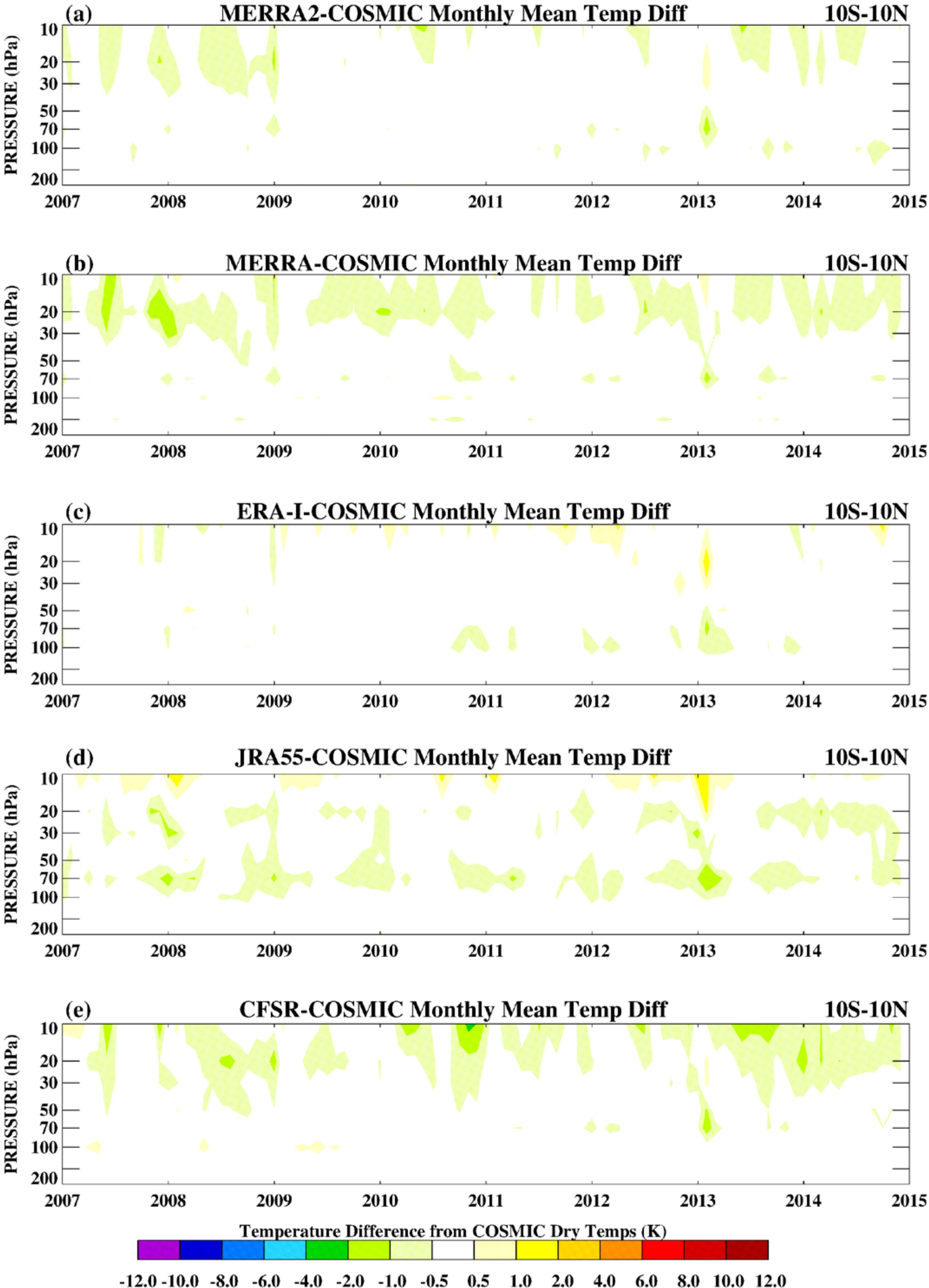

(a)
MERRA2-HIRDLS Monthly Mean Temp Diff
10S-10N
(b)
MERRA-HIRDLS Monthly Mean Temp Diff
10S-10N
(c)
ERA-I-HIRDLS Monthly Mean Temp Diff
10S-10N
(d)
JRA55-HIRDLS Monthly Mean Temp Diff
10S-10N
(e)
CFSR-HIRDLS Monthly Mean Temp Diff
10S-10N
PRESSURE (hPa)
1 2 3 5 10 20 30 50 100 200 400
Jan 05 Apr 05 Jul 05 Oct 05 Jan 06 Apr 06 Jul 06 Oct 06 Jan 07 Apr 07 Jul 07 Oct 07 Jan 08
Temperature Difference from HIRDLS Obs (K)
-14.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 14.0 18.0 22.0



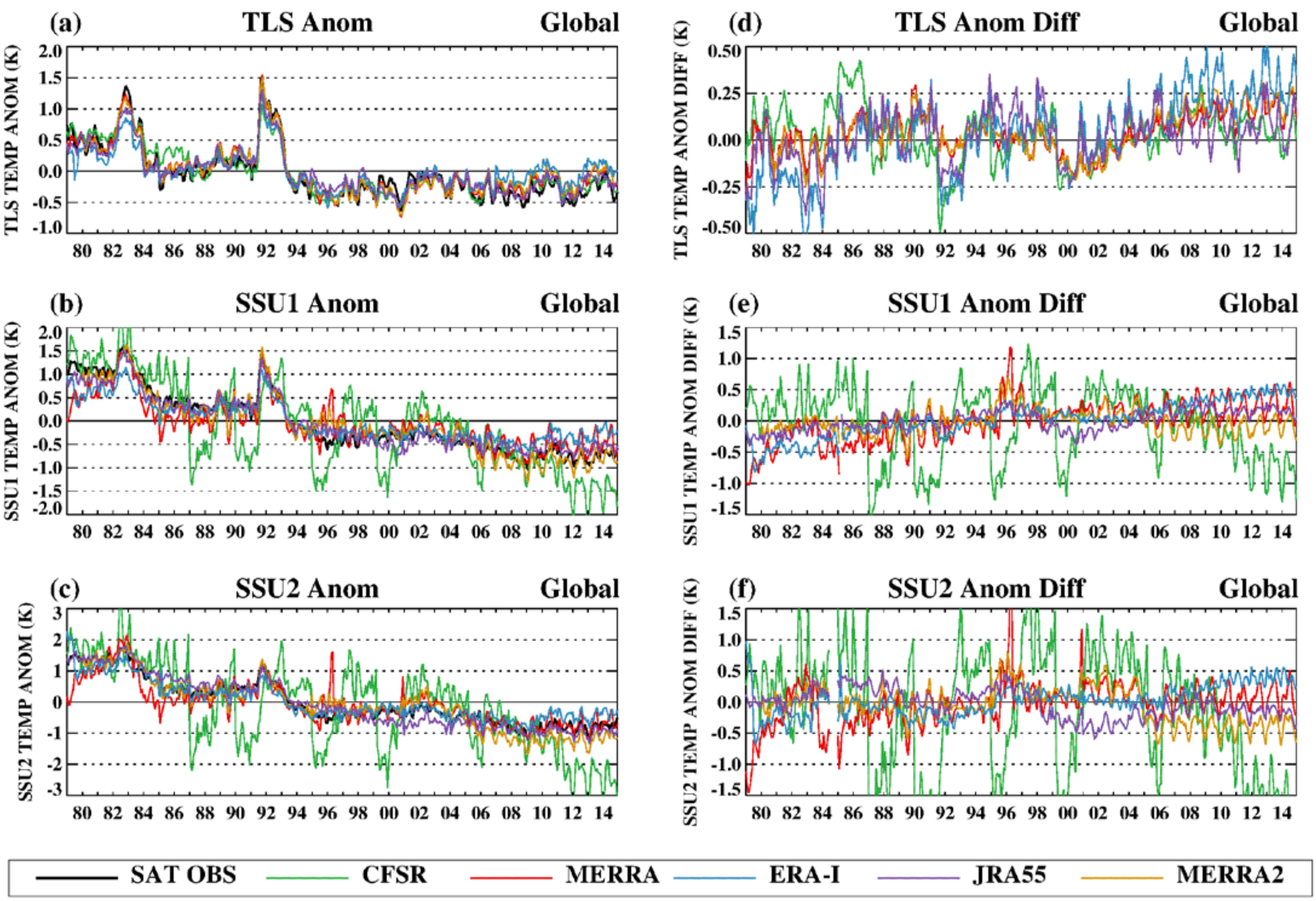
(a) MERRA2-COSMIC Monthly Mean Temp Diff 90S-60S
(b) MERRA-COSMIC Monthly Mean Temp Diff 90S-60S
(c) ERA-I-COSMIC Monthly Mean Temp Diff 90S-60S
(d) JRA55-COSMIC Monthly Mean Temp Diff 90S-60S
(e) CFSR-COSMIC Monthly Mean Temp Diff 90S-60S
PRESSURE (hPa)
10
20
30
50
70
100
200
2007
2008
2009
2010
2011
2012
2013
2014
2015
Temperature Difference from COSMIC Dry Temps (K)
-12.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 12.0



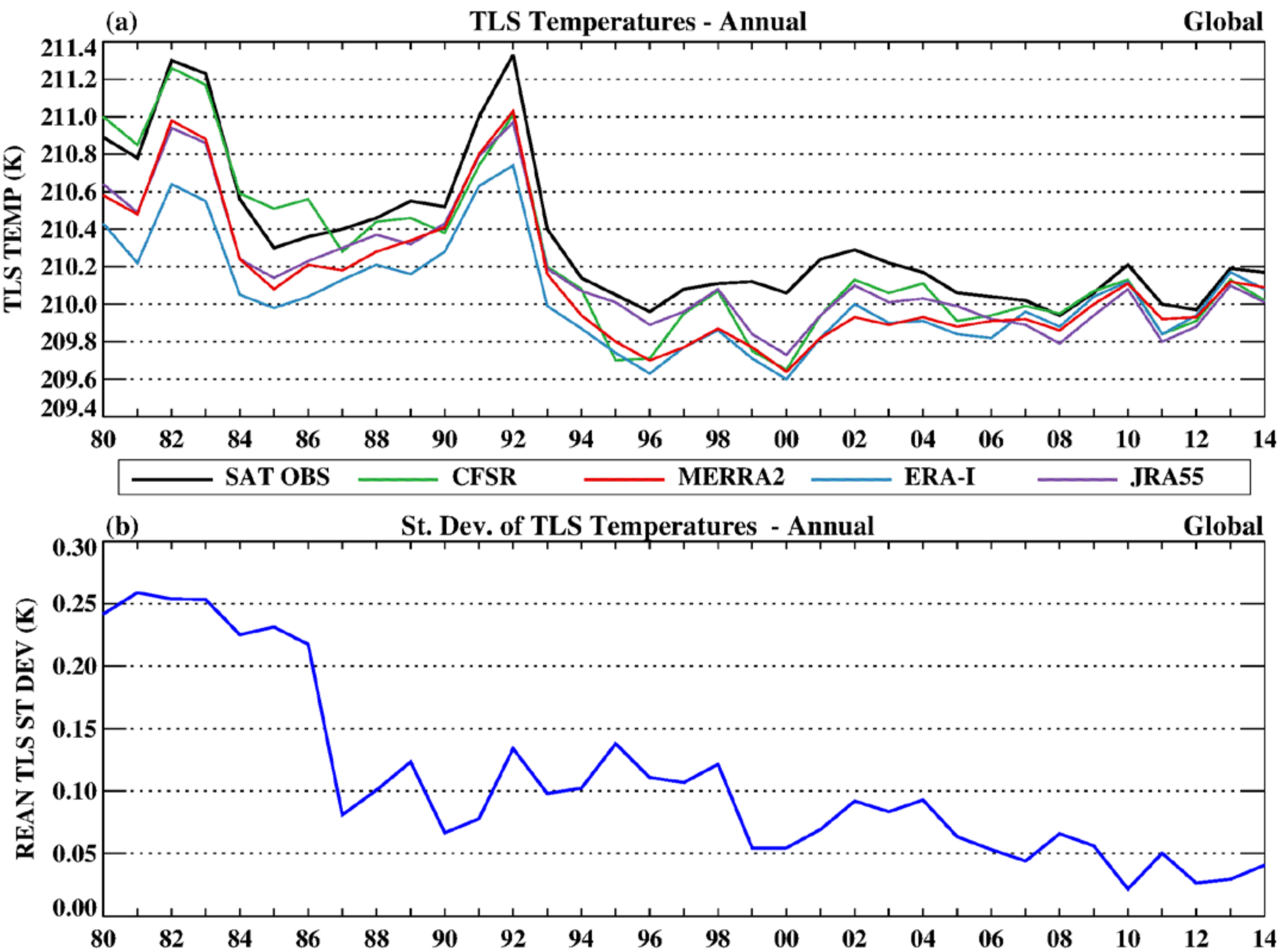
(a) MERRA2-COSMIC Monthly Mean Temp Diff
60N-90N
(b) MERRA-COSMIC Monthly Mean Temp Diff
60N-90N
(c) ERA-I-COSMIC Monthly Mean Temp Diff
60N-90N
(d) JRA55-COSMIC Monthly Mean Temp Diff
60N-90N
(e) CFSR-COSMIC Monthly Mean Temp Diff
60N-90N
PRESSURE (hPa)
10
20
30
50
70
100
200
2007
2008
2009
2010
2011
2012
2013
2014
2015
Temperature Difference from COSMIC Dry Temps (K)
-12.0
-10.0
-8.0
-6.0
-4.0
-2.0
-1.0
-0.5
0.5
1.0
2.0
4.0
6.0
8.0
10.0
12.0



(a)
MERRA2-COSMIC Monthly Mean Temp Diff
10S-10N
(b)
MERRA-COSMIC Monthly Mean Temp Diff
10S-10N
(c)
ERA-I-COSMIC Monthly Mean Temp Diff
10S-10N
(d)
JRA55-COSMIC Monthly Mean Temp Diff
10S-10N
(e)
CFSR-COSMIC Monthly Mean Temp Diff
10S-10N
PRESSURE (hPa)
10
20
30
50
70
100
200
2007
2008
2009
2010
2011
2012
2013
2014
2015
Temperature Difference from COSMIC Dry Temps (K)
-12.0 -10.0 -8.0 -6.0 -4.0 -2.0 -1.0 -0.5 0.5 1.0 2.0 4.0 6.0 8.0 10.0 12.0



(a) TLS Anom Global
TLS TEMP ANOM (K)
(b) SSU1 Anom Global
SSU1 TEMP ANOM (K)
(c) SSU2 Anom Global
SSU2 TEMP ANOM (K)
(d) TLS Anom Diff Global
TLS TEMP ANOM DIFF (K)
(e) SSU1 Anom Diff Global
SSU1 TEMP ANOM DIFF (K)
(f) SSU2 Anom Diff Global
SSU2 TEMP ANOM DIFF (K)
80 82 84 86 88 90 92 94 96 98 00 02 04 06 08 10 12 14
SAT OBS
CFSR
MERRA
ERA-I
JRA55
MERRA2



(a)
TLS Temperatures - Annual
Global
TLS TEMP (K)
211.4
211.2
211.0
210.8
210.6
210.4
210.2
210.0
209.8
209.6
209.4
80
82
84
86
88
90
92
94
96
98
00
02
04
06
08
10
12
14
SAT OBS
CFSR
MERRA2
ERA-I
JRA55
(b)
St. Dev. of TLS Temperatures - Annual
Global
REAN TLS ST DEV (K)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
80
82
84
86
88
90
92
94
96
98
00
02
04
06
08
10
12
14