

A Comparative Study of Wave Forcing Derived from the ERA-40 and ERA-Interim Reanalysis Datasets

HUA LU, THOMAS J. BRACEGIRDLE, TONY PHILLIPS, AND JOHN TURNER

British Antarctic Survey, Cambridge, United Kingdom

(Manuscript received 21 May 2014, in final form 28 November 2014)

ABSTRACT

The Eliassen–Palm (E–P) flux divergences derived from ERA-40 and ERA-Interim show significant differences during northern winter. The discrepancies are marked by vertically alternating positive and negative anomalies at high latitudes and are manifested via a difference in the climatology. The magnitude of the discrepancies can be greater than the interannual variability in certain regions. These wave forcing discrepancies are only partially linked to differences in the residual circulation but they are evidently related to the static stability in the affected regions. Thus, the main cause of the discrepancies is most likely an imbalance of radiative heating.

Two significant sudden changes are detected in the differences between the eddy heat fluxes derived from the two reanalyses. One of the changes may be linked to the bias corrections applied to the infrared radiances from the *NOAA-12* High-Resolution Infrared Radiation Sounder in ERA-40, which is known to be contaminated by volcanic aerosol from the 1991 eruption of Mt. Pinatubo. The other change may be due in part to the use of uncorrected radiances from the *NOAA-15* Advanced Microwave Sounding Units by ERA-Interim since 1998. These sudden changes have the potential to alter the wave forcing trends in the affected reanalysis, suggesting that extreme care is needed when one comes to extract trends from the highly derived wave forcing quantities.

1. Introduction

The equator to pole circulation in the winter stratosphere is primarily driven by the upward propagating waves from the troposphere. This large-scale dynamical process is called the Brewer–Dobson (B–D) circulation and can be studied using the transformed Eulerian mean (TEM) equations (Edmon et al. 1980; Andrews et al. 1987; Holton et al. 1995; Shepherd 2007; Birner and Bonisch 2011). The Eliassen–Palm (E–P) flux divergence, which represents the wave forcing that acts on the mean flow to cause wind and temperature variations, is the most important quantity in the TEM equations. Numerous studies have used the TEM equations together with the E–P flux divergence to study the interannual variation and long-term trends of the B–D circulation (Edmon et al. 1980; Seviour et al. 2012), the behavior of planetary wave activity (Hu and Tung 2002; Karpetchko and Nikulin

2004; Hu et al. 2005), the variability of the polar vortex (Vaugh et al. 1999; Newman et al. 2001), the momentum balance of the stratosphere (Dima and Wallace 2007; Monier and Weare 2011), and the annual cycle in tropical tropopause temperature (Kerr-Munslow and Norton 2006; Randel et al. 2008; Randel and Jensen 2013), among other topics. The fidelity of these studies relies crucially on the accuracy and homogeneity of the datasets that are used to derive the E–P flux divergence and the residual circulation that approximates the B–D circulation.

The E–P flux divergence is a highly derived quantity. Its calculation requires nonlocal information such as the spatial and temporal departures of the primary variables (i.e., winds and temperatures) from their mean fields. It is therefore extremely difficult to estimate the E–P flux divergence directly using station-based measurements. In addition, the calculation involves not only estimates of high-frequency fluctuation of the wave fluxes at different altitudes and latitudes but also differential operators that are applied to the slowly varying background temperature gradient. All these complications can potentially cause biases in the climatology, interannual variability, and/or long-term trends of E–P flux divergence. The

Corresponding author address: Hua Lu, British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, United Kingdom.
E-mail: hlu@bas.ac.uk

nonlocal and nonlinear operators may also amplify small errors that are associated with the primary variables to much larger errors in the E-P flux divergence. It is therefore important to gauge the uncertainties in estimating the E-P flux divergence.

The most commonly used tools to derive the E-P flux divergence are reanalysis datasets, which are normally constructed by a variety of observations that are assimilated by using numerical weather prediction models to give a coherent representation of the global atmosphere with uniform spatial and temporal coverage (Uppala et al. 2005; Dee and Uppala 2009). A major concern with the use of reanalyses is their accuracy and homogeneity in representing both the underlying dynamics and long-term trends (e.g., Sterl 2004; Bengtsson et al. 2007; Thorne and Vose 2010). In particular, regions with relatively large analysis increments (defined as the reanalysis minus the model first guess that is based on the 6-hourly model forecast) can induce errors in estimating radiative balance and temperature (Uppala et al. 2005; Dee and Uppala 2008, 2009). In addition to model errors and drifts, studies have also shown that reanalysis datasets tend to differ from each other, especially in regard to long-term trends (e.g., Bengtsson et al. 2007; Kobayashi et al. 2009). This is because low-frequency and trend uncertainties may be induced by observational errors, including instrument biases and changes in geographical coverage. Sudden changes induced by incorporating newly available radiance measurements are of a particular concern in causing biases in low-frequency variation (Simmons et al. 2014).

ERA-40 and ERA-Interim, the two major consecutive reanalysis datasets produced by the European Centre for Medium-Range Weather Forecasts (ECMWF), have been widely used for the study of atmospheric circulation and processes (Dee and Uppala 2009; Uppala et al. 2005). ERA-Interim, the newest reanalysis product of ECMWF, is known to have many improvements over ERA-40 (Dee and Uppala 2008, 2009; Dee et al. 2011a). It has much smaller analysis increments during winter at high latitudes, more realistic temperature trends and radiative budget, and more reliable low-frequency variability (Dee and Uppala 2009; Dee et al. 2011a; Screen and Simmonds 2011; Bracegirdle and Marshall 2012; Cornes and Jones 2013; Simmons et al. 2014). It also has better representations of the hydrological cycle in the tropics and subtropics and a more realistic B-D circulation in the stratosphere (Schoeberl et al. 2003; van Noije et al. 2004; Monge-Sanz et al. 2007, 2013; Dee et al. 2011b). Studies have yet to be undertaken to evaluate how the improvement may have affected the wave forcing estimates. Because it is extremely difficult to compare the wave forcing estimates directly against the observations,

a comparative study may provide some insights into the uncertainties of estimating wave forcing based on reanalysis datasets.

This study undertakes a comparative study between ERA-40 and ERA-Interim to quantify the discrepancies in wave forcing, measured by the E-P flux divergence and the associated wave fluxes. We choose to compare these two ECMWF reanalyses mainly because of the well-documented improvements of ERA-Interim over ERA-40; these help in diagnosing the possible causes of the discrepancies. Our focus is on the height region from the upper troposphere to the upper stratosphere (500–1 hPa), where the zonal mean wave forcing is the main driver of the large-scale circulation, and the Northern Hemisphere (NH) winter mean of December–February (DJF), when both the wave amplitude and variability are largest. We first detect the regions with the largest E-P flux divergence discrepancies and identify the key wave fluxes that contribute the most to them. We then examine to what extent the E-P flux divergence discrepancies are linked to discrepancies in the residual circulation. Finally, we apply a changepoint detection method called the penalized maximal t test (PMT) to investigate the temporal consistency of the poleward eddy heat flux $\overline{v'T'}$ in these reanalyses.

2. Data and methods

a. Data

The 40-yr ECMWF Re-Analysis (ERA-40) was generated by using the ECMWF Integrated Forecast System (IFS) model and its 6-hourly three-dimensional variational data assimilation (3D-Var) system (Uppala et al. 2005). It covered the period from September 1957 to August 2002 and incorporated observations from in situ measurements, including balloons, radiosondes, dropsondes, aircraft, and ships, along with satellite observations, which only provided global coverage of radiance measurements from 1979 onward. The data ingestion involved approximately $7\text{--}9 \times 10^6$ observations at each time step. The assimilation model used had a spectral T159 grid, corresponding to a 1.125° grid spacing in latitude and longitude and 60 levels in the vertical between the surface and 0.1 hPa (~ 65 km). Analysis products on the 23 standard pressure surfaces from 1000 to 1 hPa are available for general use.

Covering the data-rich satellite era of 1979–present the Interim ECMWF Re-Analysis (ERA-Interim) is the ECMWF's current comprehensive atmospheric reanalysis (Dee and Uppala 2009; Dee et al. 2011a). It makes use of the same observations as ERA-40 before September 2002, supplemented with ECMWF operational data afterward (Berrisford et al. 2011; Simmons et al. 2014) but with major

improvements over ERA-40. Especially, the ECMWF's operational four-dimensional variational data assimilation (4D-Var) system couples the dynamic variables more cohesively with the humidity and radiation than its previous 3D-Var analysis system. This ensures a realistic interaction of temperature, vertical velocity, and humidity both temporally and spatially. Improved correction of biases in satellite radiance data is also achieved through the use of an automated variational bias correction system that optimizes the consistency of multiple measurements (Dee 2005; Dee and Uppala 2009; Dee et al. 2011a). In addition, the ERA-Interim assimilation model has a spectral T255 grid, corresponding to a $\sim 0.70^\circ$ grid spacing in latitude and longitude. It represents a higher spatial resolution than ERA-40; hence smaller-scale waves are resolved explicitly. The increase in spatial resolution is one of the key factors contributing to the reduction of analysis increments of temperatures as well as to a more realistic representation of the B-D circulation, in addition to many other improvements, including better physical parameterization schemes for radiative transfer, data quality control, subgrid-scale orographic drag, humidity analysis, clouds, and surface/soil processes (Dee and Uppala 2009; Dee et al. 2011a). The ERA-Interim assimilation model uses the same vertical levels as ERA-40 but the data are made available at 37 levels between 1000 and 1 hPa, including the standard 23 levels used by ERA-40.

Our analysis is based on the overlapping 22 winters (i.e., the winters of 1979/80–2001/02) that are shared by both ERA-40 and ERA-Interim. For clarity and simplicity, the definition of a winter is based on January across this paper; for example, the DJF mean of the 1979/80 winter is numbered and stated as 1980 hereafter.

b. TEM equation and the E-P flux divergence

The momentum balance in the TEM framework provides a theoretical account of large-scale dynamics by linking the mean flow acceleration to the residual circulation and large-scale wave forcing (Andrews et al. 1987). In spherical coordinates, it is expressed as

$$\frac{\partial \bar{u}}{\partial t} = - \left[\frac{(\bar{u} \cos \phi)_\phi}{a \cos \phi} - f \right] \bar{v}^* - \bar{u}_z \bar{w}^* + \frac{\mathbf{V} \cdot \mathbf{F}}{\rho_0 a \cos \phi} + \bar{X}, \quad (1)$$

where u , v , and w are Eulerian zonal, meridional, and vertical winds, a is the mean radius of Earth, ϕ is latitude, $\rho_0 = \rho_s \exp(-z/H)$ is the standard density in log-pressure coordinates, ρ_s is the sea level reference density, z is the log-pressure height coordinate, H is the mean scale height ($=7$ km), f is the Coriolis parameter, the overbar denotes zonal average, and subscripts denote the derivatives with

respect to the given variable. In Eq. (1), \bar{v}^* and \bar{w}^* represent the residual mean meridional and vertical winds; they can be expressed as

$$\bar{v}^* = \bar{v} - \frac{1}{\rho_0} \left(\rho_0 \frac{\overline{v'\theta'}}{\bar{\theta}_z} \right)_z \quad \text{and} \quad (2)$$

$$\bar{w}^* = \bar{w} + \frac{1}{a \cos \phi} \left(\cos \phi \frac{\overline{v'\theta'}}{\bar{\theta}_z} \right)_\phi, \quad (3)$$

where θ is potential temperature and primes denote the departure from zonal mean. The term $\mathbf{V} \cdot \mathbf{F}$ on the right-hand side of Eq. (1) is the E-P flux divergence and \bar{X} represents other nonconservative mechanical forcing, such as parameterized subgrid processes including gravity wave drag. Equation (1) states that the acceleration or deceleration of zonal mean zonal wind \bar{u} is affected by the residual mean meridional circulation (the sum of the first and second terms on the right-hand side, which is denoted as Θ hereafter), the resolved or large-scale wave forcing that drives the circulation to depart from its radiative equilibrium (the third term on the right-hand side, which is denoted as Ψ hereafter), and the contribution from other nonconservative processes (the \bar{X} term). In this context, a significant difference in Ψ signifies inconsistency of wave forcing between these two datasets while a significant difference in Θ suggests a different behavior in the B-D circulation. The E-P flux divergence $\mathbf{V} \cdot \mathbf{F}$ that is the key to estimating Ψ can be further expanded into

$$\mathbf{V} \cdot \mathbf{F} = \frac{1}{a \cos \phi} [F^{(\phi)} \cos \phi]_\phi + F_z^{(z)}, \quad (4)$$

where the meridional and vertical components of the wave forcing can be calculated as

$$F^{(\phi)} = \rho_0 a \cos \phi \left(\frac{\overline{v'\theta'}}{\bar{\theta}_z} \bar{u}_z - \overline{v'u'} \right) \quad \text{and} \quad (5)$$

$$F_z^{(z)} = \rho_0 a \cos \phi \left\{ \left[f - \frac{(\bar{u} \cos \phi)_\phi}{a \cos \phi} \right] \frac{\overline{v'\theta'}}{\bar{\theta}_z} - \overline{w'u'} \right\}. \quad (6)$$

Equation (1) is assembled in this form so that the net effect of the wave forcing on the mean flow can be quantified. Its individual terms, however, may show contrasting or opposite behaviors (Edmon et al. 1980; Palmer 1981). Here, to identify the key flux terms that contribute most to the total wave forcing discrepancies,

we not only analyze the total E-P flux divergence term Ψ but also look into the individual contributing terms separately. In the latter case, we effectively employ an Eulerian approach by expanding the total wave forcing term Ψ into five additive terms according to Eqs. (4)–(6). The five terms are $\Psi_1 = (1/a \cos^2 \phi)[(\overline{v'\theta'}/\theta_z)\overline{u}_z \cos^2 \phi]_\phi$, $\Psi_2 = -(1/a \cos^2 \phi)(\overline{v'u'} \cos^2 \phi)_\phi$, $\Psi_3 = (f/\rho_0)[\rho_0(\overline{v'\theta'}/\theta_z)]_z$, $\Psi_4 = -(1/\rho_0 a \cos \phi)[\rho_0(\overline{v'\theta'}/\theta_z)(\overline{u} \cos \phi)_\phi]_z$, and $\Psi_5 = -(1/\rho_0)(\rho_0 \overline{w'u'})_z$. Theoretically, Ψ_2 and Ψ_3 should be the dominant terms that contribute to the total Ψ in the extratropics according to the quasigeostrophic approximation (Andrews et al. 1987). In the extratropical lower stratosphere where the wave forcing is primarily dominated by the vertical propagation of planetary waves from the troposphere, Ψ_3 is central to the total wave forcing calculation (Newman and Nash 2000). Near the tropics or in the regions where plane-parallel gravity waves are present, the contribution from the vertical momentum flux term Ψ_5 may also play an important role (Andrews et al. 1987).

All the wave forcing quantities are calculated using data archived at $2.5^\circ \times 2.5^\circ$ grid spacing and at the 23 pressure levels that are common to both reanalyses. As a result, the wave forcing estimated from this coarse resolution should primarily be dominated by the effect of large-scale Rossby waves. The derivatives involved in the E-P flux divergence and other quantities in Eq. (1) are all calculated using centered differences except for those at the top and bottom boundaries (i.e., 1000 and 1 hPa), where one-sided differences were employed. As such, the results at the boundaries are less reliable. In addition, all the calculations are performed on daily mean winds and temperatures and then averaged over the DJF season. We chose to use daily averages rather than the 6-hourly instantaneous records because the very high-frequency waves such as diurnal tides should make a negligible contribution toward the wave driving B-D circulation.

c. Diagnostic tools

ERA-40 and ERA-Interim describe the same circulation of Earth's atmosphere. Ideally, there should be no difference between them in all the wave forcing quantities and in the residual circulation term Θ . In reality, the reanalysis datasets differ from each other due to the dissimilarity in bias correction, physical parameterization, model resolution, and/or the ways of assimilating observations. The wave forcing as well as the circulation parameters therefore differ as a consequence of these sources of dissimilarity. We use composite analysis with a two-sided Student's t test to diagnose regions with significant differences in their climatological means across their common period (i.e., 1980–2002). The composite

differences are all performed as ERA-40 minus ERA-Interim and denoted as ERA40 – ERAInt hereafter.

We apply the penalized maximal t test (Wang et al. 2007) to detect a significant sudden shift of mean in the wave forcing differences between the two reanalyses. A brief description of the method can be found in the appendix. To examine the principal contributor to the discontinuity, the PMT identification is separately applied to the total, stationary, and transient components of the wave forcing. This is because stationary waves are excited by the topography as well as land–sea heating contrast while transient waves are dominated by synoptic-scale weather patterns (Newman and Nash 2000). At each grid point, the total eddy heat flux $v'T'_{\text{total}}$ is calculated by multiplying the daily zonal departures of meridional wind v and temperature T (i.e., v' and T') and averaging the multiplied quantity over DJF. To obtain the stationary component $v'T'_{\text{stationary}}$, we first average DJF meridional wind v and temperature T at each grid point and then multiply the zonal departures of the seasonally averaged quantities. The transient component is estimated simply by $v'T'_{\text{transient}} = v'T'_{\text{total}} - v'T'_{\text{stationary}}$. These three components are then zonally averaged in order to obtain their zonal mean fields. Also, when a winter is found to contain a significant sudden jump (i.e., a changepoint), composite analysis based on the detected changepoint winter is then used to investigate the spatial characteristics of the discontinuity. It is worth noting that the results reported here are case studies that demonstrate the usefulness of the detection technique rather than exhausting all possible discontinuities in both datasets.

3. Results

a. Discrepancies in E-P flux divergence

Figures 1a and 1b show the climatology of DJF mean E-P fluxes (arrows) and E-P flux divergence term Ψ (contours) estimated from ERA-40 and ERA-Interim respectively. Both climatologies show that the wave forcing is marked by the upward and equatorward propagation of the E-P fluxes that are associated with the mainly negative E-P flux divergence term Ψ . There are two distinct peak regions of Ψ , one in the upper troposphere [$\sim(200\text{--}300)$ hPa] and another in the upper stratosphere [$\sim(1\text{--}3)$ hPa]. Another smaller peak can also be observed at high latitudes around 5–10 hPa.

Figure 1c shows the composite difference in the DJF mean E-P fluxes and E-P flux divergence term Ψ between the two reanalyses. The main feature of $\Delta\Psi_{\text{ERA40-ERAInt}}$ is the vertically alternating positive and negative anomalies in the extratropics, which intensify and expand more toward the equator with increasing altitude. As a result, the largest $\Delta\Psi_{\text{ERA40-ERAInt}}$ appears in the upper stratosphere,

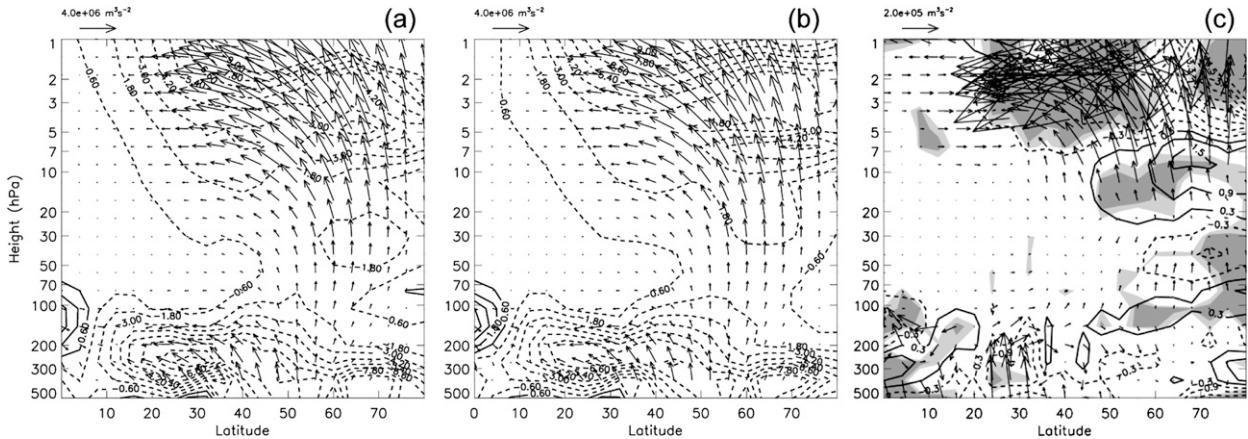


FIG. 1. Latitude–height cross section of the DJF mean E-P flux (arrows, $4 \times 10^6 \text{ m}^3 \text{ s}^{-2}$) and E-P flux divergence term Ψ (contours). Climatology from (a) ERA-40 and (b) ERA-Interim; and (c) composite difference (ERA40 – ERA-Interim). Solid (dashed) contours are positive (negative) divergence at intervals of $\pm 0.6, \pm 1.8, \pm 3, \pm 4.2, \dots \text{ m s}^{-1} \text{ day}^{-1}$ in (a),(b) and $\pm 0.3, \pm 0.9, \pm 1.5, \pm 2.1, \dots \text{ m s}^{-1} \text{ day}^{-1}$ in (c). The light (dark) shaded areas in (c) represent $p \leq 0.1$ (0.05), estimated by two-sided Student's t test. The plotted E-P flux vectors are shown as $[\alpha F^{(\phi)}, F^{(z)}]/\rho_0$, with the scaling factor $\alpha = 4.899 \times 10^{-3}$ used to adjust for the exaggerated vertical scale in the plot and a further division by the air density ρ_0 for a clearer visualization of the directions of wave propagation. The same scaling applies to the E-P flux differences.

where the negative Ψ anomalies cover poleward of 20°N . It is worth noting that the magnitudes of $\Delta\Psi_{\text{ERA40-ERA-Interim}}$ are as large as 20%–40% of the climatological Ψ in this region. Anomalous upward E-P flux vectors are found in the lower and upper stratosphere, suggesting an overall stronger wave forcing in ERA-40 than ERA-Interim in the stratosphere.

Figure 2 shows the time series of DJF mean total E-P flux divergence term Ψ that are area-averaged over $45^\circ\text{--}75^\circ\text{N}$ at 3, 10, 50, and 100 hPa (top–bottom). At 3 hPa, noticeable discrepancies in both interannual variability and climatological mean can be observed with more negative Ψ values in ERA-40 than ERA-Interim. At 10 hPa, the difference is due mainly to the climatological mean with the ERA-Interim Ψ being more negative overall than that of ERA-40. At 50 hPa, a generally similar behavior to that at 3 hPa can be seen though the magnitude of the discrepancy is relatively smaller. At 100 hPa, the discrepancy is again dominated by a difference in the climatological mean with the ERA-40 Ψ being less negative than that of ERA-Interim. Over these four pressure levels, the climatological means of Ψ estimated from ERA-40 and ERA-Interim alternately exceed each other. The discrepancies are comparable to 15% of the interannual variability of Ψ at 10 hPa; this value increases to 45% at 100 hPa. There are also apparent trends in $\Psi_{45\text{--}75\text{N}}$, especially at 100 hPa where upward trends are clearly noticeable in both ERA-40 and ERA-Interim estimates, and the trend of ERA-40 $\Psi_{45\text{--}75\text{N},100 \text{ hPa}}$ is distinctly steeper than that of ERA-Interim $\Psi_{45\text{--}75\text{N},100 \text{ hPa}}$.

Figure 3 shows the composite differences of four of the individual terms that add up to the differences of the total wave forcing term Ψ . Because the climatology of Ψ_1 is one order of magnitude smaller than the those of the other terms and no significant differences between the two reanalyses are detectable for Ψ_1 , the difference plot of Ψ_1 is not shown. It is immediately clear that Ψ_3 is the main contributor to the vertically alternating positive and negative Ψ anomalies shown in Fig. 1c. In the upper stratosphere, Ψ_2 and Ψ_4 also play a role in addition to Ψ_3 . At high latitudes, the Ψ_2 discrepancies have an opposite sign to those of Ψ_3 while the Ψ_4 discrepancies have the same sign as those of Ψ_3 . The combined effect of Ψ_3 , Ψ_2 , and Ψ_4 forms the negative $\Delta\Psi_{\text{ERA40-ERA-Interim}}$ in the high-latitude upper stratosphere. At midlatitudes (i.e., $20^\circ\text{--}45^\circ\text{N}$), Ψ_2 plays a major role in causing the wave forcing discrepancies.

In the middle-to-low latitude upper troposphere ($0^\circ\text{--}50^\circ\text{N}$, 200–500 hPa), Ψ_3 and Ψ_5 are the main contributors to $\Delta\Psi_{\text{ERA40-ERA-Interim}}$. At low latitudes, the discrepancies are dominated by the effects of Ψ_5 , which are marked by the vertically alternating negative and positive anomalies that are very similar to those of Ψ_3 in the extratropics. These tropical Ψ_5 discrepancies are associated with the vertical momentum flux $\overline{w'u'}$ to which Ψ_5 is negatively proportional. In the tropical and subtropical tropopause, the Ψ_2 term also contributes to Ψ discrepancies mainly by enhancing the Ψ_5 anomalies. In the midlatitude upper troposphere [$\sim(25^\circ\text{--}50^\circ\text{N})$, 300 hPa], significant discrepancies are found in both Ψ_3 and Ψ_5 , with positive Ψ_3 differences partially counteracting the negative Ψ_5

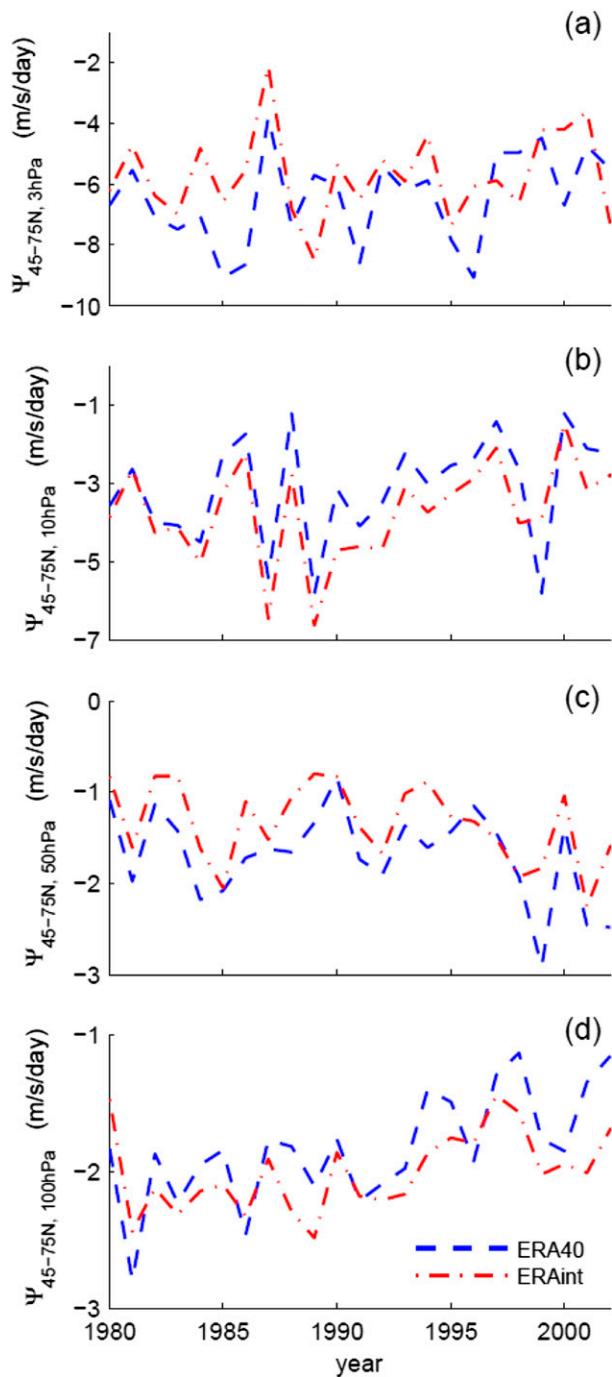


FIG. 2. Time series of DJF mean E-P flux divergence term Ψ estimated from ERA-40 (blue dashed) and ERA-Interim (red dash-dotted) area-averaged over 45° – 75° N and at (a) 3, (b) 10, (c) 50, and (d) 100 hPa.

differences. Their combined effect is insignificant negative $\Delta\Psi_{\text{ERA40-ERAInt}}$ in this region.

The poleward eddy potential heat flux $\overline{v'\theta'}$ is the most important quantity that is used to estimate Ψ_3 in the

middle-to-high latitude stratosphere. To examine the extent to which $\overline{v'\theta'}$ contributes to the high-latitude $\Delta\Psi_{\text{ERA40-ERAInt}}$, Fig. 4 shows the climatology of $\overline{v'\theta'}$ estimated from ERA-Interim as well as the $\overline{v'\theta'}$ composite difference between the two datasets, $\Delta\overline{v'\theta'}_{\text{ERA40-ERAInt}}$. The climatological $\overline{v'\theta'}$ increases with height, with larger values in the middle-to-high latitudes. Note that $\Delta\overline{v'\theta'}_{\text{ERA40-ERAInt}}$ is also mostly positive and exhibits larger values in the extratropical upper stratosphere, where $\Delta\overline{v'\theta'}_{\text{ERA40-ERAInt}}$ accounts for $\sim 10\%$ of its climatology. Vertically alternating positive and negative $\Delta\overline{v'\theta'}_{\text{ERA40-ERAInt}}$ anomalies are found at high latitudes although they are not statistically significant. The term $\Delta\overline{v'\theta'}_{\text{ERA40-ERAInt}}$ is statistically significant mainly below 200 hPa and away from the high latitudes. These results suggest that the alternating positive and negative $\Delta\Psi_{\text{ERA40-ERAInt}}$ shown in Figs. 1c and 3b cannot be explained by the differences in $\overline{v'\theta'}$ alone.

Figure 5 elaborates this point further by showing the temporal variation of poleward eddy heat flux $\overline{v'T'}$ and its long-term trends at 100 and 10 hPa respectively. It is noted that, considering each pressure level in isolation, $\overline{v'T'}$ is proportional to the poleward eddy potential heat flux $\overline{v'\theta'}$, so similar behavior would also be seen in $\overline{v'\theta'}$. In general, both datasets follow each other exceedingly well in terms of interannual variability; this holds true for the total, stationary, and transient components of $\overline{v'T'}$ both at 100 and 10 hPa. No significant trends are observed in the total $\overline{v'T'}_{45-75\text{N},100\text{hPa}}$ either in ERA-40 or ERA-Interim estimates, although there is a noticeable difference in the climatological mean in the total ERA-40 $\overline{v'T'}_{45-75\text{N},100\text{hPa}}$ estimates. However, an upward trend is shown in the stationary $\overline{v'T'}_{45-75\text{N},100\text{hPa}}$ while a downward trend is associated with the transient component, with the ERA-40 trends being generally steeper than those of ERA-Interim. Similar positive and negative trends in stationary and transient $\overline{v'T'}_{45-75\text{N},100\text{hPa}}$ are observable at 10 hPa, except that at this level the ERA-Interim trend is slightly steeper than that of ERA-40. Nevertheless, we find that the stationary and transient components of $\overline{v'T'}_{45-75\text{N}}$ show consistent trends throughout the stratosphere, in contrast to the rather confusing trend behavior of $\Psi_{45-75\text{N}}$ in the stratosphere (see Fig. 2). These results suggest that the two datasets agree with each other better for $\overline{v'T'}$ than for Ψ in the extratropical stratosphere. They also imply that something other than the eddy heat flux $\overline{v'T'}$ is responsible for the discrepancies in $\Psi_{45-75\text{N}}$.

The eddy heat fluxes $\overline{v'\theta'}$ and $\overline{v'T'}$ are not responsible for the vertically alternating feature of Ψ discrepancies, but Fig. 3b indicates that Ψ_3 is the main contributor to $\Delta\Psi_{\text{ERA40-ERAInt}}$. The other possible cause is the vertical gradient of the potential temperature θ_z . Figure 6 shows the latitude–height plane of the DJF mean of the vertical

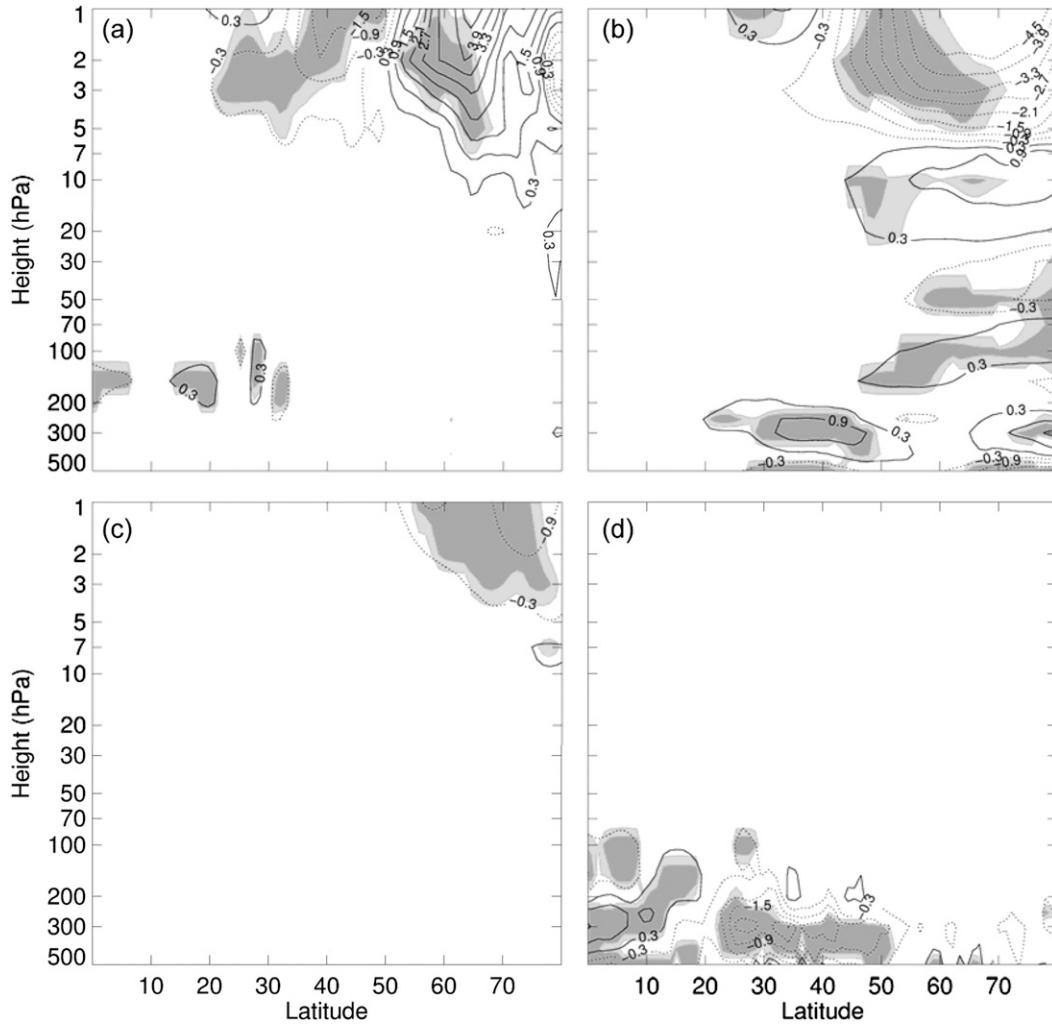


FIG. 3. Composite differences (ERA40 – ERAInt) in the DJF E-P flux divergence terms (a) Ψ_2 , (b) Ψ_3 , (c) Ψ_4 , and (d) Ψ_5 . The contours and shadings are the same as in Fig. 1c with the exception that the statistical significance shading is omitted in regions of small differences and minor dynamical significance ($<0.3 \text{ m s}^{-1} \text{ day}^{-1}$ in magnitude).

gradient of potential temperature θ_z , $\Psi_{\theta_z} = (f/\rho_0)(\rho_0/\theta_z)_z$ and $\Psi_{3,\text{Pseudo}} = (f/\rho_0)[\rho_0(\overline{v'\theta'_{\text{Clim}}}/\theta_z)]_z$. Note that $\Psi_{3,\text{Pseudo}}$ is the same as Ψ_3 except for $\overline{v'\theta'}$ being fixed as a constant that is equal to the ERA-Interim climatology. Vertically alternating anomalies are clearly noticeable in all three variables. The discrepancies in θ_z and Ψ_{θ_z} are found not only at high latitudes but also at low latitudes while the discrepancies in $\Psi_{3,\text{Pseudo}}$ are mostly confined to the middle to high latitudes. Wright and Fueglistaler (2013) recently found that net diabatic heating directly above the tropical convective regions is noticeably stronger in ERA-Interim than other reanalyses; this may be linked to the negative $\Delta\theta_z$ at 70–100 hPa and positive $\Delta\theta_z$ at 150–300 hPa. However, comparing the discrepancies in $\Psi_{3,\text{Pseudo}}$ with those in Ψ_3 (Fig. 3b), the magnitude of the $\Psi_{3,\text{Pseudo}}$ discrepancies is at most half of those of Ψ_3 in the

lower-to-middle stratosphere and differences of opposite sign are found near 1 hPa. These results suggest that differences in the vertical temperature gradient θ_z between these two datasets play the most important role in explaining the vertically alternating positive and negative anomalies of Ψ_3 and those of Ψ at high latitudes. The anomalies in the eddy heat fluxes are nevertheless not negligible in terms of their magnitudes; their contribution may be comparable to that from static stability in the upper stratosphere. Nonlinear interaction between these two may also play a role.

Figure 7 shows the vertical profile of DJF zonal mean temperature climatology \overline{T} (Fig. 7a) and differences $\Delta\overline{T}_{\text{ERA40-ERAInt}}$ (Fig. 7b) for the 0° – 35°N , 35° – 75°N , and 75° – 90°N latitude bands. In general, \overline{T} decreases from the surface to the tropopause and then

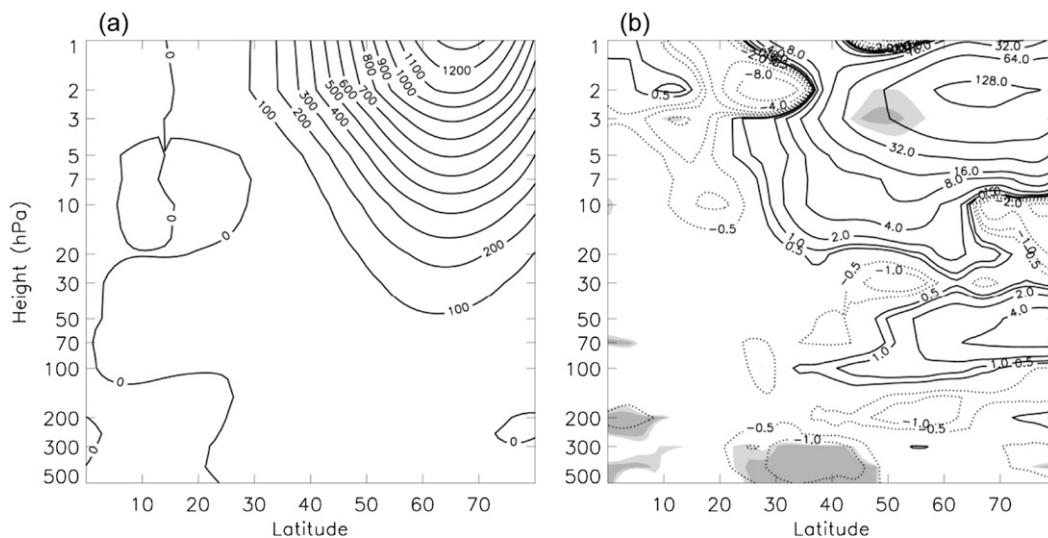


FIG. 4. Latitude–height cross section of the DJF mean of the potential heat flux $\overline{v'\theta'}$. (a) Climatology from ERA-Interim; (b) composite difference (ERA40 – ERA-Int). Solid and dashed contours indicate positive and negative values and the shadings in (b) are as in Fig. 1c.

increases from the tropopause to the upper stratosphere. Also, the vertical temperature gradient decreases with latitude with temperature gradient being the steepest at 0° – 35° N. For all three latitude bands, the magnitude of $\Delta\overline{T}_{\text{ERA40-ERAInt}}$ is relatively small (<1.5 K) below 10 hPa, but it increases sharply above 10 hPa (to ~ 5 K). Vertically alternating positive and negative $\Delta\overline{T}_{\text{ERA40-ERAInt}}$ anomalies are clearly visible at low and high latitudes; the effect is less clear for the midlatitude band below 10 hPa.

Figure 8 shows NH polar plots of DJF mean temperature differences $\Delta T_{\text{ERA40-ERAInt}}$ at various pressure levels. Two common features that are found at all the pressure levels except for 850 hPa are that 1) at a given pressure level, $\Delta T_{\text{ERA40-ERAInt}}$ tends to have the same sign hemispherically, and 2) significant $\Delta T_{\text{ERA40-ERAInt}}$ are mostly confined to the low and high latitudes with little $\Delta T_{\text{ERA40-ERAInt}}$ signal visible at midlatitudes. At 7 and 20 hPa, the low-latitude $\Delta T_{\text{ERA40-ERAInt}}$ signal peaks over the Pacific and Atlantic Oceans where the signal extends more northward. At 70 hPa, the $\Delta T_{\text{ERA40-ERAInt}}$ signal is relatively small and spatially patchy. At 100 hPa, the pattern of $\Delta T_{\text{ERA40-ERAInt}}$ is broadly similar to that at 20 hPa though it is more zonal and more confined to the tropics. At 500 hPa, there is a lack of significant $\Delta T_{\text{ERA40-ERAInt}}$ over most of the Pacific and relatively weaker $\Delta T_{\text{ERA40-ERAInt}}$ over the North Atlantic. At 850 hPa, the significant differences are found mainly over the two ocean basins and near the tropics. These results suggest that the temperature differences are zonally symmetric at some levels (i.e., 7 and 100 hPa) and asymmetric at other pressure levels. More importantly, they show that $\Delta T_{\text{ERA40-ERAInt}}$ over the Pacific and Atlantic

Oceans contributes most toward the vertical zigzag behavior of zonal mean temperature gradient difference $\Delta\overline{T}_{\text{ERA40-ERAInt}}$ in the low latitudes.

As well as showing significant discrepancies in the extratropical stratosphere, Fig. 3 also shows significant discrepancies in the upper troposphere. To illustrate the temporal variation of these tropospheric discrepancies, Fig. 9 shows the time series of DJF mean ERA-40 and ERA-Interim Ψ_3 and Ψ_5 area-averaged over 25° – 50° N at 300 hPa. The discrepancies in both Ψ_3 and Ψ_5 are due mainly to a difference of climatological mean and the magnitude of the Ψ_3 discrepancy is larger than its interannual variability. Also, the interannual variability of Ψ_5 is much larger than that of Ψ_3 ; Ψ_5 may play a dominant role in the total waving forcing for a particular winter such as 1989 in this region.

Figure 3 shows that Ψ_5 is also responsible for the total wave forcing discrepancies in the low-latitude upper troposphere. To illustrate the temporal variation of these discrepancies, Fig. 10 shows the time series of ERA-40 and ERA-Interim Ψ_5 that are area-averaged over 0° – 10° N at 300 hPa (left) and 0° – 10° N at 100 hPa (right). The discrepancies are again dominated by a difference in climatological mean. The climatological difference in Ψ_5 at 300 hPa is again larger than its interannual variability, implying that there is large uncertainty associated with the momentum budget in this region. Apart from the dominant climatological mean difference, there are also noticeable disagreements in the interannual variability in Ψ_5 at 300 hPa. It is noted that ERA-Interim Ψ_5 departs farther away from the zero line than ERA-40 Ψ_5 at both 100 and 300 hPa,

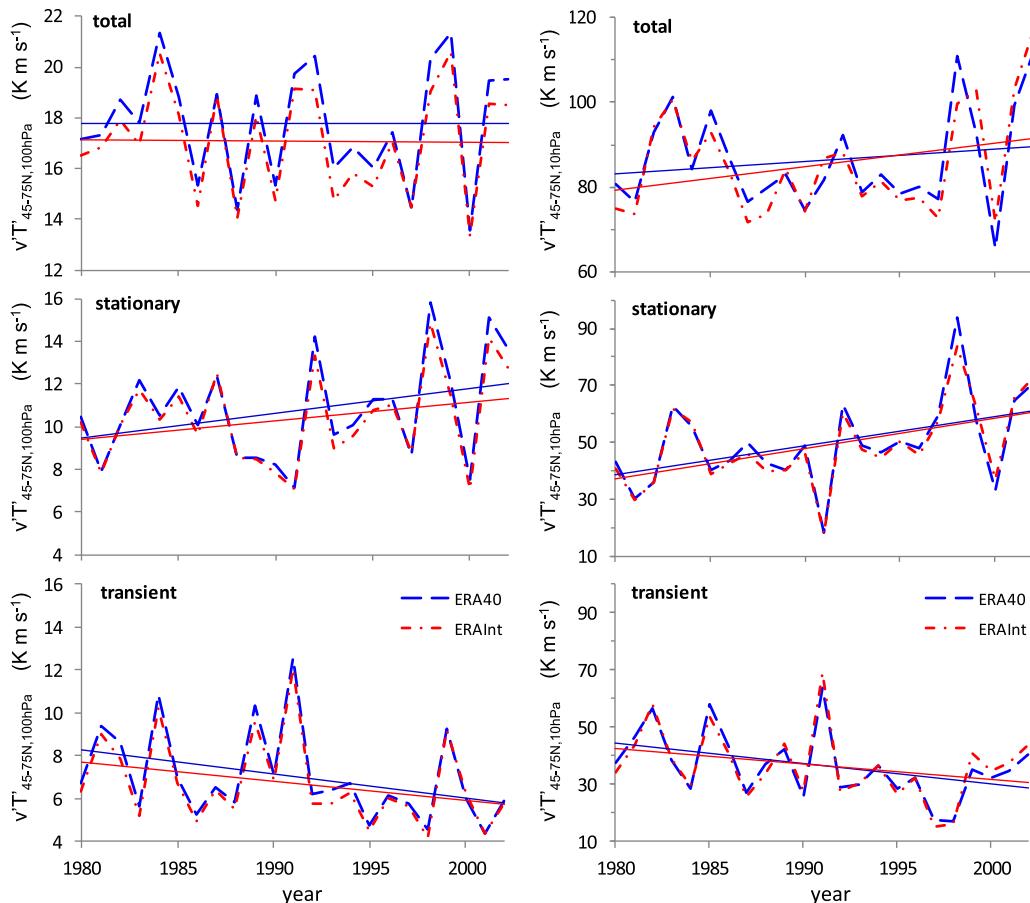


FIG. 5. Time series of DJF mean (top) total, (middle) stationary, and (bottom) transient components of the zonal mean eddy heat flux $\overline{v'T'}$ averaged over 45° – 75° N at (left) 100 and (right) 10 hPa. The solid straight lines are the associated trends in the ERA-40 (blue) and ERA-Interim (red) eddy heat flux components.

implying a larger magnitude of the vertical eddy flux $\overline{w'u'}$ in ERA-Interim than ERA-40. Small-scale processes such as gravity waves play an important role in $\overline{w'u'}$ (Lindzen 1981), and differences in model resolution and parameterization are the likely sources for the discrepancies. It has been shown that the vertical velocity of ERA-Interim is less noisy than that of ERA-40 (Dee and Uppala 2008; Iwasaki et al. 2009). This may also contribute to the larger magnitude of $\overline{w'u'}$ (or Ψ_5) in ERA-Interim than ERA-40.

b. Effect on the B-D circulation

This section investigates the extent to which the resolved wave forcing term Ψ is linked to the discrepancies in the B-D circulation by examining the momentum budget of the TEM equation. The first row of Fig. 11 shows the climatology of the DJF mean residual mean meridional circulation (\overline{v}^* , \overline{w}^*) (arrows) and the residual mean meridional circulation term Θ (contours) estimated from ERA-40 and ERA-Interim (Figs. 11a,b),

as well as the composite differences between these two datasets (Fig. 11c). The main climatological feature of the residual mean meridional circulation in both ERA-40 and ERA-Interim is the upward movement of streamlines of the flow at low latitudes that is followed by poleward movement at midlatitudes and finally downward movement at high latitudes. The residual meridional circulation term Θ is mainly positive, reflecting eastward (or westerly) acceleration and a predominantly northward apparent force on the fluid parcels. In the stratosphere, Θ peaks in the extratropical upper stratosphere and it is in approximate balance with the Ψ peak in the same region (see Figs. 1a,b). The tropospheric Θ peaks poleward of 40° N where Θ is also in rough balance with Ψ . However, Θ is not in balance with Ψ near the tropospheric subtropical jet, where gravity wave drag and upgradient eddy transport (McFarlane 1987; Birner et al. 2013) play an important role. In the TEM formulation [Eq. (1)], the effect of these processes is accounted for by the nonconservative

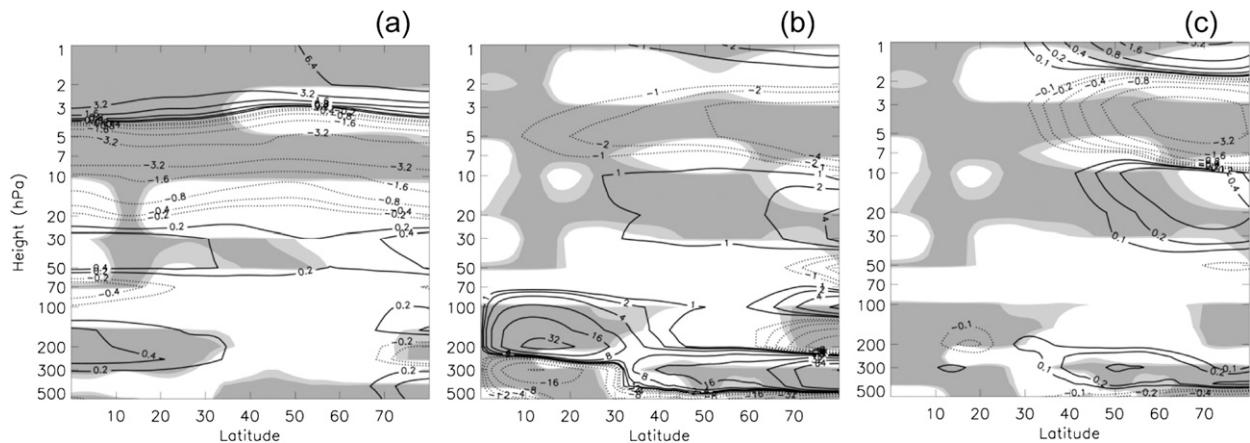


FIG. 6. Latitude–height cross section of composite differences of the DJF mean of the vertical gradient of potential temperature (a) θ_z , (b) $\Psi_{\theta_z} = (f/\rho_0)(\rho_0/\theta_z)_z$, and (c) $\Psi_{3,Pseudo} = (f/\rho_0)[\rho_0(v'\theta'_{Clim}/\theta_z)]_z$, where $\Psi_{3,Pseudo}$ is the same as Ψ_3 except for $v'\theta'$ being fixed as a constant that is equal to the ERA-Interim climatology. Solid and dashed contours are positive and negative. The actual contours are $\pm 0.0002, \pm 0.0004, \pm 0.0008, \dots \text{K m}^{-1}$ for θ_z and $\pm 0.001, \pm 0.002, \pm 0.004, \pm 0.008, \dots \text{K}^{-1} \text{day}^{-1}$ for Ψ_{θ_z} ; both are then scaled by 1000 for better visualization. The contour interval for $\Psi_{3,Pseudo}$ is $0.1 \text{ m s}^{-1} \text{ day}^{-1}$ and the shadings are as in Fig. 1c.

term \bar{X} rather than the resolved wave forcing term Ψ , suggesting that the effects of parameterized processes such as the gravity wave drag and numerical approximation play an important role in this region.

The key feature of the discrepancies in the residual circulation is the broadly positive $\Delta\Theta_{\text{ERA40-ERAInt}}$ between 2 and 200 hPa together with the poleward arrows in the same region (Fig. 11c). This implies a stronger residual circulation in ERA-40 than ERA-Interim, which is consistent with other studies (e.g., Iwasaki et al. 2009; Monge-Sanz et al. 2013). However, the regions with significant positive $\Delta\Theta_{\text{ERA40-ERAInt}}$ do not generally coincide with the regions of significant negative $\Delta\Psi_{\text{ERA40-ERAInt}}$ or vice versa (see Fig. 1c). The only exceptions are the midlatitude upper stratosphere (20°–40°N, 2–7 hPa) and the high-latitude upper troposphere and lower stratosphere (poleward of 70°N, 500–30 hPa), where $\Delta\Theta_{\text{ERA40-ERAInt}}$ partially cancels $\Delta\Psi_{\text{ERA40-ERAInt}}$. Therefore, the discrepancies in the E-P flux divergence can only partially explain the discrepancies in the residual circulation.

The second row of Fig. 11 shows the climatology of the nonconservative term \bar{X} (contours) calculated as $(d\bar{u}/dt) - \Psi - \Theta$ from ERA-40 and ERA-Interim (Figs. 11d,e), as well as the composite differences of \bar{X} between these two datasets (Fig. 11f). Above 100 hPa, the climatology of \bar{X} is mainly negative for both datasets. This implies that other processes, such as small-scale wave forcing, are also involved in driving the residual meridional circulation (Seviour et al. 2012). Note that Seviour et al. (2012) found smaller magnitudes of the nonconservative term \bar{X} than those shown in Fig. 11e for ERA-Interim. This is likely because our results are based on daily averaged data at $2.5^\circ \times 2.5^\circ$ resolution and for the period of 1979–2002 while

Seviour et al. (2012) used 6-hourly instantaneous records at $3.75^\circ \times 2.5^\circ$ resolution for the period 1989–2009.

The magnitude of stratospheric \bar{X} differs between these two datasets; it is nearly twice as large in ERA-40 than in ERA-Interim. This results in hemisphere-wide significant negative $\Delta\bar{X}_{\text{ERA40-ERAInt}}$ above 200 hPa, except for the high-latitude upper stratosphere where gravity waves may play an important role (Holton 1983). The stratospheric $\Delta\bar{X}_{\text{ERA40-ERAInt}}$ is broadly in balance with $\Delta\Theta_{\text{ERA40-ERAInt}}$ (see Fig. 11c), implying that the balance between terms other than \bar{X} is better achieved in ERA-Interim than ERA-40. Because the zonal wind tendency $\partial\bar{u}/\partial t$ term for the DJF mean is at least one order of magnitude smaller than either Ψ or Θ in terms of both the climatology and the differences (not shown) and the differences in the zonal mean zonal wind \bar{u} between these two datasets are negligibly small (Dee et al. 2011a; Lu et al. 2014), results shown in Figs. 1c and 11c,f indicate that the discrepancies in none of Θ , Ψ , or \bar{X} have corresponding differences in zonal mean flow.

In the upper troposphere, \bar{X} is largely in balance with Ψ in terms of climatology (see Figs. 1a,b). Especially, both datasets show a good balance between \bar{X} and Ψ at 15°–55°N, 150–300 hPa. As such, the TEM budget based on the resolved wave forcing becomes inadequate for the assessment of the forced variability of zonal wind in this region. Figure 11f suggests that this nonlinear interaction appears to occur lower in altitude in ERA-40 than ERA-Interim, resulting in the positive $\Delta\bar{X}_{\text{ERA40-ERAInt}}$ centered at 20°–50°N, 300 hPa; the difference may be attributed to the stronger convective motion and therefore more effective vertical heat transport in ERA-Interim (Wright and Fueglistaler 2013).

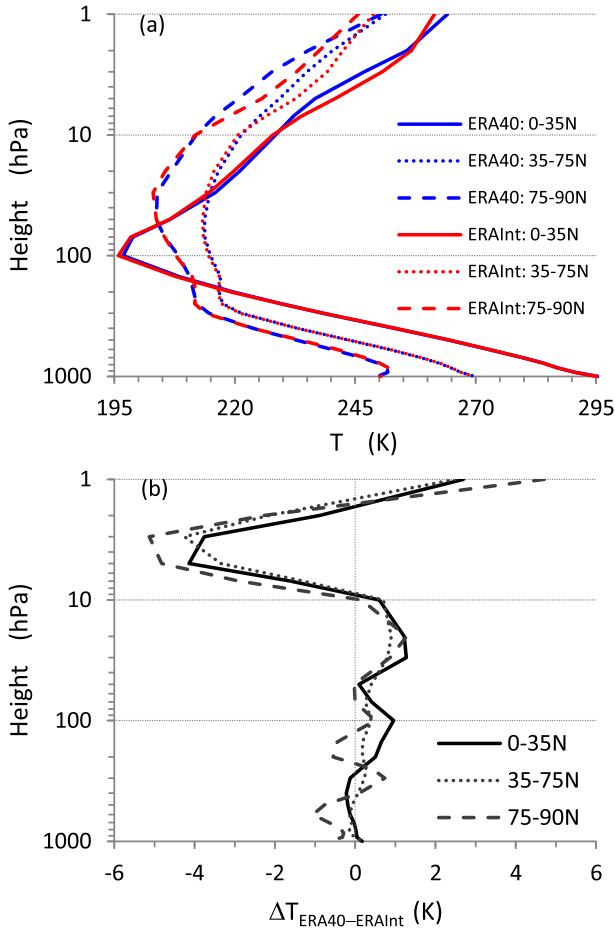


FIG. 7. Vertical profile of (a) the DJF zonal mean temperature and (b) the differences between ERA-40 and ERA-Interim, showing area-weighted averages at 0° – 35° N, 35° – 75° N, and 75° – 90° N.

In the tropical upper troposphere, $\Delta \bar{X}_{\text{ERA40-ERAInt}}$ is mostly in balance with $\Delta \bar{\Psi}_{\text{ERA40-ERAInt}}$, implying that ERA-40 and ERA-Interim account for large-scale wave forcing and the nonconservative processes differently in this region. Similar to those at 15° – 55° N at 150–300 hPa, the discrepancies are closely associated with analysis increments of temperature in the region, where the interaction of temperature, vertical velocity, and humidity is better captured by ERA-Interim than ERA-40 (Dee and Uppala 2009; Dee et al. 2011a). Differences in gravity wave parameterization may also contribute to these tropical discrepancies (McFarlane 1987).

c. Sudden change of mean in the eddy heat fluxes

Up to this point, the diagnostics have been based on the composite differences between the two datasets for their common period; they therefore do not address the discrepancies in long-term trends. Inhomogeneity in either temperature gradient θ_z or wave fluxes can induce

trend uncertainty in the wave forcing. Because the poleward eddy heat flux $\overline{v'T'}$ is the most important quantity for assessing the impact of tropospheric waves propagating into the stratosphere, it is chosen here to identify possible discontinuities that are induced by a change of instruments, or quantity and quality of observations over time. A similar analysis could also be performed for the temperature gradient θ_z , but only the results of $\overline{v'T'}$ are reported here as a demonstration.

In this section, we use the PMT technique to detect any significant sudden departure of $\overline{v'T'}$ difference between the two reanalyses [i.e., $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$]. The reason that we use $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ rather than $\overline{v'T'}_{\text{ERA40}}$ or $\overline{v'T'}_{\text{ERAInt}}$ for the detection is because the PMT technique requires that the time series under consideration is normally distributed and does not have a physically real trend. It is more likely that $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ satisfies the “no-trend” assumption because any apparent trend in $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ is more likely to be caused by a discontinuity of observations or an inhomogeneity in the treatment of observations by the data assimilation procedure in one or both of the datasets. Conversely, $\overline{v'T'}_{\text{ERA40}}$ and $\overline{v'T'}_{\text{ERAInt}}$ are more likely to combine physically real trends with instrumental-induced sudden changes, violating the no-trend requirement of the PMT. For the same reason, $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ is more likely to obey a normal distribution due to the random nature of the observational errors, except for the sudden changes. Most importantly, for each individual time series $\overline{v'T'}_{\text{ERA40}}$ or $\overline{v'T'}_{\text{ERAInt}}$, the magnitudes of the discontinuity and the trend are much smaller than that of the interannual variability, making it statistically harder to detect the changepoint. But because the two time series are very strongly covarying (see, e.g., Fig. 5), taking the difference allows us to effectively remove the interannual variability and thus to detect small discontinuities.

Figure 12 shows the time series of DJF mean total, stationary, and transient eddy heat flux $\overline{v'T'}$ averaged over 10° – 30° N, 100 hPa, from ERA-40 (blue dashed), ERA-Interim (red dash-dotted), and their difference (black solid). It appears that both total and stationary $\overline{v'T'}_{10-30\text{N},100\text{hPa}}$ in ERA-40 have long-term downward trends, which becomes noticeably steeper after the 1991 winter; those in ERA-Interim $\overline{v'T'}_{10-30\text{N},100\text{hPa}}$, however, have no obvious trends. An immediate sudden departure between ERA-40 and ERA-Interim in both total and stationary $\overline{v'T'}_{10-30\text{N},100\text{hPa}}$ can be clearly seen after the 1991 winter with ERA-Interim estimates being consistently larger than those of ERA-40 after this time. A different behavior can be observed for the transient component, with ERA-40 estimates being consistently larger than those of ERA-Interim before the 1997

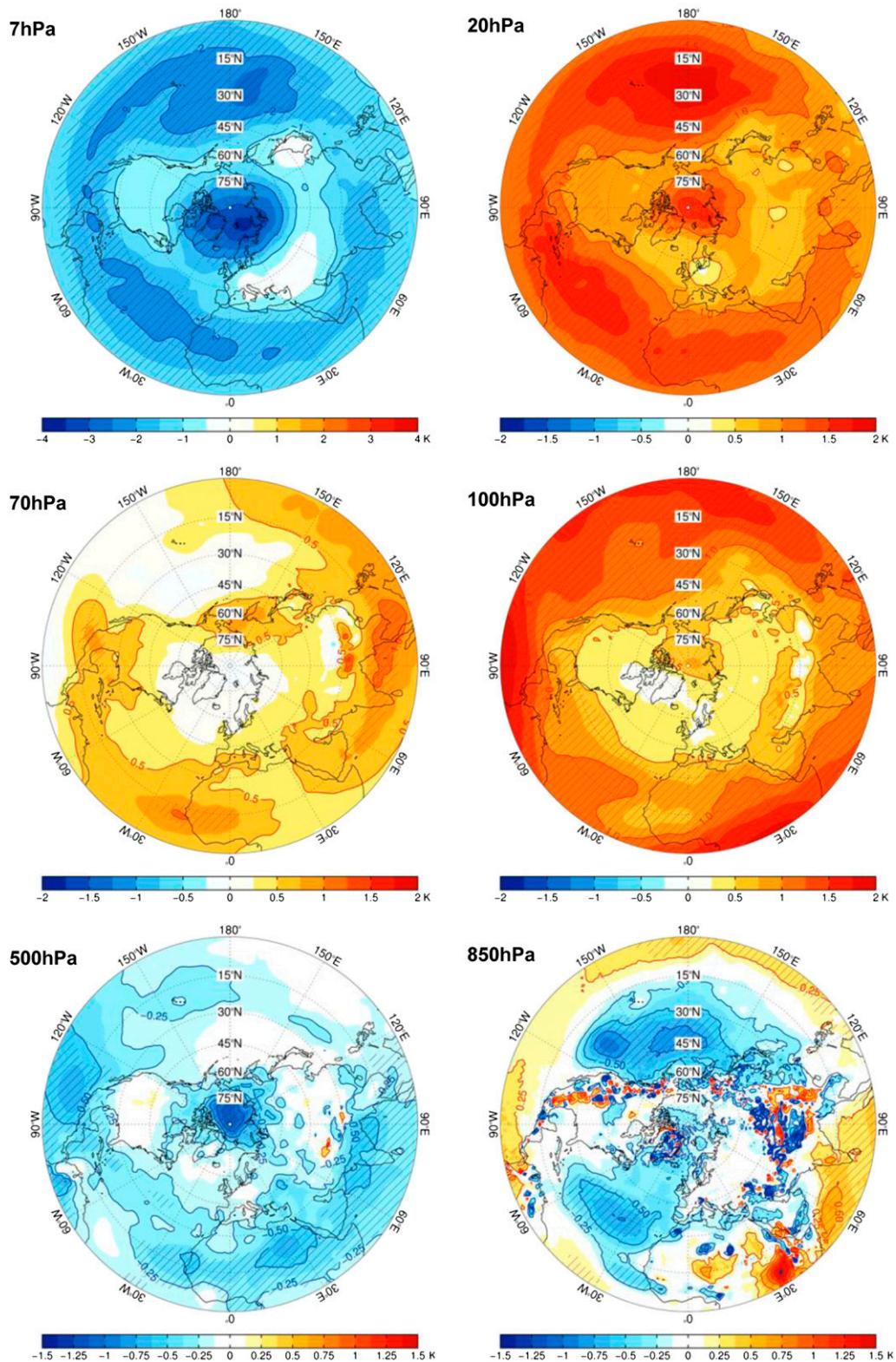


FIG. 8. Polar stereographic plot of DJF mean temperature composite differences between ERA-40 and ERA-Interim at various pressure levels from 7 to 850 hPa. The hatched regions indicate statistical significance at $p \leq 0.05$. Note that the value range of the color bars differs.

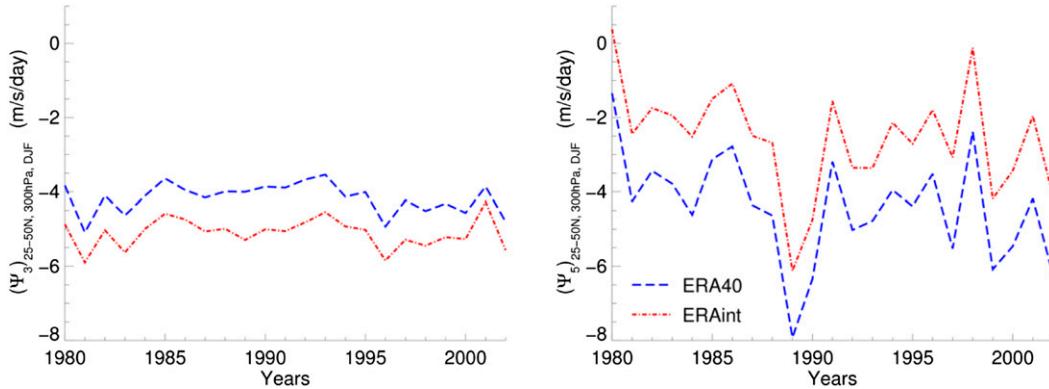


FIG. 9. Time series of DJF mean ERA-40 and ERA-Interim E-P flux divergence terms (left) Ψ_3 and (right) Ψ_5 that are area-weighted averages over 25° – 50° N at 300 hPa.

winter and the two estimates becoming more nearly identical to each other after 1997.

According to the three significance measures, a significant changepoint in the wave forcing difference $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ is detected in the 1991 winter, after which the total $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ and its stationary values dropped significantly. The drop is most noticeable in the stationary component, which has a zero mean for the period 1980–91 but a mean value of -1.5 K m s^{-1} in the period 1992–2002. The drop is about half of the amplitude of its interannual variability. There is another possible changepoint in the winter of 1997, after which the transient component of $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ appeared to drop suddenly. For the 1997 changepoint, however, only two of the three p values are significant at the 0.05 level.

Figure 13a shows a NH polar plot of DJF mean eddy heat flux $v'T'$ at 100 hPa estimated using ERA-Interim while Fig. 13b shows the composite difference of $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ at 100 hPa between the periods 1992–2002 and 1980–91 (i.e., later minus earlier periods). The climatological $v'T'$ peaks at 45° – 75° N and

attains its maximum value ($\sim 80 \text{ K m s}^{-1}$) over the North Pacific Ocean. The overall pattern of $v'T'$ resembles those of previous studies and it is known to be controlled by the stationary component that is related to near-surface topography and topographically induced perturbations (e.g., Plumb 1985; Newman and Nash 2000).

Figure 13b shows the difference plot of total $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ after and before 1991 (later minus earlier periods). The sudden change of $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ in 1991 is mainly marked by a longitudinal belt of negative anomalies centered on 20° N. The largest jump is located near the vicinity of Mt. Pinatubo and there are significant negative anomalies almost everywhere in radiosonde-data-sparse ocean regions in the latitude band of 10° – 30° N. The stationary component accounts for almost all of these negative anomalies (not shown). In the mid to high latitudes, the values of $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ are predominately positive. The magnitude of those positive anomalies is found to be noticeably enhanced in the stationary component while significant negative anomalies of transient $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ are found at

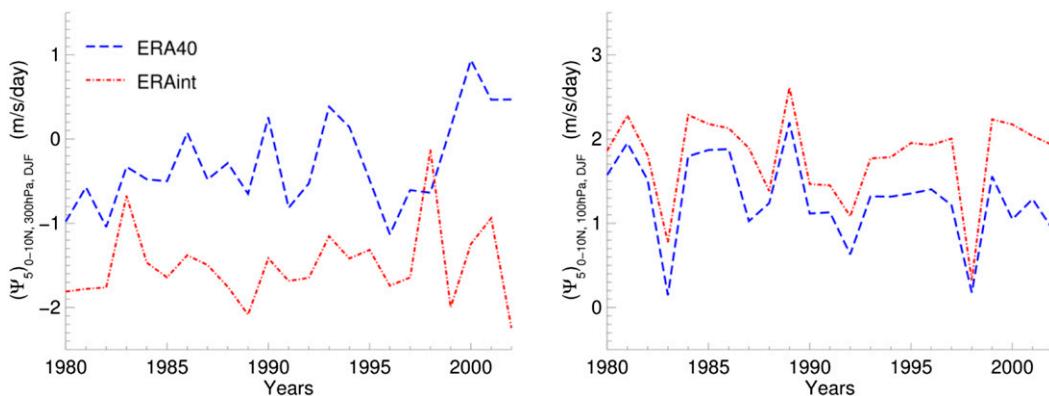


FIG. 10. Time series of DJF ERA-40 and ERA-Interim E-P flux divergence terms Ψ_5 that are area-weighted averages over 0° – 10° N at (left) 300 and (right) 100 hPa.

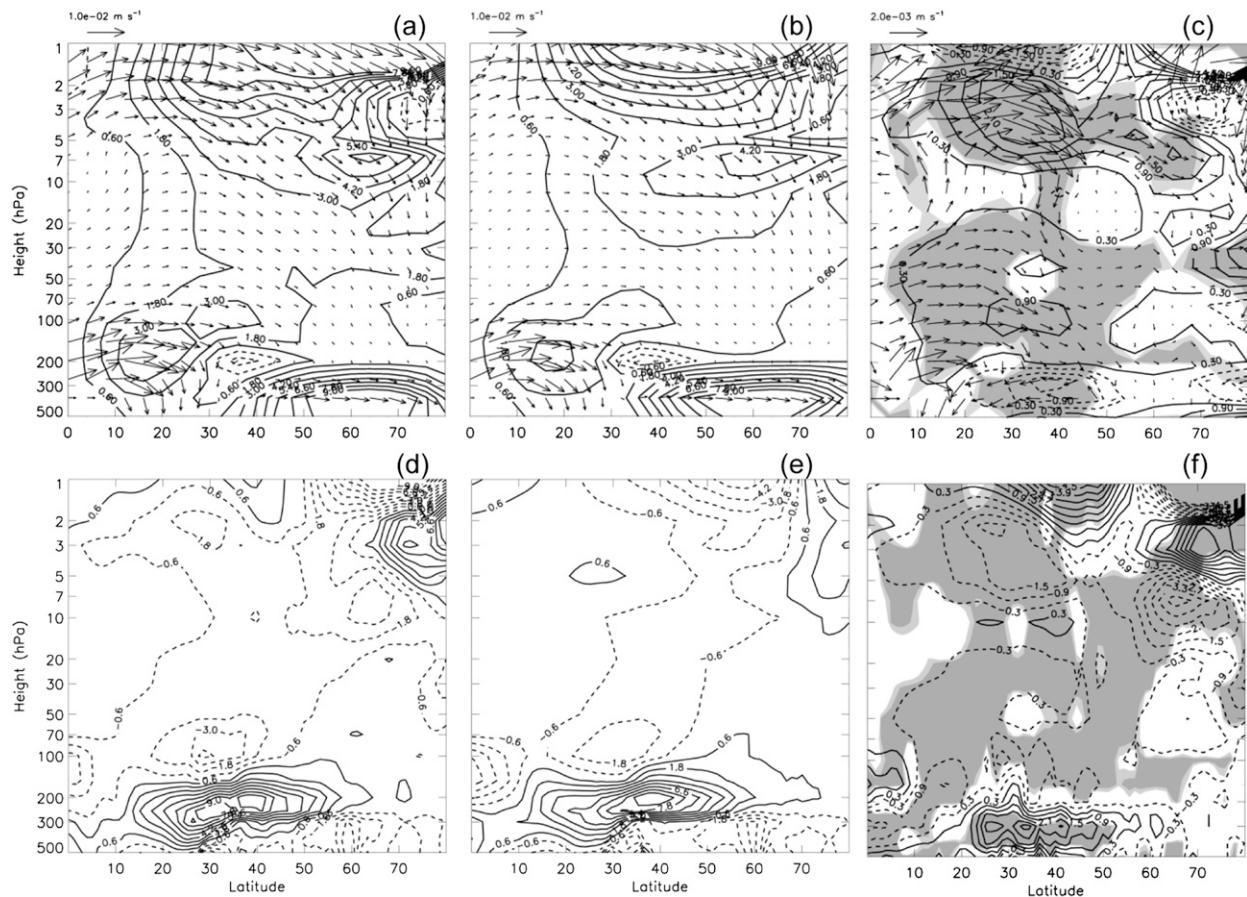


FIG. 11. Latitude–height cross section of the DJF mean residual mean meridional circulation ($\alpha \bar{v}^*$, \bar{w}^*) (arrows) and the residual circulation term Θ (contours), where the residual mean meridional velocity \bar{v}^* is scaled by $\alpha = 4.899 \times 10^{-3}$ for clearer visualization: (a), (b) climatology from ERA-40 and ERA-Interim, respectively and (c) composite differences between the two reanalyses (ERA40 – ERAInt). (d)–(f) As in (a)–(c), except that Θ is replaced by the nonconservative term \bar{X} in Eq. (1). The contour values and shadings are as in Fig. 1.

70°–80°N (not shown). These high-latitude positive (negative) differences in the stationary (transient) component imply that the 1991 sudden jump induced an upward (downward) trend in the respective component of ERA-40 $\overline{v'T'}_{45-75N,10hPa}$ (see Fig. 5).

Figure 14 shows the time series of DJF mean total, stationary, and transient $\overline{v'T'}$ averaged over 45°–75°N at 10 hPa, estimated from ERA-40, ERA-Interim, and their difference. The total, stationary, and transient $\overline{v'T'}_{45-75N,10hPa}$ show similar behaviors as $\overline{v'T'}_{45-75N,10hPa}$ (see the right-hand panels of Fig. 5), with the two datasets closely resembling each other. However, based on the three significance measures, a significant changepoint is detected in 1998 winter for the difference between these two datasets $\Delta(\overline{v'T'})_{ERA40-ERAInt}$. After the 1998 winter, the total and stationary $\Delta(\overline{v'T'})_{ERA40-ERAInt}$ dropped significantly. The transient $\Delta(\overline{v'T'})_{ERA40-ERAInt}$ also dropped after 1998 although the third measure does not attain a p value ≤ 0.05 . However, the detection of

a change after 1998 at 10 hPa involves a comparison between a 4-yr interval that exhibits large $\overline{v'T'}$ anomalies (i.e., 1999–2002) with a 19-yr period of relatively small $\overline{v'T'}$ anomalies (i.e., 1980–98). The relative size of the drop, at $\sim(5\%–7\%)$ of the climatological mean $\overline{v'T'}$, is much smaller than that at 10–30 or 100 hPa. Thus, the sudden change detection at 10 hPa may not be reliable and the effect of this sudden change on the wave forcing estimates is less of a concern than that at 100 hPa.

Figure 15 shows the spatial characteristics of the 1998 sudden change of $\Delta(\overline{v'T'})_{ERA40-ERAInt}$ at 10 hPa, displayed in a similar way as that for the 1991 changepoint at 100 hPa (i.e., Fig. 13). The climatological flux $\overline{v'T'}$ peaks at 45°–80°N, 90°E–180°, with a maximum value of $\sim 180 \text{ K m s}^{-1}$. Weaker positive $\overline{v'T'}$ fluxes are also present in the region 40°–80°N, 150°W–90°E. The effect of the 1998 changepoint in $\Delta(\overline{v'T'})_{ERA40-ERAInt}$ is mainly confined to the region poleward of 40°N. They are marked by negative differences of $\Delta(\overline{v'T'})_{ERA40-ERAInt}$

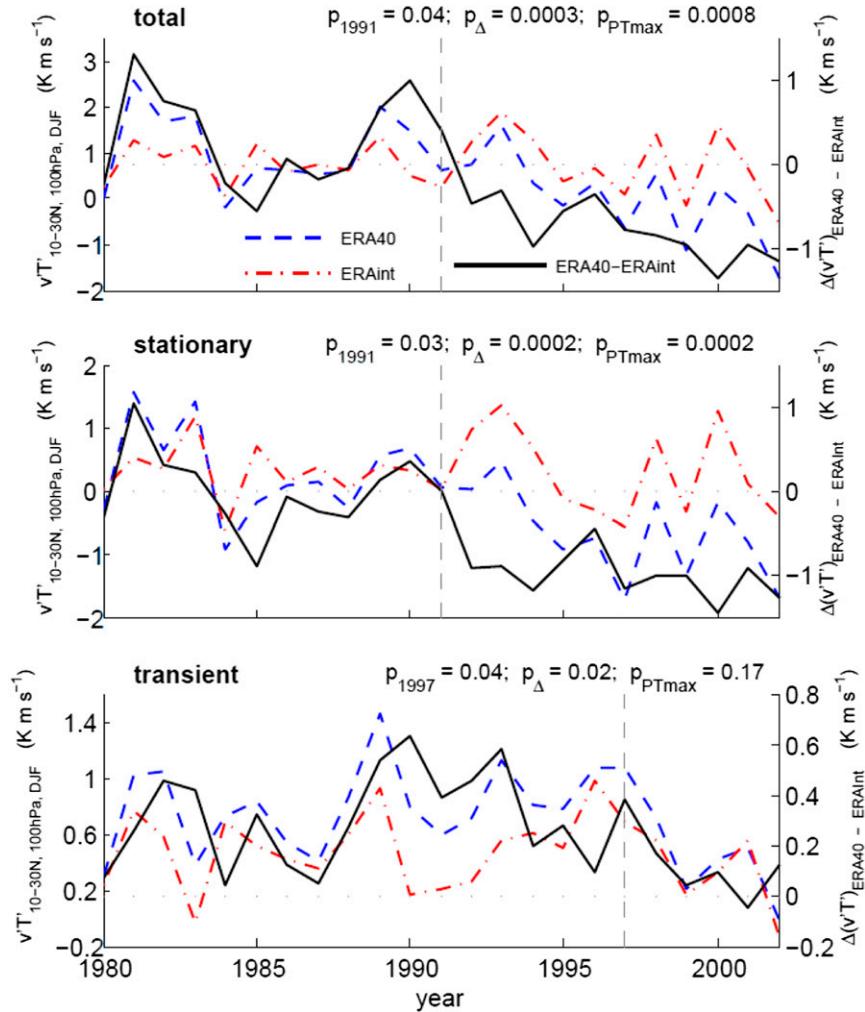


FIG. 12. Time series of DJF mean (top to bottom) total, stationary, and transient components of the zonal mean eddy heat fluxes that are area-weighted averages over 10° – 30° N at 100 hPa. The right-hand y axis is for area-weighted averages of $\Delta(v'T')$ _{ERA40–ERAInt} over 10° – 30° N, 100 hPa (black line). The vertical dashed gray lines indicate the year when a sudden change of the mean in $\Delta(v'T')$ _{ERA40–ERAInt} is identified by the penalized maximal t test of Wang et al. (2007). The significance of the drop is measured by the three p values that are calculated based on 10 000 Monte Carlo simulations and shown at the top of each subplot. In sequence, these significance measures are 1) the probability of the changepoint occurring in the identified winter; 2) the significance of the mean difference between the two periods, after and before the changepoint; and 3) the probability of the maximum values of the penalized t statistic from synthetic time series being greater than or equal to that of the original time series. See section 2c for further details.

over land surfaces at 0° – 120° E and 30° – 150° W. These high-latitude negative differences imply that the 1998 sudden jump induces an apparent upward trend in the stationary component of ERA-Interim $\overline{v'T'}_{45-75N,10hPa}$, which might consequently contribute in part to the steeper upward trend of the E-P flux divergence in this region. This may partially explain why the stationary $\overline{v'T'}_{45-75N,10hPa}$ in ERA-Interim has a steeper upward trend than that in ERA-40 (see Fig. 5).

4. Conclusions and discussion

We have here reported that significant discrepancies exist in the wave forcing estimated from ERA-40 and ERA-Interim during NH winter. When measured by the E-P flux divergence, three key regions are identified as having significant discrepancies. They are the entire high latitudes, the upper troposphere, and the extratropical upper stratosphere. The discrepancies in the high

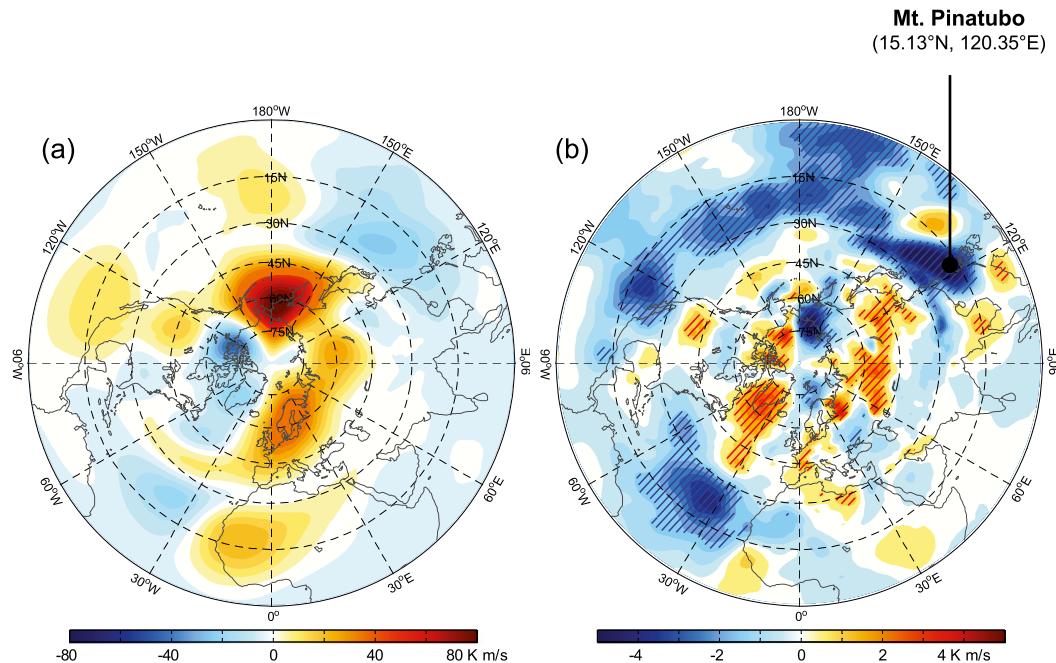


FIG. 13. Polar plots of DJF mean (a) ERA-Interim eddy heat flux at 100 hPa for the period of 1980–91 and (b) composite difference of the eddy heat flux difference $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ at 100 hPa between the periods 1992–2002 and 1980–91. The hatched regions in (b) indicate statistical significance at $p \leq 0.05$.

latitudes are marked by vertically alternating positive and negative anomalies of the E-P flux divergence. They are manifested as differences in the climatological mean between the two datasets and can account for up to 15%–45% of the interannual variability in the affected regions. Such discrepancies are due mainly to differences in the vertical gradient of potential temperature θ_z .

Similar vertically alternating positive and negative anomalies were previously found in the analysis increments of temperature in many reanalysis datasets and are known to be caused by the presence of systematic bias between the data assimilation model and the satellite measurements (Uppala et al. 2005; Dee and Uppala 2009; Kobayashi et al. 2009). Such a bias has a larger magnitude and is more persistent in ERA-40 than ERA-Interim (Simmons et al. 2007; Dee and Uppala 2009). Recent studies indicate much closer agreement to observations by ERA-Interim compared to ERA-40, which is attributed to the advances in the ERA-Interim assimilation system, especially the various improvements of the ERA-Interim 4D-Var system over the previous 3D-Var system that was used by ERA-40 (e.g., Fueglistaler et al. 2009; Dee et al. 2011a; Simmons et al. 2010, 2014; Bracegirdle and Marshall 2012). For this reason, we suggest that the E-P flux divergence discrepancies at high latitudes are most likely due to the

model drift induced by the data assimilation system, rather than observational errors.

In the middle-to-low latitude upper troposphere, the discrepancies in the E-P flux divergence are due largely to the bias in the vertical momentum flux $\overline{w'u'}$. It has been suggested that imbalance of radiative heating induced by assimilation of the observational radiance data tends to introduce noise in the vertical velocity (Schoeberl et al. 2003; Fueglistaler et al. 2009). More importantly, because small-scale processes such as convection and gravity waves may contribute significantly to the momentum budget in addition to resolved wave forcing, differences in model resolution and parameterization of subgrid processes by ERA-40 and ERA-Interim can induce discrepancies in the E-P flux divergence in this region. This may explain why the discrepancies are marked by a cancellation between the E-P flux divergence term Ψ and the nonconservative term \overline{X} . These discrepancies may also be linked to the large bias of analysis increments in the tropical upper troposphere (Dee and Uppala 2009). Furthermore, we have noted that the discrepancy in the E-P flux divergence climatology can be larger than the amplitude of its interannual variation in this region; such large uncertainty strongly discourages merging these two reanalysis products to study wave–mean flow interaction.

In the upper stratosphere, the E-P flux divergence discrepancies involve all the relevant flux terms and are

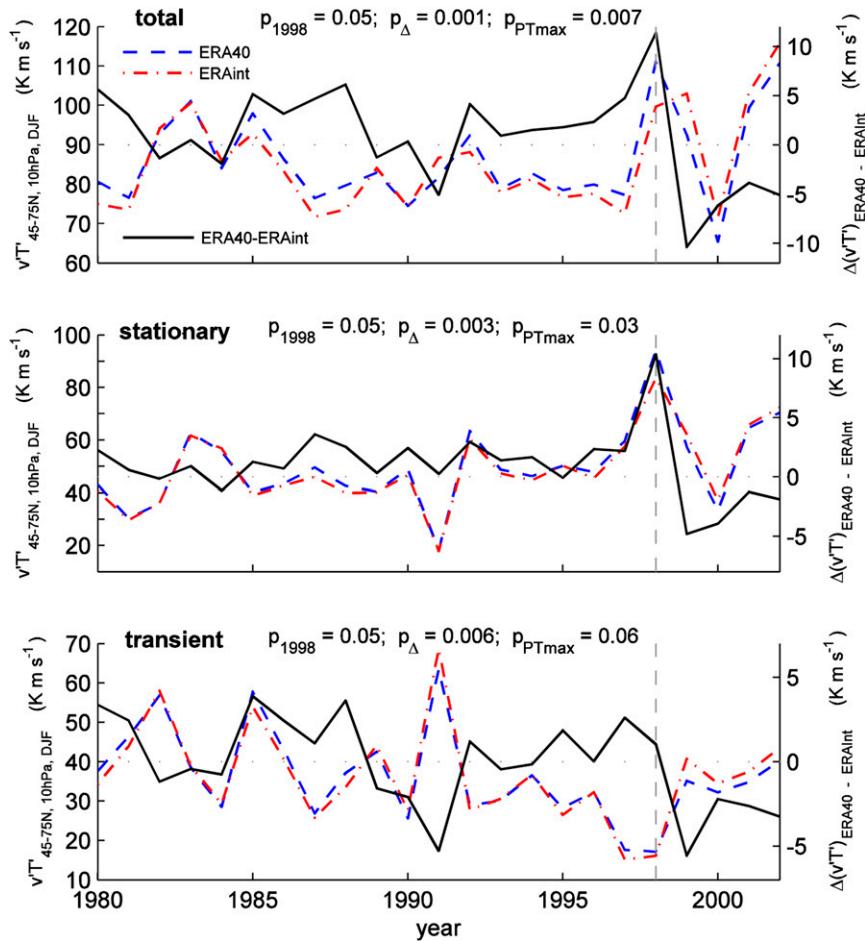


FIG. 14. As in Fig. 12, but for the eddy head flux over 45°–75°N at 10 hPa.

associated with substantial differences in temperature as well as static stability. These discrepancies may be attributed to the relatively larger model bias in the region, where observations are sparse and model errors are large (Dee and Uppala 2009). Nevertheless, we find that the discrepancies between these two datasets become much reduced both in terms of interannual variability, climatological mean, and long-term trend if the wave forcing is measured by the poleward eddy heat flux $\overline{v'T'}$.

Based on the TEM momentum budget, we have shown that a stronger residual circulation is associated with ERA-40 than ERA-Interim, agreeing with previous studies (e.g., van Noije et al. 2004; Monge-Sanz et al. 2007; Dee and Uppala 2009; Monge-Sanz et al. 2013). However, the discrepancies in the residual circulation are only partially associated with the discrepancies in the resolved large-scale wave forcing. The majority of the discrepancies in the residual circulation are associated with the nonconservative term \overline{X} ,

suggesting that the bias in wave forcing is not the main cause for the excessively strong B-D circulation in ERA-40. The excessively strong B-D circulation in ERA-40 was in part attributed to systematic analysis increments in stratospheric temperature that resulted from the biases induced by 3D-Var (Uppala et al. 2005). Apart from radiative heating, improvements in the treatment of the effects of volcanic aerosols, gravity wave drag, and better radiation schemes may also have led to an improved representation of the B-D circulation in ERA-Interim (Dee et al. 2011a). Especially, it is known that gravity wave drag plays a considerable role in driving the B-D circulation (McLandress and Shepherd 2009; Butchart et al. 2010; Seviour et al. 2012) and the effect of smaller-scale wave drag is better resolved by ERA-Interim than ERA-40 (Dee and Uppala 2009; Dee et al. 2011a). A recent study shows that the largest differences in radiative heating in the tropical upper troposphere between several reanalysis products are due primarily to differences in cloud radiative heating as well as localized biases in heating

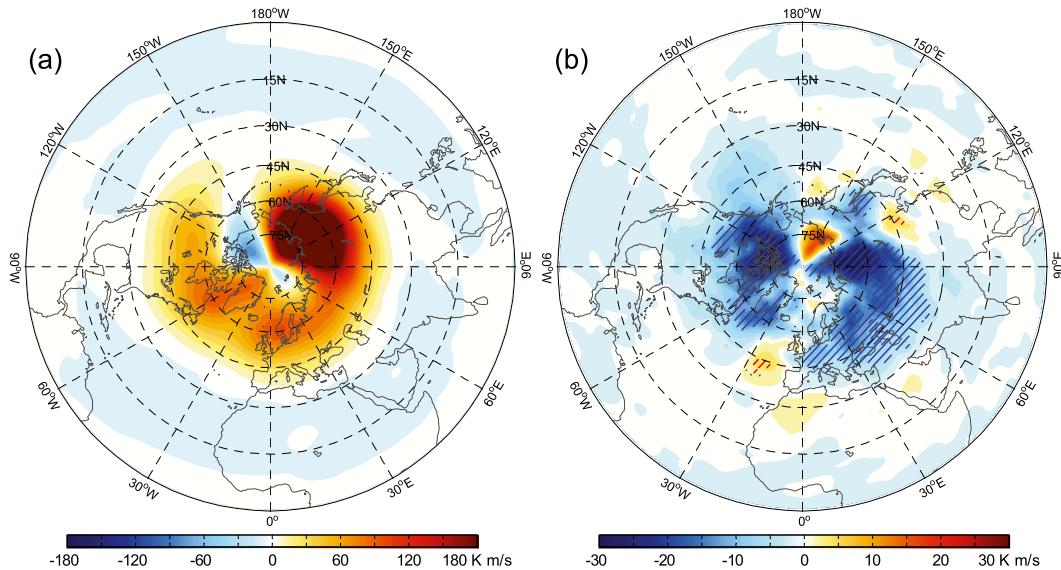


FIG. 15. As in Fig. 13, but at 10 hPa and with a changepoint in 1998.

and cooling associated with parameterized turbulent mixing (Wright and Fueglistaler 2013).

The thermodynamic balance in the stratosphere is largely a balance between the radiative heating and the dynamical heating from the advection of the residual circulation (Andrews et al. 1987). Because the dynamical heating term in the thermodynamic budget of the TEM equations is the product of the residual velocity and the temperature gradients, an enhanced residual circulation should be associated with cooling in the low-latitude stratosphere and warming at high latitudes if the radiative heating is constant. However, the temperature difference $\Delta T_{\text{ERA40-ERAInt}}$ in the tropical lower stratosphere (70–100 hPa) is characterized by significant warm anomalies at 0° – 35°N (see Fig. 8) and an enhanced B-D circulation (see Fig. 11). This is opposite to what we would expect from the augment of dynamical heating. Thus, the discrepancies in the dynamical behavior between these two datasets are more likely the result of a dynamical adjustment to a difference in radiative balance. Because the wave forcing discrepancies are mostly confined to the regions where analysis increments of temperatures are known to be largest, we suggest that they are likely to have originated from an imbalance in radiative heating that is introduced during the ingestion of observational data. However, an enhanced B-D circulation and a warmer tropical lower stratosphere previously reported for the ERA-40 cannot be explained only by differences in the static stability or radiative heating in the high-latitude stratosphere and/or in the low-latitude upper troposphere. The differences may also have originated from

the bias in the forecast model. For instance, the forecast models may have different radiative transfer scheme and/or they generate different distributions of clouds, ozone, and water vapor, which then leads to different radiative heating. Further studies are required to investigate the implications for tracer transport and ozone tendencies extracted from the reanalysis datasets.

A sudden drop of the eddy heat flux difference $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ is detected after the 1991 winter over the subtropical ocean at $\sim(10^{\circ}$ – $30^{\circ}\text{N})$ at 100 hPa. This drop could be due in part to the contamination effects of volcanic aerosols on the infrared radiances measured by the High Resolution Infrared Radiation Sounder (HIRS) on board the *NOAA-12* satellite following the eruption of Mt. Pinatubo in June 1991 (Uppala et al. 2005). Because the radiative transfer model that was used by ERA-40 did not include the effect of volcanic aerosols, the aerosol contaminated radiance measurements were absorbed by the bias corrections of the 3D-Var system, which is known to result in excessive humidity/rainfall in the tropics and subtropics (Uppala et al. 2005; Dee and Uppala 2009). A revised thinning, channel selection, and quality control of HIRS radiances assimilation was used by ERA-Interim, in which the 4D-Var analysis system couples the humidity and the dynamic variables cohesively to help ensure a realistic interaction of temperature, vertical velocity, and humidity (Dee and Uppala 2009; Dee et al. 2011a).

A subtler sudden drop in eddy heat flux difference $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ is also detected in the midlatitudes at 10 hPa. This drop may be linked to the discontinuity in upper-stratospheric temperatures associated with the

radiance measurements from the Advanced Microwave Sounding Unit-A (AMSU-A) since August 1998 in ERA-Interim (Dee and Uppala 2008). While the ERA-40 reanalyzed temperatures in the upper stratosphere inherited a consistent warm bias from the assimilation model, ERA-Interim began using uncorrected radiance data from AMSU-A channel 14 from August 1998 (Dee and Uppala 2008). This change is known to have induced a jump of the global mean temperature above 10 hPa in ERA-Interim (Dee and Uppala 2008). It remains unknown whether or not this change caused a jump of $\overline{v'T'}$ in ERA-Interim or if instead the detected sudden drop of $\Delta(\overline{v'T'})_{\text{ERA40-ERAInt}}$ after the 1998 winter is due mainly to a change of correlation coefficient between temperature T and meridional velocity v .

Several studies have found significant trends in stratospheric wave forcing (Newman and Nash 2000; Randel et al. 2002; Hu and Tung 2002) while others have found that the trends reverse in early and later winter with no significant trend in midwinter (Karpetchko and Nikulin 2004; Hu et al. 2005). Here, we have found that trends in the E-P flux divergences differ substantially between these two datasets. Sudden changes in either temperature gradient or eddy fluxes that are induced by inhomogeneity of observations are able to alter the respective trends and low-frequency variability in the wave forcing. Because of the highly derived nature of the E-P flux divergence, the trends estimated from such a quantity should be treated with extreme caution.

Nevertheless, we have found that the trends in the eddy heat flux $\overline{v'T'}$ are more consistent than those in the E-P flux divergence. In both ERA-40 and ERA-Interim in the midlatitude stratosphere there is a positive trend in the stationary $\overline{v'T'}$ and a negative trend in the transient $\overline{v'T'}$, generally in agreement with previous studies (Newman and Nash 2000; Randel et al. 2002). It must be noted that the general circulation model (GCM) used in both ERA-40 and ERA-Interim does not include the effect of solar variability; any decadal- to multidecadal-scale variation comes solely from the observations (Dee et al. 2011a). During solar maxima, the background state that is predicted by the GCM of the data assimilation system is generally biased compared to the observations, so systematic analysis increments may emerge as a result. This can affect the low-frequency variation and the trends for both datasets. Thus, further confirmation based on longer records and other reanalysis datasets is needed before we can go further into the physical explanations of the opposite trends in the stratosphere in terms of the stationary and transient $\overline{v'T'}$.

This comparative study of wave forcing estimated from ERA-40 and ERA-Interim provides an additional perspective for evaluating dynamic processes in the

stratosphere and upper troposphere. It is noted that a comparative study like this cannot make a quantitative attribution in terms of which dataset is better and/or by how much. Our results nevertheless show that bias in static stability induced by temperature differences and/or radiative heating imbalance can potentially cause large uncertainty in the E-P flux divergence, endorsing the importance of reducing the analysis increments, especially the model drift, in assimilating temperatures. Our results also demonstrate the importance of the recently established Stratospheric Processes and their Role in Climate (SPARC) Reanalysis/Analysis Intercomparison Project (S-RIP) (Fujiwara et al. 2012; Fujiwara and Jackson 2013).

Acknowledgments. This study is part of the British Antarctic Survey Polar Science for Planet Earth Programme funded by the Natural Environment Research Council. We acknowledge use of ECMWF reanalysis datasets and documentation at <http://www.ecmwf.int>. We would also like to thank the three reviewers for their detailed and constructive comments.

APPENDIX

Penalized Maximal t Test

Let $\{X_t\}$ for $t = 1, 2, \dots, N$ denote a time series with zero true trend (but potentially containing a sudden change that may give rise to an apparent linear trend in $\{X_t\}$) and identically and independently distributed (IID) Gaussian errors. To detect a changepoint in $\{X_t\}$ is to test the null hypothesis

$$H_0: \{X_t\} \sim \text{IIDN}(\mu, \sigma^2)$$

against the alternative

$$H_c: \begin{cases} \{X_t\} \sim \text{IIDN}(\mu_1, \sigma^2), & t = 1, \dots, k \\ \{X_t\} \sim \text{IIDN}(\mu_2, \sigma^2), & t = k + 1, \dots, N \end{cases}$$

where $\mu_1 \neq \mu_2$ and $\{X_t\} \sim \text{IIDN}(\mu, \sigma^2)$ stands for $\{X_t\}$ follows an IID Gaussian distribution of mean μ and variance σ^2 . When H_c is true, the entire time series $\{X_t\}$ can be viewed as two independent samples from two normal distributions of the same unknown variance σ^2 , one for all $t \leq k$ and another for all $t > k$, where the point-time $t = k$ is called a changepoint, and $\Delta\mu = |\mu_1 - \mu_2|$ is called the magnitude of the mean shift.

To detect the most probable value of k and to test whether the means of these two samples are statistically significantly different from each other, the test

statistic for detecting a mean shift by penalized maximal t test is

$$PT_{\max} = \max_{1 \leq k \leq N-1} [P(k)T(k)],$$

where

$$T(k) = \frac{1}{\hat{\sigma}_k} \left[\frac{k(N-k)}{N} \right]^{1/2} |\bar{X}_1 - \bar{X}_2|,$$

$$\bar{X}_1 = \frac{1}{k} \sum_{t=1}^k X_t, \quad \bar{X}_2 = \frac{1}{N-k} \sum_{t=k+1}^N X_t, \quad \text{and}$$

$$\hat{\sigma}_k^2 = \frac{1}{N-2} \left[\sum_{t=1}^k (X_t - \bar{X}_1)^2 + \sum_{t=k+1}^N (X_t - \bar{X}_2)^2 \right],$$

and $P(k)$ is an empirical penalty function that is constructed via Monte Carlo simulation according to N and k (Wang et al. 2007). The functional form of $P(k)$ is constructed to give the same level of confidence on the detected change points regardless of their position in the time series $\{X_t\}$ and to have the same false-alarm rate for all points if $\{X_t\}$ happens to be a homogeneous series. The empirical weight function $P(k)$ is used to diminish the effect of unequal sample sizes on the power of detection only based on $T(k)$ so that the false-alarm rate of PMT is evenly distributed across all candidate change points. The detailed formulation of $P(k)$ is given in Wang et al. (2007).

Here, we use three measures to evaluate the significance of any detected change point. The first measure is the chance of a change point occurring at the detected position, the second measure is the significance of the mean-shift magnitude $\Delta\mu$, and the third measure is the significance of the maximum value of the penalized t statistics PT_{\max} . To calculate these three measures, 10 000 synthetic random time series that share the same distribution of $\text{IIDN}(\mu, \sigma^2)$ are constructed based on Monte Carlo simulations. To calculate the first measure, the PMT detection is performed to find the position k in the time series where the maximum value of $P(k)T(k)$ exists for each synthetic series. A distribution of the random trial-based k values is then constructed accordingly. The k value estimated from the original time series is compared to this distribution and the rank of the actual k among these randomized trials determines its significance level. To calculate the second measure, we rank the true $\Delta\mu$ values against the distribution of $\Delta\mu$ calculated from the 10 000 synthetic series. Similarly, the third measure is established by ranking the maximum value of $P(k)T(k)$ of the actual series against those from the synthetic series. When the ranks for all the three measures fall in the 5% tail ends of their associated distributions, the change point is regarded as statistically significant. For simplicity, we

call the proportional ranks as p values in order to align with the traditional terminology of significance tests.

REFERENCES

- Andrews, D. G., J. R. Holton, and C. B. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, 489 pp.
- Bengtsson, L., and Coauthors, 2007: The need for a dynamical climate reanalysis. *Bull. Amer. Meteor. Soc.*, **88**, 495–501, doi:10.1175/BAMS-88-4-495.
- Berrisford, P., and Coauthors, 2011: Atmospheric conservation properties in ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **137**, 1381–1399, doi:10.1002/qj.864.
- Birner, T., and H. Bonisch, 2011: Residual circulation trajectories and transit times into the extratropical lowermost stratosphere. *Atmos. Chem. Phys.*, **11**, 817–827, doi:10.5194/acp-11-817-2011.
- , D. W. J. Thompson, and T. G. Shepherd, 2013: Up-gradient eddy fluxes of potential vorticity near the subtropical jet. *Geophys. Res. Lett.*, **40**, 5988–5993, doi:10.1002/2013GL057728.
- Bracegirdle, T. J., and G. J. Marshall, 2012: The reliability of Antarctic tropospheric pressure and temperature in the latest global reanalyses. *J. Climate*, **25**, 7138–7146, doi:10.1175/JCLI-D-11-00685.1.
- Butchart, N., and Coauthors, 2010: Chemistry–climate model simulations of twenty-first-century stratospheric climate and circulation changes. *J. Climate*, **23**, 5349–5374, doi:10.1175/2010JCLI3404.1.
- Cornes, R., and P. Jones, 2013: How well does the ERA-Interim reanalysis replicate trends in extremes of surface temperature across Europe? *J. Geophys. Res. Atmos.*, **118**, 10 262–10 276, doi:10.1002/jgrd.50799.
- Dee, D. P., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.*, **131**, 3323–3343, doi:10.1256/qj.05.137.
- , and S. Uppala, 2008: Variational bias correction in ERA-Interim. ECMWF Tech. Memo. 575, 26 pp.
- , and —, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quart. J. Roy. Meteor. Soc.*, **135**, 1830–1841, doi:10.1002/qj.493.
- , and Coauthors, 2011a: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, doi:10.1002/qj.828.
- , E. Källén, A. J. Simmons, and I. Haimberger, 2011b: Comments on “Reanalyses suitable for characterizing long-term trends.” *Bull. Amer. Meteor. Soc.*, **92**, 65–70, doi:10.1175/2010BAMS3070.1.
- Dima, I., and J. Wallace, 2007: Structure of the annual-mean equatorial planetary waves in the ERA-40 reanalyses. *J. Atmos. Sci.*, **64**, 2862–2880, doi:10.1175/JAS3985.1.
- Edmon, H., B. Hoskins, and M. McIntyre, 1980: Eliassen-Palm cross sections for the troposphere. *J. Atmos. Sci.*, **37**, 2600–2616, doi:10.1175/1520-0469(1980)037<2600:EPCSFT>2.0.CO;2.
- Fueglistaler, S., B. Legras, A. Beljaars, J.-J. Morcrette, A. Simmons, A. M. Tompkins, and S. Uppala, 2009: The diabatic heat budget of the upper troposphere and lower/mid stratosphere in ECMWF reanalyses. *Quart. J. Roy. Meteor. Soc.*, **135**, 21–37, doi:10.1002/qj.361.
- Fujiwara, M., and D. Jackson, 2013: SPARC Reanalysis Intercomparison Project (S-RIP) planning meeting. *SPARC Newsletter*, No. 41, SPARC Office, Toronto, Ontario, Canada, 52–55.
- , S. Polavarapu, and D. Jackson, 2012: A proposal of the SPARC Reanalysis/Analysis Intercomparison Project. *SPARC Newsletter*, No. 38, SPARC Office, Toronto, Ontario, Canada, 14–17.
- Holton, J. R., 1983: The influence of gravity wave breaking on the general circulation of the middle atmosphere. *J. Atmos.*

- Sci.*, **40**, 2497–2507, doi:10.1175/1520-0469(1983)040<2497:TIOGWB>2.0.CO;2.
- , P. Haynes, M. McIntyre, A. Douglass, R. Rood, and L. Pfister, 1995: Stratosphere–troposphere exchange. *Rev. Geophys.*, **33**, 403–439, doi:10.1029/95RG02097.
- Hu, Y., and K. Tung, 2002: Interannual and decadal variations of planetary wave activity, stratospheric cooling, and Northern Hemisphere annular mode. *J. Climate*, **15**, 1659–1673, doi:10.1175/1520-0442(2002)015<1659:IADVOP>2.0.CO;2.
- , —, and J. Liu, 2005: A closer comparison of early and late-winter atmospheric trends in the Northern Hemisphere. *J. Climate*, **18**, 3204–3216, doi:10.1175/JCLI3468.1.
- Iwasaki, T., H. Hamada, and K. Miyazaki, 2009: Comparisons of Brewer–Dobson circulations diagnosed from reanalyses. *J. Meteor. Soc. Japan*, **87**, 997–1006, doi:10.2151/jmsj.87.997.
- Karpetchko, A., and G. Nikulin, 2004: Influence of early winter upward wave activity flux on midwinter circulation in the stratosphere and troposphere. *J. Climate*, **17**, 4443–4452, doi:10.1175/JCLI-3229.1.
- Kerr-Munslow, A., and W. Norton, 2006: Tropical wave driving of the annual cycle in tropical tropopause temperatures. Part I: ECMWF analyses. *J. Atmos. Sci.*, **63**, 1410–1419, doi:10.1175/JAS3697.1.
- Kobayashi, S., M. Matricardi, D. Dee, and S. Uppala, 2009: Toward a consistent reanalysis of the upper stratosphere based on radiance measurements from SSU and AMSU-A. *Quart. J. Roy. Meteor. Soc.*, **135**, 2086–2099, doi:10.1002/qj.514.
- Lindzen, R. S., 1981: Turbulence and stress owing to gravity wave and tidal breakdown. *J. Geophys. Res.*, **86**, 9707–9714, doi:10.1029/JC086iC10p09707.
- Lu, H., T. J. Bracegirdle, T. Phillips, A. Bushell, and L. Gray, 2014: Mechanisms for the Holton–Tan relationship and its decadal variation. *J. Geophys. Res. Atmos.*, **119**, 2811–2830, doi:10.1002/2013JD021352.
- McFarlane, N. A., 1987: The effect of orographically excited gravity wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci.*, **44**, 1775–1800, doi:10.1175/1520-0469(1987)044<1775:TEOOEG>2.0.CO;2.
- McLandress, C., and T. Shepherd, 2009: Simulated anthropogenic changes in the Brewer–Dobson circulation, including its extension to high latitudes. *J. Climate*, **22**, 1516–1540, doi:10.1175/2008JCLI2679.1.
- Monge-Sanz, B., M. Chipperfield, A. Simmons, and S. Uppala, 2007: Mean age of air and transport in a CTM: Comparison of different ECMWF analyses. *Geophys. Res. Lett.*, **34**, L04801, doi:10.1029/2006GL028515.
- , —, D. Dee, A. Simmons, and S. Uppala, 2013: Improvements in the stratospheric transport achieved by a chemistry transport model with ECMWF (re)analyses: Identifying effects and remaining challenges. *Quart. J. Roy. Meteor. Soc.*, **139**, 654–673, doi:10.1002/qj.1996.
- Monier, E., and B. Weare, 2011: Climatology and trends in the forcing of the stratospheric ozone transport. *Atmos. Chem. Phys.*, **11**, 6311–6323, doi:10.5194/acp-11-6311-2011.
- Newman, P., and E. Nash, 2000: Quantifying the wave driving of the stratosphere. *J. Geophys. Res.*, **105**, 12 485–12 497, doi:10.1029/1999JD901191.
- , —, and J. Rosenfield, 2001: What controls the temperature of the Arctic stratosphere during the spring? *J. Geophys. Res.*, **106**, 19 999–20 010, doi:10.1029/2000JD000061.
- Palmer, T. N., 1981: Diagnostic study of a wavenumber-2 stratospheric sudden warming in a transformed Eulerian-mean formalism. *J. Atmos. Sci.*, **38**, 844–855, doi:10.1175/1520-0469(1981)038<0844:DSOAWS>2.0.CO;2.
- Plumb, R., 1985: An alternative form of Andrews’ conservation law for quasi-geostrophic waves on a steady, nonuniform flow. *J. Atmos. Sci.*, **42**, 298–300, doi:10.1175/1520-0469(1985)042<0298:AAFOAC>2.0.CO;2.
- Randel, W., and E. Jensen, 2013: Physical processes in the tropical tropopause layer and their roles in a changing climate. *Nat. Geosci.*, **6**, 169–176, doi:10.1038/ngeo1733.
- , R. Garcia, and F. Wu, 2002: Time-dependent upwelling in the tropical lower stratosphere estimated from the zonal-mean momentum budget. *J. Atmos. Sci.*, **59**, 2141–2152, doi:10.1175/1520-0469(2002)059<2141:TDUITT>2.0.CO;2.
- , —, and —, 2008: Dynamical balances and tropical stratospheric upwelling. *J. Atmos. Sci.*, **65**, 3584–3595, doi:10.1175/2008JAS2756.1.
- Schoeberl, M., A. Douglass, Z. Zhu, and S. Pawson, 2003: A comparison of the lower stratospheric age spectra derived from a general circulation model and two data assimilation systems. *J. Geophys. Res.*, **108**, 4113, doi:10.1029/2002JD002652.
- Screen, J., and I. Simmonds, 2011: Erroneous Arctic temperature trends in the ERA-40 reanalysis: A closer look. *J. Climate*, **24**, 2620–2627, doi:10.1175/2010JCLI4054.1.
- Seviour, W., N. Butchart, and S. Hardiman, 2012: The Brewer–Dobson circulation inferred from ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **138**, 878–888, doi:10.1002/qj.966.
- Shepherd, T., 2007: Transport in the middle atmosphere. *J. Meteor. Soc. Japan*, **85B**, 165–191, doi:10.2151/jmsj.85B.165.
- Simmons, A. J., S. Uppala, D. Dee, and S. Kobayashi, 2007: ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newsletter*, No. 110, ECMWF, Reading, United Kingdom, 25–35.
- , K. M. Willett, P. D. Jones, P. W. Thorne, and D. P. Dee, 2010: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets. *J. Geophys. Res.*, **115**, D01110, doi:10.1029/2009JD012442.
- , P. Poli, D. P. Dee, P. Berrisford, H. Hersbach, S. Kobayashi, and C. Peubey, 2014: Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Quart. J. Roy. Meteor. Soc.*, **140**, 329–353, doi:10.1002/qj.2317.
- Sterl, A., 2004: On the (in)homogeneity of reanalysis products. *J. Climate*, **17**, 3866–3873, doi:10.1175/1520-0442(2004)017<3866:OTIORP>2.0.CO;2.
- Thorne, P., and R. Vose, 2010: Reanalyses suitable for characterizing long-term trends: Are they really achievable? *Bull. Amer. Meteor. Soc.*, **91**, 353–361, doi:10.1175/2009BAMS2858.1.
- Uppala, S., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.*, **131**, 2961–3012, doi:10.1256/qj.04.176.
- van Noije, T., H. Eskes, M. van Weele, and P. van Velthoven, 2004: Implications of the enhanced Brewer–Dobson circulation in European Centre for Medium-Range Weather Forecasts reanalysis ERA-40 for the stratosphere–troposphere exchange of ozone in global chemistry transport models. *J. Geophys. Res.*, **109**, D19308, doi:10.1029/2004JD004586.
- Wang, X., Q. Wen, and Y. Wu, 2007: Penalized maximal t test for detecting undocumented mean change in climate data series. *J. Appl. Meteor. Climatol.*, **46**, 916–931, doi:10.1175/JAM2504.1.
- Waugh, D., W. Randel, S. Pawson, P. Newman, and E. Nash, 1999: Persistence of the lower stratospheric polar vortices. *J. Geophys. Res.*, **104**, 27 191–27 201, doi:10.1029/1999JD900795.
- Wright, J. S., and S. Fueglistaler, 2013: 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.