Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

On the value of reanalyses prior to 1979 for dynamical studies

Peter Hitchcock[1]

[1]Laboratoire de Météorologie Dynamique, École polytechnique, Palaiseau, France

Correspondence to: Peter Hitchcock (peter.hitchcock@lmd.polytechnique.edu)

Abstract. Studies of stratosphere-troposphere coupling, particularly those seeking to understand the dynamical processes underlying the coupling following extreme events such as major stratospheric warmings, suffer significantly from the relatively small number of such events in the 'satellite' era (1979 to present). This limited sampling of a highly variable dynamical system means that composite averages tend to have large uncertainties. Including years during which radiosonde observations of the stratosphere were of sufficiently high quality could substantially extend this record, potentially reducing this sampling uncertainty by up to 20%. Moreover, many open questions in this field involve aspects of tropospheric dynamics likely to be better constrained by 'conventional' (i.e. radiosonde and surface-based) observations.

Based on an inter-comparison of reanalyses, a quantitative case is made that for many purposes the improved sampling obtained by including this period outweighs the reduced precision of the reanalyses in the Northern Hemisphere. Studies of stratosphere-troposphere coupling should therefore consider the use of this period when using reanalysis data, and the community should advocate for continued attention to be focused on this period from centres producing reanalyses.

1 Introduction

One of the central challenges to the detailed study of the large-scale coupling between the stratosphere and the troposphere is the relatively limited record of high quality, global observations. In the absence of more insightful modes of analysis, quantifying the dynamical processes relevant for the coupling requires large samples to isolate them from unrelated dynamical variability. Despite the availability of nearly four decades of global satellite-based observations, the length of the observational record remains a fundamental limitation to this statistical approach. This is demonstrated explicitly here, as well as by another closely related contribution (Gerber and Martineau, submitted) to the SPARC Reanalysis Intercomparison Program (SRIP; Fujiwara et al., 2017).

The coupling between the stratosphere and the troposphere remains a significant source of uncertainty in projected climate changes over the coming century (Manzini et al., 2014; Simpson et al., 2018), as well as an important source of skill in seasonal forecasting (Sigmond et al., 2013). Global models exhibit a diversity of stratospheric circulation (Manzini et al., 2014) and variability (Charlton-Perez et al., 2013; Taguchi, 2017), and of responses to stratospheric variability (Hitchcock and Simpson, 2014). Observations of the true circulation can be used to identify which models are correctly representing these processes, but this relies on comparing the time-averaged behaviour of the models to the observations, and the large interannual variability in the observed circulation means that the sampling uncertainty remains large. Accounting for sampling error in such large-scale

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

dynamical phenomena is a major concern for many other dynamical questions, including identifying regional signals of climate change and teleconnection patterns (e.g. Deser et al., 2017).

Studies of observed stratosphere troposphere coupling often rely on reanalysis products, which combine a wide range of observations with global forecast models (see Fujiwara et al., 2017, for a comprehensive discussion). Two of the older products, ERA-40 and NCEP-NCAR R1, begin in 1957 and 1948, respectively, dates which coincide with significant extensions of the global radiosonde observing network. Many more recent products (ERA-Interim, MERRA, MERRA-2, CFSR) by contrast cover only the 'satellite' era, that is, the period after 1979. Amongst the more modern products only JRA-55 begins prior to the satellite era, in 1958.

For the purposes of the present work, 'radiosonde' era will be used to refer to the period from 1958 to 1978. There is no general consensus amongst studies of stratosphere-troposphere coupling as to whether to include the radiosonde era. This is complicated by the fact that the coverage of ERA-40 ends in 2002, leaving out the most recent (and best-observed) decade and a half. Some studies have made use of the older reanalysis products ERA-40 and NCEP-NCAR R1 alone (Charlton and Polvani, 2007; Mitchell et al., 2013) while others consider exclusively the satellite record (Dunn-Sigouin and Shaw, 2014; Kodera et al., 2015; Birner and Albers, 2017). Still others choose to merge multiple reanalyses, using an older product for the radiosonde era and a more modern product for the satellite era (Hitchcock et al., 2013; Lehtonen and Karpechko, 2016). The value of JRA-55 as a single modern product that spans both the radiosonde and satellite eras is thus evident, (and as such it will be privileged in the analysis that follows) but the question remains whether the observational record during the radiosonde era is of 'sufficiently' high quality to be worth considering.

Given that the first identification of a stratospheric sudden warming is credited to Scherhag (1952) and that much was known about their dynamics prior to the availability of a long satellite-based observational record (Matsuno, 1971; Labitzke, 1977; McIntyre, 1982, e.g.), largely on the basis of radiosonde observations, the observational record prior to 1979 would seem to be of clear value.

The immediate goal of this work is to evaluate the representation of a number of quantities of interest to the problem of stratosphere-troposphere coupling in the radiosonde era, in view of coming to a more quantitative assessment of their value. For the Northern Hemisphere the arguments given below clearly indicate their value. However, since this judgement depends on the specific quantity of interest, a broader goal is to discuss how to answer this question more generally. Indeed, the same arguments should apply to the study of many other features of the large-scale atmospheric circulation, particularly of those phenomena with large spatial scales and characteristic timescales of the order of weeks to months. The same approach could also be applied in principle to the period prior to 1958, although no effort has been made to do so here.

This evaluation is based on the availability of multiple reanalysis products. Since in general the different reanalyses assimilate subsets of the same observational record into distinct forecast models, the level of agreement provides a simple measure of how strongly the observations constrain the quantity in question. This method has caveats in that the underlying forecast models may share biases that result in them getting consistently wrong answers; more critically, the availability of only one modern reanalysis product that covers the radiosonde era (and assimilates radiosonde data) means that this comparison must be based in

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

part on older reanalyses with known deficiencies (e.g. Long et al., 2017). Nonetheless, as will be argued below, the agreement is close enough to suggest that this period has real value for carrying out many classes of dynamical studies.

The outline of this paper is as follows. The reanalysis data considered here is described in Section 2. Section 3 presents, as an initial example, a discussion of the time series of zonal mean zonal wind at 10 hPa and 60° N that is central to the identification of major sudden stratospheric warmings. Section 4 presents more general criteria for determining when the radiosonde era should be included. These criteria are then discussed in Section 5 as they apply to wider variety of zonal mean quantities. Section 6 presents conclusions and a discussion.

2 Reanalysis data

Zonally averaged output from the 12 reanalysis products listed in Table 1 are considered here. Of these reanalyses, five (JRA-55, NCEP-NCAR, ERA-40, 20CR v2, and ERA-20C) include the period from 1958 through 1978. Two reanalysis products (20CR v2 and ERA-20C) extend further back but are constrained primarily by surface observations. The JRA-55C product is also noteworthy in this context as it assimilates only 'conventional', that is to say, non-satellite based observations. It therefore provides a means of estimating of the additional value of incorporating the satellite observations. A useful comparative description of these reanalysis products including details of the underlying forecast models, the observational datasets assimilated, and the assimilation techniques used can be found in Fujiwara et al. (2017). The data used here has been re-gridded to a uniform latitude-pressure grid, and is available for download (Martineau, 2017).

Anomalies are computed from climatologies based on the years 1981 through 2001. These years are chosen since they are included in all of the reanalysis products under present consideration. Leap years are handled by omitting July 1st so that all years are treated as 365 days long. These climatologies (computed for each reanalysis) are used regardless of the period under consideration.

3 Sudden Stratospheric Warmings

As an initial example, Fig. 1a shows time series of zonal mean zonal wind at 60°, 10 hPa from the JRA-55 reanalysis for a set of 36 stratospheric sudden warming events. The central dates (lag 0) of the events are defined by when the wind at this grid point reverses from westerly to easterly, so all of the time series pass through 0 at this point. However, the inter-event variance of the winds grows rapidly both prior to and after the central date. This spread is only to a weak degree the result of the timing of the event within the cold season; a similar plot of anomalies from the climatological mean shows very similar growth in the inter-event spread (not shown). As a result of this large dynamical variability, the composite mean has a large sampling variability independent of the quality of the observations or the forecast models underlying the reanalysis products.

In contrast, Fig. 1b shows the same time series from all twelve reanalysis products for a single event that occurred on 21 Feb 1989. The inter-reanalysis spread is in general much smaller than the inter-event variability emphasized in Fig. 1a. An exception to this is ERA-20C and 20CR v2 which assimilate only surface observations. JRA-55C, which does not assimilate

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

satellite observations, is notably indistinguishable from other reanalysis products, suggesting that satellite observations are not required to closely constrain these winds.

Although there are far fewer reanalysis products that include the radiosonde period, Fig. 1c shows that the three reanalyses spanning this period which assimilate radiosonde observations (JRA-55, NCEP-NCAR, and ERA-40) exhibit a similarly close agreement, showing only a somewhat larger spread across reanalyses than in the satellite period. This again suggests that the radiosondes are providing a strong constraint on the flow, and that as a result the events that occurred during the radiosonde era are of significant potential value for constraining our knowledge of the composite mean evolution of sudden warmings.

Since sudden stratospheric warmings are typically identified by the date on which this wind reverses sign, these slight differences in reanalyzed winds can lead to the identification of central dates which differ by a day or two, and in some cases can lead to an event being identified in one reanalysis but not in others. This sensitivity is a generic feature of thresholds in the event definition, not of the particular choice of definition.

This leads to difficulties with comparing composites of events in different reanalyses: because of the large inter-event variability, the exclusion of even just one event from a given reanalysis composite mean can produce differences in the composite mean that easily overwhelm the differences in the reanalyzed flow itself. Thus small differences in the identification of events can 'alias' into relatively large apparent differences in the overall composite evolution.

Similar considerations preclude the direct comparison of composite averages of satellite-era and radiosonde-era events: they differ, but not evidently by any more than should be expected due to this dynamical sampling uncertainty. To isolate the intrinsic differences between reanalyses from this aliasing of sampling variability one must instead consider a fixed set of events across all reanalyses. This approach is followed here.

These points are illustrated in Fig. 2, which demonstrates that composites of events across reanalyses agree better when a fixed set of dates is taken then when event dates are chosen individually for each reanalysis. This is true of the full-input analyses for both the satellite era and the radiosonde era.

In contrast, the surface reanalyses (ERA20c and 20CR v2) generally agree better with the composites when event dates are chosen per-reanalysis - particularly around the central date of the event. This suggests that while the surface observations are sufficient to constrain the stratospheric flow to some extent, the break down of the stratospheric vortex is still to a significant extent determined by the behaviour of the forecast model in these products.

Considering a list of fixed event dates provides a useful starting point for quantifying the additional information contained in the radiosonde era. Using the fixed set of event dates as a basis, Fig. 3a shows estimates of the overall frequency of stratospheric sudden warmings for the satellite era alone and for the full 1958-2016 era, as well as for split and displacement events. The month-by-month frequency is shown in Fig. 3b. Confidence intervals in all cases are estimated with a bootstrapping procedure: N years are selected from the period from 1958-2016 with replacement, and the events that occurred in these N years are then used to compute event frequencies, counted multiple times for those years that are selected more than once. For the satellite era $N = N_s = 32$, while for the total period $N = N_t = N_s + N_r = 53$. This whole processes is repeated 10000 times, and the bounds of the confidence intervals are taken to be the 2.5th and 97.5th percentiles.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

As expected from the central limit theorem, the confidence intervals are scaled by factor very close to $\sqrt{N_s/N_t}$, amounting to about a 20% reduction. This improves the observational constraint on the climatological frequency of sudden stratospheric warmings. A similar reduction is obtained for the occurrence frequency of splits and displacements, as well as for the seasonal distribution of events.

Since the bootstrapping is based on the entire record, the confidence intervals for the satellite era are not centered on the mean frequencies. The use of the longer baseline results in a slight shift of the seasonal peak, suggesting that in the long term, January events are in fact more frequent than February events, in contrast to the February peak obtained using the satellite period alone. This represents a modest but useful strengthening of the observational constraints on these statistics.

4 A Statistical Criterion

Despite these promising examples, one should expect in general that the quality of the reanalyses are not as high during the radiosonde era as during the satellite era. In this light one might regard the improvement of 20% found in Fig. 3 to be an upper bound on the degree of improvement. While errors in the reanalyses will in general arise from both observational uncertainty as well as from errors in the underlying forecast model and assimilation process, these will be considered together here as 'reanalysis' uncertainty.

A simple means of quantifying this improvement is to treat the reanalysis and sampling uncertainty as uncorrelated, gaussian variance, and consider the effect on the sample mean of an inhomogeneous set. More explicitly, we consider some physical observable X (for instance, the zonal mean zonal wind at 10 hPa and 60° N) to be modeled by a normally distributed random variable with mean μ and variance $\sigma_d^2 + \sigma_o^2$. Since we are interested in the statistics of the sample mean, the central limit theorem would in principal allow the assumption of normality to be relaxed.

The variance consists of two uncorrelated components: one arising from the dynamical variability of the atmosphere, σ_d^2, and the other from the reanalysis uncertainty, σ_o^2. We further consider two sets of observations of this variable, one of N_s samples with smaller reanalysis error representing the satellite era, with $\sigma_o = \sigma_s$, and one with N_r samples and relatively larger reanalysis error representing the radiosonde era, with $\sigma_o = \sigma_r$. We take the dynamical variability to be constant across both samples. The variance of a sum of independent random variables is the sum of the variance of each variable; hence the variance of the sample mean during the satellite era is

$$\mathrm{Var}\left(\frac{1}{N_s}\sum_{i=1}^{N_s} X_i^s\right) = \frac{1}{N_s}\left(\sigma_d^2 + \sigma_s^2\right), \tag{1}$$

while that of the sample mean over the entire period is

$$\mathrm{Var}\left(\frac{1}{N_t}\left(\sum_{i=1}^{N_s} X_i^s + \sum_{i=1}^{N_s} X_i^r\right)\right) = \frac{1}{N_t^2}\left(N_s\left(\sigma_d^2 + \sigma_s^2\right) + N_r\left(\sigma_d^2 + \sigma_r^2\right)\right). \tag{2}$$

Here the superscript on X indicates the 'era' from which the sample is drawn (and thus its variance).

A simple criterion for including the both periods is that the standard deviation of the sample mean should be reduced relative to that obtained from the satellite era alone. As argued in the previous section, if the reanalysis error of the two periods

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

are equal ($\sigma_r = \sigma_s$), the standard deviation of the mean when the whole record is considered will be reduced by a factor $\sqrt{N_s/(N_s+N_r)}$. If the reanalysis error of the two periods differ, some straightforward manipulations of the formulas above can be used to show that the factor can be written $\sqrt{N_s/(N_s+\delta N_r)}$, with

$$\delta = \frac{1-\beta f}{1+(1-\beta)f}, \qquad f = \frac{\alpha_r^2 - \alpha_s^2}{1+\alpha_s^2}. \tag{3}$$

Here $\alpha_{s,r} = \sigma_{s,r}/\sigma_d$ is the ratio of the reanalysis standard deviation in each respective period to the dynamical standard deviation, and $\beta = N_s/N_t$ is the length of the satellite era as a fraction of the total length of the record. For the observational period considered here, $\beta \approx 0.6$.

The factor δ can be loosely interpreted as an efficiency factor for the sampling during the radiosonde period. Since it depends on the number of observations in both periods its value will in general change (through β) with the size of the sample; however, in the limit that the reanalysis error in both eras is small compared to the dynamical error, $\delta \approx 1 - f = 1 + \alpha_s^2 - \alpha_r^2$, in which case its value is independent of the sample size.

Figure 4 shows values of δ as a function of α_r and α_s for three values of β. One can note several properties of this factor. Firstly, so long as the reanalysis uncertainty in the radiosonde period is larger than that in the satellite era ($\alpha_r > \alpha_s$) δ will be less than 1, with $\delta = 1$ if and only if $\alpha_r = \alpha_s$. Secondly, δ can be negative for sufficiently large values of α_r, although this threshold depends on the value of β. For the present observational record (Fig. 4b), when α_s is small this occurs only when α_r is somewhat larger than 1, that is, when the reanalysis uncertainty is somewhat larger than the dynamical uncertainty. This threshold occurs at smaller values of α_r when the satellite era comprises a larger fraction of the record, as can be seen from comparing the three panels.

In practice, the reanalysis uncertainty σ_o is estimated here from the statistics of differences between different reanalysis products, while the dynamical variability is estimated from the interannual variability of the field in question. As discussed above, the reanalysis uncertainty thus includes both observational uncertainty as well as errors in the forecast model and the assimilation process. If the observations are not constraining the flow in a significant way, the reanalysis product will reflect the dynamics of the underlying forecast model and the flow across the reanalyses will become uncorrelated. If the forecast models produce reasonably accurate dynamical variability, the standard deviation of the differences should approach $\sqrt{2}$ times the dynamical variability; that is, $\alpha_{r,s} \approx \sqrt{2}$. This suggests a second criterion; if the variance of these differences approaches this value, it suggests that the observations are not providing any significant constraint on the fluctuations, and thus that the variability in the reanalysis is arising purely from the forecast model dynamics.

An important assumption that has been made is that the reanalysis uncertainty is dominated by a stochastic component that is uncorrelated across the samples. One can imagine the presence of systematic errors that remain relatively fixed in time, differing only when the assimilated observations change in a substantial way. Such a systematic error will not be reduced by a larger sample size; if such an error ϵ is present during the radiosonde era, its contribution to the overall uncertainty will be $\epsilon(1-\beta)$. However in the case that the dynamical sampling error dominates the random component of the uncertainty, the systematic error can still be negligible if $\epsilon < \sigma_d/\sqrt{N_t}$.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

Since the dynamical standard deviation is in general a function of the flow, and the reanalysis standard deviation is a function of the observational network, the relative information content present in the radiosonde period will vary both spatially and temporally, and will depend on what quantity is under consideration. A complete survey is therefore impossible, but in the next section a brief overview of some commonly used quantities of importance to stratosphere-troposphere interaction is given.

5 Results

Figure 5 shows estimates of the de-seasonalized standard deviation, σ_d, and reanalysis standard deviations σ_s and σ_r for zonal wind in boreal winter and temperature in boreal summer. The standard deviation of the anomaly from the climatology is used as an estimate of σ_d. The variability of DJF zonal winds is large in the Arctic stratospheric polar vortex, and to a lesser extent in the QBO and on the flanks of the tropospheric jets. The variance of JJA temperatures also shows enhanced variance in the winter stratosphere as well as in the deep tropical stratosphere but the structures are less pronounced. In the troposphere the largest variances are at the poles.

The reanalysis uncertainty is estimated during the satellite period (Fig. 5b) as the variance across six reanalysis products (JRA-55, NCEP-NCAR R1, ERA-40, ERA-Interim, MERRA-2, and CFSR; this choice is further justified below) after first removing the respective climatological means. The variance is of the order of 0.1 m s^{-1} through much of the extratropics with a slight increase in the winter stratosphere, and considerably larger inter-reanalysis spread in the deep tropical stratosphere where the lack of strong balance constraints reduces the utility of the thermodynamic measurements available from satellites (Kawatani et al., 2016). In contrast, the inter-reanalysis spread in temperatures is small (0.1 to 0.2 K) throughout most of the summer hemisphere below 10 hPa, but is larger in the upper stratosphere and the winter polar stratosphere. A weak maxima is also seen near the tropical and southern hemisphere tropopauses.

The reanalysis uncertainty during the radiosonde period (Figs. 5ef) is estimated similarly, but using the three full-input reanalyses that cover this period (JRA-55, NCEP-NCAR R1, and ERA-40). Above 10 hPa where data from NCEP-NCAR R1 is not available, the estimate is based on only two products. This results in some weak discontinuities apparent near 10 hPa. The structure of the inter-reanalysis spread is to first order similar to that during the satellite period, but is larger in magnitude. Interhemispheric differences are more apparent, with both wind and temperature spreads noticeably larger in the southern hemisphere, consistent with the sparser set of observational constraints. Nonetheless in many regions in remains substantially smaller than the dynamical variability. Some features with small vertical length-scales are present in the JJA temperature variance, this is likely associated with known artificial vertical temperature oscillations present in ERA-40 (e.g. Manney et al., 2005).

The ‘reanalysis’ uncertainty is, as discussed above, not associated solely with the properties of the observational data available, but also of the assimilation and forecast model used by the respective reanalysis products, and could therefore depend strongly upon which products are included in the calculation. For this reason it is not immediately obvious that the inter-reanalysis spread used here is a reasonable estimate of the reanalysis uncertainty; for instance, certain reanalyses may be outliers for a given quantity and may thus inflate the overall spread.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

Figure 6 thus shows pairwise inter-reanalysis differences, computed as a standard deviation over time of the difference between the anomalies from two different reanalyses. For example, if u_i' is the anomalous zonal mean zonal wind of reanalysis i, the difference σ_{ij} between two reanalyses i and j is

$$\sigma_{ij} = \left(\frac{1}{T} \int \left(u_i'(t) - u_j'(t) \right)^2 dt \right)^{1/2}. \quad (4)$$

Entries below the diagonal are computed for the satellite period, those above the diagonal are for the radiosonde period. Entries on the diagonal show the dynamical variability computed from the corresponding reanalysis

$$\sigma_{ii} = \left(\frac{1}{T} \int u_i'(t)^2 dt \right)^{1/2}. \quad (5)$$

The ratio of the inter-reanalysis spread to the dynamical variability (an estimate of α_r and α_s) are indicated by the colour of the off-diagonal cells.

Differences are shown for four regions in the winters of the respective hemispheres: (a,b) in the Northern and Southern Hemisphere stratosphere, respectively, and (c,d) in the Northern and Southern Hemisphere troposphere. Note that 100 hPa is used to represent the Southern Hemisphere stratosphere while 30 hPa is used for the Northern Hemisphere stratosphere. This latter is chosen to reduce the effects of the model lid in NCEP-NCAR R1 and NCEP-DOE R2; otherwise the conclusions remain essentially unchanged for 10 hPa in the Northern Hemisphere.

The agreement between reanalyses that assimilate some upper air observations (those other than 20CR v2 and ERA-20C) are in almost all cases below 10% of the dynamical variability, in both the troposphere and stratosphere. Looking more closely reanalysis products that share the same or related forecast models tend to be in closer agreement than those from different centres, and there is in general better agreement between the more modern products (JRA-55, ERA-Interim, MERRA-2, CFSR) than between older products. This confirms that the forecast model and assimilation procedure is a contributing factor to the 'reanalysis' error. The agreement between JRA-55C (which does not assimilate satellite observations) and other products is nearly as good as that of JRA-55 in the Northern Hemisphere, even in the stratosphere, while in the Southern Hemisphere the quality of agreement is degraded; interestingly the agreement in the stratosphere is still higher than with the surface reanalyses, but in the troposphere the latter are in closer agreement.

In the Northern Hemisphere troposphere, the two reanalyses that assimilate only surface observations agree broadly to within 30% of the dynamical variability. In the stratosphere and in the southern hemisphere, the differences are considerably larger, but remain smaller than dynamical variability (with the exception of 20CR v2 in the Northern Hemisphere stratosphere), suggesting that surface observations do offer some constraint on the evolution of the stratosphere.

As expected, differences in the radiosonde era are in general larger than the corresponding differences in the satellite era; the one exception to this is in the Northern Hemisphere stratosphere with 20CR v2, where agreement with JRA-55, ERA-40, and NCEP-NCAR R1 are all apparently slightly improved in the absence of satellite observations. Nonetheless, agreement between these latter three in the Northern Hemisphere remain very close, showing only a slight degradation within the troposphere, and an agreement between ERA-40 and JRA-55 to within 10% of the dynamical variability.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

Given the smaller sample size of products which represent the radiosonde period general conclusions cannot be as strong as those from the satellite period, nonetheless the choice of reanalyses used in Fig. 5 is justified in that no significant outliers are apparent. Lower values of the reanalysis uncertainty would likely be obtained if only more modern reanalyses were included, but this would make comparisons to the radiosonde era impossible. However, given the general improvement in agreement across modern reanalyses seen in the satellite era, it is plausible that further improvements within the radiosonde era are also possible.

Having justified to some extent the estimates of σ_d, σ_r, and σ_s, these can be used to estimate the ratios α_r and α_s, and from there the effective value of the radiosonde era according to the criteria discussed in the previous section. Following Fig. 5, these quantities are shown for boreal winter zonal winds and austral winter temperatures in Fig. 7.

The ratio α_s is seen to be in general smaller for the zonal winds than for temperatures, largely as a result of the larger dynamical variability of the former. Consistent with Fig. 5, values are generally smallest in the Northern Hemisphere extratropics, below 0.1 for the winds and below 0.2 for temperatures. The ratio is generally below 0.4 for the winds somewhat larger values near the surface in the deep tropics as well as above 10 hPa in the tropics and at high southern latitudes. For the temperatures values are below 0.4 or so in the extratropics up to about 50 hPa, but begin to approach 1 near the tropopause in the tropics and Southern Hemisphere, and through much of the stratosphere.

The ratio α_r shares many of the structural features present in α_s but with generally larger values. Most importantly for the present discussion, the Northern Hemisphere extratropical winds show values still in general below 0.2, although these values approach 1 in the Southern Hemisphere and stratosphere. Again, α_r for temperatures are larger, in particular near the tropopause. Values in the Northern Hemisphere extratropics remain small.

These ratios suggest that δ for the zonal wind remains in fact quite close to 1 through the Northern Hemisphere and tropics, and are in fact only somewhat reduced for the Southern Hemisphere below 10 hPa. Despite considerable additional uncertainty, this suggests that JJA winds are still well-enough constrained by observations that they may be of some value. Although not shown here, this is true also of DJF winds. This is, however, not the case for DJF temperatures in the Southern Hemisphere (or in fact JJA temperatures, though this is not shown), for which values of δ are in many cases below 0; this is notably the case near the tropical tropopause as well.

In practice these estimates are most sensitive to the dominant dynamical structures of interannual variability in the flow, which have typically relatively longer time scales and larger length scales. These bulk estimates may not therefore imply that the observational constraints on dynamical processes at shorter timescales are equally strong. To begin to assess this point, Fig. 8 compares the power spectra of winds from JRA-55 in the stratosphere and troposphere with the power spectra of pairwise differences between JRA-55 and other reanalyses. The ratio of these two spectra in the corresponding eras can thus be used as a frequency-dependent estimate of α_s^2 and α_r^2. Such spectra are shown for Northern Hemisphere winds in the stratosphere (Fig. 8a,b) and in the troposphere (Fig. 8c,d).

In all cases the raw spectrum of JRA-55 is shown as a reference; curves for all other reanalyses show the power spectrum of the differences between those reanalyses and JRA-55. During satellite era differences from most reanalyses at low frequencies are almost two orders of magnitude smaller than the spectrum (consistent with the 5-10% estimate of the raw difference) since

these plots show the variance instead of the standard deviation. However, fluctuations at higher frequencies reach the same order as the dynamical variability at timescales of a few days in the stratosphere; in the troposphere differences amongst the more modern reanalyses remain below dynamical variability down to the highest frequency considered (corresponding to a period of 6 hours). Within the stratosphere differences from NCEP-NCAR R1 and NCEP-DOE R2 are significantly larger than other reanalyses at all frequencies and differences from ERA-20C and 20CR v2 are as large as the reference spectrum. Within the troposphere the surface reanalyses are still noticeably in less good agreement with JRA-55, with difference spectra that approach the reference spectra at frequencies corresponding to periods less than half a week or so.

During the radiosonde period the differences are, as expected, larger than during the satellite period, although similar features can be noted with better agreement between JRA-55 and ERA-40, and significantly worse agreement with the surface reanalyses.

This suggests that processes with timescales even as short as a few days are still significantly constrained in the Northern Hemisphere extratropics, although this constraint is not as strong (relative to dynamical variability) as is the case for processes on timescales of a month or longer.

A similar spectral analysis could be applied spatially to determine which spatial scales which are reliable. However this has not been directly considered and would be better applied to fully three dimensional data as opposed to the zonal means considered here.

Up to this point the analysis has considered both the radiosonde and satellite eras to be to some extent uniform in their properties; of course the observational record evolved during these periods as well. To consider briefly the evolution of the observational constraint over time, the ratio α_r can be estimated for each month individually over time; in this case we take consider pairwise differences between JRA-55 and other reanalyses as an estimate of σ_o, and the standard deviation of JRA-55 itself as an estimate of σ_d. In all cases the time-series are first de-seasonalized. These ratios can then be used to estimate δ; however, to do so one must assume an appropriate reference value for α_s, here taken to be 0.1 which is roughly appropriate for both quantities based on Fig. 7a. For sufficiently high values of α_r the estimate will also depend on β; in this regime a time-dependent value of δ is not strictly meaningful, although a small value can still be considered indicative of diminished value.

Since the interest is primarily in the early part of the record, Fig. 9 shows this ratio for zonal winds at in the Northern Hemisphere stratosphere (at 60 N, 30 hPa), and in the Southern Hemisphere troposphere (at 45 S, 500 hPa), spanning from 1958 through 1986. The month by month values fluctuate considerably, but show nonetheless a distinct annual cycle with higher values of δ during the respective winter months when the dynamical variability is higher. A clearer trend can be observed by considering δ computed from 12-month running averages of α (bold lines in Fig.9, which suggests that the value of the full-input reanalyses remains high through essentially all of the radiosonde era in the Northern Hemisphere stratosphere, while in the Southern Hemisphere troposphere the value diminishes rapidly prior to 1979. The reasonably good agreement across full-input reanalyses in the Northern Hemisphere stratosphere even towards the beginning of the time period considered here suggests that even the 1950s may be of interest, however, of all full-input reanalyses considered, only NCEP-NCAR R1 includes this decade.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

The surface reanalyses, in particular 20CR v2, are nearly as good as the full-input reanalyses in the Southern Hemisphere troposphere, but their value in the Northern Hemisphere stratosphere is substantially less than those of the full-input reanalyses.

The assessment of inter-reanalysis differences presented here suggest that there is considerable value for dynamical studies in including the radiosonde era, particularly in the extratropical Northern Hemisphere. The criteria discussed suggest that for lower-frequency, large-scale processes such as those responsible for stratosphere-troposphere coupling during stratospheric sudden warmings, including the radiosonde era could reduced confidence intervals by close to 20%, despite the increase in reanalysis uncertainty during this time. To assess whether this is in fact the case, Fig. 10 presents bootstrap estimates of uncertainties on composites of several dynamical quantities fundamental to this coupling: the vertically integrated zonal wind, vertically integrated meridional momentum fluxes, and meridional heat fluxes at 100 hPa. The vertical integral is taken from 1000 hPa to 100 hPa (see, e.g., Hitchcock and Simpson, 2016). The bootstrap estimates are carried out by generating a large number of synthetic composites by selecting N events with replacement from the full period (shown in solid lines with shaded confidence intervals), and from the satellite period (shown in dashed lines with outlined confidence intervals).

Importantly, any systematic error present in these quantities during the radiosonde era will contribute to the bootstrapped confidence intervals. The fact then that in each case confidence intervals are reduced by on the order of 20% (not shown explicitly) confirms that any such systematic errors are small relative to the sampling error.

As was the case with the event frequencies shown in Fig. 3, the composite means agree nearly everywhere to within estimated confidence intervals, as should be the case. Within these uncertainties, the tropospheric jet shift is seen at somewhat lower latitudes during the full period with a less pronounced low-latitude signal; the momentum flux anomalies are somewhat more positive, and the heat-flux anomalies during the recovery phase suggest somewhat more suppression of the upward wave flux. While the differences in composite means are modest, including this period reduces the confidence intervals on these quantities by the expected amount, providing better observational constraints on dynamical understanding and modeling efforts.

6 Conclusions

The growth of satellite observations providing global coverage following 1979 resulted in major improvements in the monitoring of the detailed state of the atmosphere. However, the network of surface and radiosonde observations in the period from 1958 to 1978 were remarkably effective in constraining many features of the general circulation, even in the boreal lower stratosphere. For dynamical studies that rely on statistical composites of specific anomalous conditions, the dominant source of error is in many cases that of sampling variability, and in this context the radiosonde period represents a valuable extension of the observational record, allowing in principle a reduction of 20% in confidence intervals associated with the dynamical variability.

The value of this record towards reducing the overall sampling uncertainty in composites has been quantified (3). This depends on the ratio of the 'reanalysis' uncertainty (including errors arising both from the precision of the underlying observations as well as that arising from the assimilation process) to the dynamical uncertainty (the variability of the dynamical phenomena

themselves). In general this depends also on the relative length of the radiosonde era to the total time period considered, but when the dynamical variability dominates the overall uncertainty, this dependency drops out.

Since this quantity is in practice a function of both the physical properties of the climate system, the observations available, and of the reanalysis forecast model and assimilation system, this criteria must be applied on a case-by-case basis, and the present work cannot hope to provide a comprehensive survey. However, basic zonal mean quantities including zonal winds, temperatures, and fluxes of momentum and heat, as archived for 12 reanalysis products (see Table 1) by Martineau (2017), have been considered here.

For all quantities considered, the reanalysis uncertainty in the Northern Hemisphere extratropics from the surface up to the mid-stratosphere (about 10 hPa) is found to be sufficiently small relative to the dynamical variability to make the radiosonde era of clear value in reducing composite uncertainties. For zonal mean zonal winds, the interannual variability is such that despite larger reanalysis uncertainties, this is also the case for tropical winds (even in the stratosphere) and even Southern Hemisphere winds are of potential value. Because the dynamical variability of temperature is smaller, the reanalysis uncertainty in the radiosonde era is relatively large and suggests that much of the Southern Hemisphere is not well enough constrained to be worth including. This is also notably true of temperatures in the tropical tropopause layer.

This test has also be applied to the surface reanalyses ERA20c and 20CR v2. The statistics of differences between these products and full-input reanalyses clearly indicate that, at least for ERA20c, their stratospheric evolution bears some meaningful resemblance to reality. However, the test indicates that, relative to the constraint available from full-input reanalyses during the satellite era, their errors are too large to meaningfully constrain dynamical variability (see Fig. 9). Furthermore, while differences between other reanalyses are reduced when considering fixed dates for stratospheric sudden warmings, for the surface reanalyses the comparison is improved when considering per-reanalysis dates, suggesting that, in these surface reanalyses, stratospheric sudden warmings are more a product of the forecast model dynamics than a result of assimilated observations.

While this criteria does not consider the possibility of systematic biases in the radiosonde era, direct bootstrap estimates confirm this reduction in uncertainty of several dynamical quantities relevant to stratosphere-troposphere coupling following stratospheric sudden warmings in the Northern Hemisphere. These estimates are sensitive to systematic biases (at least any relative to those in the satellite era), suggesting that any such biases are negligible for these quantities.

Finally, while considerable improvements have been documented for more modern reanalyses during the satellite period (e.g. Long et al., 2017), there are at present not enough modern reanalyses that cover the radiosonde era to clearly document improvements over this earlier period. Nonetheless, it seems likely that further attention on this period could produce further improvements. Given the value of this period for dynamical studies, such attention from the reanalyses centres would be welcome.

7 Data availability

All analysis is based on the zonal mean dataset kindly provided by Patrick Martineau which is available online from the Centre for Environmental Data Analysis Martineau (2017).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Acknowledgements. The author thanks Sean Davis and Gloria Manney for helpful discussions, as well the lead authors of the SRIP chapter on Stratosphere-Troposphere Coupling, Patrick Martineau and Ed Gerber, for their support of this work.

References

Birner, T. and Albers, J. R.: Sudden Stratospheric Warmings and Anomalous Upward Wave Activity Flux, Sci. Online Lett. Atmos., 13A, 8–12, doi:10.2151/sola.13A-002, 2017.

Charlton, A. J. and Polvani, L. M.: A new look at stratospheric sudden warmings. Part I: Climatology and modelling benchmarks, J. Clim., 20, 449–469, 2007.

Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., Davis, N. A., Gerber, E. P., Gillett, N., Hardiman, S., Kim, J., Krüger, K., Lee, Y.-Y., Manzini, E., McDaniel, B. A., Polvani, L., Reichler, T., Shaw, T. A., Sigmond, M., Son, S.-W., Toohey, M., Wilcox, L., Yoden, S., Christiansen, B., Lott, F., Shindell, D., Yukimoto, S., , and Watanabe, S.: On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, J. Geophys. Res., 118, 2494–2505, doi:10.1002/jgrd.50125, 2013.

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

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

Deser, C., Simpson, I. R., McKinnon, K. A., and Phillips, A. S.: The Northern Hemisphere extra-tropical atmospheric circulation response to ENSO: How well do we know it and how do we evaluate models accordingly?, J. Clim., 30, 5059–5082, doi:10.1175/JCLI-D-16-0844.1, 2017.

Dunn-Sigouin, E. and Shaw, T. A.: Comparing and contrasting extreme stratospheric events, including their coupling to the tropospheric circulation, J. Geophys. Res., 120, 1374–1390, doi:10.1002/2014JD022116, 2014.

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

Gerber, E. P. and Martineau, P.: Quantifying the variability of the annular modes: Reanalysis uncertainty vs. sampling uncertainty, Atmos. Chem. Phys., submitted.

Hitchcock, P. and Simpson, I. R.: The downward influence of stratospheric sudden warmings, J. Atmos. Sci., 71, 3856–3876, doi:10.1175/JAS-D-14-0012.1, 2014.

Hitchcock, P. and Simpson, I. R.: Quantifying forcings and feedbacks following stratospheric sudden warmings, J. Atmos. Sci., 73, 3641–3657, doi:10.1175/JAS-D-16-0056.1, 2016.

Hitchcock, P., Shepherd, T. G., and Manney, G. L.: Statistical characterization of Arctic Polar-night Jet Oscillation events, J. Clim., 26, 2096–2116, doi:10.1175/JCLI-D-12-00202.1, 2013.

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, Bull. Amer. Meteor. 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), Bull. Amer. Meteor. Soc., 83, 1631–1643, 2002.

Kawatani, Y., Hamilton, K., Miyazaki, K., Fujiwara, M., and Anstey, J. A.: Representation of the tropical stratospheric zonal wind in global atmospheric reanalyses, Atmos. Chem. Phys., 83, 1631–1643, doi:10.5194/acp-16-6681-2016, 2016.

Kobayashi, C., Endo, H., Ota, Y., Kobayashi, S., Onoda, H., Harada, Y., Onogi, K., and Kamahori, H.: Preliminary results of the JRA-55C, an atmospheric reanalysis assimilating conventional observations only, Sci. Online Lett. Atmos., 10, 78–82, doi:10.2151/sola.2014-016, 2014.

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. Meteor. Soc. Japan, 93, 5–48, doi:10.2151/jmsj.2015-001, 2015.

Kodera, K., Mukougawa, H., Maury, P., Ueda, M., and Claud, C.: Absorbing and reflecting sudden stratospheric warming events and their relationship with tropospheric circulation, J. Geophys. Res., 121, 80–94, doi:10.1002/ 2015JD023359, 2015.

Labitzke, K.: Interannual variability of the winter stratosphere in the Northern Hemisphere, Mon. Wea. Rev., 105, 762–770, 1977.

Lehtonen, I. and Karpechko, A. Y.: Observed and modeled tropospheric cold anomalies associated with sudden stratospheric warmings, J. Geophys. Res., 121, 1591–1610, doi:10.1002/2015JD023860, 2016.

Long, C. S., Fujiwara, M., Davis, S., Mitchell, D. M., and Wright, C. J.: Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP), Atmos. Chem. Phys., 17, 14 593–14 629, doi:10.5194/acp-17-14593-2017, 2017.

Manney, G. L., Krüger, K., Sabutis, J. L., Sena, S. A., and Pawson, S.: The remarkable 2003-2004 winter and other recent warm winters in the Arctic stratosphere since the late 1990s, J. Geophys. Res., 110, D04 107, doi:10.1029/2004JD005367, 2005.

Manzini, E., Karpechko, A. Y., Anstey, J., Baldwin, M. P., Black, R. X., Cagnazzo, C., Calvo, N., Charlton-Perez, A., Christiansen, B., Davini, P., Gerber, E., Giorgetta, M., Gray, L., Hardiman, S. C., Lee, Y.-Y., Marsh, D. R., McDaniel, B. A., Purich, A., Scaife, A. A., Shindell, D., Son, S.-W., Watanabe, S., and Zappa, G.: Northern winter climate change: Assessment of uncertainty in CMIP5 projections related to stratosphere-troposphere coupling, J. Geophys. Res., 119, 7979–7998, doi:10.1002/2013JD021403, 2014.

Martineau, P.: Zonal-mean dynamical variables of global atmospheric reanalyses on pressure levels, doi:10.5285/b241a7f536a244749662360bd7839312, 2017.

Matsuno, T.: A Dynamical model of the stratospheric sudden warming, J. Atmos. Sci., 28, 1479–1494, 1971.

McIntyre, M. E.: How well do we understand the dynamics of stratospheric warmings?, J. Meteor. Soc. Japan, 60, 37–65, 1982.

Mitchell, D. M., Gray, L. J., Anstey, J., Baldwin, M. P., and Charlton-Perez, A. J.: The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate, J. Clim., 26, 2668–2682, doi:10.1175/JCLI-D-12-00030.1, 2013.

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. Meteor. Soc. Japan, 85, 369–432, doi:10.2151/jmsj.85.369, 2007.

Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J.-N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., Berrisford, P., Dragani, R., Trémolet, Y., Holm, E., Bonavita, M., Isaksen, L., and Fisher, M.: The Data Assimilation System and Initial Performance Evaluation of the ECMWF Pilot Reanalysis of the 20th Century Assimilating Surface Observations Only (ERA-20C), Tech. Rep. 14, ECMWF, Reading, UK, 2013.

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. Clim., 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, Bull. Amer. Meteor. Soc., 91, 1015–1057, doi:10.1175/2010BAMS3001.1, 2010.

Scherhag, R.: Die explosionsartigen Stratosphärenerwärmungen des Spätwinters, Ber. Dtsch. Wetterdienst (US Zone), 6, 51–63, 1952.

Sigmond, M., Scinocca, J. F., Kharin, V. V., and Shepherd, T. G.: Enhanced seasonal forecast skill following stratospheric sudden warmings, Nat. Geosci., 6, 98–102, doi:10.1038/NGEO1698, 2013.

Simpson, I. R., Hitchcock, P., Seager, R., and Wu, Y.: The downward influence of uncertainty in the Northern Hemisphere stratospheric polar vortex response to climate change, J. Clim., 31, 6371–6391, doi:10.1175/JCLI-D-18-0041.1, 2018.

Taguchi, M.: A study of different frequencies of major stratospheric sudden warmings in CMIP5 historical simulations, J. Geophys. Res., 122, 5144–5156, doi:10.1002/2016JD025826, 2017.

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

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

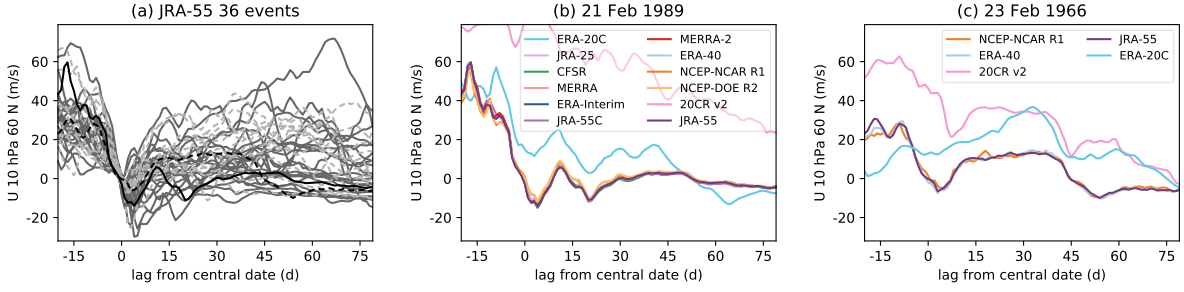

Figure 1. (a) Winds from JRA-55 for 36 sudden warmings. Events from the satellite period are in dark grey, those from the radiosonde period are in light grey and are dashed. (b) Winds for a single satellite-period event for all reanalyses; this event is shown by the black line in (a). (c) Winds for a single radiosonde-period event for all reanalyses covering this period; this event is shown by the dashed black line in (a).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

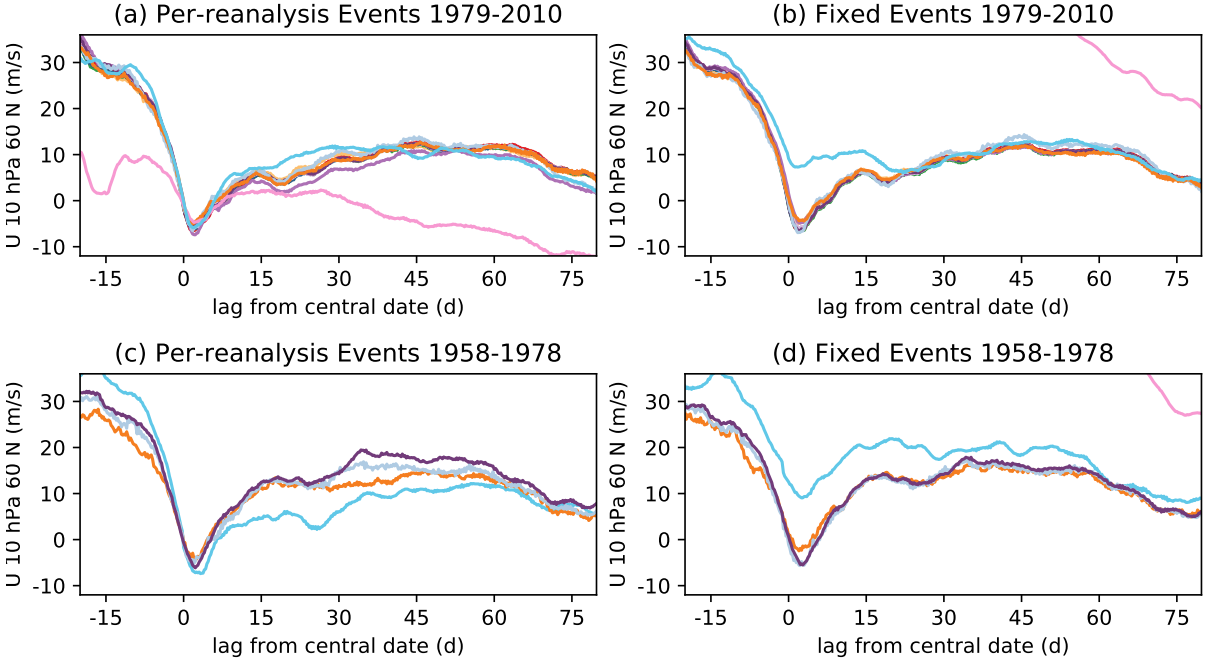

Figure 2. Composites of zonal mean zonal wind at 10 hPa, 60° N during stratospheric sudden warmings for events during the satellite era (a,b) and the radiosonde era (c,d). Events in (a,c) are determined by applying the wind reversal criteria of Charlton and Polvani (2007) to each reanalysis individually, while those in (b,d) are taken to be common across all reanalyses. Line colours are as in Fig. 1.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

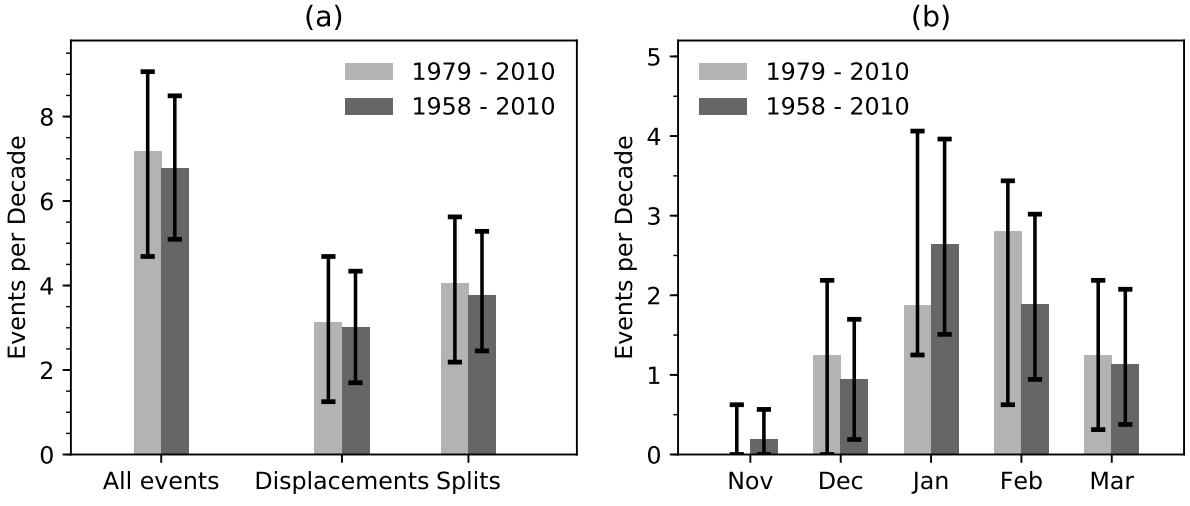

Figure 3. (a) Frequency of all events, and of events classified as splits or displacements for details for the satellite period versus for the radiosonde period. (b) Same as (a) but for each month of extended winter. Error bars indicate 95% confidence intervals, see text for details.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

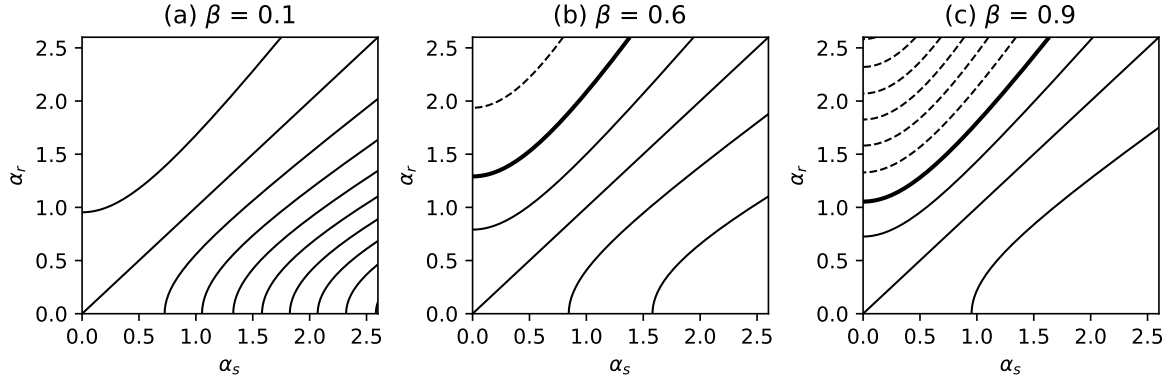

Figure 4. The effective value δ of radiosonde-era degrees of freedom relative to that of satellite-era degrees of freedom in reducing the overall uncertainty. Shown as a function of α_r and α_s for three values of β: (a) 0.1 (radiosonde era much longer than satellite era), (b) 0.6 (roughly appropriate for the observational records considered here) and (c) 0.9 (radiosonde era much shorter than satellite era). Contour interval is 0.5, with the 0 contour indicated by the bold line.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

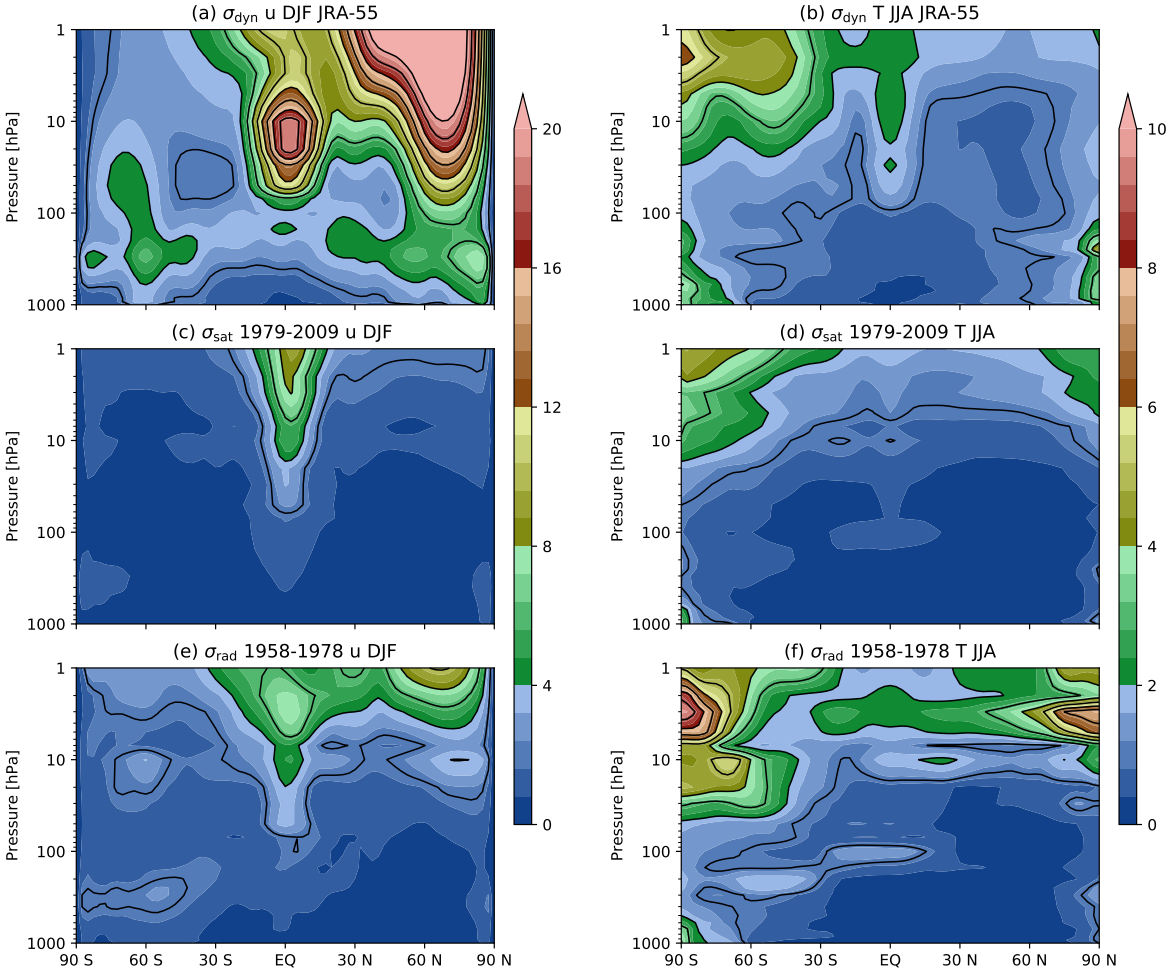

Figure 5. Standard deviation of de-seasonalized (a) winds in DJF and (b) temperatures in JJA from the JRA-55 reanalysis over the satellite period. (c,d) Standard deviation of the differences in same quantities (respectively) across six reanalysis products for the satellite period. (e,f) As in (c,d) but across three reanalysis products for the radiosonde period. See text for details.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

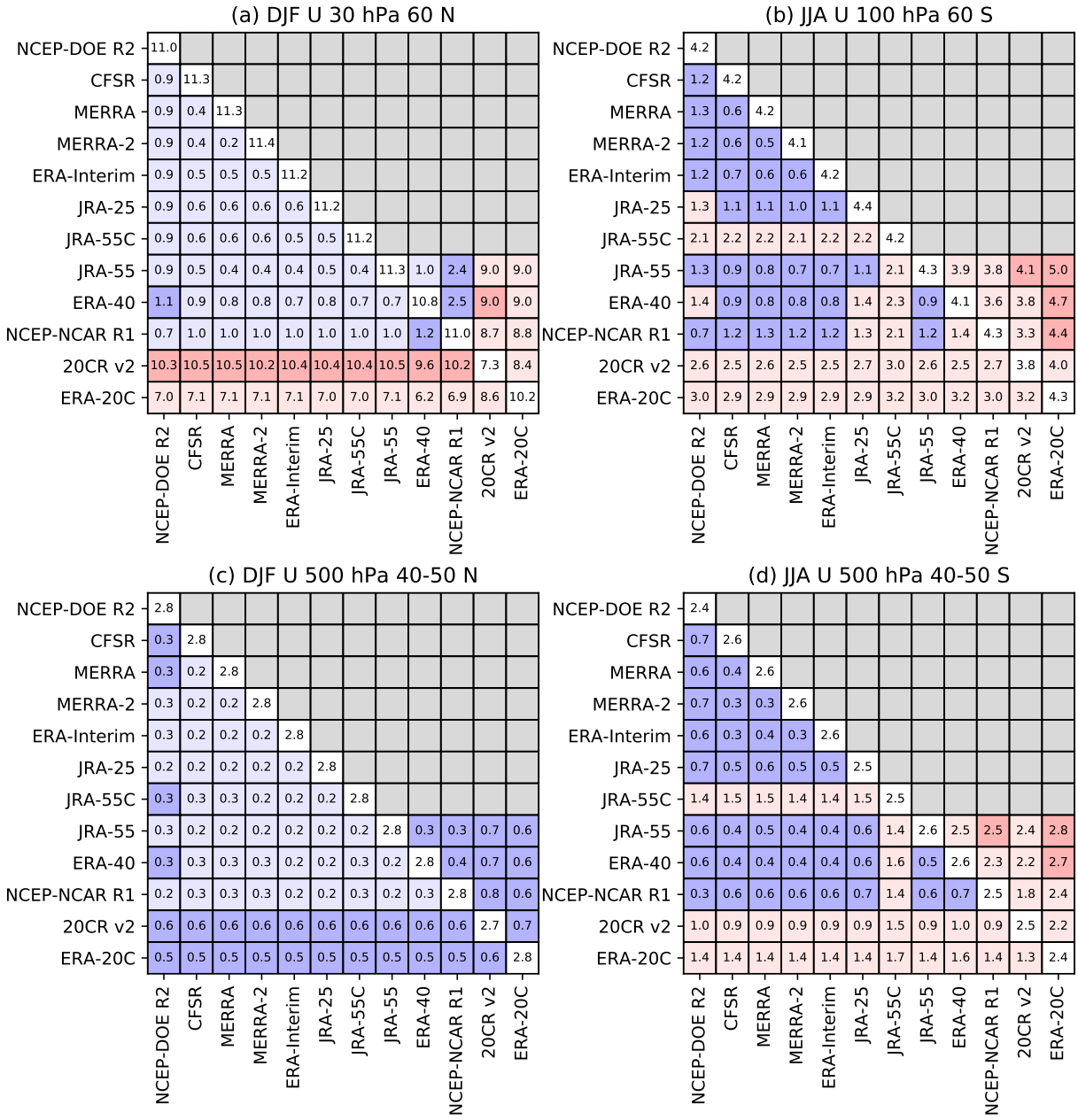

Figure 6. Standard deviations of pair-wise differences between winds in different reanalysis products at (a) 30 hPa, 60° N (DJF), (b) 100 hPa, 60° S (JJA), (c) 500 hPa, 40-50° N (DJF), and (d) 500 hPa, 40-50° S (JJA). All quantities are in m s^{-1}. The diagonal elements show the de-seasonalized standard deviation of the corresponding quantity, elements below the diagonal show differences for the satellite era, and elements above the diagonal show differences for the radiosonde era. Elements are shaded by the ratio of the difference to the mean of the dynamical standard deviations from the corresponding two diagonal elements; light blue (less than 10%), dark blue (10% to 30%), light red (30% to 100%), and dark red (greater than 100%).

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

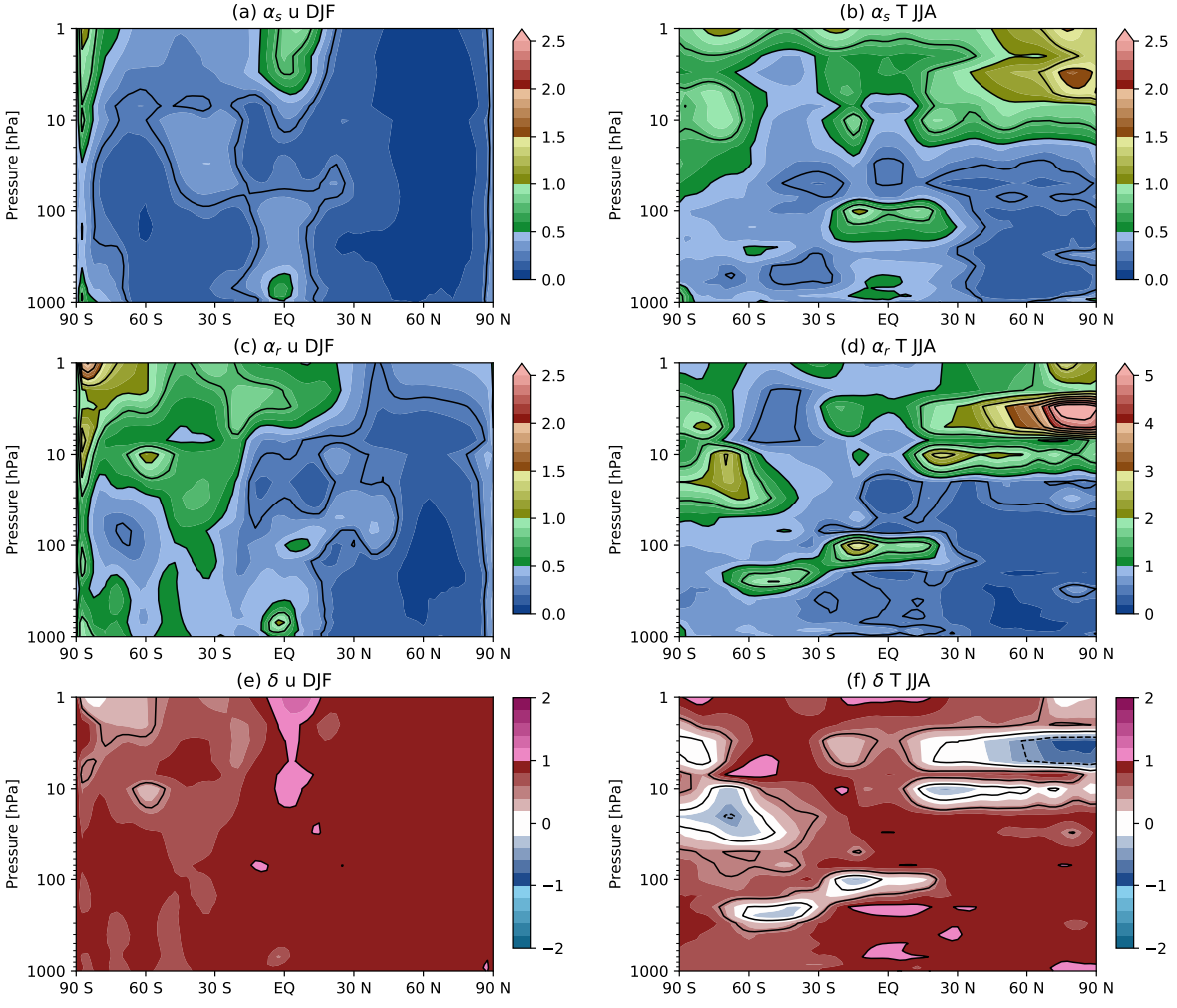

Figure 7. Ratios (a,c) γ and (b,d) α as defined in Section 3 for (a,b) zonal winds in DJF and (c,d) temperatures in JJA.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

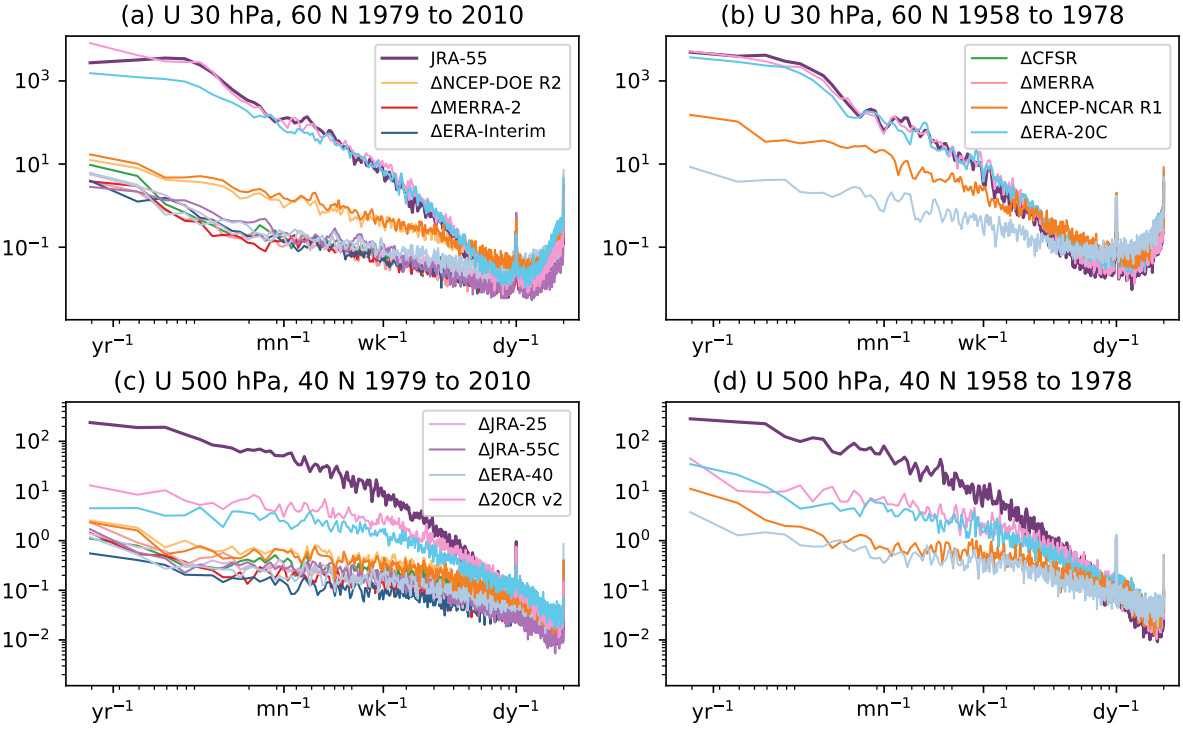

Figure 8. Power spectrum of winds in JRA-55 and of the differences in winds between other reanalyses and JRA-55 at (a,b) 30 hPa, 60 N and (c,d) 500 hPa, 40 N in the radiosonde and satellite periods, respectively. Note that the legend is divided across the panels but applies equally to each. Frequencies corresponding to periods of one year, one month (30 days), one week, and one day are indicated on the horizontal axis.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

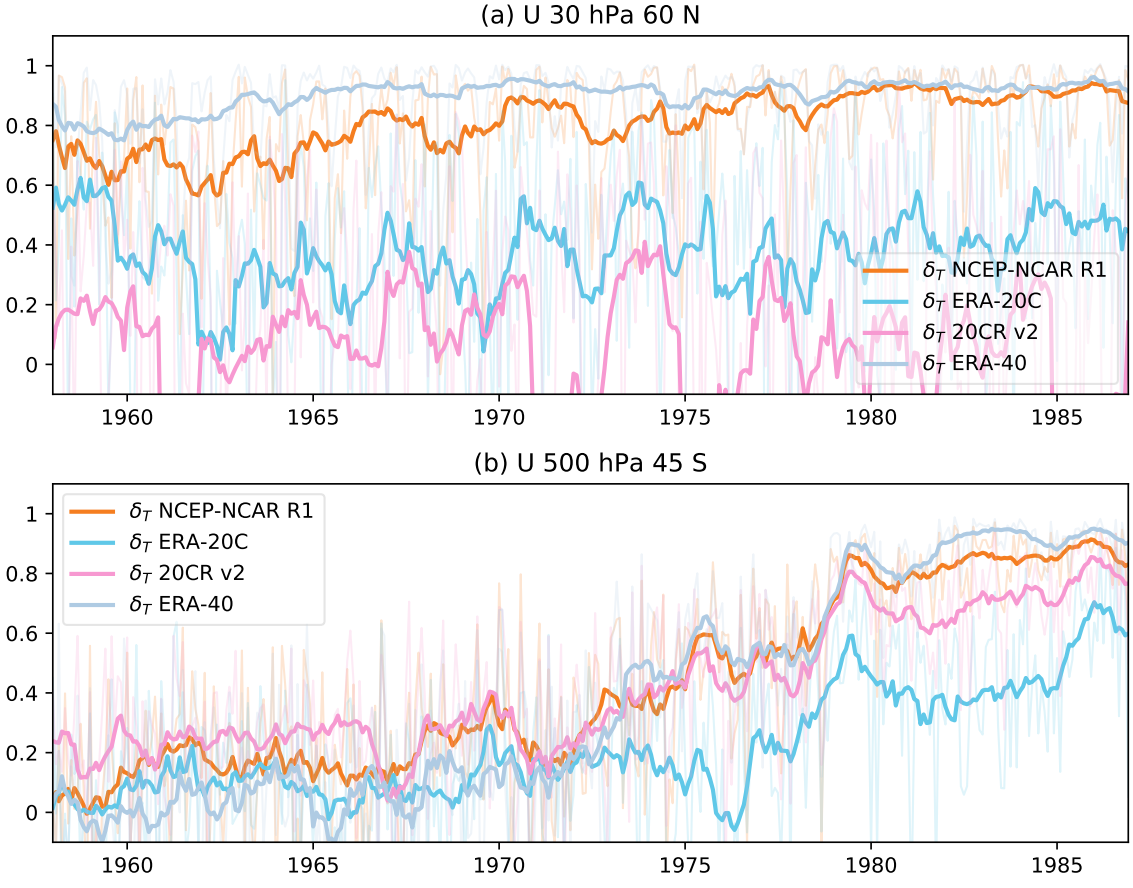

Figure 9. Time-dependent estimate of δ for (a) U at 30 hPa, 60° N and (b) U at 500 hPa, 45° S. The faint lines are computed based on month-by-month estimates of α_r, while bold lines are computed based 12-month running means of α_r. See text for details.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

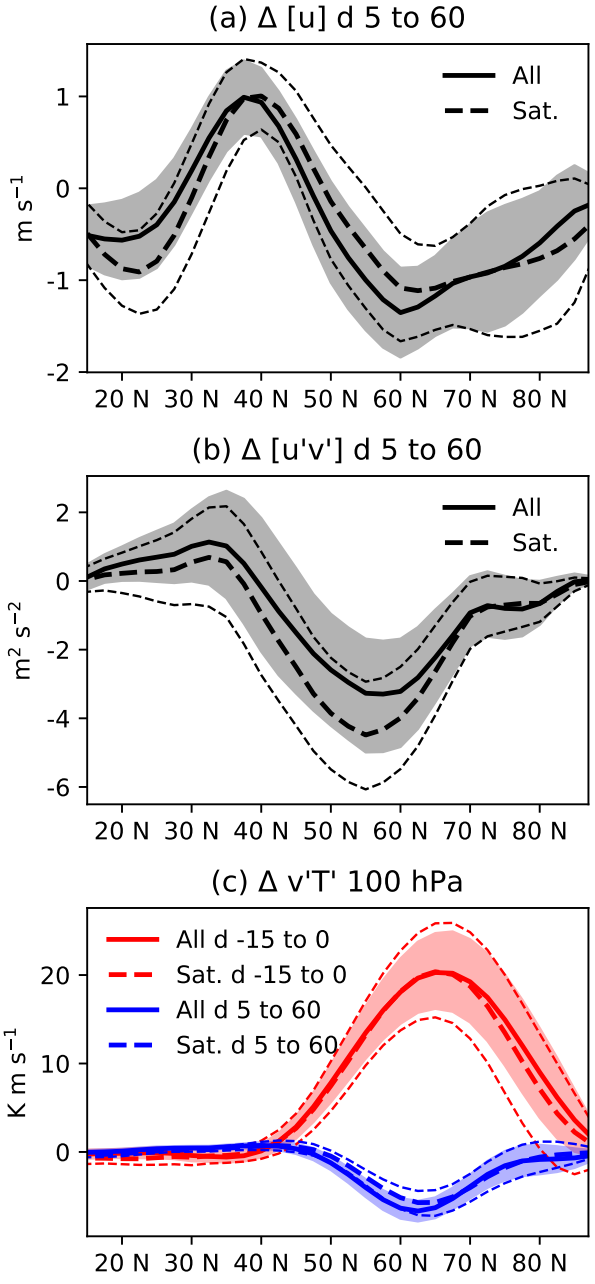

Figure 10. (a) Composite mean of vertically averaged zonal wind anomalies, averaged over lags 5 to 60 days following major warmings. Solid line shows the composite for all events while the dashed line shows the composite for the satellite era alone. Confidence intervals for the whole period are shaded while those for the satellite era are indicated by thin dashed lines. (b) Similar but for vertically integrated momentum fluxes. (c) Similar but for meridional heat fluxes at 100 hPa, averaged over lags -15 to 0 (in red), and over lags 5 to 60 (in blue). See text for details.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-879
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 11 October 2018

Atmospheric Chemistry and Physics Discussions

Table 1. Reanalysis products and dates considered in the present work. See Fujiwara et al. (2017) for a much more thorough discussion of the observations assimilated into each product. Abbreviations for certain products used within the text are indicated within parentheses.

Product (Label)	Reference	Centre	Dates considered	Classes of data assimilated
JRA-25	(Onogi et al., 2007)	JMA	01-1979 to 12-2010	All
JRA-55	(Kobayashi et al., 2015)	JMA	01-1958 to 12-2010	All
JRA-55C	(Kobayashi et al., 2014)	JMA	01-1979 to 12-2010	Conventional
MERRA	(Rienecker et al., 2011)	NASA GMAO	01-1979 to 12-2010	All
MERRA-2	(Gelaro et al., 2017)	NASA GMAO	01-1981[†] to 12-2010	All
ERA-40	(Uppala et al., 2005)	ECMWF	01-1958 to 08-2002	All
ERA-Interim	(Dee et al., 2011)	ECMWF	01-1979 to 12-2010	All
ERA-20C	(Poli et al., 2013)	ECMWF	01-1979 to 12-2010	Surface
NCEP-NCAR R1 (NCEP-NCAR)	(Kalnay et al., 1996)	NOAA/NCEP and NCAR	01-1979 to 12-2010	All
NCEP-DOE R2 (NCEP-DOE)	(Kanamitsu et al., 2002)	NOAA/NCEP and DOE	01-1979 to 12-2010	All
CFSR	(Saha et al., 2010)	NOAA/NCEP	01-1979 to 12-2010	All
NOAA-CIRES 20CR v2c (20CR v2)	(Compo et al., 2011)	NOAA and CIRES	01-1979 to 12-2010	Surface

[†] Although MERRA-2 includes 1980, there are spin-up issues in early 1980 which affect the Arctic vortex.