Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

Quantifying the variability of the annular modes: Reanalysis uncertainty vs. sampling uncertainty

Edwin P. Gerber[1] and Patrick Martineau[2]

[1]Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York NY 10012, USA
[2]Research Center for Advanced Science and Technology, University of Tokyo, Japan

Correspondence: Edwin P. Gerber (gerber@cims.nyu.edu)

Abstract. The annular modes characterize the dominant variability of the extratropical circulation in each hemisphere, quantifying vacillations in the position of the tropospheric jet streams and the strength of the stratospheric polar vortices. Their representation in all available reanalysis products is assessed. *Reanalysis uncertainty* associated with limitations in the ability to constrain the circulation with available observations, i.e., the inter-reanalysis spread, is contrasted with *sampling uncertainty* associated with the finite length of the reanalysis records.

It is shown that the annular modes are extremely consistent across all modern reanalyses during the satellite era (1979 onward). Consequently, uncertainty in annular mode variability, e.g., the coupling between the stratosphere and troposphere and the variation in the amplitude and time scale of jet variations throughout the annual cycle, is dominated by sampling uncertainty. Comparison of reanalyses based on conventional or surface observations alone with those using all available observations indicates that there is limited ability to characterize the Southern Annular Mode in the pre-satellite era. For the Northern Annular Mode, however, there is evidence that conventional observations are sufficient, at least from 1958 onward. The addition of two additional decades of records substantially reduces sampling uncertainty in several key measures of annular mode variability, demonstrating the value of more historic reanalyses. Implications for the assessment of atmospheric models and the strength of coupling between the surface and upper atmosphere are discussed.

1 Introduction

The annular modes characterize the dominant internal variability of the extratropical atmosphere (Thompson and Wallace, 2000). In the troposphere, they primarily characterize meridional shifts in the extratropical jet stream, a positive index indicating a poleward shift. The jet stream is associated with the extratropical storm tracks, such that the annular mode indices indicated shifts in storm activity, particularly in Northern Europe and eastern North America (e.g., Thompson and Wallace, 1998). In the stratosphere, the annular modes chiefly characterize variations in the strength of the polar vortex, a positive index indicating a

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

stronger than average vortex. Baldwin and Dunkerton (2001) revealed a downward influence of the stratospheric polar vortex on the tropospheric jet stream by comparing annular mode indices computed at different pressure levels.

In this study, we focus on the representation of the annular mode indices, or time series, in all available global reanalysis products. The spatial structure of the annular modes is characterized by a meridional dipole, where mass (or equivalently, momentum, as the large scale flow is geostrophically balanced) is exchanged between the high and low latitudes (Thompson and Wallace, 2000). It is can be understood as the gravest mode of a varying jet (e.g., Vallis et al., 2004; Gerber and Vallis, 2005) and is well captured across a range of models (e.g., Gerber et al., 2010). The term "annular mode" was derived from the annular structure of the patterns in longitude, where geopotential height varies in the same phase at a given latitude (Limpasuvan and Hartmann, 2000). As discussed by Deser (2000) and Ambaum et al. (2001), among others, such coherent fluctuations are generally not observed, and the annular structure is primarily statistical in nature (Gerber and Thompson, 2017). The temporal variability of the annular modes provides a convenient means to quantify coupling between tropospheric jets and the stratospheric vortices, and variations in the amplitude and persistence of the variability through the annual cycle. It has proven more challenging to understand and model, and is thus the focus of this study.

The jet streams are known to vacillate on slower time scales than individual weather systems, as first quantified with the "zonal index" (e.g., Rossby, 1939; Namias, 1950). A number of studies have investigated how this enhanced persistence results from a feedback between eddies and the mean flow (e.g., Robinson, 1996; Feldstein and Lee, 1998; Lorenz and Hartmann, 2001; Gerber and Vallis, 2007; Barnes et al., 2010; Zurita-Gotor et al., 2014). The stratospheric polar vortex exhibits greater memory than the troposphere, with implications for predictability of the tropospheric jet stream on sub-seasonal to seasonal time scales (e.g., Baldwin et al., 2003; Sigmond et al., 2013). Models, both idealized (e.g., Gerber et al., 2008b) and comprehensive (e.g., Gerber et al., 2008a), tend to exhibit too much persistence, with implications for their projected response to global warming (Kidston and Gerber, 2010; Barnes and Hartmann, 2010; Son et al., 2010). It is thus of both practical and theoretical interest to determine the limits to which we can quantify the temporal variability of the annular modes of our atmosphere with reanalyses.

After briefly describing the available reanalyses and our procedure for evaluating the annular modes in Sections 2 and 3, we explore annular mode variability in the most recent products of the four major reanalysis centers in Section 4, establishing a "Reanalysis Ensemble Mean" as a standard for comparison with all reanalysis products. In sections 5 and 6, we compare the reanalyses in the satellite and pre-satellite eras, respectively. We find that the Northern Annular Mode (NAM) can be well constrained with conventional observations, and appears consistent across a number of reanalyses since 1958. Characterization of Southern Annular Mode (SAM) variability, however, appears to depend on satellite observations, and the SAM index varies substantially between reanalysis records before 1979. In Section 7, we further explore historic reanalyses constrained only by surface measurements, and find that there is potential to constrain the annular modes in the Northern Hemisphere for earlier periods.

Given the state of the reanalyses, in Section 8 we show that sampling uncertainty is the leading source of uncertainty in our ability to constrain the temporal variability of the annular modes. We focus in particular on the downward coupling

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

associated with extreme vortex variability and the time scales of annular mode variability as a function of height and season. Our conclusions are summarized in Section 9.

2 The reanalyses

Table 1 lists the reanalyses compared in this study. Fujiwara et al. (2017) provide a detailed explanation of the data assimilation procedures and input data sets that differentiate these reanalysis products, but we highlight a few key differences here. ERA-Interim, JRA-55, MERRA2, and NCEP-CFSR represent the current state-of-the-art reanalysis products of the four major reanalysis centers. While this is an moving target, e.g., the new ERA5 reanalysis will soon replace ERA-I as the top product available from the European Center for Medium Range Weather Forecasting (ECMWF), a key finding of this paper is that all of the modern reanalyses accurately capture the variability of the annular modes.

In constructing a *re*analysis, one seeks a balance between the aim of providing a best possible estimate of the atmospheric state at any given time (as with an *analysis* for providing initial conditions for a weather forecast), with the goal of providing a homogeneous record of the atmospheric evolution. The latter requirement generally entails a choice of a fixed atmospheric model and assimilation procedure for the entire reanalysis. A more difficult decision concerns the use of new observations that become available during the record, the most notable change being the introduction of satellite measurements c. 1979. In general, the most state-of-art reanalysis make use of all observations as they become available, providing an increasingly accurate representation of the atmospheric state with time. This introduces a risk of spurious trends which are not associated with physical changes in the atmosphere, but rather an enhanced ability to observe the atmosphere.

A few reanalyses explicitly limit the observational inputs, giving us an opportunity to assess the impact of new observations. In particular, ERA-20C and NOAA-20CR v2 and v2c intentionally limit themselves to surface based observations only. This allows them to provide an estimate of the atmospheric state as far back as 1850. They provide us a unique opportunity to assess the coupling between the surface and interior circulation of the atmosphere, particularly the stratosphere.

The JRA-55C reanalysis uses the same model and data assimilation procedure as in JRA-55, but explicitly excludes all satellite measurements. It therefore provides an opportunity to assess the impact of satellite measurements on our ability to quantify the annular modes. The JRA-55AMIP "reanalysis" only considered observed sea surface temperatures as an observational input, i.e., it is the output of a free running atmospheric model constrained only by the sea surface temperature and observed changes in radiative forcing. As will be shown, it allows us to confirm the conventional wisdom that in general, the extratropical circulation is only weakly constrained by sea surface temperature.

3 The annular modes

As suggested by Baldwin and Thompson (2009), we use a simplified procedure for computing the annular mode based on the polar cap averaged geopotential height, where the cap is defined as all latitudes poleward of 65°. Following Gerber et al. (2010), we first remove the global mean geopotential height at each time step. This focusses the analysis on meridional variations in

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

geopotential height (i.e., shifts in mass), as opposed to systematic variations in the entire layer associated with changes in atmospheric temperature. In practice this only has a significant impact on the indices in the upper stratosphere and above, but also helps removes trends associated with global warming. Based on model integrations (e.g., Gerber et al., 2010), trends can become significant near the tropopause, where there is a sharp gradient in the character of the annular mode as it shifts from characterizing the tropospheric jet to the stratospheric vortex.

Fig. 1 illustrates our procedure, which we spell out precisely in the following.

1. 6 hourly output of geopotential height is averaged to form a daily time series $Z(t,\lambda,\phi,p)$, where λ, ϕ, and p refer to latitude, longitude, and pressure. For leap years, the 366 days are interpolated to 365, but we note the impact of this interpolation is inconsequential, and one could more simply remove one day from each leap year.

2. The annual cycle is formed by averaging each calendar day over the record, and then smoothing the resulting annual cycle with a 60 day low pass Lanczos digital filter. The anomalous height $Z'(t,\lambda,\phi,p)$ is formed by removing the smoothed annual cycle. The daily, zonal mean anomalous height is illustrated for a single year and pressure level in Fig. 1a.

3. We then compute the global mean geopotential height, $\overline{Z'}^{global}$, and austral and boreal polar cap heights, $\overline{Z'}^{SH}$ and $\overline{Z'}^{NH}$, as illustrated in Fig. 1b.

4. Raw SAM and NAM time series are defined by $-(\overline{Z'}^{SH} - \overline{Z'}^{global})$ and $-(\overline{Z'}^{NH} - \overline{Z'}^{global})$, respectively. The negative sign is in keeping with the standard convention of the annular mode introduced by Thompson and Wallace (1998).

5. The standardized SAM and NAM indices are obtained by dividing the raw indices by their standard deviations, yielding indices with unit variance by construction, as illustrated in Fig. 1c.

As steps 1-4 are linear, one can rearrange them to minimize the computational effort. In practice, we first compute the global and polar cap average geopotential heights, and then daily average and deseasonalize them to form time series $\overline{Z'}^{global}$, $\overline{Z'}^{SH}$, and $\overline{Z'}^{NH}$.

The formation of the annual cycle and normalization of the time series (steps 2 and 5) depend on the length of the record. If one follows this procedure to define the annular mode over two distinct periods that overlap, the resulting indices will not agree perfectly during the period of overlap. We find that that difference is inconsequential provided one uses a period of sufficient length, a decade or so in practice. This limitation could be addressed by defining the climatology and normalization constants over a set period, for instance the World Meteorological Organization (WMO) climatology spanning the last three full decades. For the purposes of this paper, we have computed the indices separately for each period of comparison, e.g., 1980-2016 for our comparison of the modern reanalyses in Section 4, or for 1981-2010 for our comparison of all reanalyses in Section 5.

We focus on a subset of the pressure levels that were shared by all reanalyses: 1000, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100, 70, 50, 30, 20, 10, 7, 5, 3, 2, and 1 hPa. Levels above 10 hPa are unavailable for NCEP-R1/R2 and NOAA 20CR v2/v2c reanalyses. The annular mode indices are fairly uniform within the troposphere and stratosphere, respectively, and such fine vertical resolution is really only needed in the tropopause region.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

Lastly, we noted that our definition of the annular mode (or equivalently, previous Empirical Orthogonal Function based definitions) require extrapolation of data below the surface. This was done by the reanalysis centers for all reanalyses with the exception of MERRA products. We have opted to omit MERRA and MERRA2 from comparisons below 700 hPa, where their data was incomplete.

4 Consistency in the representation of the annular modes in state-of-the-art reanalyses

We first compare the four modern reanalyses, ERA-Interim, JRA-55, MERRA2, and NCEP-CFSR, to justify the use of "Reanalysis Ensemble Mean" (REM) annular mode indices as a benchmark of comparison. For a comprehensive comparison, we used the full period shared by all of these reanalyses, 1980-2016. Note that the NCEP-CFSR reanalysis was upgraded substantially in 2010, but we see no evidence of a discontinuity in the annular mode variability across this break.

Fig. 2 shows the pairwise squared correlation (R^2) of the NAM and SAM indices as a function of pressure. The squared correlation coefficient indicates the fraction of variance shared by the two time series. We find that the modern reanalyses share approximately 96% of the variance or greater at all levels in both hemispheres, with the exception of the upper stratosphere in the austral hemisphere.

In the Northern Hemisphere, the agreement is 99% or better between ERA-Interim, JRA-55, and MERRA2, with slightly weaker correlation to NOAA-CFSR in the upper troposphere and lower stratosphere, where there is a minimum in correlation at 200 hPa near the extratropical tropopause. NOAA-CFSR, however, does not stand out from the other reanalyses in the austral hemisphere. There is a tendency towards better agreement in the lower and mid stratosphere as compared to the troposphere in both hemispheres, but we stress that the correlation is always extremely high.

Given the strong agreement of all the modern reanalyses, we adopted a REM annular mode index based on the average of ERA-Interim, JRA-55, and NOAA-CFSR. MERRA2 was omitted from the REM due to missing data at lower levels, but the results are nearly identical if it is included. We emphasize that this decision does not imply a value judgement of the quality of MERRA2 reanalysis.

The intensity of annular mode variability changes throughout the annual cycle, particularly in the stratosphere (e.g., Baldwin et al., 2003; Gerber et al., 2010). This can be seen in the annual cycle of annular mode variance in Fig. 3 (black contours). In the troposphere, the annular modes are most variable in the winter seasons in both hemispheres, with a greater variation over the annual cycle in the boreal hemisphere. In the stratosphere, there is considerably more variation throughout the annual cycle: the maximum is approximately 4 times the annual mean, and variability drops to near zero in the summer. There is also a notable difference in the timing of peak variability between the two hemispheres. For the NAM, the maximum is collocated with the troposphere in the winter, while for the SAM, variability is greatest in late spring.

Given this variation of the annular modes within the annual cycle, we check the consistency between the reanalyses as a function of day of year, illustrated by the shading in Fig. 3. We find that the indices agree well ($R^2 > 0.95$) at all levels and heights except when the variability is weak, where on average, the pairwise R^2 drops as low as 0.5. Thus, there is great certainty in the annular mode state except in periods when the annular mode is inconsequential.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

5 Comparing the representation of the annular mode indices across all reanalyses

Fig. 4 shows the R^2 correlation between the annular mode indices computed from each individual reanalysis with the REM annular mode index over the period 1981-2010. This period, which corresponds to the standard 30 year climatological evaluation period by WMO convention, also provides optimal overlap between the reanalyses: all but one (ERA-40) are available the entire period. The correlation for ERA-40 was based on the years when it was fully available, 1981-2001. We enumerate the key findings below.

(1) The NAM is well captured throughout the troposphere and the bulk of the stratosphere by all reanalyses that assimilate free atmospheric data. The R^2 correlation with the REM exceeds 0.975 from the surface to 10 hPa for all reanalyses excepting ERA-20C, JRA55AMIP, and NOAA 20CR v2/2c. The modern reanalyses are somewhat better correlated with the REM than earlier generation products, particularly in the lower to mid-stratosphere. While this could partly arise by construction (the REM is based on the modern reanalyses), MERRA and MERRA2 are not a part of the REM and still exhibit enhanced correlation. MERRA appears slightly more highly correlated than MERRA2, although the difference is trivial.

(2) The SAM is well captured by most reanalyses that assimilate satellite measurements, and is demonstratively better represented in the "second generation" reanalyses. Most of the comprehensive reanalyses are well correlated with the REM ($R^2 > 0.92$ up to 10 hPa), but in comparison to the Northern Hemisphere, the modern reanalyses are consistently better than earlier reanalyses; MERRA and MERRA2 are better correlated with the REM that the first generation reanalyses (ERA-40, JRA-25, NOAA-R1 and R2) at all levels.

(3) Satellite measurements are critical for the representation of the SAM, but not for the NAM. The representation of the NAM by JRA-55C, which assimilates only conventional observations, is nearly indistinguishable from the second generation reanalyses up to 30 hPa, and is still very good up to 10 hPa. We conclude that conventional observations are sufficient until the upper stratosphere, which suggests that there is potential for a skillful representation of the NAM in the pre-satellite period, as further investigated in Section 6. In the austral hemisphere, however, the representation of SAM in JRA-55C is demonstrably poorer at all levels (though it still captures more than 85% of the variance up to 3 hPa). JRA-55C is most strongly correlated with the other reanalyses in the mid-stratosphere, but decays considerably towards the surface, where it only captures 85% of the variance, substantially less than the reanalyses which assimilate only surface measurements.

(4) Reanalyses based on surface measurements alone remain highly correlated with the REM in the troposphere, but the correlation falls off sharply above the tropopause, particularly for the 20CR reanalyses. ERA-20C, however, remains remarkably well correlated through the stratosphere, capturing more than half the variance of the REM up to the stratopause at 1 hPa. As discussed in greater detail in Section 7, this indicates that the stratosphere is sufficiently coupled to the surface to extract useful information about its state from surface data alone.

We finally note that sea surface temperatures are insufficient to capture AM variability, at least in the case of the JRA-55AMIP integration. It has long been known that the extratropical circulation is not driven by local SSTs (e.g., Barsugli and Battisti, 1998). Tropical SSTs are known to influence the annular modes (e.g., El Niño; L'Heureux and Thompson, 2006), but appear insufficient to constrain the annular mode variability of either hemisphere.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

6 Pre-satellite representation of the annular modes

The ability of the JRA-55C to accurately capture the NAM without the aid of satellite measurements suggests that it should be possible to assess it during the pre-satellite era. Fig. 5 compares the 5 available reanalyses during the period 1958-1978. As it is less trivial to identify a meaningful REM during this period, we consider two comparisons against a given individual reanalysis, JRA-55 and NCEP-R1, respectively. Comparable results are found when using ERA-40 as a reference, and the selection here is not meant to be a value judgment. Based on Fig. 5a and b, we concluded that the NAM is consistently represented in ERA-40, JRA-55, and NCEP-R1, giving us confidence that there are sufficient observations to quantify the NAM as far back as 1958.

While the NAM in the surface-only reanalyses remains highly correlated in the troposphere, we note that ERA-20C is not as well correlated in the stratosphere as it was in more recent years. At 10 hPA, ERA-20C only captures about 40% of the variance, as compared to 60% between 1981 to 2010.

The situation is quite different in the austral hemisphere, with widespread divergence between the reanalyses (Fig. 5c,d). JRA-55 shares only 30% of the variance with the other reanalyses throughout the troposphere. NCEP-R1 is more strongly correlated with the NOAA 20CR and ERA-20C reanalyses in the troposphere (sharing approximately 60% of the variance), but poorly correlated with the other full reanalyses (JRA-55 and ERA-40). We conclude that the SAM is poorly constrained in the pre-satellite era. This is partly expected, given the limited ability of JRA-55C to capture the SAM during the satellite era. JRA-55C was still fairly well correlated with the REM in the satellite era, however, suggesting that the poor representation of the SAM before 1979 also reflects a dearth of conventional observations.

As seen in Fig. 4d, the SAM in JRA-55C is better correlated with the REM in the stratosphere than it is in the troposphere during the satellite era. Similarly puzzling behavior is observed in the pre-satellite era: reanalyses which do take in measurements from the free atmosphere are better correlated in the stratosphere than below. We speculate that this is due to the broadening of the spatial structure of the SAM with height (e.g., Gerber et al., 2010, their Fig. 4). Consequently, observations from the midlatitudes and tropics are sufficient to constrain the annular mode at higher elevations, and reanalyses do not suffer from limited observations over Antarctica.

7 Representation of the annular modes in surface-only reanalyses

We explore the ability of reanalyses that only incorporate surface observations to capture annular mode variability in Figs. 6 and 7. The former compares two years of the NAM indices represented by ERA-20C with those of the REM at 1000, 100, and 10 hPa. As indicated in Fig. 4a, ERA-20C captures approximately 80% of the variability of the NAM at 100 hPa, and 60% of the variability at 10 hPa during the satellite era. The accuracy of ERA-20C appears to fluctuate from year to year. In this example, the winter of 2007-8 is captured with remarkable fidelity, while during the winter of 2006-7 there is an overall negative bias at upper levels, particularly at 10 hPa, although much of the high frequency variability is still captured.

This specific period was selected in part to highlight the ability of the ERA-20C to capture Stratospheric Sudden Warming events, marked by the yellow circles in the 10 hPa indices. ERA-20C captures the two events over this period, at least within a few days of the modern reanalyses. The annular modes in the NOAA 20CR reanalyses (not shown) are comparable to ERA-20C

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

in the troposphere (ERA-20C doing slightly better in the Northern Hemisphere, NOAA-20CR slightly better in the Southern Hemisphere), but NOAA-20CR struggles to capture variability above the tropopause (Fig. 4). As found by Butler et al. (2017), the polar vortex in NOAA-20CR v2c is too strong (or equivalently, exhibits a cold bias), and exhibits only one major warming event between 1958 and 2011.

In Fig. 7, we compare the evolution of correlation between the annular mode indices in ERA-20C with other reanalyses that extend into the pre-satellite era. An 11 year moving window allows us to observe how the correlation changes over time. Focussing first on the boreal hemisphere, we find that the NAM at the surface (solid lines) is well constrained in all the reanalyses from 1950 on. The R^2 correlation with NOAA-20CR is relatively high throughout the record, albeit decaying to about 60% by the start of the century. This suggests that there may be sufficient observations in the first half of the century to capture the majority of annular variability.

The correlation between the ERA-20C NAM at 100 and 10 hPa with conventional reanalyses NCEP-R1 and JRA-55 does appear to weaken as one moves back in time. At 10 hPa, the correlation drops noticeably around 1975. We have no means to assess the skill of NCEP-R1 before 1958. Hence the drop in correlation with ERA-20C around 1950 could be due to a loss in skill in either (or both) reanalyses. As noted above, the NOAA-20CR reanalyses exhibit limited skill in the stratosphere, and correlation with ERA-20C is low at all times.

In the Southern Hemisphere, the 1000 hPa annular mode indices in ERA-20C and the NOAA-20CR reanalyses are well correlated to about 1950, at which point correlation falls off to near zero. We conclude there is insufficient information to constrain the austral circulation before this date.

ERA-20C diverges from the conventional reanalysis (JRA-55 and NCEP-R1) earlier, but recall that JRA-55C (which is equivalent to JRA-55 in the pre-satellite era) had difficulty capturing the annular mode during the satellite era. We thus have little reason to trust the conventional reanalysis over ERA-20C and NOAA-20CR. ERA-20C appears to capture much of the variability at higher levels until about 1980, at which time the correlation drops off abruptly. As discussed in the previous section, we have little reason to trust any of the conventional reanalyses above the surface in the pre-satellite era.

8 Sampling uncertainty vs. reanalysis uncertainty

We conclude our study by focusing on the factors limiting our ability to quantify the behavior of the annular modes in our atmosphere. We contrast uncertainty associated with the finite length of the observational record, or "sampling uncertainty" with spread in the statistics between the reanalyses, or "reanalysis uncertainty". We focus exclusively on the most modern reanalyses, but the results from Section 5 suggest that nearly all the reanalyses capture variability of the annular modes fairly well during the satellite era.

8.1 Coupling between the stratospheric polar vortices and tropospheric jets

We first consider the uncertainty in the downward coupling between the stratosphere and troposphere associated with weak and strong vortex events, as explored by Baldwin and Dunkerton (2001). Fig. 8a,b illustrates composites of the annular mode indices

as a function of height for weak vortex events in the JRA-55 reanalysis, chosen because it offers the longest record (1958-2016). Weak events, associated with Sudden Stratospheric Warmings (SSWs) and strong events are defined as in Baldwin and Dunkerton (2001): when the annular mode index at 10 hPa dips below (rises above) -3.0 (1.5) standard deviations, respectively. A 30 day seperation is required between events.

The extended record affords approximately twice as many events available to Baldwin and Dunkerton (2001). The key structure of the composite remains about the same, with two notable exceptions. Following weak vortex events, the response of the tropospheric annular mode appears more abruptly, trailing just by a few days, although it still clearly lags weakening of the polar vortex in the stratosphere. The faster response at the surface is consistent with composites of hundreds of events from Coupled Model Intercomparison Project, Phase 5 (CMIP5) integrations (Charlton-Perez et al., 2013, their Fig. 5).

The coupling between the troposphere and stratosphere also appears to be more simultaneous for strong vortex events, which are themselves notably less "event-like". The build up of the polar vortex is less abrupt than its destruction in SSWs, making the detection of strong vortex events more sensitive to the selection criteria. For example, the separation requirement plays a more important role in the definition of strong vortex events than for weak vortex events. The higher than average vortex preceding event onset could potentially be somewhat reduced by requiring a longer window between events.

Fig. 8c,d shows the event-to-event standard deviation of the annular mode index as a function of height and lag. White delineates the region where the variability is near unity, equal to the climatological value. Inter-event variance is reduced at 10 hPa at the time of onset by construction, but is otherwise generally at or above the climatological value. The mean impact of events on the troposphere (i.e., the signal) is approximately 0.3 to 0.6 standard deviations in amplitude, and thus easily overwhelmed by this natural variability. Hence many SSWs will not appear to be associated with a downward propagating signal.

Note that we have normalized the variance to be of order unity in the annual mean. As seen in Fig. 3, the variance is above average in the winter, when strong and weak vortex events occur. The average tropospheric value ($\sim$1.2) is consistent with the mean winter variability, and the enhanced variance of the polar vortex approximately 20-40 days after strong vortex events is comparable with the mean variability of the winter vortex.

In Fig. 9 we contrast the sampling uncertainty in these events (as determined from the JRA-55 reanalysis) with two measures of uncertainty associated with differences between the reanalyses. The sampling uncertainty (panels a,b) is expressed as a one standard deviation error bound. It is simply the event-to-event standard deviation shown in Fig. 8c,d divided by the square root of the number of events. For both weak and strong vortex events, the sampling error in the troposphere is fairly uniform in time. For weak vortex events, the average standard error over the 181 day period surrounding events is 0.21 at 300 hPa; this level is characteristic of other tropospheric levels. The average uncertainty is 0.12 for strong events, smaller because the sample size is nearly three times as large, a consequence of the weaker event threshold.

For computing the "reanalysis uncertainty" we restrict the analysis to the common period 1980-2016. Fig. 9c,d shows what we have termed the "raw" uncertainty: the standard deviation between composites of weak/strong vortex constructed independently from the four modern reanalyses. At 300 hPa, the average uncertainty over the -90 to 90 day period is just 0.025

for weak events (8 times less than the sample uncertainty) and 0.057 for strong vortex events (less than half the respective sampling uncertainty).

Much of this inter-reanalysis spread, however, arises from the fact that the composites are not necessarily built about the same events. As emphasized by Martineau et al. (2018, their Fig. 4), differences between reanalyses are much smaller than differences between events. The threshold criteria is sensitive to subtle changes in the annular mode index, such that not all reanalyses indicate the same event dates. With respect to weak vortex events, 19 were identified in all the reanalyses, 13 of which fell on the same day across all of them. For one event, the date at which the 10 hPa NAM index crossed the -3 standard deviation threshold was offset by 1 day in ERA-I reanalysis, and in 5 other events, the date in NOAA-CFSR was offset by one day.

Given the large variability of the NAM on daily time scales, these subtle offsets matter. When the event dates are fixed across all the reanalyses, the inter-reanalysis spread is further reduced, as shown in Fig. 9e,f. Here, we again plot the standard deviation between composites of weak/strong vortex events across the reanalyses, but now using identical events dates. (The dates were set to day where the event fell in the majority of the reanalysis). For weak vortex events, the average tropospheric uncertainty is reduced to 0.016, over 50% less than in the raw composite shown in Fig. 9c, and 13 times less than the sampling uncertainty.

Strong vortex "events" are less event-like by nature, leading to greater spread between the reanalyses. Analyzing each reanalysis separately, we identified 49 events in ERA-40, 52 in JRA-55 and MERRA2, and 51 in NCEP-CFSR. The spread between these composites, Fig. 9d, thus includes some sampling uncertainty (i.e. different events), and variations due to differences in timing (which were up to a week or so in many cases).

To remove this sampling uncertainty, we focus on 53 unique events which were identified by at least 2 reanalyses. There was universal agreement on the date for 20 of these events, and three reanalyses identified the same date in 11 more cases. For the other 24 cases, at least two reanalyses differed on the timing and/or missed the event all together. When composites of weak vortex events were formed for each reanalysis using the same dates (Fig. 9f), the standard deviation between reanalyses drops substantially. At 300 hPa, the average uncertainty was just 0.011, about half the level of the smallest contour on the plot, and about 10 times less than the sampling uncertainty.

For both weak and strong vortex events, then, the uncertainty associated with differences between the reanalyses is a factor of 10 less than uncertainty associated with sampling in the extended 59 year JRA-55 record. As the sampling uncertainty decays with the square of the number of events, we would need a record that is approximately 100 times longer (i.e., 6 millenia of observations) before sampling uncertainty is reduced to the same level as the reanalysis uncertainty!

As sampling uncertainty is by far the limiting factor, the fact that high quality reanalysis of the Northern Hemisphere appears possible back to 1958 is important. The addition of 21 years lengthens the record by approximately 50% percent, thereby reducing the sampling uncertainty by approximately 20% percent.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

8.2 The time scales of the annular modes

As noted in the introduction, the temporal variability of the annular modes indicates the strength of eddy-mean flow feedbacks in the troposphere, and has proven to be a challenging test for atmospheric models. Baldwin et al. (2003) quantify the persistence variations in jets by the e-folding time scale, τ, of the annular mode index. It can be computed as a function of height, to contrast variability in the stratosphere and troposphere, and calendar date, to capture seasonality of the jet variations.

The annular modes in the troposphere exhibit the longest time scales (Fig. 10a,b) during the seasons when the stratosphere is most variable (Fig. 3): January-February in the boreal hemisphere and November-December in the austral hemisphere. This suggests a downward influence of the stratosphere during the seasons when coupling is most active. Gerber et al. (2010) found that models share this behavior, although both the peak in stratospheric variability and in tropospheric time scale tend to reach a maximum later in the annual cycle, particularly in the austral hemisphere.

As noted by Keeley et al. (2009), the e-folding measure τ is influenced by variability on both synoptic and interannual time scales, and so unable to comprehensively characterize the temporal variability (cf. Osprey and Ambaum, 2011). To limit the influence of low frequency variations, here we compute the time scale on non-overlapping decadal periods, averaging them to produce Fig. 10a,b. A calculation based on the full record (thus allowing the influence of time scales greater than a decade) exhibits the same structure, but with greater τ values in regions more sensitive to low frequency variations and trends: the boreal stratosphere summer (when variability is weak) and austral spring, when stratospheric ozone is known to influence the flow.

The use of individual decades allows us to provide a rough estimate on the uncertainty associated with these time scales by computing the standard deviation between decades. In the Southern Hemisphere we only trust the satellite era, giving us only three samples, the minimum needed to estimate the variance. In the Northern Hemisphere, however, we can take advantage of reanalyses since 1958, giving us 5 full decades. The sampling uncertainty is substantial. We quantify it in relative terms in Fig. 10c,d, which allows us to effectively compare both tropospheric and stratospheric values in one plot.

For the NAM, relative uncertainty in the time scale τ is between 10 and 20% over most of the annual cycle, and approaches 40% in the lower stratosphere during summer (a time, however, when the annular mode variability is quite weak). The longer time scales in the stratospheric winter, however, are comparatively robust. For the SAM (where the three decades give us only a rudimentary estimate of uncertainty), the tropospheric time scale – notably the marked increase in November and December – still appears statistically robust. There is great uncertainty in the early spring values of τ in the stratosphere, and τ is poorly constrained in austral spring (February-March) throughout the atmosphere.

We contrast the sampling uncertainty in τ with the spread associated with different reanalyses in the lower panels. As with the vortex event composites, we quantify the reanalysis uncertainty by the standard deviation of τ values across the four modern reanalyses (averaged for three separate decadal calculations for which all reanalyses are available). The reanalyses agree very well with each other, with a relative uncertainty less than 0.05 at most pressure levels and times during the year. There is some disagreement on time scales in the boreal stratosphere during summer (when variability is weak, see also Fig. 3) and in austral

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

stratosphere during winter. This latter case is the only instance where the reanalysis uncertainty becomes comparable to the sampling uncertainty.

The fact that sampling uncertainty generally dwarfs the reanalysis uncertainty indicates that the main limiting factor is the length of the observational record. Two additional decades of high quality reanalysis appear to be possible in the boreal hemisphere, based in our analysis in Sections 6, giving the potential to increase the full record from approximately 40 years to 60. A 50% increase in sample size reduces the uncertainty in the mean by $\sqrt{1/1.5}$, or about 80%.

9 Conclusions

We have shown that the annular modes are very well represented in the modern, second generation reanalyses. ERA-I, JRA-55, MERRA2, and NOAA-CFSR can be used interchangeably for the purposes of quantifying the annular modes of our atmosphere. As detailed in Section 5, the other reanalyses are also quite good during the satellite era, albeit demonstrably less consistent in the austral hemisphere.

As discussed in Section 6, reanalysis of the Northern Hemisphere from 1958-1979 appears to be very consistent. The fact that the JRA-55C reanalysis (which explicitly excludes satellite observations) can match the full reanalyses in recent decades suggests that conventional observations are sufficient for a highly accurate representation of the Northern Annular Mode. In contrast, there is very little consistency in the representation of the Southern Annular Mode before satellite observations were available.

In Section 8, we highlighted the fact that our ability to quantify the downward coupling and temporal variability of the annular modes is limited primarily by sampling uncertainty. This suggests that two additional decades of reanalysis have substantial value for quantifying the variability of the boreal jet stream and polar vortex. This is important for both understanding the coupling between eddies and the mean flow (e.g., Lorenz and Hartmann, 2001) and testing the fidelity of atmospheric models (e.g., Gerber et al., 2010). We encourage reanalysis centers to consider extending their modern products back until 1958 or earlier.

Finally, we have documented the ability of reanalyses constrained only by surface observations to capture variability of the annular modes in the free troposphere and stratosphere. Both NOAA-20CR and ERA-20C accurately capture annular mode variability up to the extratropical tropopause. NOAA-20CR, both versions 2 and 2c, rapidly looses the ability to track the annular modes above the tropopause. ERA-20C, however, appears to capture much of the variability high into the stratosphere, notably capturing more than 50% of the variance as high as 10 hPa in both hemispheres.

The ability of surface-only reanalyses to constrain the stratosphere reflects the tight coupling between the troposphere and stratosphere. It should not be interpreted in a causal sense, i.e., that half of stratospheric variability is caused by surface observations. The issue of causality is important in that a number of studies have shown that modifying the stratosphere (for example, by nudging it toward observations or toward its climatology) can influence the tropospheric flow. While nudging is not equivalent to data assimilation, caution should be exercised in interpreting these experiments as one layer of the atmosphere driving another.

Data availability. The geopotential height data used to compute the annular modes is publicly available at the Centre for Environmental Data Analysis (Martineau, 2017).

Competing interests. There are no competing interest present.

Acknowledgements. EPG acknowledges support from the US National Science Foundation through grant AGS-1546585 to New York University. We thank the reanalysis centers for making their data publicly available, and the Statosphere-troposphere Processes And their Role in Climate (SPARC) activity of the World Climate Research Program for their help in organizing the SPARC-Reanalysis Intercomparison Project, which inspired this research.

References

Ambaum, M. H. P., Hoskins, B. J., and Stephenson, D. B.: Arctic Oscillation or North Atlantic Oscillation?, J. Climate, 14, 3495–3507, 2001.

Baldwin, M. P. and Dunkerton, T. J.: Stratospheric Harbingers of Anomalous Weather Regimes, Science, 294, 581–584, 2001.

Baldwin, M. P. and Thompson, D. W. J.: A critical comparison of stratosphere-troposphere coupling indices, Quart. J. Roy. Meteor. Soc., 135, 1661–1672, 2009.

Baldwin, M. P., Stephenson, D. B., Thompson, D. W. J., Dunkerton, T. J., Charlton, A. J., and O'Neill, A.: Stratospheric Memory and Skill of Extended-Range Weather Forecasts, Science, 301, 636–640, 2003.

Barnes, E. A. and Hartmann, D. L.: Testing a theory for the effect of latitude on the persistence of eddy-driven jets using CMIP3 simulations, Geophys. Res. Lett., 37, L15 801, https://doi.org/10.1029/2010GL044144, 2010.

Barnes, E. A., Hartmann, D. L., Frierson, D. M. W., and Kidston, J.: Effect of latitude on the persistence of eddy-driven jets, Geophys. Res. Lett., 37, L11 804, https://doi.org/10.1029/2010GL043199, 2010.

Barsugli, J. J. and Battisti, D. S.: The basic effects of atmosphere–ocean thermal coupling on midlatitude variability, J. Atmos. Sci., 55, 477–493, 1998.

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

Butler, A. H., Sjoberg, J. P., Seidel, D. J., and Rosenlof, K. H.: A sudden stratospheric warming compendium, Earth Syst. Sci. Data, 9, 63–76, https://doi.org/10.5194/essd-9-63-2017, 2017.

Charlton-Perez, A. J., Baldwin, M. P., Birner, T., Black, R. X., Butler, A. H., Calvo, N., Davis, N. A., Gerber, E. P., Gillet, N., Hardiman, S., Kim, J., Krueger, K., Lee, Y.-Y., Manzini, E., McDonald, B. A., Polvani, L. M., Reichler, T., Shaw, T. A., Sigmond, M., Son, S.-W., Tohey, 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, https://doi.org/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, Quart. J. Roy. Meteor. Soc., 137, 1–28, https://doi.org/10.1002/qj.776, 2011.

Compo, G. P., Whitaker, J., Sardeshmukh, P., Giese, B., and Brohan, P.: Intercomparison of an improved 20th Century Reanalysis version "2c" (1850-2012), American Meteorological Society Annual meeting, Phoenix, AZ, 2015.

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, Quart. J. Roy. Meteor. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Deser, C.: On the teleconnectivity of the "Arctic Oscillation", Geophysical Research Letters, 27, 779–782, https://doi.org/10.1029/1999GL010945, 2000.

Feldstein, S. B. and Lee, S.: Is the Atmospheric Zonal Index Driven by an Eddy Feedback?, J. Atmos. Sci., 55, 3077–3086, 1998.

Fujiwara, M., Wright, J. S., Manney, G. L., Gray, L. J., Anstey, J., Birner, T., Davis, S., Gerber, E. P., Harvey, V. L., Hegglin, M. I., Homeyer, C. R., Knox, J. A., Krüger, K., Lambert, A., Long, C. S., Martineau, P., Molod, A., Monge-Sanz, B. M., Santee, M. L., Tegtmeier, S., Chabrillat, S., Tan, D. G. H., Jackson, D. R., Polavarapu, S., Compo, G. P., Dragani, R., Ebisuzaki, W., Harada, Y., Kobayashi, C., McCarty, W., Onogi, K., Pawson, S., Simmons, A., Wargan, K., Whitaker, J. S., and Zou, C.-Z.: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, Atmos. Chem. Phys., 17, 1417–1452, https://doi.org/10.5194/acp-17-1417-2017, 2017.

Gerber, E. P. and Thompson, D. W. J.: What Makes an Annular Mode "Annular"?, J. Atmos. Sci., 74, 317–332, https://doi.org/10.1175/JAS-D-16-0191.1, 2017.

Gerber, E. P. and Vallis, G. K.: A Stochastic Model for the Spatial Structure of Annular Patterns of Variability and the NAO, J. Climate, 18, 2102–2118, https://doi.org/10.1175/JCLI3337.1, 2005.

Gerber, E. P. and Vallis, G. K.: Eddy-Zonal Flow Interactions and the Persistence of the Zonal Index, J. Atmos. Sci., 64, 3296–3311, https://doi.org/10.1175/JAS4006.1, 2007.

Gerber, E. P., Polvani, L. M., and Ancukiewicz, D.: Annular Mode Time Scales in the Integovernmental Panel on Climate Change Fourth Assessment Report Models, Geophys. Res. Lett., 35, L22 707, https://doi.org/10.1029/2008GL035712, 2008a.

Gerber, E. P., Voronin, S., and Polvani, L. M.: Testing the Annular Mode Autocorrelation Timescale in Simple Atmospheric General Circulation Models, Mon. Wea. Rev., 136, 1523–1536, https://doi.org/10.1175/2007MWR2211.1, 2008b.

Gerber, E. P., Baldwin, M. P., Akiyoshi, H., Austin, J., Bekki, S., Braesicke, P., Butchart, N., Chipperfield, M., Dameris, M., Dhomse, S., Frith, S. M., Garcia, R. R., Garny, H., Gettelman, A., Hardiman, S. C., Karpechko, A., Marchand, M., Morgenstern, O., Nielsen, J. E., Pawson, S., Peter, T., Plummer, D. A., Pyle, J. A., Rozanov, E., Scinocca, J. F., Shepherd, T. G., and Smale, D.: Stratosphere-Troposphere Coupling and Annular Mode Variability in Chemistry-Climate Models, J. Geophys. Res., 115, D00M06, https://doi.org/10.1029/2009JD013770, 2010.

Kalnay, E., Kanamitsua, M., Kistlera, R., Collinsa, W., Deavena, D., Gandina, L., Iredella, M., Sahaa, S., Whitea, G., Woollena, J., Zhua, Y., Leetmaaa, A., Reynoldsa, R., Chelliahb, M., Ebisuzakib, W., Higginsb, W., Janowiakb, J., Mob, K., Ropelewskib, C., Wangb, J., Jennec, R., and Josephc, D.: The NCEP/NCAR 40-year reanalysis project, Bull. Am. 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. Am. Meteor. Soc., 83, 1631–1643, 2002.

Keeley, S. P. E., Sutton, R. T., and Shaffrey, L. C.: Does the North Atlantic Oscillation show unusual persistence on intraseasonal timescales?, Geophys. Res. Lett., 36, L22 706, https://doi.org/10.1029/2009GL040367, 2009.

Kidston, J. and Gerber, E. P.: Intermodel Variability of the Poleward Shift of the Austral Jet Stream in the CMIP3 Integrations Linked to Biases in 20th Century Climatology, Geophys. Res. Lett., 37, L09 708, https://doi.org/10.1029/2010GL042873, 2010.

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

L'Heureux, M. L. and Thompson, D. W. J.: Observed Relationships between the El Niño–Southern Oscillation and the Extratropical Zonal-Mean Circulation, J. Climate, 19, 279–287, 2006.

Limpasuvan, V. and Hartmann, D. L.: Wave-Maintained Annular Modes of Climate Variability., J. Climate, 13, 4414–4429, 2000.

Lorenz, D. J. and Hartmann, D. L.: Eddy-zonal flow feedback in the Southern Hemisphere, J. Atmos. Sci., 58, 3312–3327, 2001.

Martineau, P.: S-RIP: Zonal-mean dynamical variables of global atmospheric reanalyses on pressure levels, Centre for Environmental Data Analysis, May 2018, https://doi.org/10.5285/b241a7f536a244749662360bd7839312, 2017.

Martineau, P., Chen, G., Son, S.-W., and Kim, J.: Lower-stratospheric control of the frequency of sudden stratospheric warming events, J. Geophys. Res. Atm., https://doi.org/10.1002/2017JD027648, 2018.

Namias, J.: The Index Cycle and its Role in the General Circulation, J. Meteorology, 7, 130–139, 1950.

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

Osprey, S. and Ambaum, M. H. P.: Evidence for the chaotic origin of Northern Annular Mode variability, Geophys. Res. Lett., 38, https://doi.org/10.1029/2011GL048181, 2011.

Poli, P., Hersbach, H., Tan, D., Dee, D. P., Berrisford, P., Simmons, A., Vitart, F., Laloyaux, P., Tan, D. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the 20th century, J. Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, 2016.

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., Koster, J. J. R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, J. Climate, 24, 3624–3648, https://doi.org/10.1175/JCLI-D-11-00015.1, 2011.

Robinson, W. A.: Does Eddy Feedback Sustain Variability in the Zonal Index, J. Atmos. Sci., 53, 3556–3569, 1996.

Rossby, C. G.: Relations between Variation in the Intesity of the Zonal Circulation and the Displacements of the Semi-permanent Centers of Action, J. Marine Res., 2, 38–55, 1939.

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., Delst, P. V., 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. Am. Meteor. Soc., 91, 1015–1057, https://doi.org/10.1175/2010BAMS3001.1, 2010.

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

Sigmond, M., Scinocca, J. F., Kharin, V. V., and Shepherd, T. G.: Enhanced seasonal forecast skill following stratospheric sudden warmings, Nature Geosci., https://doi.org/10.1038/NGEO1698, 2013.

Son, S.-W., Gerber, E. P., Perlwitz, J., Polvani, L. M., Gillet, N., Seo, K.-Y., Eyring, V., Shepherd, T. G., Waugh, D., Akiyoshi, H., Austin, J., Baumgaertner, A., Bekki, S., Braesicke, P., Brühl, C., Butchart, N., Chipperfeld, M. P., Cugnet, D., Dameris, M., Dhomse, S., Frith, S., Garny, H., Garcia, R., Hardiman, S. C., Jöckel, P., Lamarque, J. F., Mancini, E., Marchand, M., Michou, M., Nakamura, T., Morgenstern,

O., Pitari, G., Plummer, D. A., Pyle, J., andJ. F. Scinocca, E. R., Shibata, K., Smale, D., Teyssèdre, H., Tian, W., and Yamashita, Y.: The Impact of Stratospheric Ozone on Southern Hemisphere Circulation Change: A Multimodel Assessment, J. Geophys. Res., 115, D00M07, https://doi.org/10.1029/2010JD014271, 2010.

Thompson, D. W. J. and Wallace, J. M.: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields., Geophys. Res. Lett., 25, 1297–1300, 1998.

Thompson, D. W. J. and Wallace, J. M.: Annular modes in the extratropical circulation. Part I: Month-to-month variability., J. Climate, 13, 1000–1016, 2000.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Van De Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Holm, 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, Quart. J. Roy. Meteor. Soc., 131, 2961–3012, 2005.

Vallis, G. K., Gerber, E. P., Kushner, P. J., and Cash, B. A.: A Mechanism and Simple Dynamical Model of the North Atlantic Oscillation and Annular Modes., J. Atmos. Sci., 61, 264–280, 2004.

Zurita-Gotor, P., Blanco-Fuentes, J., and Gerber, E. P.: The Impact of Baroclinic Eddy Feedback on the Persistence of Jet Variability in the Two-Layer Model, J. Atmos. Sci., 71, 410–429, https://doi.org/10.1175/JAS-D-13-0102.1, 2014.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

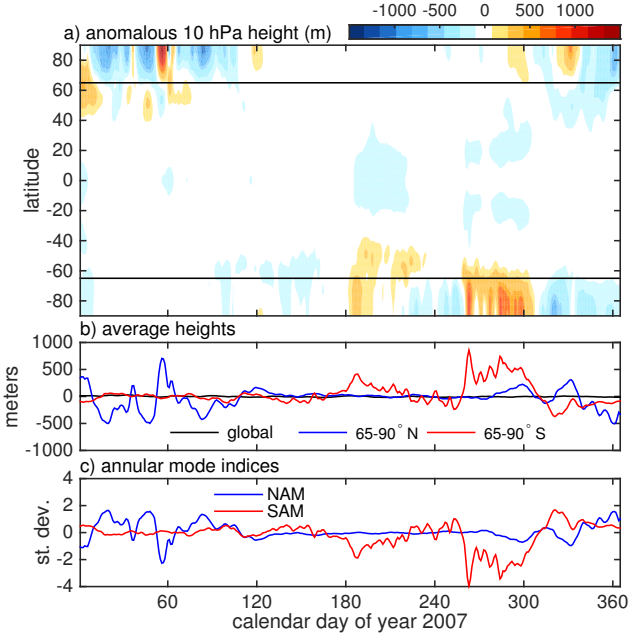

Figure 1. To illustrate our simplified method for calculating the annular mode indices, (a) shows anomalous zonal mean geopotential height, $\overline{Z'}$, at 10 hPa for one year, 2007. As detailed in the text, anomalies are defined relative to a smoothed annual cycle. Panel (b) shows global and polar cap averages of Z' for this period, and (c) the Northern and Southern Annular Mode indices at this level, defined as -1$\times$ the respective polar cap averages less the global mean, and then normalized to have unit variance. The choice of year 2007 was arbitrary, but did exhibit a Sudden Stratospheric Warming on 24 February, which is associated with high geopotential height over the polar cap and low NAM index around day 55.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

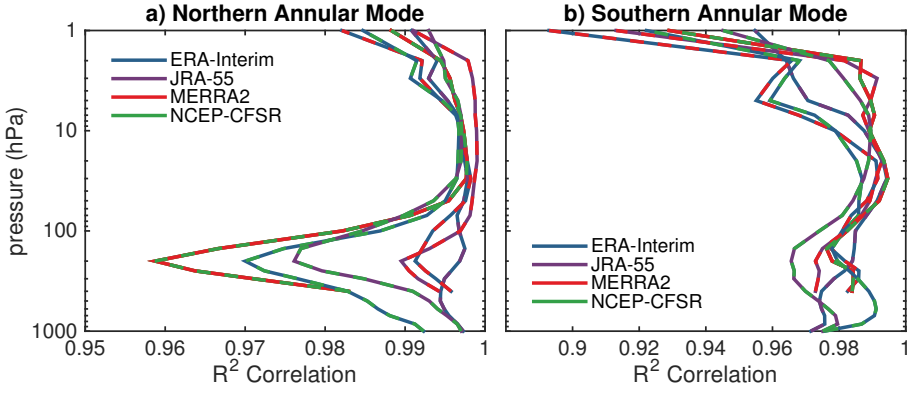

Figure 2. Consistency of (a) the Northern and (b) Southern Annular Mode indices based on the modern reanalyses, as a function of pressure. The pairwise correlation coefficient (R^2) between ERA-I, JRA-55, MERRA2, and NCEP-CFSR are plotted; the dashed colors indicating the respective pairs, e.g., green-red dashing indicates the square correlation between MERRA2 and NCEP-CFSR.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

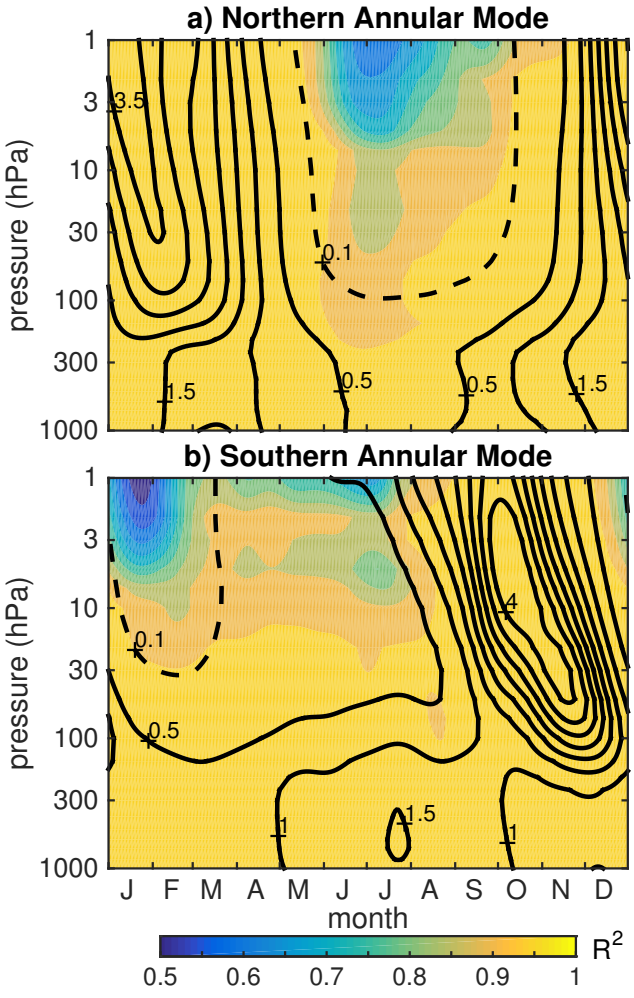

Figure 3. Consistency of (a) the Northern and (b) Southern Annular Mode indices in the modern reanalyses, as a function of pressure and time of year, for the period 1980 to 2016. The shading indicates the average pairwise R^2 correlation between ERA-I, JRA-55, MERRA2, and NCEP-CFSR. Contours indicate the variance of the REM annular mode, which has been normalized to have unit variance in the annual mean at each level; the interval is 0.5 units of variance, with an additional dashed contour at 0.1.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

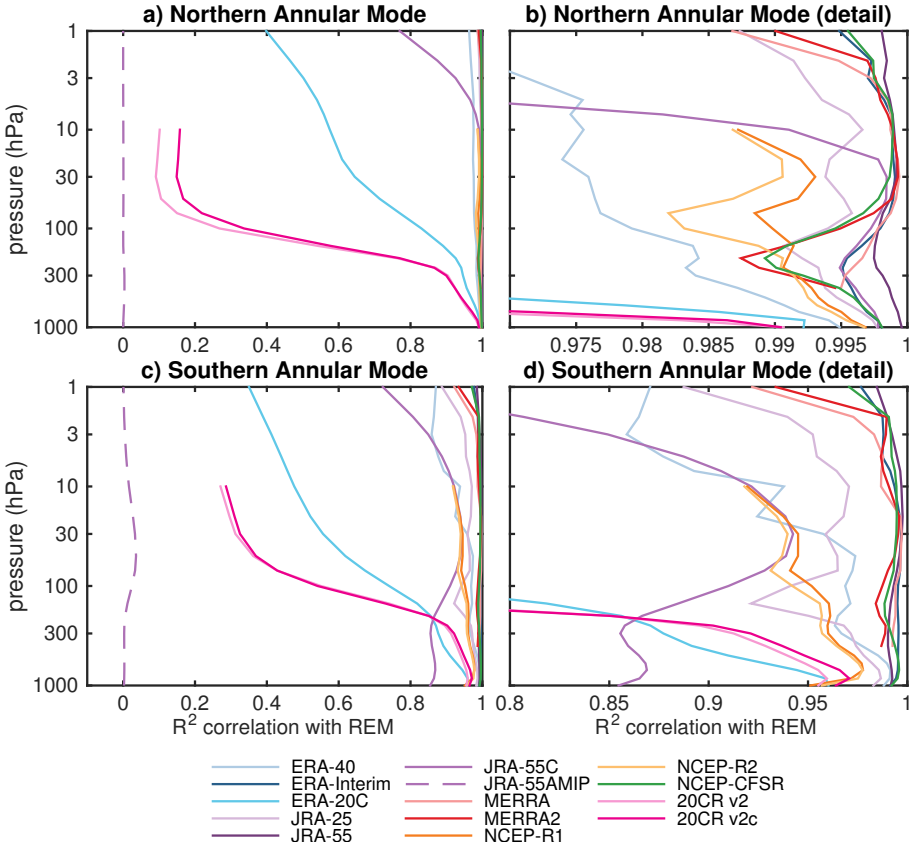

Figure 4. The squared correlation between the (a,c) Northern and (b,d) Southern Annular Mode indices computed from each individual reanalysis with the Reanalyses Ensemble Mean (REM). We use the standard WMO three decade climatological period, 1981-2010, for all reanalyses except ERA-40, where analysis is based on 1981-2001. Panels (b,d) show the same data, but zoomed in to highlight the more accurate reanalyses. Note the difference in the range of the abscissa: the NAM is very well constrained in all reanalyses that assimilate free atmospheric observations.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

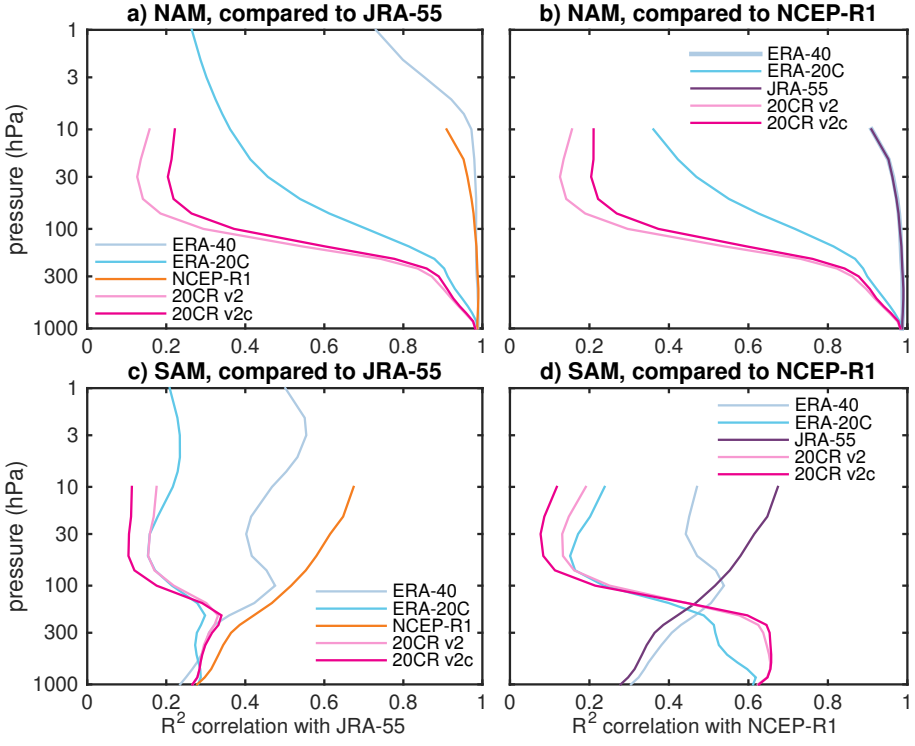

Figure 5. The squared correlation between (top) Northern and (bottom) Southern Annular Mode indices in various reanalyses with (left) JRA-55 and (right) NCEP-R1 during the pre-satellite era, 1958-1978.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

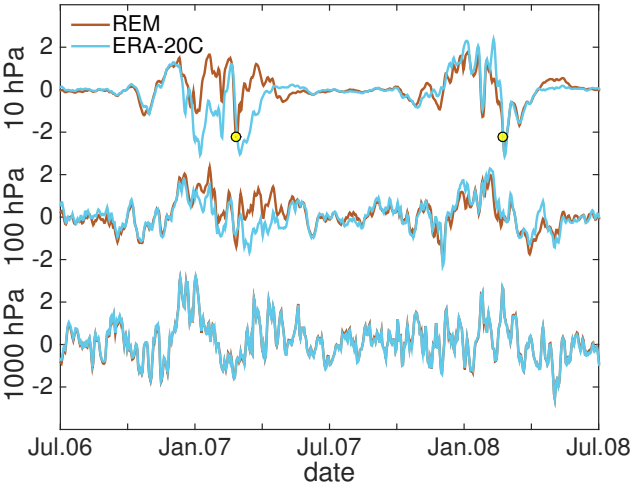

Figure 6. Comparison between the Northern Annular Mode timeseries from the REM and ERA-20C at 10, 100, 1000 hPa over a two year period, 1 July 2006 to 1 July 2008. Yellow dots indicate the occurrence of two Sudden Stratospheric Warming events during this period.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

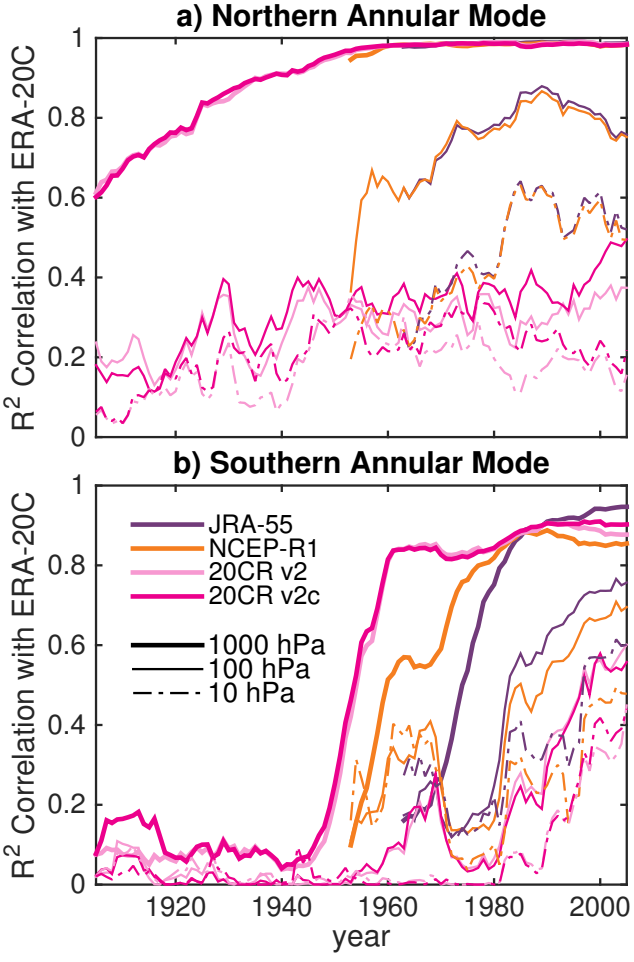

Figure 7. Correlation of the (top) NAM and (bottom) SAM computed from ERA-20C, as a function of time, with JRA-55, NCEP-R1, and the NOAA 20C reanalyses. The correlation is computed over a moving 11 year window, centered about the time given on the x-axis, i.e., the values at 1980 correspond to correlation between 1975 and 1985. ERA-40 was not included, as it provides comparable information to JRA-55 and NCEP-R1.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

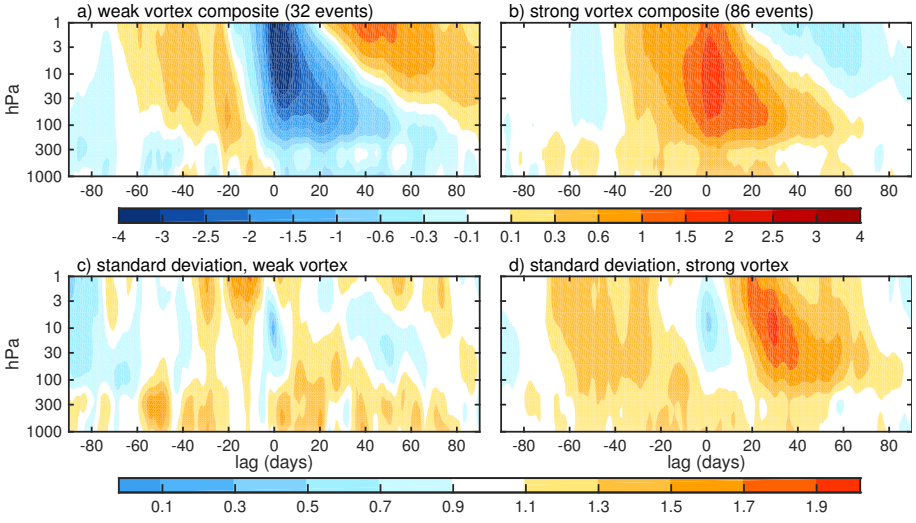

Figure 8. Composites of the Northern Annular Mode indices as a function of lag and pressure for (a) weak and (b) strong vortex events, based on JRA-55 reanalyses over the period 1958-2016. Following Baldwin and Dunkerton (2001), weak (strong) events are identified when the NAM index at 10 hPa drops below -3 (rises above 1.5), and must be separated by a minimum of 30 days. Panels (c) and (d) show the event-to-event standard deviation for weak and strong vortex events, respectively.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

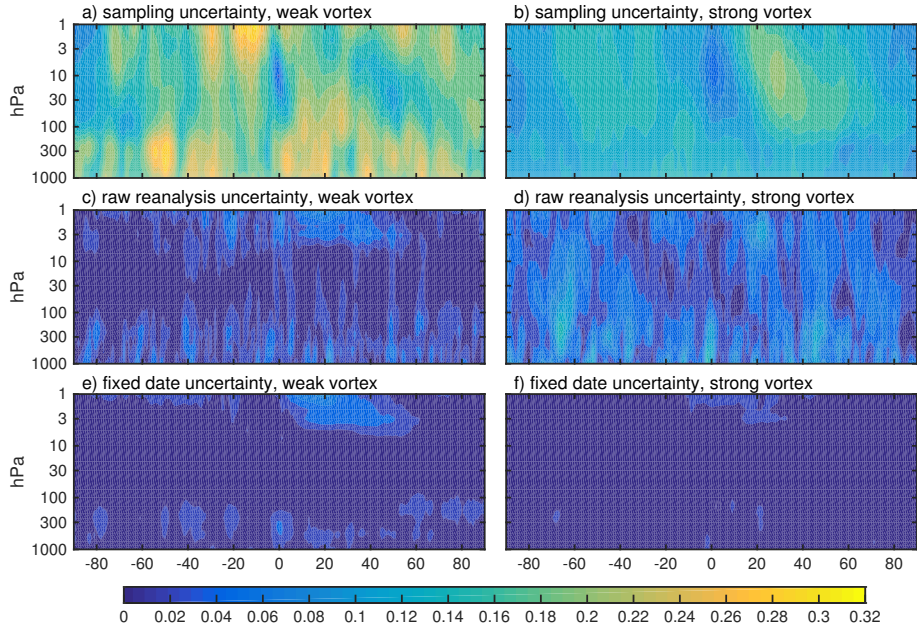

Figure 9. Three estimates of the uncertainty in the NAM index evolution around weak and strong events. (a) and (b) show the *sampling uncertainty* in the mean weak/strong composites shown in Fig. 8a,b, expressed as a 1 standard deviation error bound. Panels (c) and (d) show a first estimate of the *reanalysis uncertainty*: the standard deviation between composites of weak/strong vortex composites based on the 4 modern reanalyses (ERA-I, JRA-55, MERRA2, and NOAA-CFSR) over the period 1980-2016, where events are determined independently in each reanalysis. Panels (e) and (f) show a refined estimate of the *reanalysis uncertainty*: the standard deviation of weak/strong vortex composites based on the 4 modern reanalyses, but now using a standardized set of event dates. See the text for further details.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

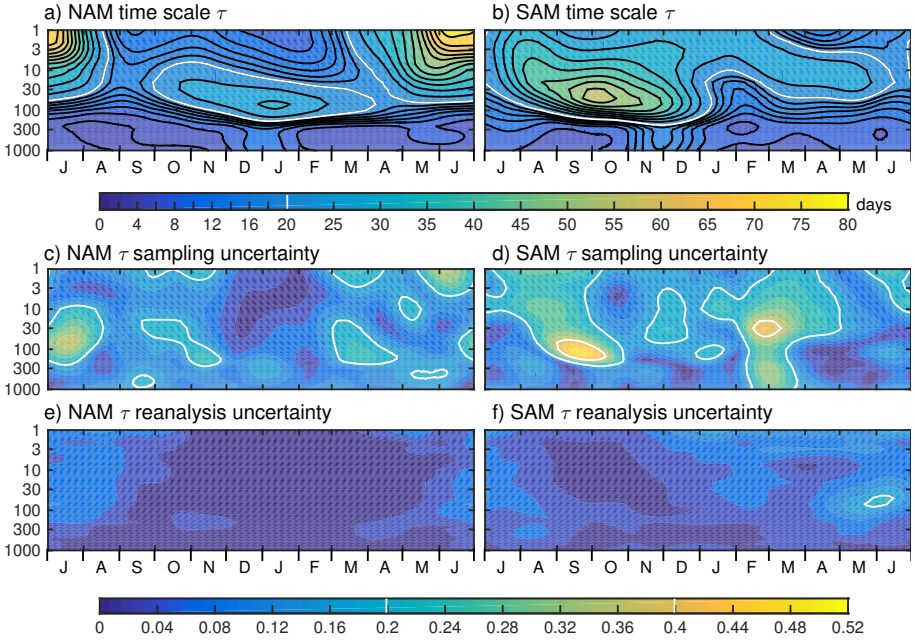

Figure 10. The annual cycle in the time scale of annular mode variability, τ, and relative uncertainty associated with the finite record and spread between the reanalyses. As detailed in the text, the top panels show the decorrelation time scale τ of the (a) Northern and (b) Southern Annular Modes in JRA-55 analysis, based on the average of 5 and 3 independent decadal calculations (1960-9, 1970-9, ..., 2000-2009 for the NH, only the last three decades for the SH). The middle panels show the sampling uncertainty of τ for (c) the NAM and (d) the SAM: the standard deviation of the decadal calculations divided by the square root of the sample size (5 and 3, respectively). We show the uncertainty in a relative sense, normalizing by the mean time scale shown in the upper panels; otherwise errors in the troposphere (where τ is smaller) appear insignificant. The bottom panels show the reanalysis uncertainty in τ for the (e) NAM and (f) SAM: the standard deviation of τ between calculations using ERA-I, JRA-55, MERRA2, and NOAA-CFSR. The standard deviation was computed across three separate decadal calculations (1980-9, 1990-9, and 2000-9) and then averaged.

Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-585
Manuscript under review for journal Atmos. Chem. Phys.
Discussion started: 27 June 2018

Atmospheric Chemistry and Physics Discussions

Table 1. The atmospheric reanalyses analyzed in this study.

Reanalysis System	Reference
ERA-40	Uppala et al. (2005)
ERA-Interim	Dee et al. (2011)
ERA-20C	Poli et al. (2016)
JRA-25	Onogi et al. (2007)
JRA-55	Kobayashi et al. (2015)
JRA-55C	Kobayashi et al. (2014)
JRA-55AMIP	Kobayashi et al. (2015)
MERRA	Rienecker et al. (2011)
MERRA2	Bosilovich et al. (2015)
NCEP-NCAR R1	Kalnay et al. (1996)
NCEP-DOE R2	Kanamitsu et al. (2002)
NCEP-CFSR*	Saha et al. (2010, 2014)
NOAA-20CR v2	Compo et al. (2011)
NOAA-20CR v2c	Compo et al. (2015)

*NOAA-CFSR underwent an upgrade starting at 1 Jan. 2011, as documented in Saha et al. (2014).