CHAPTER 4

Stratospheric Dynamics

Lead Authors: Neal Butchart & Andrew J. Charlton-Perez

Co-authors: Irene Cionni
Steven C. Hardiman
Kirstin Krüger
Paul Kushner
Paul Newman
Scott M. Osprey
Judith Perlwitz
Fabrizo Sassi
Michael Sigmond
Lei Wang

4.1 Introduction

This chapter assesses and compares the abilities of the Climate-Chemistry Models (CCMs) to reproduce the climate, circulation and associated variability of the stratosphere, though not the coherent naturally occurring variability such as that resulting from El-Niño Southern Oscillation (ENSO) events, big volcanic eruptions and variations in solar irradiance. This coherent natural variability is assessed separately in Chapter 8. The assessment in this chapter is process based, *i.e.*, the underlying dynamical processes occurring in the model stratospheres are evaluated as well as the simulation of the basic meteorological quantities such as winds and temperature. The processes and quantities considered (see **Table 4.1a, b** for a full list) are those relevant for modelling the long-term behaviour of stratospheric ozone (*e.g.*, temperature and the Brewer-Dobson circulation) and the impact of stratospheric change on surface climate (*e.g.*, annular modes). The chapter also looks at the predicted effects of climate change and ozone depletion/recovery on these modelled dynamical quantities and processes. In particular linear trends are calculated for many of the diagnostics assessed for the periods of ozone depletion (1980-1999), strong ozone recovery (2000-2049) and longer-term ozone and climate changes (2050-2099). Because of limited space, the dynamical meteorology

Table 4.1a: Climatological mean dynamical processes and/or phenomena validated in this chapter. The first column lists the processes and phenomena plus the subsection where the analysis can be found. Diagnostics used in the validation are listed in column 2 while columns 3 and 4 indicate which diagnostics will be used as quantitative metrics for the overall model assessment (see Section 4.5.3). All the diagnostics are validated against one or more of the reanalysis data sets introduced in Section 4.2. Abbreviations: NH=Northern Hemisphere; SH=Southern Hemisphere; DJF=December-January-February; MAM=March-April-May; JJA=June-July-August; SON=September-October-November; EP=Eliassen-Palm; PSC=polar stratospheric cloud; NAT=nitric acid trihydrate.

Mean climate			
Phenomena/Process	**Diagnostic**	**Metric Name**	**Metric Description**
4.3.1 Zonal mean climatology	DJF & MAM temperatures 60–90°N	**nhtemp**	DJF value at 50 hPa
	JJA & SON temperatures 60–90°S	**shtemp**	SON value at 50 hPa
	Date of the transition from eastward to westward winds at 60°S		
	Latitude & maximum wind speed of the NH jet in DJF	**umax_djf**	Maximum DJF eastward wind at 10 hPa
	Latitude & maximum wind speed of the SH jet in JJA	**umax_jja**	Maximum JJA eastward wind at 10 hPa
4.3.2 Stationary waves Zonal asymmetries	Location & maximum amplitude of the stationary wave field for the NH DJF & SH SON climatology		
	Phase & amplitude of wave-1 & wave-2 10 hPa NH DJF & SH SON stationary waves		
	Seasonal variation of the maximum amplitude of the NH & SH 10 hPa climatological stationary waves		
4.3.3 Brewer-Dobson circulation Tropical upwelling	70 hPa residual vertical velocity $\overline{w}^*$: annual mean 40°S–40°N, seasonal cycle in the turn-around latitudes where $\overline{w}^* = 0$ & tropical upwelling mass flux		
	Tropical upwelling mass flux at 70 & 10 hPa and downward control estimates of the driving from resolved (~EP-flux divergence) & parameterised (gravity) waves	**upwell_70**	Annual mean upwelling mass flux at 70 hPa
		upwell_10	Annual mean upwelling mass flux at 10 hPa
4.3.4 Extra-tropical wave driving	100 hPa meridional heat flux for January in the NH & July in the SH		
	Regression (slope & intercept) of the February & March 50 hPa temperatures 60°N–90°N on the 100 hPa January and February heat flux 40°N–80°N	**PW_nh**	Slope of the regression fit
	Regression (slope & intercept) of the August & September 50 hPa temperatures 60°S–90°S on the 100 hPa July and August heat flux 40°S–80°S	**PW_sh**	Slope of the regression fit
4.3.5 PSC threshold temperatures	Seasonally accumulated area at 50 hPa where temperatures are below 195 K (threshold for NAT clouds) & below 188 K (threshold for ice clouds)		

Table 4.1b: As Table 4.1a but for climate variability on intra-seasonal to interannual time scales. Abbreviations: EOF=empirical orthogonal function; SAO=semi-annual oscillation; QBO=quasi-biennial oscillation.

Climate variability (intraseasonal - interannual)			
Phenomena/Process	**Diagnostic**	**Metric Name**	**Metric Description**
4.4.1 Extra-tropical variability	Latitude & amplitude of the maximum interannual standard deviation of the zonal-mean zonal wind in DJF poleward of 45°N & in JJA from 30°S to 80°S		
	Eigenvalue of the leading mode of variability of the 50 hPa zonal-mean zonal wind for the NH & SH	**firsteval_nh**	Amplitude of first EOF in the NH
		firsteval_sh	Amplitude of first EOF in the SH
	Fraction of the total variance explained by first and second EOFs		
	Regression patterns of first and second mode of variability of the 50 hPa zonal-mean zonal wind for the NH & SH regions poleward of 45°		
4.4.2 Tropical variability	Vertical profile of the interannual standard deviation of the zonal-mean zonal wind, 10°S–10°N		
	Vertical profile of the amplitude of the annual cycle in the zonal-mean zonal wind, 10°S–10°N	**tann**	Amplitude at 2 hPa
	Vertical profile of the amplitude of the SAO in the zonal-mean zonal wind, 10°S–10°N	**sao**	Amplitude at 1 hPa
	Vertical profile of the amplitude of "QBO" in the zonal-mean zonal wind, 10°S–10°N	**qbo**	Amplitude at 20 hPa
4.4.3 Stratospheric sudden warmings	Frequency per year of NH major stratospheric sudden warmings, defined using reversal of the zonal-mean zonal wind at 10 hPa, 60°N	**SSW**	Mean frequency at 10 hPa, 60°N
	Monthly-distribution of NH major stratospheric sudden warmings		
4.4.4 Final warming	Mean date of the NH & SH final warmings defined using the criteria of Black and McDaniel (2007a, b)	**final_nh**	Mean date at 50 hPa, 60°N
		final_sh	Mean date at 50 hPa, 70°S

of the troposphere *per se* is not considered in any detail. Nonetheless, the stratosphere can have a direct impact on the mean climate and variability of the troposphere, and this is considered in Chapter 10.

4.2 Evaluation data sets and analyses

Although the diagnostics in this chapter are varied in their scope and time scale, their common theme is that they consider stratospheric and tropospheric dynamics on long, climate-relevant time scales, and on large spatial scales. With this in mind, a short survey of the data sets available to validate the models is useful. Individual diagnostic studies will use the data set most appropriate to their needs. Where possible at least two different data sets will be used to validate each diagnostic, so that some indication of the level of agreement between the data sets, and uncertainty in the observations, can be ascertained.

A previous SPARC report undertook an extensive comparison of middle atmosphere climatologies derived from the different data sets (Randel *et al*., 2004; SPARC, 2002). That report is used as a guide for determining biases in particular data sets. In addition to the data sets considered by Randel *et al*. (2004), a further reanalysis conducted by the Japanese Meteorological Agency (JRA-25) is now available and extends to 0.4 hPa. General conclusions from Randel *et al*. (2004) can be summarized as follows:

1. Reanalysis data sets with high model tops (ECMWF ERA-40 (Uppala *et al*., 2005) and Met Office (UKMO) Stratospheric Analyses (Swinbank and O'Neill, 1994)) have the best overall performance in comparison to rocketsonde and lidar measurements.
2. Particular uncertainties occur in the lower stratosphere and near the stratopause, and more care should be taken here.
3. The Quasi-Biennial and Semi-Annual Oscillations (QBO and SAO) are poorly captured in many of the available data sets when compared to Singapore radiosonde data.

Aspects of stratospheric data sets not considered by Randel *et al*. (2004) include the variability on both daily and interannual time scales. A brief inter-comparison of this variability in five analysis and reanalysis data sets (ECMWF ERA-40, NCEP/NCAR reanalysis, Met Office operational analyses, JRA-25, and ERA-Interim) shows a remarkable agreement between the different data sets. This agreement suggests that choosing any of the five reanalysis data sets would be appropriate when validating variability in the models.

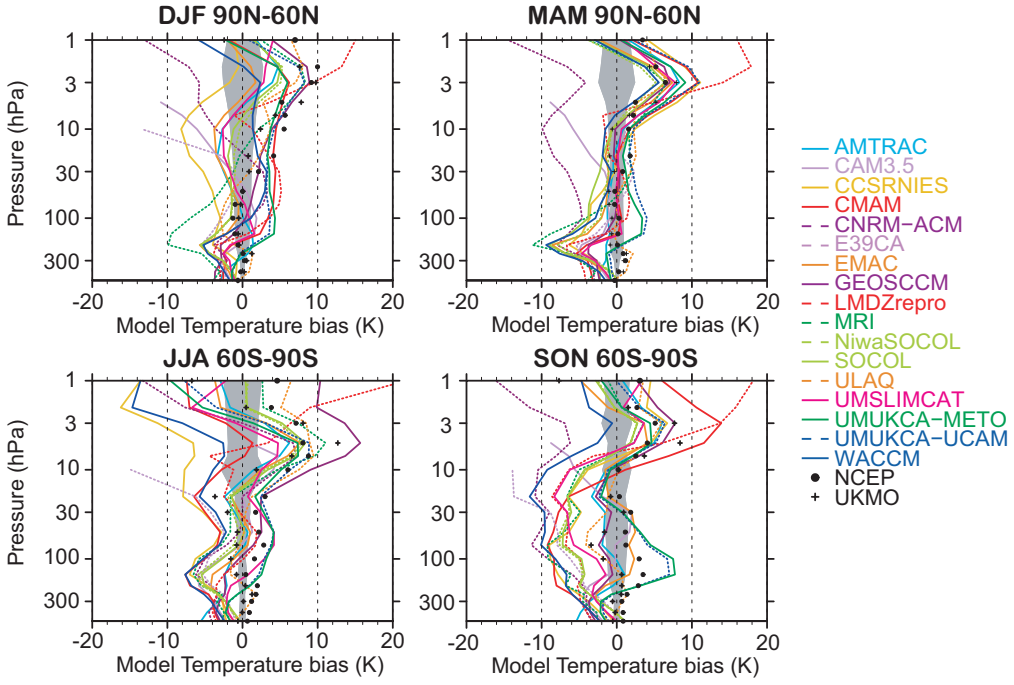

Figure 4.1: Climatological mean temperature biases for 60°N–90°N (upper panels) and 60°S–90°S (lower panels) for the winter (left) and spring (right) seasons. The climatological means for the model REF-B1 simulations and NCEP data from 1980 to 1999, and for UKMO analyses from 1992 to 2001 are included. Biases are calculated relative to ERA-40 reanalyses for 1980–1999. The grey area shows a 95% confidence interval for the 20-year mean from the ERA-40 reanalyses based on a t-distribution.

4.3 Mean climatology

4.3.1 Zonal-mean temperatures and eastward wind

In this section the zonal-mean temperature and eastward wind climatologies from the REF-B1 simulations are compared to ERA-40 and NCEP reanalyses (Uppala *et al*., 2005; Kalnay *et al*., 1996), the UKMO stratospheric analyses (Swinbank and O'Neill, 1994) and the Randel *et al*. (2004) stratospheric climatology.

Firstly, two key diagnostics of the previous CCMVal-1 inter-comparison (Eyring *et al*., 2006) are reproduced for the new CCMVal-2 simulations. **Figure 4.1** shows climatological temperature biases over the polar cap in winter and spring in the Northern (NH) and Southern (SH) Hemispheres. Eyring *et al*. (2006) highlighted the contrast between the upper and lower stratosphere in CCMVal-1. In the upper stratosphere in both hemispheres most models lie within the large range of temperatures shown in the different analyses, though there is also a very large spread between models. In the lower stratosphere, where the range of the analyses is much smaller, strong contrasts exist between the two hemispheres, with a clear cold bias for most of the models in the SH spring, and a more vertically confined cold bias between 300 and 100 hPa in the NH spring.

The results from the CCMVal-2 inter-comparison (Figure 4.1) are broadly similar to the previous Eyring *et al*. (2006) inter-comparison. The largest biases between models and observations occur in the SH spring in the lower stratosphere. For most models these biases are of the order of 5 K, though some models have biases of slightly more than 10 K. In CCMVal-1 typical biases were also around 5 K, but with two outliers with biases greater than 15 K in some places (Eyring *et al*., 2006). Both these outliers (E39CA and LMDZrepro) have reduced biases in the CCMVal-2 assessment. One new feature in the CCMVal-2 data set, in comparison to CCMVal-1, is the presence of three simulations (EMAC, UMUKCA-METO and UMUKCA-METO) with positive temperature biases between 30 and 300 hPa in the SH spring.

Figure 4.2 shows the descent of the climatological zero-wind line at 60°S (*e.g*., Scaife *et al*., 2002, Figure 7) in the models. Although this result is based on the monthly mean climatology with the transition date obtained by linear interpolation between months (assuming the monthly mean is valid for the 15th) Hardiman *et al*. (2010b) have shown that remarkably similar conclusions are obtained when the climatologies are constructed from daily data. Results from the CCMVal-1 and CCMVal-2 inter-comparisons are very similar, both showing a delayed or missing (below 10hPa) transition to westward winds in the zonal wind climatology in the SH spring. Eleven of the sixteen CCMVal-2 models analysed exhibited this delay, which is consistent with the spring-time temperature biases noted above. The date of the final warming is examined in more detail in Section 4.4.4. The large temperature biases in the lower stratosphere are strongly linked to the behaviour of ozone. Strong cold biases would tend to allow the ozone hole to persist for longer into the SH spring, and indeed the extended duration of the ozone hole is consistent with the stratospheric cooling trends seen in Chapter 10, Figure 10.12, which tend to persist well into the SH summer.

Further dynamical analysis of the basic stratospheric state in the REF-B1 simulations can be conducted by considering the structure of the zonal-mean zonal wind climatology. Instead of simply analysing the zonal wind biases

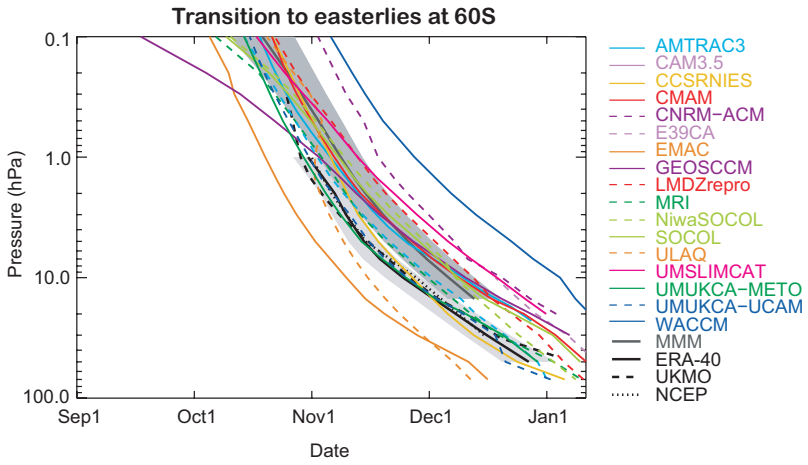

Figure 4.2: For the REF-B1 simulations. Descent of the zero zonal-mean zonal wind at 60°S based on the climatological mean annual cycle calculated from the monthly and zonal-mean zonal winds. The dark grey area shows a 95% confidence interval for the inter-model standard error, and the light grey area shows a 95% confidence interval for the 20-year mean ERA-40 transition, based on a t-distribution. Climatological means are calculated for the same period as in Figure 4.1.

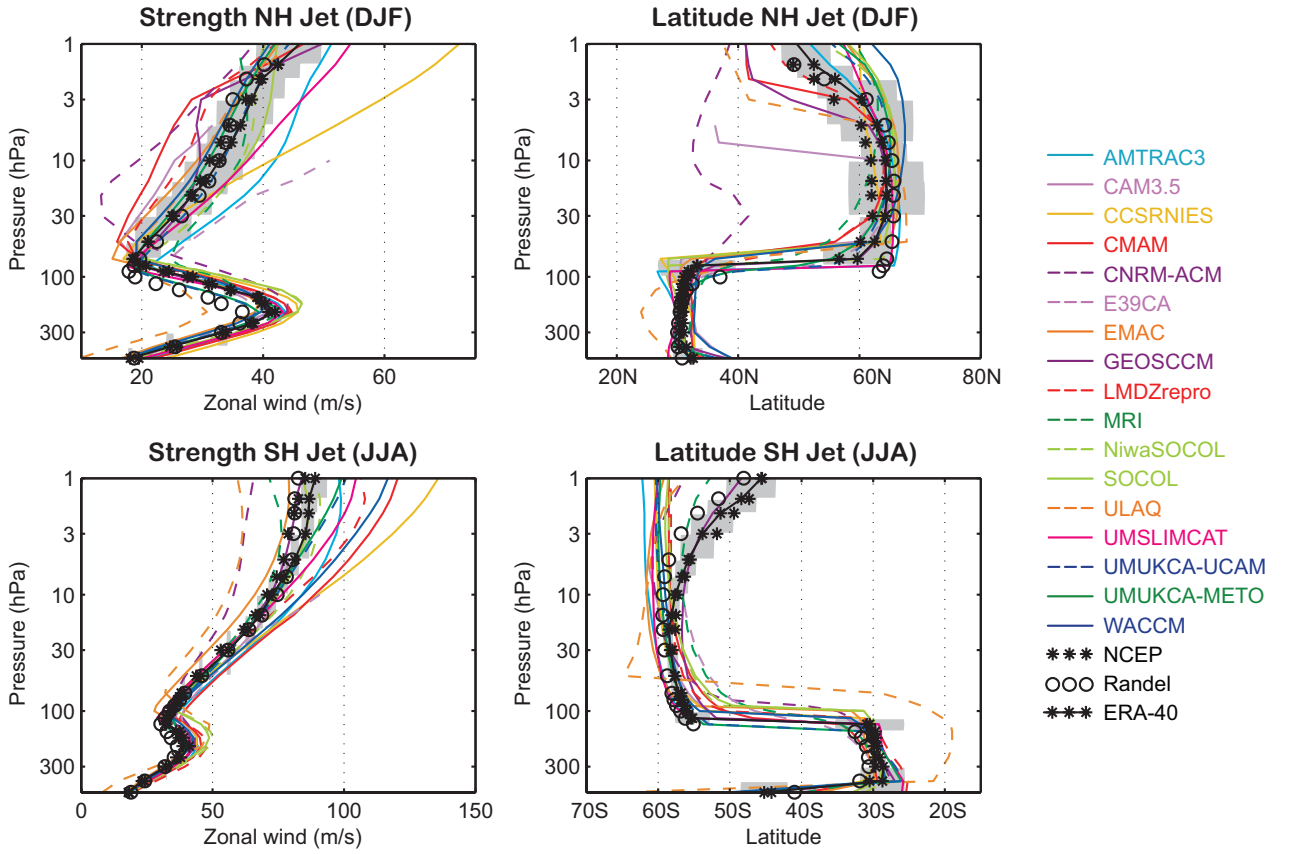

Figure 4.3: Zonal wind speed and latitude of the jet maximum of the NH DJF climatology (top), and of the SH JJA climatology (bottom) in the REF-B1 simulations. Data are based on climatological means for the models, ERA-40 and NCEP data from 1980 to 1999 and for the Randel et al. (2004) climatology that represents the time period 1992-1997. The grey area indicates a 95% confidence interval for the 20-year mean ERA-40 climatology based on a t-distribution. Where an ensemble of simulations is available, quantities are calculated for the ensemble mean zonal-mean zonal wind field.

of the models, two aspects of the zonal wind climatology are considered: the strength of the stratospheric or polar night jet and its latitudinal position. These two diagnostics are shown for the NH and SH winters in **Figure 4.3**. Very similar results are obtained for the REF-B2 simulations for the same period (not shown).

Most of the models performed extremely well in these diagnostics. In the NH the jet is generally both well positioned and of the correct strength in almost all models with the exception of the CNRM-ACM, which has its jet positioned too close to the equator with too weak winds in the lower stratosphere, the CCSRNIES model, which has a very strong jet in the upper stratosphere, and the E39CA model, which has a too strong jet near its relatively low upper boundary (10 hPa). Although the NH jet in the ULAQ model seems quite accurate in terms of position and strength, its width is too large (not shown). In the SH winter, clear biases exist for the majority of models in the upper stratosphere. Almost all the models fail to capture the observed tilt of the jet toward the equator between 10 and 1 hPa, most producing a jet with an un-tilted profile. There is a large model spread in the strength of the mid-winter jet in the SH upper stratosphere. A large number of models produce jets which are too strong, while the CNRM-ACM and ULAQ model produce jets which are too weak. The ULAQ results also show a large misplacement of the SH jet in the upper troposphere/lower stratosphere.

Predicted trends in latitudinally averaged stratospheric temperatures are also compared between the models. Since many of the more complex diagnostics later in the chapter examine changes to key stratospheric process over the 21st century, it is necessary to establish the broad context to these changes by examining changes to the stratospheric mean state. While the annual-mean global-mean temperature trends were discussed in Chapter 3, this chapter focuses on the latitudinal and seasonally averaged temperature trends. The seasonal cycle of the SH temperature trends averaged over all CCMVal-2 models is discussed in Chapter 10. **Figure 4.4** shows the trends for the three time periods: 1980-1999, 2000-2049, and 2050-2099, for the

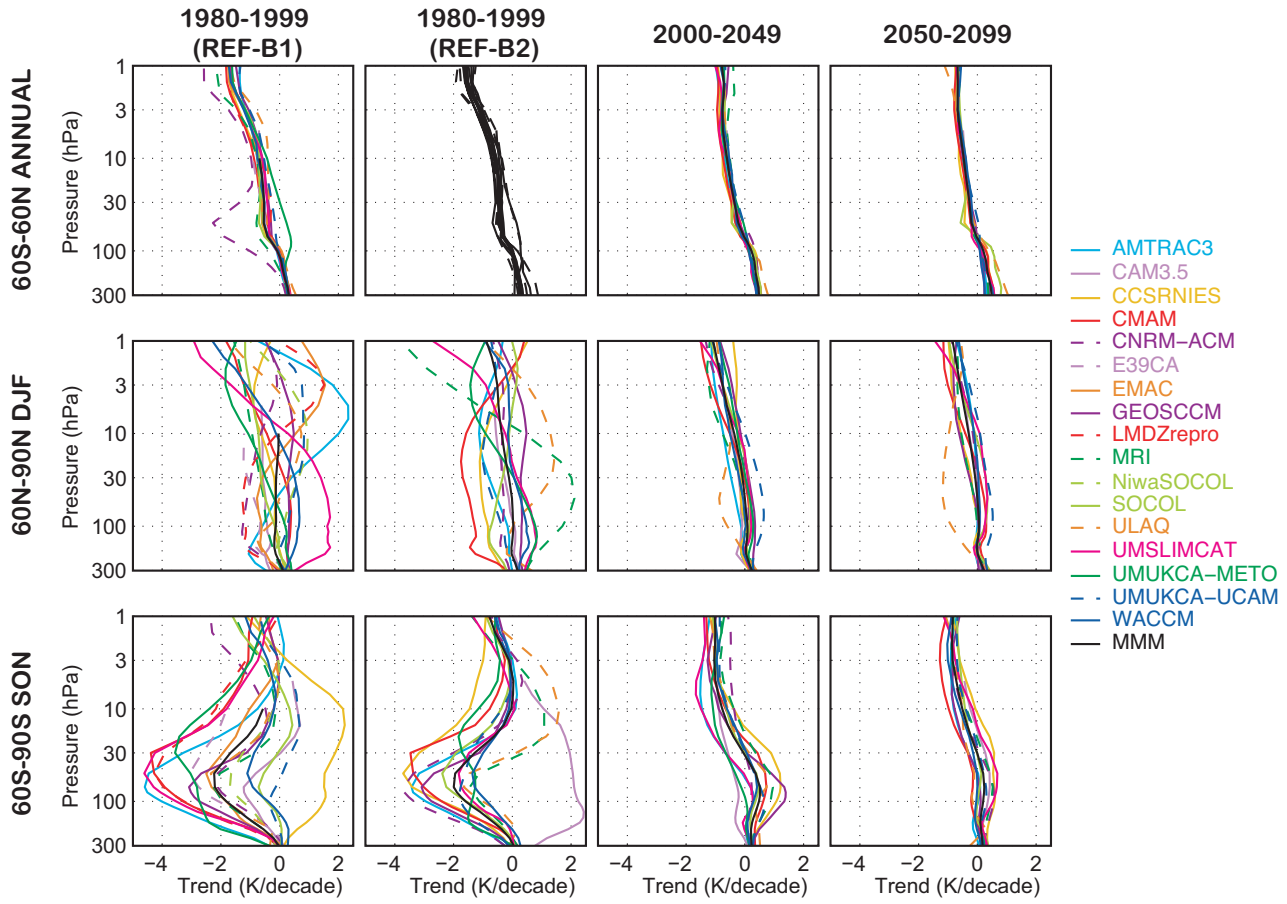

Figure 4.4: Temperature trends from 1980 to 1999 (first column for REF-B1 simulations, second column for REF-B2 simulations), from 2000 to 2049 (third column, REF-B2 simulations) and from 2050 to 2099 (fourth column, REF-B2 simulations). Top row: annual mean 60°S-60°N; middle row: December-January-February mean 60°N-90°N; lower row September-October-November mean 60°S-90°S.

regions 60°S-60°N (annual mean, first row), 60°N-90°N (December-January-February (DJF) mean, second row) and 60°S-90°S (September-October-November (SON) mean, third row). For the past, trends from both the REF-B1 (first column) and REF-B2 (second column) simulations are shown. However, due to the shorter time-period, there is a larger spread between the models for the past than for the future trends. The inter-model spread in past trends over the longer period (1960-2000, not shown) is comparable to the spread in the future trends.

For the global-mean trends, a stratospheric cooling (maximising in the upper stratosphere) occurred in all three time-periods, consistent with radiative changes due to increasing greenhouse gas (GHG) amounts. For the past, relatively large cooling trends were found in the lower stratosphere due to ozone depletion (see Chapter 3). The first row of Figure 4.4 shows that the 60°S-60°N annual mean temperature trends are very similar to the global temperature trends (c.f., Figures 4.4 and 3.4), except that the lower stratospheric cooling is smaller due to the smaller ozone depletion in this region which excludes the poles. For the past, the lower stratospheric cooling in the CNRM-ACM is a factor of four larger than the multi-model mean (in REF-B1, but not REF-B2), whereas the UMUKCA-METO shows a small warming in the lower stratosphere. There is a small spread in the model trends for the future.

The second row of Figure 4.4 shows the winter-time stratospheric temperature trends averaged over 60°N -90°N. In the lower and middle stratosphere, models on average show no significant long-term changes in the winter-time temperatures, in contrast to the small cooling that was found in the near-global, annual mean temperatures (top row of Figure 4.4). This indicates that the radiative cooling due to increasing GHG amounts is largely counterbalanced by dynamical heating resulting from stronger adiabatic compression (see Section 4.3.3). For the future, the ULAQ model shows lower stratospheric cooling that is inconsistent with the other models, whereas the UMUKCA-UCAM shows a larger than average lower stratospheric warming trend. Temperature trends in NH spring (not shown) are

similar to those in winter, except for the past when the models, on average, show small cooling trends, presumably related to the extra radiative cooling due to stratospheric ozone depletion, and possibly the absence of dynamical heating.

The last row of Figure 4.4 shows SON mean temperature trends over 60°S-90°S. Temperature trends in this region and season are clearly affected by ozone depletion in the past (1980-1999), resulting in lower stratospheric cooling, while the slower ozone recovery between 2000-2050, results in a small warming of the lower stratosphere in many of the models. Interestingly, the past cooling trends tend to be larger in the REF-B1 than in the REF-B2 simulations, reflecting the effects of different surface boundary conditions. For the past, the CCSRNIES REF-B1 and the CAM3.5 REF-B2 simulations do not show the lower stratospheric cooling. Trends in the second part of the 21st century are generally in line with the global-mean trends and are dominated by changes in GHG amounts with ozone changes having only a small impact in this period.

To summarize, the model climatological temperature biases are generally small (<5 K) apart from in the SH lower stratosphere in spring. In addition, the structure of the polar night jets is well simulated by the models with the exception of the equator-ward tilt in the SH upper stratosphere. A recurring problem from previous generations of the models is the delay in the spring-time break-down of the southern polar vortex and concomitant cold bias in the Antarctica lower stratosphere. Models predict consistent trends in the SH polar temperatures with opposite trends during the periods of strong ozone depletion (1980-1999) and ozone recovery (2000-2049). In the NH lower and middle polar stratosphere, models show no significant long-term change to the mean winter-time temperature.

4.3.2 Stationary waves / zonal asymmetries

The stationary wave field, *i.e.*, the time mean zonally asymmetric part of the circulation, is a key dynamical quantity that contributes significantly to the flux of wave activity ("EP-Flux" - see Andrews *et al.*, 1987, their Chapter 3) from the troposphere to the stratosphere and to the driving of the Brewer-Dobson circulation. The stationary wave field can be used to characterize the vertical and meridional structure of zonal asymmetries, the shape and position of the polar vortex, and long-term trends in the zonally asymmetric flow. The climatological stationary wave field, *i.e.*, the zonally asymmetric part of the climatological

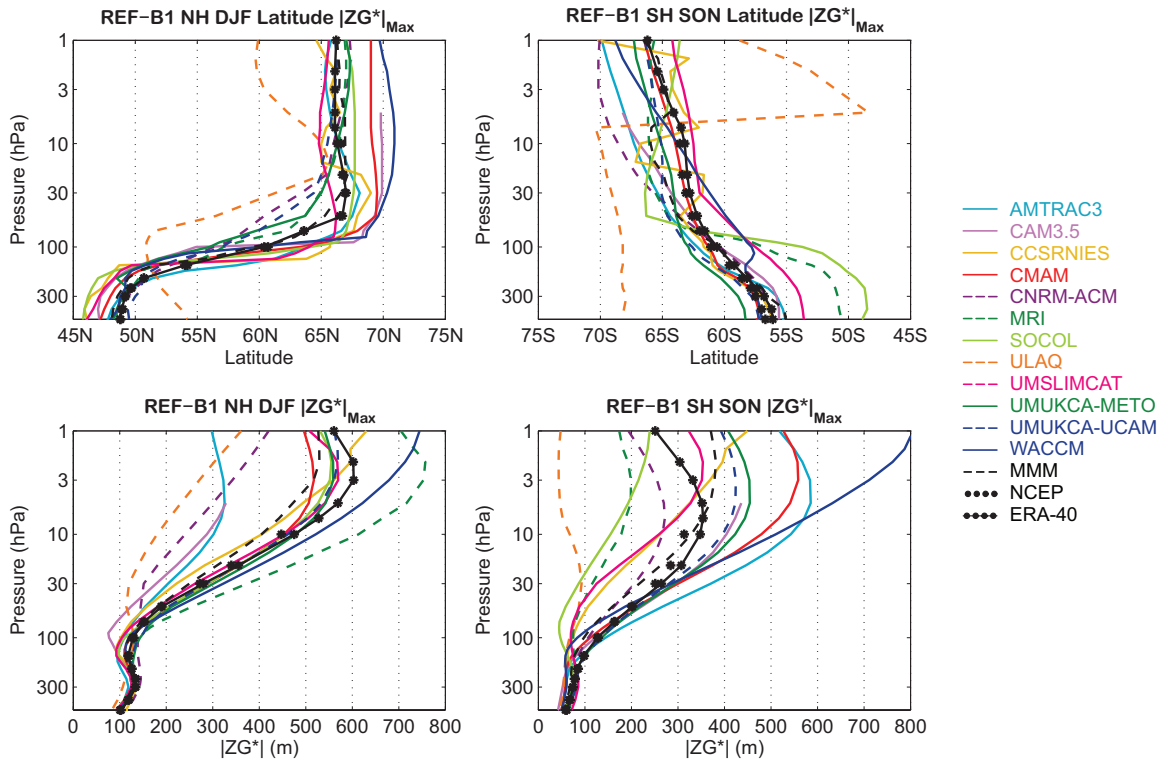

Figure 4.5: Latitudinal location and value of the maximum amplitude of the stationary wave field for the NH DJF climatology (left), and for the SH SON climatology (right). Data are based on climatological means for the models, ERA-40 and NCEP data from 1980 to 1999. Cubic spline interpolation is used to determine the latitude of the maximum and its value from the gridded data. The black dashed curve is the mean of all the model curves.

mean circulation, is observed to have a well-defined maximum in latitude at each altitude in the extra-tropical troposphere and stratosphere. **Figure 4.5** shows the simulated (REF-B1) and observed maximum amplitude in latitude of the climatological stationary wave in geopotential, and the latitudinal location of this maximum for the NH in DJF and the SH in SON. The latitudinal location of the maximum is generally well simulated by all the models, with the exception of the ULAQ model, in both hemispheres. The main biases that are robust between the REF-B1 and REF-B2 simulations (not shown) include an equator-ward bias in the upper troposphere for many of the models, a group of models with a poleward bias throughout the stratosphere, and a group of models with an equator-ward bias in the lower stratosphere. These groupings do not involve the same models for the two hemispheres, indicating that the biases are occurring for distinct reasons in each case. The models have more difficulty in simulating the observed stationary wave amplitude, with a tendency for the waves to be too weak in the NH winter and for the amplitudes to be very variable among the models in the SH spring. A systematic bias in the NH winter extends throughout the year, as seen in **Figure 4.6**, which shows the seasonal cycle of the climatological stationary wave amplitude at 10 hPa; the simulated NH stratospheric stationary waves are typically weak and have a relatively weak seasonal cycle. Figure 4.6 also shows that the amplitude and seasonal cycle of this quantity in the SH is too large and reaches its maximum amplitude too early for many of the models; differences in the seasonal timing account for the large spread in the simulations in Figure 4.5. For many of the models, the maximum NH stationary wave amplitude is weaker than the maximum SH stationary wave amplitude, in qualitative contrast to the observations. The bias toward small stationary wave amplitudes in the NH winter are consistent with the negative bias in the heat flux at 100 hPa in January (Section 4.3.4, Figure 4.12), probably because the climatological stationary wave contributes significantly to the NH heat flux (as in observations); this suggests that the NH heat flux errors are at least partially linked to problems with the stationary wave amplitude. In the SH in July, the climatological stationary wave contributes less to the SH heat flux, and this probably explains why the stationary-wave to heat-flux connection is not as straightforward: the stationary wave amplitudes are biased large (Figure 4.6) but the heat fluxes vary widely among the models (Figure 4.12).

The structure of the polar vortex is reflected in the stratospheric stationary wave field when decomposed into its dominant wave-1 component, which describes the location of the centre of the vortex relative to the pole, and its weaker wave-2 component, which describes the orientation and distortion of the vortex. **Figures 4.7a**, **b** show in polar coordinates the amplitude and phase of these components for the 50°-70° latitude climatological stationary wave at 10 hPa, for DJF in the NH and SON in the SH (the wave-2 amplitude is multiplied by a factor of four for graphical display). The amplitude biases in the figure are consistent with Figures 4.5 and 4.6. In the observations, the NH wave-1 component leads to a polar vortex centred off the pole between 0 and 30°E. Most of the models simulate this, although the UMSLIMCAT rotates the structure, and hence the polar vortex, significantly to the west, and the ULAQ simulation is almost 180° out of phase with the observations. The SH wave-1 component is more poorly simulated, corresponding to the fact that the orientation of the Antarctic polar vortex varies significantly among the models. For the NH and SH, the wave-2 component is more variable among the models, and exhibits significant differences between the REF-B1 and REF-B2 simulations (not shown). The ratio of the wave-2 to wave-1 amplitudes, which is one measure of the distortion of the vortex from a simple shifting off the pole, is shown in Figures 4.7c, d, and in the observations is about 25% in the NH and 10%

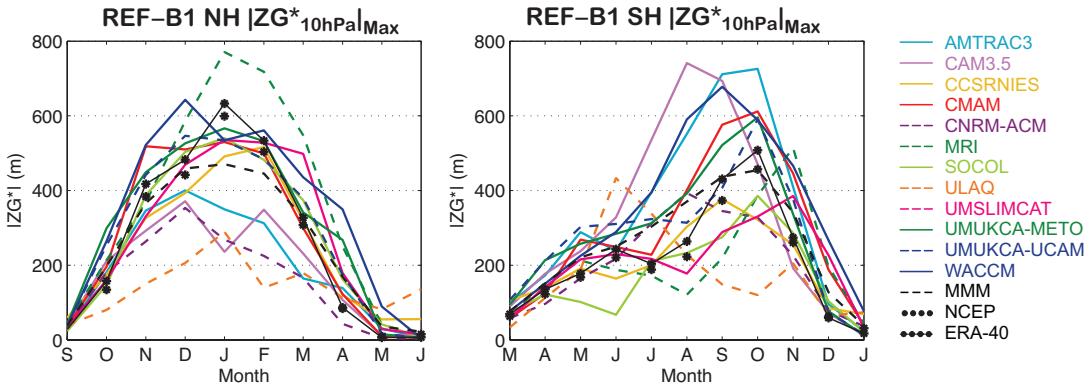

Figure 4.6: Seasonal variation of the maximum amplitude of the NH (left) and SH (right) 10 hPa climatological stationary wave. Data are based on climatological means for the model REF-B1 simulations, ERA-40 and NCEP data from 1980 to 1999. Cubic spline interpolation is used to determine the maximum value, as in Figure 4.5. The black dashed curve is the mean of all the model curves.

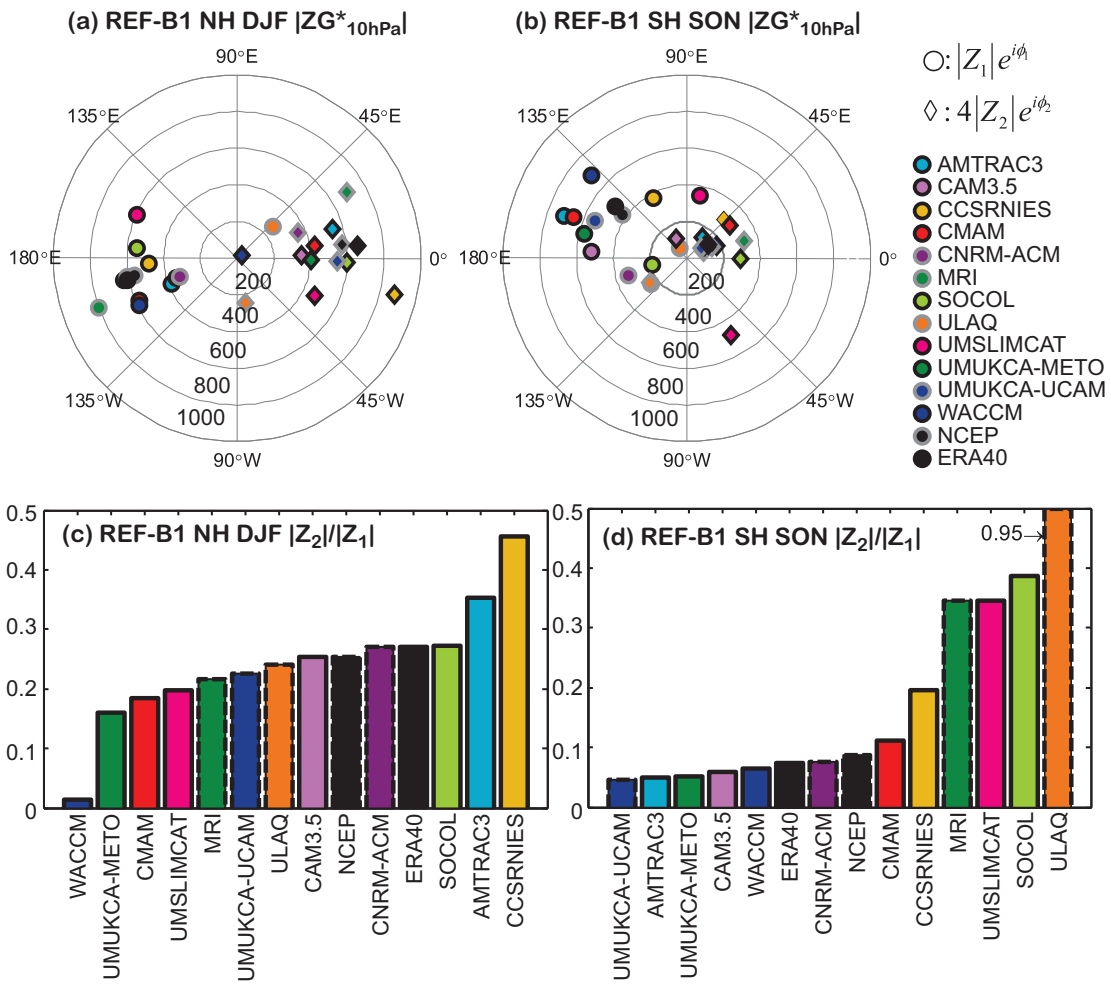

Figure 4.7: a) Phase in degrees and amplitude (contour interval 200 m), in polar coordinates, of wave-1 (circles) and wave-2 (diamonds) 10 hPa stationary waves for NH DJF. The wave-2 amplitude has been multiplied by a factor of 4. b) As in a), for SH SON. c) Ratio of wave-2 to wave-1 amplitude on 10 hPa for NH DJF. d) As in c), for SH SON. Data are based on climatological means for the model REF-B1 simulations, ERA-40 and NCEP data from 1980 to 1999.

in the SH. This ratio is generally well simulated in the NH, with a moderate bias towards small values, but is generally overestimated in the SH (the small value of the ratio in the WACCM simulation is found to increase from REF-B1 to REF-B2), suggesting that the SH vortex in the models is unrealistically distorted from circularity.

The seasonal stationary wave field is the zonally asymmetric part of the circulation for a given season and year. Trends derived from the interannual variations in this field in the period 1980-1999 of the REF-B2 simulations showed that there was no significant trend in the latitudinal location of the maximum amplitude in the NH in DJF and the SH in SON.

In the period after 2000, very few of the models project significant trends in this statistic with no significant trend in the multi-model mean. An absence of observed and simulated trends also holds for the amplitude of the NH DJF wave. However, **Figure 4.8**, shows that there is an observed significant trend in the maximum amplitude of the seasonal SH SON wave in the period 1980-1999, and that almost all models simulate trends of the same sign, although the simulated trends are generally statistically insignificant, as is the trend for the multi-model mean. Nevertheless there is a consistency between the simulations and observations suggesting that the changes to the stationary wave field in the SH are caused by ozone depletion *via* two independent, but not mutually exclusive, mechanisms. The change in the stationary wave could be a direct response to zonally asymmetric trends in the SH ozone depletion (Crook *et al*., 2008); for this mechanism, a strengthening of a displaced Antarctic vortex associated with photochemical ozone loss within the vortex would enhance the stationary wave field. This effect would reverse under ozone recovery and in all the models the positive

trend in Figure 4.8 weakens or switches sign from 1980-1999 to 2050-2099; though again, few of these trends are statistically significant. Alternatively, the impact of ozone depletion on the zonal-mean stratospheric circulation, which could also reverse as ozone recovers, could indirectly affect planetary wave properties *via* linear planetary wave dynamics.

To summarize, the models simulate the meridional and zonal location of the stationary wave field, but exhibit a bias towards weak amplitudes in the NH and a seasonal cycle that reaches its maximum too early and at too large a value in the SH. The stationary wave analysis shows that the orientation and shape of the stratospheric polar vortex is generally well captured by the models. Finally, few significant trends in the seasonal stationary wave field are found, apart from a trend towards stronger zonal asymmetry in the SH, which would be associated with either, or both of, the zonally symmetric and zonally asymmetric features of ozone radiative forcing.

4.3.3 Brewer-Dobson circulation / tropical upwelling

The Brewer-Dobson circulation plays an important role in transporting chemical species into and within the stratosphere, and also in determining the thermal structure of the stratosphere through adiabatic warming or cooling. A useful proxy for the Brewer-Dobson circulation in the models is the Transformed Eulerian Mean (TEM) residual velocity ($\overline{v}^*$, $\overline{w}^*$) (Hardiman *et al*., 2010a, Equations 22 and 23). In particular the residual vertical velocity, $\overline{w}^*$, just above the tropical tropopause can be used to deduce the mass flux entering the stratosphere and thereby provide a measure of the overall strength of the overturning meridional mass circulation in the model stratospheres (Butchart and Scaife, 2001). The rate of tropical upwelling also gives a good indication of the mean age of stratospheric air — the time elapsed since a stratospheric parcel of air was last in contact with the troposphere (Austin and Li, 2006; Butchart *et al*., 2010; also see Chapter 5).

In the REF-B1 simulations at 70 hPa there is good agreement between nine out of the fourteen models in the climatological residual vertical velocities, $\overline{w}^*$, between

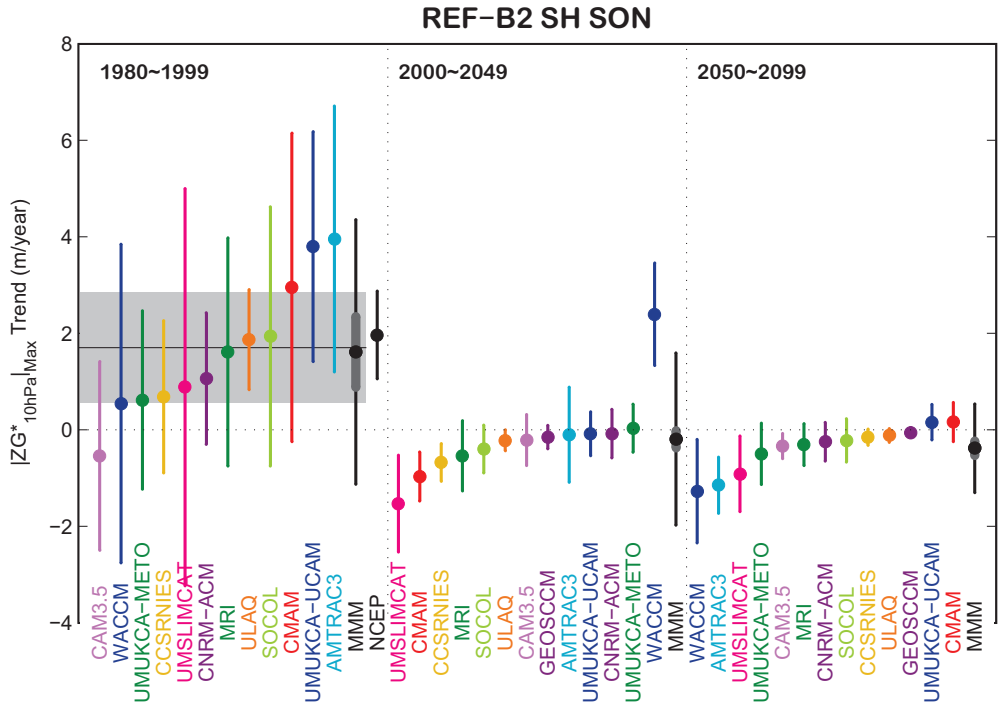

Figure 4.8: Trends in the amplitude of the seasonal-mean stationary wave for the periods 1980-1999, 2000-2049, and 2050-2099 in the REF-B2 simulations. Data are based on seasonal means for the models, ERA-40 (shaded) and NCEP data from 1980 to 1999. The linear trend is calculated using least square estimates, and the t-distribution is used to test the two-sided hypothesis that the true trend is within the estimated trend plus or minus an uncertainty for a given significance level (p = 0.05 here). The multi-model mean trend is simply the average of the trends of individual models, where its thin error bar indicates the inter-model spread (twice of the standard deviation of the trends of all models) and its thick error bar in deep grey represents the uncertainty due to the confidence intervals of individual models.

40°S and 40°N (**Figure 4.9a**). As found by Butchart *et al*. (2006, Figure 2) for a different multi-model ensemble, the latitudinal distributions of the model residual vertical velocities are remarkably similar to that derived from the UKMO analyses (Figure 4.9a). All the models apart from the SOCOL model have the characteristic local minimum in $\overline{w}^*$ at the equator with local maximum 15°-20° either side of this. Although $\overline{w}^*$ is notoriously difficult to derive from reanalysis data, these basic features were also present in the residual vertical velocities derived from ERA-40 for 1994-2002 (Randel and Wu, personal communication, 2009) hence it is possible to have some confidence that the models are behaving at least qualitatively correctly. The NiwaSOCOL model has too strong upwelling in the tropics, and the CNRM-ACM and E39CA model have downwelling there. An apparent defficiency in all the simulation occurs in the SH subtropics where the annual mean upward residual velocities ($w^* > 0.0$) extend 10°-15° further poleward than in the UKMO analyses, though there is rather good agreement (little spread) between the models at these latitudes. There is also too little upwelling in the models between 10°N and 20°N. In general model residual vertical velocities are more symmetric across the equator than those derived from the UKMO analyses (or ERA-40 — not shown).

Similar features are seen in the REF-B2 simulations (not shown), though the SOCOL results are now in better agreement with the UKMO analyses. For the multi-model mean, there is more upwelling equator-ward of ~13° and less poleward of ~13° (up to the turn-around latitudes) in the REF-B2 simulations compared to the REF-B1 simulations, but on average the total tropical upwelling is the same in both sets of simulations to within 1%.

When the seasonal movement of the tropical upwelling region toward the summer hemisphere is taken into account all the models with the exception of the NiwaSOCOL, SOCOL and ULAQ models correctly reproduce, with respect to the UKMO analyses, the locations

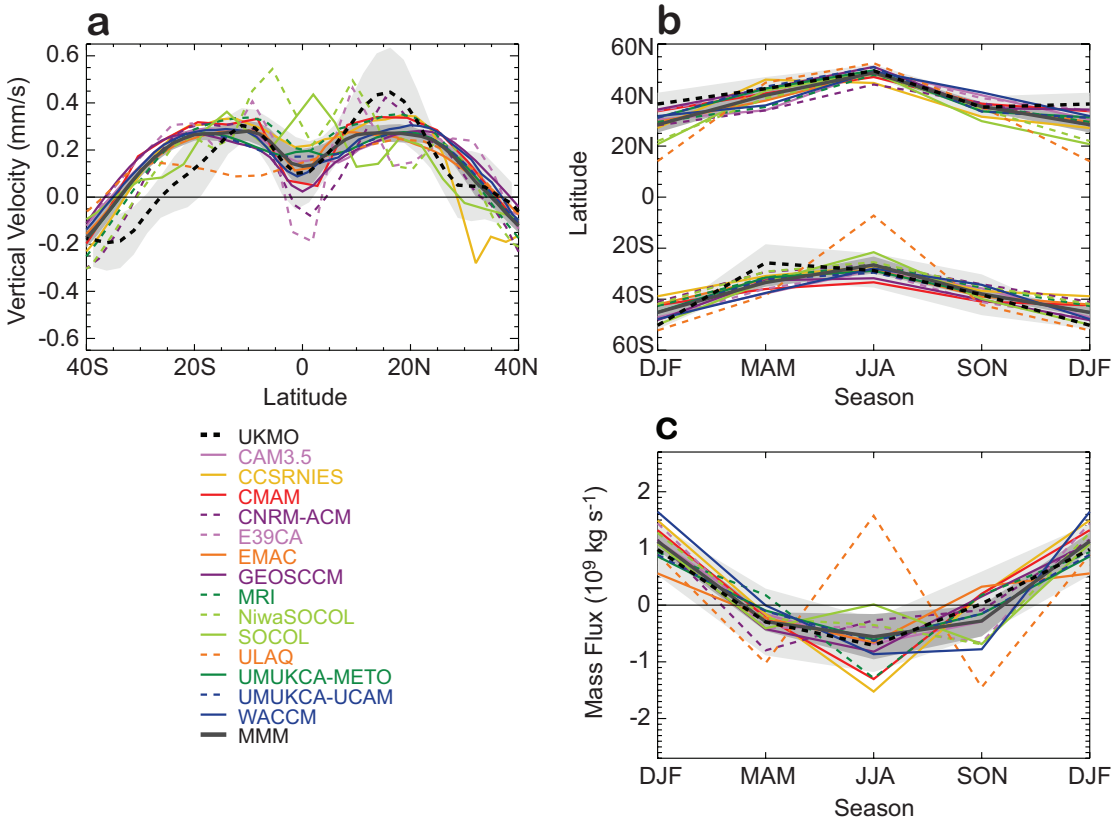

Figure 4.9: (a) Annual mean residual vertical velocities $\overline{w}^*$ at 70 hPa. (b) "Turn-around" latitudes, defined as the latitudes within 50°S-50°N where $\overline{w}^* = 0$. (c) Upward mass flux at 70 hPa calculated from the residual vertical velocity $\overline{w}^*$. Seasonal anomalies from the annual mean are shown. All panels show averages from 1980-1999 for the REF-B1 simulations. The thick dark grey solid line shows the multi-model mean. Where data from an ensemble of runs is available, the ensemble mean is used. The thick black dashed line shows the UKMO stratospheric analyses (averaged from 1992-2001; the averaging period is found not to matter). Light grey shading shows the 95% confidence interval for the 10 year mean from the UKMO analyses and dark grey shading shows the inter-model standard error.

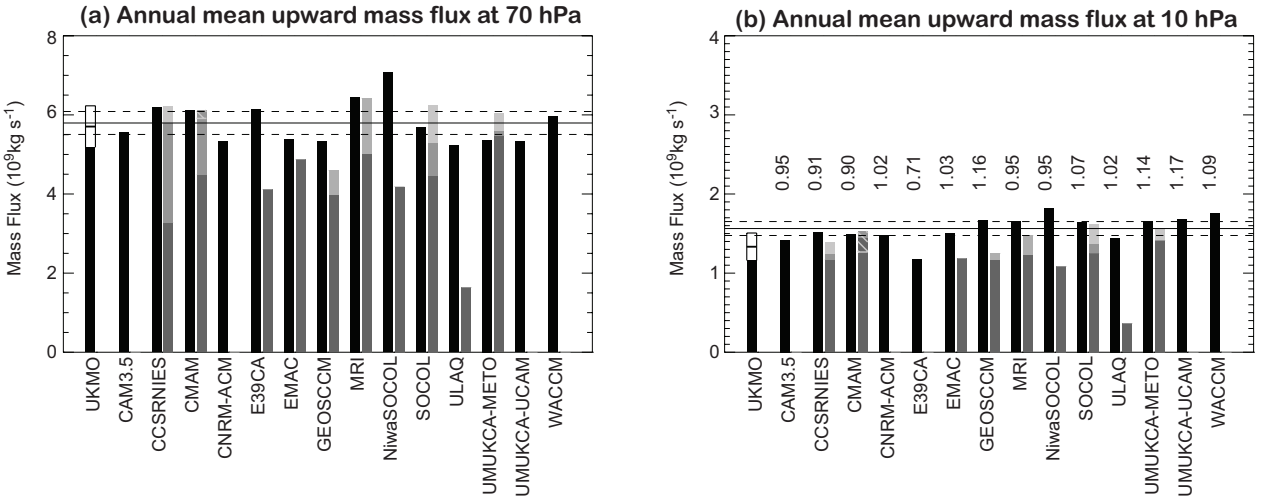

Figure 4.10: Annual mean upward mass flux averaged from 1980 to 1999 for the REF-B1 simulations and from 1992 to 2001 for the UKMO analyses. Averaging the modelled upwelling from 1992-2001 gives very similar values to those shown for 1980-1999. Upwelling calculated from $\overline{w}^*$ is shown by black bars. Upwelling calculated by downward control is split into contributions from: resolved waves (dark grey), orographic gravity wave drag (OGWD) (grey), and non-orographic gravity wave drag (NOGWD) (light grey). OGWD and NOGWD are shown combined for the GEOSCCM and MRI model. For some models only the resolved wave contributions are shown due to the unavailability of gravity wave drag diagnostics. In the CMAM, NOGWD produces a negative upwelling and so cancels some of the upwelling produced by the OGWD and the resolved waves. This cancellation is shown by diagonal lines. The black horizontal lines show the multi-model mean and the inter-model standard error. The 95% confidence interval for the UKMO analyses is shown by the unshaded part of the bar with the horizontal line at the mid-point being the multi-year (10-year) mean. Values shown at (a) 70 hPa, and (b) 10 hPa. The numbers above the bars in (b) are the ratio for that model of the upwelling mass flux (normalised by the multi-model mean) at 10 hPa to upwelling mass flux (normalised by the multi-model mean) at 70 hPa.

of the "turn-around latitudes" where $\overline{w}^*$ is zero (*i.e.*, the latitudes where the tropical upwelling changes to extra-tropical downwelling — Figure 4.9b). The annual cycle in the integrated upward mass flux between these turn-around latitudes was also generally well reproduced, though again the SOCOL and ULAQ models did not perform as well as the other models (Figure 4.9c). In the REF-B2 simulations, the turn-around latitudes are, on average, the same as in the REF-B1 simulations to within 0.5°. The multi-model mean REF-B2 upwelling is 0.1-0.2 mm/s greater in DJF and SON and 0.1-0.2mm/s less in March-April-May (MAM) and June-July-August (JJA) than that for REF-B1, though the annual mean upwelling is the same in both sets of simulations to within 1%.

On average the annual-mean tropical upwelling mass fluxes in the REF-B1 simulations, calculated between the turn-around latitudes at 70 hPa and following the seasonal movement of those latitudes, agrees with the mass fluxes derived from the UKMO analysis (**Figure 4.10a**, black bars). The standard error in the multi-model mean is less than the interannual variability in the analysed mass fluxes (not shown). The contributions of resolved and parameterised wave drag in driving this upward mass flux can be estimated using the Haynes *et al*. (1991) Downward Control Principle (*e.g*., Butchart *et al*., 2010). These contributions are shown by the grey bars in Figure 4.10a. With the exception of the UMUKCA-METO there is a significant contribution from the parameterised orographic gravity wave drag (OGWD) (for those models that supplied OGWD data), which on average accounts for 21.1% of the driving of the upwelling at 70 hPa decreasing to 4.7% at 10 hPa (Figure 4.10b). At 70 hPa the resolved waves accounted for 70.7% (71.6% at 10 hPa) and non-orographic gravity wave drag (NOGWD) 7.1% (10.9% at 10 hPa) of the driving again averaged over those models which provided these diagnostics. In general, however, there was a wide spread between the models in the contributions from the wave drags. At 70 hPa, the contributions from the resolved waves ranged from 31.4% (ULAQ) to 102.1% (UMUKCA-METO), while the range for OGWD and NOGWD was 2.0 (UMUKCA-METO) to 40.9% (CCSRNIES) and -3.4 (CMAM) to 16.8% (SOCOL), respectively. It is also worth noting that the models generally overestimate the 100 hPa heat flux (~vertical component of the EP-Flux) between 20°S and 40°S (Section 4.3.4, Figure 4.12), which includes the southern latitude (*i.e*., the turn-around latitude, *c.f*.,

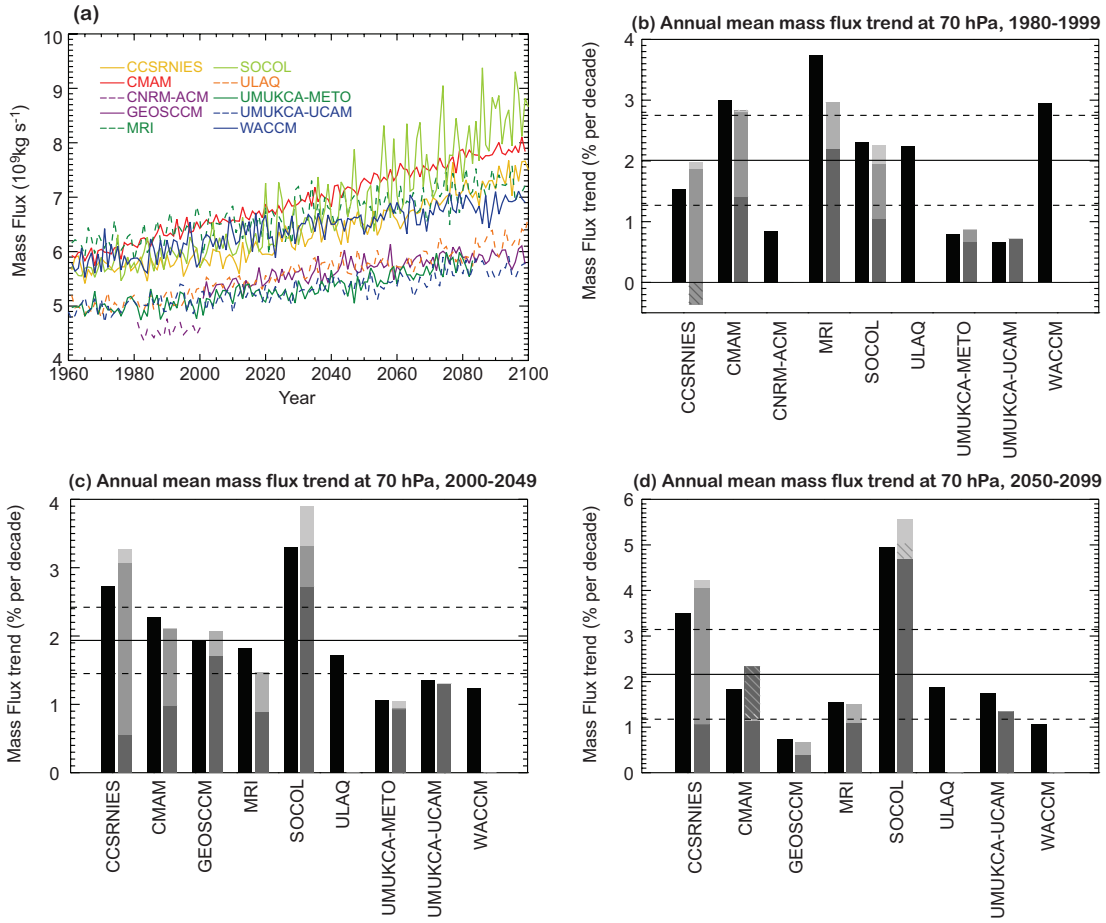

Figure 4.11: For the REF-B2 simulations. (a) Annual mean upward mass flux at 70 hPa, calculated from $\overline{w}^*$. Also shown is the annual mean mass flux trend at 70 hPa from (b) 1980-1999, (c) 2000-2049 and (d) 2050-2099. The grey shading and diagonal lines are as in Figure 4.10. The values are calculated from a linear fit to the annual mean mass flux from (b) 1980-1999 shown as a percentage of 1980 mass flux values, (c) 2000-2049 and (d) 2050-2099 shown as a percentage of 2000 mass flux values. The horizontal lines show the multi-model mean and the inter-model standard error.

Figure 4.9b) at which the downward control integral is performed, though it is unclear what impact this would have on the upwelling estimated from the EP-flux divergence.

Similar results were obtained for the REF-B2 simulations (not shown) with the multi-model mean upwelling within 3% of that in the REF-B1 simulations. The largest differences occurred for the CNRM-ACM which had over 15% less upwelling for the REF-B2 than for the REF-B1 simulation.

In REF-B1 simulation, the ratio of the upwelling (as calculated from $\overline{w}^*$) at 10 hPa to that at 70 hPa (weighted by the multi-model means) gives some indication of the relative leakiness of the tropical pipe in the lower stratosphere with respect to the multi-model mean (Neu and Plumb, 1999; see also Chapter 5, Section 5.2.1.2). This ratio is shown in Figure 4.10b (see figure caption for details). The UMUKCA and GEOSCCM simulations show too little upwelling at 70 hPa and too much upwelling at 10hPa, with the ratio of upwelling at 10 hPa to that at 70 hPa being around 115% of that of the multi-model average. Conversely, the CCSRNIES model, and the CMAM show too much upwelling at 70 hPa and too little upwelling at 10 hPa, with a ratio of 90% or less of that of the multi-model average.

For all the models the annual mean upward mass flux at 70 hPa increased from the start (1960) to the end of the REF-B2 simulations (see Figure 4.11a). On average the trend in the upward mass flux was about 2% per decade (Figures 4.11b, c, d) with the largest trends occurring in JJA (not shown). With the exception of the SOCOL model, interannual variability in the annual mean upward mass flux is less than the multi-model spread (Figure 4.11a). For the end of the 20th century (1980-1999) the trends predicted by the REF-B2 simulations (Figure 4.11) were very similar to those for the REF-B1 simulations (not shown). The largest difference was found for the CCSRNIES model, which had

a negative trend from the resolved waves in the REF-B2 simulation but not for the REF-B1 simulation. The multi-model mean trend for the period 1980-1999, was 2% per decade for REF-B2, and 2.3% per decade for REF-B1. It should also be noted that the partitioning of the downward control estimate of the trends into resolved and parameterised drag contributions (grey bars in Figures 4.11b, c, d) is rather sensitive to the calculation of the location of the turn-around latitudes (McLandress and Shepherd, personal communication, 2009). This sensitivity results from the strong latitudinal dependence of the OGWD in the NH subtropics (*i.e.*, near the turn-around latitudes) and will most likely impact the calculations for models with a coarse horizontal resolution.

To summarize, the strength of tropical upwelling and position of the turn-around latitudes are well represented in general, but in all models the annual mean upwelling in the SH extends 10°-15° further poleward than in the analysis, and there is too little upwelling between 10°N and 20°N. Tropical upwelling at 70 hPa is within observational uncertainty whilst at 10 hPa there is slightly too much upwelling. There is disagreement across models as to the relative contributions from resolved waves and parameterised gravity waves to driving this upwelling, through apart from one model there was a significant contribution from orographic gravity wave drag at 70 hPa. The strength of tropical upwelling, and thus the Brewer-Dobson circulation is projected to increase throughout the 21st century by

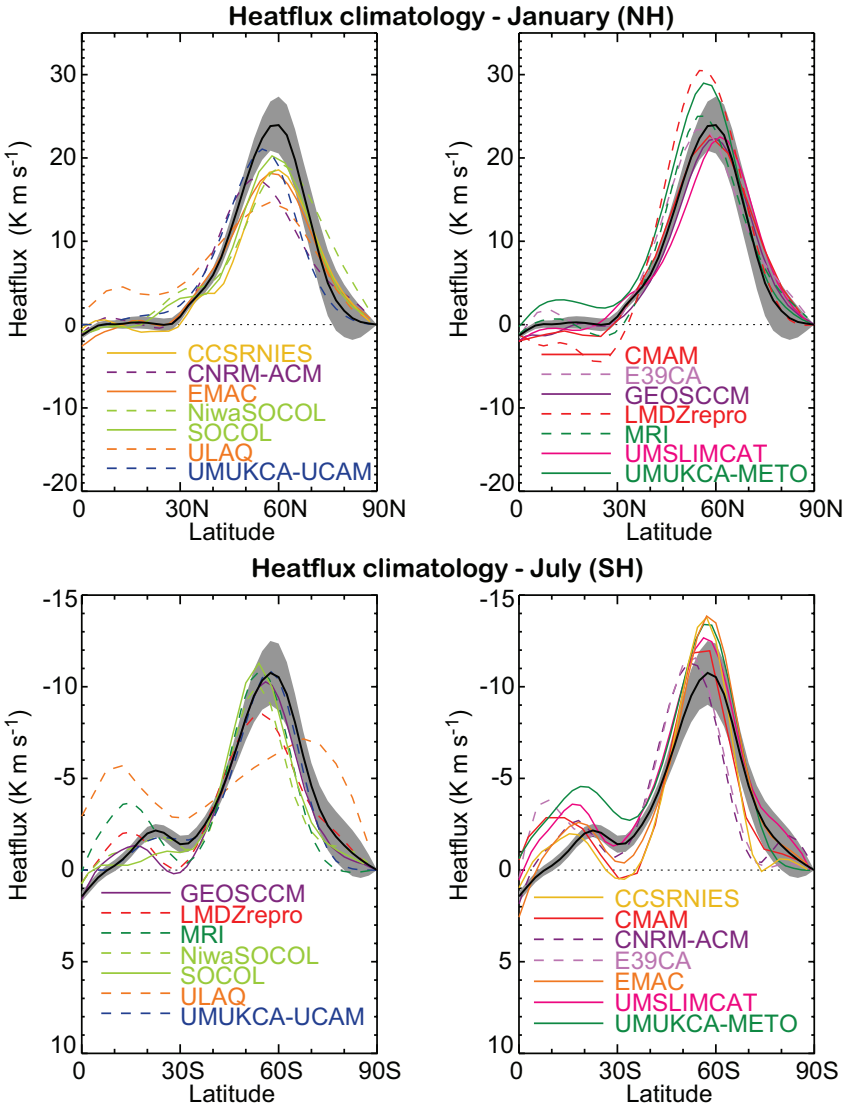

Figure 4.12: Monthly mean climatology of the eddy meridional heat flux at 100 hPa for the months of January and July, 1980-1999. Data from ERA-40 reanalysis is shown in the black line. Grey shading shows the 95% uncertainty estimate for the 20-year mean of the ERA-40 based on a t-distribution. Where ensemble simulations are available, the climatology is derived by taking the mean across all ensemble members.

around 2% per decade.

4.3.4 Heat flux / heat flux-temperature correlations

In this section the climatology of the eddy meridional heat flux at 100 hPa and the relationship between year-to-year variability of this quantity and spring-time polar cap temperatures is assessed. The meridional heat flux is the zonal-mean of the product of the eddy components of temperature and meridional wind. For studies of stratospheric dynamics, it is a useful proxy for the vertical component of the EP-flux (see Andrews *et al*., 1987) due to planetary-scale Rossby waves. Newman *et al*. (2001) noted the strong correlation between the eddy meridional heat flux at 100 hPa, averaged over a band between 40°N-80°N during January and February, and the subsequent temperature of the polar cap at 50 hPa in February and March. A simple linear fit to a scatter plot of 100 hPa heat flux and 50 hPa polar temperature provides information about the way in which the polar stratosphere responds to anomalous tropospheric wave activity propagating into the stratosphere.

Figure 4.12 shows the monthly mean climatological heat flux for mid-winter in the NH and SH in the REF-B1 simulations. In the NH, most models reproduce the latitudinal distribution of the mid-winter heat flux climatology well, with a strong peak located around 55°N. However, in many of the models, including all those in the top left panel, the maximum heat flux is substantially smaller than that seen in the ERA-40 reanalysis data, with peak values well outside the estimated 95% confidence interval for the reanalysis data. The ULAQ model has a particularly low heat flux maximum. In the top right panel, most of the models perform well in simulating the heat flux climatology, although the LMDZrepro model and UMUKCA-METO have peak values larger than in the ERA-40 reanalysis.

In contrast in the SH, the climatological heat flux tends to be close to or slightly larger than that derived from the ERA-40 reanalysis. There are also some significant differences in the structure of the mid-winter heat flux climatology between the models and the reanalysis. In particular, models seem to over-emphasize the heat flux in the region between 20°S and 40°S associated with the subtropical jet. This is a particular problem in the ULAQ model which also mis-positions and under-estimates the strength of the main region of large negative meridional heat flux centred around 60°S.

Linear trends in meridional heat flux in the three periods 1980-1999, 2000-2049 and 2050-2099 of the REF-B2 simulations are shown in **Figure 4.13**. For this calculation, the mean heat flux between 40°N/S and 80°N/S during the northern and southern mid-winter is considered. Although there is a great deal of variability between models in each period, overall there is little sign of a consistent trend in the mid-winter heat flux in any of the three periods in either hemisphere. The only period with a statistically significant trend in the multi-model mean heat flux is the period 2050-2099 in the NH. However, sensitivity tests of the multi-model mean show that if either the strong negative trends present in the ULAQ model or the UMUKCA-METO are removed, the multi-model trend is no longer significant. This indicates the multi-model trend should be treated with caution, particularly since the UMUKCA-METO REF-B2 simulation ends in the mid-2080s. Similar analyses were also performed for the REF-B1 simulations and for heat flux trends in the meteorological spring, but neither produced significant, consistent trends in the multi-model average.

Finally, the response of stratospheric temperatures to variations in heat flux is considered. This analysis follows directly from the work of Newman *et al*. (2001) and was reproduced by the CCMVal-1 inter-comparison (Eyring *et al*., 2006). Here, to provide a more succinct way of comparing different models, only the parameters of the linear fits to the scatter plots of 100 hPa heat flux *vs*. 50 hPa temperatures are presented. **Figure 4.14** shows these for the REF-B1 simulations. The slope and intercept of the regression lines diagnose different properties of the model stratosphere. The intercept of the regression line (x-axis) gives an indication of the temperature that the polar cap would have if no resolved wave-driving were present. The slope of the regression line (y-axis) gives an indication of the strength of the stratospheric temperature response to a unit amount of resolved tropospheric wave-driving.

In the NH, almost all of the models produce linear fit parameters within the sampling uncertainty of the linear fit parameters in the ERA-40 reanalysis. Only the UMSLIMCAT and UMUKCA-UCAM, which has a significantly large stratospheric temperature response to the 100 hPa heat flux, are outliers in this diagnostic. It is interesting however, that the 95% confidence limits for the two UMUKCA simulations (UMUKCA-UCAM and UMUKCA-METO) overlap, and the UMUKCA-METO has linear fit parameters that are not significantly different from the ERA-40 reanalysis. In general in the NH, the cluster of model points is shifted toward the upper left quadrant of the plot, indicating a tendency toward lower polar temperatures and an enhanced response of the lower stratosphere to tropospheric wave-driving. The tendency towards a cold bias in the lower stratosphere during spring is consistent with previous model assessments (*e.g*., Eyring *et al*., 2006) and with Figure 4.1.

In the SH, there is a similar spread in the properties of the linear fit to the NH. Several of the models show properties statistically distinct from those in the ERA-40 reanalysis, including the EMAC model, which shows both a cold bias and an enhanced sensitivity to tropospheric wave-driving, the AMTRAC3, which has an enhanced sensitivity

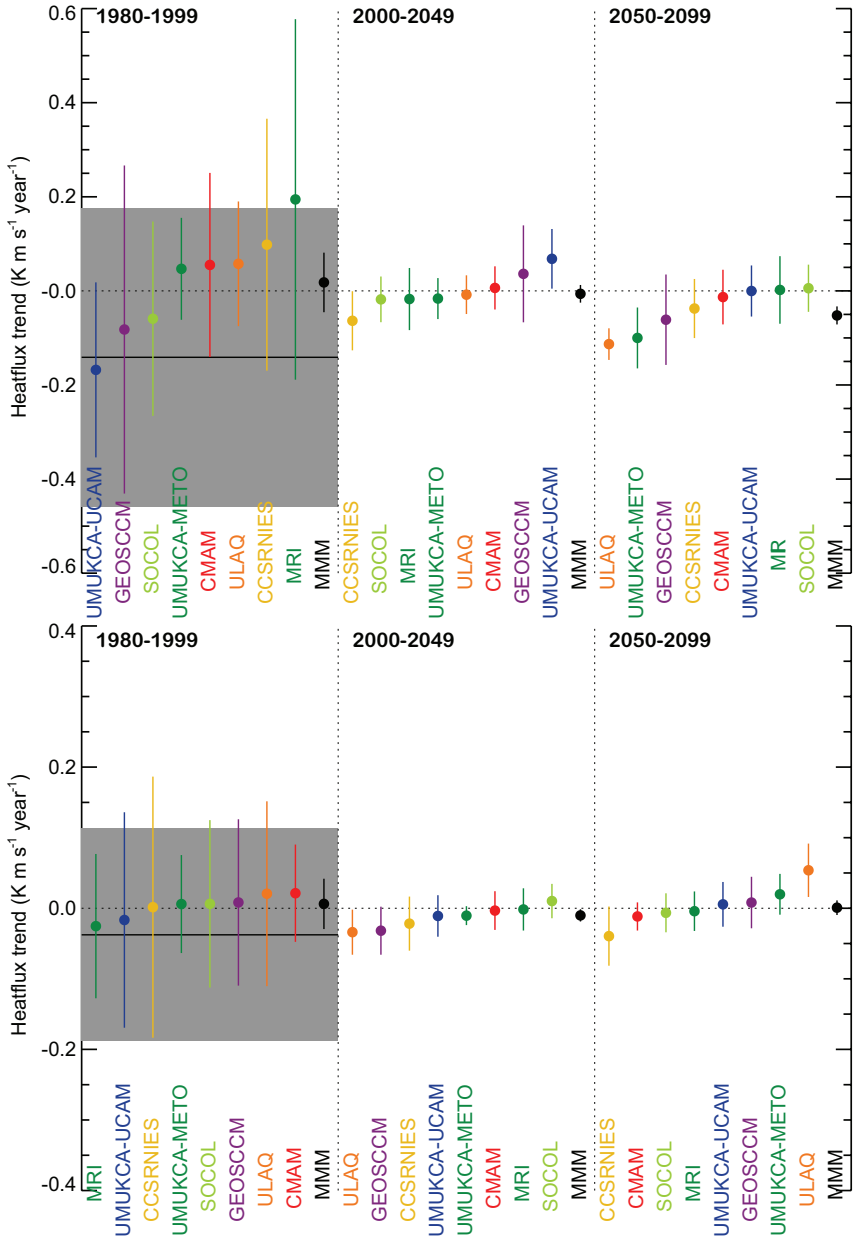

Figure 4.13: Linear trends in the mean meridional heat flux averaged between 40°N/S and 80°N/S for the winter seasons. Top panel shows trends for the January heat flux in the NH. Each section shows trends for a different time period, 1980-1999 in the leftmost panel, 2000-2049 in the middle panel and 2050-2099 in the right panel. Each model is represented by its mean trend shown by a dot and an estimate of the 95% confidence estimate on the trend, shown by plotting two standard errors either side of the mean estimate. In the 1980-1999 panel, trends and confidence limits for the ERA-40 reanalysis are shown by the solid black line and grey shading. In each panel, a multi-model mean estimate is given. The multi-model mean is calculated by weighting each model's trend by its uncertainty. The bottom panel shows the same information for the trends in the SH in July.

to tropospheric wave-driving, the CMAM, which displays a cold bias, and the GEOSCCM, which displays a warm bias. As an aside. it is important to remember that in this analysis (in both the NH and SH), results from the CMAM are from a three member ensemble average while many of the other models supplied only one realization, which can be observed from the relatively small error bars for the CMAM fit parameters in Figure 4.14.

In contrast, the CNRM-ACM has very large error bars on its fit parameters because only 10 years of heat flux data were supplied by this model. Given the difference in the amount of data considered it is therefore easier to distin-

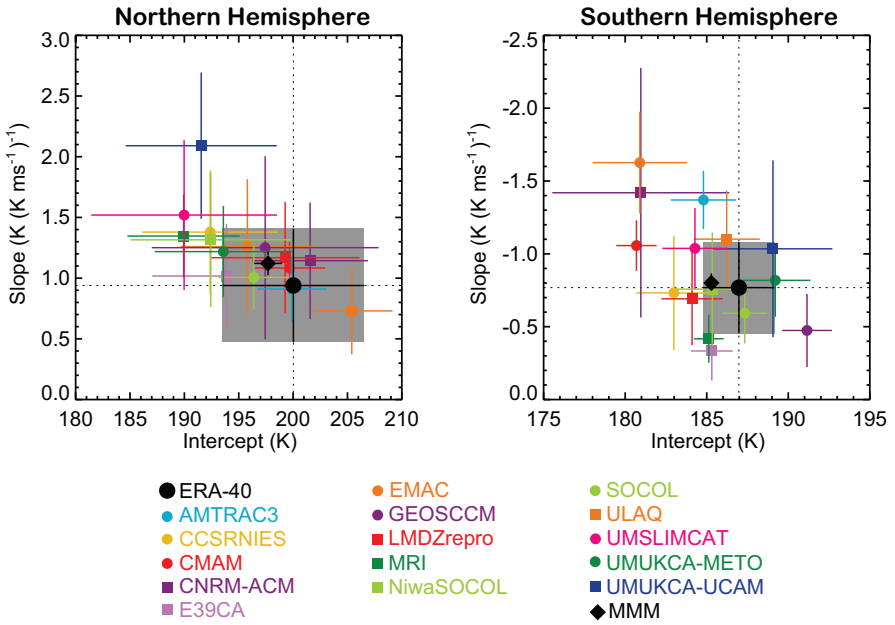

Figure 4.14: Parameters of the linear fit to the scatter plot of the 100 hPa heat flux vs. the 50 hPa temperature (for more details of procedure see Newman et al. (2001) and Eyring et al. (2006)). The panels show the intercept of the linear fit (x-axis) plotted against the slope of the regression line (y-axis) for the NH and the SH. Black symbols show the same diagnostic for the ERA-40 reanalysis data. Estimates of 95% confidence limits for the two parameters are included for each estimate in the thin coloured lines. Grey shading indicates the 95% confidence estimates for the ERA-40 reanalysis data.

guish significant differences between models and reanalysis for those models which supplied multiple realizations.

To summarize, the models reproduce the observed connection between lower stratospheric heat flux and temperature reasonably well in both the NH and SH, and the generally good performance of models in this diagnostic is consistent with the previous CCMVal-1 generation of models (Eyring *et al*., 2006). In Chapter 8 the relationship between heat flux and the seasonal ozone gain/loss is diagnosed in more detail and shows larger spread between models than is evident from the heat flux *vs*. temperature correlations diagnosed here, because of the additional model spread introduced by the chemistry and transport.

4.3.5 Polar stratospheric cloud threshold temperatures

Changes in stratospheric temperatures are expected to have a large impact on polar ozone loss through their influence on the formation and occurrences of polar stratospheric clouds (PSCs). Polar stratospheric cloud formation is related to both the mean climatological structure of the polar vortex and its variability. To broadly assess the ability of models to reproduce mean conditions suitable for the formation of nitric acid trihydrate (NAT) and ice PSCs, the accumulated area where temperatures are below the appropriate thresholds (195 K for NAT and 188 K for ice) are calculated. Although these diagnostics provide a useful estimate of the potential for PSC formation they do not take into account microphysical factors which are considered in more detail in Chapter 6. However, the simple diagnostics based on accumulated areas as described here are used in Chapter 9 when assessing the amounts of polar ozone depletion in the models (Section 9.5.4).

Following Pawson *et al*. (1999) and Austin *et al*. (2003), the potential for PSC formation in the models and ERA-40 reanalysis is estimated by calculating, each day, the percentage of the horizontal area of the hemisphere where the 50 hPa temperatures poleward of 50° are below the NAT and ice PSC formation thresholds. These daily percentage areas are then accumulated over the course of the winter and spring (92 days from July to September in the SH; and 90 days from December to February in the NH) to provide, for that year, an estimate of the total amount of NAT ($\tilde{A}_{NAT}$) and ice ($\tilde{A}_{ice}$) PSCs in units of %-days.

The climatological mean $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ for the REF-B1 simulations for 1980-1999 are shown in **Figure 4.15**. The linear trends for the same period are shown in **Figure 4.16**, though there is considerable uncertainty in the trend estimates due to the large interannual variability in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$, particularly in the Arctic. In the Antarctic, the multi-model mean $\tilde{A}_{ice}$ (grey bars) agrees well with the ERA-40 estimate, but the multi-model mean $\tilde{A}_{NAT}$ is significantly smaller than the ERA-40 estimate over the same

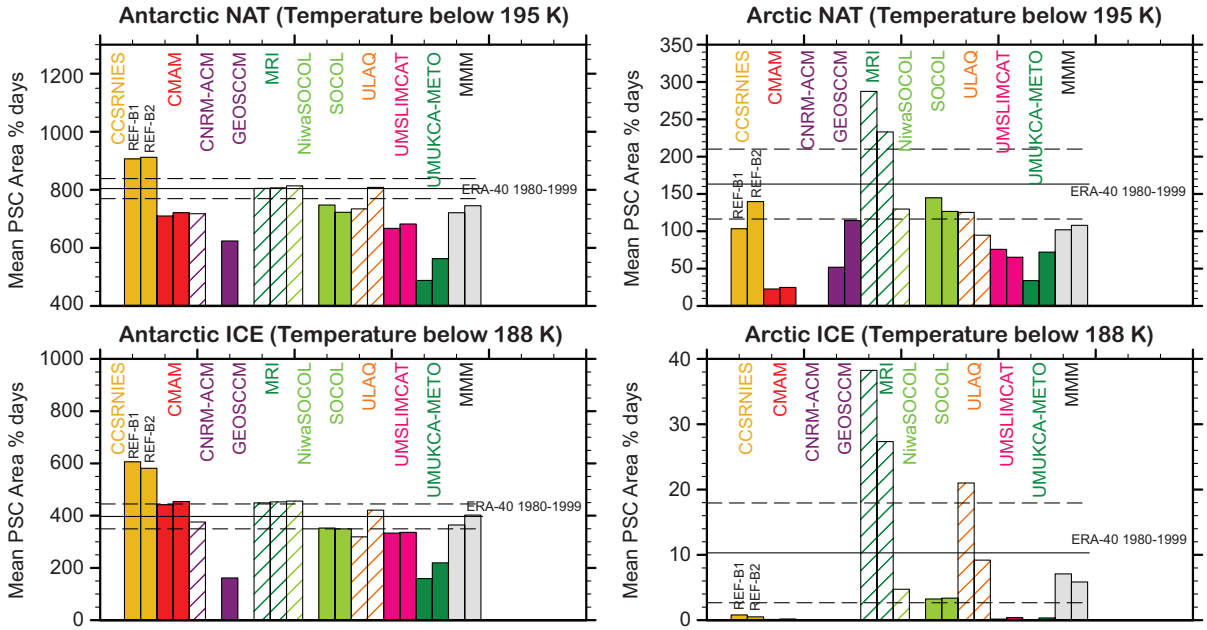

Figure 4.15: Mean (1980-1999) for the Antarctic (left) and the Arctic (right) of the seasonally accumulated area at 50 hPa where daily temperatures are below 195 K (approximate threshold temperature for NAT formation, top panels) and below 188 K (approximate threshold temperature for ice formation, bottom panels) for the REF-B1 (first column) and REF-B2 (second column) simulations. The dashed black line is for the ERA-40 reanalysis (1980-1999). The units are the percentage of the hemisphere where the daily temperature is below the threshold multiplied by the duration in days.

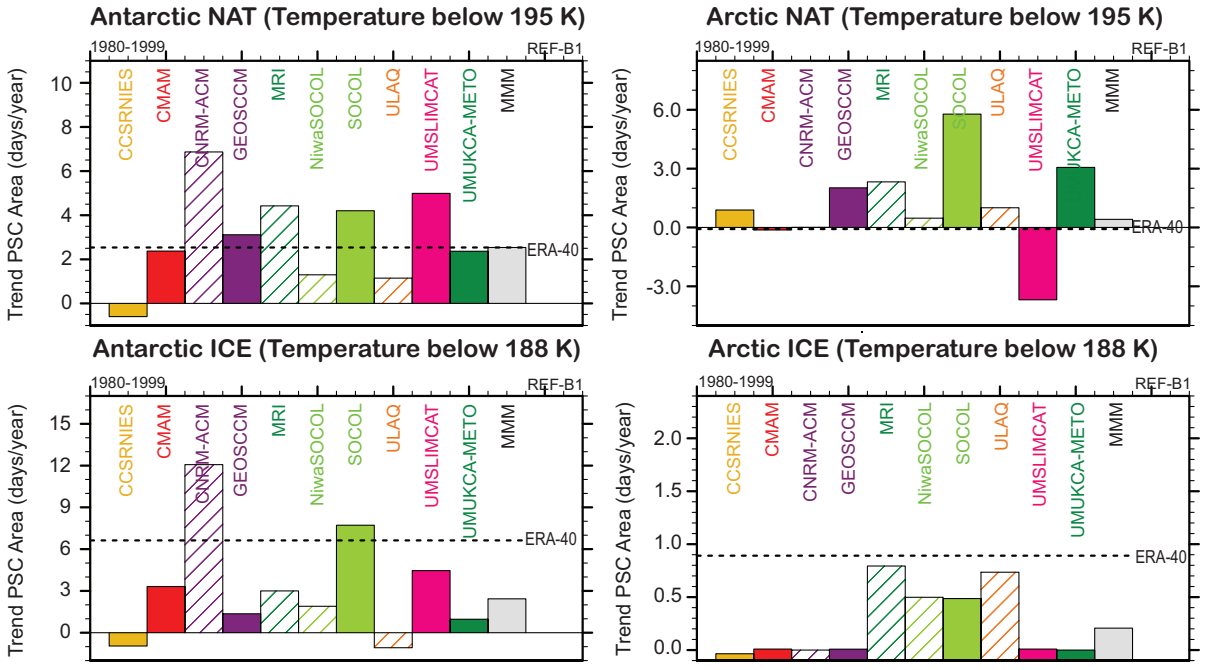

Figure 4.16: Linear trend (1980-1999) for the Antarctic (left) and the Arctic (right) of the seasonally accumulated area at 50 hPa where daily temperatures are below 195 K (approximate threshold temperature for NAT formation, top panels) and below 188 K (approximate threshold temperature for ice formation, bottom panels) for REF-B1 simulations. The dashed black line is the trend for the ERA-40 reanalysis (1980-1999).

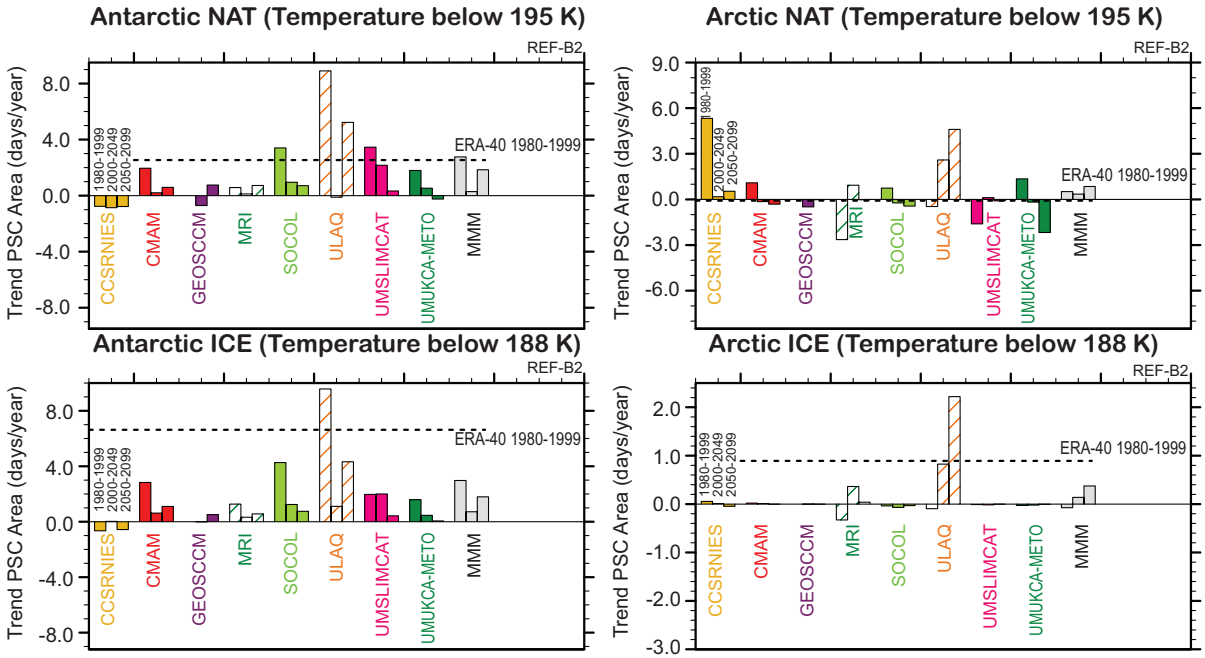

Figure 4.17: Linear trend (1980-1999, first column; 2000-2049, second column; 2050-2099 third column) for the Antarctic (left) and the Arctic (right) of the seasonally accumulated area at 50 hPa where daily temperatures are below 195 K (approximate threshold temperature for NAT formation, top panels) and below 188 K (approximate threshold temperature for ice formation, bottom panels) for the REF-B2 simulations. The dashed black line is the trend for the ERA-40 reanalysis (1980-1999).

period. There is a large spread among the models, with the UMUKCA-METO and the GEOSCCM being particular outliers with low values of both $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ (consistent with the warm bias at this height in both models — see Figure 4.1). The majority of the models simulate an increase in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ throughout the time period (Figure 4.16) leading to a positive trend of 2.5 and 2 %-days per year respectively, for the multi-model mean. In the case of $\tilde{A}_{NAT}$, this is close to the trend estimate from ERA-40, but for $\tilde{A}_{ice}$ is significantly smaller than the ERA-40 trend. Again there is a large model spread in the trends with some models indicating zero or small negative trends over the 1980-1999 period, and others simulating large positive trends.

In the Arctic there are large differences among the simulations in both the climatological mean values (Figure 4.15) of $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ and the trends (Figure 4.16) for the period 1980-1999. In general, the models simulate lower values of $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ than those derived from the ERA-40 reanalysis with the exception of the MRI and ULAQ models, which both have large cold biases in the NH winter (see Figure 4.1). Although the multi-model mean estimate of $\tilde{A}_{ice}$ is not significantly different from the ERA-40 reanalysis, there is large uncertainty in the ERA-40 estimate because of the large NH variability, and the multi-model mean estimate is dominated by the two significant outliers (MRI and ULAQ) with above average $\tilde{A}_{ice}$. The Arctic multi-model mean trend reflects almost no trend in $\tilde{A}_{NAT,}$ whereas for $\tilde{A}_{ice}$, a positive trend is simulated by all the models.

Trends in PSC quantities are hard to derive from global assimilation data and estimates of $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ can be quite different depending on the analysis or reanalysis data sets used (*e.g.*, Manney *et al.*, 2003, 2005a, b; Austin and Wilson, 2010). Therefore, those based on the ERA-40 reanalyses have to be interpreted cautiously. The ERA-40 reanalyses also show unrealistic vertical temperature oscillations in the Antarctic lower stratosphere in recent years, affecting PSC area calculations (Manney *et al.*, 2005a, b). In the Arctic, such behaviour is much less pronounced, and limited to the upper stratosphere and the last few years of the ERA-40 time series. For Arctic winters, ERA-40 temperatures also have a cold bias compared to Freie Universität Berlin data, which are able to capture the temperature extremes relevant for the PSC derived quantities (Manney *et al.*, 2003). Nonetheless, the time evolution and the positive trend between 1980 and 1999 for $\tilde{A}_{NAT}$ at 50 hPa exist in both observational data sets (Manney *et al.*, 2005a). Trends in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ for the REF-B2 simulations are shown in **Figure 4.17**. The trends for the period 1980-1999 can be compared with the corresponding trends in the REF-B1 simulations shown in Figure 4.16. With the

exception of the CCSRNIES model, all the other models which provided output from the REF-B2 simulations show an increase in Antarctic $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ over the period 1980-1999. A positive, statistically significant trend of $\tilde{A}_{NAT}$ of 2.5%-days per year is simulated for the multi-model mean. This positive trend has a similar size to the trend calculated from the ERA-40 reanalysis, however the trend in the ERA-40 data is not significant at 95% confidence using a standard t-test for the regression slope. For the period 2000-2049, all the models project a smaller positive or slightly negative trend in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$; these trends are statistically significant only for the SOCOL model and the UMSLIMCAT. In the following 50 years (2050-2099) all models, except for the CCSRNIES model and UMUKCA-METO, simulate a positive trend for Antarctic $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$, leading to a multi-model mean trend for both quantities of 2%-days per year with 95% significance. These trends are smaller than the observed trend in 1980-1999. The change in magnitude and sign of the trend for the different time periods is consistent with the radiative effects of simulated ozone depletion between 1980 and 1999 and slower ozone recovery in the 21st century (see Chapter 9), in addition to continued cooling of the stratosphere due to the prescribed increases in GHG amounts (Section 4.3.1; Eyring *et al*., 2007; Butchart *et al*., 2010).

In the Arctic there are large differences among the models in the trends in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ obtained from the REF-B2 simulations. Over the period 1980-2099 only a few of the REF-B2 simulations show an increase in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$, reflecting the role of interannual variability in influencing temperatures over the polar cap during Arctic winters (see Section 4.4). During the period 2050-2099, the CMAM, GEOSCCM, SOCOL and UMUKCA-METO time series show a decrease in $\tilde{A}_{NAT}$, whereas the CCSRNIES and ULAQ models simulate an increase, leading to a non-significant positive trend of 1%-days per year in the multi-model mean. All of the models simulate a decrease in the magnitude of the trend in $\tilde{A}_{NAT}$ from 1980-1999 to 2050-2099. In contrast, most models show an increase in the magnitude of the trend in $\tilde{A}_{NAT}$ from 2000-2049 to 2050-2099. Trends in Arctic $\tilde{A}_{ice}$ are very small in most models, with only those with large cold biases (ULAQ and MRI) showing a significant trend over the period 2000-2099.

To summarize, the multi-model mean predicts a significant increase in $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ for the Antarctic and no significant changes for the Arctic during the 1980-2099 period. On average in the Antarctic, the models show greater agreement in their climatological estimates of $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$ and their estimates of trends from 1980-1999. This gives confidence in model future projection for the SH. There is little agreement between the models concerning future Arctic amounts of $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$. This is almost certainly related to the large spread in climatological accumulated PSC amounts between the models. Much of this spread in the model ensemble is likely related to the differences in variability between the models (see Section 4.4) which has a strong influence on both the mean temperature climatology and year-to-year variability which is important for $\tilde{A}_{NAT}$ and $\tilde{A}_{ice}$.

4.4 Variability

4.4.1 Extra-tropical variability of the zonal-mean zonal wind

A realistic simulation of the winter-time extra-tropical stratospheric circulation indicates a more realistic description of dynamical troposphere-stratosphere coupling with implications for chemistry-climate interactions. The variability of the winter-time extra-tropical stratospheric circulation is mainly characterized by variations in the strength and location of the polar night jet. As with the climatological zonal-mean state (Section 4.3.1), the zonal wind variability is first assessed in terms of the strength and latitude of the maximum interannual standard deviation in the zonal-mean zonal wind. Results for the extra-tropical regions (see Section 4.4.2 for tropical variability), 45°N-90°N for the Boreal winter, and 30°S-80°S for the Austral winter are shown in **Figure 4.18**. Most models do not simulate the variability as well as they do the mean climatology. For the NH, the reanalyses show maximum variability close to the climatological mean jet maximum. All the models fail to capture the equator-ward tilt with height, and the maximum variability in the AMTRAC3 and the GEOSCCM occurs too far equator-ward in the upper stratosphere. The CNRM-ACM has too little stratospheric variability, whereas the MRI model, UMUKCA-UCAM and WACCM show too much variability, especially in the upper stratosphere. For the REF-B2 simulations for the same period (not shown), the results are very similar, except for the CAM3.5 and ULAQ model which have too little variability, and the UMUKCA-METO and UMSLIMCAT which have too much variability compared to the reanalysis. For the SH, the reanalysis shows the maximum variability occurs on the equator-ward side of the jet, fairly close to the QBO region. Most of the models have variability that is too weak and located too far poleward compared to the reanalyses. The REF-B2 simulations (not shown) have very similar biases.

The nature of the variability of the polar night jet can be further investigated by applying an Empirical Orthogonal Function (EOF) analysis to the extra-tropical zonal-mean zonal wind (*e.g*., Feser *et al*., 2000; Black and McDaniel, 2009). Here, an EOF analysis of the 50 hPa zonal-mean zonal wind is used to provide a more detailed assessment of the lower stratospheric variability in the

REF-B1 simulation than that obtained above using the interannual standard deviation. Furthermore, by considering all months, this EOF analysis captures seasons when the variability maximises. In the reanalysis data, variability of extra-tropical zonal-mean zonal wind in the stratosphere maximised during January to March in the NH, and during mid-October to mid-December in the SH (Thompson and Wallace, 2000). In general, the models capture this seasonality reasonably well, though the period when there is large variability is extended in several of the models compared to the reanalysis (not shown, but see Chapter 10, Section 10.3.2 and Figure 10.10 for the multi-model mean seasonal cycle in the annular mode variance).

In both the reanalysis and the models, the extra-tropical variability of the zonal-mean zonal wind in the stratosphere can be mainly described by two modes, with the first mode dominating. In the reanalysis data the leading mode clearly dominates in the NH, explaining 87% of the variance. In the SH, both modes contribute explaining 59% and 35% of the variance, respectively. **Figure 4.19** shows the eigenvalues of the first mode which is a measure of the variance described by this mode. The error bars indicate sampling error (see figure caption for details). **Figure 4.20** shows the spatial regression patterns of the first and second mode. The leading mode describes the variations in the strength of the eastward polar night jet while the second mode represents the meridional shift of the jet (Figure 4.20).

In the NH, nine of the sixteen models agree well with the reanalysis after allowing for the uncertainties given by the error bars, while two (five) models have larger (smaller) values (Figure 4.19). The two models with the largest low-bias (CNRM-ACM, ULAQ) also show the smallest maximum variability of the zonal-mean wind in mid-latitudes (Figure 4.20). Furthermore, the results for the ULAQ model are consistent with that model having the lowest frequency of stratospheric sudden warmings (SSWs) (see Section 4.4.3, Figure 4.25). The SSW frequency was not diagnosed for the CNRM-ACM due the absence of the appropriate data. A comparison of Figures 4.19 and 4.20 (top right) also indicates that the two models with the largest positive bias in the eigenvalue of the leading EOF (MRI, UMUKCA-UCAM) show larger than average maximum in the zonal wind variability in mid-latitudes, but more im-

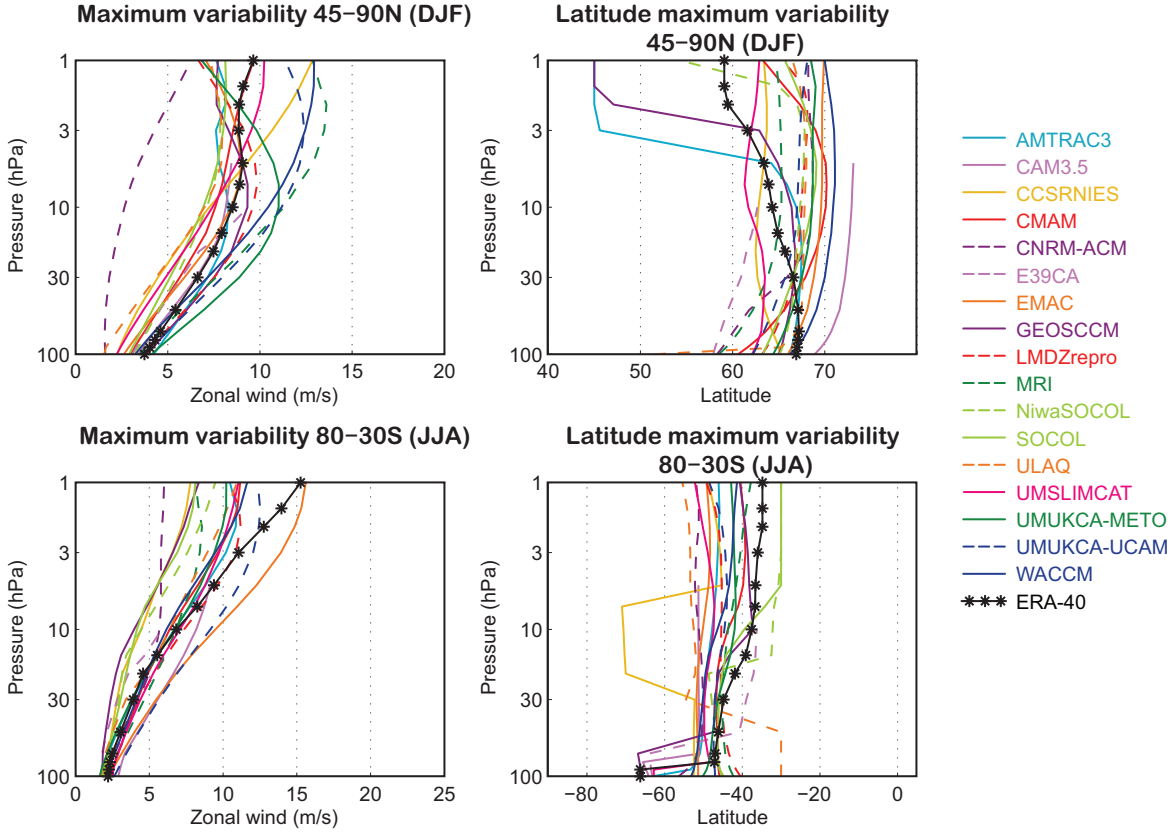

Figure 4.18: Location and amplitude of the maximum interannual standard deviation of the zonal-mean zonal wind in the NH in DJF poleward of 45°N (top) and in the SH in JJA between 80°S and 30°S (bottom). Data are based on the period 1980-1999 for the REF-B1 simulations and the ERA-40 reanalysis. Where an ensemble of simulations is available, the time series of the different members are concatenated before the calculation of the interannual standard deviation field.

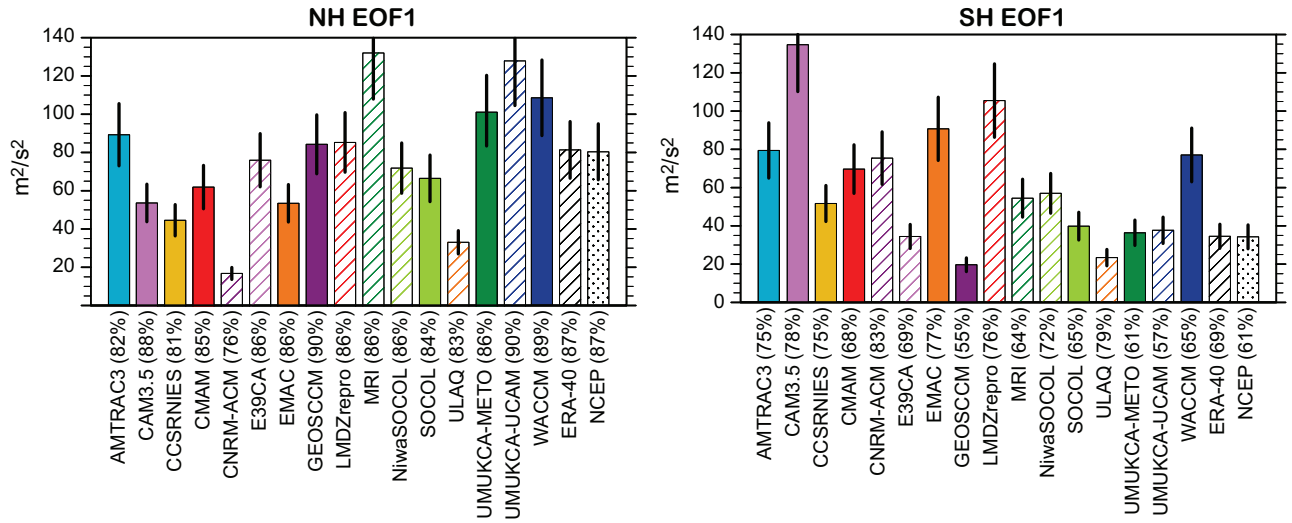

Figure 4.19: Eigenvalue of the leading mode of variability of the 50 hPa zonal-mean zonal wind (m^2/s^2) for the SH (right) and NH (left). Numbers in brackets (tick labels of the x-axes) indicate the fraction of the total variance explained by the leading mode. Error bars $2\Delta\lambda$ indicate the sampling error determined after North et al. (1982): $\Delta\lambda=\sqrt{(2/N)}$, where N is the sample size. With N = 60, a conservative estimate of the effective sample size is used considering long persistence (two months) in the stratosphere and weak zonal wind variations during 50% of the year. The EOF analysis was carried out for the NH (SH) 50 hPa zonal-mean zonal wind anomalies poleward of 45°N (°S). Monthly mean fields for all months from 1980 to 1999 are included with the seasonal cycle and linear trends removed. Data are also weighted with the square root of the cosine of latitude.

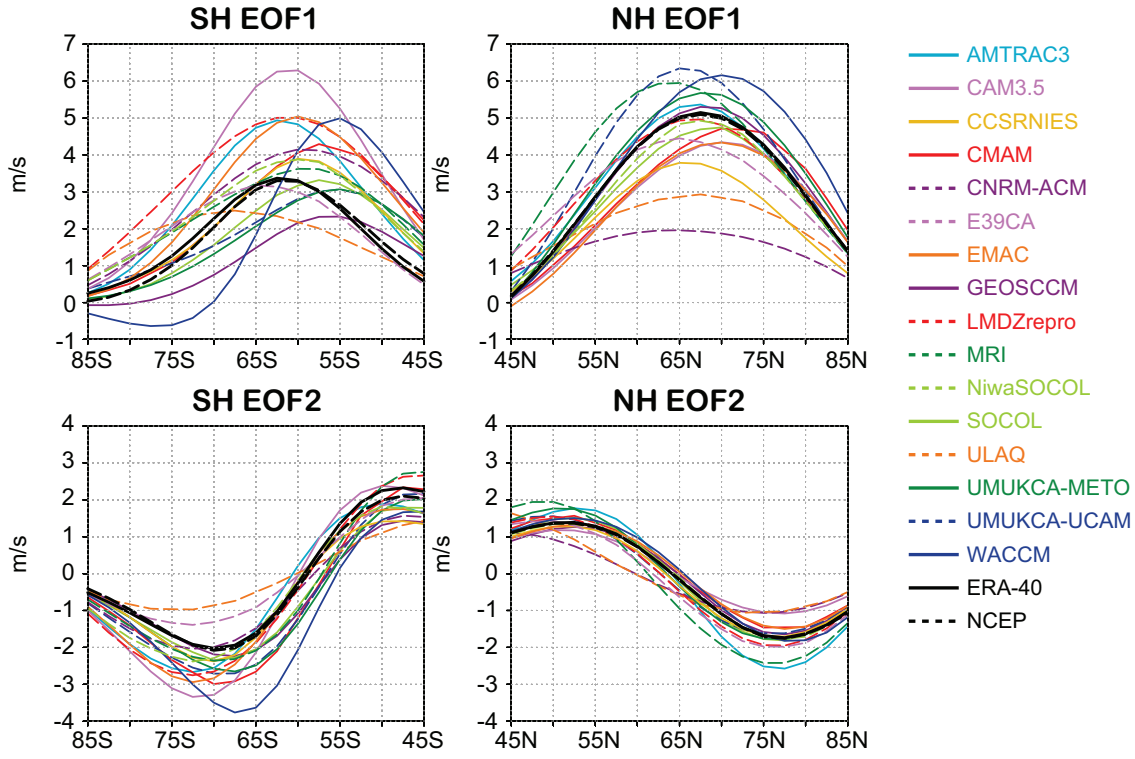

Figure 4.20: Regression patterns (m/s) of first (top) and second (bottom) mode of the 50 hPa zonal-mean zonal wind determined for regions poleward of 45°; (left) SH and (right) NH.

portantly, the variability of the jet maximises at a lower latitude than in the reanalysis, by up to ~5°.

The situation is quite different in the SH where only five models agree with observations in respect to the magnitude of the leading eigenvector (Figure 4.19). Ten models overestimate the variations in the strength of the polar night jet and two under-estimate this parameter. Figure 4.20 reveals a large spread both in the magnitude and location of maximum zonal wind variability between the models. The bias in the magnitude (and eigenvalue) is largest and positive in the CAM3.5. Several models show maximum zonal wind variations 10° further equator-ward than the reanalysis, while the ULAQ model shows maximum zonal wind anomalies shifted poleward by about 8°.

In summary, the analysis of the structure of the monthly mean zonal wind variability in the lower stratosphere reveals that most models simulate well both the magnitude and zonal structure of the variability in the NH. The largest biases result from an under-estimation of variability of the jet in the CNRM-ACM and ULAQ model, and a shift in the location of the maximum variability to lower latitudes in the UMUKCA-UCAM and MRI model. In the SH, most models exhibit large biases in the leading mode of variability of the zonal-mean zonal wind with subsequent implications for the second mode. Positive biases in the magnitude of the leading coupled mode and, hence, the variability of the polar jet, are related to a delayed break-up of the polar vortex in spring (Fogt *et al*., 2009) which is a common problem in most of the models (see Section 4.3.1). Several models also overestimate the variability in the SH compared to the NH.

4.4.2 Tropical variability of the zonal-mean zonal wind

A faithful representation of tropical variability above the tropopause has broad scientific relevance. In the tropical middle and upper stratosphere, the direction of the zonal-mean zonal winds, *e.g*., phase of the quasi-biennial oscillation (QBO), has been linked with the frequency of a disturbed polar vortex (Holton and Tan, 1980; Lu *et al*., 2008) and the rate of transport to higher latitudes of trace gases such as ozone (Li *et al*., 2008). Tropical variability can also influence processes thought to be relevant for maintaining the extra-tropical mean stratospheric climate and its variability.

Tropical variability in the REF-B1 simulations is first assessed in terms of the vertical profile of the interannual standard deviation in the de-trended zonal-mean zonal wind averaged between 10°S and 10°N (**Figure 4.21**). Below ~48 km (~1 hPa) all the models under-estimate tropical variability in comparison to ERA-40, with the exception of the WACCM at levels below 20 hPa. Five models exhibit particularly low stratospheric variability, largely due to the absence of either an internally generated or explicitly prescribed QBO.

Figure 4.22 shows the vertical profile of the amplitude of the variability in zonal wind at periods between 2 and 5 years (see figure caption for details). This range of periods captures possible QBO-like variability and it is evident from the figure which models neither prescribe nor internally-generate a QBO (*c.f*., Chapter 2, Table 2.8; Chapter 8, Table 8.4). Interestingly enough, there are still differences seen between those models that prescribe a QBO, possibly related to the fact that these models do not include any feedback mechanisms between the simulated ozone and the imposed artificial forcings. Furthermore, all models show a weaker peak amplitude for the QBO compared with ERA-40 (1980-1999).

The representation of the semi-annual oscillation (SAO) in the models can be seen in **Figure 4.23**. This shows the amplitude of the SAO, calculated using the same method as in Figure 4.22 but now including only the 6-month harmonic. Unlike for the QBO, the models shows a spread in peak amplitude of the SAO about the amplitude seen for ERA-40. The CAM3.5 and ULAQ model, exhibit SAO amplitudes significantly less than that obtained for ERA-40. On the other hand the GEOSCCM and LMDZrepro model have peak SAO amplitudes significantly larger than that for ERA-40 (*c.f*., Figure 4.29). Most likely this is due to a lack of a QBO in these models: the QBO in the lower stratosphere winds would act periodically to filter out parts of the model's resolved and/or parameterised gravity waves, responsible for driving the eastward phase of the SAO. The significance of any net model bias above ~32 km (10 hPa) must, however, be treated with caution due to the paucity of observations assimilated there by ERA-40.

The amplitude of annual cycle in tropical zonal-mean zonal wind in the REF-B1 simulations is shown in **Figure 4.24**, again derived using the same method as for Figure 4.22, but now including only the 12-month harmonic. The amplitude of the ERA-40 annual cycle shows two peaks; in the upper troposphere and at the stratopause. All the models exhibit a peak in the amplitude in the upper troposphere, with the ULAQ model and the CNRM-ACM having unrealistically large and small amplitudes, respectively.

All the models significantly under-estimate the amplitude of the annual cycle near the stratopause, although the MRI and EMAC models and the UMSLIMCAT perform better than the others. A weak annual cycle in the models may be linked with the overly strong SAO. Similar features are referred to in Osprey *et al*. (2010), using the a high-top version of the Met Office's global climate model. They link a reduced annual cycle at the tropical stratopause to an overly strong SH summer jet and stronger than observed westward circulation during JJA.

A brief comparison of the tropical zonal wind variability in the REF-B1 and REF-B2 simulations from 1980-

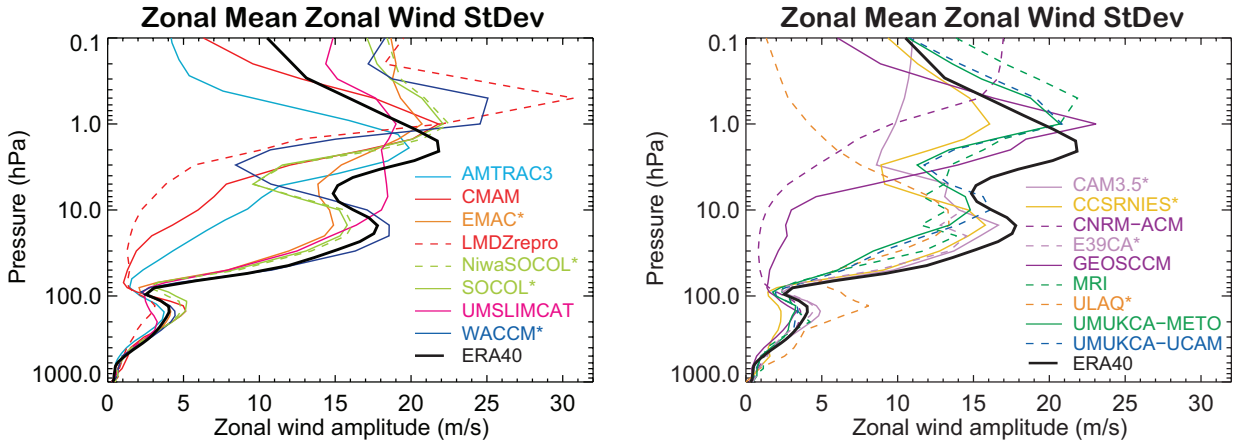

Figure 4.21: Profiles of the standard deviation in the de-trended zonal-mean zonal wind averaged from 10°S-10°N for the REF-B1 simulations. An asterisk after a model name indicates that that model has an externally forced (i.e., artificial) QBO.

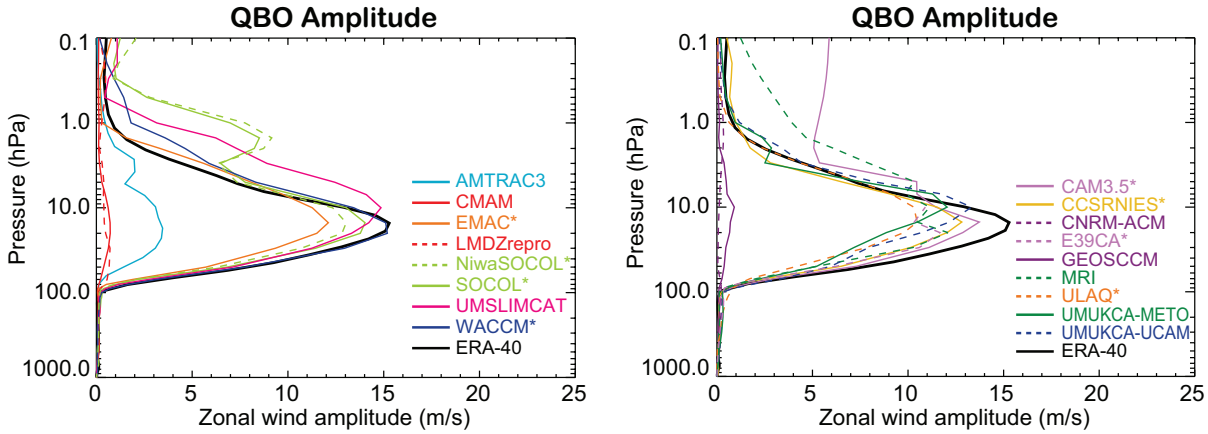

Figure 4.22: Profiles of the amplitude of the "QBO" (i.e., variability with periods between 2 and 5 years) in the zonal-mean zonal wind averaged between 10°S-10°N for the REF-B1 simulations. Methodology follows that of Pascoe et al., (2005). The amplitude is the ratio of the definite integral of the zonal wind power spectrum (between periods of 2 and 5 years) to the standard deviation of the zonal-mean zonal wind. The data was first detrended by removing the linear fit. An asterisk after a model name indicates that that model has an externally forced (i.e., artificial) QBO.

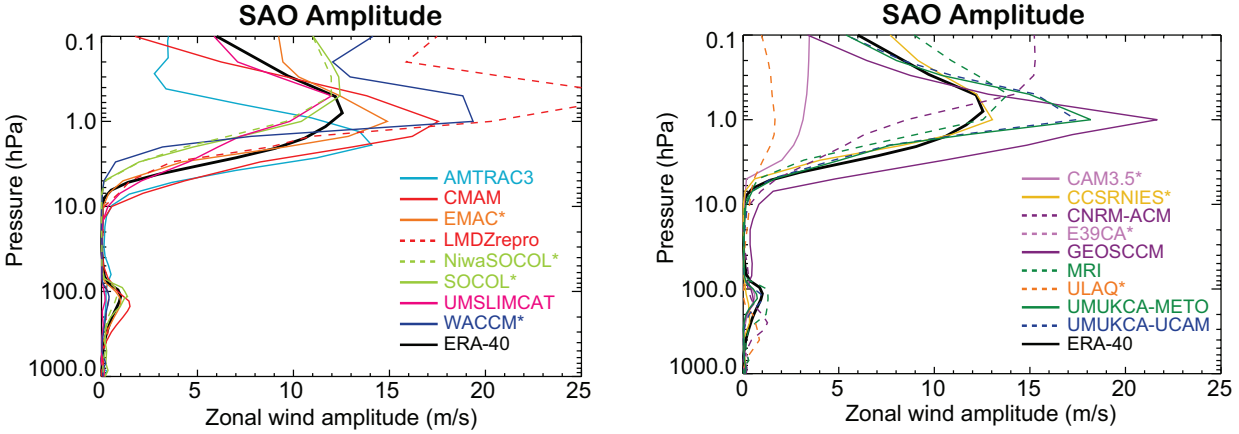

Figure 4.23: Profiles of the amplitude of the SAO in the zonal-mean zonal wind averaged between 10°S-10°N for the REF-B1 simulations. Method as for Figure 4.22, but including only the 6-month harmonic. An asterisk after a model name indicates that that model has an externally forced (i.e., artificial) QBO.

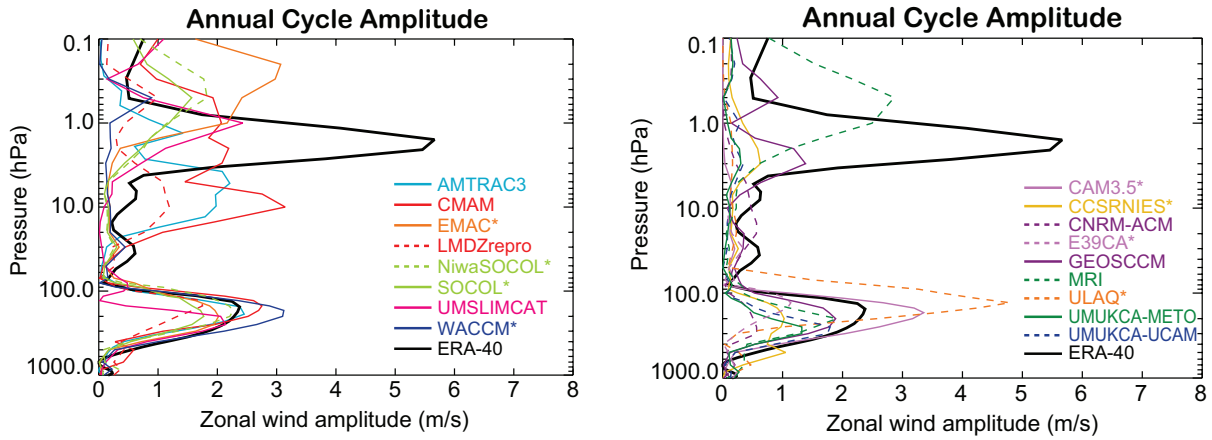

Figure 4.24: Profiles of the amplitude of the annual-cycle in the zonal-mean zonal wind averaged between 10°S-10°N for the REF-B1 simulations. Method as for Figure 4.22, but including only the 12-month harmonic. An asterisk after a model name indicates that that model has an externally forced (i.e., artificial) QBO.

2000 shows differences throughout the stratosphere, which are associated with a lack of a QBO in most of the REF-B2 simulations and a strengthened SAO (not shown). For those models with an internally generated QBO, only the UMUKCA-METO shows a slightly weaker QBO and SAO in the REF-B2 simulations. Neither the UMSLIMCAT nor MRI model show significant differences. As with the REF-B1 simulations, all the REF-B2 simulations exhibit a poor annual cycle in the upper stratosphere. Little systematic change or trend is seen in the magnitude of the tropical variability in the zonal wind in the stratosphere across the REF-B2 simulations. However, of the ten models compared, six showed a larger amplitude SAO from 2050-2099 compared with 2000-2049 (two smaller amplitude and two no change).

In summary, considering the large fraction of REF-B1 simulations that included nudging toward observations, there was an unexpected spread in tropical zonal wind variability. Most models under represented the amplitude of the QBO, while there was a large spread in the multi-model ensemble in the amplitude of the SAO. An excessively weak stratospheric annual cycle was common across all models. Finally, little trends were seen in the future REF-B2 simulations, although a significant fraction of the models showed an increase in the amplitude of the SAO.

4.4.3 Frequency of major stratospheric sudden warmings

Sections 4.4.1. and 4.4.2 considered interannual variability in monthly and seasonal mean fields. However a novel feature of this assessment compared to most previous assessments of stratosphere resolving models is the evaluation of variability on sub-monthly time scales using daily data. This is an important advance since much of the variability of the stratospheric polar vortex occurs on short time scales and spans the boundary between months. Moreover this intra-seasonal variability is known to contribute significantly to the interannual variability in the monthly and seasonal means. Its main manifestations are mid-winter major stratospheric sudden warmings and variability in the timing of the final warming or transition from winter to summer conditions. The simulation of these two phenomena is considered here and in Section 4.4.4, respectively.

In the extra-tropics major stratospheric sudden warmings (SSWs) play a key role in determining the mean climate and chemistry of the region. Obvious differences exist between the northern and southern winters due to the differences in the number of major SSWs. For the models and ERA-40 major SSWs are identified using the methodology of Charlton and Polvani (2007), based on reversals of the zonal-mean zonal wind at 60°N and 10 hPa, for the months November to March. **Figure 4.25** shows the mean frequency of major SSWs for both the REF-B1 and REF-B2 simulations (where results were available) compared to the frequency of major SSWs in the ERA-40 reanalysis. In contrast to a previous inter-comparison of stratosphere resolving general circulation models (Charlton *et al.*, 2007) most of the CCMs produce approximately the correct number of major SSWs over the second half of the 20th century (1960-1999). This result should not be taken to mean that models with interactive chemistry produce better dynamical variability (although a detailed investigation of this idea would be interesting), merely that this selection/generation of models appears to produce an improved simulation of major SSWs than those without interactive chemistry analysed in Charlton *et al.* (2007).

The only models with a significantly different frequency of major SSWs (at 95% confidence) when compared to the ERA-40 reanalysis are the AMTRAC3 and SOCOL and ULAQ models (which have a lower frequency of SSWs than the reanalysis) and the CMAM (which

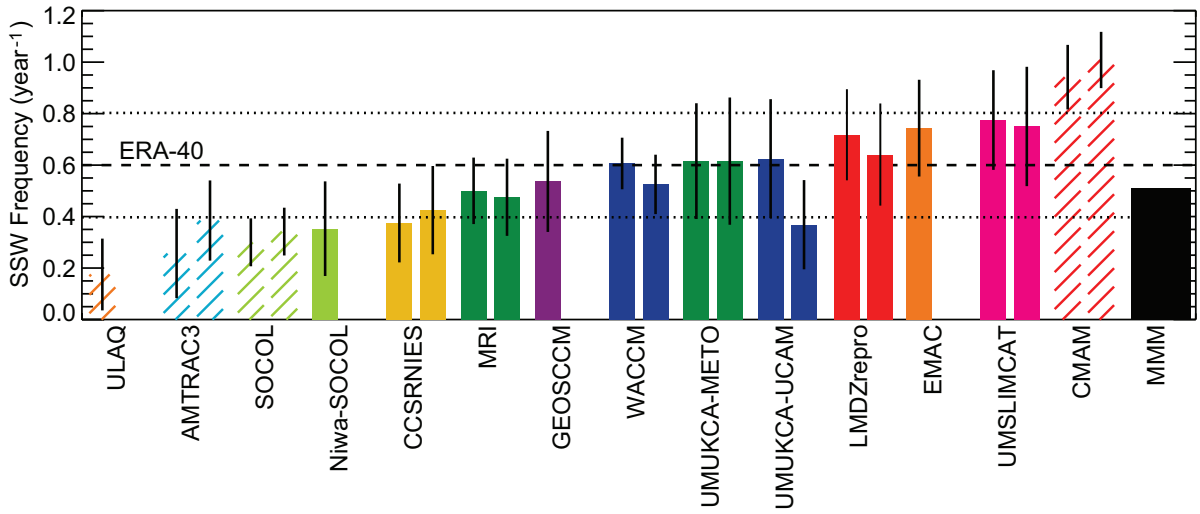

Figure 4.25: Mean frequency of NH major SSWs per year for the REF-B1 and REF-B2 simulations between 1960 and 2000. Dashed black line shows the mean frequency in the ERA-40 data set and 95% confidence interval (dotted lines). For each model, bars to the left of the model name indicate REF-B1 simulations and bars to the right of the model name indicate REF-B2 simulations. Where ensemble simulations are available, the mean frequency is calculated by combining all ensemble members. Bars are sorted according to major SSW frequency in the REF-B1 simulations. Where the frequency of major SSWs in the model and ERA-40 data set is significantly different at 95% confidence the bars are shown with hatched shading. Whiskers on each bar indicate a 95% confidence interval for the major SSW frequency.

has a higher frequency of SSWs than the reanalysis). The CMAM, which has a large number of major SSWs also has a mid-winter stratospheric jet with significantly reduced strength (see Figure 4.3). There is no significant difference between the REF-B1 (with prescribed sea-surface temperatures (SSTs)) and REF-B2 (with an interactive ocean) simulations of the CMAM, which suggests that coupling that model to an interactive ocean does not have a large impact on SSW variability. It should, however, also be noted that the version of the CMAM used for CCMVal-1 produced a realistic simulation of the number of major SSWs (McLandress and Shepherd, 2009).

A useful comparison of the impact of the full observed variability in SST forcing on SSW frequency can be made by comparing models which have REF-B1 and REF-B2 simulations available for the 1960-2000 period (REF-B1 simulations are run with observed SSTs while REF-B2 simulations, apart from those from the CMAM, are run with SSTs generated by atmosphere-ocean general circulation models). In Figure 4.25 models with two plotted bars show the frequency of major SSWs in the REF-B1 simulations in the left bar and in the REF-B2 simulations in the right bar. There appears to be little systematic difference between the number of major SSWs in the REF-B1 and REF-B2 simulations, except for the UMUKCA-UCAM. In all cases however, 95% confidence intervals for the SSW frequency (shown in black lines extending from the top of each bar) overlap for the REF-B1 and the REF-B2 simulations, suggesting that the differences largely result from sampling variations, even in the case of the UMUKCA-UCAM. For the SH winter period between 1960 and 2000 in the REF-B1 simulations no examples of a major SSW, similar to that observed during September 2002 (Shepherd *et al*., 2005), were found for any of the models based on the same criteria for major SSW occurrence as used for the NH.

A more detailed comparison of SSW variability in the models and the ERA-40 reanalysis can be made by plotting histograms of major SSW frequency in each month from November to March. **Figure 4.26** shows the climatology of major SSW events for the REF-B1 simulations. In general, models which produce SSWs with a frequency close to that of the reanalysis also tend to produce more realistic SSW climatologies, although there are some notable exceptions. In particular, the EMAC model tends to produce lots of dynamical variability at the start of winter and little during the mid-winter period. Further analysis of these events, suggests that they occur after the initial spin-up of the vortex in mid-September, although during the period in which the vortex is still relatively weak. In several other models, noticeably the CCSRNIES and two SOCOL models, the climatology of major SSWs is shifted toward the end of winter. This problem was noted in previous studies of SSW climatologies in models (Charlton *et al*., 2007) and may be related to the late final warming, which occurs in some of the models (particularly SOCOL and NiwaSOCOL).

Any significant trends in SSW frequency in the NH, could have important consequences for both ozone chemistry and the signal of climate change in the lower stratosphere. Recently, Charlton-Perez *et al*. (2008) showed that simulations of the AMTRAC predicted a small increase in major SSW frequency over the 21st century. McLandress and Shepherd (2009) also note a similar trend in the CMAM, but suggest that this trend may simply reflect the change in stratospheric climatology rather than a real increase in stratospheric variability. In general, when comparing the frequency of major SSWs in the period 1960-2000 with the frequency of major SSWs in the period 2060-2100 projected by the models, a mixed result is found. The majority of models simulate either no change in the SSW frequency or a small increase in the late 21st century. The multi-model mean projection is therefore for a slight increase in the SSW frequency in the later half of the 21st century, although the trends in heat flux (January, 100 hPa) over the 21st century show a slight decrease between 2050-2099 (see Section 4.3.4). However, it is emphasized that some caution should be exercised when considering projected trends in heat flux since the majority of models do not suggest a significant trend. In addition, the subset of models considered here is not the same as the subset used for the heat flux comparison in Section 4.3.4 due to data availability. It is also important to note that changes in the SSW frequency are likely to be strongly influenced by changes in heat flux variability rather than just by changes to the mean heat flux climatology.

4.4.4 Timing of final warmings / winter-summer transition

The timing of the final warming (the date at which the winter-time polar vortex breaks down and is replaced by the summer-time stratospheric westward circulation (Andrews *et al*., 1987)) is an important diagnostic related both to the climatological vortex breakdown shown in Figure 4.2 and the study of stratospheric variability in Section 4.4.3. Studies by Black and McDaniel (2007a, b) and Black *et al*. (2006) have shown that there is an important dynamical link between the stratosphere and troposphere as the final warming takes place and that the timing of the final warming is highly variable from year to year. In addition, there is a clear trend toward later final warming dates over the 22 years between 1979 and 2001 (Waugh *et al*., 1999; Waugh and Rong, 2002) in the SH which is related to coupling between dynamics and ozone depletion and therefore should be captured by the CCMs.

Final warming dates in both the NH and SH are calculated using the method of Black and McDaniel (2007a, b) and 5-day low-pass filtered zonal-mean zonal wind data at 50 hPa from the models and reanalysis. This method defines the final warming as occurring when the zonal-

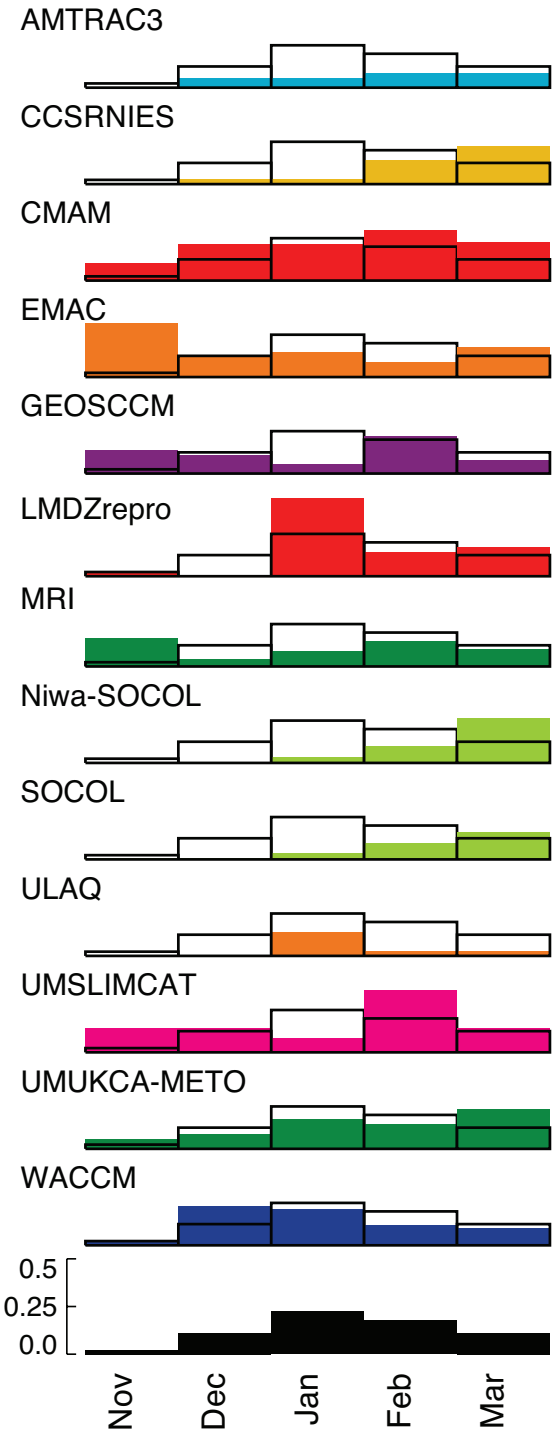

Figure 4.26: Histograms showing the frequency of major SSWs (in events per year) in the REF-B1 simulations (1960-2000, coloured bars) in comparison to ERA-40 reanalysis (open bars). Where an ensemble of simulations is available, the plot reflects major SSWs observed in all ensemble members. ERA-40 reanalysis climatology is reproduced at the bottom of the plot for comparison.

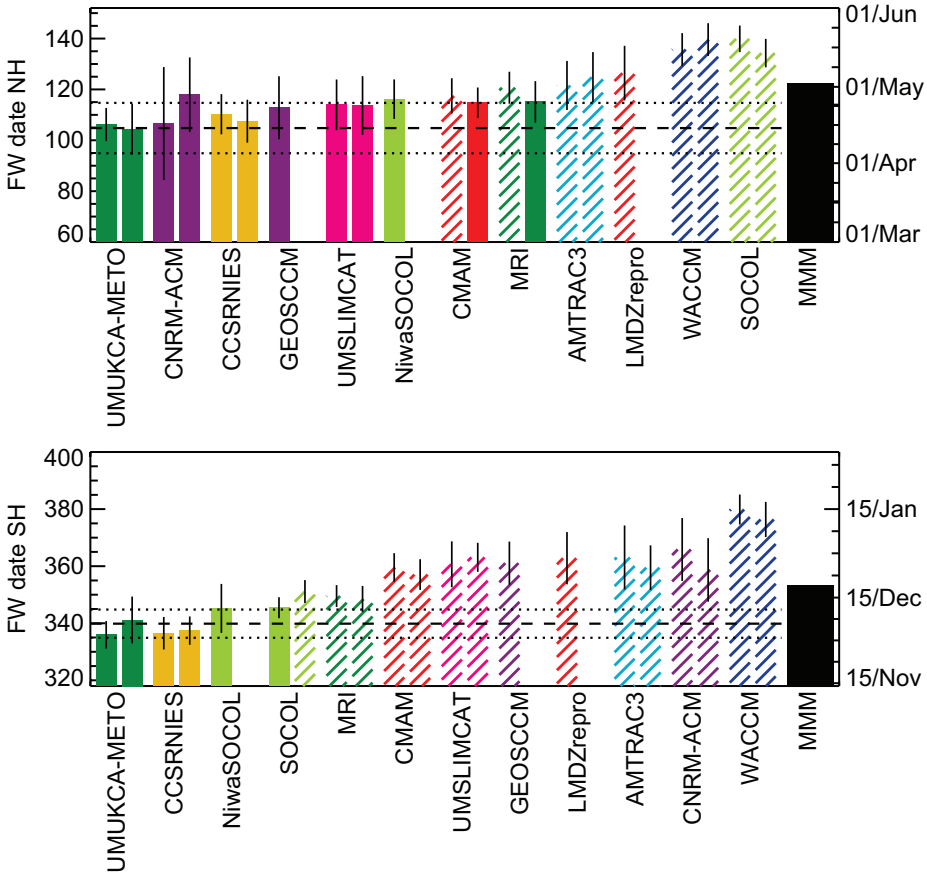

Figure 4.27: Mean date of the NH and SH final warmings (day number) for the REF-B1 (leftmost bars for each model) and REF-B2 (rightmost bars for each model) simulations (1980-1999). Black dashed line shows the mean final warming date for the ERA-40 data with 95% confidence estimates shown in dotted lines. Models are ordered by the mean date of their final warming in the REF-B1 simulation. Where a significant difference between models and the ERA-40 reanalysis estimate is observed, the bar is plotted with a hatched shading. Where an ensemble of simulations is available, the statistic reflects the mean of all three ensemble members. Black whiskers on each bar indicate twice the standard error for each estimate. Approximate comparable calendar dates for a non-leap year are included on the right-hand axis.

mean zonal winds at a specified latitude cross a low-wind threshold (0.0 m/s in the NH and 10 m/s in the SH) and do not return to eastward values before the next winter (see Black and McDaniel (2007a, b) for further details). For some models, zonal-mean zonal winds never cross the low-wind threshold in some years, these years are ignored in the analysis. The occurrence of these years is not frequent enough in any of the models (typically of the order 1 or 2 winters in a given 20 year period) that it would impose a significant bias requiring modification of the identification technique.

Figure 4.27 shows the mean date of the final warming for the NH and SH for the REF-B1 and REF-B2 simulations over the period (1980-1999). The black dashed line shows an estimate of the final warming date for the ERA-40 reanalysis data set for the same period. Figure 4.27 shows that models have final warming dates in both the NH and SH which are generally either at or later than the date of the final warming in the reanalysis data. This result is consistent with the diagnoses in Figure 4.2 of the descent of the mean, climatological zero-wind line. Of the models considered, in both hemispheres, more than half have mean final warming dates significantly later than those in the ERA-40 reanalysis. There are particularly large differences in the SOCOL model and the WACCM in the NH and the CNRM-ACM and the WACCM in the SH.

Trends in the date of the final warming in the SH for the three periods 1980-1999, 2000-2049 and 2050-2099 are shown in **Figure 4.28** along with the multi-model trend estimate for each period. Although there is some spread between models, the multi-model mean trend shows the expected pattern of large positive values over the recent past (~+1 day per year, consistent with the ERA-40 estimate) and smaller negative values (~-0.3 days per year) during

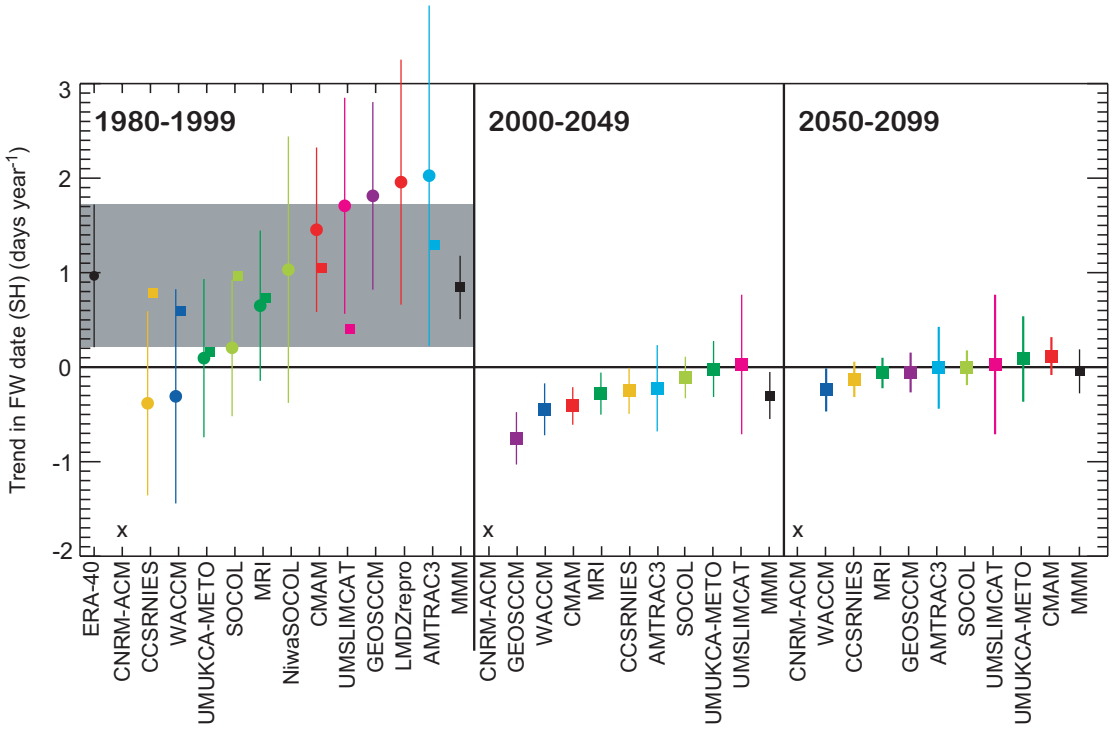

Figure 4.28: Linear trend in the date of the SH final warming from the REF-B1 and REF-B2 simulations. Where an ensemble of simulations is available the trend is calculated for the ensemble mean final warming date in each year. Grey shading shows 95% confidence limits for the ERA-40 estimate. Whiskers on each point show respective 95% confidence limits. Left panel shows trends for the period 1980-1999, middle panel shows trends for 2000-2049 and right panel shows trends for 2050-2099. If not enough data exists to estimate a consistent trend, model is plotted with an X near to the x-axis. Models are shown in order of the trends for the REF-B1 simulations in the left panel and the REF-B2 simulations in the middle and right panels. Where both REF-B1 and REF-B2 simulations are available for the 1980-1999 period, REF-B1 trends are shown with a circle and REF-B2 trends shown with a square. Multi-model means are calculated by weighting each trend with its standard error.

the period of ozone recovery and is statistically significant at 95% confidence in both cases.

Part of the inter-model spread in the estimated final warming trend may be related to mean final warming biases over the same period. Of the models which fail to capture a strong trend in the recent past (**Figure 4.28**), two of them (SOCOL and WACCM) have very late mean final warming dates (see Figure 4.27). During the period of ozone recovery (2000-2049) there is reduced inter-model variability in the trend estimate, almost all the models show a reversed negative trend in the SH final warming date, toward earlier final warmings (Figure 4.28). These results reinforce the idea that the spring-time ozone concentration plays a large role in determining the final warming date and that the models are able to capture this coupling between chemistry and climate satisfactorily. In the following period (2050-2099), when the rate of ozone recovery is smaller in most models and the effects of changes in GHG forcing are larger, trends in the final warming date are much smaller and not statistically significant. In the NH (not shown), during the period 1980-1999, a significant, positive multi-model trend is simulated similar to that seen in the SH, however this trend is not consistent with the corresponding, observed reanalysis trend for the same diagnostic. In the two future periods, no significant individual or multi-model trends were simulated in the NH.

4.5 Conclusions

4.5.1 Multi-model summary

In Sections 4.3 and 4.4 the reproduction by the climate-chemistry models of those stratospheric dynamical processes and phenomena that are considered important for modelling the long term evolution of stratospheric ozone and the impact of the stratosphere on climate, have been assessed by comparison with observation and/or through model inter-comparisons. The response and robustness of the response (*i.e.*, model independence) of the various dynamical processes and phenomena to climate and long-

term ozone changes was also noted (see Section 4.5.4).

For the mean climate there was generally good agreement among the majority of the models both in their strengths and weaknesses when evaluated against the reanalysis data sets. Overall the models reproduce the extra-tropical temperature and zonal-mean climatology very well with the notable exception of some key phenomena. Common problems for many of the models were a SH polar night jet that is too strong, lacks the observed equatorward tilt with height, and that persists too long into spring before changing to the summer-time westward circulation. Concomitantly, many models had a cold bias in the Antarctic lower stratosphere in spring. These problems in the SH are almost certainly inter-related and existed in earlier versions of many of the models (*e.g.*, Eyring *et al.*, 2006). On the other hand the causes of the weaknesses are not well understood. Possibly they are related to a deficiency in the wave driving from the troposphere though, in both hemispheres the climatological upward flux of resolved wave activity entering the stratosphere in mid-winter was remarkably well reproduced in nearly all the models. Moreover, most models displayed, more or less, the correct temperature response to variations in this wave flux. Again the majority of models accurately reproduce the strength of the Brewer-Dobson circulation, or at least the tropical upwelling mass flux at 70 hPa, though there was considerable uncertainty across the models as regards the contribution of the different wave drags (*i.e.*, EP-flux divergence and parameterised orographic and non-orographic gravity wave drag) driving the upwelling. The models also showed, on average, less inter-hemispheric asymmetry than is observed in both the wave driving and also the stationary waves.

Variability in the stratosphere on all time scales from the intra-seasonal to interannual was, in general, less well reproduced by the models, with a large inter-model spread for some diagnostics. Most obviously, many models place the transition between winter and spring conditions significantly later than observed. This deficiency is seen in both hemispheres and in diagnostics of the mean climate (see above) and individual final warming dates. In the northern winter, although most models accurately simulate the frequency of major stratospheric sudden warmings, the climatology of these key events is poorly simulated with the worst performing models producing only early and late winter warmings. Nonetheless in mid-winter in both hemispheres the polar night jet has the correct modes (EOFs) of variability in the lower stratosphere in all the models. However, in agreement with the other diagnostics, there is considerable inter-model spread in the amount of variability in the strength of the polar night jet (the first EOF of the zonal-mean zonal wind).

In the tropics, there remains a large divergence between model design and/or experimental setup and the subsequent ability to simulate tropical variability, particularly the QBO. Models can be split into three groups: those that have very little or no stratospheric variability, those that impose an artificial QBO and therefore can not be considered as free running climate-chemistry models, and those with an internally generated QBO. Even with a QBO (artificial or not), the variability in the models is less than observed. An assessment of the impact of these modelling choices on the extra-tropical stratospheric variability or the variability of ozone is presented in Chapter 8.

An implicit assumption of the above assessment is that any dynamical biases in the models are the result of a poor representation of dynamical processes in the models. However, because the models considered are coupled climate-chemistry models, errors in the simulated radiatively active gases and, in particular, ozone can lead to an apparent error in the dynamical fields when evaluated against observations where no dynamical bias in the model exists. The representation of other dynamical process in the models such as the QBO may also be quite sensitive to details of the simulated ozone (*e.g.*, Bushell *et al.*, 2010), even if this lies within the observational errors. In general however, it is thought that any errors in ozone are unlikely to affect the overall conclusions of this chapter apart from over Antarctica where there is rapid ozone change and significant model spread in ozone behaviour (see Chapter 9). Because the Antarctic ozone change projects strongly on to the Southern Annular Mode (SAM), uncertainties in the simulated ozone can lead to uncertainties in the representation and behaviour of the SAM in the models. Quantitative uncertainties in the projected long-term trends of ozone (Chapter 9) will also introduce quantitative uncertainties in the long-term trends in some of the dynamical fields and processes reported in this chapter.

4.5.2 Summary by model

In this section, a summary of the performance of each model in the range of dynamical processes assessed in the chapter is provided. This summary and, in particular, the overall assessment at the end of each paragraph is subjective and thereby gives a more complete picture of the overall dynamical performance of each model than can be obtained from a limited number of metrics. In the next section this subjective assessment is complemented by a quantitative summary of several key processes in terms of appropriately chosen metrics.

The **AMTRAC3** simulates the stratospheric mean temperatures in winter and spring well in both hemispheres, but the strength of its mid-winter jet is significantly stronger than the reanalysis climatology in both hemispheres, and the model is an outlier in the multi-model ensemble in the NH. The mean meridional circulation was not assessed. Stratospheric variability in the AMTRAC3 is somewhat

weaker than observed, particularly in terms of the number of major SSWs and is consistent with its stronger than average winter jet in the NH. The final warming is significantly late in both hemispheres. In the tropics the AMTRAC3 exhibits weak, internally generated variability in the QBO region but simulates well the variability in the SAO region. Overall, the AMTRAC3 simulates stratospheric dynamics adequately, but has some dynamical biases particularly in the SH.

The **CAM3.5** provided only limited diagnostics. In the diagnostics produced, the CAM3.5 has significant biases in its mean state close to the upper boundary at 3.5 hPa, particularly in the NH. The mean meridional circulation is adequately simulated. In the tropics the model imposes a QBO *via* external forcing.

The **CCSRNIES** model has significant large biases in its mean state during winter and spring in both hemispheres. In general, the stratospheric vortex in this model is too strong and has very low temperatures at its core. The CCSRNIES model produces an accurate simulation of the mean meridional circulation. Stratospheric variability in the NH is slightly weaker than the multi-model mean which may be related to the relatively weak meridional heat flux climatology at 100 hPa in the model. In the tropics the model imposes a QBO *via* external forcing. Overall, the CCSRNIES model has an adequate representation of stratospheric dynamics.

The **CMAM** suffers from significant biases in its mean state of opposite sign in the two hemispheres. In the NH, the mid-winter polar vortex is too weak above 30 hPa and is too dynamically active. In the SH, the mid-winter polar vortex is too strong, and breaks up too late in the spring season. The mean meridional circulation is consistent with that derived from meteorological analyses. In the tropics, the CMAM does not simulate a QBO, but has an SAO of large amplitude. Overall, the CMAM has a mixed dynamical performance, and is sometimes an outlier in the multi-model distribution.

The **CNRM-ACM** produces a stratospheric mean state with significant biases in temperature and jet strength and position. It has particularly large biases in the NH, positioning its mean jet too far equator-ward. The meridional circulation has generally the correct strength but has an unusual structure, with downwelling seen at the equator, and a seasonal cycle with a minimum in MAM as opposed to JJA. Only a limited assessment of stratospheric variability was possible, but significant problems, including a lack of 100 hPa heat flux during the NH mid-winter and a late break-up of the vortex in the SH were identified. In the tropics, the CNRM-ACM does not simulate a QBO, but has an SAO of large amplitude. Overall, the CNRM-ACM has significant dynamical problems, particularly in the SH and is frequently an outlier from the multi-model mean.

The **E39CA** model provided only limited diagnostics. In the diagnostics produced, the E39CA model has a large cold bias in its mean state in the middle stratosphere of both the NH and SH near the rather low model top (~10 hPa) and large biases in the spring SH lower stratosphere. Linked to these biases, the model has very strong jets during both NH and SH winters. The mean meridional circulation is weaker than the multi-model mean and that derived from analysis data and exhibits mean downwelling at the equator. Limited diagnosis of stratospheric variability in the extra-tropics was carried out. In the tropics the model imposes a QBO through external forcing, although the model has a top boundary at 10 hPa. Overall, the E39CA model has significant dynamical biases in both hemispheres and is a particular outlier in the multi-model ensemble in the middle stratosphere, especially near the model top.

The **EMAC** model simulates the stratospheric mean state in winter and autumn well in both hemispheres, despite relatively small values of 100 hPa heat flux in the NH mid-winter and large values of 100 hPa heat flux in the SH mid-winter. In the SH, the EMAC model exhibits a stronger relationship between heat flux and spring polar temperatures than most models or observations, and a very cold bias for undisturbed vortex conditions. This suggests compensating errors help to produce its accurate SH mean state. The EMAC model simulates stratospheric variability well, although with too much dynamical variability in November in the NH. In the tropics the model imposes a QBO through external forcing. Overall, the EMAC model has an adequate representation of stratospheric dynamics.

The **GEOSCCM** simulates the stratospheric mean state in winter and autumn well in both hemispheres. Its meridional circulation in the middle stratosphere is somewhat weaker than the circulation estimated from observations. Stratospheric variability in the GEOSCCM is generally well simulated, although it does exhibit a late SH vortex break up. In the tropics, the model does not produce a QBO but has a large amount of variability in the SAO region. Overall, the GEOSCCM simulates stratospheric dynamics well, with better performance than the multi-model mean in most diagnostics.

The **LMDZrepro** model simulates the stratospheric jet strength in NH mid-winter well, but has significant warm biases in winter and spring in the upper stratosphere and in winter in the lower stratosphere. In the SH, the LMDZrepro model has similar warm biases in winter and spring above 5 hPa and a strong bias in jet strength above and at 10 hPa.

The mean meridional circulation was not assessed. Mid-winter variability in the stratosphere is well simulated, but vortex break-up in both hemispheres is too late. In the tropics, no QBO is simulated but the SAO region has a large amount of variability. Overall, the LMDZrepro model produces an adequate representation of stratospheric dynamics.

The **MRI** model simulates the stratospheric mean state in winter and spring well in the middle and upper stratosphere but has a significant cold bias in the lower stratosphere (below 50 hPa) in all seasons. The structure of the mean meridional circulation is well represented in the MRI model, although it is one of the strongest in the multi-model ensemble. Stratospheric variability is also well simulated by the model but the final warming is slightly too late in both hemispheres. In the tropics, the model has an internally generated QBO with an amplitude comparable to the observations and has a large amount of variability in the SAO region. Overall, the MRI model simulates stratospheric dynamics well with better than or similar performance to the multi-model mean in most diagnostics.

SOCOL and **NiwaSOCOL** simulations are considered together, since they use versions of the same model. However, in several diagnostics there are considerable differences between simulations by the two model versions. An important point to note here is that many of the SOCOL diagnostics are based on the mean performance of a three member ensemble, whereas for NiwaSOCOL there was only one realization, which may explain some of the discrepancy. The SOCOL models simulate the stratospheric mean state in winter and spring well in both hemispheres although they both have significant biases in the SH lower stratosphere in spring. Tropical upwelling in the SOCOL model shows a maximum on the equator, and the annual cycle of the mean meridional circulation is also qualitatively wrong. The NiwaSOCOL model does little better, and shows the strongest circulation in the multi-model ensemble. Stratospheric variability in the model is weak, perhaps linked to the small amounts of heat flux at 100 hPa and to the late final warming in the NH. In the SH the relationship between heat flux and lower stratospheric temperatures is well simulated. In the NH, the final warming simulated by the SOCOL model is significantly later than that observed. In the tropics, the model imposes a QBO through external forcing. Overall, the SOCOL model has a moderate representation of stratospheric dynamics with notable shortcomings.

The **ULAQ** model has significant biases both in the position and strength of the stratospheric jet in the SH and is a large outlier in the multi-model ensemble. Its meridional circulation is too weak and exhibits an incorrect seasonal cycle. Stratospheric variability is generally too weak. In the tropics, the model imposes a QBO through external forcing. Overall, the ULAQ model has limited success in simulating stratospheric dynamics and is an outlier in the multi-model ensemble for some diagnostics.

The **UMSLIMCAT** simulates the stratospheric mean state in winter and spring well in the NH but has a slight strong bias in the strength of the winter jet in the upper stratosphere of the SH. It also has a large cold bias in the middle stratosphere in the SH spring. The mean meridional circulation was not assessed. The model produces a good simulation of stratospheric variability, although its final warming in the SH is significantly too late. In the tropics, the model internally generates a good simulation of the QBO and SAO. Overall, the UMSLIMCAT simulates stratospheric dynamics well, with better performance than the multi-model mean in most diagnostics.

The **UMUKCA-METO** and **UMUKCA-UCAM** have the same dynamical core. Although the two models only had small differences in the experimental setup there are considerable differences between the simulations for several of the diagnostics. In general, however, the UMUKCA models simulate the stratospheric mean state in winter and spring well in both hemispheres, apart from a large warm-bias near 100 hPa in the SH spring. The mean meridional circulation is well simulated by the models, although weaker than the multi-model mean. Unlike the other models, parameterised orographic gravity wave drag does not contribute significantly to driving the meridional circulation. Stratospheric variability is well simulated by the model in all diagnostics. In the tropics, the model internally generates a good simulation of the QBO and SAO. Overall, the UMUKCA models simulate stratospheric dynamics well, with better performance than the multi-model mean in most diagnostics.

The **WACCM** simulates the stratospheric mean state well in the NH but has a very large cold bias throughout much of the stratosphere in the SH spring and a very strong stratospheric jet in the SH in winter. This cold bias is linked to a very late break-down of the vortex in both the climatological annual cycle and final warming diagnostics, causing the model to be a significant outlier in the multi-model ensemble. The model has a good simulation of the mean meridional circulation. Stratospheric variability in the NH is well simulated by the model. In the tropics, the model imposes a QBO through external forcing. Overall, the WACCM simulates a great deal of stratospheric dynamics well but has a significant problem in simulating the SH vortex break-down.

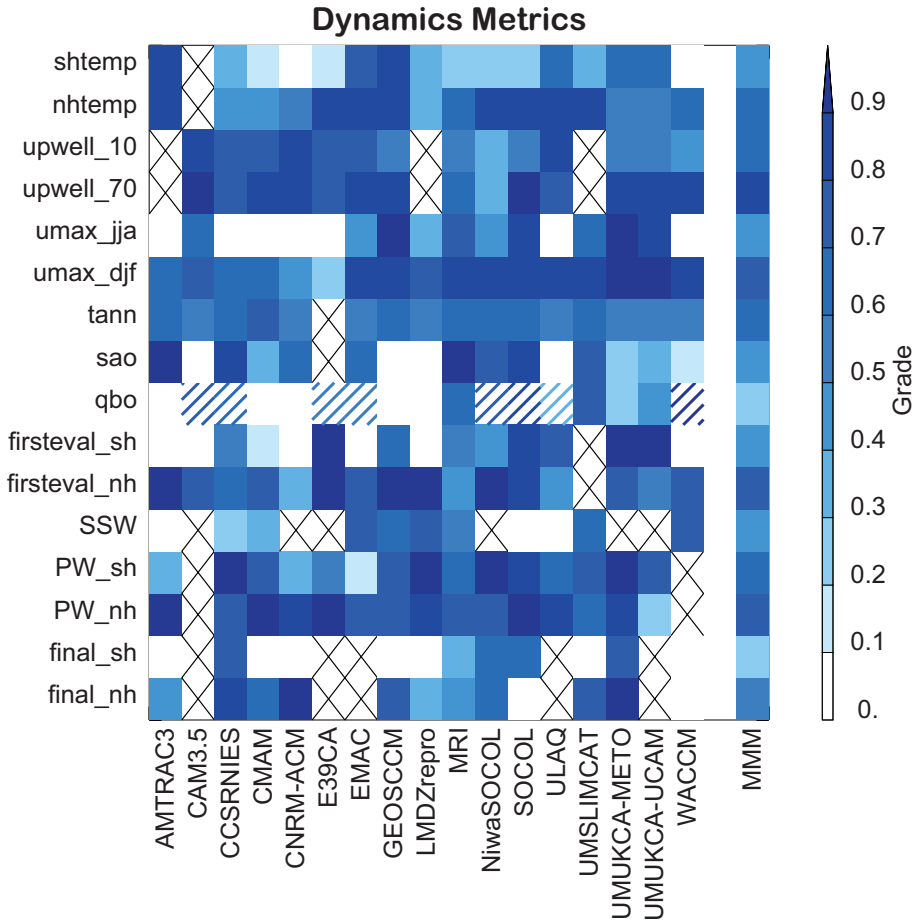

Figure 4.29: Matrix showing the performance of the model ensemble in a variety of metrics described in Table 4.1a, b after Waugh and Eyring (2008). See text for details of the matrix calculation. Each pixel represents the performance of a model in a given metric, darker colours indicate better performance. Crosses in a pixel indicate the metric was not evaluated for that model. On the right of the diagram the average metric score for each diagnostic is shown. Hatching in the QBO metric indicates the model uses a relaxation scheme to produce QBO variability, models with a relaxation scheme are not included in the calculation of the mean metric.

4.5.3 Quantitative assessment / metrics

To establish the fidelity and quantify the assessment of the simulation of stratospheric dynamics by the models, "metrics" representing most of the dynamical diagnostics analysed in Sections 4.3 and 4.4 have been identified. The full list of metrics is presented in Tables 4.1a and b. The list has some metrics in common with the study of Waugh and Eyring (2008), but also extends that list, particularly in the area of stratospheric variability.

This list captures the main dynamical processes in the stratosphere. A pragmatic approach has, however, been used and for many diagnostics the metric opted for requires the least input of dynamical fields or complex analysis. There are a number of cases where diagnostics might be replaced with more dynamically meaningful alternatives if more data from more model runs were available. For example, the diagnostics "shtemp" and "nhtemp" might be replaced by the area of temperatures below PSC formation thresholds. Metrics are calculated as listed in the table and normalised using Equation 4 of Waugh and Eyring (2008). In order to maintain some consistency with the Waugh and Eyring analysis, scores are standardized using the standard deviation of the observed quantity in question. For the tropical variability, estimating the uncertainty in the ERA-40 reanalysis is more complex. To estimate the uncertainty, the data set was re-sampled for several 10-year periods, and the range of possible values for the amplitudes of the annual cycle, SAO and QBO was used in the metric calculation. An attempt is also made to assess where a model's performance (as assessed by the metric) is different from the observations and different from the multi-model mean.

Figure 4.29 shows the metric portrait of the models. For each model, every metric is assigned a box on the diagram, and the box is shaded according to model performance. Darker colours indicate that a metric is closer to 1,

i.e., a very good performance in that metric. Black crosses indicate that insufficient data was available to calculate the metric for that model. On the right of the diagram, mean metrics across the multi-model ensemble are calculated for each metric. For the QBO metrics, models which relaxed the tropical winds to the observed QBO are shown hatched. The calculation of the mean metric for the QBO does not include those models with tropical wind relaxation.

As well as impacting on the QBO metric ("qbo"), constraining the winds in the tropics towards observations in some models is likely to impact on many of the other processes and phenomena considered in this chapter and, indeed, other chapters too (*e.g.*, Chapter 8). With the data and simulations available it was not possible to quantify what effect this would then have on the other metrics though there is the possibility that prescribing the QBO in a model will artificially enhance the score for some of these metrics. Consequently there is a risk that the confidence that is placed in the future projections made with these models will also be overestimated since the observational constraints can not be included in simulations for the future, for obvious reasons.

A broad conclusion which can be drawn from the metric portrait is that all models have deficient performance in some metrics. No model produces an excellent simulation of stratospheric dynamics in all metrics. However, it is also clear that both the performance of the multi-model ensemble in some metrics is better than in others and that the overall performance of some models in these metrics is better than others (note that some caution should be exercised here since some of the differences between metrics can arise from the different diagnostics used to produce them). Particularly poor performance is seen across models in metrics for the SH temperature bias, the tropical annual cycle and diagnostics of the final warming date in both the NH and SH. One obvious point to note is the difference in model performance in the NH and SH. In general, metrics for the SH circulation (shtemp, ummax_jja, firstevl_sh, PW_sh and final_sh — see Tables 4.1a and b for a description of the metrics) have lower values and more outliers than those for the NH circulation. In general, metrics of tropical variability highlight the significant difference in performance related to the different choices for model design described in Section 4.4.2 (see also Chapter 2). Improvement of the simulation of tropical dynamics, particularly the QBO, remains a pressing need for the models.

A more detailed examination of the performance of the models in terms of the metrics in Tables 4.1a and b clearly indicates that, when considering the multi-model ensemble performance as a whole, the assessment of poor metric performance can be further refined. For some metrics in which model performance is generally poor, model biases tend to have the same sign indicating a systematic difference between the models and the observations.

For the metrics considered here,

- metrics with a systematic negative bias are those for the SH temperature, tropical annual cycle and the QBO (excluding models with a nudged QBO, although some caution is necessary for the tropical diagnostics),
- metrics with a systematic positive bias are those for upwelling at 10 hPa, the final warming date in the NH, the final warming date in the SH, the amplitude of the first EOF in the SH and the slope of the fit between lower stratospheric heat flux and lower stratospheric temperature in the NH.

For other metrics, there are large numbers of models with significant biases, but these tend to be evenly distributed between positive and negative signs and hence while indicating poor performance for individual models, they do not indicate systematic biases amongst the multi-model ensemble.

Additionally, it is also possible to examine the statistical distribution of models within individual metrics to determine those models which perform significantly differently from the multi-model mean. In general, this tends to indicate particularly poor performance of a model for that metric (although this conclusion assumes that the multi-model mean performance is good). There are some models which are outliers in significantly more diagnostics than others, *e.g.*, ULAQ (5; 35% of the model's submitted metrics), CAM3.5 (3; 33%), CNRM-ACM (5; 33%).

One concern about the assignment of metrics, also noted by Waugh and Eyring (2008), is that they may not be independent measures of dynamical performance. Following Waugh and Eyring (2008), the correlation between metrics was examined as a simple measure of this non-independence. Of the 120 possible correlations between the 16 metrics considered, only 7 had correlations above 0.5, suggesting that there is not a large degree of dependence between these metrics. The highest correlation between metrics was 0.7, between metrics for the amplitude of the first EOF in the SH and the final warming date in the SH. The amplitude of the first EOF in the SH also had large correlations with several other metrics, including the slope of the fit between lower stratospheric heat flux and lower stratospheric temperature in the SH and strength of the winter jet in the SH.

Due to the large diversity in the formulation of the models (see Chapter 2) it is difficult to systematically assess what role resolution by itself plays in the qualitative and quantitative performance of the models. Nonetheless, a brief subjective assessment of the results presented in the chapter suggests that no systematic improvement in the ability to represent the stratospheric dynamical processes and/or phenomena was obtained with finer horizontal or vertical resolution, or the position of the model top.

However, it is clear that some of the models with very coarse horizontal resolution, or a rather low upper boundary in the middle stratosphere, have worse performance than the multi-model average. For models with upper boundaries above the stratopause and horizontal resolution greater than a moderate threshold (which might tentatively be set at 4° or a spectral truncation at T30) there is no obvious link between increased model resolution and dynamical performance in stratosphere. This suggests that beyond this threshold, the suitability of other model components dominates model performance, though some of these other components, such tropospheric dynamics or tracer transport, are highly likely to be resolution dependent.

4.5.4 Future projections

For many of the dynamical quantities and processes assessed in this chapter, past and future trends were calculated for the periods 1980-1999, 2000-2049 and 2050-2099. Using these three periods gave some indication of the different roles of ozone depletion/recovery and GHG induced climate change on the long-term secular changes in stratospheric dynamics. In general, and as expected, the signal of ozone depletion/recovery on the trends in dynamical quantities is broadly stronger in the SH than in the NH, particularly over the Antarctic. This is reflected both in the strength and significance of the trends in individual models and in the consistency of the trends across the multi-model ensemble. Particularly strong trends in the periods 1980-1999 and 2000-2049 were found in diagnostics of stratospheric polar temperatures, and final warming dates. In both these cases, opposite trends were found in the ozone depletion and recovery periods. Strong positive trends in Antarctic accumulated PSC area diagnostics were found during the 1980-1999 and 2050-2099 periods, consistent with the above trends in temperature, but trends during the 2000-2049 period were small, consistent with the changing influence of ozone concentrations on lower stratospheric temperatures. In the NH both past and predicted trends become more uncertain as the region of interest becomes smaller or the diagnostic more complex. Nonetheless, in the lower and middle Arctic stratosphere the models, on average, projected no significant long-term change to the mean winter-time temperature in contrast to the predicted annual global mean cooling at these levels shown in the previous chapter.

It is also clear that there is a strong consensus amongst the models that the strength of the Brewer-Dobson circulation is simulated to have increased over the recent past and will continue to increase in strength by about 2% per decade over the 21st century. However, there is little consensus amongst the models about the contribution of different types of tropospheric wave-forcing toward this trend. Similarly there was no significant multi-model trend in the extra-tropical mid-winter meridional heat flux at 100 hPa or in the amplitude of the stationary wave field in the upper troposphere. In contrast, a recent analysis of a few observations from the NH mid-latitude lower stratosphere (32-51°N and 24-35 km) shows a weak increase in estimates of the age-of-air (Engel *et al*., 2009) implying a deceleration of the Brewer-Dobson circulation, though the large uncertainties in this observational trend estimate mean it is not inconsistent with the model trends (Engel *et al*., 2009). Differences between the Engel *et al*. (2009) results and the models are considered in more detail in the following chapter.

Acknowledgements

The chapter authors thank John Austin, Mark Baldwin, Natalia Calvo, Ed Gerber, Gloria Manney, Elisa Manzini, Katja Matthes, Charles McLandress, Jessica Neu, Dieter Peters, Adam Scaife, John Scinocca, Ted Shepherd, and Susan Strahan for helpful discussions. Also rather helpful were the written reviews from Martin Dameris, Marv Geller, Nathan Gillett, Peter Haynes, John Thuburn, nine anonymous reviewers and various model PIs. Bill Randel and Fei Wu very kindly provided Transformed Eulerian Mean diagnostics derived from ERA-40 data.

References

Andrews, A. G., J. R. Holton, and C. B. Leovy, 1987. Middle Atmosphere Dynamics, Academic, 489pp.

Austin, J., and F. Li, 2006. On the relationship between the strength of the Brewer-Dobson circulation and the age of stratospheric air, *Geophys. Res. Lett*., **33**, doi:10.1029/2006GL026867.

Austin, J., and R. J. Wilson, 2010. Polar ozone: sensitivity to sea surface temperature and bromine amounts, *J. Geophys. Res*., in press.

Austin, J., D. Shindell, S. R. Beagley, C. Brühl, M. Dameris, E. Manzini, T. Nagashima, P. Newman, S. Pawson, G. Pitari, E. Rozanov, C. Schnadt, and T. G. Shepherd, 2003. Uncertainties and assessments of chemistry-climate models of the stratosphere, *Atmos. Chem. Phys*., **3**, 1-27.

Black, R. X., and B. A. McDaniel, 2009. Submonthly polar vortex variability and stratosphere-troposphere coupling in the Arctic, *J. Clim*., **22**, 5886-5901.

Black, R. X., and B. A. McDaniel, 2007a. The dynamics of Northern Hemisphere stratospheric final warming events, *J. Atmos. Sci*., **64**, 2932-2946.

Black, R. X., and B. A. McDaniel, 2007b. Interannual variability in the Southern Hemisphere circulation organized by stratospheric final warming events, *J. Atmos. Sci.*, **64**, 2968-2974.

Black, R. X., B. A. McDaniel, and W. A. Robinson, 2006. Stratosphere-troposphere coupling during spring onset, *J. Clim.*, **19**, 4891-4901.

Bushell, A. C., N. Butchart, S. C. Hardiman, T. J. Hinton, D. R. Jackson, S. M. Osprey, and L. J. Gray, 2010. Sensitivity of GCM tropical middle atmosphere variability and climate to ozone and parameterized gravity wave drag changes, *J. Geophys. Res.*, in press.

Butchart, N., and A. A. Scaife, 2001. Removal of chlorofluorocarbons by increased mass exchange between the stratosphere and troposphere in a changing climate, *Nature*, **410**, 799-802.

Butchart, N., A. A. Scaife, M. Bourqui, J. de Grandpré, S. H. E. Hare, J. Kettleborough, U. Langematz, E. Manzini, F. Sassi, K. Shibata, D. Shindell, and M. Sigmond, 2006. Simulations of anthropogenic change in the strength of the Brewer-Dobson circulation. *Climate Dyn.*, **27**, 727-741.

Butchart, N., I. Cionni, V. Eyring, T. G. Shepherd, D. W.Waugh, H. Akiyoshi, J. Austin, C. Brühl, M. P. Chipperfield, E. Cordero, M. Dameris, R. Deckert, S. Dhomse, S. M. Frith, R. R. Garcia, A. Gettelman, M. A. Giorgetta, D. E. Kinnison, F. Li, E. Mancini, C. McLandress, S. Pawson, G. Pitari, D. A. Plummer, E. Rozanov, F. Sassi, J. F. Scinocca, K. Shibata, B. Steil, and W. Tian, 2010. Chemistry-climatemodel simulations of the 21st century stratospheric climate and circulation changes, *J. Clim.*, in press.

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

Charlton, A. J., L. M. Polvani, J. Perlwitz, F. Sassi, E. Manzini, K. Shibata, S. Pawson, J. E. Nielsen, and D. Rind, 2007. A new look at stratospheric sudden warmings. Part II. Evaluation of numerical model simulations, *J. Clim.*, **20**, 470-488.

Charlton-Perez, A. J., L. M. Polvani, J. Austin, and F. Li, 2008. The frequency and dynamics of stratospheric sudden warmings in the 21st century, *J. Geophys. Res.*, **113**, doi:10.1029/2007JD009571.

Crook, J. A., N. P. Gillett, and S. P. E. Keeley, 2008. Sensitivity of Southern Hemisphere climate to zonal asymmetry in ozone, *Geophys. Res. Lett.*, **35**, doi:10.1029/2007GL032698.

Engel, A. T. Möbius, H. Bönisch, U. Schmidt, R. Heinz, I. Levin, E. Atlas, S. Aoki, T. Nakazawa, S. Sugawara, F. Moore, D. Hurst, J. Elkins, S. Schauffler, A. Andrews, and K. Boering, 2009. Age of stratospheric air unchanged within uncertainties over the past 30 years, *Nature Geosci.*, **2**, 28-31.

Eyring, V., N. Butchart, D. W. Waugh, H. Akiyoshi, J. Austin, S. Bekki, G. E. Bodeker, B. A. Boville, C. Brühl, M. P. Chipperfield, E. Cordero, M. Dameris, M. Deushi, V. E. Fioletov, S. M. Frith, R. R. Garcia, A. Gettelman, M. A. Giorgetta, V. Grewe, L. Jourdain, D. E. Kinnison, E. Mancini, E. Manzini, M. Marchand, D. R. Marsh, T. Nagashima, P. A. Newman, J. E. Nielsen, S. Pawson, G. Pitari, D. A. Plummer, E. Rozanov, M. Schraner, T. G. Shepherd, K. Shibata, R. S. Stolarski, H. Struthers, W. Tian, and M. Yoshiki, 2006. Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past, *J. Geophys. Res.*, **111**, doi:10.1029/2006JD007327.

Eyring, V., D. W. Waugh, G. E. Bodeker, E. Cordero, H. Akiyoshi, J. Austin, S. R. Beagley, B. A. Boville, P. Braesicke, C. Brühl, N. Butchart, M. P. Chipperfield, M. Dameris, R. Deckert, M. Deushi, S. M. Frith, R. R. Garcia, A. Gettelman, M. A. Giorgetta, D. E. Kinnison, E. Mancini, E. Manzini, D. R. Marsh, S. Matthes, T. Nagashima, P. A. Newman, J. E. Nielsen, S. Pawson, G. Pitari, D. A. Plummer, E. Rozanov, M. Schraner, J. F. Scinocca, K. Semeniuk, T. G. Shepherd, K. Shibata, B. Steil, R. S. Stolarski, W. Tian, and M. Yoshiki, 2007. Multi-model projections of ozone recovery in the 21st century, *J. Geophys. Res.*, **112**, doi:10.1029/2006JD008332.

Feser, F., H.-F. Graf, and J. Perlwitz, 2000. Secular variability of the coupled tropospheric and stratospheric circulation in the GCM ECHAM 3/LSG, *Theor. Appl. Climatol.*, **65**, 1-15.

Fogt, R. L., J. Perlwitz, S. Pawson, and M. A. Olsen, 2009. Intra-annual relationships between polar ozone and the SAM, *Geophys. Res. Lett.*, **36**, doi:10.1029/2008GL036627.

Hardiman, S. C., D. G. Andrews, A. A. White, N. Butchart, and I. Edmond, 2010a. Using different formulations of the transformed Eulerian-mean equations and Eli-

assen-Palm diagnostics in general circulation models, *J. Atmos. Sci.*, in press.

Hardiman, S. C., N. Butchart, S. M. Osprey, L. J. Gray, A. C. Bushell, and T. J. Hinton, 2010b. The climatology of the middle atmosphere in a vertically extended version of the Met Office's climate model. Part I: Mean state, *J. Atmos. Sci.*, **67**, 1509-1525.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, 1991. On the "downward control" of the extratropical diabatic circulation by eddy-induced mean zonal forces, *J. Atmos. Sci.*, **48**, 651-678.

Holton, J. R., and H.-C. Tan, 1980. The influence of the equatorial quasi-biennial oscillation on the global circulation at 50 mb, *J. Atmos. Sci.*, **37**, 2200-2208.

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

Li, F., J. Austin, and J. Wilson, 2008. The strength of the Brewer-Dobson circulation in a changing climate: Coupled chemistry-climate model simulations, *J. Clim.*, **21**, 40-57.

Lu, H., M. P. Baldwin, L. J. Gray, and M. J. Jarvis, 2008. Decadal-scale changes in the effect of the QBO on the northern stratospheric polar vortex, *J. Geophys. Res.*, **113**, doi:10.1029/2007JD009647.

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

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

Manney, G. L., D. R. Allen, K. Krüger, J. L. Sabutis, S. Pawson, R. Swinbank, C. E. Randall, A. J. Simmons, and C. Long, 2005b. Diagnostic comparison of meteorological analyses during the 2002 Antarctic winter, *Mon. Wea. Rev.*, **133**, 1261-1278.

McLandress, C., and T. G. Shepherd, 2009. Impact of climate change on stratospheric sudden warmings as simulated by the Canadian middle atmosphere model, *J. Clim.*, **22**, 5449-5463.

Neu, J. L., and R. A. Plumb, 1999. Age of air in a "leaky pipe" model of stratospheric transport, *J. Geophys. Res.*, **104**, 19,243-19,255.

Newman, P. A., E. R. Nash, and J. E. Rosenfield, 2001. What controls the temperature of the Arctic stratosphere during spring? *J. Geophys. Res.*, **106**, 19,999-20,010.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982. Sampling errors in the estimation of empirical orthogonal functions, *Mon. Wea. Rev.*, **110**, 699-706.

Osprey, S., L. Gray, A. Bushell, N. Butchart, S. Hardiman, and T. Hinton, 2010. The climatology of the middle atmosphere in a vertically extended version of the Met Office's climate model. Part II: Variability, *J. Atmos. Sci.*, in press.

Pascoe, C. L., L. J. Gray, S. A. Crooks, M. N. Juckes and M. P. Baldwin, 2005. The quasibiennial oscillation: Analysis using ERA-40 data, *J. Geophys. Res.*, **110**, doi:10.1029/2004JD004941.

Pawson, S., K. Krüger, R. Swinbank, M. Bailey, and A. O'Neill, 1999. Inetercomparison of two stratospheric analyses: temperatures relevant polar stratospheric cloud formation, *J. Geophys. Res.*, **104**, 2041-2050.

Randel, W., P. Udelhofen, E. Fleming, M. Gelman, K. Hamilton, D. Karoly, D. Ortland, R. Swinbank, M.-L. Chenin, P. Keckhut, K. Labitzke, E. Remsberg, A. Simmons, D. Wu, M. Geller, S. Pawson, F. Wu, and M. Baldwin, 2004. The SPARC Intercomparison of Middle-Atmosphere Climatologies, *J. Clim.*, **17**, 986-1003.

Scaife, A. A., N. Butchart, C. D. Warner, and R. Swinbank, 2002. Impact of a spectral gravity wave parametrization on the stratosphere in the Met Office Unified Model, *J. Atmos. Sci.*, **59**, 1473-1489.

Shepherd, T. G., R. A. Plumb, and S. C. Wofsy (Eds.), 2005. *J. Atmos. Sci.*, **62**, Issue #3.

SPARC, 2002. SPARC Intercomparison of Middle Atmosphere Climatologies, W. Randel, M.-L. Chanin and C. Michaut (Eds.), SPARC Report No. 3, WCRP-116, WMO/TD-No. 1142, 96pp.

Swinbank, R., and A. O'Neill, 1994. A stratosphere-tro-

posphere data assimilation system, *Mon. Wea. Rev.*, **122**, 686-702.

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

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

Waugh, D. W., and V. Eyring, 2008. Quantitative performance metrics for stratosphere-resolving chemistry-climate models, *Atmos. Chem. Phys.*, **8**, 5699-5713.

Waugh, D. W., and P. P. Rong, 2002. Interannual variability in the decay of lower stratospheric Arctic vortices, *J. Meteor. Soc. Japan*, **80**, 997-1012.

Waugh, D. W., W. J. Randel, S. Pawson, P. A. Newman, and E. R. Nash, 1999. Persistence of lower stratospheric polar vortices, *J. Geophys. Res.*, **104**, 27,191-27,201.