Influence of the upper stratosphere on polar vortex variability



Matthew Brown

Oriel College

University of Oxford

This dissertation is submitted for the degree of

Master of Science by Research

Hilary Term 2019

Abstract: The polar vortex is a band of strong westerly winds that forms each winter and encircles the pole, high up in the atmosphere in a region known as the stratosphere. However this band of winds is not stable in the northern hemisphere, and often dramatically disintegrates in events known as sudden stratospheric warmings (SSWs). When SSWs happen they influence the weather occurring below, often leading to cold winters for northern Europe. SSWs are driven by the vertical propagation of waves from lower down in the atmosphere. Recent studies have suggested that the winds in the upper tropical and subtropical stratosphere may affect how these waves propagate, and whether they are able to influence the development of SSWs.

In this thesis I investigate the various contributions different parts of the atmosphere make to the evolution of the SSW in the winter of 2008/9 by fixing different areas of the atmosphere to reality in a computer model, whilst letting the polar vortex evolve freely and comparing its evolution to reality. Fixing the lower atmosphere in this way prescribes the wave forcing entering the stratosphere, but this is not found to be enough on its own to generate an SSW close to that seen in 2008/9. Neither is fixing the upper stratosphere, above the polar vortex. It is only when the winds in the upper tropical stratosphere are fixed in addition to that of the lower atmosphere that a similar SSW occurs, consistent with the idea that this area of the atmosphere is important for determining the evolution of SSWs by affecting wave propagation.

Statement of authorship

I confirm that the work presented here is entirely my own. I carried out the analysis and produced the figures, except where clearly indicated in the figure caption. I also confirm that none of this work is being considered for any other degree or qualification.

Acknowledgements

Well, here we are. The last bit of the thesis that needs to be written, and it's a tricky one! It will always be impossible to do justice in words to the many people who have helped me through the past three years, and for the most part I like to think that it doesn't need to be said here – you all know already. Nonetheless, I'll feel bad if I don't at least make an effort, so, in no particular order, thank you to:

The office of 113, for not working too hard and having so many great conversations, about work or more usually, otherwise. Specifically: Verena – for many invaluable chats about love and life over a bottle of wine, Stergios – for being my companion through many late nights in the office, and Helena – for encouraging me to escape into fun outreach activities, chats about theme parks and volcanoes, and that amazing day where we had an office clearout by throwing old posters out the 1st floor window.

The AOPP Friday Pub Crew, not limited to: KarSteN, Jan, Fenwick, Al and Sam for many deep, entertaining and frequently bizarre conversations that I will never regret. Fenwick and Al in particular for encouraging me to do right by me when it came to leaving academia.

Mum, Dad, for supporting me in all my decisions and just, well, being there!

Calvin Smith, AKA 'gay dad extraordinaire', for helping process and normalise this whole 'being gay' and whole 'academia' thing at the same time.

My Oxford supervisor, Lesley Gray for her unerring patience and understanding, particularly when things were not going well. I would not be writing this if it was not for you.

William Ingram for help running the UM and Jeff Knight and Martin Andrews for the same, and a sense of humour that kept me sane whilst writing this thesis.

The DTP Fam, not limited to: Izzy - for so many good ideas, encouraging me to write more, and helping me visit every Oxford pub… which I now need to write about, Frey - for inspiration to explore the world outside academia, introducing me to the most amazing homemade bread and allowing me to therapeutically shake some rocks when all was going wrong.

The many bands who have produced the music that has kept me sane and inspired me to change over the past three years.

Trevor, for never making me feel like I eat too much pasta and sharing our experiences in how to (or not) get through this whole PhD thing.

Wei-Ling and Charlena – for turning David Patterson House into our own little community.

All of you, and so many more that listing you all would take me well over a page, have helped me complete this journey and close this chapter of my life in a way that worked for me, and been responsible for some of the best days I have ever had. Thank you.

Contents

Section 1: Introduction 8

Section 2: Background 10

2.1: Climatology and seasonal cycle 11

2.2: Wave-driven dynamics 14

2.3: The Quasi-Biennial and Semi-Annual Oscillations 25

2.4: Sudden Stratospheric Warmings and connections to the equatorial stratosphere 29

Section 3: Data, Models and Techniques 36

3.1: Reanalyses 36

3.2: The model 41

3.2.1 Ensemble Setup 45

3.3: Nudging 46

3.4: Experiment Setup 47

Section 4: Results 50

4.1: Introduction to the winter of 2008/9 in the stratosphere 50

4.2: Experiments 52

4.2.1: Control experiment 52

4.2.2: Tropospheric Sensitivity Experiments 53

4.2.3: Stratospheric Sensitivity Experiments 57

4.2.4: Tropical Stratosphere Sensitivity Experiments 63

4.2.5: Analysis of Tropical Stratosphere Sensitivity Experiments 68

Section 5: Summary and Conclusions 75

5.1: Further Work 77

References 80

Section 1: Introduction

The atmosphere is an extremely complicated system, with many links and feedbacks between different regions and processes. In this study I focus on one specific region of the atmosphere, the polar winter stratosphere of the Northern Hemisphere, which consists largely of a strong but frequently disturbed jet of westerly winds encircling the pole known as the polar vortex.

Historically the stratosphere has not been thought to exert a significant influence on the troposphere below, where the weather happens, and thus has been an under-studied region of the atmosphere relative to the troposphere for a long time, due also to the difficulty of obtaining observations this high up in the atmosphere. This started to change with observations from satellites in the 1980s allowing more comprehensive analyses than from observations taken by weather balloons and rocketsondes. After a while it became clear that whilst the stratosphere does respond strongly to forcing from the troposphere the relationship is not all one-way, with the polar winter stratosphere also able to influence the troposphere on monthly to seasonal timescales. This is particularly the case when the polar vortex breaks up dramatically in events known as 'Sudden Stratospheric Warmings'. Although these events themselves are largely driven by waves that propagate vertically from the troposphere, the propagation and momentum deposition of the waves is also influenced by the background wind conditions in the stratosphere, including the winds over the equator and subtropics.

I present a case study of the 2008/9 winter which featured a particularly strong sudden stratospheric warming (SSW) in which the polar vortex broke up completely and split into two

separate parts. I 'fix' different regions of a modelled atmosphere to their observed ('correct') state whilst allowing the polar vortex region to evolve freely, to demonstrate the effects that these different regions can have on the vortex's evolution through the winter. I show that fixing the tropospheric wave forcing is not enough to reproduce the observed warming, and that this warming can be triggered from above by fixing only the upper stratosphere. I also show that getting the upper equatorial stratosphere correct is a crucial factor in accurately simulating the polar vortex when combined with the forcing from the troposphere. This area is currently under-developed in computer models of the atmosphere due to the lack of consistent observations in this region and, until recent years, lack of demonstrated importance for seasonal climate prediction.

A comprehensive introduction to the dynamics and flow regimes of the stratosphere is given in section 2 along with the history of the link between the equatorial stratosphere and the polar vortex. I then describe the datasets, models and techniques used in the experiments in section 3 and present results and analysis in section 4. Section 5 gives conclusions and the directions further work could take based on the results to finish.

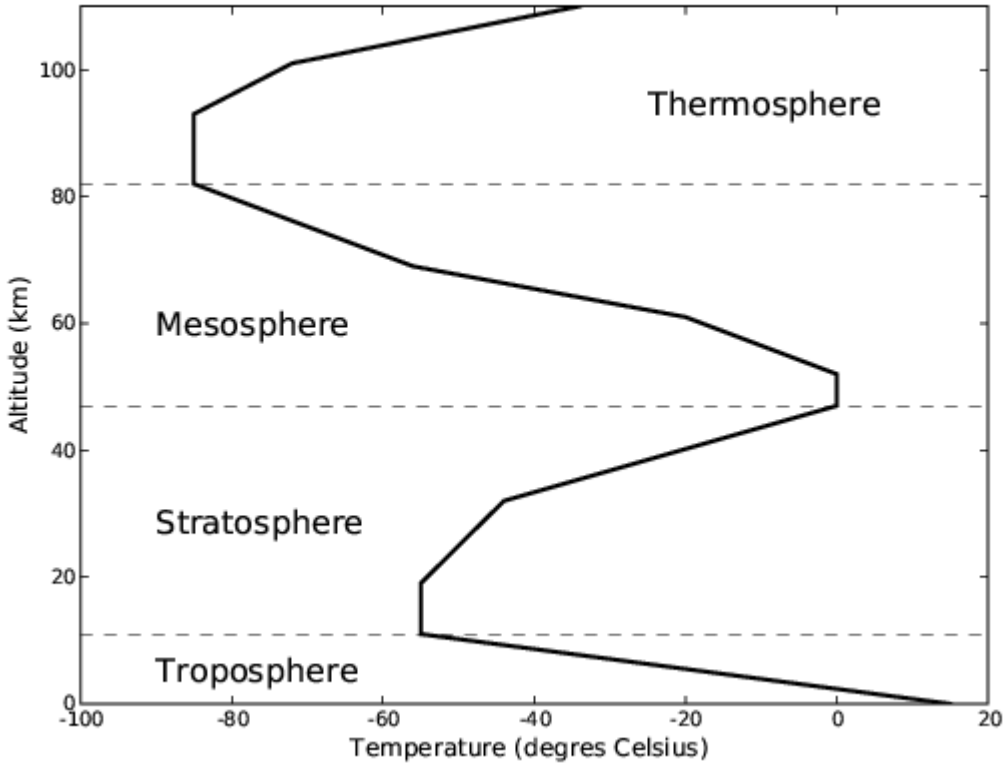


Figure 1: The approximate vertical temperature structure of the atmosphere, and the corresponding layers. Figure from https://commons.wikimedia.org/wiki/File:Profil_temperature_atmosphere.png

Section 2: Background

My research this past year has primarily focused on the stratosphere and to a lesser extent the lower mesosphere, collectively known as the 'middle atmosphere'. The atmosphere is divided into sections based on the vertical temperature structure shown in Figure 1. The troposphere is the lowest region of the atmosphere where the global average temperature decreases with height and the majority of what is colloquially thought of as 'weather' occurs. The depth of the troposphere varies with latitude from an average of 17 km at the equator to 9 km over the poles (Hoinka 1999). Above this, up to about 50 km, lies the stratosphere

where the temperature increases with height instead, with a shallow region of approximately constant temperature known as the tropopause separating the two. Above the stratosphere lies the mesosphere, where the temperature once again begins to decrease with height. Whereas the troposphere is primarily heated from below, the stratosphere is heated by high concentrations of ozone which form what is colloquially known as the ozone layer. The ozone in the stratosphere absorbs the ultra-violet part of the radiation that comes from the sun - heating up this region of the atmosphere and reversing the vertical temperature gradient of the troposphere (Brasseur and Soloman 2005). This source of heating is known as diabatic heating – heating from a source external to the body of interest.

2.1: Climatology and seasonal cycle

Although the stratosphere does not have any 'weather' in the colloquial sense, it nonetheless exhibits a remarkable variety of dynamical behaviour on a range of spatial and temporal scales, some of which can influence the weather and climate in the troposphere (e.g. Baldwin and Dunkerton 2001). Some of the flow observed in the stratosphere can be explained by the dynamic response of the atmosphere to the heating from ozone UV absorption. Just like the troposphere, the stratosphere has a seasonal cycle, driven in part by the shift in the location of the maximum daily-average solar heating between the Northern and Southern Hemispheres. Whilst this heating, and thus the temperature, is fairly constant over the equator, it varies significantly in the polar regions from summer to winter. The pole in the summer hemisphere experiences constant daylight, and therefore constant heating in the stratosphere from ozone absorption of UV. The opposite is true in the winter polar

stratosphere. This sets up a latitudinal temperature gradient between the poles that reverses approximately every 6 months (Haynes 2005 and references therein).

As warmer air is less dense than colder air, the pressure of the air above a particular height within cold air is lower than that at the same height in warm air. This pressure gradient between the warm and cold air increases with height. The relationship governing how the pressure varies with height is hydrostatic balance, where the rate of change of pressure due to the weight of the air above (P) with height (z), $\partial P/\partial z$ is related to the density of the air (ρ): $\frac{\partial P}{\partial z} = -\rho g$. Here g is the force a kilogram of air feels due to Earth's gravity. The pressure gradient between warm and cold air as a result of this relationship causes the air to flow from higher to lower pressure, from the warm air to the cold air (so poleward in the winter hemisphere), increasingly strongly with height, according to the pressure gradient acceleration equation: $\frac{\partial v}{\partial t} = -\frac{1}{\rho}\frac{\partial P}{\partial y}$ Here $\partial v/\partial t$ represents the change of v (∂v), the north-south or meridional wind, at a particular point in space, with respect to time (∂t) - in other words, the acceleration of the meridional wind. Similarly, $\partial P/\partial y$ is the pressure gradient in the same direction (i.e. how much the pressure changes in the 'y' or meridional direction).

However in reality we do not see this flow because of the Earth's rotation. This exerts a force (known as the Coriolis force) perpendicular to the direction of the flow with strength proportional to both the speed of the flow itself, and the latitude, via the relationship: $\frac{\partial u}{\partial t} = fv = 2\Omega sin\theta . v$ Ω is the rotation rate of the Earth, θ is latitude and v is the meridional wind velocity. This Coriolis force acts to deflect the wind until it is perpendicular to the pressure gradient, at which point the pressure gradient force balances the Coriolis force, a situation

that is known as geostrophic balance. This balance means that, given the pressure gradient force and the Coriolis parameter, f, the geostrophic zonal (east-west) velocity can be found according to

$$u_g = -\frac{1}{\rho f}\frac{\partial P}{\partial y}$$

Note that here I have only discussed the equations for u, the zonal wind. The balance between pressure-gradient and Coriolis forces can be defined similarly for the meridional wind.

I have shown that because of hydrostatic balance the result of a temperature gradient in the atmosphere is an increasingly strong pressure gradient with height. This can be combined with geostrophic balance to form a relationship known as thermal wind balance, relating the horizontal temperature gradient (e.g. $\partial T/\partial y$) to the change in the geostrophic wind velocity with height (e.g. $\partial u/\partial z$). The equation describing thermal wind balance for the zonal geostrophic wind is:

$$\frac{\partial u_g}{\partial z} = -\frac{g}{fT}\frac{\partial T}{\partial y}$$

where here T is the vertically-averaged temperature in the layer of the atmosphere being considered, g is the acceleration due to gravity and u_g is the geostrophic zonal wind velocity. This a key relationship that governs elements of the circulation in the stratosphere. Eddies and waves propagating up from the troposphere, described in section 2.2, also play a large role in the stratospheric flow (Andrews et al. 1987a).

Figure 2 shows a climatology (long-term average) of the stratosphere in winter and summer. The temperature gradient set up by the cooling of the winter polar stratosphere relative to the heating in the tropics from ozone absorption leads to a strong westerly jet in the atmosphere at around 60 degrees North/South. This jet extends throughout the depth of the stratosphere and is known as the polar vortex. In the summer, the temperature gradient between the pole and equator is reversed, leading to easterly winds dominating throughout this region of the stratosphere. This figure also shows that the Southern Hemisphere polar vortex is stronger than its Northern Hemisphere counterpart, an important distinction that requires the addition of the effects of atmospheric waves to explain. The tropospheric jet streams can also be seen around the tropopause (the boundary between the troposphere and the stratosphere) in Figure 2.

2.2: Wave-driven dynamics

The equatorial stratosphere has 3 predominant modes of variability of which the annual cycle described above is only one. The other modes exist in the tropics and are shown in Figure 3 to be a semi-annual oscillation (i.e. where the winds reverse from westerly to easterly and back every 6 months) in the upper stratosphere and mesosphere and a quasi-biennial oscillation (where the winds cycle approximately every 28-29 months) in the lower stratosphere. To understand these and the differences in the seasonal cycle between hemispheres, we need to go beyond the basic model of thermal wind and geostrophic balance described above and consider the role of atmospheric waves in the stratosphere. This requires the introduction of some more complex mathematical equations.

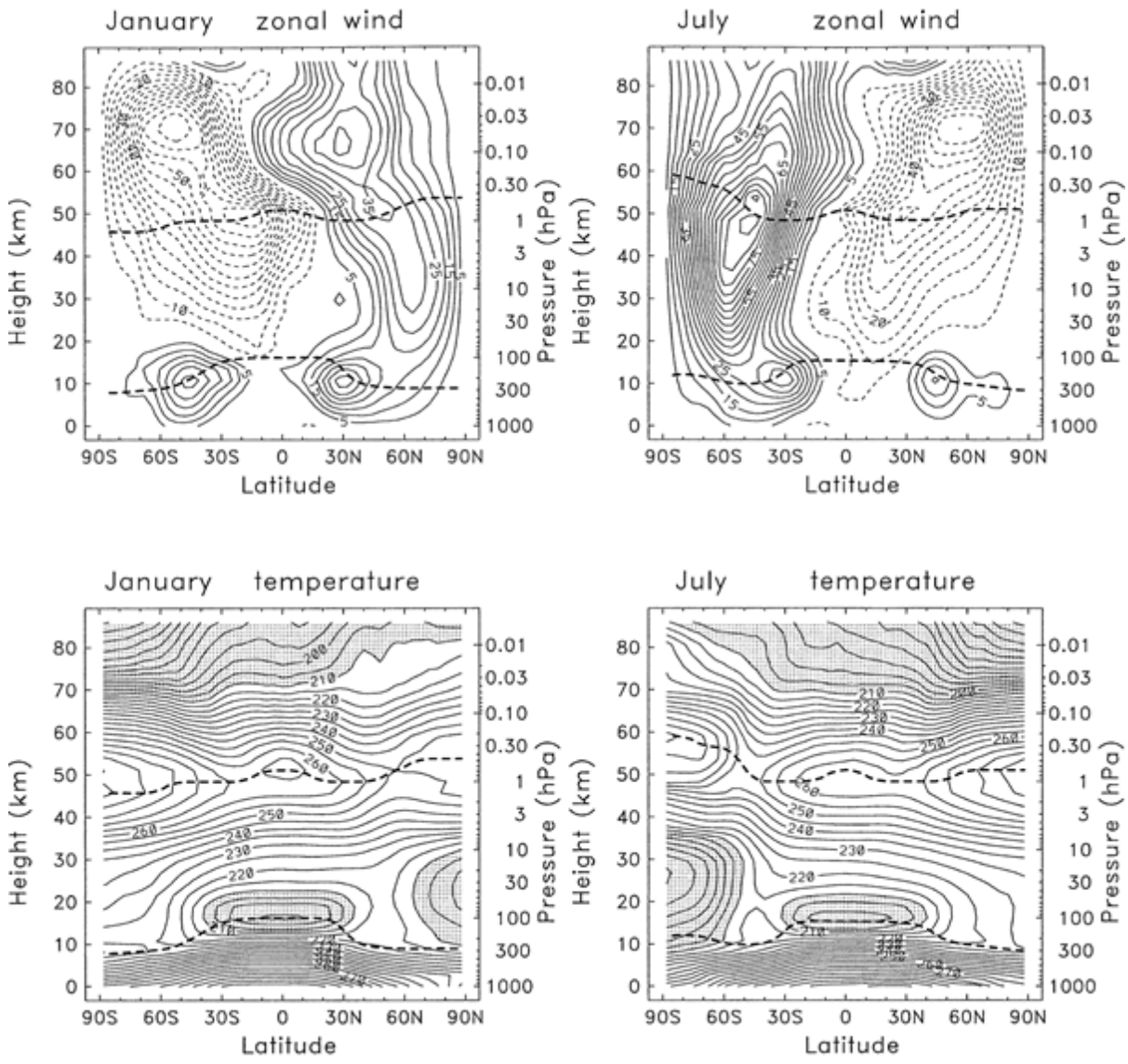


Figure 2: Climatological zonal mean zonal winds (top) and temperatures (bottom) for January (left) and July (right). Zonal winds are from the URAP (UARS (Upper Atmosphere Research Satellite) Reference Atmosphere Program) dataset (contour interval 5ms^{-1}, with zero contours omitted). Temperatures are from Met Office analyses (1000–1.5 hPa), and a combination of HALOE (HALogen Occultation Experiment) plus MLS (Microwave Limb Sounder) data above 1.5 hPa. The heavy dashed lines denote the tropopause (taken from the NCEP-NCAR (National Centers for Environmental Prediction National Center for Atmospheric Research) reanalysis) and stratopause (defined by the local temperature maximum near 50 km). Reproduced from Randel et al. (2004).

The equations that describe the complete flow of the atmosphere are:

$$\frac{\partial \boldsymbol{u}}{\partial t} = -\boldsymbol{u} \cdot \nabla \mathbf{u} - 2\Omega \times \mathbf{u} - \frac{1}{\rho}\nabla P - g\boldsymbol{k} + \nu\nabla^2 \boldsymbol{u} \quad \text{(a)}$$

$$\frac{\partial \rho}{\partial t} = \nabla.(\rho \boldsymbol{u}) \quad \text{(b)}$$

$$P = \rho R T \quad \text{(c)}$$

$$\frac{\partial \theta}{\partial t} = -\boldsymbol{u}.\nabla\theta + \kappa\nabla^2\theta + H \quad \text{(d)}$$

(a) is the equation of motion, where each term on the right hand side represents a force acting to accelerate the flow: $-\boldsymbol{u} \cdot \nabla \boldsymbol{u}$ is the advection (transport) of u, v (the horizontal wind velocities) and w (the vertical wind velocity) by itself, $2\Omega \times \boldsymbol{u}$ is the combined Coriolis and centrifugal forces that arise from the rotation of the Earth, $1/\rho \, \nabla P$ is the pressure gradient force, $g\boldsymbol{k}$ is gravity (the **k** indicates that it acts in the vertical only) and $\nu\nabla^2 \boldsymbol{u}$ is friction. Here bold characters denote vectors, where each component of the vector represents a different direction e.g. $\boldsymbol{u} = (u, v, w)$. (b) represents conservation of mass (i.e. mass cannot be created or destroyed) for a compressible fluid, (c) is the equation of state which relates the temperature, pressure and density of a fluid and (d) represents conservation of heat. Each term on the right hand side of (d) is a source or sink of heat: $-\boldsymbol{u}.\nabla\theta$ is the transport of heat by the wind, $\kappa\nabla^2\theta$ is the source of heat through friction and H represents any external (diabatic) sources of heating. θ, the 'potential temperature' is used instead of temperature in this equation as it accounts for the fact that the temperature of air changes as air rises and falls. These equations as they stand cannot be easily solved due to the large range of spatial scales spanned by the forces described and large non-linearities involved (e.g. the advection

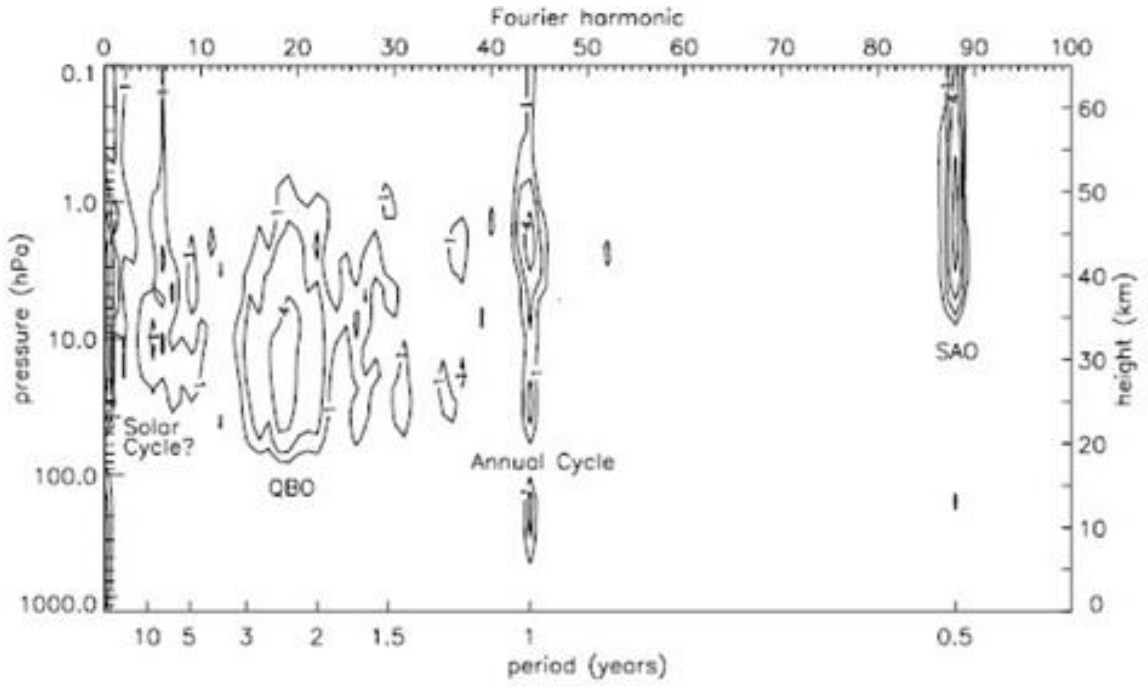


Figure 3: Vertical profile spectra of zonal mean zonal wind at the equator. Contour lines drawn at Fourier amplitudes of 1, 2, 4, 8, 16 m/s. The main modes of variability, namely Semi-Annual Oscillation (SAO), Annual Cycle and Quasi-Biennial-Oscillation (QBO), can clearly be seen at periods of 0.5, 1 and about 2 years. Reproduced from Pascoe et al. (2005).

of the wind by itself). However, for many situations assumptions can be made which dramatically simplify these equations, allowing solutions which explain particular atmospheric situations/phenomena to be obtained.

Solutions that describe the key aspects of some of the wave-driven phenomena in the middle atmosphere can be obtained by making the beta-plane and/or quasi-geostrophic (QG) approximations. The beta-plane approximation replaces the curvature of Earth's surface with a 2D-plane tangent to a point on Earth's surface, allowing horizontal cartesian coordinates (x, y and z) to be used instead of spherical. The main consequence of this assumption is that f, the Coriolis parameter, now varies with y instead of latitude. If we assume that this variation is linear we obtain the equation $f = f_0 + \beta y$ where β is the rate of change of f with y and f_0

is f at the reference latitude, ϕ_0, the latitude at which the tangent plane touches the Earth's surface. The beta-plane assumption is considered valid when $|y| \ll \mathrm{r}\cot\phi_0$ where r is the radius of the Earth.

Quasi-geostrophy can be simply understood as an extension of geostrophic balance allowing for vertical motion. It results from the following assumptions: Rotational effects are significantly larger than non-linear effects, the net acceleration of the flow is small compared to the acceleration by the Coriolis force, magnitude of the variation of f with y is much smaller than f_0 and frictional effects are negligible.

The momentum and mass conservation equations obtained after the above assumptions are:

$$D_g u_g - f_0 v_a - \beta y v_g = 0$$

$$D_g v_g + f_0 u_a + \beta y u_g = 0$$

$$\frac{\partial u_a}{\partial x} + \frac{\partial v_a}{\partial y} + \frac{1}{\rho_0}\left(\frac{\partial}{\partial z}\left(\rho_0 w_a\right)\right) = 0$$

where $D_g = \frac{\partial}{\partial t} + u_g \frac{\partial}{\partial x} + v_g \frac{\partial}{\partial y}$ and is the rate of change following the *geostrophic* winds. The subscript 'a' denotes ageostrophic components of the winds where, for example, $u_a = u - u_g$

As many features of the flow in the stratosphere are either driven or strongly affected by atmospheric waves, it is helpful to decompose the flow into a 'background mean' and 'eddy' or 'wave' flow, so that the effect of one on the other can be more easily studied, in the process obtaining an important criterion for the propagation of waves. To obtain wave

solutions that are supported by the above equations we first have to linearise them about a background state. This is a process where we assume that the waves are small perturbations (denoted by a ‘) from the background winds, and that any forces acting that result from multiples of these perturbations (the non-linear forces) are small enough to be ignored. These equations can then be solved for waves by first assuming a general wave-type solution. An example of such a solution for a 1-dimensional wave in the zonal wind (u) in the x-direction only is $u' = U(\cos(kx + \omega t) + i.\sin(kx + \omega t)) = Ue^{i(kx-\omega t)}$. This equation represents any wave, it is only by substituting it into the linearised equations for the flow we are interested in (such as those shown above) that we can find the unknowns and the specific types of waves that said equations support. In other words we are ‘searching’ for the types of waves that these equations support. The main unknown that we need to solve for is ω, the (angular) frequency of the wave.

If we make our wave equation 3-dimensional before substituting in, a solution that describes Rossby waves capable of vertical propagation can be found, following Charney and Drazin (1961):

$$\omega = k\overline{u} - \beta k\left[k^2 + l^2 + \frac{f_0^2}{N^2}\left(m^2 + \frac{1}{4H^2}\right)\right]^{-1}$$

$\omega = ck$ which implies:

$$\overline{u} - c = \beta\left[k^2 + l^2 + \frac{f_0^2}{N^2}\left(m^2 + \frac{1}{4H^2}\right)\right]^{-1}$$

where β is the horizontal gradient of the Coriolis parameter (the component of Earth's rotation in the local vertical), N describes how buoyant the air is, H is a scale height (approximately 8 km in the atmosphere), c is the phase speed (the propagation speed of the individual waves) and k, l and m are the wavenumbers in the x, y and z directions respectively. The wavenumbers simply describe the number of wavelengths that exist over a unit length.

All waves result from a restoring force acting to restore some equilibrium, which overshoots to the other 'side' of equilibrium, setting up an oscillation – the wave. Rossby waves ultimately result from the variation of the Coriolis parameter with latitude, which is the restoring force acting to reduce the disturbances of air moving north and south. Since β, H, f_0^2 and N^2 are always positive in Earth's atmosphere, and the squared wavenumbers can also only be positive, $u - c$ can never be less than 0. Since m^2 has to be greater than 0 for vertically propagating Rossby waves to exist, we can form a condition in u which must be satisfied for such Rossby waves to exist:

$$0 < \overline{u} - c < u_c = \beta \left(k^2 + l^2 + \frac{f_0^2}{4N^2H^2} \right)^{-1}$$

If we only consider stationary waves where the phase speed is zero relative to the ground ($c = 0$) the above condition can be simplified to:

$$0 < \bar{u} < u_c$$

This criterion for vertically propagating Rossby waves was first proposed by Charney and Drazin (1961). Surfaces on which the background zonal wind is either equal to c or u_c are

known as critical surfaces, through which waves cannot propagate. In this region the linear assumption made in deriving the Charney-Drazin criterion is no longer valid, as non-linear behaviour becomes dominant. When waves reach such critical layers they can 'break' and accelerate the background mean flow. This acceleration takes place just below/before the critical surface and in the case of significant wave activity the surfaces can start to reflect waves instead (Andrews et al. 1987b).

The critical velocity decreases with the zonal and meridional wavenumbers, which means that the 'propagation window' decreases as these wavenumbers increase. Therefore faster background zonal winds can only support waves of smaller wavenumbers/longer wavelengths. As the westerly winds in the stratosphere tend to be stronger than their tropospheric counterpart (e.g. the polar vortex), the stratosphere is dominated by long waves of spherical zonal wavenumbers ~1-3. Additionally, the Charney-Drazin criterion explains the absence of vertically propagating waves from the troposphere in the summer stratosphere, when the winds are easterly.

These waves are responsible for the weaker Northern Hemisphere polar vortex relative to its Southern Hemisphere counterpart (Hirota et al. 1983, Shiotani and Hirota 1985). They are forced primarily by the flow of the tropospheric mid-latitude jet streams over large mountain ranges, and also by large temperature contrasts e.g. in areas of tropical convection and strong land-sea contrasts (Andrews et al. 1987a). The jet stream in the Southern Hemisphere is largely uninterrupted by mountain ranges, whereas the Northern Hemisphere jet encounters the Rockies. The waves generated then propagate vertically from the troposphere into the

stratosphere where they can break in the strong westerly winds of the vortex, weakening it as described further in section 2.4**.**

Rossby waves are only one type of the many waves that are present in the atmosphere. Solutions for other waves are obtained by using a similar process to that outlined above, but with different assumptions made, and different combinations of forcing terms in the governing equations. Planetary-scale waves (such as Rossby waves) are particularly important to the flow in the Northern Hemisphere winter stratosphere where the winds are westerly. In the equatorial stratosphere however, equatorially-trapped waves give rise to two key flow regimes – the quasi-biennial oscillation (QBO) and semi-annual oscillation (SAO). The two dominant equatorial waves here are Kelvin and Rossby-gravity waves (Andrews et al. 1987c). Kelvin waves can only exist in areas where geostrophic balance does not hold, as is the case at the equator where the Coriolis parameter is zero, and have eastward phase velocity. Rossby-gravity waves have features of both internal gravity waves and Rossby waves and typically have westward phase velocity. When the zonal wavenumber k is large (small zonal scales) these reduce to gravity waves, and to Rossby waves when k is small (Holton 2013). Also important are inertial internal gravity waves (Dunkerton 1997) which can have either eastward or westward phase velocity. Equatorial gravity waves are buoyancy oscillations either generated by flow over orography (orographic gravity waves) or from convection (non-orographic gravity waves).

To understand how all of these types of waves influence the background mean flow it is necessary to introduce a quantity known as the Eliassen-Palm (EP) flux, after Eliassen & Palm (1960). To derive the EP-flux, we once again assume a QG-system, this time including an

equation for temperature as well as wind. The following equations are derived by splitting u, v and T into zonal-mean components denoted by the overbar, and eddy components (deviations from the zonal-mean, the waves) denoted by the primed quantities, and then taking the zonal-mean of the whole equation:

$$\frac{\partial \overline{u}}{\partial t} = f\overline{v} - \frac{\partial}{\partial y}\overline{u'v'} + \bar{\dot{u}}$$

$$\frac{\partial \overline{b}}{\partial t} = -N^2\overline{w} - \frac{\partial}{\partial y}\overline{v'b'} + \overline{\dot{b}}$$

where we have used buoyancy, b instead of temperature to simplify the equations and $\bar{\dot{u}}$ and $\overline{\dot{b}}$ represent the dissipative forces acting on the flow (e.g. drag) and diabatic (external) heating respectively. Buoyancy (b) represents the vertical acceleration of the flow due to perturbations in density and hence temperature from that of the surroundings: $b = g\,\frac{T - T_r}{T_r}$ where T_r denotes the temperature of some reference state, usually dependent only on height (horizontal variations are relatively small and thus neglected). $\overline{u'v'}$ and $\overline{v'b'}$ are the eddy fluxes of momentum and heat respectively, and can be understood as the transport of momentum (wind speed) and temperature (and thus buoyancy) by waves.

Using these equations as they stand it is hard to understand how waves will influence the mean flow due to the strong coupling between the terms. Take the first equation. It appears that acceleration of the zonal mean zonal wind ($\partial\bar{u}/\partial t$) is dependent on the zonal mean meridional wind through the Coriolis force, convergence of the eddy momentum flux, and dissipative forces. However, the zonal mean meridional wind is not actually independent of

either the eddy momentum flux or eddy heat flux, coupling the equations and meaning we cannot easily diagnose $\partial\overline{u}/\partial t$ from the eddy/wave fluxes.

To solve this problem, we can rewrite the equations in terms of a residual time-mean meridional circulation ($\overline{v_r}, \overline{w_r}$), which is the time-mean path an air parcel follows:

$$\overline{v_r} = \overline{v} - \frac{1}{\rho_0}\frac{\partial}{\partial z}\left(\rho_0 \frac{\overline{v'\theta'}}{\partial\theta_0/\partial z}\right)$$

$$\overline{w_r} = \overline{w} + \frac{\partial}{\partial y}\left(\frac{\overline{v'b'}}{\partial\theta_0/\partial z}\right)$$

and essentially decouples the equations:

$$\frac{\partial}{\partial t}\overline{u} - f_0\overline{v_r} = \frac{1}{\rho_0}(\nabla \cdot \boldsymbol{F}) + \overline{\dot{\boldsymbol{u}}}$$

$$\frac{\partial}{\partial t}\overline{b} + \overline{w_r}N^2 = \overline{\dot{b}}$$

where $\boldsymbol{F} = -\rho_0\overline{u'v'}\boldsymbol{j} + f\rho_0\frac{\overline{v'\theta'}}{\partial\theta_0/\partial z}\boldsymbol{k}$ is the Eliassen-Palm flux vector. $\boldsymbol{j}$ and $\boldsymbol{k}$ represent the direction that each term acts in: y and z respectively. These equations are known as the Transformed Eulerian Mean (TEM) equations after Andrews and McIntyre (1976).

From this we can more clearly see that the wave fluxes work together to change the mean flow, and that in the absence of dissipative processes represented by $\overline{\dot{\boldsymbol{u}}}$, only *divergence* of these wave fluxes ($\nabla.\boldsymbol{F}$) changes the mean flow. A divergence of wave flux ($\nabla.\boldsymbol{F} > 0$) accelerates the zonal mean zonal wind ($\overline{u}$) and results in a time-mean meridional and vertical circulation ($\overline{v_r}, \overline{w_r}$). Buoyancy, in the absence of diabatic heating, is transported solely

through the vertical time-mean residual circulation, $\overline{w_r}$. Convergence of these wave fluxes results from wave dissipation, such as the wave breaking that occurs around critical surfaces.

Further to this Eliassen and Palm (1960) showed that $\nabla \cdot \boldsymbol{F} = 0$ for steady (unchanging in time), linear waves in purely zonal flow, with no dissipative forces or diabatic heating. Substituting this into the above equations yields the result:

$$\frac{\partial \overline{u}}{\partial t} = \frac{\partial \overline{\theta}}{\partial t} = v_r = w_r = 0$$

i.e. when $\nabla \cdot \boldsymbol{F} = 0$ there is no residual circulation and the zonal mean zonal winds do not change. In reality, the waves do still induce a change in the circulation, but this change precisely cancels out the original effect of the waves. This result is known as the non-circulation theorem and is not at all obvious from the un-transformed equations. A large number of assumptions are made to arrive at this result which therefore holds only approximately in the real atmosphere. Nonetheless the EP-flux is an extremely useful tool, used by many studies of the dynamics of the stratosphere to diagnose the effects of waves on the background flow (e.g. Watson and Gray (2014), Garfinkel (2012), Shiotani and Hirota (1985)), particularly in the case of the Northern Hemisphere polar vortex, which is discussed in section 2.4 and is the focus of this study.

2.3: The Quasi-Biennial and Semi-Annual Oscillations

The QBO consists of downward-propagating regions of westerly and easterly winds with a period of approximately 2 years, varying between 22 and 40 months (Pascoe et al. 2005). The QBO exists throughout the depth of the equatorial stratosphere with a maximum amplitude

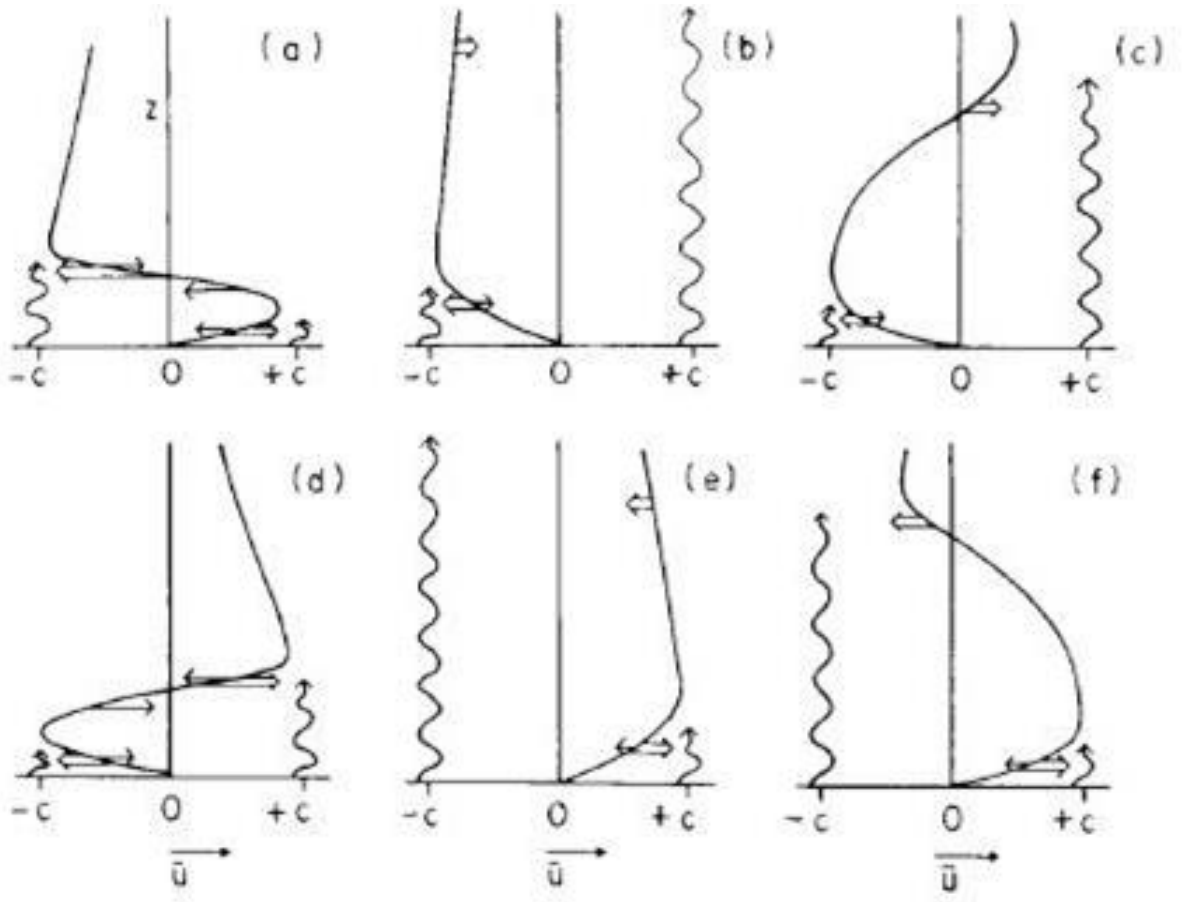


Figure 4: Schematic representation of the interaction of the mean flow vertical profile (solid line) with vertically propagating waves of differing phase speed (wavy lines) by Plumb (1984). Double arrows represent wave-driven, single arrows viscously driven acceleration.

around 15 mb. It is approximately zonally symmetric about the equator with a half width of around 12 degrees (Baldwin et al. 2001). The easterly winds on average descend more slowly than the westerlies. It was first discovered by Reed et al. (1961) and Ebdon (1960), however it was not named as such until Angell and Korshover (1964), who also noted that each phase (QBO-East/QBO-West) was associated with characteristic temperature fluctuations and a mean meridional circulation from the wave activity we now know drives the oscillation.

A gravity wave theory describing the mechanism for the QBO was first proposed by Lindzen and Holton (1968)**.** This was extended to include Kelvin and Rossby-gravity waves by Holton

and Lindzen (1972), and is illustrated in Figure 4. The theory showed that a QBO-like oscillation could be obtained in a simple one-dimensional model from wave-mean flow interactions of vertically propagating Kelvin and Rossby-gravity waves that contribute westerly and easterly force, respectively. Much like the criterion obtained for the propagation of Rossby waves, Kelvin waves can only propagate vertically in easterly winds and Rossby-gravity in westerly. As these waves propagate upwards they are absorbed by the mean flow, below their critical layer, depositing momentum and hence accelerating the mean flow in the (phase-velocity) direction of the wave. Crucially, the maximum acceleration occurs below the maximum in the easterly or westerly wind (Plumb 1984), causing the maxima to descend in time (Figure 4a, b). At low levels, wind layers are damped out by viscous diffusion (Figure 4c). In the meantime, an easterly regime propagates downward (Figure 4d, e), allowing eastward waves to travel upward and build a new westerly zone (Figure 4f)). Though this theory explained the periodicity and downward propagation of the anomalies, the different descent speeds of the two regimes was not accounted for. This difference between the two regimes is a result of the time-mean meridional circulation associated with the wave breaking that drives the QBO (Haynes 1998). The easterly phase is associated with upward motion over the equator, slowing the descent of the winds, with the opposite true of the westerly phase (Kinnersley and Pawson 1996). Note that whilst this theory is analogous to the wave-mean flow interaction theory introduced at the end of section 2.2, that theory was based on a quasi-geostrophic (QG) framework, which is not valid over the equator. The wave breaking (a convergence of wave flux) that drives the QBO *accelerates* the zonal mean flow, whereas in the QG framework wave breaking is associated with a deceleration.

Above 10 mb a different oscillation between westerly and easterly winds becomes dominant – the stratospheric Semi-Annual Oscillation (SAO) (Baldwin et al. 2001). The SAO peaks in the upper stratosphere near 50 km (1 mb) with velocities of about 30 m/s and a regular period of 6 months. It was first discovered in rocketsonde observations by Reed (1965) and has since been confirmed in many rocketsonde analyses (e.g. Naujokat 1986) and satellite retrievals (e.g. Hitchman and Leovy 1985). The SAO is a cycle between easterly flow at the solstices generated by the latitudinal shift in the region of maximum solar heating (through thermal wind balance) and westerly flow generated by wave activity around the equinoxes through the same mechanism as the westerly phase of the QBO (Hitchman and Leovy 1985). As a result of the seasonal driving of the oscillation it is locked to the annual calendar, unlike the QBO, with the westerly and easterly phases occurring at the same time each year. Kelvin waves account for up to 70 % of the wave activity (Hitchman and Conway 1988) with the rest coming from smaller-scale equatorial inertia-gravity waves excited by convective heating (i.e. non-orographic gravity waves) (Garcia and Sassi 1999). This is consistent with Shepherd (2000) who emphasized the fact that GCMs (general circulation models) are generally able to simulate an SAO without a non-orographic gravity wave parameterisation scheme (i.e. without the smaller scale waves). The SAO has also been observed to have an annual cycle component with a reduction in the strength of the peak easterlies at 1 mb during July compared to December. This is related to a seasonal cycle in the wave driving of the oscillation, with stronger driving from planetary waves in northern hemisphere winter (Garcia et al. 1997).

2.4: Sudden Stratospheric Warmings and connections to the equatorial stratosphere

A crucial feature of the winter polar vortex in the Northern Hemisphere is its variability. It is frequently disturbed by the planetary waves that arise from interaction of the tropospheric jet stream with the mountain ranges in the Northern Hemisphere. As such orography is not present in the Southern Hemisphere, the polar vortex in the Southern Hemisphere is far less disturbed (Baldwin et al. 2003).

The variability of the Northern Hemisphere polar vortex manifests largely through events known as Sudden Stratospheric Warmings (SSWs), where the vortex rapidly warms and weakens over the course of a few days. The exact definition of SSWs is somewhat arbitrary, however it is common to use the reversal of either zonal mean temperature or zonal wind gradient at 10 mb between 60 and 90 °N after Andrews et al. (1987a). Major SSWs therefore essentially occur when the polar jet reverses direction completely. Such an event was first observed by Scherhag (1952) who described the events of the winter of 1951/2 as 'explosive warmings', with subsequent studies by Quiroz (1975) and Labitzke (1977) among the first to document their occurrence.

The first study to attempt to explain the observed SSWs was Matsuno (1970), which remarkably is still largely accepted today. They proposed that planetary scale waves propagate upwards into the stratosphere where they break as a result of the decreasing density of the air and deposit easterly momentum. Where this wave breaking happens in the vortex, the deposition of easterly momentum acts to slow the westerly jet. In the extreme cases where the jet reverses, a critical layer for these waves is formed, above which they are

unable to propagate. This leads to a descending region of easterly winds as the waves break underneath the critical layer, lowering it and causing the region of wave breaking to shift downward. Above this critical layer the radiative processes that drive the polar vortex dominate, and the westerly winds re-establish from the top until the vortex completely reforms on a timescale of 2-4weeks (which is approximately the radiative timescale in this part of the atmosphere).

There are two main types of SSW, displacement and split. A displacement SSW occurs when the polar vortex is pushed off the pole by an area of high pressure and easterly winds known as the Aleutian high. A split SSW is when the polar vortex splits into two daughter vortices, with higher pressure and easterly winds again over the pole. Examples are shown in Figure 5. These different types of SSWs are associated with the wave breaking of spherical zonal wavenumber 1 and 2 Rossby waves respectively. The type of SSW has been shown to influence the effect on the tropospheric circulation below, with split SSWs found to be more barotropic with a greater surface influence (Mitchell et al. 2013).

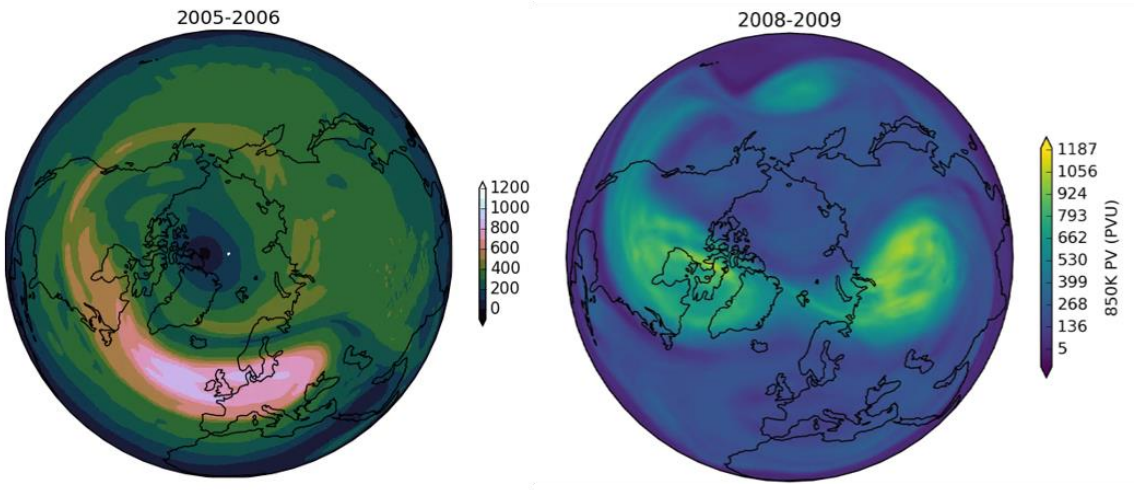


Figure 5: A displacement sudden stratospheric warming event of the northern hemisphere polar vortex from the 2005-6 winter around the 21st Jan 2006 (left) and a split event from the 2008-9 winter around 24th Jan 2009 (right). The left plot was produced by Dr Will Seviour, John Hopkins University. The filled contours are of potential vorticity on the 850K potential temperature surface using ERA-Interim reanalysis data.

Many observational studies have shown there to be a link between the equatorial zonal winds in the low-middle stratosphere and the variability of the polar vortex. Such a link was first observed by Holton and Tan (1980) and Holton and Tan (1982) who found that the vortex was stronger when the QBO was in its west phase in the lower stratosphere (40-50 mb, specifically) (QBO-W) than when it was in its east phase here (QBO-E). This relationship has since been named after them, with many subsequent observational studies confirming this relationship (e.g. Baldwin & Dunkerton 1991**,** Pascoe et al. 2005 and Lu et al. 2008), although there is some evidence that this relationship is disrupted when the 11-year sunspot solar cycle is at a maximum (see e.g. Labitzke 1987, Gray et al. 2004).

Despite the large number of studies confirming the Holton-Tan relationship, there is still some disagreement on the physical mechanism that underlies it. The original mechanism suggested by Holton and Tan (1980) was that the planetary wave critical zero zonal-mean wind line in

the equatorial stratosphere is shifted poleward in QBO-E, with waves propagating up from the troposphere more latitudinally confined and thus on average directed more towards the pole and the vortex. The waves which reach the vortex subsequently break and deposit easterly momentum, weakening it. This mechanism would result in a convergence of EP-flux near the pole in QBO-E conditions (with the opposite in QBO-W) however they were careful to state that this was purely a speculative mechanism.

Subsequent studies have shown mixed support for this mechanism. Hamilton (1998) found that on average over all months there is greater EP-flux in the polar stratosphere in QBO-E, whilst Hu and Tung (2002) and Gray et al. (2004) concluded that this mechanism only applied in early winter. Calvo et al. (2007) found that the enhanced EP-flux in the polar stratosphere was only present in the months with a diagnosed Holton-Tan relationship, and then only in the upper stratosphere. More recently, Garfinkel (2012) found that whilst the mechanism associated with the shift in the zero wind line was present in the Whole Atmosphere Community Climate Model (WACCM), it had less of an impact on the polar vortex than a mechanism involving the changing meridional circulation associated with each QBO phase. Naoe and Shibata (2010) found support for a similar mechanism in observational data, but none for the original Holton-Tan mechanism at all. White et al. (2016) found that whilst the Holton-Tan mechanism was present it was only part of the picture, with many other mechanisms also at play. Watson and Gray (2014) found support for the Holton-Tan mechanism in the Hadley Centre climate model HadGEM2, but again stressed that this was likely part of a larger picture.

The observational study of Gray et al. (2001b) using rocketsonde data was the first to suggest that the upper stratosphere may play a role in driving the Holton-Tan (HT) relationship, with a strong, albeit not statistically significant, correlation between the vortex strength and wind shear in the upper equatorial stratosphere, which was suggested to be associated with the SAO. The HT relationship was also confirmed in the lower stratosphere. A similar result was obtained in the subsequent modelling study of Gray et al. (2001a). In this study a stratosphere-mesosphere only model (SMM) was used where the winds were relaxed to the observations used in Gray et al. (2001b). The study found sensitivity of the polar vortex to the equatorial winds at all levels of the equatorial stratosphere, with the HT relationship only found when the whole equatorial stratosphere was relaxed, not when just the upper stratosphere was, indicating it was the whole equatorial wind profile that was important rather than specific levels. This study was similar to Hamilton (1998) who relaxed up to 40 km (not quite as high as Gray et al. (2001a)) and didn't achieve such a strong HT relationship.

Gray (2003) then did a simplified modelling study with the same SMM to better determine the nature of the influence (in terms of region and timing). The experiment was split into separate upper and lower equatorial stratosphere easterly nudging areas rather than lower and combined lower and upper as in Gray et al. (2001a). Sensitivity of the polar vortex was found to the winds in the lower equatorial stratosphere in early winter only and in the upper equatorial stratosphere throughout winter, but particularly during mid-winter, a time when most sudden stratospheric warmings occur. It was suggested that this may have something to do with the timing of the switchover between the westerly and easterly phases of the SAO, which is influenced by the QBO (Baldwin et al. 2001) and the 11-year solar cycle.

A similar study using the same model by Sparrow et al. (2004), and later by Gray et al. (2003) with confirmation from EP-flux diagnostics, then found that there may be a role for the equatorial and subtropical upper stratosphere on SSWs in early winter as well, through the modification of the development of the Aleutian High, which is a precursor to SSWs. Gray et al. (2004) extended Gray (2003) by investigating the impact of easterly winds in the equatorial and subtropical stratosphere on the timing of the SSWs themselves (as well as considering the impact of different height bands in the stratosphere rather than simply 'upper' and 'lower'). They found that the upper subtropical stratosphere in particular was associated with earlier SSWs in their idealised model experiments.

Pascoe et al. (2006) conducted an investigation using a full GCM (i.e. one including the troposphere as well as the stratosphere and mesosphere), in which they applied momentum forcing in the stratosphere corresponding to a realistic QBO and SAO in their respective regions. This momentum forcing was applied individually in separate experiments and then combined in a further experiment. The model itself did not have a realistic (self-generated) QBO or SAO and when the corresponding momentum forcing was introduced the simulation of the interannual variability in polar stratospheric temperatures was significantly improved. It was also found that the timing of mid-winter SSWs in the model was sensitive to the momentum forcing in the SAO region (the upper stratosphere), with the timing of the warmings shifting forward by a month, confirming the earlier results of the more idealised experiments of Gray et al. (2004). The QBO region (lower stratosphere) was found to influence the early winter polar variability. This series of studies demonstrated the influence of the upper stratosphere on the polar vortex and the need for a realistic representation of

the QBO and SAO in models for realistic simulation of the polar vortex and SSWs. However just like the Holton-Tan relationship, which is more focused on the lower stratosphere, the mechanism behind the influence is not yet clear.

The variability of the polar vortex is influenced by many regions of the atmosphere. Whilst some of these are well understood, e.g. the wave forcing from the troposphere and the barotropic 'top-down' development of disturbances to the vortex, the influence of the upper tropical and subtropical stratosphere is less well understood. It is clear that this region does play some role in polar vortex variability, but the mechanisms underlying this are still open to debate, with little further investigation since Pascoe et al. 2006.

A common theory for the influence of the tropical and subtropical stratosphere is that by changing the winds here waves coming up from the troposphere can be deflected more towards or away from the polar vortex. This can influence the variability of the polar vortex, as it is the breaking of these waves in the vortex that weakens it and sometimes causes SSWs. The ability of GCMs to reproduce this influence is unclear, with studies using different models obtaining varying results. This is in part due to differing QBOs and SAOs in these models.

This study demonstrates the different influences the various regions of the atmosphere have on the polar vortex and how they combine, highlighting the importance of that of the upper tropical and subtropical stratosphere. A preliminary analysis of the demonstrated influence of this region is then conducted.

SSWs have been shown to influence the weather in the troposphere on monthly to seasonal timescales. By better understanding the drivers of variability in the polar vortex and SSWs we can better predict these events and thus their influence on our weather.

The next section details the datasets and techniques used in the study, along with a brief assessment of the representation of the polar vortex and the upper tropical and subtropical stratosphere in the GCM used.

Section 3: Data, Models and Techniques

In this section I describe the experimental setup along with the data, models and techniques used.

3.1: Reanalyses

The ensemble study makes use of two reanalysis datasets, ERA-I (a Re-Analysis produced by the European centre for medium-range weather forecasting, (Dee et al. 2011)) and MERRA2 (Modern Era Retrospective-analysis for Research and Applications version 2, produced by NASA, (Gelaro et al. 2017)). A reanalysis dataset is, put simply, a 'best guess' at the state of the atmosphere at a given time. It is a combination of observational data from satellites, radiosondes and observation stations (among others) and a numerical model of the atmosphere. The model is used to 'fill in the gaps' in the observations, and the observations are used to constrain the model and keep it in line with reality. Different reanalyses use different methods to merge the observational and model data (a process known as data assimilation), as well as using different models and different observational data. This

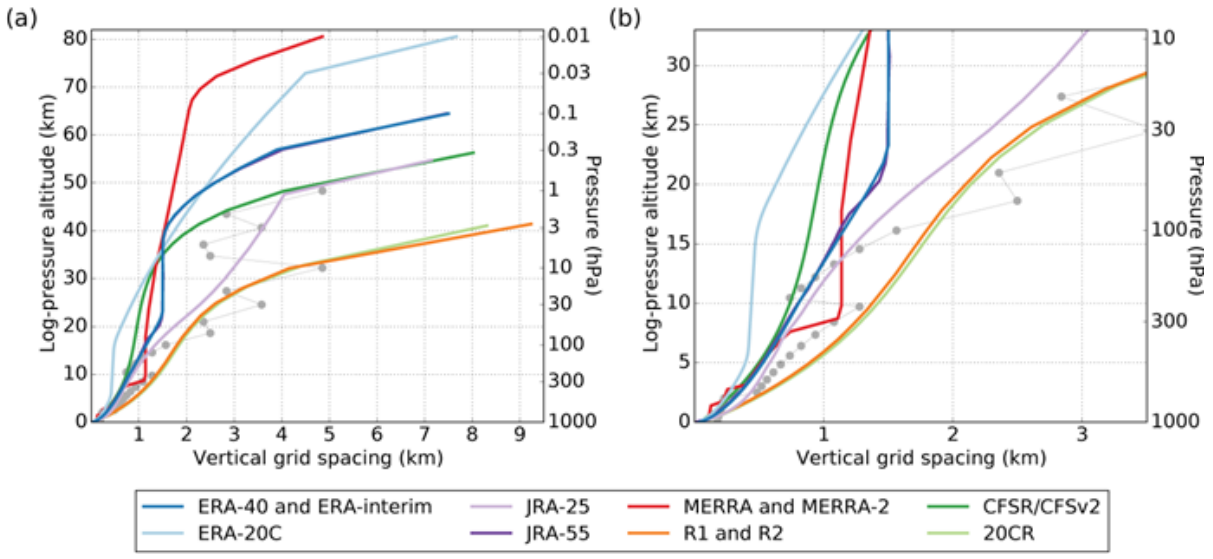


Figure 6: Approximate vertical resolutions of a selection of reanalysis forecast models for (a) the full vertical range of the reanalyses and (b) the surface to 33km (~10 hPa). Altitude and vertical grid spacing are estimated using log-pressure altitudes, where the surface pressure is set to 1000 hPa and the scale height is set to 7 km. The grid spacing indicating the separation of two levels is plotted at the altitude of the upper of the two levels, so that the highest altitude shown in (a) indicates the lid location. Some reanalyses use identical vertical resolutions; these systems are listed together in the legend. Approximate vertical spacing associated with the isobaric levels on which ERA-40 and ERA-Interim reanalysis products are provided (grey discs) is shown in both panels for context. Reproduced from Fujiwara et al. (2017).

unsurprisingly leads to differences between the reanalyses, particularly in areas where there is little observational data to constrain the model, for example the upper stratosphere.

A comparison of the vertical resolution of several reanalyses is shown in Figure 6. It is clear that from approximately 3 mb/40 km and upwards, MERRA2 has a higher vertical resolution than any of the other reanalyses. However, this by itself does not mean that MERRA2 has a more accurate representation of this region. This depends to a large extent on the observations that are assimilated.

The main operational observations of the stratosphere come from satellites, typically microwave or infra-red sounders. The instruments of these satellites are often nadir-pointing

(i.e. directly facing the ground) and measure spectral radiances – the amount of radiation/energy coming from different wavelengths. As different layers of the atmosphere emit strongly at different wavelengths, this is equivalent to measuring the radiances at different heights in the atmosphere, from which a vertical profile can be derived. A particular instrument will consist of several 'channels', each sensitive to a different wavelength. The greater the number of channels, the greater the vertical resolution of the instrument. The stratosphere is generally poorly sampled from such instruments, as the vertical resolution tends to degrade with height.

A quantitative measure of the vertical resolution of such measurements is hard to calculate because the layers of the atmosphere that each channel is sensitive to generally overlap, meaning that a particular layer of the atmosphere will actually be measured by several different channels, which have to be combined to obtain a temperature profile. A weighting function for each channel describes how the proportion of the total signal (from all channels) varies with height, with the vertical width of the weighting functions indicating the vertical extent of the atmosphere that was observed by each channel. An example is shown in Figure 7 (from Fujiwara et al. 2017). These are the weighting functions for the channels on the SSU (Stratospheric Sounding Unit 1979-2005, left) and AMSU-A (Advanced Microwave Sounding Unit A 1998 – present, right) instruments. The vertical resolution is higher for AMSU-A as there are more channels covering the same height and their weighting functions are narrower, however this is still limited by the degree to which the functions overlap.

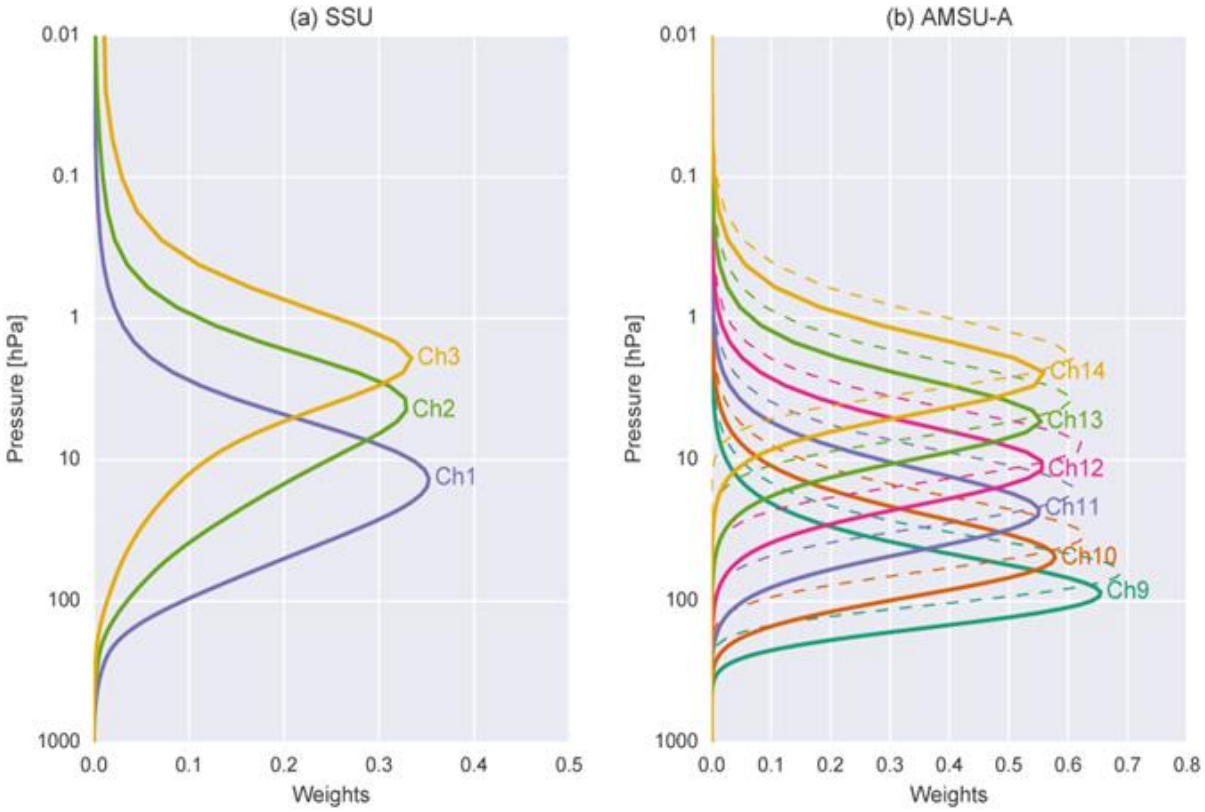


Figure 7: Vertical weighting functions of radiance measurements for (a) the TOVS-suite Stratospheric Sounding Unit (SSU) instrument (1979–2005) channel 1 (centred at ~15 hPa), channel 2 (~5 hPa), and channel 3 (~1.5 hPa) and (b) the ATOVS-suite Advanced Microwave Sounding Unit A (AMSU-A) instrument (1998– present) temperature channels 9–14 at near-nadir (1.67◦, solid lines) and limb (48.33◦, dashed lines) scan positions. Reproduced from Fujiwara et al. (2017).

More recent 'hyperspectral' instruments such as IASI (Infrared Atmospheric Sounding Interferometer), CrIS (Cross-track Infrared Sounder), and ATMS (Advanced Technology Microwave Sounder) have thousands of channels. However, MERRA2 is almost the only reanalysis so far to assimilate any of these datasets, with IASI assimilated from mid-2008 and CrIS and ATMS from the start of 2012. The CFSR (Climate Forecast System Reanalysis) is the only other reanalysis that assimilates IASI data.

More recently limb sounders have been developed and made operational. These instruments point at an angle, looking through the atmosphere out to space, instead of vertically down.

Because the background radiation of the Earth no longer needs to be taken account of, there is no need to consider what height the different channels correspond to. Instead, the different channels of the instrument correspond to the radiance at specific wavelengths coming from the portion of the atmosphere that is in the field-of-view of the instrument. The vertical resolution is therefore determined by the precision with which the instrument can change its field-of-view, instead of the weighting functions. The Microwave Limb Sounder (MLS) on the Aura satellite has a vertical resolution of approximately 5 km above 5 mb (Livesey et al. 2011). MERRA2 is the only reanalysis that assimilates MLS temperatures, and does so above 5 mb.

Whilst most of the available reanalyses have similar features to MLS and rocketsonde data in the upper stratosphere, including an annual cycle in the winds, there are still large differences in the amplitude and downward propagation of the winds (Long et al. 2017, Figure 9). The assimilation of MLS data above 5 mb in MERRA2 should make it more accurate, with differences between the other reanalyses largely due to differences in the internal models used. We can see from Long et al. (2017) Figure 9 that MERRA2 has climatological winds that are ∼6 m/s more westerly than MERRA (which does not assimilate MLS data) between 5 and 1 mb throughout the year. This is more likely to be a result of changes made to the non-orographic gravity wave scheme of the internal model between MERRA and MERRA2 which may cause a bias in this region (Molod et al. 2015, Coy et al. 2016). It also has by far the strongest westerly winds in this region of all the reanalyses analysed in Long et al. (2017). Over the equator the winds are not strongly constrained by geostrophic or gradient wind balance, which means equatorial zonal mean zonal wind profiles derived from temperature

observations, including MLS, should also be treated with caution. The lack of multiple independent observations in this region therefore makes it hard to ascertain the 'truth' of the winds in this region. However, in this study I have chosen to mainly use MERRA2 for its assimilation of MLS data, lack of gravity wave-absorbing sponge layers below approximately 0.2 mb, and non-orographic gravity wave scheme.

3.2: The model

The experiments presented here use a computer model of the atmosphere. Such models are typically highly complex, but ultimately all solve the equations of motion, conservation of heat and mass on a sphere by discretising said equations over a 3-dimensional grid. The values of the prognostic variables (e.g. wind speed and temperature, or variables from which these can be calculated) are stored at each grid point and updated to new values according to the equations every timestep. Few global scale models have a horizontal resolution of less than a few kilometres, meaning that the physical processes occurring on scales smaller than this (sub-grid scales) such as convection, gravity waves and cloud microphysics are not represented. Instead, these processes are parameterised, which means that the effect that they would have on the grid-box values of the relevant variables is calculated and added to said variables during the model timestep. Parameterisations are crucial in the representation of gravity-wave-driven phenomena in models, in particular the QBO.

In this study I use the atmosphere-only HadGEM3 configuration of the UK Met Office's Unified Model (UM, Brown et al. 2012) with UM version 10.3 and 'Global Atmosphere' version 7.0 (GA7.0) code (Walters et al. 2017). GA7.0 uses the 'Even Newer Dynamics for General

atmospheric modelling of the environment' (ENDGame) dynamical core. HadGEM3-GA7.0 comes with a choice of horizontal and vertical grid resolutions. I use the highest vertical resolution version, 'L85', which has 85 levels in the vertical on a hybrid-height terrain-following coordinate system. The 85 levels are split roughly 50-35 between the troposphere and the upper atmosphere (above 18 km) respectively, with a model top at 85 km (0.0053 mb). I use a horizontal resolution of 1.875° x 1.25°, which has $2N$ longitude grid points and $1.5N$ latitude grid points where $N = 96$. This is referred to as 'N96'.

GA7.0 includes the Ultra Simple Spectral Parameterization (USSP) of non-orographic gravity wave drag (Scaife et al. 2000, Scaife 2002), allowing the model to simulate its own spontaneous QBO, as the QBO is partially driven by the non-orographic gravity waves generated by convection in the tropics. The model also has a spontaneous SAO.

The large number of UM model versions and lack of externally published assessments for the stratosphere, including the most recent versions (such as HadGEM3-GA7.0) make it hard to assess the biases of the model in the stratosphere compared to observations. The most recent comprehensive assessment of the stratosphere in the UM model is Hardiman et al. (2010) and Osprey et al. (2010) which used the earlier HadGEM1 configuration. The version of HadGEM1 in Hardiman et al. (2010) has the same model top height but lower vertical resolution (L60). Whilst the dynamical core is different from HadGEM3-GA7.0 the non-orographic gravity wave scheme, crucial in generating a spontaneous QBO, is the same. Changes were made to the parameters in the scheme to keep it consistent between model versions as a result of changes to other aspects of the model code so overall the QBO in HadGEM1-L60 is likely to be similar to HadGEM3-GA7.0. Less heed has been paid to changes

in the SAO, however Andrew Bushell (personal communication) has found that the amplitude of the SAO has changed little between model versions and resolutions.

In Hardiman et al. (2010) the winter polar vortex (referred to as the polar night jet) of HadGEM1 was found to be approximately 10 m/s too weak with half the interannual (year-to-year) variability relative to the ERA-40 reanalysis (i.e. it is consistently too weak). This was attributed to excessive vertical planetary wave propagation. The summer easterlies were found to be too strong. Whilst the tropical tropopause temperature and moisture bias in HadGEM1 has been significantly reduced in HadGEM3-GA7.0 (Walters et al. 2017), it is unclear how the other climatological biases have changed.

Osprey et al. (2010) found the QBO of HadGEM1 to be in good agreement with ERA-40 in terms of the height and magnitude of the peak amplitude, with a slightly longer QBO period. Similar to Bushell et al. (2010), the amplitude of the SAO was found to be overestimated in the model compared to the respective reanalyses used, with a much weaker annual cycle. A comparison of the climatological equatorial winds in HadGEM3-GC3.1 (HadGEM3-GC3.1 is the coupled atmosphere-ocean version of HadGEM3 using GA7.1) is shown alongside that of the ERA-Interim and MERRA2 reanalyses in Figure 8. The observed annual cycle in the SAO can clearly be seen in the easterly phases of both reanalyses whereas in the model there is little change in the amplitude of the phases throughout the year. There is a clear easterly bias in the model, with the westerly phase much weaker in the model than both the reanalyses and the easterly phase slightly stronger.

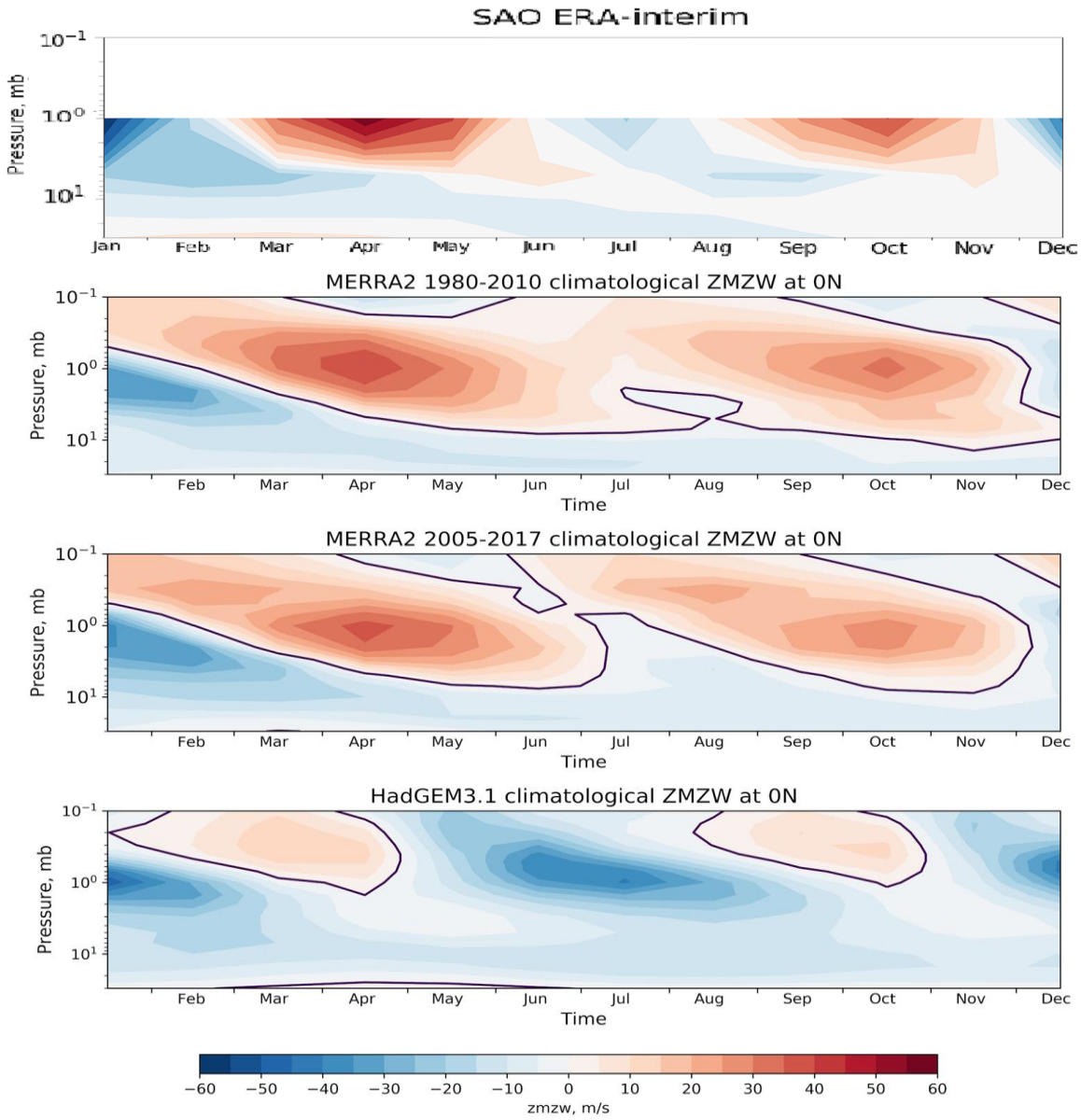


Figure 8: The climatology of the zonal mean zonal wind (zmzw) at the equator (zero degrees latitude) for (from top) ERA-Interim for 1979-2016, MERRA2 for 1980-2010, MERRA2 for 2005-2017 and HadGEM3.1. The HadGEM3.1 climatology was taken from a 500 year long pre-industrial control run of the model. The black contour is the zero wind line and the unlabelled x-axis tick marks represent the month boundaries. The ERA-Interim plot was provided by Oscar Dimdore-Miles at the University of Oxford. See main text for expansion of acronyms.

3.2.1 Ensemble Setup

The atmosphere is a chaotic system. This means that simulations of the atmosphere in computer models are extremely sensitive to initial conditions, with simulations starting from atmospheric states differing by a tiny amount diverging after a few days and looking completely different after a few weeks. As our measurements of the atmosphere are far from perfect, all simulations of the atmosphere will suffer from the uncertainty that this introduces. To get a handle on this uncertainty simulations are typically run as an ensemble, with each ensemble member starting from a slightly different initial state. The resulting spread of the ensemble members over time gives an idea of the uncertainty in the simulation. Some models, such as the UM, also deliberately introduce perturbations into their ensembles to achieve the same effect, and account for the additional uncertainty introduced by the necessary parameterisation schemes.

The GA7.0 code in the UM uses SKEB2 (Stochastic Kinetic Energy Backscatter 2) and SPT (Stochastic Perturbed Tendencies) detailed in Sanchez et al. (2016) to stochastically perturb each ensemble member. SKEB2 applies spatially and temporally correlated random noise to the model winds, and SPT applies variations correlated in time and space multiplicatively to the increments to the prognostic variables that result from the parameterisations only. Physically, SKEB2 is attempting to add back in the kinetic energy that is lost because of underrepresentation of upscale kinetic energy processes in the model. Upscale kinetic energy processes are those by which kinetic energy is transferred from small to large scale flow. SPT on the other hand accounts for the local (in time and space) uncertainty in the parameters of parameterisations that results from the gridbox-scale processes.

3.3: Nudging

MERRA2 is used to constrain the HadGEM3 ensemble simulations in different regions of the atmosphere by a simple nudging or Newtonian-relaxation process, where the model temperature and winds are 'nudged' to the reanalysis values at each timestep using the following equation:

$$\Delta X = G\Delta t(X_{analysis} - X_{model})$$

after Jeuken et al. (1996). Here X can be the zonal or meridional wind velocities or temperature. ΔX is the increment applied to X by the nudging over one model timestep, Δt. $X_{analysis}$ is the value of the nudged variable in the analysis, X_{model} is the model value of X, and G, a constant, is the relaxation parameter. G can be viewed as $\frac{1}{relaxation\ timescale}$. The default relaxation timescale in the UM nudging scheme is 6 hours. According to the original study, Telford et al. (2008), this timescale was originally chosen arbitrarily in keeping with other nudged models at the time. The main requirement of the nudging in this study is to constrain different regions of the atmosphere to the reanalysis dataset, and this default relaxation timescale achieves this in the troposphere and stratosphere, despite the variation in the radiative timescales of different levels of the stratosphere.

The MERRA2 reanalysis data are provided to the model on the standard 72 hybrid-pressure levels. The reanalysis data can be downloaded from GES DISC (The Goddard Earth Sciences Data and Information Service Center, https://disc.gsfc.nasa.gov/) and then converted to both N48 and N96 horizontal resolutions through bi-linear interpolation. At each timestep the reanalysis data are then linearly interpolated in log-pressure to the model hybrid-height grid

and linearly interpolated in time to the model timestep before being applied to the nudged parameters according to the nudging equation above at the end of the model timestep (after all the other model physics has been applied). The reanalysis data are provided at timesteps of 6 hrs to correspond with the relaxation timescale.

3.4: Experiment Setup

The results presented below investigate the constraints that different regions of the atmosphere place on the Northern Hemisphere polar vortex in the winter of 2008/9 with a particular focus on the role of the upper tropical and subtropical stratosphere. This is achieved by fixing the horizontal winds and temperature of different regions of the atmosphere to the MERRA2 reanalysis but always leaving the polar vortex region unconstrained.

All the experiments are initialised to ERA-I, as the appropriate files were readily available on the supercomputer used. The stochastic forcing described in section 3.2.1 is used in a 1-month run from 1st August to 1st September 2008 with global nudging to ERA-I horizontal winds and temperature to generate a set of initial conditions that are slightly perturbed about the reanalysis. From here, a set of experiments with nudging to MERRA2 horizontal winds and temperature in different regions of the atmosphere are started. Each experiment consists of 50 ensemble members. The experiments are detailed below and summarised in Table 1. Except where specified otherwise all nudging regions have a linear taper applied over the lower 4 levels by multiplying the nudging increment by a taper factor between 0 and 1.

Experiment description	Name	Troposphere region	Stratosphere region
No nudging	Control	-	-
Tropospheric Sensitivity Experiments			
Nudging in the lower troposphere	LowTrop	Up to 6 km	-
Nudging throughout the troposphere	AllTrop	Up to tropopause	
Stratospheric Sensitivity Experiments			
Nudging in the upper stratosphere	UpStrat	-	Above 5 mb level
Nudging in the lower troposphere and upper stratosphere	LowTrop+UpStrat	Up to 6 km	Above 5 mb level
Tropical Stratosphere Sensitivity Experiments			
Nudging in the lower troposphere and the tropics and subtropics of the upper stratosphere	LowTrop+UpStrat-Tropics	Up to 6 km	Above 5 mb level, 0-30 °N only
Nudging throughout the troposphere and the tropics and subtropics of the upper stratosphere	AllTrop+UpStrat-Tropics	Up to tropopause	Above 5 mb level, 0-30 °N only

Table 1: Table showing the region(s) of the atmosphere relaxed/nudged to reanalysis data in the model simulations presented in this report.

- **Control**: No nudging anywhere. The model is allowed to run freely, constrained only by the prescribed sea surface temperatures (SSTs) at the lower boundary and 1st September initial conditions.
- **LowTrop**: Same as **Control** except with the addition of nudging between model levels 12 and 29 roughly corresponding to between the boundary layer and 6 km above the surface.
- **AllTrop**: Same as **Control** except with the addition of nudging between model levels 12 and the model tropopause, defined at each horizontal grid point as the lowest model level with a lapse rate of less than 2 K/km.
- **UpStrat**: Same as **Control** except with the addition of nudging from model level 69 (approximately 5 mb) up to the model top.
- **LowTrop+UpStrat**: Same as **LowTrop** but with the addition of nudging in the same region as in **UpStrat**.
- **LowTrop+UpStrat-Tropic**s: Same as **LowTrop+UpStrat** but with the nudging in the upper stratosphere restricted to 0-30 °N with a linear taper approximately 10 degrees either side. Where the vertical and latitudinal tapers overlap, the taper factors are multiplied together before being applied to the nudging increment.
- **AllTrop+UpStrat-Tropics**: Same as **LowTrop+UpStrat-Tropics** except with the top of the tropospheric nudging region extended up from 6 km to the model tropopause, defined as in **AllTrop**.
- Note that both 'UpStrat-Tropics' experiments will constrain the wave-breaking region in the subtropics quite strongly as well as the tropical zonal winds.

Section 4: Results

4.1: Introduction to the winter of 2008/9 in the stratosphere

The evolution of the 2008/9 zonal mean zonal wind (zmzw) in the stratospheric polar vortex is shown in Figure 9 (bottom right)**.** From September the radiatively-driven development of the westerly winds of the polar vortex occurs. This process is disrupted at various points before the warming by planetary wave activity. This can be seen clearly in the first half of November and start of December by rapid weakenings of the zonal mean zonal wind.

The winter of 2008/9 was characterised by a strong, split SSW (see Figure 5 (right)) that occurred at the end of January, following a particularly strong vortex from the end of December. The reversal of the westerly winds in the warming starts high up in the mesosphere, above 0.1 mb and rapidly descends through the stratosphere. These easterly winds persist in the lower stratosphere for most of February, whilst higher up the radiatively-driven westerlies are re-established from the top-down in the absence of planetary wave activity from below (Iida et al. 2014).

Over the equator the QBO (Figure 9, top right) is in a westerly phase, with westerly winds between 70 and 20 mb. Above this the next easterly phase is forming but does not descend much during the winter. Above this the SAO evolves from its climatological westerly phase in autumn (September to November) through an easterly phase in mid-winter (December to February) and back to westerly by late winter (March onwards). The 11-year solar cycle is at a minimum (Figure 9, top left), which means that the equatorial temperatures are lower than average, reducing the equator-pole temperature gradient, which then implies a small easterly

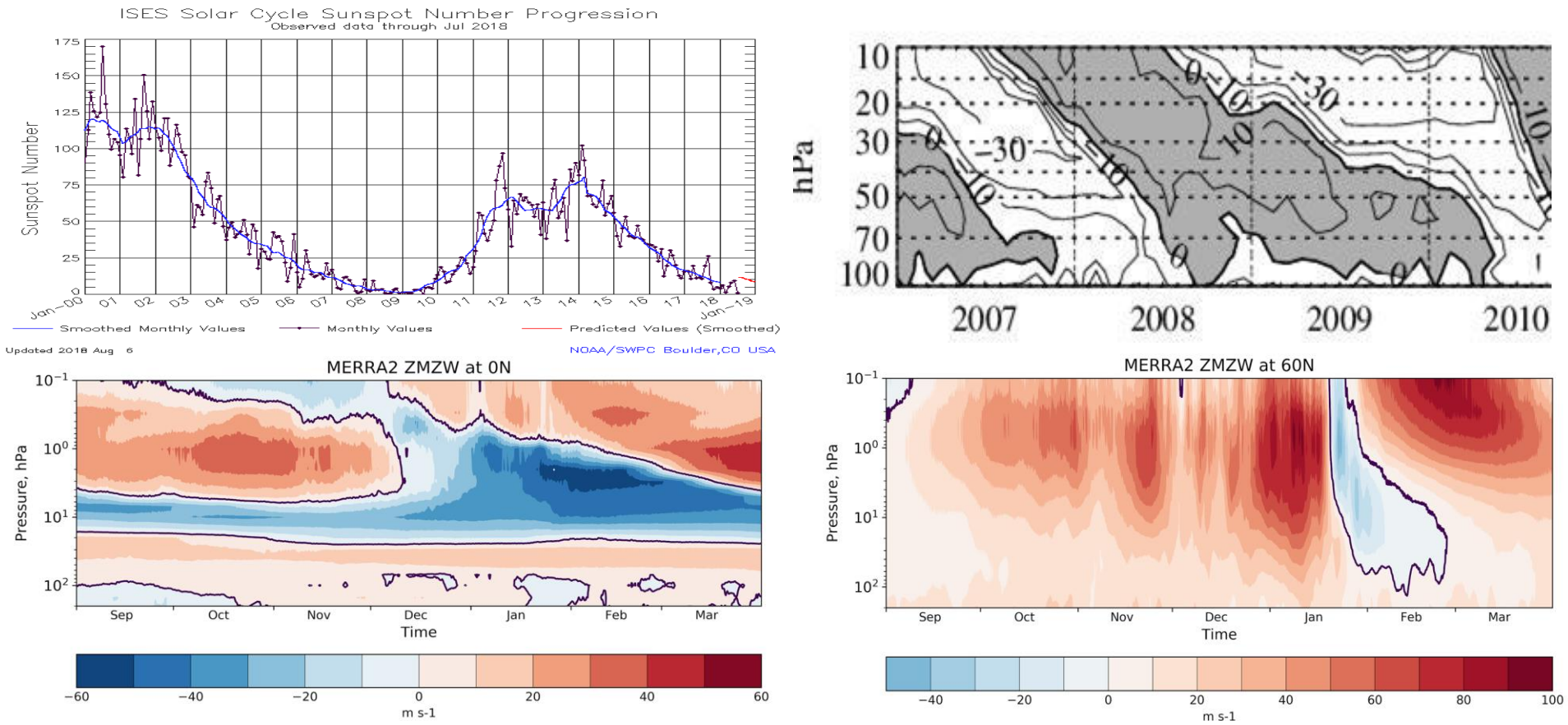


Figure 9: Top left: The monthly average number of sunspots observed from 2000 to present (black line) and the 13-month running average (blue line), as a proxy for the solar cycle. Figure obtained from the National Oceanic and Atmospheric Administration Space Prediction Centre (https://www.swpc.noaa.gov/products/solar-cycle-progression). Top right: Time-height section of monthly mean zonal winds (m/s) at Singapore, 1 °N/104 °E for 2007-2010. Isopleths are at 10 m/s intervals; westerlies are shaded. Figure adapted from Freie Universität Berlin Institute for Meteorology (http://www.geo.fu-berlin.de/en/met/ag/strat/produkte/qbo/index.html#compilationhttp://www.cpc.ncep.noaa.gov/products/stratosphere/strat_int/). Bottom right: Time-height plot of the zonal mean zonal wind (m/s) from MERRA2 at a latitude of 60 °N for September 2008 through March 2009. Red represents westerly winds whilst blue represents easterly. The solid black contour is the zero wind line. The x-axis tick marks represent the month boundaries. Bottom left: As bottom right but for the equatorial winds (zero degrees latitude).

anomaly in the subtropical upper stratosphere (Gray et al. 2010).

4.2: Experiments

4.2.1: Control experiment

Control: Before conducting the other nudging experiments in Table 1 I needed to first examine whether the model was able to simulate the warming when it was only constrained by the SSTs and the September initial conditions.

The zonal mean zonal wind at 10 mb & 60 °N is employed as the standard diagnostic for monitoring the evolution of SSWs. Figure 10 shows the 50-member Control ensemble together with the ERA-I/MERRA2 evolution and the ensemble-mean, for comparison. There is a large spread between the ensemble members, which grows with time as the constraining effect of the initial conditions fades. Although some do show SSWs through a reversal of the zonal mean zonal wind, there is a large range of timings throughout the winter. The timing is essentially random between the members as evidenced by the ensemble mean showing no SSWs and an evolution similar to climatology (not shown). We can therefore conclude that the SSTs do not constrain the stratosphere on the daily timescales needed to simulate a warming and the September initial conditions only determine the subsequent evolution for the first month or so**.** Therefore, the model cannot simulate the observed January SSW with only September initial conditions and prescribed SSTs.

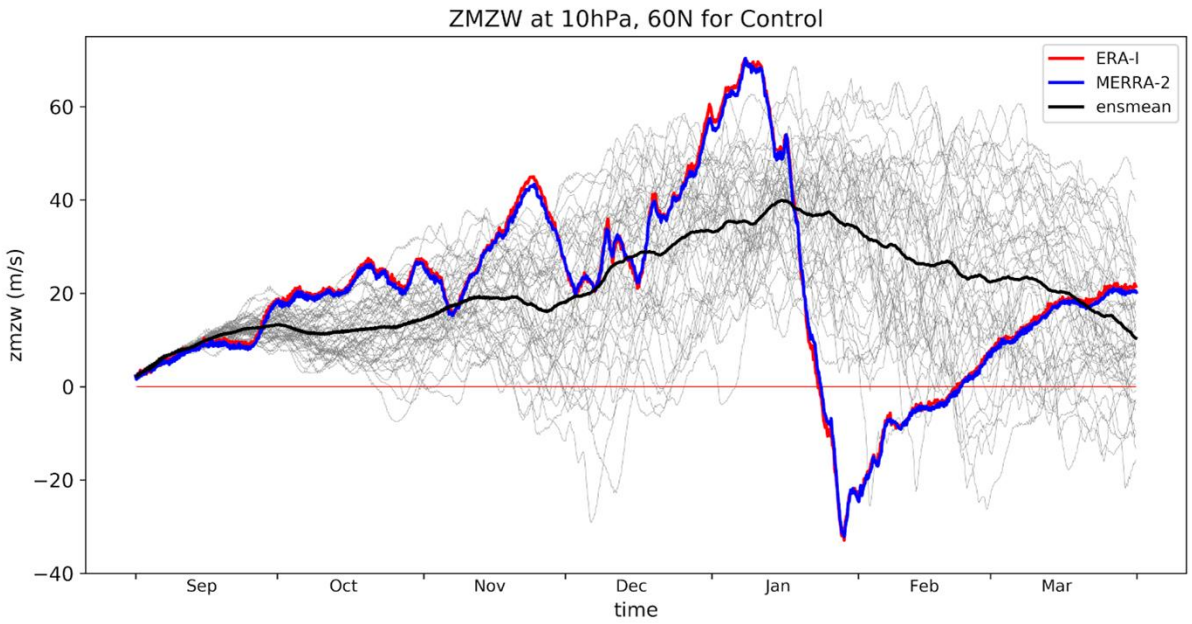


Figure 10: Zonal mean zonal wind (m/s) at 10 hPa and 60 degrees north for September 2008 through March 2009. The tick marks on the x axis represent the month boundaries. The grey lines are the individual ensemble members of the Control experiment and the black line is the ensemble mean. The thick red and blue lines show the reanalysis data (ERA-I and MERRA2 respectively). The thin, horizontal red line is the zero wind line.

4.2.2: Tropospheric Sensitivity Experiments

LowTrop: This next experiment is identical to the control except for the addition of nudging towards MERRA2 in the lower troposphere, below 6 km. This experiment is designed to test how much the troposphere controls the influence of the stratosphere through the upward propagation of planetary waves this winter, and whether the warming can be simulated by prescribing these alone. The upper boundary of the nudging was originally chosen as 6 km to avoid the risk of nudging the bottom of the SSW.

Figure 11 (left) shows that the ensemble members at 10 mb are more constrained than the control throughout the simulation, particularly up to around early November. They all follow the reanalysis during this period. After this the model diverges from the reanalysis where

there is a notable split in the ensemble members between those following the strengthening of the polar vortex and those remaining constant or weakening slightly. There is a similar occurrence at the start of December, with the larger portion of the ensemble members undergoing a large SSW that is not present in the observations. As a result, the model ensemble-average remains consistently weaker than the observed stratosphere until the warming in mid-January, although the rate of the increase in the zonal mean zonal wind over this time period is similar. Around the time of the actual warming, some of the ensembles do show an SSW, but none as strong as that observed. In Figure 11 (left) we can see that the ensemble mean does experience a significant burst of planetary wave activity in the stratosphere, as the strong westerly winds present as the vortex recovers from the earlier spurious warming suddenly weaken around the time of the observed warming. After this, whilst the observed stratosphere is recovering from the SSW, the majority of the ensemble members experience a later SSW, which again is not as strong as the earlier observed SSW. So the model ensemble mean does, just, experience a warming but far too late.

The timings of the LowTrop model warmings are no longer random though - their evolution has been more tightly constrained to three distinct times: early December, late January (in which the observed warming occurs) and late February through March. Constraining the troposphere up to 6 km is therefore not enough for an accurate simulation of the winter stratosphere initialised in September, in particular the January SSW, despite this fixing the wave activity propagating from the lower troposphere. This indicates that it is not just the source of vertically propagating waves, but also the atmosphere above 6 km through which they propagate, and how the stratosphere responds to such waves, that may be important in

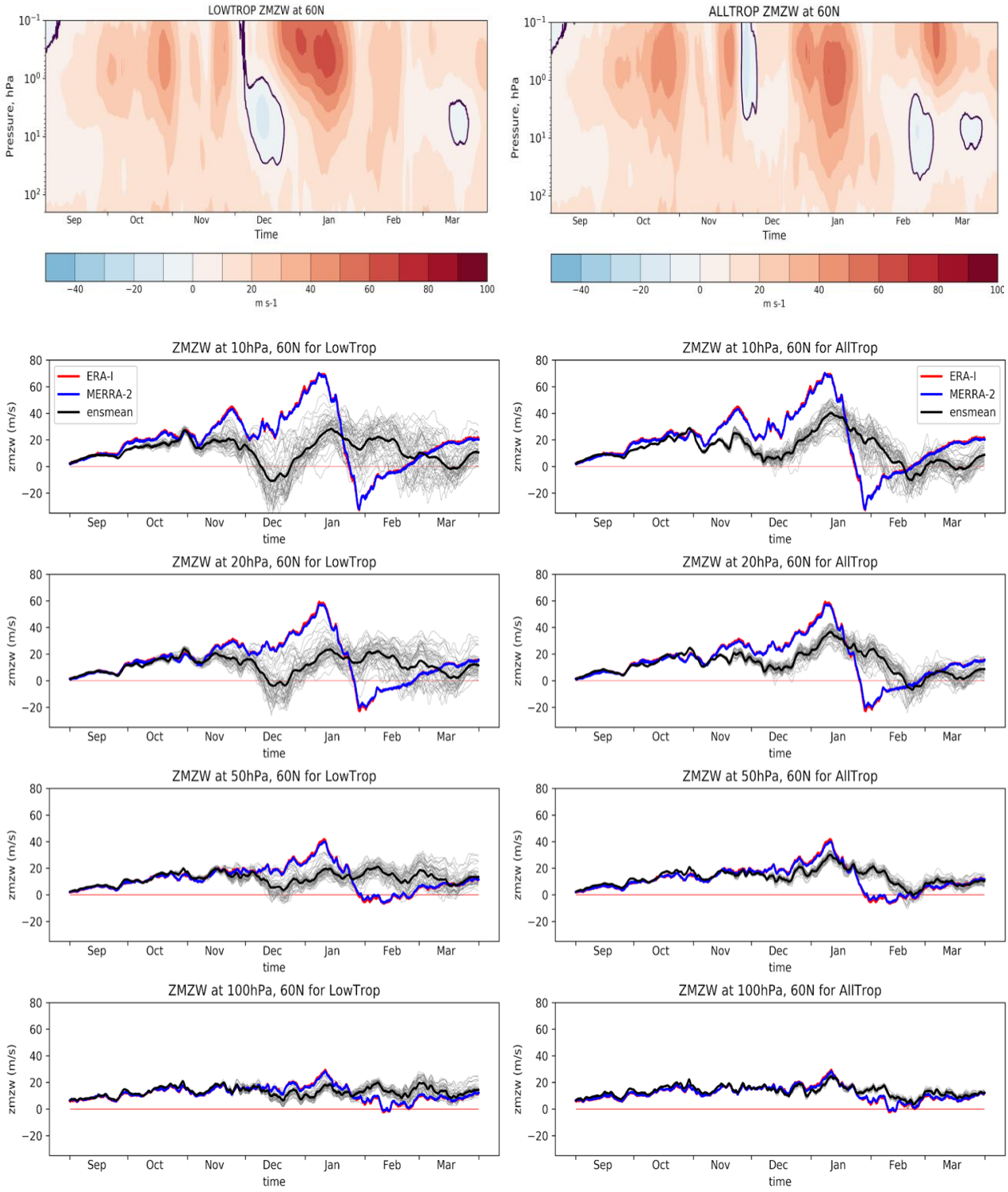


Figure 11: All plots on the left are for the LowTrop experiment, the right for AllTrop. Top: Time-height section of zonal mean zonal wind (m/s) at 60 degrees north from September 2008 through March 2009. Red colours represent westerly winds, blue represents easterly. The black contour is the zero wind line. Bottom four plots: The zonal mean zonal wind (m/s) at 60 degrees north for the individual ensemble members (grey lines), ensemble mean (black line), ERA-I (red line) and MERRA2 (blue line). From top to bottom the plots are for 10, 20, 50 and 100 hPa. The thin red line is the zero wind line. The tick marks on the x axis represent the month boundaries.

determining the evolution of the stratosphere. Figure 11 (left) also shows the evolution at different depths in the atmosphere, showing how deeply the observed SSW penetrates into the lower stratosphere. This is another useful diagnostic for comparison between the observed and simulated evolution, discussed in more detail in later experiments.

AllTrop: In this experiment I nudge throughout the troposphere all the way up to the model tropopause as opposed to 6 km and compare the two experiments. This experiment therefore explores the sensitivity to the flow in the upper troposphere, in the vicinity of the tropopause.

Figure 11 (right) shows that nudging in the whole troposphere does constrain the model ensemble members in the stratosphere more at 10 mb although they all still diverge from the reanalysis during the latter half of November. The spurious warming in mid-December is much weaker than in **LowTrop** (see evolution of the ensemble-mean in the top panel) and does not descend as far. Below 10 mb the ensemble members of **AllTrop** have stronger westerly winds and remain closer to the reanalysis than in **LowTrop** until the warming in January. At ∼1 mb the **AllTrop** vortex in January is weaker than in **LowTrop**, presumably because there is reduced shielding of the waves by the easterly winds at lower levels. The model does show a weakening of the vortex around the same time as the observed SSW, but the weakening of the westerly winds is much slower, with the winds of the ensemble mean not reversing until around 20 days after the reanalysis (in late February), by which point the reanalysis winds have recovered to westerly again. Compared to the **LowTrop** experiment this late-winter warming does occur sooner (late February compared to mid-March), but still nowhere near the observed warming.

A similar story occurs at 20 and 50 mb, with the ensemble members more tightly constrained and following the evolution of the zonal mean zonal wind more closely up until the observed warming. Again the warming starts around the same time as in the reanalysis, but evolves much too slowly. At 100 mb the **AllTrop** ensemble members are strongly constrained by the proximity of the tropospheric nudging region but a gap between the reanalysis and ensemble mean warming is still present.

This experiment shows that prescribing the wave flux from the entire troposphere is still not enough to accurately simulate the evolution of the stratosphere, in particular the SSW. Whilst the timing of the initialisation of the warming is accurately simulated the other key aspects such as the strength, speed and depth of the warming are not. Internal stratospheric dynamics must therefore be playing an important role in these aspects of the warming.

We can therefore conclude that constraining the troposphere in this way does not completely determine the wave flux in the stratosphere, although further diagnostics such as EP-fluxes would be helpful to confirm this.

4.2.3: Stratospheric Sensitivity Experiments

UpStrat: We have so far examined the influence of the troposphere on the evolution of the polar vortex. Next we consider the influence that the upper atmosphere (up to 85 km) may have on the evolution below in **UpStrat** by nudging only above the 5 mb level (approximately 36 km) in the model. Note that this nudging is performed at all latitudes, in contrast to some of the experiments described later on. Because SSWs are usually initiated in the upper

stratosphere and then descend rapidly to lower levels (see section 2.4) we might expect this nudging to help with the timing and initial speed of the SSW.

Firstly, we can see from Figures 12 and 11 that the spread of the ensemble members at 10 mb is smaller than in **LowTrop**, and similar to **AllTrop** indicating that the model stratosphere winds at this level are similarly constrained by the upper atmosphere and the troposphere individually.

In addition, the ensemble mean in **UpStrat** is much closer to the reanalysis at 10 mb than either **LowTrop** or **AllTrop**, with the ensemble spread encompassing the reanalysis through to mid-January (i.e. until the observed SSW occurs). The divergence from the reanalysis that occurs in both the tropospheric nudging experiments around the start of November at 10 mb and later for lower levels does not occur in the **UpStrat** experiment, meaning that the ensemble mean continues to track and ensemble spread continues to encompass the reanalysis winds at all levels down to 100 mb until the SSW.

Regarding the warming, the model does simulate a warming at close to the right time and at 10 mb at least gets the speed of the warming right with the rate of decrease in the zonal mean zonal wind almost the same as the reanalysis. However, the model warming is not strong enough (the ensemble mean winds at 10 mb do not achieve a large enough easterly maximum) nor deep enough. The latter is particularly clear in Figure 12 (top) where the zonal mean zonal wind reversal at 60 °N does not extend nearly as far into the troposphere or persist for as long as in the reanalysis. Nonetheless it is a distinct improvement over the

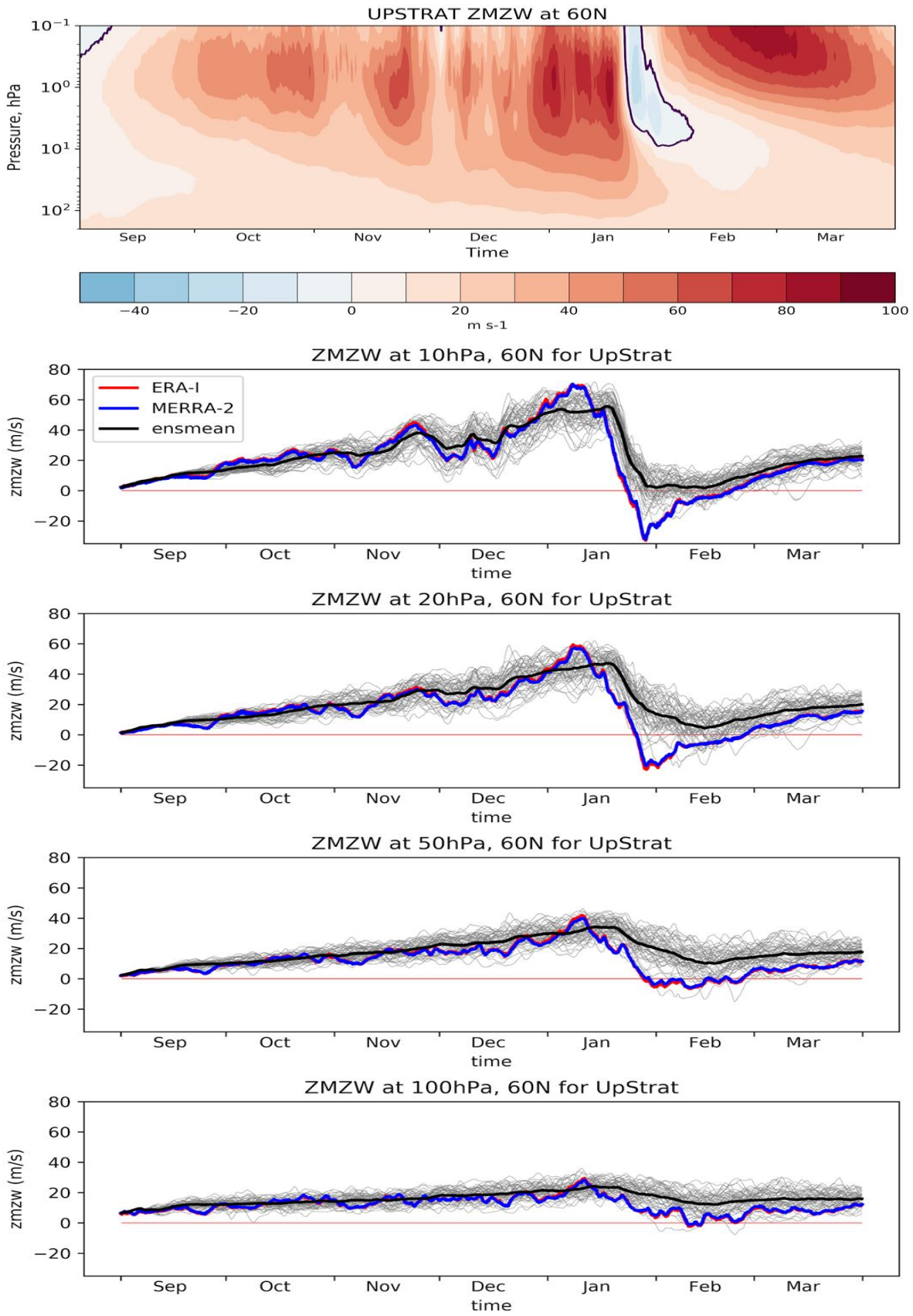


Figure 12: As one side of Figure 11, but for the UpStrat experiment.

simulation of the warming in **AllTrop**. From this, we can conclude that the upper atmosphere is more important for determining the evolution of the polar vortex at 10 mb than the troposphere. It is however important to note that 5 mb is arguably not that far away from 10 mb so it is perhaps not surprising that constraining above 5 mb has such a strong influence, at least at 10 mb. Both the troposphere and upper atmosphere are able to start the warming at the correct time when constrained however constraining the upper atmosphere yields a much more accurate simulation of the strength and depth. As with the full troposphere, nudging in the upper atmosphere is not enough on its own to get a fully accurate simulation of the warming, with inaccuracies in the strength and depth, if not the speed.

LowTrop+UpStrat: We have so far seen that for the 2008/9 winter both the troposphere and upper stratosphere are, separately, important for accurately simulating the timing of the warming, with only the upper stratosphere nudging able to accurately simulate the speed of the warming. Additional improvements in the strength and depth of the warming over **AllTrop** are also present in **UpStrat**. Next I combine the **LowTrop** nudging with the **UpStrat** nudging, shown in Figure 13.

When the model is nudged both up to 6 km and above 5 mb, the evolution of the zonal mean zonal wind at 10 mb (60 °N) is much closer to the observations throughout most of the simulation period, for all the ensemble members. This is also the case at the other levels shown, but with a small westerly deviation of the model apparent after around mid-October, particularly at 50 and 100 mb.

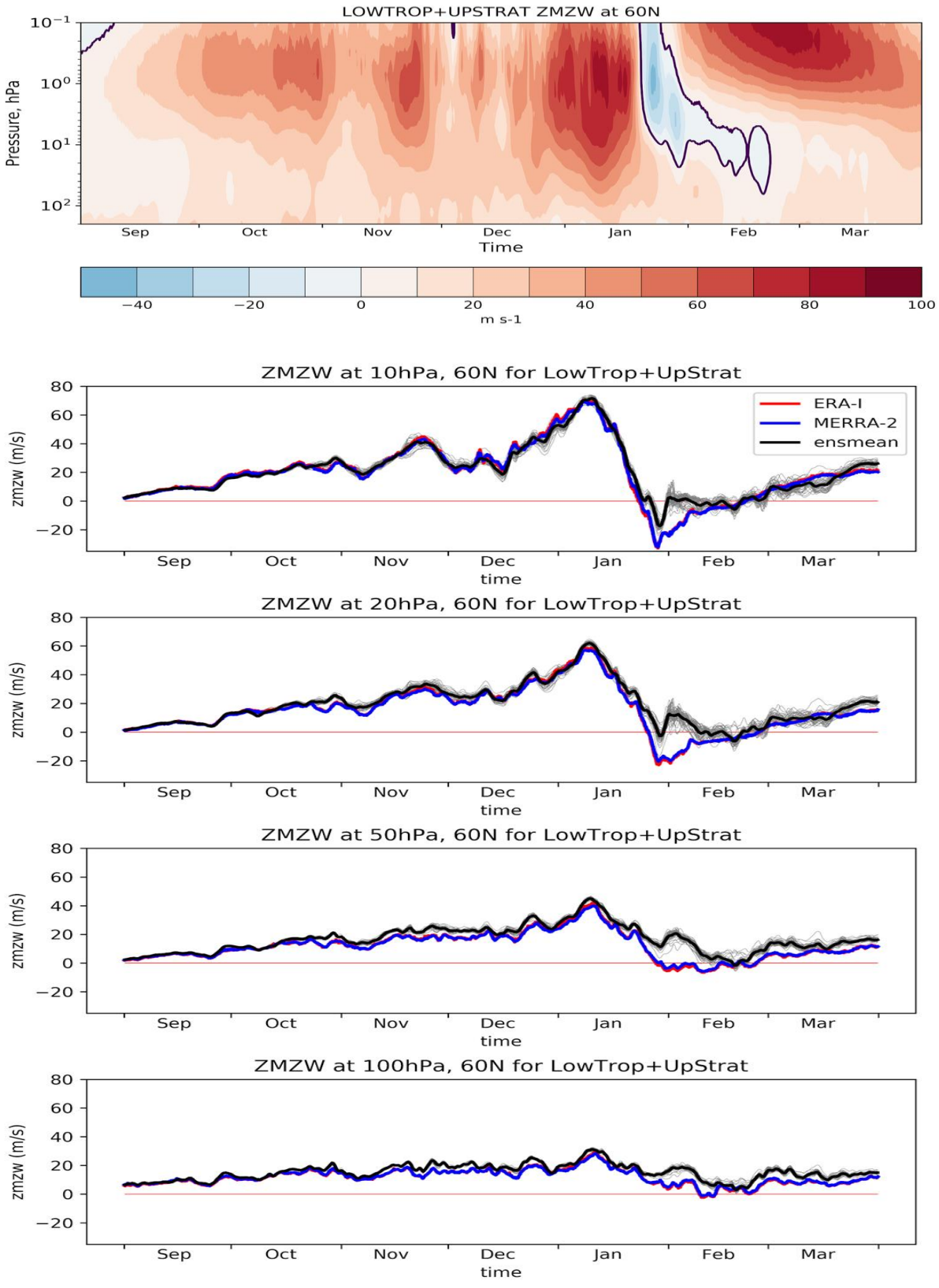


Figure 13: As Figure 12, but for the LowTrop+UpStrat experiment.

The timing and speed of the warming is accurately simulated at all levels shown, up until around the peak of the warming, where the decrease of the model winds slows and then reverses. The increase in the winds occurs around the same time as in the reanalysis, except it is much stronger in the model, and particularly abrupt at 10 and 20 mb. This leads to the model zonal mean zonal winds being more westerly during the final two weeks of the warming. During this time the model winds decrease again slowly until they meet the reanalysis winds, closing the gap between them. The quicker recovery of the model winds compared to the reanalysis is consistent with Osprey et al. (2010) who found faster timescales in the northern hemisphere annular mode in the winter stratosphere of HadGEM1 (an older version of the model used in this study) compared to the ERA-40 reanalysis (see their Figure 13).

Overall the easterly winds of the warming do not penetrate as deep or persist for as long in the lower stratosphere (Figure 13 (top)) – the model is still unable to simulate the full depth and strength of the warming. It is important to note that the ensemble spread has collapsed significantly in this experiment – it is not just the ensemble mean that is following the reanalysis as described above, but each member. This indicates that the model is being strongly constrained between 10 and 100 mb by the nudging above and below. This demonstrates that a large amount of the evolution of the polar vortex can be explained by influences from the lower troposphere and upper stratosphere, with deficiencies only in the finer details of the SSW.

4.2.4: Tropical Stratosphere Sensitivity Experiments

LowTrop+UpStrat-Tropics: In the **LowTrop+UpStrat** experiment we saw clear improvements in the model simulation of the warming. A more interesting question to ask though is which part of the upper stratosphere nudging region is causing the substantial improvement of **LowTrop+UpStrat** over just **LowTrop**. Due to the largely barotropic (height independent) nature of the circulation in the stratosphere it is perhaps not all that surprising that controlling the top of the stratosphere directly above where the warming occurs is able to generate a warming. Therefore the next experiment repeats **LowTrop+UpStrat** but restricts the nudging region in the upper stratosphere to between latitudes 0 and 30 °N, to remove this effect, and see if the link between the tropical upper stratosphere and the polar vortex suggested by Pascoe et al. (2006) and previous studies is evident in this particular winter.

Figure 14 shows the evolution of the **LowTrop+UpStrat-Tropics** (right panels) compared with the **LowTrop** (left panels – reproduced from Figure 11). As expected, we do not see as much of an improvement between **LowTrop** and **LowTrop+UpStrat-Tropics** as we did between **LowTrop** and **LowTrop+UpStrat** (Figures 11, 13). Nonetheless, there is still a notable improvement in the simulation of the warming. The model zonal mean zonal wind at 60 °N in **LowTrop+UpStrat-Tropics** follows the reanalysis closely, particularly when we compare the period from the start of November onwards. In **LowTrop+Upstrat-Tropics** at 10 and 20 mb the ensemble members no longer diverge from the reanalysis here. The speed and strength of the spurious warming in mid-December is much reduced at all levels. Easterly winds briefly appear around mid-December but are not persisting. In the subsequent months up to the warming the ensemble mean is closer to the reanalysis.

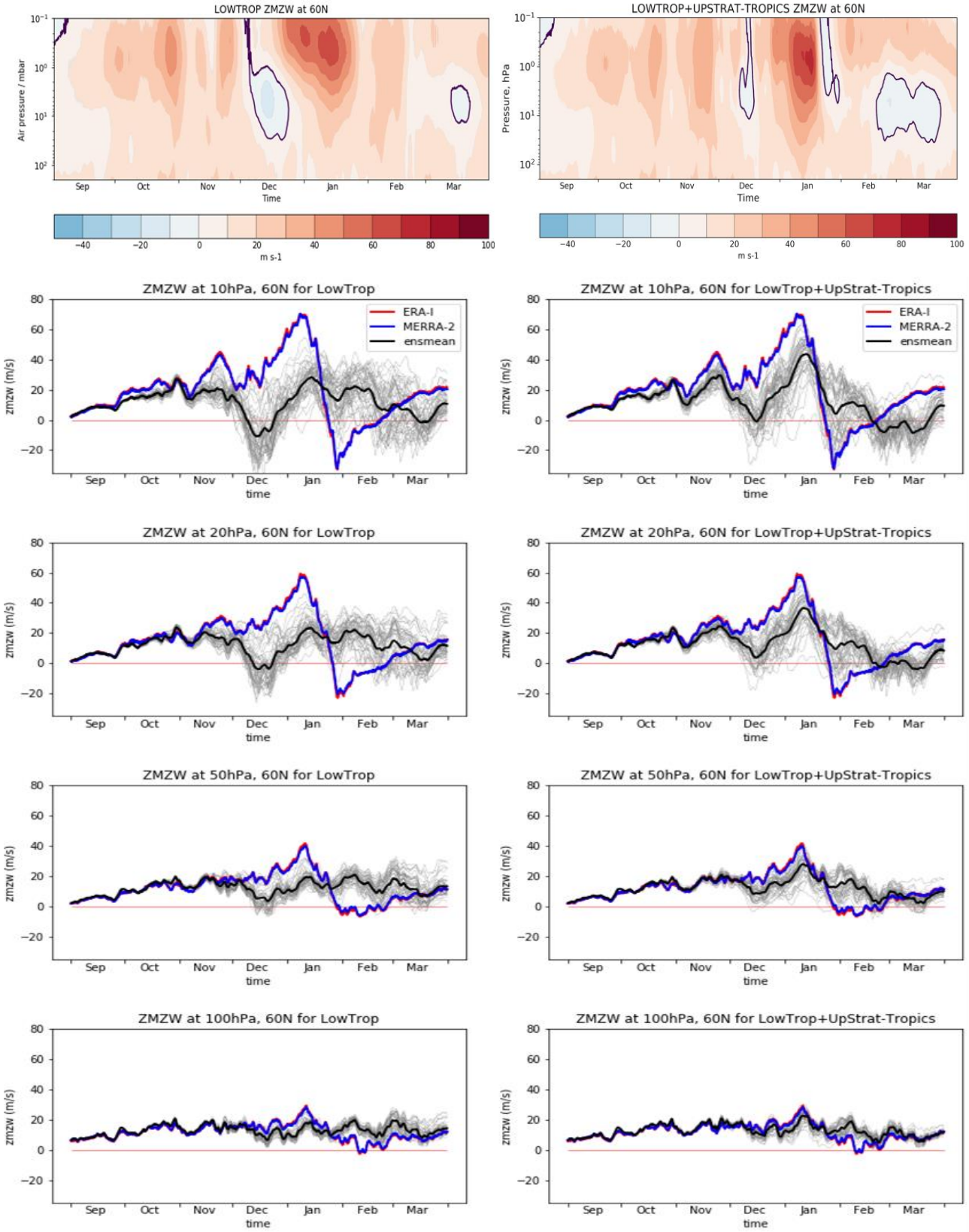


Figure 14: As Figure 11 but for the LowTrop (left) and LowTrop+UpStrat-Tropics (right) experiments.

There is distinctly more of a warming in January in **LowTrop+UpStrat-Tropics** compared to **LowTrop** with the winds in the upper stratosphere now reversing to easterly. A reduction in the strength of the westerlies at the start of the warming can be seen at all levels whereas in **LowTrop** it is not present beneath approximately 20 mb.

This experiment shows that the Northern tropics and subtropics of the upper stratosphere appears to be at least partly responsible for the timing of the start of the warming, as well as the initial few days of the evolution of the warming, as this aspect of the warming is simulated accurately without the barotropic influence of the polar upper stratosphere. It also plays some role in reducing the spurious warming during November and December that knocks the evolution of the stratosphere off-course in **LowTrop**.

AllTrop+UpStrat-Tropics: Next I wanted to see whether I obtained similar results when I used the AllTrop nudging region instead of LowTrop, i.e. comparing **AllTrop** against a simulation where I add in the northern tropical and subtropical upper stratosphere (**AllTrop+UpStrat-Tropics**). In the **AllTrop** experiments I am prescribing the wave forcing the stratosphere receives from the troposphere whereas in the **LowTrop** experiments there is some variability in the wave forcing received by the stratosphere. This experiment is shown in Figure 15 (right panel), along with the **AllTrop** results from Figure 11 for comparison.

The addition of the nudging in the northern tropics of the upper stratosphere seems to have a remarkable effect. The model is able to replicate the evolution of the stratosphere and the SSW at all levels, correctly simulating the timing, strength, speed and depth of the warming at all levels in the stratosphere at 60 °N. There is some slight deviation of the model from the

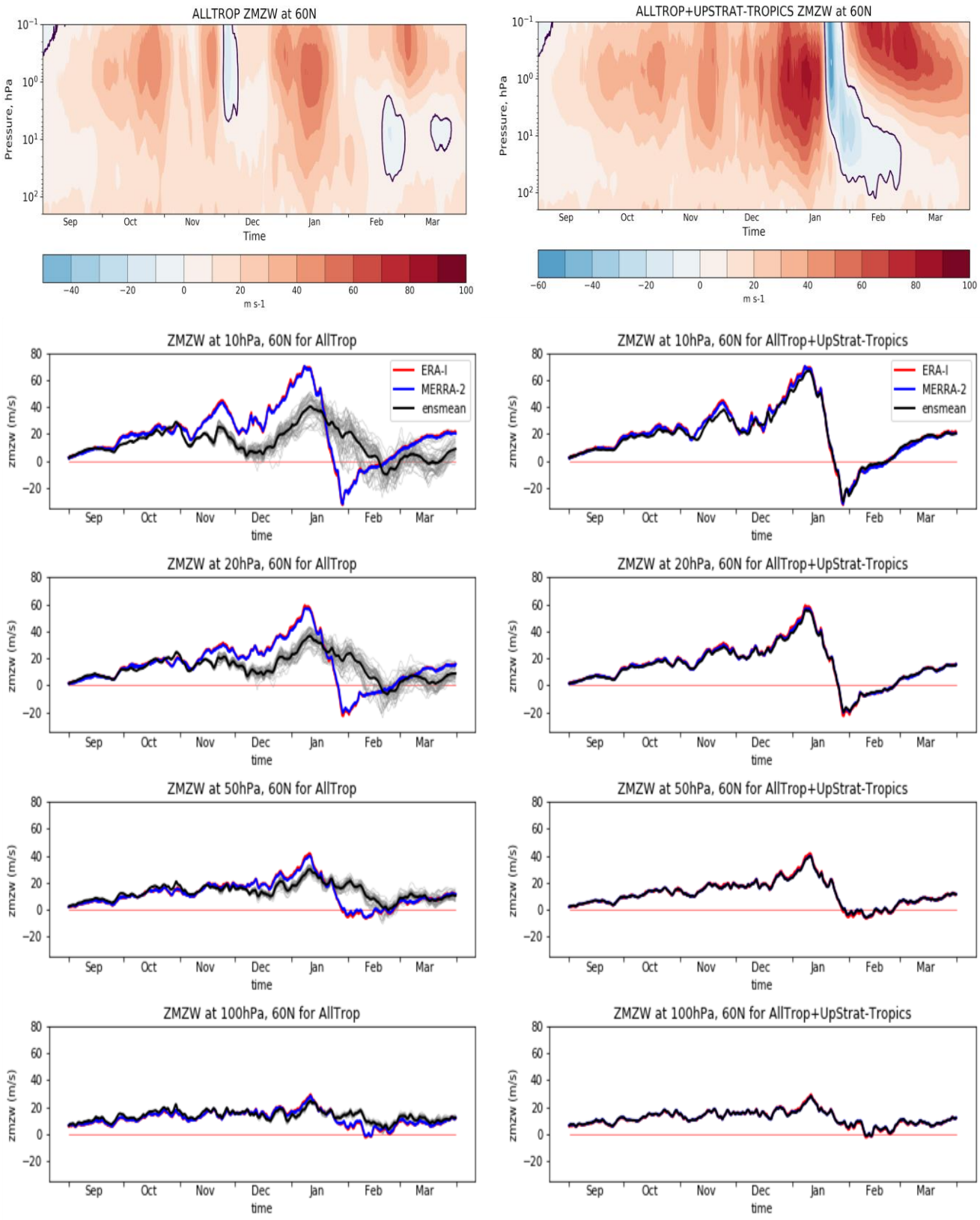


Figure 15: As Figure 11 but for the AllTrop (left) and AllTrop+UpStrat-Tropics (right) experiments.

reanalysis around the end of October/start of November, which is where the model has started to deviate substantially from the reanalysis in most of the other simulations.

Otherwise, it tracks the reanalysis close to perfectly, with negligible spread between the ensemble members. At 100 mb this is not so surprising as we are close to the troposphere nudging region, but at the upper levels this is a remarkable result.

This experiment shows that when we are prescribing the forcing of the stratosphere from the troposphere we only need additional nudging in the northern tropical and subtropical upper stratosphere to get an almost perfect simulation of the winter stratosphere and SSW, without the need for barotropic control from the polar upper stratosphere. The lack of spread between the different ensemble members is particularly surprising, as it indicates that the modelled polar vortex has very little internal variability and is not sensitive to any barotropic forcing from above or variations in the model QBO in the lower equatorial stratosphere (which is not directly constrained by the nudging in these runs), and thus variations in wave breaking in the lower stratosphere. We would expect there to be some variation between members from at least some combination of the above! It is likely that in some way the model polar vortex is too strongly constrained by the nudging extending out to around 40 °N (once the latitudinal taper has been accounted for).

Nonetheless, this result does confirm the hints from the **LowTrop+UpStrat-Tropics** experiment that the northern tropics and subtropics of the upper stratosphere are important for the polar vortex evolution. To start to understand why, we first need to look at the

difference between these two experiments in the UpStrat-Tropics region to see how the nudging is forcing the model here.

4.2.5: Analysis of Tropical Stratosphere Sensitivity Experiments

There is a large difference between the **AllTrop** and **AllTrop+UpStrat-Tropics**, where the UpStrat-Tropics region is constrained to MERRA2. Somehow, the differences introduced in the model in this region must be responsible for the marked improvement in the simulation of the warming in this year at least. By looking at the location and structure of these differences, and how they evolve, we can start to investigate the mechanism behind their effect on the polar vortex.

Shown in Figure 16 are monthly-mean latitude-height plots of the difference between **AllTrop** and **AllTrop+UpStrat-Tropics**. What becomes clear is that the structure of the difference between the two is more complex than a simple westerly or easterly anomaly in the equatorial upper stratosphere. There is a large easterly anomaly present over the equator above ~0.5 mb in September. This anomaly appears to strengthen and grow poleward and downward into to the polar vortex region from October-December, which is consistent with the 'upper stratosphere Holton-Tan' relationship identified by Pascoe et al. (2006) and previous studies, where a more easterly upper stratosphere in early winter is associated with a weaker polar vortex. It is possible that the more easterly winds might affect the wave forcing propagating through the stratosphere in this region - if the winds are anomalously easterly then some of the planetary wave activity from the troposphere may be diverted away from this region toward the pole, thus weakening the vortex too early in the winter, which is what is seen in the **AllTrop** ensemble.

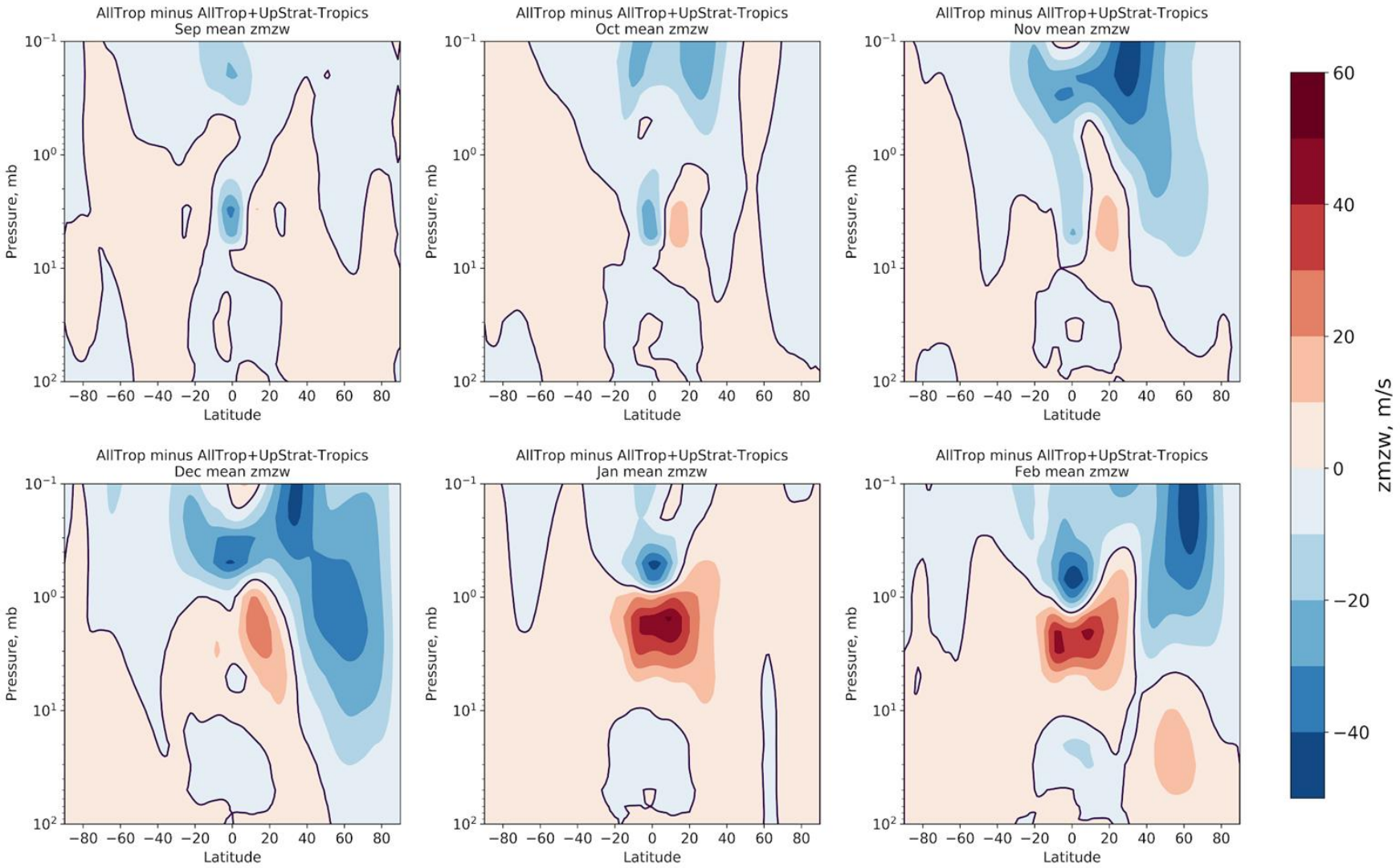


Figure 16: Latitude-height section of the monthly mean zonal mean zonal wind (m/s) of the AllTrop ensemble mean minus that of AllTrop+UpStrat-Tropics for September (top left) through February (bottom right)

During the month of the warming itself (January) the anomaly in the polar vortex region switches sign as the winds in **AllTrop+UpStrat-Tropics** become easterly. The anomaly is not particularly large as the warming only happens towards the end of January and is largely averaged out. In February a large easterly anomaly reappears here as a result of the winds having recovered after the warming in **AllTrop+UpStrat-Tropics** that does not happen as strongly in the **AllTrop** experiment.

A large easterly anomaly also develops in the equatorial upper stratosphere/lower mesosphere in January and February, particularly in the Northern Hemisphere, between about 5 and 0.3 mb. It consists of a strong (∼40 m/s) easterly anomaly atop a similarly strong westerly anomaly, which is indicative of a difference in the phase of the Semi-Annual Oscillation (SAO) between the two experiments. Figure 17 compares the zonal mean zonal winds at the equator from the **AllTrop** (bottom plot) and **AllTrop+UpStrat-Tropics** experiment (top plot (although note MERRA2 is shown, because this region of the atmosphere is being nudged to MERRA2 and is therefore identical)). We can see the stronger easterly phase of the model SAO in the **AllTrop** experiment (bottom plot) in November. This progresses slowly downward until at the start of January the descent stalls, and the easterly winds remain separate from the easterly winds of the top of the QBO. In the **AllTrop+UpStrat-Tropics** experiment (top plot) the descent of strong easterly winds is not seen. Instead we see a rapid descent of weak easterly winds at the start of December which then quickly strengthen at the start of January, merging with the easterly winds of the top of the QBO below and continuing to descend. This is consistent with the model climatological SAO not

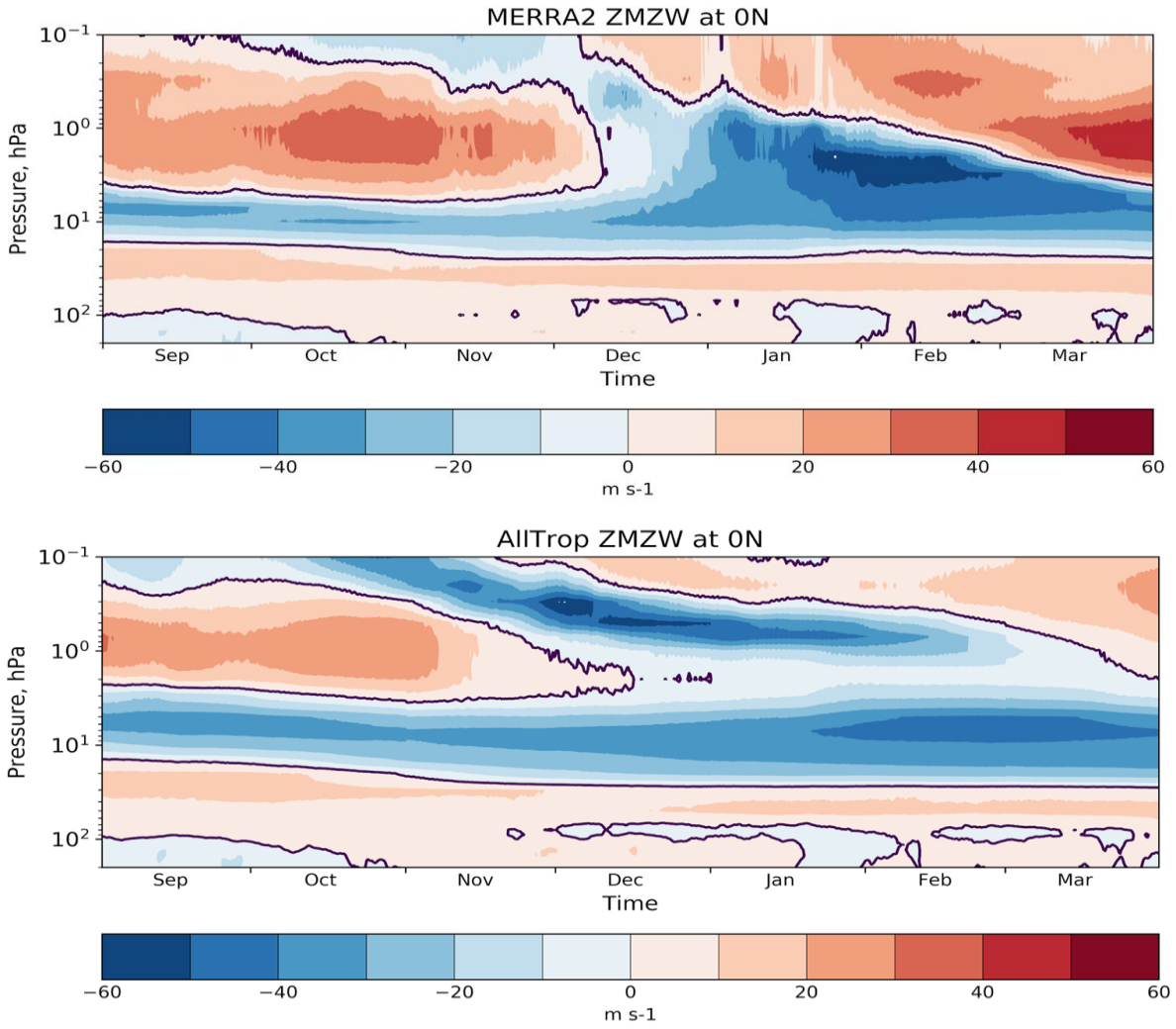


Figure 17: Time-height section of zonal mean zonal wind (m/s) at the equator (zero degrees latitude) for MERRA2 (top) and the AllTrop experiment (bottom). Red colours represent westerly winds, blue represent easterly. The black contour is the zero wind line. The tick marks on the x axis represent the month boundaries.

extending as deep as that of MERRA2 and having a stronger easterly phase (Figure 8, section 3.2, and section 3.1). It is likely that when the model is not constrained by the nudging here the phase of its SAO quickly drifts back towards its climatological state, since the radiative timescales at these levels is quite short (of order days or less).

Throughout the first 3 months the **AllTrop** experiment is therefore more easterly in the SAO region. We have seen from most of the previous experiments that the model starts to diverge from the reanalysis around the start of November, including in **AllTrop**. An easterly anomaly in the model before and at this time is consistent with the Holton-Tan idea of waves being unable to propagate into this region and diverted towards the polar vortex, causing a weakening of the winds or even an early warming (**LowTrop**, Figure 11 (left)) in all the experiments where this part of the stratosphere is not constrained. The time-height plot for MERRA2 (Figure 9 bottom right, section 4.1) shows a brief easterly blip very high up (around 0.1 mb) towards the end of November, which develops downward into a warming in the **LowTrop, AllTrop and LowTrop+UpStrat-Tropics** experiments, where the upper polar stratosphere is not constrained. SSWs and smaller vortex disturbances propagate downward through a wave-mean flow interaction (see section 2.4), so are less likely to develop when the wave forcing is not present or as strong. Therefore weaker wave forcing in the real atmosphere and the reanalysis at this time could be an explanation for why this easterly blip does not develop into a warming. In the experiments where the troposphere is not fully constrained the wave forcing from the troposphere in the model may be stronger, encouraging the downward development of the easterly winds. However, the easterly wind does develop downward a bit in **AllTrop**, but not as much as in **LowTrop**, as the model winds are simply constrained to the reanalysis higher up in **AllTrop**, so the model physically cannot develop the easterly winds below the tropopause as there are no easterly winds in the reanalysis here. This points to the model wanting to develop these easterly winds regardless

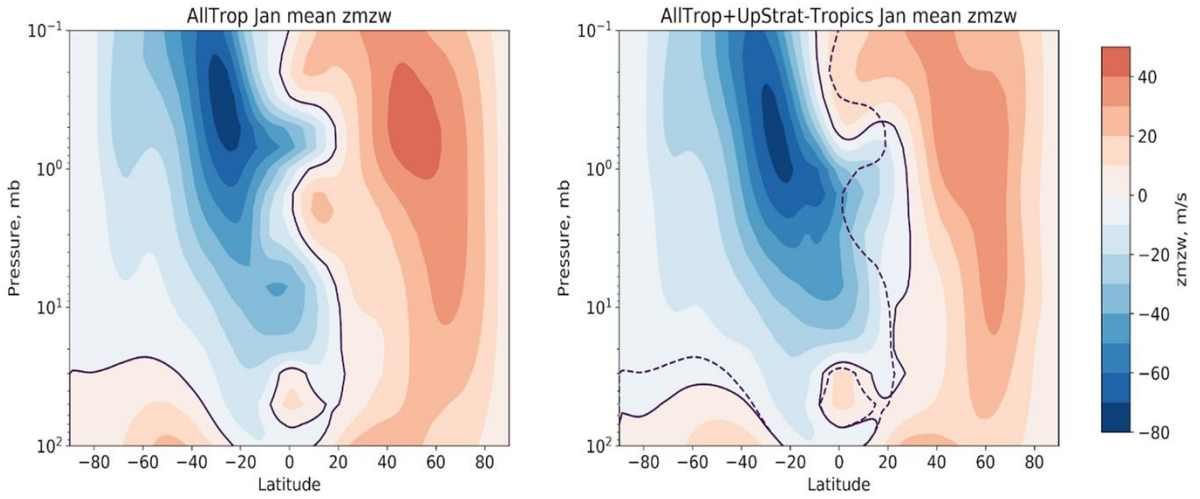


Figure 18: Latitude-height section of January mean zonal mean zonal wind (m/s) for the AllTrop (left) and AllTrop+UpStrat-Tropics (right) experiments. Red colours represent westerly winds, blue represent easterly. The solid black contour is the zero wind line for each plot. The dashed black contour on the right hand plot is the zero wind line of the left hand plot.

of the forcing from the troposphere, with only forcing in the stratosphere in this region preventing it from doing so, as again the easterly winds are not present in the reanalysis apart from the very tiny blip in the mesosphere.

In Figure 18 which shows the January average zonal mean zonal wind for the **AllTrop** and **AllTrop+UpStrat-Tropics** experiments, we can see the equatorward shift in the zero-wind line between **AllTrop** and **AllTrop+UpStrat-Tropics** between 1 and 8 mb as a result of the differing SAO in January. This is likely to affect the vertically propagating planetary wave activity, with some of the activity which would usually head equatorward through this region now being reflected by the zero-wind line back towards the pole. The increased wave breaking on the edge of the polar vortex would then mean an SSW is more likely, and lead to a faster descent of the easterly winds were one to occur, consistent with what we see happening in the **AllTrop** and **AllTrop+UpStrat-Tropics** experiments.

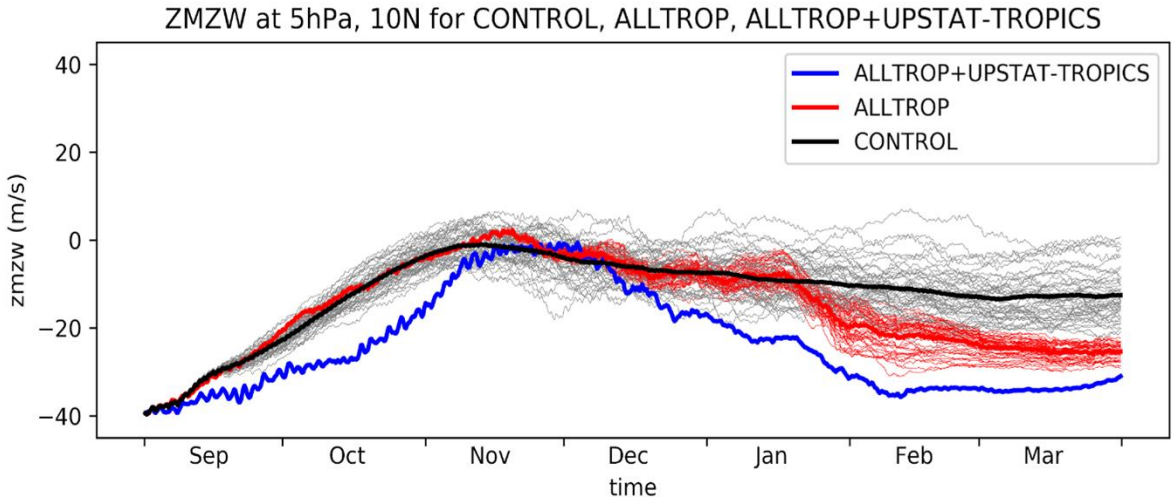


Figure 19: Zonal mean zonal wind (m/s) at 5 hPa, 10 degrees north for September through March of each ensemble member of the Control (grey lines), AllTrop (red lines) and AllTrop+UpStrat-Tropics (blue lines).

The zonal mean zonal wind for 5 mb, 10 °N (within the UpStrat-Tropics nudging region) is shown in Figure 19 for the **Control**, **AllTrop** and **AllTrop+Upstrat-Tropics** experiments. This figure shows that the **AllTrop** experiment remains closer to the **Control**, particularly during the first 3 months. This was also found to be the case throughout the UpStrat-Tropics nudging region (not shown). This indicates that the tropospheric nudging, i.e. the wave forcing, is not affecting the evolution of the model equatorial upper stratosphere. This indicates that the inability of the model to correctly simulate the tropical upper stratosphere (relative to MERRA2) without nudging is not a result of inaccuracies in the troposphere, but differences in the model tropical upper stratosphere, particularly the SAO. It seems the model wants to go a particular way in this region, regardless of the wave forcing. This is most likely because the SAO is determined primarily by the advection of easterlies from the summer hemisphere (in the easterly phase) and the momentum transfer associated with equatorial wave driving (westerly phase). It is possible that the tropospheric nudging in **AllTrop** (and similar

experiments) could affect the former via the indirect wave forcing of the strength of the Brewer-Dobson circulation, or the latter via modification of the tropospheric wave driving, but it appears that this is not the case.

Overall, there are two main anomalies in the model corrected by the nudging in the UpStrat-Tropics region, an easterly anomaly in the equatorial lower mesosphere and one lower down around 5 mb. Both of these are consistent with a weaker polar vortex through planetary wave activity in **AllTrop+UpStrat-Tropics**. These anomalies are likely a result of the large difference between the SAO of the model and MERRA2.

The climatology of the model SAO region is very different to that of MERRA2 as discussed in sections 3.1 and 3.2**.** Whilst MERRA2 might be quite different from other reanalyses in this region it is clearly 'right enough' compared to the model to yield this accurate simulation of the polar vortex.

Section 5: Summary and Conclusions

This study has found that prescribing the wave forcing from the troposphere is not enough to accurately simulate the evolution of the winter polar vortex of 2008/9. This indicates that the internal dynamics of the stratosphere plays a role and that the wave forcing from the troposphere is a necessary but insufficient condition for accurately simulating the stratosphere. For example, only the experiments with nudging regions in both the troposphere and the stratosphere are able to get close to the observed strength and depth of the warming.

With tropospheric nudging up to only 6 km the ensemble polar vortex evolution at 10 mb, 60 °N was more tightly constrained than in the control ensemble that had only SST forcing at the lower boundary, but it was nevertheless unable to capture the timing or characteristics of the observed SSW. When the tropospheric nudging was extended up to the tropopause level the evolution was substantially improved, with a warming event that started at the correct time. This highlights the importance of an accurate simulation of the wave dynamics in the upper troposphere in the vicinity of the tropopause. However, the simulated SSW was still much weaker and slower than that observed.

The SSW in this particular study could also be triggered at the correct time by nudging towards observations only in the upper stratosphere, although the SSW did not extend as deep as was observed. However, when the upper stratospheric nudging was combined with tropospheric forcing up to 6 km the resulting polar vortex evolution was remarkably realistic. Subsequent experiments highlighted the important role of the equatorial/subtropical region of the upper stratosphere (above 5 mb, equatorward of 30 °N). This region is therefore important for initiating the SSW in 2008/9, but accurate tropospheric wave forcing is also required for the warming to continue propagating downwards.

It is also important to note that the influences on the polar vortex of the separate nudging regions do not combine linearly, in other words it is not as simple as 'adding' the effect of, for example, the upper stratosphere to that of the troposphere to get the influence shown in the experiment with both. Each region of the atmosphere affects different aspects of the polar vortex evolution and with different strengths.

In summary, the most striking conclusion from this study is that with fixed, accurate wave forcing prescribed from the troposphere, the addition of nudging to observations in the tropical upper stratosphere is all that is needed for an accurate simulation of the polar vortex evolution. The 'top-down' triggering of the warming from the region directly overhead at polar latitudes, seen in the **UpStrat** experiment, is not required. The upper tropical stratosphere must therefore play a role in modifying the forcing of the polar vortex from the troposphere, as suggested by Gray (2003) and subsequent studies (see section 2.4).

Preliminary analysis of the impact of the nudging in the upper equatorial stratosphere highlighted the possible role of the semi-annual oscillation (SAO). This points to the potential for improving the forecasting of SSWs if a more accurate simulation of this previously neglected region could be achieved. The role of the SAO appears to be especially important in early winter when the polar vortex is established and in the timing of any subsequent SSWs. However, further analysis is required to fully understand the mechanisms involved.

5.1: Further Work

The barotropic nature of the winter polar vortex circulation in the stratosphere means that conclusions from this 2008/9 case study are likely valid for warming events in other years, but it would be important to verify this by repeating a selection of the experiments for many other years i.e. 1979-onwards if possible. Wave-flux diagnostics are also needed to verify the statements made about the effects of the nudging on the wave forcing. One concern for extending this experiment to earlier years is the lack of MLS data assimilated into MERRA2. Although the assimilation of MLS should make the upper level winds in MERRA2 more

accurate, I did find that the climatology of the SAO region did not vary much between the years in which MLS data was and wasn't assimilated (Figure 8). Further investigation is therefore needed to see if the improved simulation is dependent on nudging towards MERRA2 that includes MLS or whether another reanalysis (such as ERA-Interim) would give similarly improved results.

As mentioned in section 4.2.4, the lack of ensemble spread in the AllTrop+UpStrat-Tropics experiment is surprising. One possible reason for this is a strong constraint applied by the nudging extending out to 40 °N. To confirm this it would be useful to run an identical experiment with the nudging region latitudinally confined to, say, $\pm$10 °. It would also be useful to see how strongly the model QBO is constrained by the nudging above it, as a lack of ensemble variation in the QBO would be another possible contributor to the lack of variation in the polar vortex.

It would also be interesting to examine whether there are distinctions between the two main types of SSWs: split (like the one in 2008/9) and displacement (where the vortex is pushed away from the pole), as Mitchell et al. (2013) found that the circulation associated with split SSWs was more barotropic than that of displacement SSWs. It is possible that the SAO region may be more important in one type of warming than the other.

Analysis of the tropical upper stratosphere and lower mesosphere showed that there are large differences between the model and reanalysis in 2008/9. Whilst it is difficult to say with any certainty that any one reanalysis has a more accurate representation of the SAO region it is clear that at the very least the model SAO has an easterly bias, with westerly phases that

are too weak. It also lacks the annual cycle component of the SAO winds. This bias contributes to an inaccurate simulation of the 2008/9 warming. The cause of this bias also requires further investigation.

A further experiment where the SAO region is relaxed towards a climatological representation of the SAO, rather than to the observed SAO, would remove the underlying easterly bias in the model and help shed light on whether it is the climatology of the SAO that is important for accurately simulating a warming, or the specifics of any one year (the interannual variability). The climatology that is relaxed to could also be varied to further examine the sensitivity of the polar vortex to this region, and ascertain which of the climatologies (including that of the model) yields the most accurate simulation of the polar vortex. Similarly, adjusting details of the east/west phases of the SAO and the timing of the phase change could help identify which aspects of the SAO are important.

Finally, the nudging of the SAO region also had an impact on the region immediately below it, in the height region dominated by the QBO, so it would be interesting to consider whether experiments could be designed to test the direct impact of the SAO region on the polar vortex evolution compared with its indirect impact via the QBO.

References

Andrews, D. G., and M. E. McIntyre, 1976: Planetary Waves in Horizontal and Vertical Shear: The Generalized Eliassen-Palm Relation and the Mean Zonal Acceleration. *J. Atmos. Sci.*, **33**, 2031–2048, https://doi.org/10.1175/1520-0469(1976)033<2031:PWIHAV>2.0.CO;2.

——, J. R. Holton, and C. B. Leovy, 1987a: *Middle Atmosphere Dynamics*. Academic Press,.

——, ——, and ——, 1987b: Planetary-Wave Critical Layers. *Middle Atmosphere Dynamics*, 253–257.

——, ——, and ——, 1987c: Linear Wave Theory: Equatorial Waves. *Middle Atmosphere Dynamics*, 200–215.

Angell, J. K., and J. Korshover, 1964: Quasi-Biennial Variations in Temperature, Total Ozone, and Tropopause Height. *J. Atmos. Sci.*, **21**, 479–492, https://doi.org/10.1175/1520-0469(1964)021<0479:QBVITT>2.0.CO;2.

Baldwin, M. P., and T. J. Dunkerton, 1991: Quasi-biennial oscillation above 10mb. *Geophys. Res. Lett.*, **18**, 1205–1208.

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric Harbingers of Anomalous Weather Regimes. *Science (80-.).*, **294**, 581–584, https://doi.org/10.1126/science.1063315.

Baldwin, M. P., and Coauthors, 2001: THE QUASI-BIENNIAL OSCILLATION. 179–229.

Baldwin, M. P., D. W. J. Thompson, E. F. Shuckburgh, W. A. Norton, and N. P. Gillett, 2003: ATMOSPHERIC SCIENCE: Weather from the Stratosphere? *Science (80-.).*, **301**, 317–319, https://doi.org/10.1126/science.1085688.

Brasseur, G. P., and S. Soloman, 2005: *Aeronomy of the Middle Atmosphere*. Springer-Verlag, 646 pp. http://link.springer.com/10.1007/1-4020-3824-0.

Brown, A., S. Milton, M. Cullen, B. Golding, J. Mitchell, and A. Shelly, 2012: Unified modeling and prediction of weather and climate: A 25-year journey. *Bull. Am. Meteorol. Soc.*, **93**, 1865–1877, https://doi.org/10.1175/BAMS-D-12-00018.1.

Bushell, A. C., D. R. Jackson, N. Butchart, S. C. Hardiman, T. J. Hinton, S. M. Osprey, and L. J. Gray, 2010: Sensitivity of GCM tropical middle atmosphere variability and climate to ozone and parameterized gravity wave changes. *J. Geophys. Res. Atmos.*, **115**, 1–14, https://doi.org/10.1029/2009JD013340.

Calvo, N., M. A. Giorgetta, and C. Peña-Ortiz, 2007: Sensitivity of the boreal winter circulation in the middle atmosphere to the quasi-biennial oscillation in MAECHAM5 simulations. *J. Geophys. Res. Atmos.*, **112**, 1–13, https://doi.org/10.1029/2006JD007844.

Charney, J. G., and P. G. Drazin, 1961: Propagation of Planetary-Scale Disturbances from the Lower into the Upper Atmosphere. **66**, 83–109.

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

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the

data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597, https://doi.org/10.1002/qj.828.

Dunkerton, T. J., 1997: The role of gravity waves in the quasi-biennial oscillation. *J. Geophys. Res. Atmos.*, **102**, 26053–26076, https://doi.org/10.1029/96jd02999.

Ebdon, R. A., 1960: Notes on the wind flow at 50 mb in tropical and sub-tropical regions in January 1957 and January 1958. *Q. J. R. Meteorol. Soc.*, **86**, 540–542, https://doi.org/10.1002/qj.49708637011.

Eliassen, A., and E. Palm, 1960: On the Transfer of Energy in Stationary Mountain Waves. **22**, https://doi.org/10.1098/rstl.1884.0016.

Fujiwara, M., and Coauthors, Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the renalysis systems. 1–52.

Garcia, R. R., and F. Sassi, 1999: Modulation of the mesospheric semiannual oscillation by the quasibiennial oscillation. *Earth, Planets Sp.*, **51**, 563–569, https://doi.org/10.1186/BF03353215.

——, T. J. Dunkerton, R. S. Lieberman, and R. A. Vincent, 1997: Climatology of the semiannual oscillation of the tropical middle atmosphere. *J. Geophys. Res. Atmos.*, **102**, 26019–26032, https://doi.org/10.1029/97JD00207.

Garfinkel, C. I., 2012: Does the Holton – Tan Mechanism Explain How the Quasi-Biennial Oscillation Modulates the Arctic Polar Vortex ? *J. Atmos. Sci.*, 1713–1733, https://doi.org/10.1175/JAS-D-11-0209.1.

Gelaro, R., and Coauthors, 2017: The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.*, **30**, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1.

Gray, L. J., 2003: The influence of the equatorial upper stratosphere on stratospheric sudden warmings. **30**, 30–33, https://doi.org/10.1029/2002GL016430.

Gray, L. J., E. F. Drysdale, B. N. Lawrence, and T. J. Dunkerton, 2001a: Model studies of the interannual variability of the northern-hemisphere stratospheric winter circulation: The role of the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.*, **127**, 1413–1432, https://doi.org/10.1002/qj.49712757416.

——, J. Phipps, T. J. Dunkerton, M. P. Baldwin, R. Allen, and E. F. Drysdale, 2001b: A data study of the influence of the equatorial upper stratosphere on northern-hemisphere stratospheric sudden warmings. *Q. J. R. Meteorol. Soc.*, **127**, 1985–2003.

——, S. Sparrow, M. Juckes, A. O'Neill, and D. G. Andrews, 2003: Flow regimes in the winter stratosphere of the northern hemisphere. *Q. J. R. Meteorol. Soc.*, **129**, 925–945, https://doi.org/10.1256/qj.02.82.

Gray, L. J., S. Crooks, C. Pascoe, S. Sparrow, and M. Palmer, 2004: Solar and QBO Influences on the Timing of Stratospheric Sudden Warmings. *J. Atmos. Sci.*, **61**, 2777–2796, https://doi.org/10.1175/JAS-3297.1.

Gray, L. J., and Coauthors, 2010: Solar influence on climate. *Rev. Geophys.*, **48**, RG4001, https://doi.org/10.1029/2009RG000282.1.INTRODUCTION.

Hamilton, K., 1998: Effects of an imposed quasi-biennial oscillation in a comprehensive troposphere

- stratosphere - mesosphere general circulation model. *J. Atmos. Sci.*, **55**, 2393–2418, https://doi.org/10.1175/1520-0469(1998)055<2393:EOAIQB>2.0.CO;2.

Hardiman, S. C., N. Butchart, S. M. Osprey, L. J. Gray, A. C. Bushell, and T. J. Hinton, 2010: 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, https://doi.org/10.1175/2009JAS3337.1. http://journals.ametsoc.org/doi/abs/10.1175/2009JAS3337.1.

Haynes, P., 2005: Stratospheric Dynamics. *Annu. Rev. Fluid Mech.*, **37**, 263–293, https://doi.org/10.1146/annurev.fluid.37.061903.175710. http://www.annualreviews.org/doi/10.1146/annurev.fluid.37.061903.175710.

Haynes, P. H., 1998: The latitudinal structure of the quasi-biennial oscillation. *Q. J. R. Meteorol. Soc.*, **124**, 2645–2670, https://doi.org/10.1002/qj.49712455206.

Hirota, I., T. Hirooka, and Shiotani, 1983: Upper stratospheric circulations in the two hemispheres observed by satellites. *Q. J. R. Meteorol. Soc.*, **109**, 443–454.

Hitchman, M. H., and C. B. Leovy, 1985: Diurnal tide in the equatorial middle atmosphere as seen in LIMS temperatures.

——, and L. B. Conway, 1988: Estimation of the Kelvin Wave Contribution to the Semiannual Oscillation. *J. Atmos. Sci.*, **45**, 1462–1474.

Hoinka, K. P., 1999: Temperature, Humidity, and Wind at the Global Tropopause. *Mon. Weather Rev.*, **127**, 2248–2265, https://doi.org/10.1175/1520-0493(1999)127<2248:THAWAT>2.0.CO;2.

Holton, J. R., and R. S. Lindzen, 1972: An Updated Theory for the Quasi-Biennial Cycle of the Tropical Stratosphere. *J. Atmos. Sci.*, **29**, 1076–1080, https://doi.org/10.1175/1520-0469(1972)029<1076:AUTFTQ>2.0.CO;2.

Holton, J. R., and H.-C. Tan, 1980: The influence of the equatorial QBO on sudden stratospheric warmings.

Holton, J. R., and H.-C. Tan, 1982: The Quasi-Biennial Northern Hemisphere Oscillation in the Lower Stratosphere. *J. Meteorol. Soc. Japan*, 140–148.

Holton, J. R., G. J. Hakim, J. R. Holton, and G. J. Hakim, 2013: Chapter 11 – Tropical Dynamics. *An Introduction to Dynamic Meteorology*, 377–411.

Hu, Y., and K. K. Tung, 2002: Tropospheric and equatorial influences on planetary-wave amplitude in the stratosphere. *Geophys. Res. Lett.*, **29**, 1–6, https://doi.org/Artn 1019\rDoi 10.1029/2001gl013762.

Iida, C., T. Hirooka, and N. Eguchi, 2014: Circulation changes in the stratosphere and mesosphere during the stratospheric sudden warming event in January 2009. *J. Geophys. Res. Atmos.*, **119**, 7104–7115, https://doi.org/10.1002/2013JD021252.

Jeuken, A. B. M., P. C. Siegmund, L. C. Heijboer, J. Feichter, and L. Bengtsson, 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. **101**.

Kinnersley, J. S., and S. Pawson, 1996: The Descent Rates of the Shear Zones of the Equatorial QBO.

J. Atmos. Sci., **53**, 1937–1949, https://doi.org/10.1175/1520-0469(1996)053<1937:TDROTS>2.0.CO;2.

Labitzke, K., 1977: Interannual Variability of the Winter Stratosphere in the Northern Hemisphere. *Mon. Weather Rev.*, **105**, 762–770, https://doi.org/10.1175/1520-0493(1977)105<0762:IVOTWS>2.0.CO;2.

——, 1987: Sunspots, the qbo, and the stratospheric temperature in the north polar region. **14**, 535–537.

Lindzen, R. S., and J. R. Holton, 1968: A Theory of the Quasi-Biennial Oscillation. *J. Atmos. Sci.*, **25**, 1095–1107, https://doi.org/10.1175/1520-0469(1968)025<1095:ATOTQB>2.0.CO;2.

Livesey, N. J., and Coauthors, 2011: Earth Observing System (EOS), Aura Microwave Limb Sounder (MLS), Version 3.3 Level 2 data quality and description document. *Tech. Rep. JPL D-33509.*, https://doi.org/Tech. Rep. JPL D-33509.

Long, C. S., M. Fujiwara, S. Davis, D. M. Mitchell, and C. J. Wright, 2017: Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP). *Atmos. Chem. Phys.*, **17**, 14593–14629, https://doi.org/10.5194/acp-17-14593-2017.

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. Atmos.*, **113**, 1–14, https://doi.org/10.1029/2007JD009647.

Matsuno, T., 1970: Vertical Propagation of Stationary Planetary Waves in the Winter Northern Hemisphere. *J. Atmos. Sci.*, **27**, 871–883, https://doi.org/10.1175/1520-0469(1970)027<0871:VPOSPW>2.0.CO;2.

Mitchell, D. M., L. J. Gray, J. Anstey, M. P. Baldwin, and A. J. Charlton-Perez, 2013: The influence of stratospheric vortex displacements and splits on surface climate. *J. Clim.*, **26**, 2668–2682, https://doi.org/10.1175/JCLI-D-12-00030.1.

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister, 2015: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2. *Geosci. Model Dev.*, **8**, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015.

Naoe, H., and K. Shibata, 2010: Equatorial quasi - biennial oscillation influence on northern winter extratropical circulation. *J. Geophys. Res.*, **115**, 1–13, https://doi.org/10.1029/2009JD012952.

Naujokat, B., 1986: An Update of the Observed Quasi-Biennial Oscillation of the Stratospheric Winds over the Tropics. *J. Atmos. Sci.*, **43**, 1873–1877.

Osprey, S. M., L. J. Gray, S. C. Hardiman, N. Butchart, A. C. Bushell, and T. J. 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.*, **67**, 3637–3651, https://doi.org/10.1175/2010JAS3338.1.

Pascoe, C. L., L. J. Gray, S. A. Crooks, M. N. Juckes, and M. P. Baldwin, 2005: The quasi-biennial oscillation: Analysis using ERA-40 data. *J. Geophys. Res. D Atmos.*, **110**, 1–13, https://doi.org/10.1029/2004JD004941.

——, L. J. Gray, and A. A. Scaife, 2006: A GCM study of the influence of equatorial winds on the timing of sudden stratospheric warmings. *Geophys. Res. Lett.*, **33**, 33–36, https://doi.org/10.1029/2005GL024715.

Plumb, R. A., 1984: The quasi-biennial oscillation. *Dynamics of the Middle Atmosphere*, J.R. Holton and T. Matsuno, Eds., Terrapub, Tokyo, 217–251.

Quiroz, R., 1975: The Stratospheric Evolution of Sudden Warmings in 1969-74 Determined from Measured Infrared Fields. *J. Atmos. Sci.*, **32**, https://doi.org/10.1175/1520-0469(1975)032<0211:TSEOSW>2.0.CO;2.

Reed, R. J., 1965: The Quasi-Biennial Oscillation of the Atmosphere Between 30 and 50 km Over Ascension Island. *J. Atmos. Sci.*, **22**, 331–333.

——, W. J. Campbell, L. a. Rasmussen, and D. G. Rogers, 1961: Evidence of a downward-propagating, annual wind reversal in the equatorial stratosphere. *J. Geophys. Res.*, **66**, 813, https://doi.org/10.1029/JZ066i003p00813.

Sanchez, C., K. D. Williams, and M. Collins, 2016: Improved stochastic physics schemes for global weather and climate models. *Q. J. R. Meteorol. Soc.*, **142**, 147–159, https://doi.org/10.1002/qj.2640.

Scaife, A., 2002: Impact of a Spectral Gravity Wave Parameterization on the Stratosphere in the Met Office Unified Model. 1473–1489.

Scaife, A. A., N. Butchart, C. D. Warner, D. Stainforth, and W. Norton, 2000: Realistic Quasi-Biennial Oscillations in a simulation of the global climate. **27**, 3481–3484.

Scherhag, R., 1952: Die explosionsartigen Stratosphärenerwärmungen des Spätwinter 1951/1952 (The explosive warmings in the stratosphere of the late winter 1951/1952). *Ber. Dtsch. Wetterdienstes U.S. Zo.*, **38**, 51–63.

Shepherd, T. G., 2000: The middle atmosphere. *J. Atmos. Solar-Terrestrial Phys.*, **62**, 1587–1601, https://doi.org/10.1016/S1364-6826(00)00114-0.

Shiotani, M., and I. Hirota, 1985: Planetary wave-mean flow interaction in the stratosphere: A comparison between northern and southern hemispheres. *Q. J. R. Meteorol. Soc.*, **111**, 309–334, https://doi.org/10.1002/qj.49711146404.

Sparrow, S. N., L. J. Gray, M. Juckes, A. O'Neill, and D. G. Andrews, 2004: Simulations of stratospheric flow regimes during northern hemisphere winter. *Adv. Sp. Res.*, **34**, 337–342, https://doi.org/10.1016/j.asr.2003.03.042.

Telford, P. J., P. Braesicke, O. Morgenstern, and J. A. Pyle, 2008: and Physics Technical Note : Description and assessment of a nudged version of the new dynamics Unified Model. 1701–1712.

Walters, D., and Coauthors, 2017: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations. *Geosci. Model Dev. Discuss.*, 1–78, https://doi.org/10.5194/gmd-2017-291.

Watson, P., and L. Gray, 2014: How does the quasi-biennial oscillation affect the stratospheric polar vortex? *J. Atmos. Sci.*, **71**, 391–409, https://doi.org/10.1175/JAS-D-13-096.1.

White, I. P., H. Lu, and N. J. Mitchell, 2016: Seasonal evolution of the QBO-induced wave forcing and circulation anomalies in the northern winter stratosphere. 411–431, https://doi.org/10.1002/2015JD024507.Received.