Review: the predictability of the extra-tropical stratosphere on monthly timescales and its impact on the skill of tropospheric forecasts

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Open Access


It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1002/qj.2432
To link to this article DOI: http://dx.doi.org/10.1002/qj.2432

Publisher: Royal Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law,
including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR
Central Archive at the University of Reading
Reading’s research outputs online
Extreme variability of the winter- and spring-time stratospheric polar vortex has been shown to affect extratropical tropospheric weather. Therefore, reducing stratospheric forecast error may be one way to improve the skill of tropospheric weather forecasts. In this review, the basis for this idea is examined. A range of studies of different stratospheric extreme vortex events shows that they can be skilfully forecasted beyond 5 days and into the sub-seasonal range (0–30 days) in some cases. Separate studies show that typical errors in forecasting a stratospheric extreme vortex event can alter tropospheric forecast skill by 5–7% in the extratropics on sub-seasonal time-scales. Thus understanding what limits stratospheric predictability is of significant interest to operational forecasting centres. Both limitations in forecasting tropospheric planetary waves and stratospheric model biases have been shown to be important in this context.

Key Words: stratospheric predictability; tropospheric forecast; seasonal predictability

Received 6 February 2014; Revised 12 July 2014; Accepted 16 July 2014; Published online in Wiley Online Library

1. Introduction

The skill of numerical weather prediction (NWP) on weekly to monthly time-scales is limited both by errors in atmospheric initial conditions provided by data assimilation, and by chaotic growth of errors in model forecasts launched from those initial conditions. For NWP model runs in real time, the additional constraint of limited computational resources forces modelling centres to prioritize model configurations that can most effectively reduce both types of error growth. In the past, and in some current NWP models, the top atmospheric level has conventionally been placed somewhere in the middle to upper stratosphere. These so-called low-top models were used, based on the assumption that the stratosphere did not contribute significantly to the predictability of surface conditions and therefore the stratosphere did not necessitate model computational resources.

Early efforts to extend the upper boundaries of NWP models were driven by the desire to reduce errors in atmospheric initial conditions (Lorenz, 1963). For example, microwave and infrared radiances acquired from nadir sounders on operational meteorological satellites have vertical weighting functions that typically peak at tropospheric or lower stratospheric altitudes, but have long tails that extend deep into the stratosphere. With the advent of operational radience assimilation, higher upper boundaries were needed in NWP systems to provide forecast backgrounds at all contributing altitudes, in order to accurately assimilate the temperature information contained in these radiances (Gerber et al., 2012). In this review, we will concern ourselves less with...
these and other influences of a well-resolved stratosphere in improving the accuracy of atmospheric initial conditions used by NWP systems, and more on how a well-resolved stratosphere improves NWP model forecasts of dynamical coupling pathways that in turn can lead to improved predictability of both the stratosphere and the troposphere.

Over the past 30 years, it has been increasingly recognized that during periods in which its state is far from its climatological norm the stratosphere can contribute significantly to extratropical tropospheric predictability and that forecasts might be improved by representing the stratosphere with greater fidelity in NWP models (e.g. Thompson and Wallace, 1998; Baldwin and Dunkerton, 1999, 2001; Kuroda and Kodera, 1999). In a recent review of the current state of seasonal and decadal forecasting skill of current operational NWP systems, Smith et al. (2012) highlighted the importance of stratospheric sudden warming (SSW) events as a potential source of additional predictability in long-range forecasts of cold winter weather in Europe and the eastern USA (e.g. Thompson et al., 2002; Marshall and Scaife, 2010).

In this article we assemble evidence that shows the extent to which extreme events in the extratropical stratosphere can be predicted, and quantify their potential impact on the tropospheric state. Though the main focus of this article is on major midwinter SSWs, we also consider the role of other relevant stratospheric extremes. The aim is to provide a clear picture of our understanding of the influence of the stratosphere on tropospheric predictability on time-scales up to 30 days covering sub-seasonal variability. We also discuss some sources of predictability which predominantly play a role on seasonal time-scales because often the boundaries between seasonal and sub-seasonal forecasts are close and these sources make contributions on both time-scales. The review is organized as follows. In the remainder of section 1, we briefly review the proposed mechanisms by which the stratosphere might influence tropospheric circulation. In section 2 we discuss the predictability of the stratosphere and how this has evolved as NWP models have increased in complexity with higher upper boundaries and finer horizontal and vertical resolution. Section 3 discusses the dynamical origins of stratospheric predictability. Finally, in section 4, we attempt to quantify the impact of the stratosphere on tropospheric predictability. We end the review with a discussion of current issues and ideas for future experiments.

There are a number of proposed mechanisms by which stratospheric variability might influence the troposphere. These can be broadly divided into three groups: (i) influences of the stratosphere in tropospheric baroclinic systems, (ii) large-scale adjustment in the troposphere to stratospheric potential vorticity anomalies, and (iii) planetary wave–mean flow interaction. Before discussing the impact of the stratosphere on tropospheric predictability, we briefly review the evidence underpinning each of these mechanisms.

Within the context of an idealized modelling study, Garfinkel et al. (2013) compared various mechanisms of the influence of the stratospheric vortex on the eddy-driven (midlatitude) tropospheric jet. Echoing the previous result of Song and Robinson (2004), they showed that, in order to explain the magnitude of tropospheric jet shifts in response to stratospheric perturbations, it was necessary to invoke purely tropospheric feedbacks between eddies and the jet. It is important to note that Garfinkel et al. did not benchmark their model with reanalyses to ensure that various mechanisms were present in their simulations. For example, they did not assess the role of planetary wave coupling, which has been linked to the position of the Atlantic jet stream (Shaw et al., 2014). The importance of their finding (also expressed by Song and Robinson (2004)), however, is to suggest that although the mechanisms listed below highlight visible dynamical coupling pathways linking the stratosphere and troposphere, the ultimate tropospheric outcome of the coupling remains strongly influenced by internal tropospheric processes.

1.1. Stratospheric influence on tropospheric baroclinic systems

Several different processes have been proposed whereby stratospheric changes influence the development or structure of tropospheric baroclinic systems. These are mostly related to the so-called index of refraction for Rossby waves (Matsuno, 1970), and include:

- influences on eddy phase speed (Chen and Held, 2007);
- influences on eddy length scales (Kidston et al., 2010; Rivièere, 2011);
- changes to the index of refraction for baroclinic systems (Simpson et al., 2012);
- changes to the structure of baroclinic systems leading to modified heat and momentum fluxes (Thompson and Birner, 2012);
- changes to the type of wave-breaking (Wittman et al., 2007; Kunz et al., 2009).

It is difficult to separate these different and possibly complementary effects, but Garfinkel et al. (2013) reviewed diagnostics for each of these effects independently in their idealized experiments. In particular, many of the processes were able to account for the nonlinear state dependence of their modelled response of the tropospheric jet to stratospheric perturbations in their experiments.

1.2. Large-scale adjustment in the troposphere in response to the stratospheric PV distribution

The second mechanism describes the balanced geostrophic and hydrostatic response of the tropospheric flow to stratospheric potential vorticity (PV) anomalies. As shown by Hoskins et al. (1985), a PV anomaly associated with a change in strength of the polar vortex leads to large-scale changes in the tropopause height as isentropic surfaces bend towards or away from a positive or negative PV anomaly, respectively. Ambaum and Hoskins (2002) calculate that about 10% change in the strength of the stratospheric jet leads to a 300 m change in the position of the Arctic tropopause height. These numbers obtained from theoretical calculations might not be realistic for the real atmosphere but they do highlight the importance of stratospheric variations on the tropospheric circulation patterns. Other studies (Hartley et al., 1998; Black, 2002) use piecewise PV inversion techniques to show, similarly, that lower stratospheric PV anomalies induce circulations in the upper troposphere of similar magnitude to those produced by purely tropospheric PV anomalies (Hartley et al., 1998) and that at least some of the variability of the tropospheric jet down to the surface is related to stratospheric PV anomalies (Black, 2002; Hines, et al., 2010).

1.3. Planetary wave–mean flow interaction

This third mechanism involves the fate of upward propagating planetary-scale waves due to wave–mean flow interaction in the stratosphere (Matsuno, 1970; Chen and Robinson, 1992; Song and Robinson, 2004; Harnik, 2009; Plumb, 2010). Whether the vertically propagating waves are reflected, propagated, or absorbed in a certain region of the atmosphere, depending on the zonal wind structure, is determined by the vertical part of the index of refraction squared ($N^2_{ref}$) (Harnik, 2009). If $N^2_{ref}$ is negative, waves are propagated unhindered and if positive they are reflected back. In the critical case of being $N^2_{ref}$ zero, the waves are absorbed in the region. The reflected planetary waves propagate downward, cross the tropopause and continue to the troposphere, thereby impacting the tropospheric conditions. The potential reflection of upward propagating planetary waves occurs due to anomalous gradients in the stratospheric zonal wind when the stratospheric polar vortex is in certain states. This idea was initially explored...
through singular-value decomposition of re-analysis data (e.g. Perlwitz and Harnik, 2003, 2004) but more recently other authors have used cross-spectral correlation analysis (Shaw et al., 2010) to show the impact of the downward propagating reflected wave energy on planetary wave structure in the troposphere to derive a detailed life cycle of ‘downward wave coupling’ events (Shaw and Perlwitz, 2013). In a recent study, Shaw et al. (2014) demonstrated that such extreme planetary wave–mean flow interaction events are linked to high-latitude planetary-scale wave patterns in the troposphere and zonal wind, temperature and mean-sea-level pressure anomalies in the Atlantic basin.

An obvious question is therefore, which of the mechanisms discussed above is the dominant one? At present there is no consensus in the literature. It may be the case that more than one of the mechanisms mentioned above is important.

2. How predictable is the winter stratosphere?

In this section our focus is on the predictability of stratospheric events which represent a significant departure of the extratropical stratospheric state from its climatological norm. This category mainly includes stratospheric sudden warmings and polar vortex intensification events which, collectively, we term Extreme Vortex Events (EVEs). Final warmings (FWs) may also be considered EVEs because they often involve a strong dynamical component that determines their timing and vertical structure. Although most work on stratospheric predictability has focussed on SSW events, there is evidence in the literature that dynamically driven FW and rapid polar vortex intensification events might be similarly important sources of tropospheric predictability. Hardiman et al. (2011) show that the significant variation in the timing of FW can result in significant changes to the tropospheric state. Shaw and Perlwitz (2014) show that dynamical processes contribute to rapid polar vortex intensification on time-scales relevant to the forecasting problem. The EVE category may also include extreme planetary wave heat-flux events (Shaw et al., 2014) that are linked to weather and climate in the North Atlantic and were prevalent during the winter of 2014.

We first discuss the predictability of the stratosphere in comparison to the tropospheric predictability. Under normal climatological conditions, the stratosphere is extremely stable and predictable on long time-scales when compared to the troposphere. For example, Waugh et al. (1998) used an NWP system to quantify the forecast skill in the troposphere (500 hPa) and lower stratosphere (50 hPa) for the Southern Hemisphere vortex. They found that the forecast skill for the lower stratosphere at 7 days lead time was comparable to the tropospheric skill at 3 days lead time when the vortex was undisturbed. Lahoz (1999) compared the predictive skill of the UK Met Office (UKMO) Unified Model in the stratosphere and troposphere for both Northern and Southern Hemisphere winters. He found that the model has higher forecast skill in the lower stratosphere than in the mid-troposphere and also showed that it has higher skill in northern winter than in southern winter. He attributed the differences in the model skill to the flow regime in the lower stratosphere which was dominated by lower wave numbers than in the mid-troposphere, and to larger initialization errors in the Southern Hemisphere. Similarly, Jung and Leutbecher (2007) presented an analysis of the historical forecast skill of the European Centre for Medium-range Weather Forecasts (ECMWF) forecast for all winters between 1995/1996 and 2006/2007, showing that 10-day forecasts of the 50 hPa geopotential height field have comparable skill to 5-day forecasts of the 500 hPa geopotential height field over the Arctic.

During large departures from climatology, the stratospheric predictability varies greatly. Table 1 lists studies which quantified the lead time at which forecasts of EVEs were considered skillful. Early attempts to understand EVEs often used so-called mechanistic models (Labitzke, 1965; Matsumo, 1971; Clark, 1974; Geisler, 1974; Holton, 1976; Holton and Mass, 1976). By 1970, one of the first true forecasts of an SSW event using a general circulation model (GCM) was performed by Miyakoda et al. (1970). They attempted to simulate the vortex-splitting SSW event of March 1965 and were able to predict the tendency of the polar vortex toward a breakdown, but failed to fully capture the splitting event, even when initialized only 2 days prior to the event. Since the work of Miyakoda et al., there have been a number of studies related to the predictability of EVEs as summarized in Table 1 which quantify the lead time at which forecasts of EVEs are considered skillful.

The advent of higher-resolution, more sophisticated NWP models combined with a reinvigoration of interest in SSW events following observations of the 22 February 1979 SSW event by satellites (McIntyre and Palmer, 1983) led to a number of studies re-examining stratospheric predictability. The February 1979 event was well predicted by contemporary NWP models at the time. Simmons and Strüting (1983) showed that the event was captured by the ECMWF model at 10-day lead times. Mechoso et al. (1985) reported more-limited skill for this event: for a coarser model resolution they found good forecast skill at 5-day lead times but their model failed to capture the SSW event at 7-day lead times. They also noted strong sensitivity to resolution and initial condition of their forecasts, with the model’s forecast skill improved as the horizontal resolution was increased from 4°(latitude) × 5°(longitude) to 2.4°(latitude) × 3°(longitude).

There was little work on the dynamical predictability of EVEs using NWP models until the late 1990s and early 2000s, perhaps linked to the lack of SSW events in the 1990s (Pawson and

<table>
<thead>
<tr>
<th>Year</th>
<th>Model</th>
<th>Event (SSW)</th>
<th>Predictability</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>GFDL GCM</td>
<td>March 1965</td>
<td>2 days (captured only tendency)</td>
<td>Miyakoda et al. (1970)</td>
</tr>
<tr>
<td>1979</td>
<td>ECMWF</td>
<td>February 1979</td>
<td>10 days</td>
<td>Simmons and Strüting (1983)</td>
</tr>
<tr>
<td>1979</td>
<td>UCLA GCM</td>
<td>February 1979</td>
<td>5 days</td>
<td>Mechoso et al. (1985)</td>
</tr>
<tr>
<td>2002</td>
<td>ECMWF</td>
<td>September 2002 (Antarctic)</td>
<td>7 days</td>
<td>Simmons et al. (2005)</td>
</tr>
<tr>
<td>2001</td>
<td>JMA NWP</td>
<td>December 2001</td>
<td>14 days</td>
<td>Mukougawa et al. (2005)</td>
</tr>
<tr>
<td>2002</td>
<td>NOGAPS-ALPHA</td>
<td>September 2002 (Antarctic)</td>
<td>5 days</td>
<td>Allen et al. (2006)</td>
</tr>
<tr>
<td>2007</td>
<td>ECMWF</td>
<td>Various</td>
<td>10 days</td>
<td>Jung and Leutbecher (2007)</td>
</tr>
<tr>
<td>2003</td>
<td>JMA NWP</td>
<td>December 2003</td>
<td>9 days</td>
<td>Hirooka et al. (2007)</td>
</tr>
<tr>
<td>2009</td>
<td>NCEP SFSIE</td>
<td>Various</td>
<td>15 days</td>
<td>Stan and Straus (2009)</td>
</tr>
<tr>
<td>2009</td>
<td>NOAA/ARL</td>
<td>January 2009</td>
<td>5 days</td>
<td>Kim and Flatau (2010) and Kim et al. (2011)</td>
</tr>
<tr>
<td>2013</td>
<td>Met Office</td>
<td>January 2013</td>
<td>14 days</td>
<td>Scife (2013)</td>
</tr>
</tbody>
</table>
| 2013 | GEOS-5 | January 2013 | 5 days | Lawrence Coy and Steven Pawson (http://gmao.gsfc.nasa.gov/researchhighlights/SSW/)

The predictability limits shown here are those quoted by the original study and therefore are not all calculated using the same methodology.

The advent of higher-resolution, more sophisticated NWP models combined with a reinvigoration of interest in SSW events following observations of the 22 February 1979 SSW event by satellites (McIntyre and Palmer, 1983) led to a number of studies re-examining stratospheric predictability. The February 1979 event was well predicted by contemporary NWP models at the time. Simmons and Strüting (1983) showed that the event was captured by the ECMWF model at 10-day lead times. Mechoso et al. (1985) reported more-limited skill for this event: for a coarser model resolution they found good forecast skill at 5-day lead times but their model failed to capture the SSW event at 7-day lead times. They also noted strong sensitivity to resolution and initial condition of their forecasts, with the model’s forecast skill improved as the horizontal resolution was increased from 4°(latitude) × 5°(longitude) to 2.4°(latitude) × 3°(longitude).

There was little work on the dynamical predictability of EVEs using NWP models until the late 1990s and early 2000s, perhaps linked to the lack of SSW events in the 1990s (Pawson and

© 2014 The Authors. Quarterly Journal of the Royal Meteorological Society published by John Wiley & Sons Ltd on behalf of the Royal Meteorological Society.
Figure 1. The Goddard Earth Observing System model, version 5 (GEOS-5) 5-day forecast of the stratospheric sudden warming of 7 January 2013. Latitudinal cross-sections of forecast (blue) and observational analyses (red and green) of (a) stratospheric temperatures (K), (b) zonal winds (m s$^{-1}$), and (c) profiles of zonal winds at 10 hPa. The left and middle panels (a,b) show how temperature gradient and wind at 10 hPa reversed from 2 to 7 January and how the model successfully forecasted the reversal 5 days in advance. Winds and temperatures at 1200 UTC on 2 January 2013 show pre-warming conditions, and the winds and temperatures at 1200 UTC on 7 January 2013 show that the 1200 UTC 2 January 5-day forecast predicted the event very well. The rightmost panel (c) shows the vertical profile of zonal wind at 60°N on 2 January (red) and on 7 January (blue is forecast and green is observational analysis).

Naujokat, 1999). Using the Japan Meteorological Agency (JMA) NWP model, Mukougawa and Hirooka (2004) showed the warming in the stratospheric polar region associated with the SSW event of 15 December 1998 could be predicted by an NWP model from 1 month in advance. This extended predictability was based upon control forecasts without any perturbation to the initial condition and therefore it is unlikely that the model would have practical probabilistic skill at such long lead-times. This was demonstrated, though for a different SSW case, by Mukougawa et al. (2005) who found skill up to only 2 weeks when considering probabilistic predictions of the December 2001 SSW event. They emphasized that the predictability of SSW events was sensitive to the predictions of the planetary wave structures causing the warmings and also to whether the major warming was preceded by a minor warming (Hirooka et al., 2007). For example the extended predictability of the December 1998 and the December 2001 SSW events were attributed to their dominant wave-1 precursors. In contrast, the SSW event of the winter 2003/2004 had a significant contribution from smaller-scale waves (wave-2 and wave-3) and therefore could only be predicted about 9 days in advance (Hirooka et al., 2007). These authors also suggested that skill is enhanced by successfully predicting the rate and location of amplification of planetary waves in the troposphere prior to the SSW, and that has a larger impact on forecast skill than accurately predicting the zonal flow configuration in the lower stratosphere.

Kim and Flatau (2010) and Kim et al. (2011) performed a detailed sensitivity study of the predictability of the 2009 Arctic SSW using the Navy Operational Global Atmospheric Prediction System (NOGAPS) and showed significant predictive skill at 5-day lead-times. For this case, the skill of NOGAPS was very sensitive to the orographic wave drag parametrization schemes, which influenced the zonal mean state. The SSW event in the Southern Hemisphere in 2002 was successfully predicted a week in advance by the ECMWF operational forecasting system (Simmons et al., 2005) and 6 days in advance by the NOGAPS-ALPHA system (Allen et al., 2006). Simmons et al. (2005) also included examples of three successful forecasts of Northern Hemisphere vortex-splitting cases when the ECMWF model was initialised using ERA-40 re-analysis data (i.e. the SSW events of 29 January 1958, 21 February 1979 and 17 February 2003). Coy et al. (2009) highlighted significant sensitivity of NOGAPS-ALPHA forecasts of the January 2006 SSW event to horizontal resolution, which they attributed to the strong influence of planetary wave activity emanating from a compact upper tropospheric ridge over the North Atlantic.

More recent studies have attempted to take a broader perspective on the predictability of SSW events by considering the forecast skill of a model for a larger number of events. Stan and Straus (2009) showed that the SSW predictability time (the time for the normalized error in the 50–70°N zonal wind to become 0.5) was about 15 days for wave-1 events and significantly smaller (about 10 days) for wave-2 events (see their Fig. 8). They suggested that the limited SSW predictability was mainly due to the inability of the model to correctly simulate the phase and flux of upward propagating planetary waves. Marshall and Scaife (2010) compared the predictability of four SSW events in a 38-level low-top and a 60-level high-top version of the Hadley Centre Atmospheric General Circulation Model (AGCM). They found improved predictability with the high-top version (9–15 days) in comparison to the low-top version (6–8 days). However, they did not find any difference in the tropospheric wave activity during the growth stage in the two model versions. They suggested that the high-top model showed improved predictability because it could capture downward propagating SSW signals in the upper stratosphere a few days earlier than for the low-top model. Jung and Leutbecher (2007) showed that the stratospheric predictive skill of ECMWF with a 10-day lead-time has significantly improved from the low resolution (about 180 km) version to the high resolution (40 km) version. They also showed that the downward propagation of stratospheric circulation anomalies, which constitutes a potential source of tropospheric forecast skill, was realistically represented in the seasonal integration.

As discussed in the introduction, enhanced vertical and horizontal model resolution in the stratosphere benefit the assimilation of observations, both affecting the skill of the resulting operational forecasts and the quality of widely used re-analysis products. This source of forecast skill was recognised by Simmons et al. (1989) and motivated the increase in the number of vertical model levels from 16 to 19 in the ECMWF operational system with increased resolution in the stratospheric and model top at 10 hPa. Later, the number of vertical levels in ECMWF assimilation and forecasting system was increased to 50 with model top at 0.1 hPa. This increase in vertical resolution was shown to have improved the quality of stratospheric analysis and stratospheric predictability at the levels up to 10 hPa in comparison to the prior 31-level system (Untch et al., 1999).

Figures 1 and 2 show the typical predictability of current operational models for the major SSW event of 7 January 2013. Figure 1 shows the 5-day forecast of the SSW event produced by the Goddard Earth Observing System Model, Version 5 (GEOS-5) model (blue line) and the Global Modeling and Assimilation...
3. The origins of stratospheric predictability

The extratropical stratosphere is influenced by a number of processes which occur on a variety of spatial and temporal scales. Conceptually, stratospheric predictability in the models arises in two areas: (i) initial value predictability, which is derived from a model’s ability to capture the dynamical processes and mechanisms that characterize the evolution and life cycle of a specific EVE; and (ii) boundary value predictability, which derives from a model’s ability to capture the propensity of the wintertime stratosphere to produce an EVE. This section first summarises our knowledge of the dynamics of EVEs, and then elaborates on the processes which provide initial value and boundary value predictability in the stratosphere and the limitations associated with their modelling.

3.1. Dynamics of EVEs

For a detailed review of stratospheric dynamics and stratosphere–troposphere coupling in particular, other review papers are available (e.g. Shepherd, 2002; Haynes, 2005; Gerber et al., 2012). In this subsection we confine the discussion to the aspects of stratospheric dynamics most relevant to our understanding of stratospheric predictability.

The interaction of planetary waves and the mean westerly stratospheric flow is fundamental to our understanding of EVEs. The first detailed numerical model of the interaction of vertically propagating planetary waves with the mean zonal flow was developed by Matsuno (1971). The abstract of this paper succinctly described why this interaction is important for the stratosphere, as apparent in the four key sentences reproduced below:

If global-scale disturbances are generated in the troposphere, they propagate upward into the stratosphere, where the waves act to decelerate the polar night jet through the induction of a meridional circulation. Thus, the distortion and the break-down of the polar vortex occur. If the disturbance is intense and persists, the westerly jet may eventually disappear and an easterly wind may replace it. Then ‘critical layer interaction’ takes place.

Figure 3 illustrates the co-evolution of the polar vortex (a,b) and planetary wave activity (c,d) during the major SSW in early January 2013. The data presented in this Figure is taken from the 6-hourly ERAI re-analysis fields. Planetary wave propagation is diagnosed using Eliassen–Palm (EP) flux vectors, under the quasi-geostrophic and linear approximations (Edmon et al., 1980). The EP-flux vectors shown in Figure 3(c,d)∗ represent both the magnitude and net group propagation of this planetary-wave activity flux (McIntyre, 1982). Where there is large convergence of EP fluxes, there is irreversible exchange of wave momentum into the mean flow, which produces a deceleration of the zonal mean flow (Andrews et al., 1987). Figure 3(a) illustrates a characteristic configuration of the wintertime stratospheric polar vortex, as represented on 2 January 2013. The corresponding analysis of the implied propagation of planetary wave activity from the

Figure 2. Predictability of the northern hemispheric stratospheric sudden warming of 7 January 2013. The zonal mean zonal wind at 60°N and 10 hPa is diagnosed from the ERA-Interim re-analysis (black line) and for three different forecasting systems: (a) The Centre for Australian Weather and Climate Research (CAWCR) forecast system (red), (b) Meteorological Research Institute, Japan (MRI) (blue), and (c) Korea Polar Research Institute, Korea (KOPRI) (green). Forecasts initialised on 23 December 2012 are shown in solid lines, 28 December in dotted lines and 2 January in dashed lines.
troposphere to the stratosphere and subsequent refraction of wave activity equatorward implied by the deflection of EP-flux vectors is shown in Figure 3(c). Over the next few days to 7 January, the orientation of EP-flux vectors in the middle stratosphere changes as waves begin to propagate into the polar region, leading to a large EP-flux convergence around 70°N (Figure 3(d)) and deceleration of the zonal mean jet associated with the vortex splitting into two pieces at 10 hPa (Figure 3(b)).

Changes to the zonal mean state which allow poleward focussing of planetary waves are normally termed ‘vortex preconditioning’ (McIntyre, 1982). Typically, a preconditioned vortex should be weaker and smaller than normal and centred over the Pole. In the zonal mean, this onset stage appears as anomalously weak flow equatorward of 60° latitude and anomalously strong flow poleward of 60° latitude (McIntyre, 1982; Andrews et al., 1987; Limpasuvan et al., 2004). The preconditioning stage is one part of the typical SSW life cycle which can be exploited by NWP models for the purpose of predicting SSW occurrence.

If planetary wave forcing is large and persistent then, as noted by Matsuno (1971), zonal winds can reverse sign and a critical layer for planetary waves is formed. Typically, this process occurs first in the upper stratosphere and mesosphere (well above 10 hPa: Coy et al., 2011) and then the zonal wind reversal migrates downwards, over a period of a few weeks, through the stratosphere toward the tropopause as waves dissipate at successively lower levels. This ‘downward propagation’ of the zonal mean flow anomaly is a critical aspect for stratospheric predictability since it provides the means by which the flow in the upper stratosphere might influence the troposphere at some later point in the future on time-scales of several days to weeks.

As the zonal mean wind reversal propagates to the lower stratosphere, wave activity in the upper stratosphere weakens significantly; easterly winds prevent any further vertical propagation of planetary waves (Charney and Drazin, 1961). The lack of planetary wave activity allows radiative recovery of the vortex described as a ‘vacillation cycle’ by Kodera et al. (2000), Kodera and Kuroda (2000), and Kuroda (2002).

Although SSWs are always complex events, they may be arbitrarily classified as either vortex displacement events, characterized by a shift of the vortex off the Pole or vortex-splitting events, when the vortex splits into two distinct vortices (O’Neill, 2003; Charlton and Polvani, 2007). There is some evidence, beginning with the work of Simmons (1974), Tung and Lindzen (1979), and Plumb (1981) that vortex-splitting SSWs are produced by a distinct ‘resonant excitation’ mechanism which does not depend upon anomalous tropospheric wave activity or favourable stratospheric ‘preconditioning’ (Esler and Scott, 2005; Esler and Matthewman, 2011; Matthewman and Esler, 2011). According to the ‘resonant excitation’ mechanism, SSW events may occur when planetary waves resonantly excite either a barotropic mode of the vortex (in the case of vortex-splitting...
events) or baroclinic mode of the vortex (in the case of vortex-displacement events). These ideas have important consequences for the predictability of split-vortex SSWs, in that vortex splitting might be initiated by very small changes in tropospheric wave forcing and/or changes to the stratospheric state. The implication of this result is that the vortex-splitting type of SSW events might have lower predictability than displacement events, in line with the results of Stan and Strauss (2009). To quantify the relative predictability of the vortex split and displacement types of SSW events, a series of such events need to be evaluated by using multiple models; this is an important topic for future research.

If the SSW occurs during late winter or early spring, the seasonal increase of radiative heating in the polar region may prevent the reformation of the polar vortex. These events are thus termed FW events. The major contributor to the variation in the stratospheric FW date is the planetary wave activity (Waugh and Rong, 2002; Black et al., 2006; Salby and Callaghan, 2007). This FW concludes the stratospheric winter season, and, as suggested by Waugh and Rong (2002), its timing is highly variable from year to year in the Northern Hemisphere. They found that a change of EP flux from the troposphere by ±2 standard deviations can vary the timing of the Northern Hemisphere FW by as much as 2 months, thus advancing the warming to as early as February or delaying it to as late as May. A similar sensitivity of EP-flux anomalies was also found to be associated with warm and cold winters (Salby and Callaghan, 2002, 2007). Black et al. (2006) showed that the weakening of stratospheric westerlies occurs much more rapidly for stratospheric FW events in contrast to the climatological seasonal cycle. In another study Hardiman et al. (2011) found that in some years FW events start in the mid-stratosphere and in others FW events start in the upper stratosphere. The difference in the vertical evolution of FW events depends on the strength of the winter stratospheric polar vortex, the refraction of planetary waves, and the altitudes at which the planetary waves break in the northern extratropics. The large variations in the FW dates and initiation altitude result in significant year-to-year variability in tropospheric spring climate and may have implications for tropospheric predictability in the spring season (Black et al., 2006; Hardiman et al., 2011).

It is also possible to observe EVEs in which the polar vortex becomes unusually strong and a significant reduction in the polar cap temperature occurs. These vortex intensification events are similar in some ways to vortex weakening events but opposite in sign. They are associated with anomalously weak tropospheric wave activity and enhanced radiative cooling of the polar cap region (Limpasuvan et al., 2005). However, the changes in wind and polar cap temperature are weaker, slower and much less dramatic than during SSWs. Although these events are linked to a lack of tropospheric wave activity in the polar cap, similar problems limit their predictability, as discussed in the next section.

3.2. Initial value problem

Given the dynamics discussed above, predicting EVEs in the stratosphere depends both on the ability of models to reproduce the mean stratospheric state prior to an EVE and on their ability to predict both the forcing and propagation of planetary wave activity through the troposphere and stratosphere. In this section, we first consider the case where a model is able to capture properties of the flow present in the initial state, for example an enhancement of tropospheric wave activity, which ultimately allows it to predict an individual EVE. In section 3.3 we broaden our discussion to include factors which influence the stratospheric mean state on longer time-scales and so may lead to a greater or lesser likelihood of EVEs and enhanced predictability on longer time-scales.

3.2.1. Modelling wave propagation and EVEs

Polvani and Waugh (2004) clearly demonstrated the anomalous enhancement of 40-day integrated eddy heat fluxes, which are strongly correlated with the upward propagation of planetary waves, prior to extreme stratospheric events. The composite Northern Annular Mode (NAM) index and corresponding heat flux anomaly at 100 hPa for 25 high heat flux events and 24 low heat flux events from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP-NCAR) re-analysis data is shown in Figure 4. The Figure also shows 40-day integrated average of heat flux anomaly for both
composites. From Figure 4 it is clear that positive (negative) NAM index anomalies are preceded by positive (negative) heat flux anomalies for high heat flux (low heat flux) events. Polvani and Waugh argued that although the NAM anomalies appear to be originating from the upper stratosphere and propagating downward to the troposphere according to the downward control hypothesis presented by Baldwin and Dunkerton (2001), the fact that upper-stratospheric NAM anomalies are preceded by anomalies in the upward wave activity (as shown in the lower panel of the Figure) indicates otherwise: that the control of stratospheric anomalies lies in the troposphere. This study clearly demonstrated a strong link between stratospheric extreme events and tropospheric wave activities.

The understanding of variability in the lower troposphere that leads to anomalous upward propagation of wave activity is important for accurate prediction of events in the stratosphere. Our understanding of wave propagation through the tropopause into the stratosphere is based on the detailed mathematical treatment of atmospheric wave propagation by Charney and Drazin (1961) and Matsuno (1970, 1971) as well as the review of the dynamics of stationary waves in the troposphere by Held et al. (2002). Forced planetary waves are excited mechanically from perturbations in the mean flow over mountain ranges, and the differential heating of the atmosphere over the continents and oceans. Additionally, stirring of the atmosphere by baroclinic instability also generates Rossby waves, although typically at small horizontal wavelengths. Charney and Drazin (1961) showed that wave energy can only propagate vertically when the mean zonal velocity is positive (westerly), but less than a critical velocity, which is dependent upon the wavelengths of the waves. The stratosphere acts as a selective short-wave filter and only long planetary waves with wave numbers up to zonal wave number 2 can typically penetrate into the middle and upper stratosphere. Since there is a radiatively driven reversal of the stratospheric zonal mean flow between winter and summer, this also means that planetary waves are almost entirely absent in the summertime stratosphere.

One of the interesting consequences of the filtering of planetary waves by the mean flow is that the propagation of planetary waves into the extratropical stratosphere can vary even when the amplitude of tropospheric wave forcing is constant. This effect, often known as stratospheric vacillation, was first demonstrated by Holton and Mass (1976) in a very simple channel model, but has since been shown in a range of models with different levels of complexity (Yoden, 1987; Christiansen, 1999; Scott and Haynes, 2006; Scaife et al., 2005; Scott and Polvani, 2006; Scott et al., 2008). As described by Scott and Polvani (2006) this means that the lower stratosphere can act as a ‘valve’ which opens and closes for upward propagating waves according to the current state of the stratospheric polar vortex. This means the stratosphere itself controls the amount of wave energy entering into the stratosphere from the troposphere. Scott et al. (2008) showed that the temporal and spatial structure of the vacillations in the stratosphere is independent of whether tropospheric forcing is in the form of transient pulses or steady if the forcing amplitude exceeds a critical value. Sjoberg and Birner (2012) showed in a modelling experiment that the time-scale over which tropospheric planetary-wave forcing is applied can be more important than its amplitude in determining whether they cause an SSW. However, they also showed that the required time-scale of tropospheric forcing to produce an SSW is set by the internal stratospheric characteristics such as time-scales of radiative relaxation.

In the context of understanding what limits stratospheric predictability, it is clear that not only should a model capture processes that lead to the amplification of the tropospheric planetary wave field, but it should also accurately represent the mean flow in the lower and middle stratosphere. The role of model configuration on the simulation of wave propagation and dissipation in the stratosphere is also discussed by Shaw and Perlwitz (2010). They showed that reflection of waves from the model top in low-top models could severely compromise the ability of models to simulate the propagation of the stationary wave field. They found that the effects of the model lid can be significantly mitigated by forcing any remaining parameterized gravity-wave momentum to deposit at the upper boundary, since this conserves column-integrated momentum and leads to realistic downward-control circulations.

As noted by Haynes (2005), it is likely that variability in the stratosphere is determined both by the ‘valve’ effects described in this section and also by transient changes to tropospheric planetary waves driven by a range of tropospheric processes (Garfinkel et al., 2010; Kolstad and Charlton-Perez, 2011). Recently, Sun et al. (2012) performed idealized studies to investigate the relative role of these two effects and concluded that stratospheric preconditions was much less important than tropospheric precursor effects in determining the timing of warming events. In the following sections we explore processes in the troposphere which affect the initiation and propagation of these waves.

### 3.2.2. Tropical wave sources: MJO

The Madden–Julian Oscillation (MJO) is characterized by the eastward propagation (4–8 m s$^{-1}$) of large-scale clusters of deep convective activity over the tropical oceans occurring on intraseasonal time-scales (30–60 days) associated with anomalous rainfall and coupled to the large-scale atmospheric circulation. Cassou (2008) shows this coupling as an asymmetric tropical–extratropical lagged relationship with the North Atlantic Oscillation (NAO) where MJO pre-conditioning occurs for positive (negative) NAO events as a mid-latitude wave train initiated by the MJO in the western-central tropical Pacific (eastern tropical Pacific and western Atlantic). Garfinkel et al. (2012) show that, as the MJO influences the tropospheric North Pacific sector, which is strongly associated with SSWs, then SSWs tend to follow certain MJO phases. Garfinkel et al. also demonstrate that the MJO’s influence on the vortex is comparable to the QBO (see section 3.3.1) and El Niño, and could be used to improve NAM forecasts out to 1 month.

### 3.2.3. Extratropical wave sources: atmospheric blocking

Atmospheric blocks are stationary weather patterns (usually high-pressure systems) in the troposphere which typically persist beyond a week. Stratospheric warming episodes are often accompanied by blocks (Andrews et al., 1987; O’Neill et al., 1994; Kodera and Chiba, 1995; Coy et al., 2009; Nishi et al., 2009), but the causal link between blocking and SSWs, if any, has always been in question. For the purposes of understanding limits to stratospheric predictability it is important to understand if blocking play any role in triggering SSWs. Although there have been significant recent improvement in simulating blocks (e.g. Scaife et al., 2011), there are still substantial biases. In recent studies, Scaife et al. (2011) and Dunn-Sigouin and Son (2013) showed that both Coupled Model Intercomparison Project CMIP3 and CMIP5 models have significant biases in duration and frequency of the simulated blocks. A similar bias is also found in NWP models (e.g. Dunn-Sigouin et al., 2013).

Martius et al. (2009) found that out of 27 SSW events in ERA–40 data from 1957 to 2001, 25 events were preceded by atmospheric blocking, in line with previous studies by Quiroz (1986) and O’Neill and Taylor (1979). Furthermore they found evidence that vortex displacement SSW events were preceded by Atlantic basin blocking and vortex-splitting SSW events were preceded by blocking in the Pacific basin or in both the Atlantic and Pacific basins. A broad correspondence between the amplification of the wave number 1 planetary wave prior to SSW vortex displacement events and the amplification of the wave number 2 planetary wave prior to SSW vortex splitting events has also been found (Martius et al., 2009)
et al., 2009; Cohen and Jones, 2011). The findings of Castanheira and Barriopedro (2010) supported this result, showing that Atlantic blocks caused in-phase forcing and amplification of the zonal wave number 1 planetary wave, while Pacific blocks cause in-phase forcing and amplification of the zonal wave number 2 planetary wave. Castanheira and Barriopedro also noted that the connection between the amplification of the wave number 2 planetary wave and vortex-splitting SSWs is more complex than that between amplification of wave number 1 planetary waves and vortex displacement SSWs, as noted in other prior studies of SSWs (Labitzke and Naujokat, 2000). Using a different diagnostic of blocking, Woollings et al. (2010) found evidence that European blocking was linked to the amplification of the zonal wave number 2 planetary wave. In contrast to these studies, Taguchi (2008) analysed 49 years of NCEP-NCAR reanalysis data from 1957/1958 to 2005/2006 and found no evidence of preferential blocking either pre- or post-SSWs. Since Taguchi (2008) did not separate vortex displacement and split events, this might explain the different result of his study compared to others in the literature.

3.3. Boundary value problem

This section will focus on the predictability of the stratosphere on weekly to sub-seasonal time-scales and the sources that can impact the statistical likelihood of an extreme stratospheric event in a given winter season.

3.3.1. Quasi-biennial oscillation (QBO)

The stratospheric QBO in the Tropics arises from the interaction of the stratospheric mean flow with eddy fluxes of momentum. The eddy fluxes are carried upward by Kelvin waves, mixed Rossby–gravity waves, and small-scale gravity waves which are excited by tropical convection. The QBO is characterized as downward propagating easterly and westerly wind regimes with an average period of about 28 months and as such has the potential to exert a significant regulating effect on atmospheric predictability. For an in-depth account of the QBO, readers are referred to the review paper by Baldwin et al. (2001).

Since its discovery, there have been a number of attempts to search for links between the QBO and tropospheric weather (e.g. Ebdon, 1975). As noted by Anstey and Shepherd (2014), the first study to examine the relationship between the QBO and variability in the high-latitude stratosphere with a reasonably long record was that of Holton and Tan (1980). The so-called Holton–Tan relationship revealed by this and subsequent studies predicts that a weaker polar vortex and more SSWs are expected during the easterly phase of QBO (Holton and Tan, 1980, 1982; Labitzke, 1982). This relationship has been largely supported by numerous subsequent studies (Kodera, 1991; O’Sullivan and Young, 1992; O’Sullivan and Dunkerton, 1994; Niwano and Takahashi, 1998; Kinnersley and Tung, 1999; Hu and Tung, 2002; Ruzmaikin et al., 2005; Hampson and Haynes, 2006; Calvo et al., 2007; Naoe and Shibata, 2010; Watson and Gray, 2014), although there is some evidence that the strength of the relationship has varied over the observed period (Lu et al., 2008).

The mechanism for this link proposed by Holton and Tan involves the presence or absence of the zero-wind line in the subtropical lower stratosphere which influences extratropical planetary waves propagating into the stratosphere. The region between the zero-wind line and the Pole acts as a waveguide for these waves. Upward and equatorward propagating planetary waves when encountering the zero-wind layer either converge or reflect back towards the polar region depending on the vertical and meridional component of the wave number (Garfinkel et al., 2012). Large-amplitude waves tend to dissipate at the critical line whereas for smaller-amplitude waves the zero-wind line may act as a reflecting surface. In either scenario wave activities are limited in a region between the zero-wind line and the Pole.

During the easterly phase the associated zero-wind line in the subtropics acts as a critical line for equatorward-propagating planetary waves. Waves dissipate at the equatorward flank of the polar night jet leading to a stronger residual circulation which weakens the polar vortex. In the case of the westerly phase of the QBO, waves propagate to the Tropics unhindered without much dissipation or impact on the residual circulation or polar vortex. The weaker vortex in the easterly phase of the QBO is also found to be associated with an increased upward component of EP flux (Dunkerton and Baldwin, 1991; Garfinkel and Hartmann, 2008; Yamashita et al., 2011).

The review of Anstey and Shepherd (2014) notes that subsequent studies have proposed alternative means by which the QBO influences the high-latitude stratosphere. The studies by Gray (2003) and Pascoe et al. (2006) suggest that wind anomalies in the tropical upper stratosphere are responsible for the Holton–Tan effect. Garfinkel et al. (2012) suggested that the extratropical influences of the QBO may be more strongly related to the mean meridional circulation induced by the QBO itself rather than associated critical-line effects. They pointed out, for example, that the easterly QBO phase reduces the planetary-wave refractive index in the mid-stratosphere near 40–50°N, which induces a residual circulation by altering the wave propagation and warms the polar vortex (see Fig. 1 of Garfinkel et al. (2012)). However, the nudging experiments of Watson and Gray (2014) produce anomalies in the EP flux and EP-flux convergence consistent with the original Holton–Tan mechanism. As summarised by Anstey and Shepherd, there is still no definitive picture of the mechanism of QBO–polar vortex coupling.

Climate models often struggle to resolve or represent the QBO (Thompson et al., 2002; Boer and Hamilton, 2008). However, a few general-circulation models have been shown to be able to simulate the evolution of the QBO (Takahashi, 1996; Scaife et al., 2000; Giorgetta et al., 2002; Kim et al., 2013). Fine vertical resolution in the stratosphere is known to be important in attempting to simulate the vertical propagation of waves and momentum deposition which drives the QBO (Schmidt et al., 2013). Similarly, models that are able to simulate the QBO typically employ a non-orographic gravity-wave drag parametrisation (e.g. Scaife et al., 2000). The impact of the QBO on the extratropical stratosphere is also sensitive to the stratospheric representation of the model. Marshall and Scaife (2009) showed the weakening of the vortex in response to the easterly QBO phase was in better agreement with observations in simulations with a high-top model than with a low-top model.

In the context of medium-range and monthly forecasts, however, it is important to note that the impact of the QBO on the extratropics is well captured simply by accurate data assimilation in the tropical stratosphere. The radiative relaxation rates in the tropical lower stratosphere are very slow. This means that if the model does not generate its own QBO the assimilated QBO in the model will ‘die out’ but only very slowly over a period of many days and it will not realistically evolve by slowly descending, but instead will just sit there. Since the QBO has such a long time-scale, there will be little change in tropical winds during the course of a 15- or 30-day forecast and so even models which simply preserve tropical winds will capture QBO effects in the extratropics. Note that it is also possible for data assimilation systems to fail to adequately capture the QBO, given the sparse sampling of winds in the tropical stratosphere (e.g. Saha et al., 2010).

3.3.2. El Niño–Southern Oscillation (ENSO)

ENSO has been shown to influence the northern stratospheric polar vortex (Bromwich et al., 2004; Bell et al., 2009; Cagnazzo and Manzini, 2009; Ineson and Scaife, 2009) through enhancement of tropospheric planetary wave activity. Early
studies showed that planetary wave activity is enhanced and the stratospheric vortex is weaker than normal during the El Niño phase (van Loon and Labitzke, 1987; Sassi et al., 2004; García-Herrera et al., 2006; Manzini et al., 2006; Free and Seidel, 2009). It might therefore be expected that SSW events would occur more frequently during the El Niño phase. Taguchi and Hartmann (2006) found that the SSWs were twice as likely to occur in El Niño winters as in La Niña winters, in a perpetual winter integration of a climate model.

More recent studies suggest that the connection between ENSO and SSW is more complex. An example is the study of Butler and Polvani (2011) which showed that SSW events are almost equally associated with both phases of ENSO. Figure 5 updates Fig. 1 of Butler and Polvani (2011) to include data up to June 2013 and has slight changes to both the Niño3.4 and NCEP-NCAR re-analysis fields that were used. The NCEP CPC (NCEP Climate Prediction Center) has updated the Niño3.4 index (both the instrument being used and the climatology, now 1981–2010). These make very slight differences in El Niño/La Niña classifications. Here we are also using the classification scheme that CPC follows (3-month averages of this index must stay above 0.5°C for five consecutive seasons). In addition, NCEP changed their NCEP-NCAR re-analysis version slightly in the last two years. This meant that using the Charlton and Polvani (2007) SSW definition on the new data produced no warming in 1968 and two warmings in 2010. Those changes are reflected on the Figure. The Figure shows that there are frequent SSW events in both the El Niño and La Niña phases of ENSO with almost equal frequency over the period studied (although it should be noted that the sample size is small as with most observational studies of stratospheric variability). These updates are included in a recent Butler et al. (2014) study relating SSW to the northern hemispheric winter climate in ENSO active years. Garfinkel et al. (2012) also found in the ERA–40 re-analysis that both La Niña and El Niño lead to similar anomalies in the region associated with precursors of SSWs leading to a similar SSW frequency in La Niña and El Niño winters.

The mechanism behind the ENSO–SSW teleconnection is still unclear. The nonlinear interaction between ENSO and other dynamical phenomena like the QBO (Calvo et al., 2009) make it difficult to untangle their specific influence on the polar stratospheric state. Earlier studies based on the observational record were also largely inconclusive mainly because of the difficulty in isolating the ENSO signal from the QBO due to the concurrence of warm ENSOs with easterly QBOs in the observed record (Wallace and Chang, 1982; Baldwin and O’ Sullivan, 1995). Experiments using different combinations of QBO and variable sea-surface temperature (SST) show that ENSOs interact nonlinearly with the QBO to produce the observed number of SSW events per decade (Richter et al., 2011). They showed that individual forcing factors of SSW and ENSO (variable SST) do not add up linearly to produce the observed result of the combined forcing. In fact only one forcing (either QBO or ENSO) alone was sufficient to produce most of the observed SSWs whereas absence of both drastically reduced the number of SSW events. This underlines the difficulties in attributing ENSO and QBO influences on SSWs.

3.3.3. Surface forcings

The state of the land, ocean and ice surface conditions and their seasonal and interannual variability influence sea-level pressure and surface temperature. This variability in surface conditions modulates the planetary wave structures and ultimately influences the extratropical stratosphere. Major contributions of surface-induced anomalies in the local and large-scale circulation and weather patterns may originate from variability in snow cover, SST, and sea–ice extent at high latitudes (Petoukhov and Semenov, 2010; Tang et al., 2014). Anomalously large October snow extent over Eurasia is associated with enhanced wintertime upward propagating planetary waves which lead to a weaker polar vortex (Cohen et al., 2007; Orsolini and Kvamstø, 2009; Allen and Zender, 2010; Smith et al., 2010). It has been postulated that the above-normal snow cover in October leads to the intensification of the Siberian high and colder surface temperatures that increase wave activity flux in late autumn and early winter leading to the weaker vortex and an increased probability of a stratospheric warming. The substantial lag between anomalies in October snow cover and winter weakening of the stratospheric vortex may be explained by the linear interference between the climatological stationary wave field and the snow-forced transient wave field (Smith et al., 2011). Smith et al. (2011) show that waves associated with the snow anomalies are initially out of phase with the climatological wave, but later in the winter interfere constructively to increase upward wave flux. However, the reason for this phase change between October and midwinter remains unclear. Experiments in which snow anomalies have been prescribed in general circulation models have shown some dynamical response in the stratosphere (Gong et al., 2003; Fletcher et al., 2009) but failed when snow was allowed to evolve freely in the model (Hardiman et al., 2008). Cohen and Jones (2011) also suggest that the surface precursors of vortex-displacement and vortex-splitting SSW events are distinct, with displacement events more strongly linked to changes over Eurasia associated with the Siberian High. A similar wave-induced forcing mechanism was also associated with the sea-surface temperature in the North Pacific with cold sea-surface temperatures appearing to weaken the polar vortex (Fereday et al., 2008; Hurwitz et al., 2011, 2012).

3.3.4. Volcanic aerosols and solar radiation

Tropical stratospheric temperatures are also sensitive to other long time-scale climate forcings, including radiative impacts of the injection of sulphur dioxide and other materials from explosive volcanic eruptions into the stratosphere (which leads to the production of large quantities of sulphate aerosol), and changes in short-wave solar forcing related to the 11-year solar cycle. The impact of these forcings on the tropical stratosphere and their links to the high latitudes and to the troposphere are covered in detail by the reviews of Robock (2000) and Gray et al. (2010).

Volcanic eruptions give rise to an enhanced Equator-to-Pole temperature gradient in the lower stratosphere and consequently a stronger polar vortex (Robock, 2000). There is evidence that volcanic eruptions might lead to enhanced predictive skill for the troposphere on seasonal time-scales (Marshall et al., 2009), but due to the known problems climate models often have in

Figure 5. The association of SSWs with the Niño3.4 index monthly time series. The data are taken from the NCEP-NCAR re-analysis from January 1958 to June 2013. The red markers show the time of observed SSW events. Markers are placed at position +2 if the SSW occurs during an El Niño winter, and at −2 if it occurred during a La Niña winter, and at 0 if during a neutral winter. Two SSWs in the same winter are indicated by triangle marker. The shading indicates the ENSO neutral phase. Updated from Butler and Polvani (2011).
simulating the extratropical response to volcanic forcing (Driscoll et al., 2012) it is thought that the enhanced skill originates from the initial conditions rather than the model capturing the dynamical response to the eruption (Marshall et al., 2009).

There is some limited evidence that the phase of the solar cycle gives rise to enhanced North Atlantic surface seasonal predictability, with a lag of 3–4 years which may involve coupling between the atmosphere and ocean (Gray et al., 2013; Scaife et al., 2013). We are not aware of studies which examine the predictability of the extratropical stratosphere during different phases of the solar cycle, although mechanistic studies (e.g. Kodera and Kuroda, 2002) suggest predictability should exist. There is still significant uncertainty about the mechanism by which solar variability influences the troposphere, although the studies of Simpson et al. (2009, 2012) strongly suggest an important role for tropospheric eddy processes and the tropical lower stratosphere which may not involve coupling to the extratropical stratosphere.

We also emphasise that for both volcanic and solar forcing, their long time-scale in comparison to the time-scale of medium-range and sub-seasonal forecasts means that their impact on predictability can be largely captured by accurate data assimilation of the tropical stratospheric state and representation of the solar cycle in the model.

4. Stratospheric predictability and tropospheric forecast skill

In the last several years, there has been a number of research efforts focussed on quantifying the impact of stratospheric dynamical variability on the predictability of the troposphere. Several different methods have been used to construct experiments to quantify the impact of stratospheric variability on the troposphere. In the following sections, we group experiments by type and summarize their collective results.

4.1. Comparison of high-top and low-top model experiments

This first method directly compares the forecast skill of the high- and low-top models. The experiments are constructed in which forecasts of the same periods are made using high-top and low-top models and the resulting differences in forecast skill are attributed to the presence of the full stratosphere in the high-top model. One difficulty with these experiments is that running two model versions which differ only in their stratospheric representation is often difficult to achieve in practice. Nonetheless, this type of experiment can be a very successful way to assess the impact of the stratosphere on tropospheric forecast skill.

For example, Kuroda (2008) demonstrated, using the Japan Meteorological Agency (JMA) model, that the lead time for the correct prediction of tropospheric zonal mean winds was increased to lead times of 2 months in the high-top model from 15 days in the low-top for the SSW event during the 2003–2004 winter. Marshall and Scaife (2010) performed a similar study with a high-top and low-top version of the Met Office model and found that the high-top model gave improved predictability. Furthermore, the low-top model was unable to capture enhanced cooling over Europe after SSW events seen in both observations (e.g. Thompson et al., 2002) and simulated in the high-top models. A comparison of high-top and low-top seasonal forecasts for the northern winter of 2009–2010 (Fereday et al., 2012) showed that the low-top models respond to El Niño forcing in the same way as the high-top models, but more weakly due to the limited stratospheric representation. The high-top runs also showed the SSW impact on surface climate, with a descending signal in zonal mean zonal wind reaching the troposphere in late winter and leading to cold, blocked conditions in the middle and high latitudes.

As already discussed, Marshall and Scaife (2010) suggested that the enhanced predictability in the high-top models may be the result of earlier initialisation of the downward propagating SSW signal and preconditioning of the stratosphere. Their results are consistent with Xu et al. (2009), who demonstrated a clear SSW signal in the upper mesosphere that precedes the stratospheric signal at 10 hPa by 1–2 days. Furthermore, Lee et al. (2009) showed that in the case of the 2006 SSW event significant negative NAM signals appeared in the mesosphere during early January, but after mid-January in the stratosphere below 10 hPa. Coy et al. (2011) used a surface to 90 km data assimilation system to examine the 2009 SSW event and showed that wind reversals at high northern latitudes occurred first in the upper mesosphere, about a week prior to those at 10 hPa. Thus, resolving the upper stratosphere and lower mesosphere is important to improved predictability. Indeed, McTaggart-Cowan et al. (2011) demonstrate that a better representation of the stratosphere in an NWP model improves tropospheric forecasts on time-scales of 2–5 days, based on a case-study of the 2007 vortex displacement event.

Models used for medium-range weather forecasts (i.e. lead time less than 30 days) have also demonstrated the benefits of the inclusion of the stratosphere. However, there are fewer studies demonstrating the additional skill at shorter time-scales or the benefit of using the horizontal resolutions appropriate to weather forecasting. Mahmood (2013) compared results from high-top and low-top versions of a higher-resolution NWP model and showed the benefits for the 2009–2010 SSW event after as little as 5 days into the forecast. Gerber et al. (2012) show significant improvements in 1000 hPa geopotential height anomaly correlations out to 2–5 days in both the Northern and Southern Hemisphere from a major stratosphere-focused upgrade to the operational NOGAPS NWP system.

Roff et al. (2011) focused on the extended-range forecast skill that may be gained by the inclusion of a stratosphere, in Southern Hemisphere spring when there is a strong coupling between stratosphere and troposphere (e.g. Graversen and Christiansen, 2003; Thompson et al., 2005). Figure 6 shows the percentage improvement in the prediction of polar cap geopotential height (south of 60°S latitude) in high-top vs. low-top versions of the model. The experimental configuration consisted of running 30-day ensemble forecasts over three decades for two model configurations which differed only in the vertical resolution in the stratosphere (above 10 hPa): low-top configuration (L38) has 10 levels between 100 hPa and model top at ~5.8 hPa and high-top configuration (L50) has 22 levels between 100 hPa and model top at ~0.2 hPa (see Roff et al. (2011) for more details). Below 100 hPa both configurations had the same 28 levels. The high-top model showed improved forecast skill in the troposphere 3–4 weeks into the forecast as shown in Figure 6. Relative to the low-top model, the high-top model had ~5–7% lower forecast error in the geopotential height field in the troposphere. Tropospheric improvements are significant during most of the days but not all along as shown in Figure 6(b). Son et al. (2013) also recently showed that Southern Hemisphere spring prediction could be improved by considering stratospheric variability a month in advance. These results suggest that the improved representation of the stratosphere adds skill to tropospheric predictions.

4.2. Perturbation experiments

The perturbation set of experiments involve examining the transient response of the troposphere to stratospheric perturbations of some description. There are numerous ways in which this has been performed, from changing the diffusion parameter in the stratosphere (e.g. Boville, 1984; Boville and Baumhefner, 1990), to applying varying heating rates to force changes to the stratospheric zonal mean wind (e.g. Kodera et al., 1990), to directly damping the zonal wind within the polar vortex (e.g. Scaife et al., 2005).

Charlton et al. (2004) examined changes to the tropospheric forecast skill of the ECMWF model for three case-studies in which stratospheric initial conditions were artificially degraded
to represent the opposite phase of the stratospheric annular mode. The forecasts with degraded stratospheric initial conditions produced less skillful tropospheric forecasts, with an average decrease of the 500 hPa geopotential height anomaly correlation of between 5 and 10% after 5 days of the forecast. Jung and Barkmeijer (2006) extended this result by applying forcing optimised to produce rapid changes to the stratospheric vortex for an ensemble of sixty different 40-day forecasts. Their results showed a statistically significant tropospheric response to the stratospheric perturbation in just a few days that projected onto the NAO and a shift in the storm-track regions. However, studies like these raise the question of the realism of the experimental forcing. Thus, although studies of this type highlight the fact that the stratosphere does change the tropospheric circulation, they also suggest that this is only true in extreme cases. Interestingly Jung and Barkmeijer suggest that the tropospheric response is linear to the stratospheric perturbations (which is consistent with analysis of NAM index data by Baldwin et al. (2003) and Charlton et al. (2003)) and that large-scale dynamics mediate the stratosphere–troposphere link. The experiments of Cheung et al. (2014) showed that, even for some moderately large stratospheric forcings (in their case differences between climatological and observed ozone during the anomalously cold northern stratosphere in March 2011) differences in the skill of tropospheric forecasts can be small for individual case-studies.

Nonetheless, Scaife and Knight (2008) showed that perturbation experiments can be used to study the impact of stratospheric variability on the troposphere for case-studies of particular interest to forecasting centres. In their specific example, by adding artificial perturbations to the stratospheric zonal wind they were able to simulate the SSW event that occurred in January 2006. Artificially imposing this warming in the stratosphere was seen to lead to a strong cooling effect over northern Europe in the late winter similar to that observed in this and other events (e.g. Charlton et al., 2004; Jung and Barkmeijer, 2006) and more than 2 °C colder than a simulation which did not simulate the SSW.

4.3. Relaxation experiments

Relaxation experiments involve nudging certain regions of the atmosphere towards re-analysis data and so artificially suppress the development of forecast error. For the purposes of estimating the impact of stratospheric conditions on a tropospheric forecast, this type of experiment makes it possible to estimate an upper bound on the impact of an improved stratospheric forecast on tropospheric forecast skill. The underlying assumption made is that improving the stratospheric representation and reducing stratospheric model error would lead to improved tropospheric forecasts.

On the seasonal time-scale, Douville et al. (2009) showed a strong improvement of the simulation of wintertime European climate and the NAO in simulations in which stratospheric conditions were nudged toward the ERA-40 re-analysis. Jung et al. (2010) applied similar techniques to study the origin of forecast error on sub-seasonal time-scales. They showed that, even with moderate stratospheric relaxation, there was a more than 10% reduction in forecast error on forecast ranges beyond 7 days for a series of winter forecasts using the ECMWF model. Similar relaxation experiments by Greatbatch et al. (2012) suggest that the impact of stratospheric variability is much stronger in the Atlantic sector than in the Pacific sector.

Jung et al. (2010, 2011) used similar techniques to diagnose the origin of the cold winters of 2005–2006 and 2009–2010, respectively. For the 2005–2006 winter, they agree with Scaife and Knight (2008) that a midwinter SSW may have played a role in the extreme cold in Europe, but argue that conditions in the tropical stratosphere (QBO-E) and in the tropical troposphere (La Niña) were more important in this event. For the 2009–2010 winter, they find no evidence that this event was linked to stratospheric variability. In contrast, both Ouzeau et al. (2011) and Fereday et al. (2012) show a significant role for stratospheric variability in producing the very cold anomalies over Europe in winter 2009–2010.

4.4. Conditional hindcasting

In contrast to the perturbation and relaxation experiments described in the previous section, this final approach does not involve any artificial perturbations to the stratosphere or changes in its representation. Instead, the stratospheric impact on tropospheric forecast skill is quantified by contrasting hindcasts with different stratospheric conditions. Mukougawa et al. (2009) found that the hindcast skill of upper tropospheric circulation anomalies is significantly larger when initialized at times when the stratospheric vortex is weak compared to similar hindcasts initialized when the vortex is strong. Gerber et al. (2009) considered hindcasts around SSW events in an idealized atmospheric model. They found that a negative shift in the

Figure 6. (a) The percentage improvement in polar cap (south of 60°S) geopotential height prediction in the high-top (L50) model in comparison to the low-top (L38) model calculated as skill scores ratio. (b) The significance of the improvement. Reprinted from Roff et al. (2011) with permission from the American Geophysical Union.
NAM throughout the troposphere could be forecast at long lead times (about a month), but only if the hindcasts were initialized sufficiently close to the SSW so that the event itself could reliably be captured.

These exploratory studies were expanded in a systematic analysis of a comprehensive seasonal forecasting system by Sigmond et al. (2013) to assess the model’s ability to capture observed NAM indices, sea-level pressure, surface temperature and precipitation following SSWs, as shown in Figure 7. They compared hindcasts initialized at the time of SSW events with hindcasts initialized on similar dates with normal stratospheric conditions. In addition to capturing the influence of SSWs on the tropospheric NAM, they show that there is significant, improved sub-seasonal to seasonal hindcast skill of surface temperature and precipitation for the forecasts initialized around the time of SSWs in comparison to forecasts that were not initialized during SSWs (see e.g. Figure 7).

Finally, we also note that there has been some suggestions, using composite analysis, that the tropospheric response to splitting vs. displacement types of SSW may be different (Mitchell et al., 2013). This would be an interesting topic to explore in future with an appropriately large number of cases.

5. Discussion and current issues

Since at least the early 1960s, there has been an interest in understanding how stratospheric variability might be connected to the troposphere (e.g. Labitzke, 1965). However, in the last 15 years, as NWP models have improved with increased sophistication in their representation of the stratosphere, there has been interest in how stratospheric EVEs might be used to improve tropospheric weather forecasts and how the predictability of stratospheric EVEs might be improved and extended. In this review, we synthesized this body of work to provide guidance on our current, quantitative understanding of EVEs and how the stratosphere could be exploited to improve tropospheric predictability.

Studies have shown that stratospheric EVEs are predictable in modern NWP models, out to 10 days or more in some cases. However, the predictability of EVEs is limited both by the skill of forecasting the tropospheric planetary waves, which are the precursors of these events, and by model biases in the stratosphere. On longer time-scales there is a plethora of processes in the Earth system providing boundary condition forcing which increase the probability of EVEs. In this sense, EVEs may act as an important bridge linking sub-seasonal forcing in one part of the globe with impacts elsewhere because of the large horizontal scales typical of the wintertime stratospheric flow.

Several modelling techniques are available to assess the impact of stratospheric conditions on tropospheric forecasts at lead times of greater than 5 days. Methods which directly assess the overall impact of the stratosphere on tropospheric forecast skill include degrading the representation of the stratosphere by restricting the stratospheric resolution and raising/lowering the top level of the model. Other experimental designs are more suited to assessing the importance for tropospheric forecast skill of the evolution of the stratospheric state in different regions and at different times. Both perturbation techniques (adding additional artificial forcing to the stratosphere) and relaxation techniques (damping the stratospheric state towards observations) have been used to quantify the role of the stratosphere in recent extreme winter seasons in the Northern Hemisphere. An emerging area is the use of large forecast archives to contrast the performance of different models or sets of models based on their stratospheric skill or initialization time.

Operational forecasting centres have recognised the potential of improving the resolution of the stratosphere in their numerical models to enhance tropospheric forecast skill. Among the operational centres, the ECMWF took an early lead in improving the representation of the stratosphere in its model by raising the model top and introducing more levels in the stratosphere (Simmons et al., 1989, 1999; Untch et al., 1999). In later years a number of operational models increased stratospheric resolution and raised their model top. Currently, it is very common to see a model top above 1 hPa. The increased resolution of the stratosphere resulted in enhanced model predictive skill particularly during extreme stratospheric events in winter such as SSWs. Recent examples include Met Office forecasts of the 2012–2013 winter which explicitly made use of forecasts of a midwinter SSW to change their monthly-range guidance and recent developments of the Canadian Global Weather Forecasting System which showed gains in skill from raising the model lid above the 0.1 hPa level (Charron et al., 2012).

However, what is also clear from this review of the literature is that understanding of the predictability of EVEs is still relatively limited. Many studies report significant differences in the ability of a given numerical model to simulate similar SSWs. To improve the ability of numerical models to represent the stratosphere and benefit from improved stratospheric forecast skill, a wider study of stratospheric predictability with comparisons among different numerical models is required.

Specifically, the following questions need to be addressed:

- To what extent is the coupling of stratospheric EVEs to the troposphere determined by the state of the troposphere during the event?
- Are some EVEs more predictable than others? For example, is it easier to predict vortex-splitting SSWs than vortex-displacement SSWs or vice versa?
- How far in advance can EVEs be predicted such that resolving the EVE in a forecast can add skill to tropospheric forecasts?
- Which stratospheric processes, both resolved and unresolved, need to be captured by models to gain optimal stratospheric predictability?

Figure 7. The forecast skill (quantified by the Correlation Skill Score or CSS) of (a) the NAM at 100 and 1000 hPa, (b) the NAO index, (c) the surface temperature over northern Russia and eastern Canada, (d) the North Atlantic precipitation gradient, and the forecast skill averaged over 20–90° N for (b) SLP, (c) surface temperature and (d) precipitation. Pink bars represent forecasts initialized during SSWs and blue bars represent forecasts that are not initialized during SSWs. The forecast range is 16–60 days. The difference between the forecast skills are statistically significant at the 95% level where the confidence interval indicated by the thick brown lines do not overlap (see Sigmond et al. (2013) for more details).
Addressing these scientific and technical questions requires collaboration among the parts of the scientific community interested in stratospheric predictability (both stratospheric dynamists and forecast providers). It requires planned experiments that objectively compare the stratospheric predictability skills of different numerical models to understand its source. To achieve these objectives, a Stratospheric Processes and their Role in Climate (SPARC) supported project, the Stratospheric Network for the Assessment of Predictability (SNAP), has recently been initiated (Charlton-Perez and Jackson, 2012; Tripathi et al., 2013). SNAP provides a central forum by which expertise can be pooled and information and knowledge centralized and regularly shared (http://www.sparsenap.org) and involves all the authors of this study. The aim of SNAP is to design and perform an intercomparison of stratospheric predictability by examining multiple EVEs using multiple operational NWP models.

Acknowledgements

This work is supported by the Natural Environmental Research Council (NERC) funded project Stratospheric Network for the Assessment of Predictability (SNAP) (Grant H15476600) and partially supported by the SPARC. ACP and RGH acknowledge funding through the EU ARISE project (Grant 284387) (EU-FP7). We also acknowledge Steven Pawson and Lawrence Coy from NASA for providing Figure 1. We wish to thank Lorenzo Polvani from Columbia University for providing Figure 4 and Amy Butler from NOAA for her contribution to Figure 5. We thank Adrian Simmons of ECMWF for his insightful review and two anonymous reviewers for their comments and suggestions that improved the quality of the manuscript.

References


Baldwin MP, O’Sullivan D. 1995. Stratospheric effects of ENSO related tropospheric predictability (SNAP), has recently been initiated (Charlton-Perez and Jackson, 2012; Tripathi et al., 2013). SNAP provides a central forum by which expertise can be pooled and information and knowledge centralized and regularly shared (http://www.sparsenap.org) and involves all the authors of this study. The aim of SNAP is to design and perform an intercomparison of stratospheric predictability by examining multiple EVEs using multiple operational NWP models.

Acknowledgements

This work is supported by the Natural Environmental Research Council (NERC) funded project Stratospheric Network for the Assessment of Predictability (SNAP) (Grant H15476600) and partially supported by the SPARC. ACP and RGH acknowledge funding through the EU ARISE project (Grant 284387) (EU-FP7). We also acknowledge Steven Pawson and Lawrence Coy from NASA for providing Figure 1. We wish to thank Lorenzo Polvani from Columbia University for providing Figure 4 and Amy Butler from NOAA for her contribution to Figure 5. We thank Adrian Simmons of ECMWF for his insightful review and two anonymous reviewers for their comments and suggestions that improved the quality of the manuscript.

References


Kuroda Y, Kodera K. 1999. Role of planetary waves in the strato-


