The potential impact of upper stratospheric measurements on sub-seasonal forecasts in the extra-tropics

Christopher F. Lee, Pieter Smets, Andrew J. Charlton-Perez, Lőslo Evers, R. Giles Harrison, Graeme Marlton

Abstract This chapter examines the potential improvements in tropospheric weather forecasts that might arise from an improved representation of the upper stratospheric state. First, the chapter reviews current operational practice regarding observation of the atmosphere and the relative paucity of observations in the altitude range 40 to 70km. Then we describe in detailed some idealised model calculations to quantify the potential gain in skill available from improved monitoring in this region. The idealised model experiments use a relaxation technique with the Hadley Centre General Environment Model, to assess the potential gain in skill from improved knowledge both the whole stratosphere and the upper stratosphere. At weather forecasting time-scales (up to forecast day 30), improved knowledge of the stratosphere, close to the onset of a sudden stratospheric warming, improves forecasts of the tropospheric norther annular mode. Whole-stratosphere information significantly improved average surface temperature anomalies over northern North America, whilst upper-stratosphere information improved anomalies over Central Siberia. These results suggest any new observational technique which can improve montioring of the 40 to 70km region would likely benefit tropospheric forecast skill during winter time.

Lee, Charlton-Perez, Harrison and Marlton
Department of Meteorology, University of Reading e-mail: a.j.charlton-perez@reading.ac.uk

Smets and Evers
Seismology Division, Royal Netherlands Meteorological Institute, Wilhelminalaan 10, 3732 GK De Bilt, The Netherlands

Department of Geoscience and Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

1 Introduction

Variability in the extra-tropical stratosphere is known to influence tropospheric weather conditions, and their predictability in numerical models. The origins of this predictability were reviewed by Tripathi et al. (2014), and have been demonstrated to have an impact on tropospheric forecast skill on the timescales of routine weather forecasting by, for example, Jung et al. (2010). Our ability to exploit this predictability is limited both by the availability of routine stratospheric observations and uncertainties in the mechanism for stratosphere-troposphere coupling and its representation in numerical models. Additional stratospheric observations could be provided by a suite of novel observation techniques (for example, those studied by the Atmospheric Dynamics Research Infrastructure in Europe - ARISE - project and discussed in this book). Before investing in new operational monitoring networks and the data assimilation methods needed to incorporate new observations into Numerical Weather Prediction (NWP) suites, it is necessary to assess the potential benefit of the new technique to fill current or future gaps in the observing system and its potential impact on forecast skill. This chapter attempts to take the first step towards assessing the benefit of the ARISE measurement network. First we review the current observing system and identify the gap that exists in the region 40-70km. Secondly, we perform idealised experiments to quantify the potential value of new observations in this region for NWP. An idealised approach to estimating the value of additional stratospheric observations is adopted because, at this early stage, their operational characteristics are unknown and performing an experiment in which pseudo-observations were assimilated into a full Numerical Weather prediction (NWP) system would be premature.

2 Atmospheric measurements used for NWP

A broad diversity of observational techniques are used to capture the state of the atmosphere. Only a limited number offer observation of the higher stratosphere, and above ($<$ 100 hPa). In this section the following instruments, commonly assimilated in NWP models or used for middle atmosphere research, are briefly described :

- Radiosondes
- Remote Sounding : nadir, zenith, and limb.
- Stratospheric Wind Profers.
- Mesosphere Radars
- Atmospheric Lidar

Following the overview of their operation an indication of the vertical range that the observation types cover is given in Fig. 1

2.1 Radiosondes

A radiosonde is a small instrument package which communicates with a ground station using radio signals. The precursors of modern radiosondes were developed in the 1880s, and the use of sondes was routine by the beginning of the 20th century. They are now the longest record of the "higher" atmosphere (Jeannet et al., 2008). Early observations were limited to altitudes below 16 km (100 hPa). Since then radiosondes have become one of the few instruments delivering in-situ observations of pressure, geopotential height, temperature, dewpoint depression, wind direction and wind speed. The development of cheap acquisition systems that can be attached to standard radiosondes (for example Harrison et al., 2012) allows the incorporation of additional instruments with standard packages; for example, improved resolution of cloud-to-air transitions can be achieved using radiometers (Nicoll and Harrison, 2012), with other instruments measuring aerosol optical depth, light scattering and absorption, gas concentrations (for example, ozone, carbon dioxide, and methane (Gcos-134, 2009), atmospheric charge, and electric fields).

The instrument package hangs approximately 25 m below a weather balloon. As the radiosonde rises at about 5 m/s, sensors on the radiosonde perform different measurements. The sensors are linked to a battery powered radio transmitter, broadcasting the measurements to a sensitive ground tracking antenna on a VHF radio frequency. Wind velocity and direction are obtained by tracking the sonde in flight using GPS or a radio direction finding antenna. The latest radiosonde systems are equipped with a standard C/A code-correlated GPS receivers and a ground station differential GPS receivers. The resulting data is used to accurately calculate the wind velocity and direction of the sonde's ascent. Wang et al. (2009) shows periodic variations of the vertical wind velocity in the stratosphere, resulting from radiosonde data, related to gravity waves with an averaged horizontal wavelength of 15 km; however, below 5 km, the vertical velocity is overestimated by the sondes. Radiosondes are typically released every 6 to 24 hours, at several hundred locations globally, with the majority in the Northern Hemisphere (**?**).

Temperature is typically measured by a small rod or bead thermistor. The temperature instrument is usually silver coated to minimize radiational heating effects. Radiosondes make use of thermodynamic variables to determine the altitude of the balloon, and use satellite-based Global Positioning Systems (GPS) to determine their position. Reference radiosonde instruments, used in the Global Climate Observing System (GCOS) Reference Upper-Air Network (GRUAN), have a temperature accuracy of 0.1 K (after correction) in the troposphere and 0.2 K in the stratosphere; with a vertical resolution of 0.1 km from ground level to 30 km, and 0.5 km above 30 km (Gcos-134, 2009). These instruments make use of GPS double differencing for positioning and wind velocity, to increase the accuracy. The wind speed accuracy is 0.5 m/s with a wind direction accuracy of 5°, the pressure accuracy is 0.1 hPa and the geopotential height accuracy is $<$5 m.

Related in-situ measurements have in the past been provided by rocket-sondes (although their use has declined significantly since the 1970s), from drop-sondes

released from aircraft during measurement campaigns and from Aircraft Meteorological Data Relay (AMDAR) on commercial aircraft.

2.2 Remote Atmospheric Sounding

Remote atmospheric "sounding" is the process of deriving (vertical) distributions of temperature, and atmospheric constituents such as humidity or ozone, from radiative signals (Okamoto, 2008). It is a commonly a passive satellite based technique, essential for obtaining global meteorological observations.

2.2.1 Nadir Temperature Profiling

A satellite equipped with a sounding instrument (a passive sensor) captures the upward radiation from both atmosphere and surface. The observed radiation information is the sum of all radiated signals from ground up to the top of the atmosphere (TOA). The vertical distribution of radiation for specific wavenumbers can be inferred where multiple receiver channels, each characterized by a different sensitivity range for a specific frequency range, are used. For each channel the contribution of a layer of the atmosphere varies according to a bell-shaped weighting function, with the largest contribution at an intermediate altitude. A drawback of discretisation by WF is smoothing of the profile. Smoothing can be reduced by using more channels, with maxima at different altitudes. The maximum number of channels is limited by the frequency, bandwidth, and radiative characteristics of the atmosphere. Therefore, nadir sounding instruments with two different frequency ranges exist, namely infrared (IR) and microwave (MW) sounders. IR sounders can obtain many more channels, but they cannot penetrate clouds and have a smaller footprint. Desired resolution, and application, determine whether MW or IR sounders are used.

Nadir sounding is very important for numerical weather prediction. With five operational satellite platforms, the AMSU-A instruments on board the NOAA polar orbiting and European Metop satellites series is the flagship for NWP data assimilation, having delivered global temperature records for more than a decade (Kidwell et al., 2009). AMSU, a 20-channel microwave sounder, comprises three separate instruments: AMSU-A1 and AMSU-A2, which have the 15 lower frequency channels (23.889.0 GHz) primarily for temperature sounding (surface to 2 hPa or 43 km); and AMSU-B which has five higher frequency channels (89190 GHz) primarily for humidity sounding (surface to 200 hPa or 12 km).

The Atmospheric Infrared Sounder (AIRS), flown aboard Aqua, is the first of a new generation of advanced satellite-based atmospheric sounders, with the capability to obtain high-vertical resolution profiles of temperature and water vapor (Divakarla et al., 2006). The instrument is a cooled grating spectrometer that provides 2378 channels covering the IR spectrum, from 650 to 2675 cm^1, delivering high-accuracy temperature profiles from the ground up to 2 hPa (43 km). Tobin et al.

(2006) compared the retrievals with well-characterised ground based sites, showing that AIRS retrievals in the tropics have very good performance, with RMS errors approaching the theoretical limit of 1 K in one kilometre layers below 100 hPa (16 km). Midlatitude retrievals indicate poorer RMS errors ranging from 1 to 2 K in 1 km layers.

Horizontal resolution is characterized by the sensor field-of-view (FOV), the channel frequency, and the viewing angle. Temporal resolution depends on the orbit of the satellite, the maximum off-nadir looking angle, and the number of satellites. For AMSU, flown by the NOAA KLM spacecrafts, the temporal resolution using 1 satellite varies from 90 minutes at the poles up to 6 days at the equator. The temporal resolution is decreased using multiple satellites. Using 7 satellites, (as assimilated for example by ECMWF) the global coverage from equator to 60°varies between 76% and 84%. EUMETSAT METEOSAT geostationary satellites have a much better temporal resolution and are used (for example) in the ECMWF IFS to improve the water vapour resolution near the equator and to reduce the bias of IR sounders such as AMSU-B (Szyndel et al., 2005).

A major problem with nadir sounders is bias correction. Since most AMSU channels have a beam position-dependent bias, it is crucial to remove such a bias to provide useful profiles of the atmosphere. Typically limb-sounding data (for example GPS-RO) is used for this purpose.

2.2.2 Limb Sounding

Limb sounding is quite similar to nadir sounding, except that the satellite measures only the limb of the atmosphere. Limb sounders do not point towards the Earth's surface, but look at the atmosphere from the side, measuring the attenuation of, for example, solar radiation (passive) or radio waves (active). Limb sounding radiometers measure in some channels whose characteristics are comparable to those used in vertical sounding and also use radio frequencies (for example from GPS signals in the CHAMP constellation) (Okamoto, 2008).

Vertical information is extracted using a progressive scan of many paths, resulting in a much finer vertical resolution (0.52 km), but poorer horizontal resolu on ($>$100 km), compared to nadir sounding. Limb sounding is also not influenced by surface interactions. Examples of limb-sounding sensors are ENVISAT/MIPAS, Aura/MLS, and GPS Radio Occultation (RO). For NWP, GPS-RO is most widely used. The accuracy of GPS-RO derived temperature depends on precise orbit determination and the thermal noise of receiver, along with other atmospheric sources of error. GPS-RO temperature data is globally measured with an accuracy of 1 K from the lower troposphere to 40 km (Kuo et al., 2005). The vertical resolution of the processed data is high, 0.5 km in the troposphere up to 1.5 km in the stratosphere (Steiner et al., 2007), and with a horizontal resolution of 100 to 300 km.

2.3 Wind Profiling Radars in the Stratosphere

Radar wind profilers exploit Bragg scatter (scattering from slight changes in atmospheric refractive index) to infer wind velocities. Echoes from a moving scatterer are Doppler shifted by an amount proportional to their speed, so a spectrum of atmospheric motions can be constructed from the echo frequencies observed by the radar. Different radars have been designed to concentrate on different regions of the atmosphere, with most Mesosphere-Stratosphere-Troposphere (MST) radars observing at least the lower Stratosphere; ceilings vary, with some reaching only around 7 km, but others measuring above 20 km. The latter operate at VHF frequencies (usually around 40 MHz), with UHF instruments concentrating on the boundary layer/lower troposphere. The radar antenna comprises many (typically hundreds) of individual antennae. The signals from each of the antennae constructively/destructively interfere to create the radar beam. Altering the phase of the signal, transmitted at different regions of the array, alters the beam pointing direction - a 'Phased Array'.

Most MST radars can achieve a vertical resolution of 150 m to 300 m; temporal resolutions vary, but typically fall between 2 and 5 minutes for routine measurements. Temporal resolution is determined by the number of radar beams. MST radar data is assimilated into some models (for example, the UK Met Office assimilates winds from the Capel Dewi MST radar up to 15 km, as well as measurements from the Canadian Q-NET instruments). A clear limitation, though, is the ceiling of the instruments. Echo power is proportional to the square of the range to the target; doubling a radar's transmitting power only increases altitude coverage by a few kilometres at lower stratsospheric heights.

2.4 Mesosphere Radars

2.4.1 MST Radars

There is a large gap between the MST observation ceiling in the stratosphere, and the measurement floor in the mesosphere. From approximately 70 km, electron densities can be large enough to generate radio refractive index changes observable by radar. The processes governing enhancements in radar detectable electron density is a subject of much debate; many days can pass without the detec on of any mesospheric echoes, and winds can only be derived where such echoes exist. Nevertheless, where echoes do occur, the multi-beam approach used in the stratosphere and troposphere can be used to infer wind velocities. Echo layers are more common and widespread in summer (Polar Mesosphere Summer Echoes - PMSE), though winter echoes (PMWE) have been routinely observed. MST observations of the mesosphere are largely restricted to altitudes between 70 and 110 km.

2.4.2 MF Radars

Medium Frequency (MF) radars also exploit scattering from radio refractive index fluctuations in the upper mesosphere. Unlike MST radars though, they examine the correlation of echoes received by several groups of antennae (the Spaced Antenna technique - Doviak et al. (1996)), to infer wind velocities. An MF radar's altitude floor is limited by the same electron density effects that constrain MST radar observations in the mesosphere; nevertheless, they routinely provide wind velocities between 80 and 100 km, with most instruments providing coverage down to 70km.

2.4.3 Meteor Radars

Meteor radars use interferometry techniques (similar to the spaced antenna method) to measure the ionisation trails left by meteors. Those trails are advected by the background wind, and so can be used to quantify wind speed and direction. Meteor radars typically measure between 80 and 100 km; outside that range there are typically too few meteors to obtain reliable measurements (Day and Mitchell, 2010).

2.5 Atmospheric Lidar

Lidar operates on the same principle as radar, using light rather than radio signals to exploit different scattering mechanisms of gasses and aerosols; different techniques are used to measure differing aspects, and regions, of the atmosphere. Of particular interest to stratospheric and mesospheric dynamics are lidar temperature observations, typically exploiting Rayleigh and/or Raman scatter: Rayleigh scatter (an elastic scattering process) can be used to infer temperatures where aerosols concentrations are negligible (typically above 30 km). Raman scatter is an inelastic process that alters the energy state of the scattering molecules. The lidar pulse stimulates rotational energy levels of Oxygen and Nitrogen; the strength of the resulting backscatter varies with temperature, with strengthening scatter from some modes, and weakening from others. Comparing scattering intensities from modes with opposing tendencies gives a measure of temperature that is independent of illuminating signal strength (i.e. independent of atmospheric attenua on along the signal path).

Typically the lidar techniques listed here are restricted to nightime observations, but some high power systems can operate in daylight. Rayleigh temperature measurements reach from around 30 km to the mesopause, but can be extended to lower altitudes with additional observations. A summary of lidar temperature uncertainties can be found in Keckhut et al. (2004) and are typically of the order of 1K.

2.6 Summary

For the instruments described above that form a routine part of most NWP modelling systems, a summary of the vertical sampling range is presented in Fig. 1. It highlights the lack of stratospheric wind measurements, and higher stratospheric and mesospheric temperature observations. Since lidar observations are not assimilated by NWP models, the data gap between 40 and 70 km is particularly stark. Radiosondes provide the longest data set, with routine soundings performed from the early 1900s. Temperature biases (< 1 K) have remained relatively constant over the last 20 years. The accuracy and bias of geopotential height and pressure has improved significantly since the introduction of GPS. Radiosondes are assumed to be unbiased, and are therefore often used as reference observations. Remote sounding is the only technique that can providing global coverage from one platform. Nadir sounding has the highest global coverage with the smallest temporal resolution. Although it has a bias problem, it is important for NWP. To reduce the bias a small set (a few hundred points distributed around the globe) of anchor points are used, comprising mainly measurements from radiosondes (in the troposphere) and GPS-RO (in the stratosphere up to 35 km). Mesosphere-Stratosphere-Troposphere (MST) radars are able to observe from ground to the lower stratosphere, and occasionally in the mesosphere. MST, MF, and Meteor radars are, at present, the only instruments capable of observing mesospheric winds. Currently only tropospheric and some lower stratospheric winds from MST radars are assimilated in some NWP models. Atmospheric lidar are capable of providing fine detail temperature observa ons of the stratosphere and (lower) mesosphere, typically operating only at night. Lidar data is not yet assimilated by NWP models. The instruments commonly assimilated by NWP models are remote sounders (GPS-RO, MW and IR) and radiosondes, alongside some wind profiling radar data.

In the near future (end-2017), the launch and operation of the ADM-Aeolus satellite, a space-borne, wind-profiling LIDAR is likely to have a significant impact on the data coverage for NWP in the 40-70km region. Nonetheless, any additional opportunity to increase observational coverage in this critical region would be welcome. In the context of this book, a number of ARISE measurement techniques, particularly infrasound inversion has the potential to fill this observational gap. In the next part of the chapter we present the results of a series of experiments which are designed to attempt to quanify the extent and location of potential gains in predictability which might result from additional observational capability in this region of the upper stratosphere.

3 Sudden Stratospheric Warmings

The influence of the stratosphere on tropospheric variability and predictability is strongest during mid-winter and occurs when dramatic changes to the climatological stratospheric polar vortex occur. The stratospheric polar vortex forms at the

beginning of each winter, because the ozone-rich air loses its solar heat source, and cools, during the polar night. The temperature gradient induces strong westerly winds, sometimes in excess of 40 m/s, around the cold pool. The vortex breaks down each spring as the solar heat source returns, but it may also be disturbed before this final, seasonal, warming.

When these changes are large, and lead to warming of the polar cap they are known as sudden stratospheric warmings (SSWs). SSW events occur due to the breaking of planetary Rossby waves in the upper and middle stratosphere leading to a weakening of the polar night jet at the edge of the vortex and the formation of a critical line for further Rossby wave packets. SSW events are initiated in the upper stratosphere and typically descend into the middle and lower stratosphere over the course of a week.

Several definitions of an SSW are used in the literature (see, for discussion, Butler et al. (2015)); in this chapter we use the zonal mean of the zonal wind velocity (ZMZW) at 10hPa and 60N as a measure of the vortex winds (Charlton and Polvani (2007)). A major SSW is defined as a reversal of the ZMZW to an easterly direction (i.e. negative velocities).

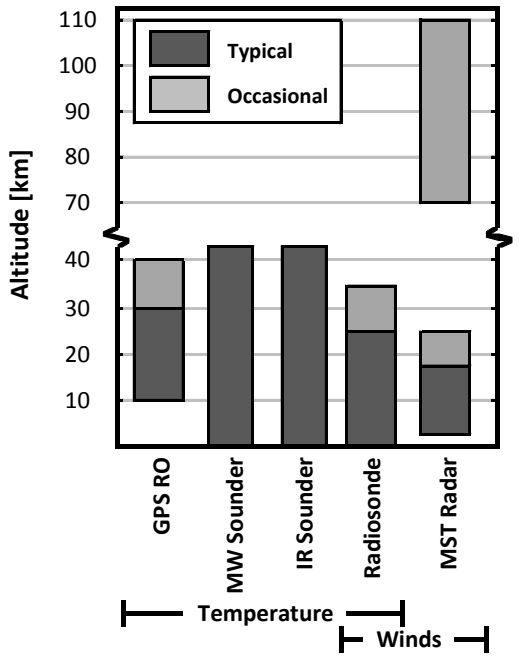

Fig. 1 Typical and occasional altitude ranges for the reviewed instruments. Note that wind profiling radars, both stratospheric and mesospheric, and lidar are not commonly assimilated by NWP models and so are not included in this figure

SSWs are observed at a frequency of around 6.2 per decade, in the Northern Hemisphere (Charlton and Polvani (2007)). In the Southern Hemisphere only one SSW has been recorded, in 2002 (Charlton et al., 2005). It is thought that the apparent stability of the Southern Hemisphere vortex is due to an absence of significant topography, capable of forcing planetary waves, at southern mid-latitudes. In this study we only examine the Northern Hemisphere.

SSWs can be classified into two types: a vortex split, or a vortex displacement (Charlton and Polvani (2007)). A vortex displacement is characterized by a displacement of the vortex away from the pole, followed by its distortion into a comma shape. A vortex split describes the division of the vortex into two parts. Vortex splits and displacements occur at roughly the same frequency, with splits mostly observed in January and February, and displacements occurring throughout the winter.

SSWs are known to influence surface weather conditions several months after the their onset (e.g. Thompson et al. (2002); Sigmond et al. (2013). They are characterised by a breakdown of the stratospheric polar vortex. Over a period of several weeks these anomalous flows, which exhibit a pattern similar to the negative phase of the Northern Annular Mode (NAM), can propagate to the surface. They are associated with surface temperatures anomalies of -1 to -2 K in Northern Europe and eastern North America Thompson and Wallace (2001); Thompson et al. (2002). Experiments with numerical weather prediction (NWP) models have shown a shift in the tropospheric jet and extra-tropical cyclone tracks, to lower latitudes, after SSWs. Such shifts are associated with enhanced South Atlantic, and reduced North Atlantic, precipitation; along with an increased frequency of severe storms in the UK and Southern Europe (Baldwin and Dunkerton, 2001; Gerber et al., 2009; Garfinkel et al., 2013; Sigmond et al., 2013). Consequently, improved stratospheric data assimilation may show promise for weather prediction, following SSWs and also following periods when the stratospheric vortex is anomalously strong (Tripathi et al., 2015).

Propagation of SSW circulation anomalies from 10 hPa to the surface takes around 10 days. These signals can persist in the troposphere and lower stratosphere for up to two months, with anomalies at 10hPa persisting for around 40 days (Baldwin and Dunkerton, 2001; Sigmond et al., 2013). The stratospheric signal does not always propagate into the troposphere, and modeling studies by Gerber et al. (2009) found that such propagation only occurred where the tropospheric NAM was initially negative or neutral. However, where the signal does propagate into the troposphere, its persistence has been shown to improve tropospheric forecast skill at the medium-range (Gerber et al., 2009; Jung et al., 2010; Sigmond et al., 2013). A detailed discussion of the origins of this skill is given by Tripathi et al. (2014). To exploit this skill, an accurate estimate of the initial stratospheric state is required, and the forecast must be initialised, at most, a couple of weeks prior to the SSW onset Nishii and Nakamura (2010), because at longer lead times the SSW may not be forecast. As an example, constrast the two ensemble forecasts of the SSW which occured during March 2016 in Fig. 2 initialised only four days apart but which show a significnatly different evolution for the polar vortex and a significantly different ensemble spread a few days after initialisation.

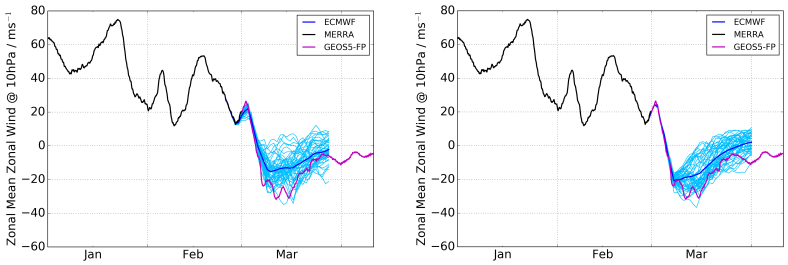

Fig. 2 ECMWF Ensemble forecasts (blue lines) of the zonal mean zonal wind at 10hPa and 60°N during winter and spring 2016. Black line shows MERRA re-analysis and GEOS operational analysis for the same quanitty. Left panel shows forecast initialised on 25th February and the right panel shows forecast initialised on the 29th February

The majority of stratospheric observations assimilated in NWP models are provided by satellite and GPS Radio Occultation (GPSRO) instruments, with some lower-stratospheric measurements from radar wind profilers and radiosondes; however, the upper stratosphere is not as well observed. This study seeks to assess the potential impact of the upper-stratosphere on tropospheric predictability, using relaxation experiments performed with the Met Office Unified Model. We focus on forecasts initialised close to SSWs, where a clear tropospheric signature has been demonstrated in previous studies. The experimental approach is presented in the next section, with the results and conclusions in Sections 3 and 4 respectively.

4 Potential Predictability Experiments

4.1 Methodology and Model used

In the experiment reported in this chapter we have used relaxation experiments to identify the response of the troposphere to changes in the stratospheric state. This approach allows multiple forecasts of the same SSW to be run, with small changes made to the stratospheric state. Any differences in the forecasts are directly attributable to the stratospheric changes.

The model used in the experiment is the Met Office HadGEM2 (Hadley centre General Environment Model version 2) high-top version (N96, L60), extending to 0.01hPa (approximately 85 km), with a slab ocean, and run at a temporal resolution of 20 minutes. The SSWs studied in the experiment are generated from an un-constrained 50-year long model run. SSWs were identified using the method of Charlton and Polvani (2007), where the SSW onset is defined as the date where the daily zonal mean zonal wind (ZMZW) at 10hPa and 60° N, becomes negative. Subsequent transitions to negative ZMZW in the following 30 days are regarded as the same event. 15 SSW cases were selected from the long model run for further study.

To avoid final warmings, and those associate with a weak vortex at the beginning of winter, only SSWs with onset dates lying between December and February were chosen. A five member ensemble forecast was generated for each case, initialised 6 - 10 days before the SSW onset, with ensemble members produced by perturbing the initial conditions at all levels. To generate the ensemble, perturbations were made to the vorticity field, at wavenumbers 4 - 10. Rather than add or subtract vorticity as has been done in some studies, e.g. Gerber et al. (2009); Simpson et al. (2011), we re-distribute vorticity by randomly shifting the phase of each wave; this preserves the spectral power at a given wavenumber. For a given wavenumber and ensemble member, the phase shift is the same at each model level. Examination of the corresponding ensemble spread indicated that they were comparable to those of forecasts produced by the ECMWF, and so are representative of typical weather forecast behavior. To quantify how an improved knowledge of the stratosphere would enhance tropospheric forecast skill, the ensembles were re-run (from the same perturbed initial conditions) with stratosphere wind velocities relaxed back to the truth run. Forecast wind velocity components (u) are relaxed towards the velocities of the truth run (u_T - Equation 1, where u_r is the relaxed wind velocity component). λ is a constant used to specify the strength of relaxation.

$$u_r = \lambda(u - u_T) + u \tag{1}$$

We follow the approach used in previous studies, such as Jung et al. (2010); Hitchcock and Simpson (2014), and nudge at every time step (i.e. every 20 minutes). In our experiment we use the nudging method of Telford et al. (2008), with $\lambda = 1$, such that $u_r = u_T$ at each time step. This approach effectively introduces a perfect representation of the stratospheric state, in the relaxed region (with respect to the truth run). Consequently these experiments identify the upper-limits of forecast improvement arising from additional stratospheric information. Even if stratospheric observations could generate a perfect representation of stratosphere in operational weather forecasts, this state would not persist throughout the model run. This approach helps to identify the regions of the stratosphere where such information is likely to be particularly valuable. Relaxation is performed above 100 hPa and above 40 km in separate experiments, to estimate the potential impacts of improved knowledge of the whole stratosphere, and upper stratosphere only, respectively. At the bottom of the relaxation region, λ is tapered from 1 to zero over three model levels. The impact of the relaxation experiments on tropospheric forecast skill is quantified over three regions in the Northern Hemisphere, using the Correlation Skill Score (CSS, Equation 2). The CSS compares the forecast (f) with observations (T), with respect to the climatology (c). (Note that the observations in our experiments are represented by un-constrained model runs; henceforth these are referred to as the truth runs.) In this study the climatology is taken from the long 50-year run, and the forecast and observations are the control (forecasts with no relaxation) and relaxed ensemble means respectively. The CSS is evaluated for 10 day forecast intervals, and over the following three regions (where stratospheric impacts are known to be important from previous studies): Eastern Canada, 45° - 70°N, 100° -

60°W; Western Europe, 45° - 70°N, 12°W - 35°E; and Northern Russia, 50° - 70°N, 30° - 130°E. Significance levels are calculated by bootstrapping, using random sampling with replacement of ensemble means from the 15 cases. The Northern Annular Mode (NAM) index is calculated using zonal mean zonal winds, following the method of Gerber et al. (2009). The Empirical Orthogonal Function (EOF), onto which the winds are projected, is calculated from the 50-year climate run.

$$CSS = \frac{\overline{(f-c)(T-c)}}{\sqrt{\overline{(f-c)^2}\,\overline{(T-c)^2}}} \tag{2}$$

4.2 Results and Discussion

First, we examine the nature of the truth states, generated by the climate run of the model, to assess their similarity to the observed behaviour of the stratosphere, and associated tropospheric phenomina. Charlton-Perez et al. (2013) compared the frequency of observed SSWs (no average, 6.2 per decade) with those generate in various CMIP5 (Coupled Model Intercomparison Project 5) models. The high-top HadGEM2 CC (Carbon Cycle) model use in the comparison generated an average of 8.0 SSWs per decade. However, the HadGEM2 model used in this comparison does not include the carbon cycle component, and generates SSWs at a frequency of 6.8 per decade. This is well within the confidence intervals of observed SSWs.

Second, we examine the tropospheric impact of SSWs, in the model, at time scales out to 60 days. The focus of this experiment is on shorter forecasts (out to 30 days), however these longer runs allow us to make an initial comparison to previous studies. Figure 3 shows the surface temperature, pressure, precipitation, and 500 hPa geopotential height anomalies between 15 - 60 days after SSW onset. Warm and cold surface temperature anomalies (Figure 3a) over eastern North America and Northern Europe, respectively, are consistent with the negative phase of the NAM, and the temperature signatures shown by Sigmond et al. (2013). The authors presented surface temperature maps from re-analysis of the Canadian Middle Atmosphere Model (CMAM). Their results showed that the surface temperature annomailes peaked between the Greenland and Eastern Canada, at around 1.4 - 1.8°C, similar to the values shown in Figure 3.a. Temperature anomalies over Western Canada are also consistent between the two studies, peaking at around 1.4°C. Over Northern Russia, CMAM forecast runs presented by Sigmond et al. (2013) produce large temperature anomalies associated with SSW events and this is replicated in our model runs, with temperature annomalies peaking at $<$ -2.2°C. Our model results show a slight over-estimate of surface temperature annomalies over Central Europe (of around 0.4°C), but warmer annomalies over Eastern Europe are captured in our results. Sigmond et al. (2013) also examined precipitation signals following SSWs, and found a shift in North Atlantic precipitation, consistent with an equator-ward shift in storm tracks. That feature is replicated in our results (Figure 3.b). Though the annoma-

lies observed in the two studies are comparable, Sigmond et al. (2013) showed that there can be considerable differences between modeled and observed precipitation anomalies, suggesting that the precipitation signal associated with SSWs is weaker than that of surface temperature. Surface pressure and geopotential height anomalies (Figure 3c and d) exhibit similar patterns to surface temperature, with positive and negative values over North America and Northern Europe.

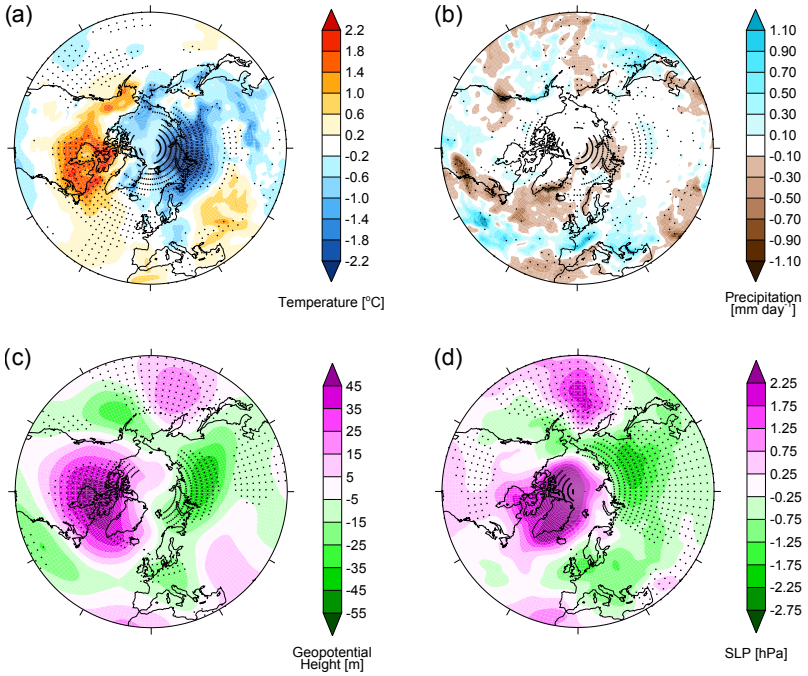

Fig. 3 Average anomalies of (a) surface temperature, (b) precipitation, (c) 500 hPa geopotential height, and (d) sea level pressure, 15 - 60 days after the onset of 15 SSWs in a 50-year unconstrained climate run of the HadGEM2 Unified Model. Stippling indicates regions where the anomalies are significant at the 95% level.

The correlation between surface pressure, surface temperature, and 500 hPa geopotential height anomalies, is indicative of the downward propagation of the SSW signals. This propagation, from the upper stratosphere into the troposphere, is evident, at sub-seasonal timescales, in the NAM index. Figures 4a - d show the NAM calculated for the different relaxed and control model runs. On average, the control runs underestimate the negative NAM index of the stratosphere seen in the truth runs (Panel e), though the difference is largely insignificant. This under-estimate is slightly reduced, in the first 15 days of the forecast, when relaxation of the upper-stratosphere is used (Panel g). In the lower troposphere, after 25 days, there are significant differences between the NAM indicies of the control and truth runs. The truth runs return to an approximately neutral NAM index after 25 days; but this is not

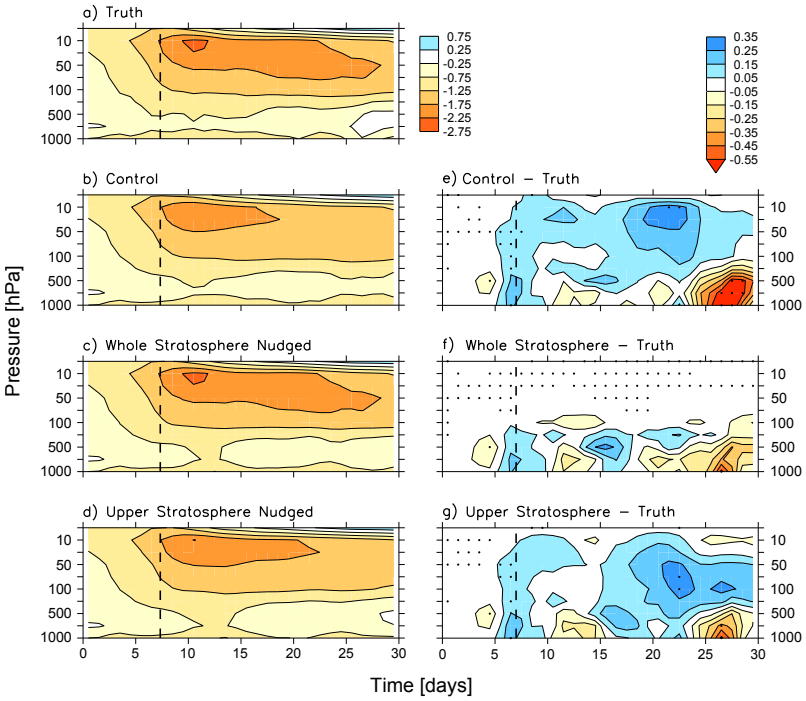

Fig. 4 NAM index averaged over 15 SSW cases for (a) truth run, (b) the control forecast ensemble mean, (c) forecasts with whole-stratosphere relaxation, (d) forecasts with upper-stratosphere relaxation, (e) control NAM index minus truth NAM index, (f) whole-stratosphere relaxation minus truth, (g) upper-stratosphere relaxation minus truth. The vertical dashed line indicates the average SSW onset date. Stippling indicates a statistically significantly difference at the 95% confidence level, calculated with bootstrapping, in figures e - g.

captured by the control, which remains significantly negative. This under-estimate is reduced with either relaxation of the whole- or upper-stratosphere, such that it is insignificant at most times and altitudes.

The negative NAM index in the stratosphere and upper troposphere, in the last 10 days of the forecasts, is associated with positive geopotential height anomalies over the pole, corresponding to the absence of the polar vortex (Figure 5a, which shows geopotential heights at 100hPa). Positive anomalies are also observed at 500hPa at the same time (Figure 5e), with a large negative anomaly over North Western Europe, which extends towards polar latitudes. The corresponding control runs fail to capture this structure at 500 hPa (Figure 5f); instead, the positive anomaly extends into the Arctic Ocean, and the negative anomaly over Western Europe is too weak. There is also a large negative anomaly over the coast of eastern North America. These anomalies contribute to the persistence of the negative NAM signal shown in Figure 4. Relaxing the whole- or upper-stratosphere reduces these anomalies (Figures 5g & h), consistent with the more neutral NAM indicies for these runs, shown in Figure 4. The impact of stratospheric relaxation at the surface is illustrated by the

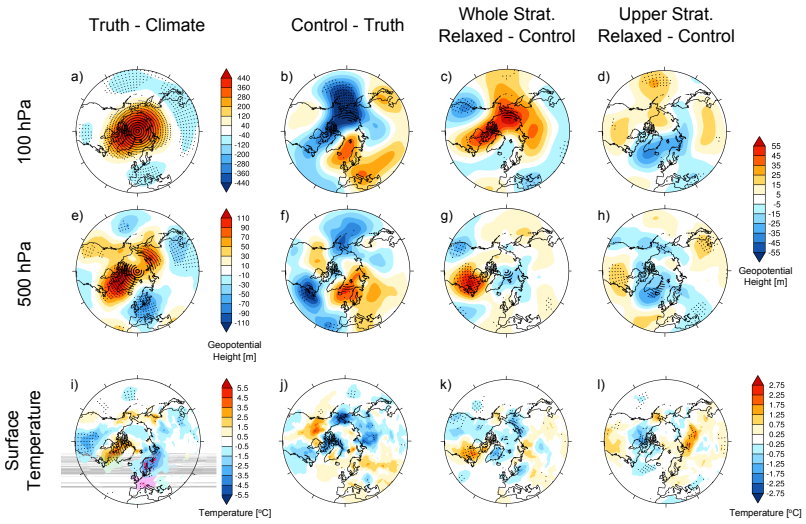

Fig. 5 Geopotential height and surface temperature anomalies, averaged over 15 SSW cases and forecast days 21 - 30. The top two rows (panes a - d and e - h) correspond to 100, and 500 hPa geopotential height anomalies, respectively. The bottom row shows surface temperature anomalies. The columns correspond to anomalies of (i) the truth runs with respect to the climate; (ii) the control runs with respect to the truth; (iii) the whole-stratosphere relaxed forecasts, and (iv) the upper-stratosphere relaxed forecasts, with respect to the control runs.

CSS of surface temperature (Figure 6). Over both eastern North America and Western Europe, relaxation of the whole-stratosphere improves the surface temperature CSS. These differences are not significant, compared to either the control or upper-stratosphere relaxation runs; however there is a significant improvement in the local temperatures over eastern North America, in the whole-stratosphere relaxed runs (Figures 5k).

The most significant improvement in surface temperature CSS is seen over Northern Russia, for the upper-stratosphere relaxed runs (Figure 6). Over much of this region the control forecast predicts surface tempeartures that are too cold (Figures 5j); however, this anomaly is corrected, over Central Siberia, in the upper-stratosphere relaxed runs (Figures 5l), contributing to the improvement in CSS. The origin of this improvement is not clear, however it strongly influences the CSS calculations. We have varied the area over which the Northern Russia CSS is evaluated (not shown), and where the evaluation area encompasses the Central Siberian region, the upper-stratosphere relaxation experiments consistently return a larger CSS than the other forecasts. It is important remember that the forecasts used here are shorter than those used in similar studies. This is particularly important in the context of the surface temperature patterns over Northern Russia: the typical SSW-related surface temperature anomalies are not as prominent at Day 30 as they are at Day 60 of the

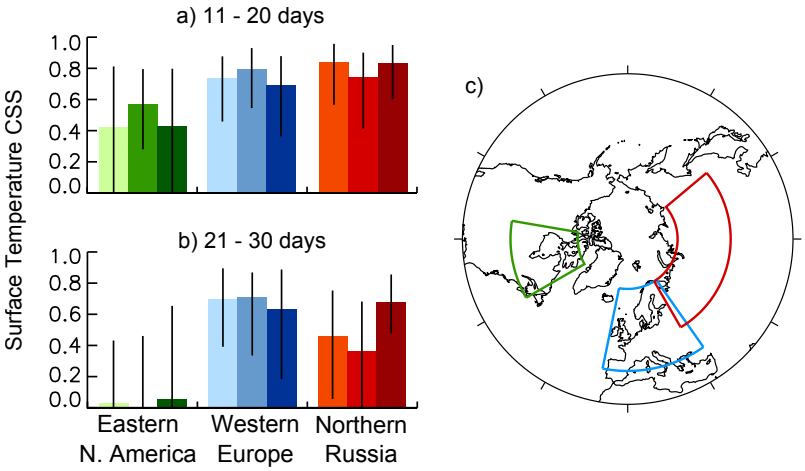

Fig. 6 CSS of surface temperature averaged over (a) days 11-20 and (b) 21-30; for eastern North America, Western Europe, and Northern Russia. The three boxes for each region give the CSS for (i) the control (light boxes), (ii) whole-stratosphere relaxed, (and (iii) upper-stratosphere relaxed (dark boxes) forecasts, from left to right. Black lines give the 95% confidence intervals, calculated with bootstrapping. The three analysis regions are highlighted in (c).

truth run (compare Figures 3a and 5i; note that Figure 3a uses a finer temperature resolution, for consistency with Sigmond et al., 2013). Over North-western Russia, the 21-30 day surface temperatures do not exhibit the cold anomalies seen between 15-60 days, suggesting that this feature appears later in the truth runs. However, the cold anomalies do begin to appear in control runs, so upper-stratosphere relaxation has moved the forecast away from this familiar pattern, towards the truth.

5 Conclusions

This chapter has shown that improved characterisation of the stratosphere in the region where stratopsheric observations are sparse can improve tropospheric forecast skill at the timescales of routine weather forecasting. The relaxation experiments presented here have focused on forecasts initialised close to SSWs, when there is known to be a correlation with surface weather patterns at sub-seasonal timescales (e.g. Sigmond et al., 2013). In this study we have examined shorter, 30-day, forecasts, where the average surface temperature anomalies that follow SSWs (warmer and cooler temperatures in eastern North America, and Northern Europe, respectively) are apparent, but not fully established.

The impact of additional information in the whole-stratosphere (above 100 hPa), and the upper-stratosphere only (above 40 km), have been examined separately. In both cases the additional information improved the forecast of the tropospheric NAM, capturing a return to a near-neutral NAM index, 25 days into the forecast

(15 days after SSW onset). Additional information in the whole stratosphere significantly improved the forecasting of local surface temperatures in eastern North America, over forecast days 21-30. Relaxation of the upper-stratosphere improved surface temperature forecasts over Central Siberia, over the same period.

A perfect model approach has been used in these experiments, so the results presented here represent the upper-limit of improvements that might be expected from more comprehensive measurements of the stratosphere. Nevertheless, the relative paucity of routine observations in the upper-stratosphere means that additional measurements in this region could make a significant contribution to the improvement of tropospheric forecasts. As noted elsewhere in this volume, ARISE measurements are one potential way to improve routine monitoring in the region between 40 to 70km, but much more work needs to be done to understand the reliability and error chararcteristics of such measurements before more detailed experiments which assess the direct benefit for NWP could be conducted.

Acknowledgements This work made use of the facilities of HECToR, the UK's former national high-performance computing service. Access was provided by the Natural Environment Research Council's High Performance Computing Facility. This research is part of the Atmospheric Dynamics Research Infrastructure in Europe (ARISE) project, funded by the European Union's Seventh Framework Program. We would like to thank Kirsty Hanley, Grenville Lister, Scott Osprey, Paul Telford, and Peter Watson, for their help setting up the experiments discussed here.

References

Baldwin, M. P. and Dunkerton, T. J. (2001). Stratospheric harbingers of anomalous weather regimes. Science, 294(5542):581–584.

Butler, A. H, Seidel, D. J. and Hardiman, S. C., Butchart, N., Birner, T. and Match, A. (2015). Defining sudden stratospheric warmings. Bull. Am. Met. Soc., doi: http://dx.doi.org/10.1175/BAMS-D-13-00173.1.

Charlton-Perez, A. J. and others (2013). On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models, J. Geophys. Res., doi:10.1002/jgrd.50125

Charlton, A. J. and Polvani, L. M. (2007). A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. Journal of Climate, 20(3):449–469.

Charlton, A. J., O'Neill, A., Lahoz, W. A. and Berrisford, P. (2005). The splitting of the stratospheric polar vortex in the Southern Hemisphere, September 2002: Dynamical evolution. J. Atmos. Sci., doi:10.1175/JAS-3318.1

Day, K. A. and Mitchell, N. J. (2010). The 16-day wave in the arctic and antarctic mesosphere and lower thermosphere. Atmospheric Chemistry and Physics, 10(3):1461–1472.

Divakarla, M. G., Barnet, C. D., Goldberg, M. D., McMillin, L. M., Maddy, E., Wolf, W., Zhou, L., and Lio, X. (2006). Validation of Atmospheric Infrared Sounder temperature and water vapor retrievals with matched radiosonde measurements and forecasts. Journal of Geophysical Research, 111:D09S15.

Doviak, R. J., Lataitis, R. J., and Holloway, C. L. (1996). Cross correlations and cross spectra for spaced antenna wind profilers .1. theorecal analysis. Radio Science, 31(1):157–180.

Garfinkel, C. I., Waugh, D. W., and Gerber, E. P. (2013). The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere. Journal of Climate, 26(6):2077–2095.

Gcos-134 (2009). Implementation Plan for the Global Climate Observing System Reference Upper Air Network 2009-2013. Technical report, World Meteorological Organization, WMO/TD No. 1506, Geneva, Switzerland.

Gerber, E. P., Orbe, C., and Polvani, L. M. (2009). Stratospheric influence on the tropospheric circulation revealed by idealized ensemble forecasts. Geophysical Research Letters, 36.

Harrison, R. G., Nicoll, K. A., and Lomas, A. G. (2012). Note: Programmable data acquisition system for research measurements from meteorological radiosondes. The Review of Scientific Instruments, 83(3):036106.

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

Jeannet, P., Bower, C., and Calpini, B. (2008). Global criteria for tracing the improvements of radiosondes over the last decades. Technical report, World Meteorological Organizaon, Instruments and Observing Methods Report No. 95, WMO/TD-No. 1433.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. J. Hydrometeor, 5(3):487–503.

Jung, T., Miller, M. J., and Palmer, T. N. (2010). Diagnosing the origin of extended-range forecast errors. Monthly Weather Review, 138(6):2434–2446.

Keckhut, P., McDermid, S., Swart, D., McGee, T., Godin-Beekmann, S., Adriani, A., Barnes, J., Baray, J. L., Bencherif, H., Claude, H., di Sarra, A. G., Fiocco, G., Hansen, G., Hauchecorne, A., Leblanc, T., Lee, C. H., Pal, S., Megie, G., Nakane, H., Neuber, R., Steinbrecht, W., and Thayer, J. (2004). Review of ozone and temperature lidar validations performed within the framework of the network for the detection of stratospheric change. Journal of Environmental Monitoring, 6(9):721–733

Kidwell, K. B., Goodrum, G., and Winston, W. (2009). NOAA KLM user's guide, with NOAA-N, -N' supplement. Technical report, National Environmental Satellite, Data, and Information Services, Silver Spring, MD.

Kuo, Y. H., Schreiner, W. S., Wang, J., Rossiter, D. L., and Zhang, Y. (2005). Comparison of GPS radio occultation soundings with radiosondes. Geophysical Research Letters, 32(5):L05817+.

Nicoll, K. A. and Harrison, R. G. (2012). Balloon-borne disposable radiometer for cloud detection. The Review of scientific instruments, 83(2):025111.

Nishii, K. and Nakamura, H. (2010). Three-dimensional evolution of ensemble forecast spread during the onset of a stratospheric sudden warming event in January 2006. Quarterly Journal of the Royal Meteorological Society, 136(649):894–905.

Okamoto, K. (2008). L5 Satellite Remote Sensing by Temperature and Humidity Sounders. In The Eighteenth IHP Training Course, Nagoya, Japan. Satellite Remote Sensing of Atmospheric Constituents.

Seidel, D. J., Gillett, N. P., Lanzante, J. R., Shine, K. P., and Thorne, P. W. (2011). Stratospheric temperature trends: our evolving understanding. WIREs Clim Change, 2(4):592–616.

Sigmond, M., Scinocca, J. F., Kharin, V. V., and Shepherd, T. G. (2013). Enhanced seasonal forecast skill following stratospheric sudden warmings. Nature Geoscience, 6(2):98–102.

Simpson, I. R., Hitchcock, P., Shepherd, T. G., and Scinocca, J. F. (2011). Stratospheric variability and tropospheric annular-mode timescales. Geophysical Research Letters, 38(20).

Steiner, A. K., Kirchengast, G., Borsche, M., Foelsche, U., and Schoengassner, T. (2007). A multi-year comparison of lower stratospheric temperatures from CHAMP radio occultation data with MSU/AMSU records. Journal of Geophysical Research, 112(D22):D22110+.

Szyndel, M., Kelly, G., and Thpaut, J. N. (2005). SEVIRI radiance assimilation at ECMWF. In ITSC XIV Proceedings, Bejing, China. International ATOVS working group.

Tan D. G. H. and Andersson E. (2005) Simulation of the yield and accuracy of wind profile measurements from the Atmospheric Dynamics Mission (ADM-Aeolus). Q. J. R. Met. Soc. doi:10.1256/qj.04.02.

Telford, P. J., Braesicke, P., Morgenstern, O., and Pyle, J. A. (2008). Technical Note: Description and assessment of a nudged version of the new dynamics Unified Model. Atmospheric Chemistry and Physics, 8(6):1701–1712.

Tobin, D. C., Revercomb, H. E., Knuteson, R. O., Lesht, B. M., Strow, L. L., Hannon, S. E., Feltz, W. F., Moy, L. A., Fetzer, E. J., and Cress, T. S. (2006). Atmospheric Radiation Measurement site atmospheric state best estimates for Atmospheric Infrared Sounder temperature and water vapor retrieval validation. Journal of Geophysical Research, 111(D9):D09S14+.

Thompson, D. W. J., Baldwin, M. P., and Wallace, J. M. (2002). Stratospheric connection to Northern Hemisphere wintertime weather: Implications for prediction. Journal of Climate, 15(12):1421–1428.

Thompson, D. W. J. and Wallace, J. M. (2001). Regional climate impacts of the Northern Hemisphere annular mode. Science, 293(85):85–89.

Tripathi, O. P., Baldwin, M., Charlton-Perez, A., Charron, M., Eckermann, S. D., Gerber, E., Harrison, R. G., Jackson, D. R., Kim, B.-M., Kuroda, Y., Lang, A., Mahmood, S., Mizuta, R., Roff, G., Sigmond, M., and Son, S.-W. (2014). The predictability of the extratropical stratosphere on monthly time-scales and its impact on the skill of tropospheric forecasts. Quarterly Journal of the Royal Meteorological Society

Tripathi, O. P., Charlton-Perez, A. J., Sigmond, M. and Vitart, F. (2015) Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. Env. Res. Lett., doi: 10.1088/1748-9326/10/10/104007

Wang, J., Bian, J., Brown, W. O., Cole, H., Grubisic, V., and Young, K. (2009). Vertical Air Motion from T-REX Radiosonde and Dropsonde Data. J. Atmos. Oceanic Technol., 26(5):928–942.