University of Reading

# Stratosphere-Troposphere Coupling on Subseasonal Timescales

---

Simon H. Lee

*A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy*

**PhD Atmosphere, Oceans and Climate**

**Department of Meteorology**

**School of Mathematical, Physical and Computational Sciences**

**October 2021**

University of Reading

# Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Simon H. Lee

# Abstract

Variability in the Arctic stratospheric polar vortex can significantly influence the behaviour and predictability of wintertime tropospheric weather conditions on subseasonal timescales (lead-times between 2 weeks and 2 months). This thesis improves our understanding of tropospheric processes leading to subseasonal stratospheric variability and how to characterise the tropospheric impact of stratospheric variability in ways useful for subseasonal prediction.

First, the predictability onset of the February 2018 sudden stratospheric warming is linked to the occurrence of synoptic-scale anticyclonic wave breaking in the northeast Atlantic. This event is diagnosed through the Scandinavia-Greenland dipole pattern in mean sea-level pressure. Analysis of previous cases shows the Scandinavia-Greenland pattern is associated with significantly enhanced vertically propagating wave activity and a subsequently weakened polar vortex. The representation of the Scandinavia-Greenland pattern in hindcasts from the S2S Prediction Project database is then analysed. Substantial biases are found in the representation of the pattern in the troposphere and its relationship with a weakened polar vortex. These biases likely contribute to limiting subseasonal stratospheric forecast skill and highlight the need to improve the representation and predictability of the Scandinavia-Greenland pattern for improved polar vortex forecasts.

Next, the downward influence of the stratosphere on the troposphere over North America is characterised using four weather regimes, defined through clustering analysis. Significant observed differences in regime behaviour are found for three of the four regimes depending on the strength of the lower-stratospheric vortex. By considering the empirical orthogonal functions which define the regimes, a linear theory is developed to explain how changes in the stratospheric vortex strength can lead to regime transitions. The theory is supported by the results from a set of stratospheric relaxation model experiments. These results provide a framework to explain how uncertainty or improvements in the subseasonal stratospheric forecast can translate to changes in large-scale tropospheric forecasts over North America.

# Acknowledgments

Foremost, I would like to thank my supervisors – Andrew Charlton-Perez, Steve Woolnough, and Jason Furtado – for making this work possible. Our Tuesday afternoon meetings were more often than not the highlight of my working week; I will miss them! Always encouraging, always interested, and always willing to let me take the project off in my own direction, yet able to "constrain" me sufficiently to focus myself into one body of work – I have learned so much from you, and I hope we can collaborate more in the future. Thank you also to my Monitoring Committee members Maarten Ambaum and Joy Singarayer for helping me to think about my project in different ways and keeping me on track. I would also like to thank the wider atmospheric sciences community for engaging, friendly and enthusiastic discussions alongside several pieces of related work over the last few years – particularly, but not exhaustively, Amy Butler (NOAA CSL), Alexey Karpechko (FMI), Daniela Domeisen (ETH Zurich), and Zachary Lawrence (NOAA PSL).

I am grateful to the Natural Environment Research Council, via the SCENARIO Doctoral Training Partnership, for funding this project, and to all organisers and contributors to the S2S Prediction Project across the world. Glenn Carver and Marcus Koehler (ECMWF) deserve special thanks for their assistance with OpenIFS and preparing the initial conditions and relaxation data. A big thank you to everyone in the Department of Meteorology in Reading, and the School of Meteorology at the University of Oklahoma, for providing such a friendly and supportive environment in which to work (even during the pandemic).

To my family – thank you for always fostering my curiosity about the world throughout my life, for listening to me as I often went on at length about atmospheric science, and for always reminding me just how interesting this subject is.

And finally, thank you to the 2018 sudden stratospheric warming for providing a rich source of inspiration for much of this work.

# Authorship of Papers

This thesis is structured around four papers. While they have been reformatted for use here along with minor typographical adjustments, they are otherwise unmodified from the accepted manuscripts. The first three papers have been published; the fourth has been revised following review for *Journal of Climate*. All Supporting Information/Supplementary Material for each paper is also included in this thesis. The estimated percentage contribution of the candidate (SHL) is provided below.

**Lee, S. H.**, A. J. Charlton-Perez, J. C. Furtado, and S. J. Woolnough, 2019: Abrupt stratospheric vortex weakening associated with North Atlantic anticyclonic wave breaking. *Journal of Geophysical Research: Atmospheres*, **124**, 8563-8575, `https://doi.org/10.1029/2019JD030940`

*Estimated contribution: 80%. SHL developed the methodology, performed all analysis and wrote the manuscript with input and suggestions from AJCP, JCF and SJW. All authors commented on the manuscript and discussed the results at all stages. Three anonymous reviewers provided comments on one earlier version of the manuscript.*

**Lee, S. H.**, A. J. Charlton-Perez, J. C. Furtado, and S. J. Woolnough, 2020: Representation of the Scandinavia-Greenland Pattern and its Relationship with the Polar Vortex in S2S Models. *Quarterly Journal of the Royal Meteorological Society*, **146**, 4083-4098, `https://doi.org/10.1002/qj.3892`.

*Estimated contribution: 80%. SHL developed the methodology, performed all analysis and wrote the manuscript with input and suggestions from AJCP, JCF and SJW. All authors commented on the manuscript and discussed the results at all stages. Two anonymous reviewers provided comments on one earlier version of the manuscript.*

**Lee, S. H.**, J. C. Furtado, and A. J. Charlton-Perez, 2019: Wintertime North American Weather Regimes and the Arctic Stratospheric Polar Vortex. *Geophysical Research Letters*, **46**, 14892-14900, https://doi.org/10.1029/2019GL085592.

*Estimated contribution: 80%. SHL developed the methodology, performed all analysis and wrote the manuscript with input and suggestions from JCF and AJCP. All authors commented on the manuscript and discussed the results at all stages. Two anonymous reviewers provided comments on two earlier versions of the manuscript.*

**Lee, S. H.**, A. J. Charlton-Perez, S. J. Woolnough, and J. C. Furtado, 2021: How do stratospheric perturbations influence North American weather regime predictions? *Journal of Climate*, in review following revision.

*Estimated contribution: 85%. SHL developed the methodology, performed all analysis and wrote the manuscript with input and suggestions from AJCP, SJW and JCF. All model experiments were performed by SHL. All authors commented on the manuscript and discussed the results at all stages. Two anonymous reviewers provided comments on one earlier version of the manuscript.*

# Table of Contents

# Chapter 1

# Introduction

## 1.1 The Subseasonal Problem

### 1.1.1 Bridging the gap

The subseasonal timescale lies between the medium-range (lead-times typically $\leq 2$ weeks) and the seasonal range (lead-times typically $> 2$ months) and has been described as *"bridging the gap"* between the two (e.g. Vitart et al., 2012; Lang et al., 2020). In contrast to medium-range and seasonal forecasting, subseasonal prediction is still in its infancy – in terms of both understanding the fundamental predictability problem and the ability to produce skilful subseasonal forecasts. The moniker 'S2S' (subseasonal-to-seasonal) has been widely adopted after the ongoing World Weather Research Program (WWRP) and World Climate Research Program (WCRP) prediction project of the same name (Vitart et al., 2017)[1]. The S2S Prediction Project includes the S2S database[2], an archive of forecasts and hindcasts from 11 different international prediction centres.

Certain weather events, or sequences of individual weather events, have timescales of 1–2 weeks, characteristic of the subseasonal timescale and can be described as the 'S2S phenomena' that subseasonal forecasts aim to predict (Robertson et al., 2018). Examples include: coldwaves in winter (e.g. Ferranti et al., 2018b), heatwaves in summer (e.g. Wulff and Domeisen, 2019), atmospheric blocking (which can be associated with either

[1] `https://s2sprediction.net`

[2] Originally hosted at ECMWF as an extension of the TIGGE archive of medium-range forecasts: `https://apps.ecmwf.int/datasets/data/s2s`

of the former), the large-scale anomalous circulation associated with tornado outbreaks (Gensini et al., 2019) or flooding and wind impacts from periods of increased or suppressed tropical and extratropical cyclone activity (Vitart and Robertson, 2018). These all have the potential for large socio-economic impacts (in part owing to their longevity compared with individual or synoptic-scale weather systems), motivating the need for both accurate and useful S2S forecasts to give advanced warning.

The S2S timescale possesses some features of both medium-range and seasonal timescales. Medium-range forecasts usually have a time resolution[3] of at least daily and include both deterministic and probabilistic ensemble predictions, while seasonal forecasts are probabilistic predictions from an ensemble covering monthly or seasonally-averaged statistics. In order to be useful, subseasonal forecasts must be at a higher temporal resolution than seasonal forecasts (both the output data and forecast initialisation frequency), and thus range from weekly to monthly-averaged probabilities.

In a similar sense, medium-range prediction can be considered as an initial-value problem (where the source of predictability is the initial state of the atmosphere) and seasonal prediction as a boundary-value problem (where predictability arises from slowly-varying components like sea-surface temperatures or the cryosphere). Subseasonal prediction is effectively a hybrid of these, covering both the spatio-temporal aggregation of synoptic-scale weather events and the background state upon which they form. This is where the predictability challenge lies, because the S2S timescale is beyond the limit implied by an initial-value prediction of the atmosphere (e.g. Lorenz, 1969) but both too short in lead-time and too fine in temporal resolution for seasonal boundary conditions to have substantial impact. When accounting for the different forecast timescales or averaging periods, the average skill on the subseasonal timescale is currently much lower than that on the medium-range or seasonal timescales (Figure 1.1; White et al. (2017)).

[3]Here, 'time resolution' refers to the *usable* (potentially skilful) processed model output that is used in forecasting.

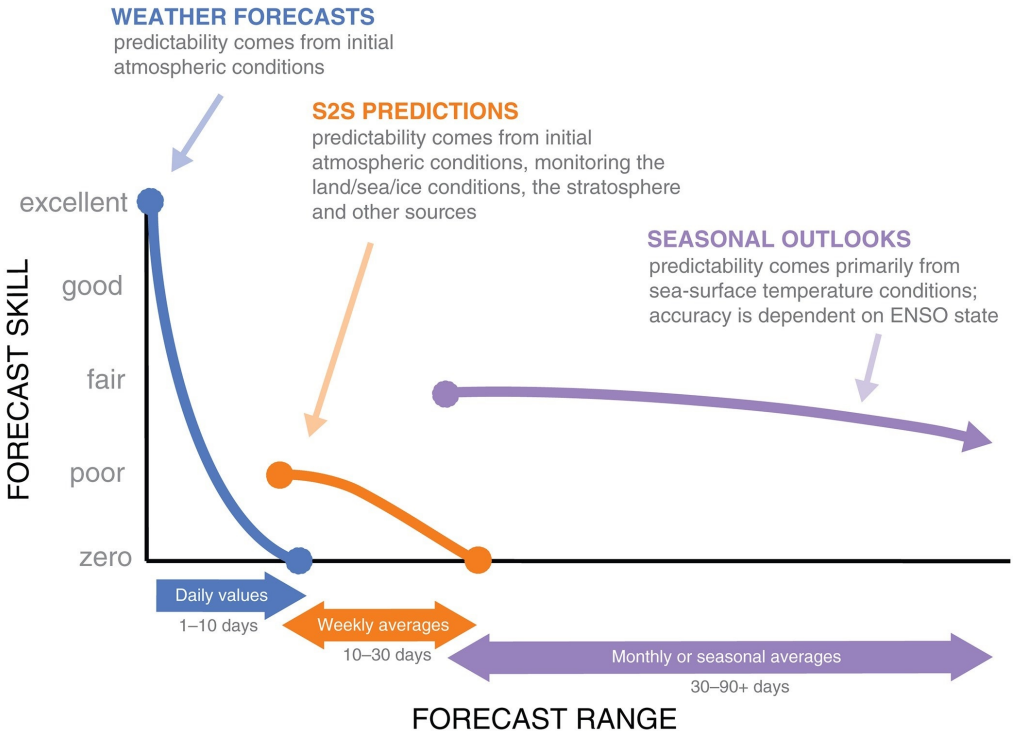


**Figure 1.1:** Schematic of relative forecast skill for the medium, subseasonal and seasonal scales alongside example sources of predictability. The relative skill depends on the averaging period (i.e., daily, weekly, and seasonal). Figure 1 (a) in White et al. (2017). Reprinted with permission from Wiley.

Furthermore, at subseasonal lead-times, model mean-state biases (owing to systematic model errors) can become substantial as models drift toward their own climatology. The model climate can differ markedly from the observed climate, leading to significant effects on the representation of key processes and overall forecast performance (e.g. Kim et al., 2019; Son et al., 2020; Furtado et al., 2021). A key objective of the current Phase II of the S2S Prediction Project is to better understand S2S model biases and their dynamical and predictive consequences using the set of hindcasts from the S2S database (WMO, 2018).

### 1.1.2 Sources of subseasonal prediction

Sources of S2S prediction skill (Figure 1.2) are components of the Earth system which vary on timescales of several weeks and therefore possess potential skill horizons beyond two weeks, providing usable information on the S2S scale. During Northern Hemisphere winter, two leading examples are the 30–60 day cycles of the Madden-Julian Oscillation (MJO; e.g. Madden and Julian (1972); Woolnough (2019)) and variability in the stratospheric

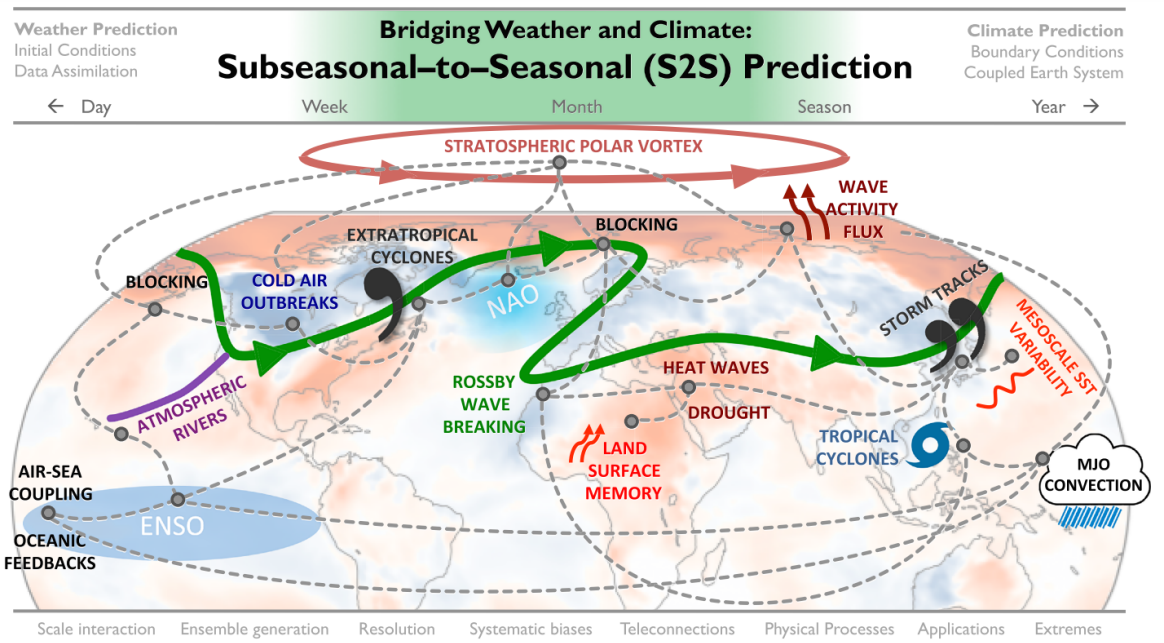


**Figure 1.2:** Schematic representation of S2S phenomena and sources of S2S predictability. Figure 1 in Lang et al. (2020). Reprinted with permission from Wiley.

polar vortex (SPV; e.g. Butler et al. (2019a); Domeisen and Butler (2020)). Other sources of S2S predictability include land-surface processes (Dirmeyer et al., 2019), sea ice (Chevallier et al., 2019) and atmosphere-ocean interaction (Saravanan and Chang, 2019), which can all act to modulate the atmospheric circulation on extended timescales.

The MJO – an eastward-propagating envelope of alternating enhanced and suppressed tropical convection – is the dominant mode of subseasonal variability within the tropics, but also contributes to subseasonal extratropical variability. For example, there is a lagged modulating influence of the MJO on North Atlantic-European weather regime occurrence (Cassou, 2008; Lee et al., 2019a). The convection associated with the MJO can act as a Rossby wave source, exciting tropospheric Rossby wave trains which can both directly and indirectly (via the stratosphere) influence tropospheric weather patterns (Barnes et al., 2019). When combined with the timescale of the MJO, the lagged relationship extends the potential regime forecast skill horizon into the S2S range. In addition, forecasts initiated during strong MJO events tend to exhibit extended skill in the North Atlantic Oscillation (NAO; the leading mode of tropospheric variability in the North Atlantic) (Lin et al., 2010; Feng et al., 2021). This planetary-scale influence has

motivated work to understand and improve the prediction of the MJO and the fidelity of the MJO-extratropical teleconnection in S2S forecast models (e.g. Vitart, 2017; Lim et al., 2018) to improve S2S forecast skill. However, when the MJO is inactive, the source of S2S predictability is lost, which can fundamentally limit the available forecast skill.

This intermittence highlights a further S2S prediction challenge: at times, when sources of predictability are active, S2S forecasts have the *potential* for much greater accuracy, known as 'windows of opportunity' (Mariotti et al., 2020). Subseasonal windows of opportunity can arise when parts of the Earth system are in a long-lasting anomalous state, associated with driving subsequent anomalous (and thus often high-impact) weather patterns (in some ways analogous to the seasonal-scale effect of the El Niño-Southern Oscillation). A recent example of a window of opportunity is the persistent strongly positive NAO and closely-related Arctic Oscillation (AO) during winter 2019/20 (Lawrence et al., 2020), which was associated with extreme weather conditions across the Northern Hemisphere (e.g. Davies et al., 2021). This was well-forecast by subseasonal and seasonal prediction systems (Hardiman et al., 2020; Lee et al., 2020b; Rao and Garfinkel, 2021), which Hardiman et al. (2020) related to anomalous Rossby wave forcing from an extremely positive Indian Ocean Dipole event.

Identifying such windows of opportunity in advance, or knowing when and why the subseasonal range is more or less predictable, is thus a major hurdle for successful S2S prediction with large potential benefits. Like the MJO, anomalously weak or strong Arctic SPV states are not always present, and can exert a remote influence on the behaviour and predictability of the large-scale extratropical tropospheric circulation. The subseasonal prediction and tropospheric impact of variability in the strength of the SPV are the focus of this thesis.

## 1.2 Subseasonal Prediction and the Stratosphere

### 1.2.1 Why the stratosphere?

During winter, the timescale of variability in the Arctic stratosphere diagnosed through the Northern Annular Mode (NAM; the leading mode of extratropical variability associated with mass fluctuations between the middle and high-latitudes) is 3–5 weeks (Figure 1.3) (Baldwin et al., 2003; Simpson et al., 2011). Stratospheric NAM variability primarily arises due to the upward propagation and subsequent breaking in the stratosphere of long-wavelength Rossby (planetary) waves (Matsuno, 1970, 1971; McIntyre and Palmer, 1983); zonal wavenumbers 1–3 dominate due to the zonal wind and stability profile (Charney and Drazin, 1961).

The Arctic vortex is more variable than its Antarctic counterpart owing to higher-amplitude stationary waves in the Northern Hemisphere, which are in turn driven by the zonal asymmetries arising from the distribution of continents, orography, and land-sea temperature contrasts (e.g. Garfinkel et al., 2020). Breaking Rossby waves exert a westward (i.e., easterly) drag on the westerlies into which they propagate, which can distort and decelerate the westerly flow of the SPV. The meridional circulation induced by the wave breaking (a transient enhancement of the mean Brewer-Dobson circulation; reviewed by Butchart (2014)) leads to polar stratospheric warming and tropical stratospheric cooling, which, by thermal wind, balances the westward wind tendency.

Stratospheric NAM variability is confined to the winter hemisphere, where mean stratospheric winds are westerly, as Rossby waves cannot propagate into easterlies. Extremely strong or weak SPV states, including sudden stratospheric warmings (SSWs)[4], can persist for over two months in the lower stratosphere, stretching across the S2S timescale (Figure 1.4; Baldwin and Dunkerton (2001)). Moreover, the state of the stratospheric

[4]SSWs are sometimes called *stratospheric sudden warmings* or simply *sudden warmings*. The term "major SSW" is used to specifically refer to a subset of extreme mid-winter events (e.g. Charlton and Polvani, 2007; Butler et al., 2015).

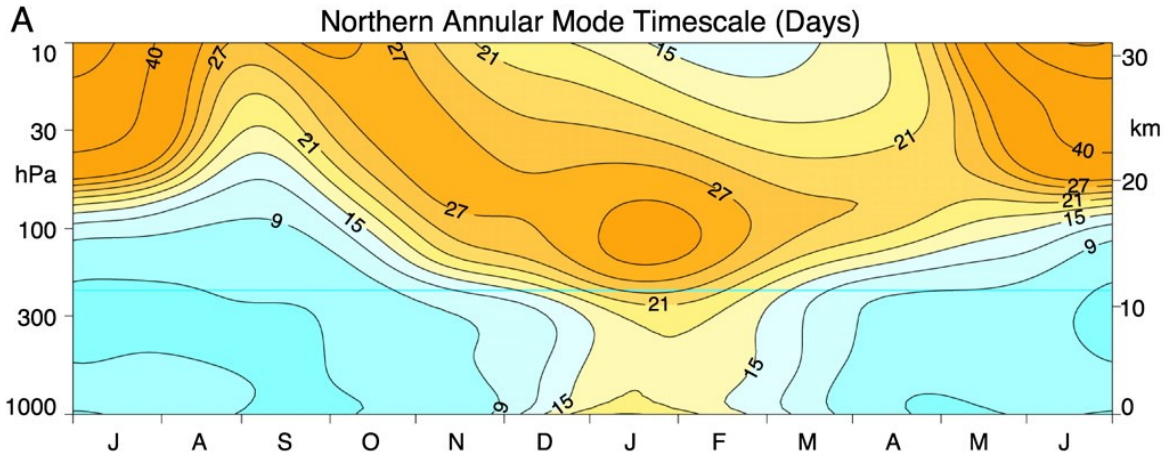


**Figure 1.3:** Time-height cross section of the *e*-folding timescale (in days) of the NAM time series. The maximum in the wintertime lower stratosphere is around 30 days. Figure 1 (a) in Baldwin et al. (2003). Reprinted with permission from AAAS.

NAM is on average most closely related to the *integrated* wave activity on subseasonal scales (1–2 months), rather than instantaneous, synoptic-scale wave pulses (Polvani and Waugh, 2004). Thus, the stratospheric state can be viewed as a subseasonal integrator of higher-frequency (and less predictable) tropospheric variability.

When combined with the relatively recent ability to fully represent the stratosphere due to increased computational and observational resources, there has been a significant increase in attention devoted to the model representation, predictability, and impacts of SPV variability on S2S timescales (Domeisen et al., 2020b,c). This has extended earlier work which demonstrated the impact of SPV variability on seasonal scale prediction (e.g. Sigmond et al., 2013; Scaife et al., 2016; Jia et al., 2017) and model experiments confirming the downward impact of stratospheric variability on the troposphere (Gerber et al., 2009; Hitchcock and Simpson, 2014). Extended-range model performance has also been shown to be higher when forecasts are initialised with, or correctly predict, an extreme SPV state (Tripathi et al., 2015a,b; Butler et al., 2019b; Domeisen et al., 2020c), which further supports considering the SPV as a potential source of subseasonal predictive skill.

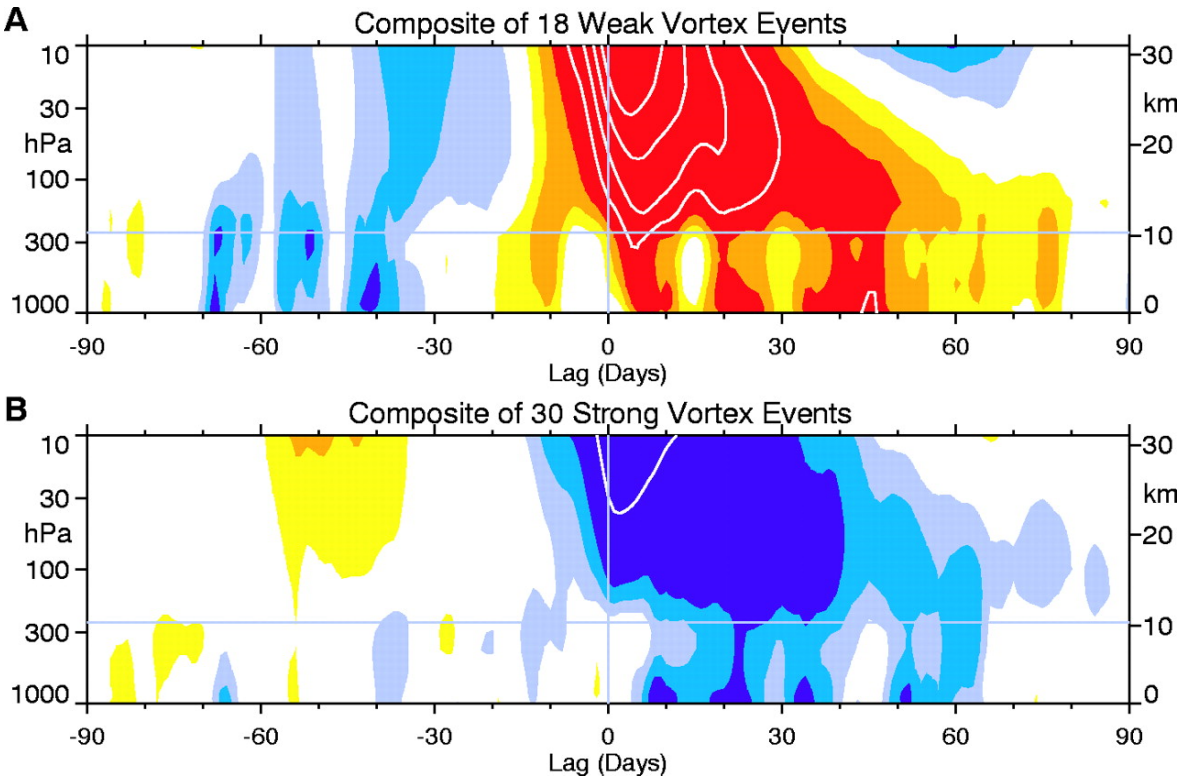


**Figure 1.4:** Time-height cross sections of the NAM for 90 days before and after the onset of (a) 18 weak and (b) 30 strong vortex events during 1958-1999. Colour contours every 0.25, white contours every 0.5, with values between −0.25 and 0.25 unshaded. Figure 2 in Baldwin and Dunkerton (2001). Reprinted with permission from AAAS.

### 1.2.2 The tropospheric response to the stratosphere

The downward influence of variability in the Arctic stratosphere is most commonly viewed through lagged changes to the tropospheric NAM. On average, the sign of the tropospheric NAM is shifted toward that of the stratospheric NAM in the weeks-to-months following strong (positive NAM) or weak (negative NAM) SPV states, characterised by latitudinal shifts in the mid-latitude storm tracks (Baldwin and Dunkerton, 2001; Kidston et al., 2015) and associated anomalous weather (e.g. Kolstad et al., 2010; Lawrence et al., 2020). Additionally, there is usually a lag of 1–3 weeks between the onset of the stratospheric event and the maximum surface response (Hall et al., 2020).

Several frameworks can be used to explain the downward propagation of stratospheric circulation anomalies into the troposphere (see reviews by Kidston et al. (2015) and Baldwin et al. (2021)). In the stratosphere following an SSW, the development of

a critical layer opaque to Rossby waves (i.e., where zonal winds are easterly) forces upward propagating Rossby waves to break at successively lower altitudes, progressively decelerating the zonal winds through the depth of the stratosphere and bringing the anomalous easterly circulation lower (Matsuno, 1971). The remote effects of stratospheric wave driving have been shown to yield so-called 'downward control' (Haynes et al., 1991): the meridional circulation induced by the stratospheric wave drag must close downwards due to the requirement of a frictional sink – the planetary boundary layer – to balance the anomalous wave momentum deposition.

The anomalous heating (or cooling) in the stratosphere can be viewed as a potential vorticity (PV) anomaly, and hence the balanced geostrophic and hydrostatic response can be inferred from PV inversion (Hartley et al., 1998; Ambaum and Hoskins, 2002; Black, 2002). In isolation this effect would decrease with distance from the PV anomaly, but the maximum mass (pressure) anomalies are seen near the surface (Baldwin et al., 2021). Hence, it is generally recognised that in order to explain the full magnitude and duration of the surface response, tropospheric weather systems – in the form of synoptic and planetary-scale eddies – must feed back onto the initial large-scale balanced response (Kushner and Polvani, 2004; Charlton et al., 2005; Hitchcock and Simpson, 2016; White et al., 2020). A contribution to this effect has been shown to arise from the effect of changes to lower-stratospheric shear associated with the strength of the SPV on baroclinic growth rates (Wittman et al., 2007; Smy and Scott, 2009).

Changes to planetary wave propagation in the stratosphere can also directly influence surface weather and climate. When certain conditions are met (usually requiring curvature in the vertical zonal wind profile), upward-propagating Rossby waves in the stratosphere can be reflected downward, influencing the tropospheric wave field (Perlwitz and Harnik, 2003, 2004; Kodera et al., 2013, 2016). Wave reflection often accompanies stratospheric vortex intensification (e.g. Lawrence et al., 2020) since it is associated with divergence of wave activity in the stratosphere.

However, while the *average* tropospheric response is robust, there is large intra- and inter-event variability in the tropospheric evolution following extreme stratospheric events (especially SSWs), with not all events seeing a persistent and coherent downward coupling of the NAM (see, for example, the intermittent 'drips' from the stratosphere to the troposphere in Figure 1.4). A prominent example is the contrasting tropospheric responses to SSWs in February 2018 and January 2019 (Butler et al., 2020; Rao et al., 2020). While the former was followed by a strongly negative tropospheric NAM and extreme cold in northwest Europe (the so-called 'Beast from the East'; Greening and Hodgson (2019)), the latter was not. Approximately two-thirds of SSWs are followed by the canonical negative tropospheric NAM, a ratio which is found in both observations and climate model simulations (Karpechko et al., 2017; White et al., 2019; Afargan-Gerstman and Domeisen, 2020). The specific reasons for these differences has yet to be fully explained.

Some studies have attributed such tropospheric variation to the morphology of the SPV (e.g. vortex split versus displacement events; Charlton and Polvani (2007); Matthewman et al. (2009); Mitchell et al. (2013); Choi et al. (2019)), the wave dynamics involved in the SSW (Barriopedro and Calvo, 2014; Kodera et al., 2016), or the magnitude of the lower-stratospheric anomaly associated with the event (Hitchcock et al., 2013; Maycock and Hitchcock, 2015). These distinctions are not necessarily exclusive; vortex splits usually have a stronger lower-stratospheric signature due to their near-barotropic nature (Matthewman et al., 2009). Other studies have viewed the varied response of the troposphere as arising due to differences in the troposphere itself, rather than differences in the stratospheric perturbation (Charlton-Perez et al., 2018; Domeisen et al., 2020a; Afargan-Gerstman and Domeisen, 2020; Knight et al., 2021). These may also not be independent; for example, there are differing tropospheric precursor patterns to downward and non-downward propagating SSWs (White et al., 2019) which may themselves induce different SPV states.

Furthermore, for the purposes of subseasonal prediction, planetary-scale modes of variability (like the NAM) are of limited use as they only explain a fraction of the total variability (e.g. 20-30% for the surface NAM/AO; Thompson and Wallace (1998)). More regional flow anomalies and their impacts (e.g., a specific blocking anticyclone over northwest Europe) are not well-represented by such large-scale modes, and there can be a diverse range of weather within a particular phase of the NAM. To address this problem, regional weather regimes (often defined using clustering methods) have become increasingly common as a subseasonal analysis and forecasting framework that lies somewhere between the synoptic and planetary scale (e.g. Cassou, 2008; Grams et al., 2017, 2020). Regimes provide a spatially and temporally filtered version of the full field (suitable for extended-range prediction), but also usually higher dimensionality than, for example, simply the two phases of the NAM. Some recent studies (Beerli et al., 2017; Charlton-Perez et al., 2018; Maycock et al., 2020; Goss et al., 2021) have proposed that the tropospheric response to stratospheric perturbations in the Euro-Atlantic sector is a manifestation of changes to the probabilities of regime persistence and transition. The use of regimes may therefore provide a framework to characterise the downward impact of stratospheric variability in a way that is well-suited to subseasonal prediction, including its applications (e.g. Huang et al., 2020; Bloomfield et al., 2021).

### 1.2.3 Predictability of the stratosphere

While the long persistence timescale of stratospheric anomalies is useful for S2S prediction and extends the already-long average stratospheric prediction skill (Son et al., 2020), the strength of the SPV can change abruptly. The most striking example is during a major SSW, which can involve a change of state from a climatologically strong to an extremely weak vortex within a few days. As an example, a vertical cross-section of the evolution of the February 2018 SSW is shown in Figure 1.5. The interaction between the synoptic and subseasonal timescales at the onset of extreme stratospheric events presents another challenge for S2S prediction. Once developed, the extreme SPV state can persist across the S2S timescale (and is therefore more predictable), but the onset is on a much smaller

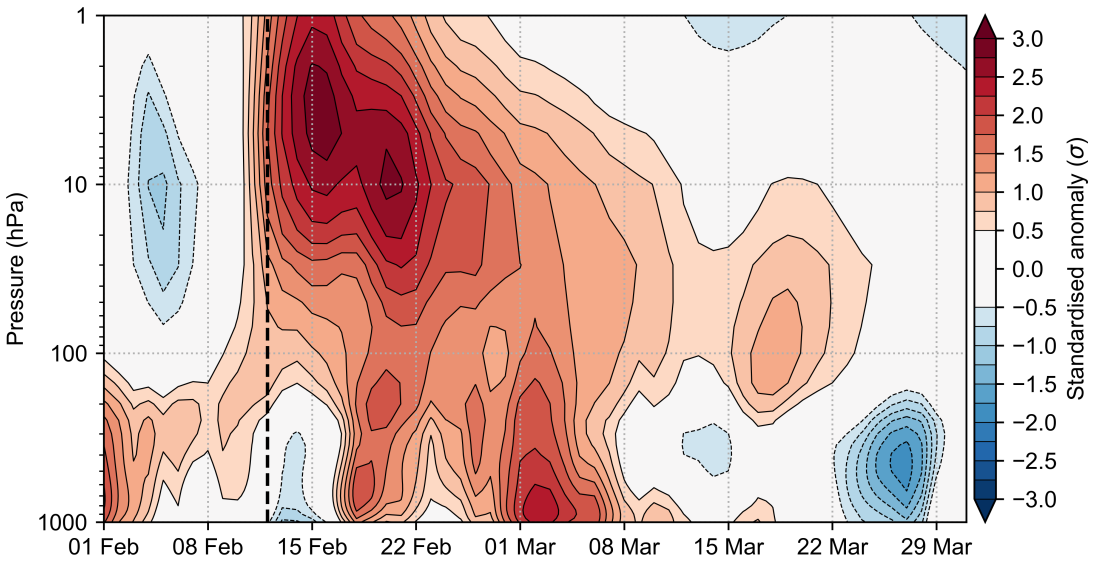


**Figure 1.5:** Evolution of 60-90°N-average standardised geopotential height anomalies (proportional to the NAM) from 1 February–31 March 2018 according to ERA5 reanalysis (Hersbach et al., 2020). The onset of the major SSW on 12 February (Karpechko et al., 2018) is shown with a vertical black dashed line. Anomalies are standardised with respect to the daily mean and standard deviation over 1979–2020.

(and less predictable) timescale. It follows that missing the less predictable onset could lead to an erroneous tropospheric forecast in the weeks thereafter, as has been shown for the 2018 SSW (Karpechko et al., 2018; Kautz et al., 2020).

Current forecast models are generally unable to accurately predict the onset of SSWs or strong vortex events more than 2 weeks in advance (Tripathi et al., 2015b; Karpechko, 2018; Taguchi, 2018; Domeisen et al., 2020b). However, seasonal-scale probabilistic skill (i.e., whether a strong vortex or SSW is more or less likely than normal to occur during the season) has been demonstrated (Scaife et al., 2016). The predictability of individual events is also highly variable (Karpechko, 2018; Butler et al., 2020). A component limiting stratospheric predictability is the predictability of the troposphere itself as a source of vertically-propagating wave activity. The effect is compounded by the response of the stratosphere to a vertically-propagating wave packet depending on the state of the stratosphere (Scott and Polvani, 2004; Birner and Albers, 2017; de la Cámara et al., 2019; Lawrence and Manney, 2020), which itself depends on the preceding wave activity (Polvani and Waugh, 2004). Together, this may partly explain why the current pre-

dictability limit of extreme stratospheric events is not much beyond that of the troposphere.

Therefore, an increased understanding of which tropospheric processes contribute to vertically-propagating wave activity, and thus stratospheric variability, will be an important component in exploiting the stratosphere for S2S forecasts. Better model representation and prediction of these may then help extend stratospheric forecast skill, with the potential for subsequent increases in tropospheric S2S skill.

## 1.3 Aims of the Thesis

As motivated in the previous section, two key challenges in subseasonal prediction arising from the stratosphere are: (i) the predictability of processes driving stratospheric variability on medium-to-subseasonal timescales, and (ii) understanding how the stratosphere can change tropospheric weather patterns in a way that is useful for subseasonal forecasting. This thesis therefore addresses the following three key questions (KQs):

1. What tropospheric processes influence stratospheric variability on subseasonal timescales?

2. How well do current S2S forecast models represent these tropospheric processes, and is there evidence that model biases limit stratospheric skill?

3. How do circulation anomalies in the stratosphere influence subseasonal tropospheric weather patterns?

KQs 1 and 2 deal with the problem of predicting stratospheric anomalies on S2S timescales, while KQ3 deals with problem of utilising the stratosphere for S2S prediction of the troposphere. In answering KQ3, this thesis focuses on understanding the relationship between changes to the strength of the SPV and the large-scale weather over North America. This has so far received less attention in the literature than the equivalent for the Euro-Atlantic sector, despite increasing evidence of a significant stratospheric influence on North American weather with potentially large societal impacts (e.g. Kodera et al., 2016; Kretschmer

et al., 2018a; Cohen et al., 2021).

## 1.4 Thesis Structure

The main body of the thesis consists of four individual papers, presented here as (reformatted) original manuscripts. Chapters 2 and 3 focus on the development and prediction of anomalous SPV states on subseasonal timescales (KQ1 and 2), while Chapters 4 and 5 consider when and how the stratospheric state influences the troposphere (KQ3). Further specific literature is reviewed in each of these chapters.

The first paper, Chapter 2 (*"Abrupt stratospheric vortex weakening associated with North Atlantic anticyclonic wave breaking"*, first published in *Journal of Geophysical Research: Atmospheres* in July 2019), analyses the relatively short lead-time of predictions of the February 2018 major SSW and demonstrates the importance of a synoptic-scale tropospheric event – anticyclonic wave breaking in the northeast Atlantic – in driving both this event and prior cases of SPV weakening (KQ1). Chapter 3 (*"Representation of the Scandinavia-Greenland pattern and its relationship with the polar vortex in S2S forecast models"*, first published in *Quarterly Journal of the Royal Meteorological Society* in August 2020) builds on these results by considering this tropospheric variability more generally and assesses its representation in hindcasts from 10 S2S models. The chapter focuses on quantifying model biases in the representation of this variability in the troposphere and the stratosphere which may contribute to limiting stratospheric S2S forecast skill (KQ2).

Chapter 4 (*"Wintertime North American weather regimes and the Arctic stratospheric polar vortex"*, first published in *Geophysical Research Letters* in December 2019) defines four large-scale tropospheric weather regimes over North America, and analyses their statistical relationship with the strength of the SPV and potentially high-impact cold air outbreaks using reanalysis data (KQ3). Chapter 5 (*"How do stratospheric perturbations influence North American weather regime predictions?"*, revised following review for *Journal of Climate*, October 2021) extends the results of Chapter 4 by building an understanding

of how changes in the stratosphere couple to changes in weather regimes through the underlying modes of variability, and then uses model experiments to understand the effect of improved stratospheric forecasts on regime predictions (KQ3).

A summary and unification of the main findings of the four research papers follows in Chapter 6, including a discussion of subsequent related research since their publication. Suggestions for future work arising from this thesis are also presented.

# Chapter 2

# Abrupt Stratospheric Vortex Weakening Associated with North Atlantic Anticyclonic Wave Breaking

This chapter has been published in *Journal of Geophysical Research: Atmospheres* with the following reference:

**Lee, S. H.**, A. J. Charlton-Perez, J. C. Furtado, and S. J. Woolnough, 2019: Abrupt stratospheric vortex weakening associated with North Atlantic anticyclonic wave breaking. *Journal of Geophysical Research: Atmospheres*, **124**, 8563-8575, https://doi.org/10.1029/2019JD030940

## Abstract

The sudden stratospheric warming (SSW) of 12 February 2018 was not forecast by any extended-range model beyond 12 days. From early February, all forecast models that comprise the subseasonal-to-seasonal (S2S) database abruptly transitioned from indicating a strong stratospheric polar vortex (SPV) to a high likelihood of a major SSW. We demonstrate that this forecast evolution was associated with the track and intensity of a cyclone in the northeast Atlantic, with an associated anticyclonic Rossby wave break, which was not well forecast. The wave break played a pivotal role in building the Ural high, which existing literature has shown was a precursor of the 2018 SSW. The track of the cyclone built an anomalously strong sea level pressure dipole between Scandinavia and Greenland (termed the S-G dipole), which we use as a diagnostic of the wave break. Forecasts that did not capture the magnitude of this event had the largest errors in the SPV strength and did not show enhanced vertical wave activity. A composite of 49 similarly strong wintertime (November–March) S-G dipoles in reanalysis shows associated anticyclonic wave breaking leading to significantly enhanced vertical wave activity and a weakened SPV in the following days, which occurred in 35% of the 15-day periods preceding observed major SSWs. Our results indicate a particular transient trigger for weakening the SPV, complementing existing results on the importance of tropospheric blocking for disruptions to the Northern Hemisphere extratropical stratospheric circulation.


## 2.1 Introduction

The major mid-winter sudden stratospheric warming (SSW) event of 12 February 2018 was the first major SSW since January 2013 (defined as a reversal of the daily-mean 10 hPa 60°N zonal-mean zonal winds; Charlton and Polvani (2007)), a 5 year gap which was the longest since 1989–1998 according to the SSW Compendium (Butler et al., 2017). It produced a split of the stratospheric polar vortex (SPV) into two smaller vortices. Following the metric of Karpechko et al. (2017) the event was downward-propagating with the negative phase of the stratospheric Northern Annular Mode (NAM) (Thompson and

Wallace, 2000; Baldwin and Dunkerton, 2001) accompanied by a strong and persistent negative tropospheric NAM in the 45 days following the event. The negative tropospheric NAM and associated negative North Atlantic Oscillation (NAO) produced extremely cold conditions across Europe and northern Asia, with a large anticyclone over Scandinavia generating a cold easterly flow (Ferranti et al., 2018a). With such a high impact response, the ability to predict the onset of an SSW like in 2018 is of vital importance for sub-seasonal forecasting.

The 2018 SSW was the first to occur following the development of the subseasonal-to-seasonal (S2S) database of extended-range forecasts from 11 international forecast models (Vitart et al., 2017). None of the S2S model forecasts issued at the time indicated a major SSW until early February (Karpechko et al., 2018), giving a predictability window less than the medium-range timeframe ($\sim$2 weeks). Although this lies within the window typical of predicting major SSWs (Taguchi, 2014; Tripathi et al., 2015a, 2016), S2S model forecasts abruptly transitioned from projecting a strong SPV to a weak SPV/major SSW in late January-early February, with a corresponding transition in forecasts of tropospheric conditions (such as from forecasts of a positive NAO to a negative NAO). Karpechko (2018) showed several SSWs were poorly forecast in ECMWF hindcasts at lead-times beyond 7–10 days, but most were generally associated with a longer-range signal of SSW likelihood.

Specifically for the February 2018 event, Karpechko et al. (2018) examined S2S model forecasts from 1 February onwards and showed a strong relationship between the accuracy of stratospheric wind forecasts and the intensity of an anticyclone over the Urals (named the 'Ural high'). However, they did not assess the longer-term predictability of the event, the mechanism driving the onset of the Ural high, or its influence on the stratosphere, leaving open questions about the abrupt predictability onset. The Ural high has also been shown to drive SPV variability (Peings, 2019; White et al., 2019) by projecting onto the climatological stationary wave pattern.

Most studies of SSW precursors use a 'top-down' perspective, where the tropospheric features are analyzed in the period preceding observed stratospheric events. These approaches typically discern stationary or longer-lived features through the process of averaging anomalies in the build-up to SSWs. Tropospheric blocking is one such feature (e.g. Colucci and Kelleher, 2015; Garfinkel et al., 2010; Julian and Labitzke, 1965; Martius et al., 2009; Quiroz, 1986). For example, Bao et al. (2017) used cluster analysis to assess 500 hPa geopotential height patterns in the month before 37 SSWs in reanalysis, and found the patterns to be associated with linear interference with climatological stationary waves. Kolstad and Charlton-Perez (2011) used reanalysis alongside climate model simulations and found a particularly strong signal for a height anomaly dipole over northern Eurasia preceding 'weak vortex months'.

Other studies have considered more transient features associated with specific stratospheric events. Coy et al. (2009) noted the importance of zonal wavenumbers 4-5 associated with synoptic-scale systems preceding the SSWs of January 2006 and 2003. They implicated tropospheric systems over the North Atlantic and subtropical wave breaking; forecasting experiments showed a realistic SSW only occurred when a North Atlantic weather system was correctly represented in the model. On the other hand, a study of the January 2013 SSW (Coy and Pawson, 2015) suggested a rapidly-deepening cyclone in the North Atlantic played only a minor role, acting as a transient source of vertical wave activity that was not crucial to forcing the event. The authors also remark on the dynamical link between the initial stratospheric vortex state and the track of the cyclone, suggesting a two-way relationship. O'Neill et al. (2017) demonstrated a link between extratropical tropospheric cyclogenesis occurring at the edge of the SPV and split-type SSWs through a potential vorticity framework. Although mainly focusing on the Southern Hemisphere SSW of 2002 (e.g. Krüger et al., 2005), they briefly show a similar mechanism with cyclogenesis over the eastern seaboards of the Northern Hemisphere continents. Most recently, Attard and Lang (2019) approach the problem by looking at the meridional eddy heat flux, and demonstrate the different responses for blocks and 'bomb' cyclones in the Atlantic and Pacific sectors.

They conclude cold-season Atlantic bomb cyclones and Pacific blocks were associated with negative heat flux anomalies (and vice versa) whilst also noting that only a relatively small number of blocks and bombs are actively associated with SSWs.

Thus, there exist both transient and stationary drivers of stratospheric variability (including but not limited to SSWs), the predictability of which plays a role in the onset of SSW prediction. In this study, we provide a dynamical explanation for the abrupt transition in the forecasts of the February 2018 event, building upon existing analysis. We demonstrate that this is a characteristic of historical cases of vortex weakening, rather than unique to the flow configuration driving the 2018 event, through a 'bottom-up' approach (analysing the response of the stratosphere to tropospheric events). Our results have implications for extended-range predictability of SSWs and thus sub-seasonal tropospheric forecasts.

## 2.2 Data and Methods

We use forecast data from the European Centre for Medium Range Weather Forecasts (ECMWF) and National Centers for Environmental Prediction (NCEP) models, as these provide a combination of both large ensemble sizes and frequent launch dates - ECMWF launches twice weekly (Tuesday and Thursday) with 51 members, and NCEP launches daily with 16 members. The predictability onset of the SSW was common across all the S2S models (Karpechko et al., 2018), so our analysis is not sensitive to the choice of model. For verification, we use the ECMWF ERA-Interim reanalysis (Dee et al., 2011). The strength of the SPV is defined using the zonal-mean zonal wind at 10 hPa and 60°N ($U10_{60}$). We use 45–75°N meridionally-averaged zonal-mean eddy heat flux (denoted as [v*T*] where the star notation indicates a departure from the zonal-mean, and square brackets indicate a zonally-averaged quantity) at 300 hPa as a proxy for upper-tropospheric wave activity. This is proportional to the vertical component of the Eliassen-Palm flux (Andrews et al., 1987). Standardized polar cap (60–90°N) geopotential height anomalies are used as a proxy for the NAM index (Karpechko et al., 2017); the anomalies are inversely proportional to the index. Unless otherwise stated, standardized anomalies are computed with respect to the

climatological daily-mean and standard deviation in ERA-Interim. Historical composites use data from January 1979–March 2017 inclusive, and statistical significance is assessed using a bootstrap re-sampling method with replacement ($n = 50{,}000$) for November–March in the period January 1979 to March 2017. Potential vorticity is analyzed on the 315 K isentropic surface in ERA-Interim data and the 320 K isentropic surface in model forecast data as these are the nearest tropospheric levels available in both datasets. All data are re-gridded to 2.5° horizontal resolution for consistency.

## 2.3 Results

### 2.3.1 Characterizing the Onset of SSW Predictability

To demonstrate the evolution of forecasts of the zonal-mean state, Figure 2.1 (a, c) shows forecasts of $U10_{60}$ for the first 5 days of the verifying SSW (12–16 February) for all forecasts in which those dates featured. There is an abrupt transition in late January-early February from forecasts of a strong vortex to a weakened vortex or major SSW. In both ECMWF and NCEP systems, the 29 January ensembles showed no members indicating mean easterlies during this period, with a tightly clustered ensemble. The following day, forecasts from NCEP substantially changed, with some members suggesting a mean zonal wind reversal and the entire ensemble forecasting weaker $U10_{60}$ than the 25th percentile of the ensemble from the previous day – a change which also occurred in the 29 January and 1 February ECMWF ensembles. There is also an increase in spread despite the reduced lead-time. Ensemble spread was then much reduced by 5-6 February, a lead-time of only 6-7 days before the major SSW. A similar predictability evolution is found in other S2S models (not shown), indicating this was not related to the ability of certain models to capture the event. Moreover, we see the abrupt transition of vortex strength was associated with an abrupt increase in 300 hPa [v*T*] preceding the wind reversal (Figure 2.1b and 2.1d). The increase in forecast heat flux suggests the low predictability of the SSW was dependent on poorly-forecast tropospheric wave-driving, rather than the response of the stratospheric vortex to a wave pulse or the sensitivity of the $U10_{60}$ metric.

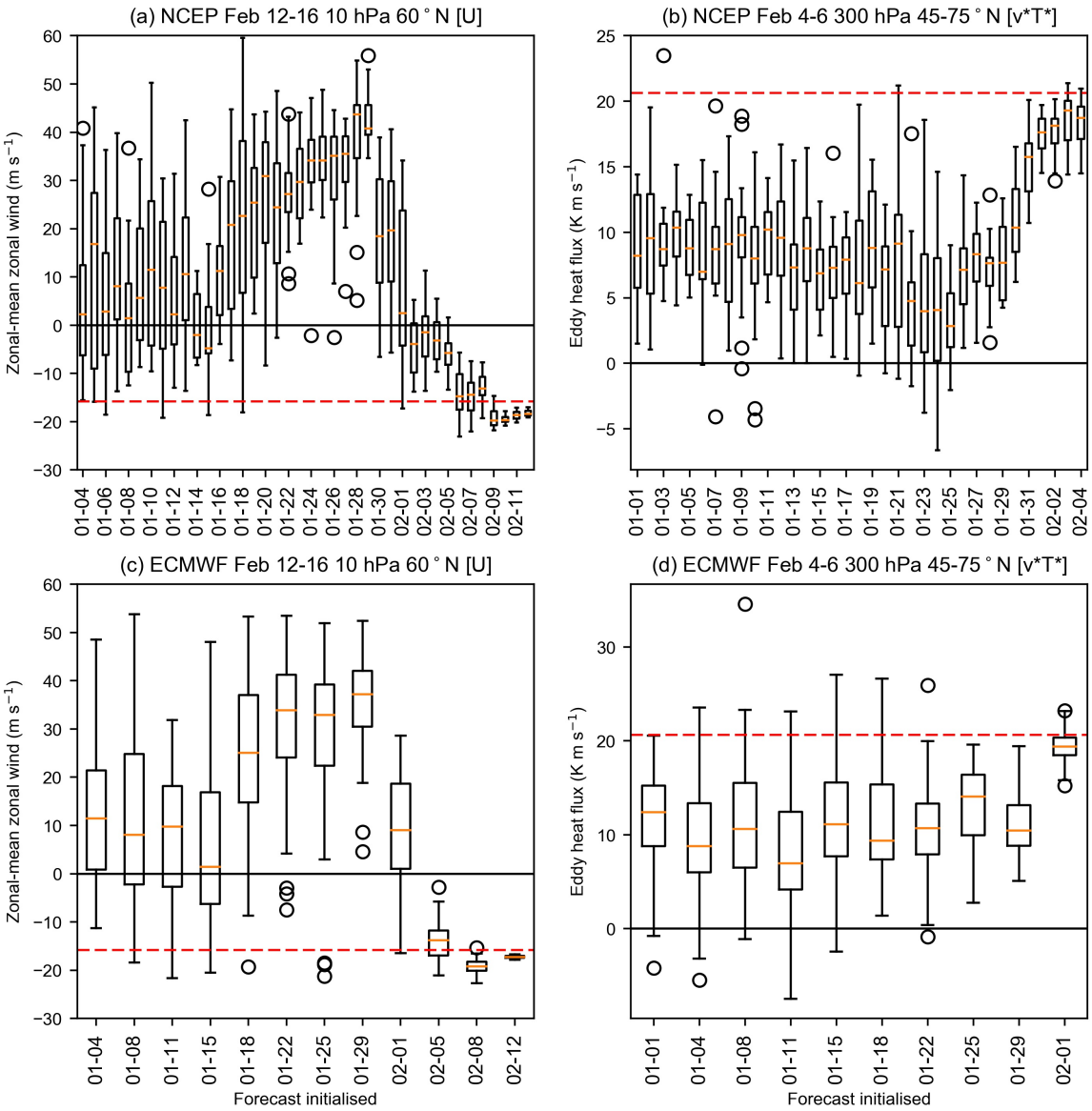


**Figure 2.1:** Boxplots showing (a), (c) average 10 hPa 60°N zonal-mean zonal-winds for 12–16 February 2018 and (b), (d) 300 hPa 45–75°N meridional eddy heat flux averaged over 4–6 February 2018 in (a), (b) NCEP and (c), (d) ECMWF models for all ensemble members as a function of initialisation date. Boxes indicate the interquartile range (IQR), whiskers extend to the last point less or greater than 1.5 times the IQR, with circles indicating outliers. The dashed red lines indicate verifying values according to ERA-Interim reanalysis.

To investigate this evolution further, we look at the 29 January and 1 February ensembles from ECMWF, which cover the spread of evolutions from strong vortex to weak vortex (Figure 2.2). The ensembles systematically diverge after 5 February – with the forecasts from 29 January showing low wave activity and strengthening zonal-mean zonal winds, whilst the opposite is true for forecasts from 1 February. This is an even greater divergence

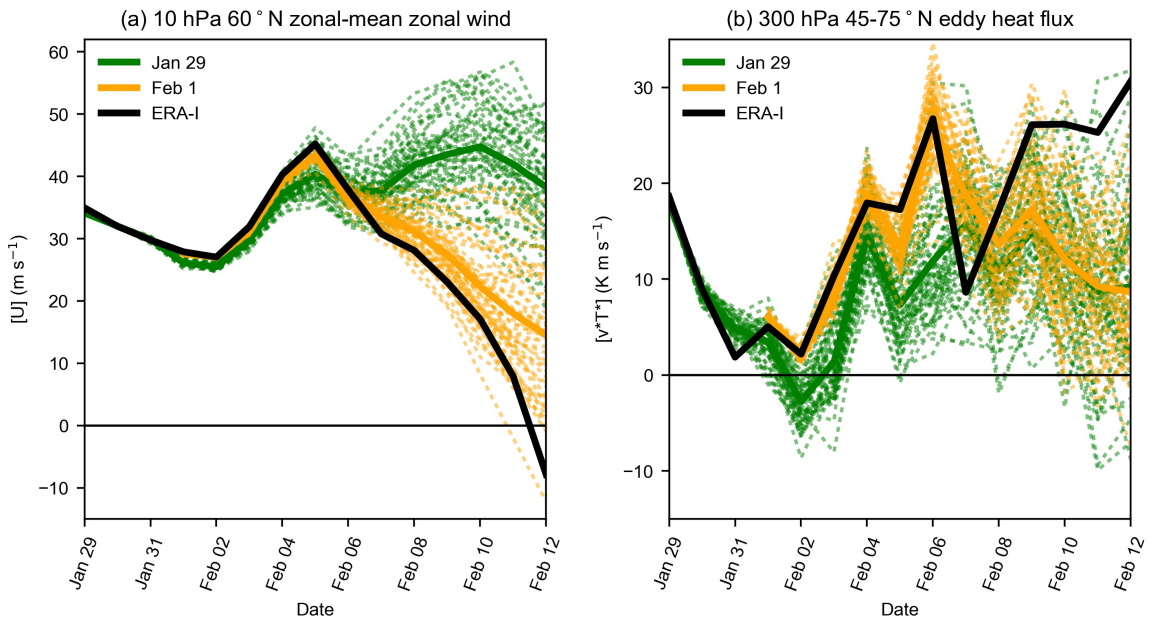


**Figure 2.2:** ECMWF ensemble forecasts from 29 January (dashed green) and 1 February (dashed orange) for (a) 10 hPa 60°N zonal-mean zonal wind and (b) 300 hPa 45–75°N meridional eddy heat flux for 29 January–12 February 2018. Ensemble means are shown with thick lines. Verifying evolution from ERA-Interim is shown with the thick black line.

in zonal wind intensity than day-15 forecasts for the January 2013 SSW shown in Tripathi et al. (2016). Despite the 3 day difference in lead-time, the systematic difference between the two ensembles motivates considering them together to capture the uncertainty. Analysis in the corresponding NCEP ensembles gives similar results (see Figure 2.12).

Thus, there are two alternative scenarios demonstrated in ensemble forecasts from late January and early February: (a) enhanced vertical wave activity around 5 February leading to SPV weakening, and (b) suppressed wave activity with little subsequent change in SPV strength. In the next section, we discern the tropospheric drivers for these divergent stratospheric evolutions.

### 2.3.2 Characterizing Tropospheric Uncertainty

Figure 2.3 depicts the linear correlation between the mean $U10_{60}$ forecast for 9–11 February (a period where ensemble members either projected a quiescent vortex or strong

deceleration, c.f. Figure 2.2) and the mean sea-level pressure (MSLP) for 3–5 February (i.e., just before the onset of enhanced vertical wave flux). This correlation is calculated in joined ensembles from 29 January to 1 February in NCEP (to increase ensemble sample size and incorporate a larger range of SPV strengths), and 29 January and 1 February in ECMWF; independent calculations (not shown) for the separate ensembles suggest this is not a result of the difference in character of the forecasts or a facet of the differences in lead-time. We average across forecasts initialized during the onset of predictability of the vortex weakening event (c.f. Figure 2.1) to determine what changed during this window. The results show the strongest correlations between the preceding MSLP field and the strength of the SPV form a dipole between Scandinavia and Greenland. Secondary regions of strong correlation are also located upstream and downstream of the main dipole. The correlation field indicates that ensemble members with lower MSLP over eastern Greenland and higher MSLP over Scandinavia forecast weaker $U10_{60}$.

Based on this correlation analysis, we define the Scandinavia-Greenland dipole in MSLP (hereafter, the S-G dipole) to describe the evolution. We calculate this by subtracting the area-average MSLP in a grid box over Scandinavia (60-70°N, 12.5-42.5°E) from that in a grid box over eastern Greenland (72.5-90°N, 2.5-42.5°W). The MSLP in each grid box is cosine-weighted to account for the convergence of meridians at higher latitudes. The two nodes, primarily based on the track of a cyclone and the development of a Scandinavian ridge (see Figure 2.4), are shown as black dashed lines in Figure 2.3. The Ural high, also shown Figure 2.3, is defined as the area-average MSLP in the grid box 45-60°N, 50-80°E.

To discern the tropospheric drivers of the vertical wave flux, we assess the MSLP evolutions of ECMWF ensemble members from 29 January and 1 February 2018 with the top and bottom 10%-mean 300 hPa [v*T*] for 4–6 February. Results (Figure 2.4) support the correlation analysis from Figure 2.3; a cyclone near Iceland on 3 February progresses up the eastern coast of Greenland and deepens to $<$970 hPa by 5 February in the high-flux members, with a ridge extending from the Azores through Scandinavia, whilst in the

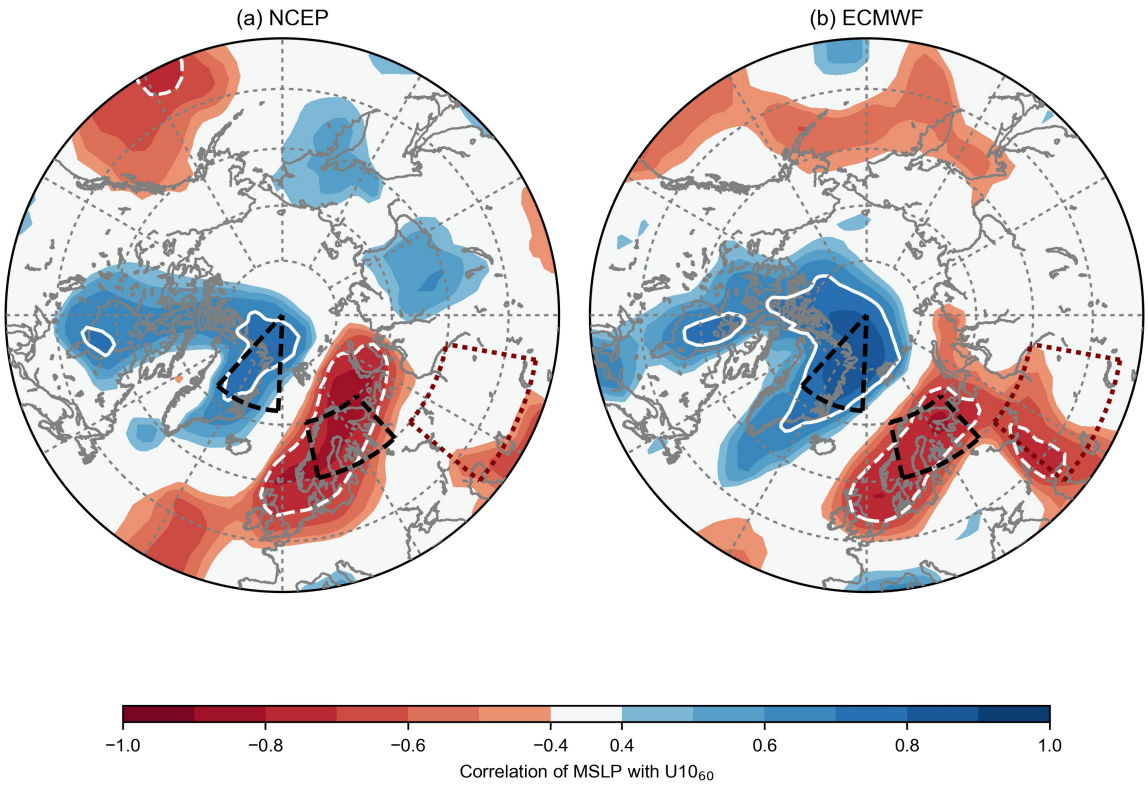


**Figure 2.3:** Linear correlation between average 9–11 February $U10_{60}$ and average 3–5 February mean MSLP from (a) 29 January to 1 February NCEP forecasts and (b) 29 January and 1 February ECMWF forecasts. White lines delineate where the magnitude of the correlation exceeds 0.7. The two nodes of the S-G dipole are shown with black dashed lines, and the location of the Ural high is shown with maroon dotted lines.

low-flux members the cyclone moves south-east towards Europe and weakens without any ridge development. Figure 2.4c demonstrates the dipole structure; pressures are >20 hPa higher (lower) over Scandinavia (Greenland) in the top 10% versus the bottom 10% heat flux members. A similar result is found when the same analysis is performed in the NCEP forecasts (see Figure 2.13).

Next we compare the evolution of the S-G dipole with that of the Ural high (after Karpechko et al. (2018)) (Figure 2.5). The S-G dipole peaked at 52 hPa in ERA-Interim on 5 February. There is rapid divergence after 3 February in accordance with Figure 2.4 (due to discrepancies in both nodes of the dipole), whilst it is also shown that the ensemble members with the largest heat flux more closely follow ERA-Interim verification. There is also a lagged relationship between the S-G dipole evolution and

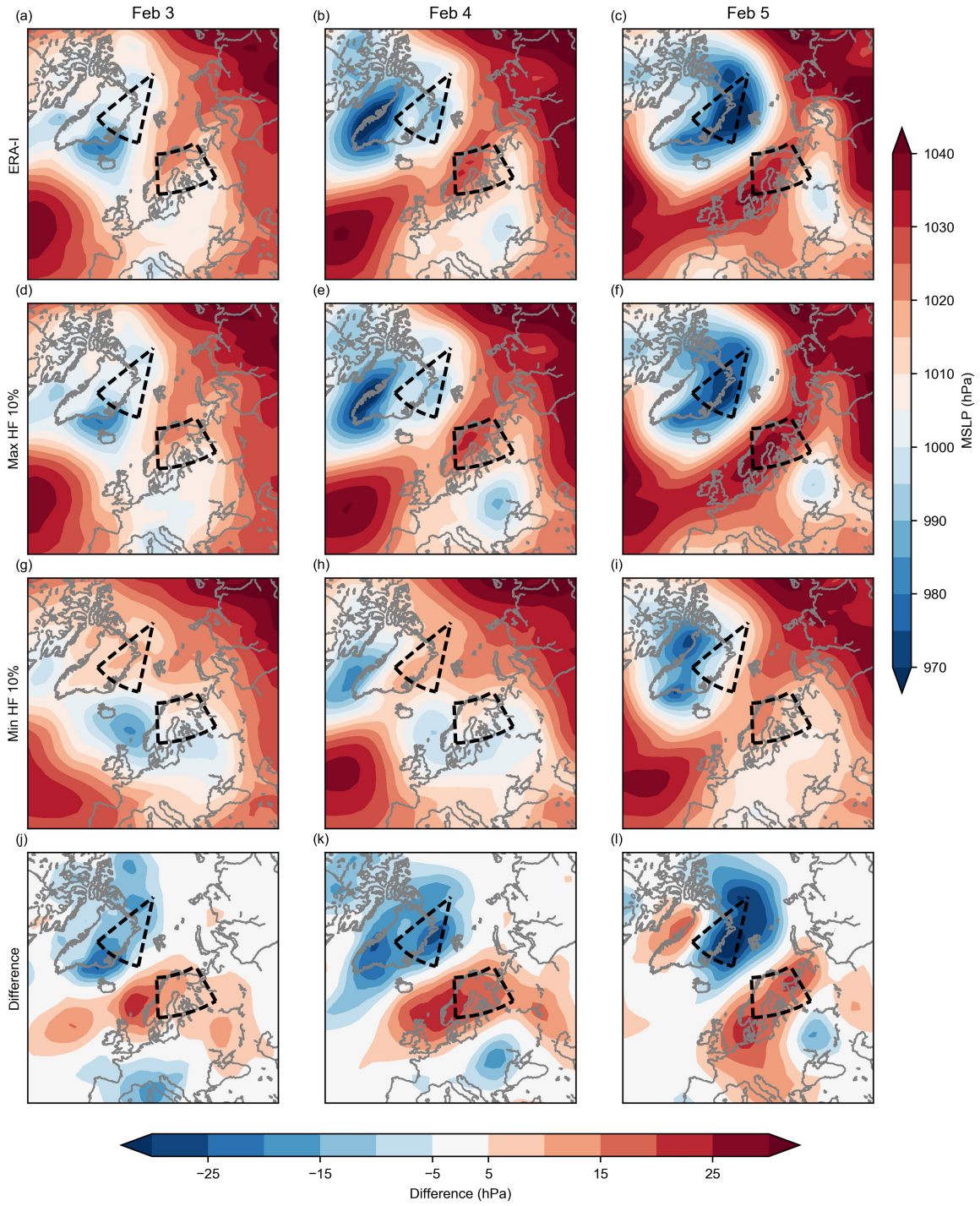


**Figure 2.4:** MSLP for 3–5 February from (a-c) ERA-Interim reanalysis, (d-f) the mean forecast from members of the ECMWF 29 January and 1 February joined ensemble with the top 10% 300 hPa 45–75°N heat flux on 4–6 February, (g-i) the bottom 10%, and (j-l) the difference (d-f – g-i). The two nodes of the S-G dipole are shown with black dashed lines.

the Ural high, and ensemble members with lowest mean 300 hPa heat flux lack both a strong S-G dipole and Ural high. Inspecting potential vorticity (PV) on the 320 K

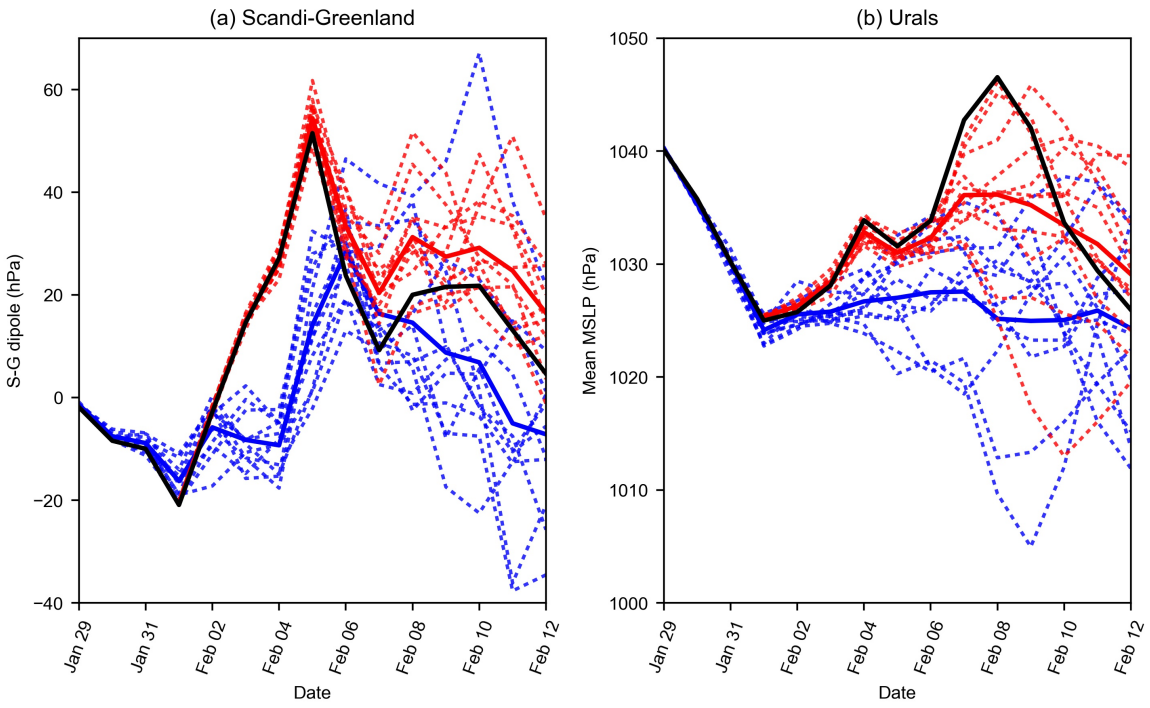


**Figure 2.5:** Time series of ECMWF ensemble forecasts from 29 January and 1 February for (a) the S-G dipole and (b) the Ural high from members with the top (dotted red) and bottom (dotted blue) 10% mean 4–6 February 300 hPa 45–75°N [v*T*]. Their respective means are shown with thick lines coloured accordingly. The verifying evolution from ERA-Interim is shown in black.

isentropic surface in ensemble members with the top 10% mean 300 hPa heat flux (Figure 2.6a) shows a tongue of low PV air (<2 PVU) protruding polewards in the Atlantic sector east of Greenland on 5 February before becoming cut-off and overturning on 8 February, indicative of an anticyclonic Rossby wave break. This evolution is spatially and temporally coherent with both the cyclone track/S-G dipole development and the 300 hPa heat flux. The wave break is not present in ensemble members with the lowest 10% heat flux, which correspondingly lacked a strong S-G dipole (Figure 2.6b). Thus, the predictability of the wave breaking event and its impact on the stationary wave pattern indicates a possible explanation for the abrupt forecast transition, as well as a dynamical mechanism by which the S-G dipole in MSLP relates to both enhanced wave activity and amplification of the Ural high downstream (through the attendant upper-level PV anomaly).

The relationship between the S-G dipole, the Ural high, and 300 hPa heat flux in February 2018 is shown in Figure 2.7. Forecast heat flux increases approximately linearly with S-G

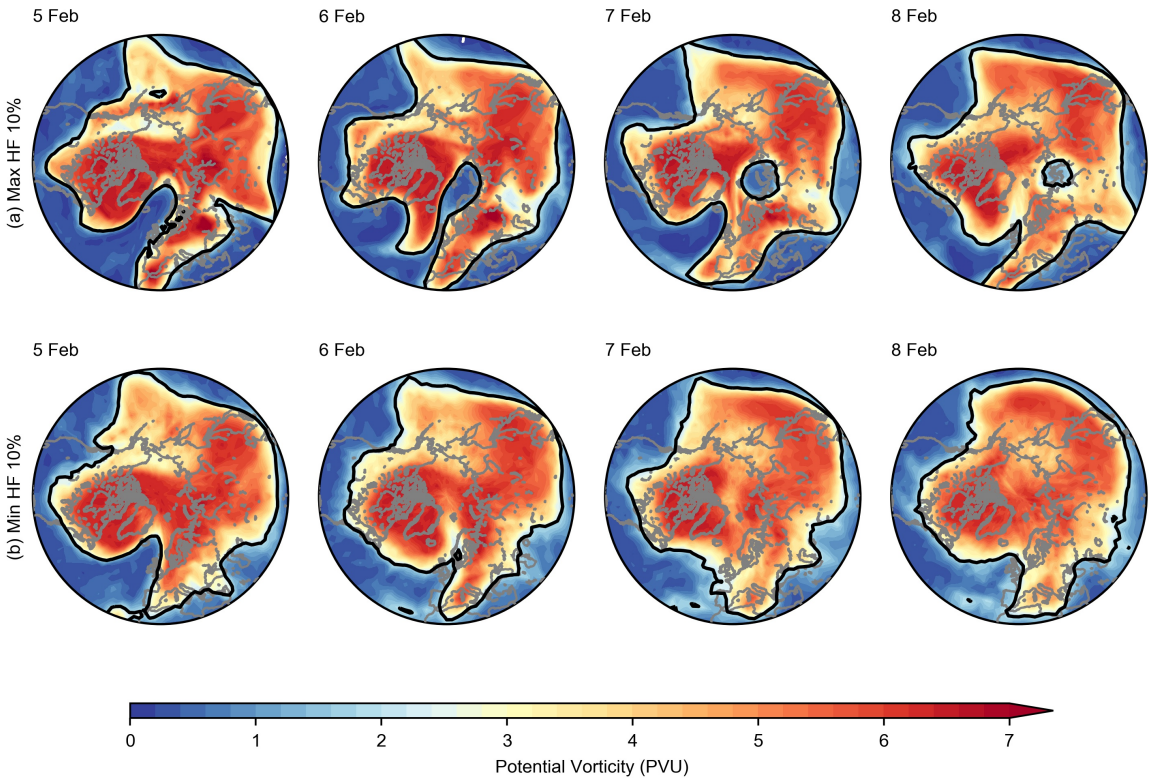


**Figure 2.6:** Forecasts of PV on the 320 K isentropic surface for 5–8 February from members of the ECMWF 29 January and 1 February joined ensemble. (a) shows the mean forecast from members with the top 10% 300 hPa 45–75°N heat flux on 4–6 February and (b) the bottom 10%. The 2 PV unit isoline (PVU, where 1 PVU $= 10^{-6}$ m$^2$ s$^{-1}$ K kg$^{-1}$) is contoured in black.

dipole strength ($r = 0.79$ in NCEP vs. 0.75 in ECMWF) and heat flux is only enhanced for values of the S-G dipole above ∼40 hPa. However, not all ensemble members with an enhanced dipole produce enhanced heat flux; members with the strongest heat flux feature both an amplified S-G dipole *and* a strengthened Ural high. Thus, the enhancement of wave activity and amplification of the Ural high was dependent upon the prior occurrence of the S-G dipole/wave breaking event as well as specifics of the wave break and its interaction with the stratosphere. Figure 2.8 illustrates the surface evolution of the Ural high over 6–8 February in high vs. low heat flux members (c.f. Figure 2.4 and Figure 2.14). The anticyclone that develops over the Urals on 8 February in the high heat flux members is the same system that is present over Scandinavia in the preceding days associated with one node of the S-G dipole; this anticyclone is absent in the low heat flux members, and thus directly links the evolution of the Ural high to the S-G dipole. Therefore, the two

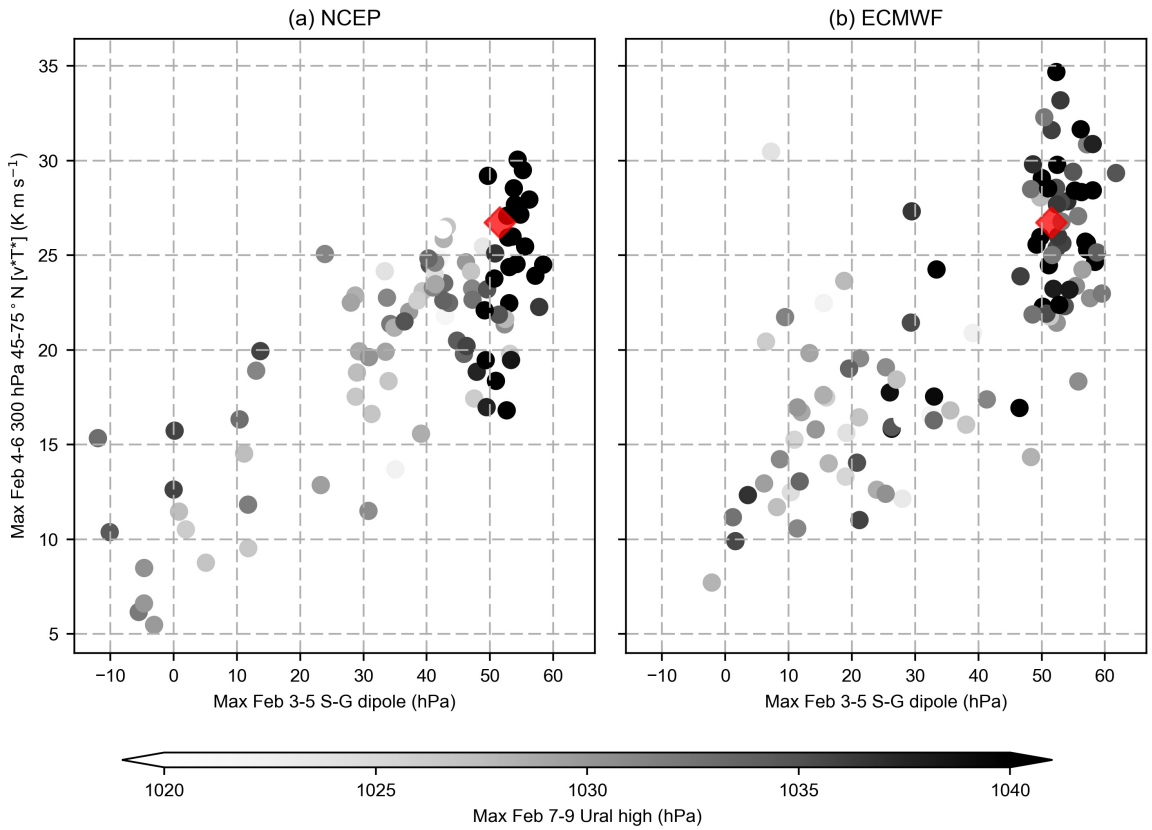


**Figure 2.7:** Scatter plots of maximum 4–6 February 300 hPa [v*T*] versus maximum 3–5 February mean S-G dipole for (a) NCEP ensembles from 29 January to 3 February ($n$ = 96), and (b) ECMWF ensembles from 29 January and 1 February ($n$ = 102). The points are coloured by the corresponding maximum 7–9 February Ural high strength. The verifying ERA-Interim value is shown with a red diamond.

precursors are not independent.

### 2.3.3 Historical S-G Dipoles

In this section we consider historical cases in extended-winter (November–March) where the S-G dipole exceeds 40 hPa (similar in magnitude to the 2018 event, and approximately equal to the 99th percentile of daily November–March 1979-2017 ERA-Interim climatology) to discern whether the dipole is a characteristic of previous cases of SPV weakening. This threshold is not influenced by the time of year of an individual event, as there is little day-to-day variability in the daily-mean and standard deviation of the S-G dipole through the extended winter period. The Ural high is also considered, and by using these previous examples we seek to understand whether the dipole or the Ural high was the root cause of

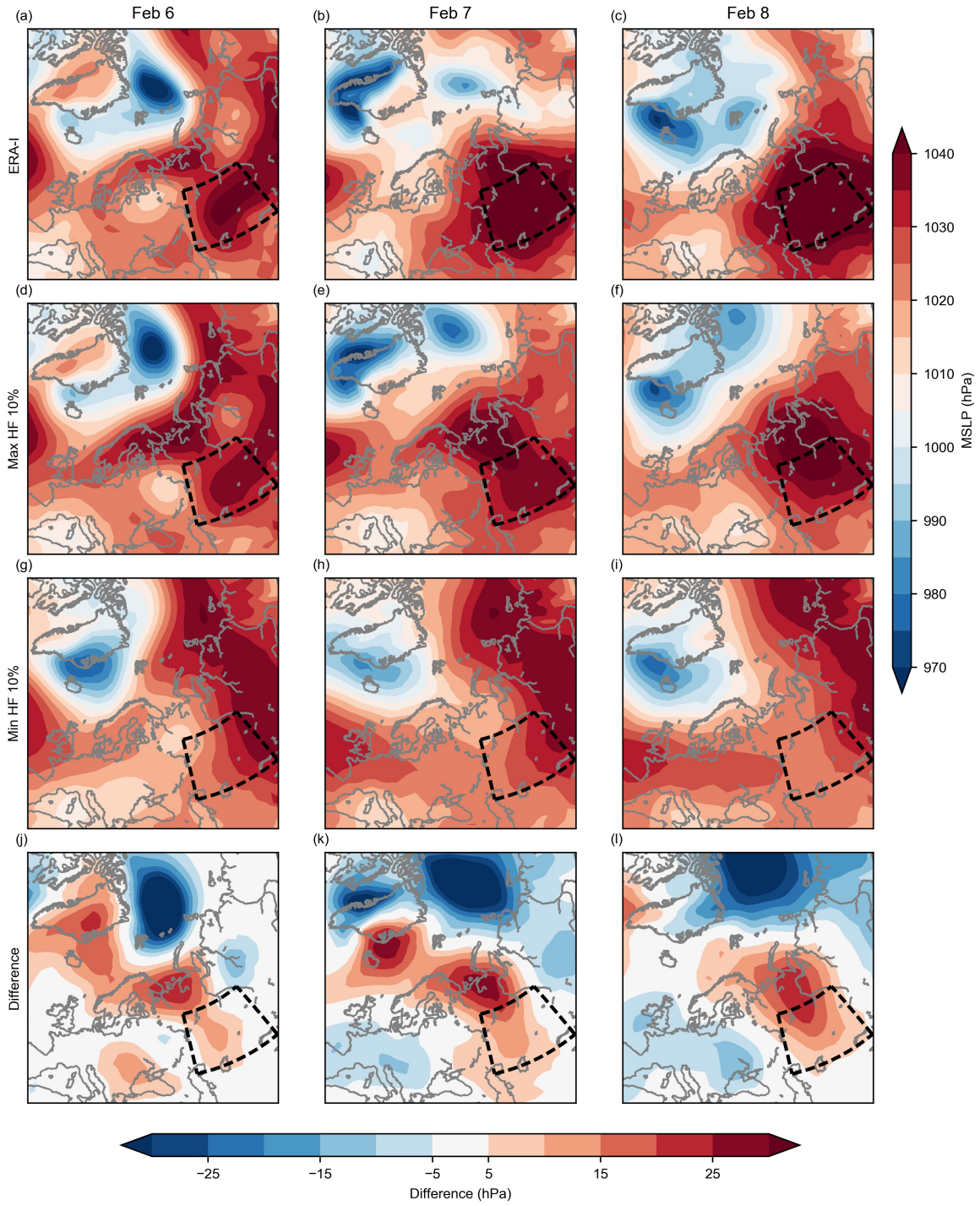


**Figure 2.8:** As in Figure 2.4 but for 6–8 February. The Ural high is indicated with black dashed lines.

the enhanced vertical wave activity.

Motivated by Charlton and Polvani (2007) and their consideration of stratospheric radiative timescales, we use a window of 20 days to separate individual events, yielding a total of

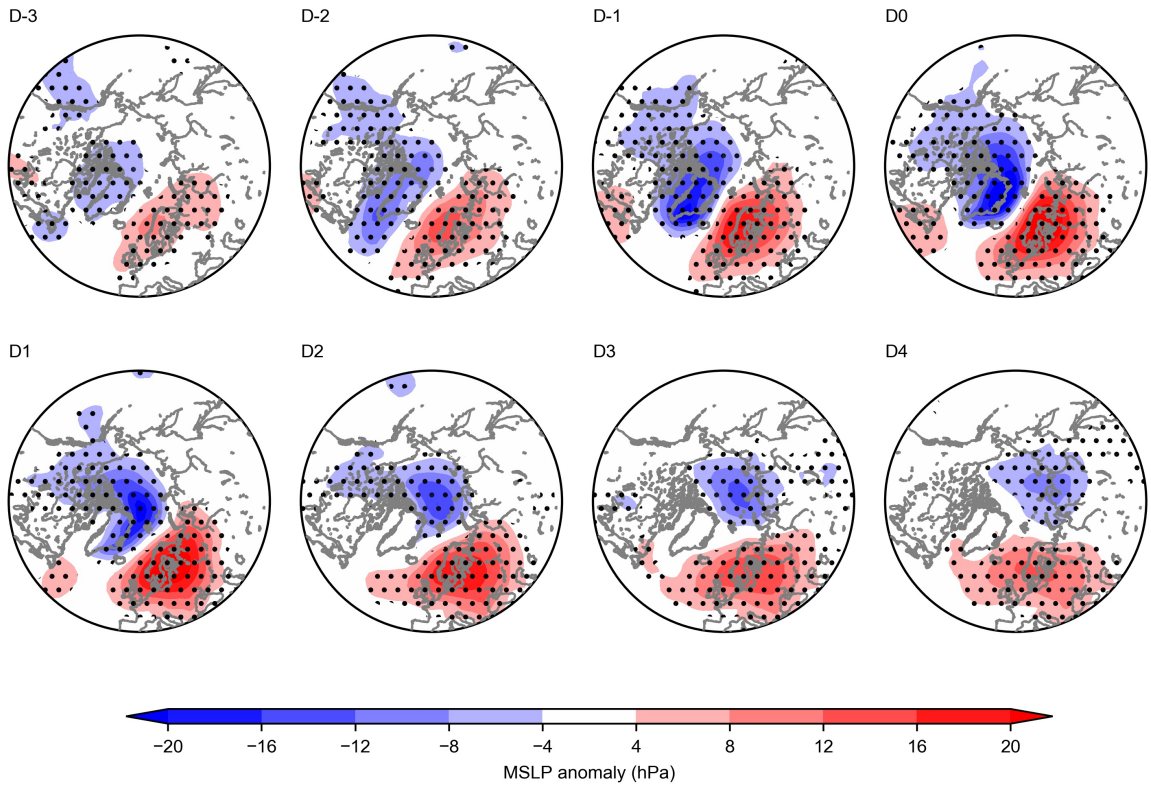


**Figure 2.9:** Composite of MSLP anomalies (with respect to January 1979–March 2017 climatology) from ERA-Interim for the period 3 days before to 4 days after 49 historical S-G dipole events exceeding 40 hPa. Stippling indicates areas significant at the 95% confidence level (for details see Section 2.2).

49 cases (listed in Table 2.2). The strongest S-G dipole, 56 hPa, occurred on 15 March 2015. In Figure 2.9 we see the MSLP lag-composite anomaly evolution, with a cyclone tracking up eastern Greenland into the Arctic region and a concomitant anticyclone over Scandinavia. Notably the anticyclone is of greater persistence throughout this period with transient amplification upon the passage of the cyclone. Following the dipole peak, the anomaly field resembles the Scandinavian blocking Atlantic weather regime (Cassou et al., 2004; Charlton-Perez et al., 2018) and is similar to the precursors to weak SPV episodes shown in Kolstad and Charlton-Perez (2011).

These historical events are also associated with anticyclonic wave breaking (Figure 2.10) similar to that which occurred in 2018, with the wave break in the Atlantic and northern Europe evident through the reversal of the meridional PV gradient on the 315 K isentropic surface in this region. Composites of 45–75°N [v*T*] and 60–90°N geopotential height

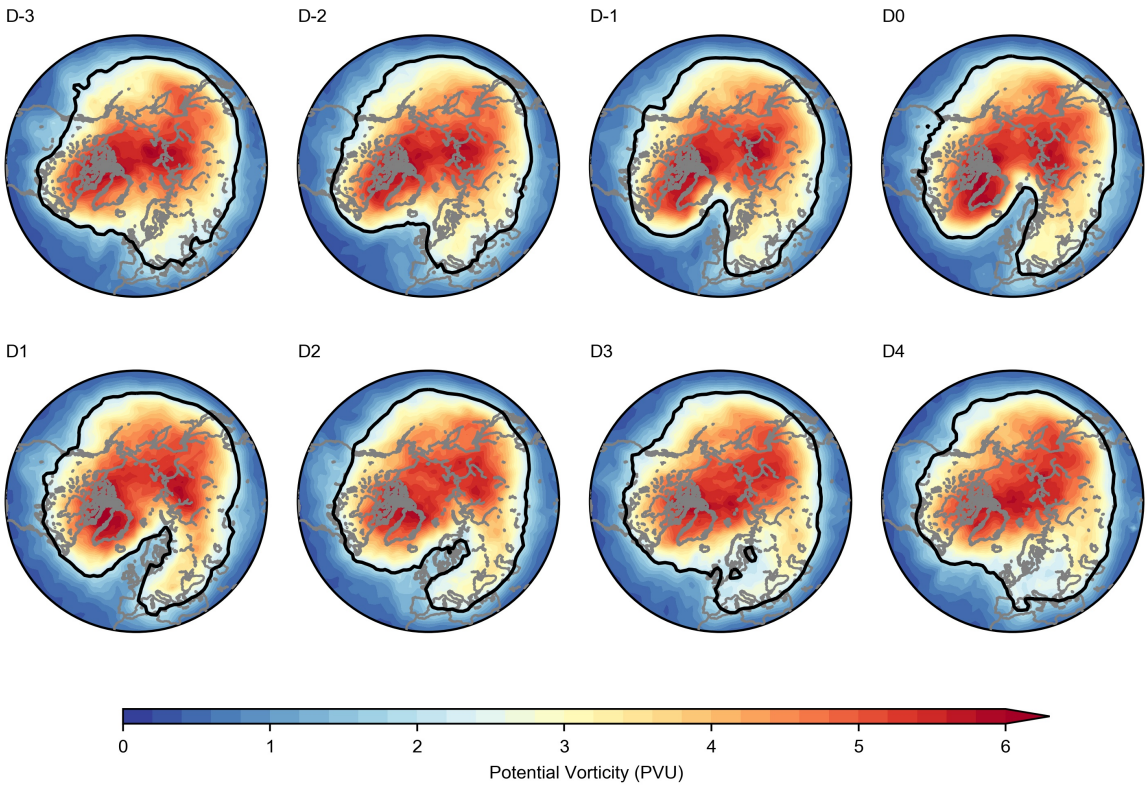


**Figure 2.10:** As in Figure 2.9 but for PV on the 315 K isentropic surface. The thick black line indicates the 2 PVU contour.

for 30 days before and after the peak of the dipole are shown in Figure 2.11. Strong S-G dipoles are associated with a significant vertical wave pulse, and a weakening of the SPV (increasing polar cap geopotential heights indicating a negative stratospheric NAM tendency) in 10–15 days. It should be emphasized that these results show relative vortex weakening, rather than the development of a climatologically weak vortex. Indeed, some cases show a weakening of a strong SPV, or a temporary reduction in the rate of vortex intensification, following an S-G dipole event. The evolution in 2018 (not shown) is very similar to the composites, albeit with increased magnitude.

To discern whether these historically strong S-G dipoles were also associated with enhanced Ural highs, we analyse the change in the Ural high at a 3-day lag from the dipole peak (motivated by the evolution in 2018). There is no clear tendency toward either a strengthening or a weakening Ural high ($\mu = 0.3$ hPa, $\sigma = 10.3$ hPa). Splitting the composites by whether the Ural high weakens or strengthens does not significantly alter the composites: strong

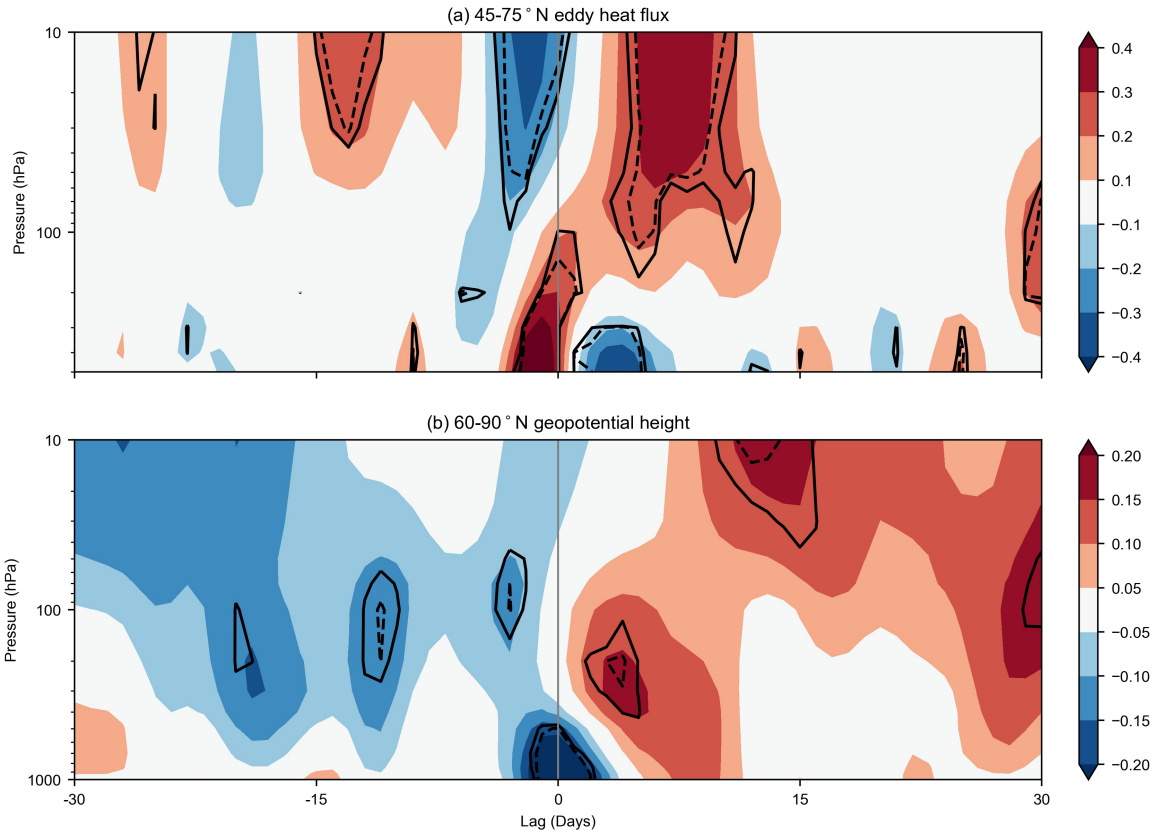


**Figure 2.11:** Composite of anomalies in (a) 45–75°N meridional eddy heat flux [v*T*] and (b) 60–90°N geopotential height for 30 days before and after 49 events in ERA-Interim 1979–2017 where the S-G dipole exceeded 40 hPa. Anomalies are standardized departures and are filtered using a 1 $\sigma$ Gaussian smoother; in (b) these are shown relative to the mean for the 61-day window to show relative tendency. The gray vertical line indicates the day on which the 40 hPa threshold was exceeded. Solid (dashed) black contours indicate regions significant at the 90% (95%) confidence level (for details see Section 2.2).

S-G dipoles followed by a weakening of the Ural high still show enhanced heat flux and a weakened polar vortex in the following days. This indicates it is the wave break associated with the S-G dipole, not the resultant Ural high, which drives the enhanced vertical wave flux – and that instead, in 2018, the Ural high was a consequence of the preceding evolution.

Next, we assess the association between the S-G dipole and observed major mid-winter SSWs prior to 2018 (Table 2.1). Of the 23 SSWs (Karpechko et al., 2017), we find 8 (35%) followed a similar evolution to 2018 and were preceded by an S-G dipole exceeding 40 hPa within 15 days of the start date of the SSW. Given the total of 345 days preceding the 23 events (and assuming independence), this is 2.3 times larger than the climatological

**Table 2.1:** Major SSWs in the period 1979–2017 (following Karpechko et al. (2017)) and the peak value of the S-G dipole in the 15 days before the event according to ERA-Interim reanalysis. Those events exceeding 40 hPa are shown in bold.

| SSW event | Peak S-G index (hPa) |
|---|---|
| **February 1979** | **41** |
| **February 1980** | **43** |
| **March 1981** | **48** |
| December 1981 | 13 |
| **February 1984** | **45** |
| **January 1985** | **44** |
| **January 1987** | **43** |
| December 1987 | 26 |
| March 1988 | 21 |
| February 1989 | 39 |
| December 1998 | 34 |
| February 1999 | 18 |
| March 2000 | 24 |
| February 2001 | 31 |
| December 2001 | 24 |
| January 2003 | 2 |
| January 2004 | 12 |
| **January 2006** | **42** |
| February 2007 | 17 |
| **February 2008** | **40** |
| January 2009 | 22 |
| February 2010 | 36 |
| January 2013 | 27 |

likelihood (since 40 hPa is approximately the 99th percentile, it would be expected that it was exceeded on 3-4 days). We note that the 2018 event was stronger than any of these prior events associated with major SSWs (the previous strongest being 48 hPa preceding the major SSW in March 1981), possibly a facet of 2018 being the event used to define the index. Although the major SSW in February 2018 was a vortex split, 6 of the observed SSWs with a strong S-G dipole precursor were displacement events (Karpechko et al., 2017) suggesting this pattern does not itself induce a specific stratospheric evolution but acts to amplify an existing planetary wave structure. When 2018 is included, 78% of the major SSWs preceded by an amplified S-G dipole were downward-propagating (Karpechko et al., 2017), with only March 1981 and February 2008 otherwise. This is larger than the observed ratio of 57% (although the sample is too small to draw robust conclusions), but is in agreement with Birner and Albers (2017) who note larger tropospheric impacts following

SSWs preceded by enhanced tropospheric wave activity. We further note that 33% (90%) of the S-G dipole events considered here were associated with daily 500 hPa [v*T*] exceeding 2 $\sigma$ (1 $\sigma$) within 5 days either side of the events, indicating the dipole is an important contributor to anomalously high zonal-mean tropospheric wave flux in general.

## 2.4 Discussion and Conclusions

In this study we have shown that the abrupt onset of predictions of stratospheric polar vortex (SPV) weakening and sudden stratospheric warming (SSW) in February 2018 was driven by an anticyclonic Rossby wave break (Figure 2.6) associated with the track and intensity of a cyclone over eastern Greenland, and an associated ridge over Scandinavia. From this, we define a Scandinavia-Greenland (S-G) dipole index in MSLP to describe the evolution, and show that this was not well-forecast at long lead-times. The location and intensity of the cyclone was a rare occurrence, with the mean MSLP in the Greenland node and the dipole itself exceeding the 99th percentile of extended winter months from 1979-2017. Occurrences of similarly strong S-G dipoles in reanalysis are shown to be associated with anticyclonic wave breaking in the Atlantic sector (Figure 2.10), which induces anomalously strong vertical wave activity (Figure 2.11a), and a rapid tendency towards a weakened SPV/negative stratospheric NAM (Figure 2.11b) within 10 days. This indicates that the evolution in 2018 was not a characteristic of the specific flow configuration but a more general mechanism for vortex weakening present in other events.

We have also shown that the S-G dipole and wave breaking event was important for amplifying a high pressure system over the Urals, first described in Karpechko et al. (2018) as a surface-pressure precursor of the 2018 SSW. Our results suggest that the Ural high was likely a consequence of the wave breaking event which drove stratospheric wave activity leading to the SSW, rather than a primary driver itself. The initial divergence in the evolution of the SPV strength and the onset of the enhanced vertical wave flux occurred around 5 February (Figure 2.2), preceding the amplification of the Ural high which followed on 8 February (Figure 2.5b), which further indicates the Ural high was a

secondary response. It is likely that the persistence of the Ural high might have resulted in its detection in the averaging used in Karpechko et al. (2018), rather than the transience of the S-G dipole/wave break.

Our results differ from previous work through using a 'bottom-up' perspective, assessing the stratospheric response to tropospheric events. We provide a particular *transient* trigger which would not be easily distinguished through 'top-down' time-mean composites (where the tropospheric configuration prior to stratospheric events is considered). This helps illuminate mechanisms by which persistent tropospheric blocking, including Scandinavian blocking which has previously been shown to precede SSWs (Martius et al., 2009; Kolstad and Charlton-Perez, 2011; Cohen and Jones, 2011), can produce *sudden* changes in the stratospheric circulation. Furthermore, our results apply to a wider range of SPV variability than major SSWs – even the case of weakening a climatologically strong vortex towards an average state – which helps describe precursors of a larger proportion of the sub-seasonal behaviour. We therefore suggest the S-G dipole should be monitored operationally as a precursor to SPV weakening. Changes and uncertainty in its forecasts may help to qualitatively identify sources of uncertainty in stratospheric forecasts.

The intensity of the S-G dipole in 2018 was not well-forecast, driven by uncertainty in the track and intensity of an Atlantic cyclone. At longer lead-times, model biases in storm track and intensity may negatively impact the skill in predicting such events. For example, Frame et al. (2015) showed cyclone intensity decayed with lead-time up to 15 days, which would constrain the ability of forecast models to produce strong S-G dipoles sufficient for strong wave breaking and vortex weakening, whilst Gray et al. (2014) and Saffin et al. (2017) also showed biases in tropopause PV and Rossby wave structure which may limit the ability to capture these types of wave breaking episodes and associated stratospheric variability. These considerations are consistent with a deterministic limit on SSW predictability (Karpechko, 2018; Taguchi, 2018).

The occurrence of strong S-G dipoles requires a poleward-shifted Atlantic storm track, which is associated with the positive NAM/NAO pattern. This is often related to the prior occurrence of a strengthened SPV (e.g. Baldwin and Dunkerton, 2001) and SSWs are typically preceded by strong SPV conditions (e.g. Charlton and Polvani, 2007). This behaviour could imply a two-way coupling in which the vortex drives its own variability - akin to a self-sustaining oscillator. Several studies (e.g. Lorenz and DeWeaver, 2007; Tamarin and Kaspi, 2017) have indicated a poleward shift in the North Atlantic storm track during winter under future climate change. This may lead to an increased frequency of strong S-G dipoles and thus more frequent wave breaking events and stratospheric vortex weakening, but the aforementioned biases may reduce the ability of climate models to fully represent this source of sub-seasonal variability.

## Acknowledgments

S.H.L. was funded by the Natural Environment Research Council (NERC) via the SCENARIO Doctoral Training Partnership (NE/L002566/1) at the University of Reading. S.J.W. was supported by the National Centre for Atmospheric Science, a NERC collaborative centre, under contract R8/H12/83/001. This work is based on S2S data, a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP). S2S and ERA-Interim reanalysis data are available online at `https://apps.ecmwf.int/datasets/`. Pre-processing of data used here was completed as part of the WCRP SPARC SNAP (Stratosphere-troposphere Processes and their Role in Climate: Stratospheric Network for the Assessment of Predictability) programme. The authors acknowledge three anonymous reviewers for their helpful suggestions.

## 2.5 Supporting Information

**Table 2.2:** Dates and magnitudes of the 49 S-G dipole events exceeding 40 hPa in the period January 1979 to December 2017 according to ERA-Interim reanalysis. Only events separated by 20 days are included. Those events which are followed by a major SSW within 15 days are shown in bold.

| Date | S-G dipole (hPa) | Date | S-G dipole (hPa) |
|---|---|---|---|
| **20 February 1979** | **40** | 12 March 1995 | 41 |
| 21 December 1979 | 40 | 4 December 1995 | 54 |
| **21 February 1980** | **43** | 19 January 1996 | 44 |
| 2 November 1980 | 40 | 21 February 1996 | 52 |
| 28 December 1980 | 41 | 13 March 1996 | 48 |
| **18 February 1981** | **48** | 25 January 1997 | 51 |
| 2 March 1983 | 53 | 16 December 1997 | 41 |
| **11 February 1984** | **45** | 7 January 2002 | 44 |
| 28 November 1984 | 43 | 7 December 2002 | 47 |
| **12 January 1985** | **43** | 21 November 2005 | 40 |
| 8 March 1985 | 41 | **9 January 2006** | **42** |
| 17 November 1985 | 42 | 24 March 2007 | 46 |
| 3 February 1986 | 43 | 1 January 2008 | 43 |
| 11 March 1986 | 45 | **10 February 2008** | **40** |
| 25 December 1986 | 41 | 25 December 2008 | 46 |
| **19 January 1987** | **43** | 25 February 2011 | 41 |
| 7 March 1987 | 51 | 30 January 2012 | 44 |
| 2 January 1989 | 44 | 16 February 2015 | 50 |
| 17 December 1990 | 46 | 15 March 2015 | 56 |
| 3 February 1991 | 50 | 29 December 2015 | 46 |
| 8 December 1991 | 46 | 12 March 2016 | 41 |
| 22 January 1992 | 43 | 14 November 2016 | 42 |
| 14 February 1993 | 40 | 17 January 2017 | 42 |
| 11 November 1993 | 48 | 12 February 2017 | 42 |
| 12 February 1994 | 41 | | |

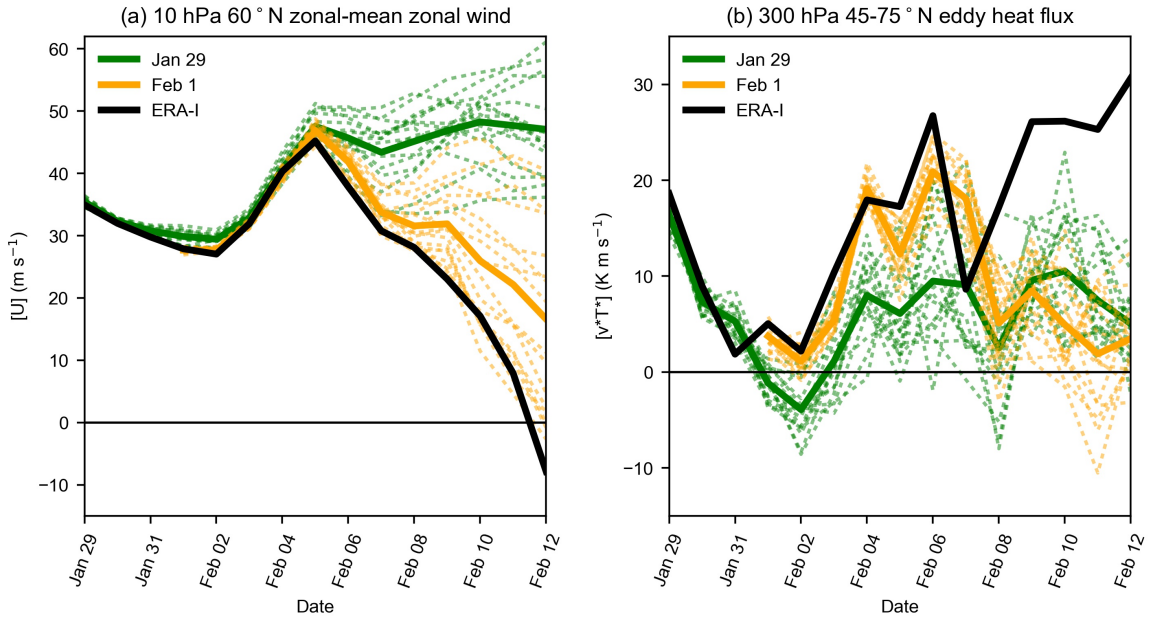


**Figure 2.12:** As Figure 2.2 in the main text but for the corresponding NCEP ensembles.

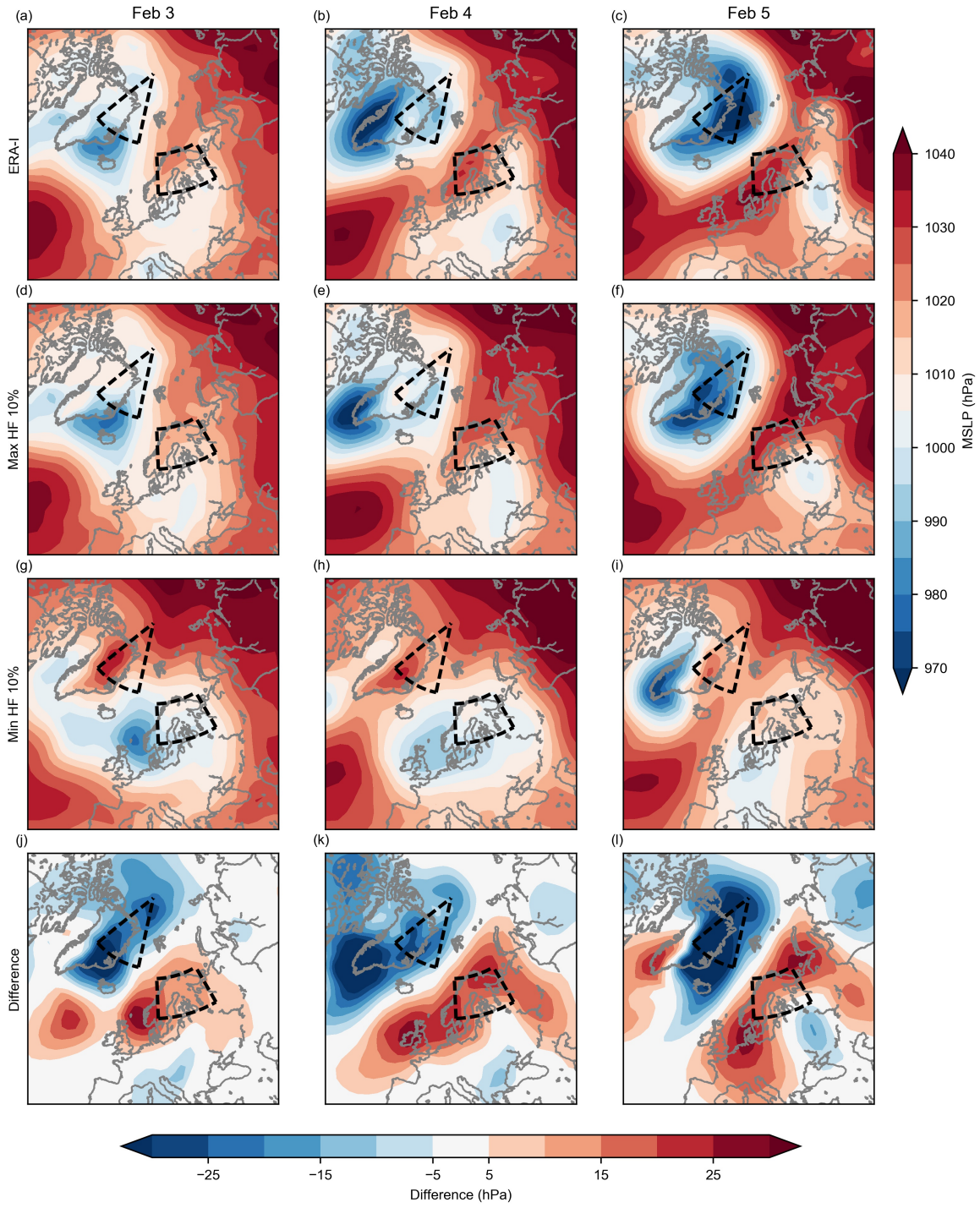


**Figure 2.13:** As Figure 2.4 in the main text but for the corresponding NCEP ensembles.

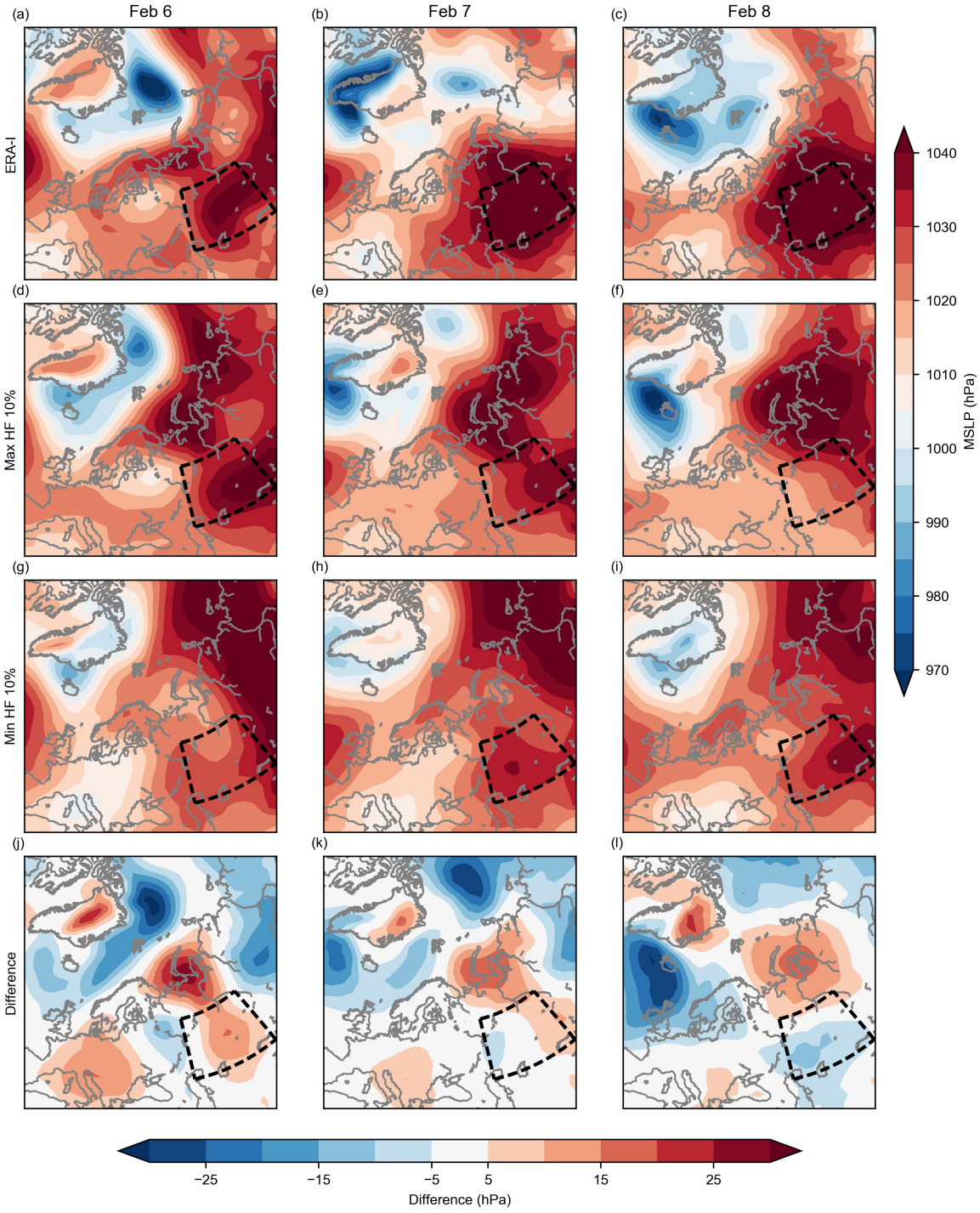


**Figure 2.14:** As Figure 2.8 in the main text but for the corresponding NCEP ensembles.

# Chapter 3

# Representation of the Scandinavia-Greenland Pattern and its Relationship with the Polar Vortex in S2S Models

This chapter has been published in *Quarterly Journal of the Royal Meteorological Society* with the following reference:

**Lee, S. H.**, A. J. Charlton-Perez, J. C. Furtado, and S. J. Woolnough, 2020: Representation of the Scandinavia-Greenland Pattern and its Relationship with the Polar Vortex in S2S Models. *Quarterly Journal of the Royal Meteorological Society*, **146**, 4083-4098, https://doi.org/10.1002/qj.3892.

## Abstract

The strength of the stratospheric polar vortex is a key contributor to subseasonal prediction during boreal winter. Anomalously weak polar vortex events can be induced by enhanced vertically propagating Rossby waves from the troposphere, driven by blocking and wave breaking. Here, we analyse a tropospheric pattern – the Scandinavia-Greenland (S-G) pattern – associated with both processes. The S-G pattern is defined as the second empirical orthogonal function (EOF) of mean sea-level pressure in the northeast Atlantic. The first EOF is a zonal pattern resembling the North Atlantic Oscillation. We show that the S-G pattern is associated with a transient amplification of planetary wavenumber-2 and meridional eddy heat flux, followed by the onset of a weakened polar vortex which persists for the next 2 months. We then analyse 10 different models from the S2S database, finding that while all models represent the structure of the S-G pattern well, some models have a zonal bias with more than the observed variability in their first EOF, and accordingly less in their second EOF. This bias is largest in models with the lowest resolution. Skill in predicting the S-G pattern is not high beyond week 2 in any model, in contrast to the zonal pattern. We find that the relationship between the S-G pattern and enhanced eddy heat flux and a weakened polar vortex is initially well-represented but significantly decays with lead time in most S2S models. Our results motivate improved representation of the S-G pattern and its stratospheric response at longer lead-times for improved subseasonal prediction of the stratospheric polar vortex.


## 3.1 Introduction

Variability in the strength of the wintertime Arctic stratospheric polar vortex (SPV) significantly impacts both the behaviour *and* the predictability of Northern Hemisphere (NH) tropospheric weather patterns on subseasonal to seasonal (S2S) timescales (∼2 weeks to 2 months) (Domeisen et al., 2020c; Kidston et al., 2015; Tripathi et al., 2015b). Of particular importance is the development of a weakened SPV, which includes major sudden stratospheric warming (SSW) events – when the mean westerly circulation of the

SPV entirely reverses to easterlies during mid-winter (e.g. Butler et al., 2015; Charlton and Polvani, 2007). In the subsequent weeks-to-months following the onset of a weakened SPV, there is an increased likelihood of NH cold air outbreaks (Kautz et al., 2020; Kolstad et al., 2010; Kretschmer et al., 2018a,b), through the development of a negative tropospheric Northern Annular Mode (NAM) and associated equatorward eddy-driven jet shift (Baldwin and Dunkerton, 2001) or by favouring particular regional weather regimes (Charlton-Perez et al., 2018; Lee et al., 2019c). Weak or disrupted SPV states are driven by the vertical propagation and subsequent breaking of large-scale (planetary wavenumbers 1-3) Rossby waves from the troposphere to the stratosphere (Charney and Drazin, 1961; Matsuno, 1971; McIntyre and Palmer, 1983), although the tropospheric wave activity need not be anomalously large with pre-conditioning of the SPV playing an important role (Birner and Albers, 2017; de la Cámara et al., 2019; Lawrence and Manney, 2020).

Despite substantially longer predictability and persistence timescales in the stratosphere than in the troposphere (e.g. Baldwin et al., 2003; Son et al., 2020), predicting the *onset* of significant SPV circulation anomalies on S2S timescales remains challenging and is likely to be a contributing factor to poor skill in wintertime NH subseasonal forecasts. Several recent studies have analysed the performance of prediction systems contributing to the World Climate Research Program and World Weather Research Program S2S Prediction Project database (Vitart et al., 2017), comprising of operational forecasts and hindcasts from various modelling centres around the world. Domeisen et al. (2020b) determined that S2S model skill in predicting both weak and strong SPV states was generally limited to the medium-range (1-2 weeks). They found particularly poor predictability for major SSWs classified as "split" vortex events, usually associated with amplification of a wavenumber-2 disturbance. Models with the greatest vertical resolution in the stratosphere and highest lid ("high-top" models) performed best. However, models with the highest vertical resolution also generally have higher horizontal resolution, making disentangling the respective influences difficult. These results are consistent with those of Taguchi (2018) and Karpechko (2018), who assessed the predictability of major SSWs in S2S model and European

Centre for Medium Range Weather Forecasts (ECMWF) hindcasts respectively. Additionally, Butler et al. (2019b), who also used S2S hindcasts, found dynamically-driven final stratospheric warmings (akin to major SSWs) were typically poorly predicted beyond week 2.

A component limiting the skill in predicting these stratospheric events may arise from the predictability of tropospheric wave activity, or its interaction with the stratospheric mean state in the model. On subseasonal timescales, sources of enhanced tropospheric wave activity linked to SPV variability are typically associated with processes which constructively interfere with the mean planetary wave pattern. These include the Madden-Julian Oscillation (MJO) (Garfinkel et al., 2012, 2014; Barnes et al., 2019; Green and Furtado, 2019), tropospheric blocking (Quiroz, 1986; Martius et al., 2009; Attard and Lang, 2019; Peings, 2019), anticyclonic Rossby wave breaking (Lee et al., 2019b), and extratropical cyclones (Attard and Lang, 2019; Coy et al., 2009; Coy and Pawson, 2015). The latter three are dynamically interconnected; Rossby wave breaking is a key driver of blocking (Masato et al., 2012), while warm sector processes associated with extratropical cyclones play a significant role in block onset (e.g. Maddison et al., 2019). Moreover, these are also tropospheric processes known to have particularly poor predictability. Quinting and Vitart (2019) assessed S2S model representation of Rossby wave packets, finding that all models generally overestimated the propagation distance of wave packets in the eastern North Atlantic, with attendant negative blocking biases. Models with the coarsest resolution (both horizontal and vertical) exhibited the largest biases. Block-onset biases in the northeast Atlantic are in agreement with the results of Ferranti et al. (2015). They found that the ECMWF ensemble forecasts exhibited the worst skill when transitioning from a positive North Atlantic Oscillation (NAO+) to a Scandinavian blocking regime (Cassou, 2008), overestimating the persistence of the prior zonal state.

Various studies have explored the impact of model resolution (including vertical resolution) on the development and maintenance of blocking (e.g Anstey et al., 2013; Berckmans et al., 2013; Davini et al., 2017; Matsueda, 2009; Schiemann et al., 2017). There is good

agreement of significant improvements, especially in the Euro-Atlantic region, at higher resolution, owing to better representation of several of the processes involved directly and indirectly in blocking. These processes include, but are not limited to: orography and its impact on the generation of planetary waves (including the climatological tilt of the North Atlantic jet); the strength of the North Atlantic jet and associated wave breaking in its exit region; the representation of the trimodal variability of the North Atlantic jet; warm sector ascent and tropopause outflow; and the overall model mean state. Given the link between Scandinavian/Ural blocking and SSWs (Martius et al., 2009; Bao et al., 2017; Peings, 2019), model errors in this region may impact subseasonal skill through SPV variability.

In a case-study of the February 2018 SSW, Lee et al. (2019b) (hereafter, L19) related the relatively short predictability onset ($\sim$12 days) to a poorly-predicted anticyclonic wave break in the North Atlantic (diagnosed as the "Scandinavia-Greenland (S-G) dipole"), which drove enhanced tropospheric wave activity and led to the onset of Ural blocking (Karpechko et al., 2018). While L19 found similar S-G dipoles were evident prior to previously observed SSWs, they did not assess its forecast predictability beyond the 2018 event. Motivated by this, in the present study we define a similar, but more generalised pattern, and assess its representation and predictability in 10 extended-range models from the S2S database. The ultimate aim is to determine whether there exist lead-time or model dependent biases in the S-G pattern and its influence on the SPV, which may then contribute toward limiting subseasonal stratospheric skill.

The remainder of the paper is thus laid out as follows. In Section 3.2, we introduce the datasets and methods used. Section 3.3 defines the S-G pattern and its relationship with the SPV in reanalysis. We then analyse the representation and predictability of the S-G pattern in S2S model hindcasts in Section 3.4. Section 3.5 analyses the modelled relationship of the S-G pattern with the SPV in S2S hindcasts. Conclusions of our work follow.

## 3.2 Data and Methods

We use hindcasts (i.e., re-forecasts for dates in the past) launched between November and March (NDJFM) from 10 models from the S2S database, the details of which are shown in Table 3.1. The chosen model versions were used to produce operational forecasts during NDJFM 2018-19, though the models range in age from 2011 (NCEP) to contemporary. The hindcast ensemble sizes vary from 3 (KMA) to 33 (BoM) and are smaller and/or less frequent than their operational counterparts (with between 15 (JMA) and 151 (CMA and NCEP) initialisations per NDJFM period), yielding a different number of total forecasts for each model. Moreover, the hindcast periods for each model differ. For inter-comparison purposes, we select only the period common to all models, yielding 11 full winters from 2000-2010, where the year refers to that of the January. For historical analysis and verification, we use the ECMWF ERA5 reanalysis (Hersbach et al., 2020). All data are sampled once-daily at 00Z and are re-gridded to 2.5° horizontal resolution for inter-comparison purposes (hindcasts are stored in the S2S database at 1.5° resolution, except BoM which is stored at 2.5°).

**Table 3.1:** Details of the S2S model hindcasts used in this study. These model versions were chosen as they were all used to produce real-time forecasts during November–March 2018-19. The ECMWF and JMA models switch to a lower horizontal resolution after 15 and 18 days respectively.

| Centre | Model Version | Ensemble size | Resolution | Model Top | Runs per NDJFM |
|---|---|---|---|---|---|
| ECMWF | CY45R1 | 11 | Tco639/319 L91 | 0.01 hPa | 43 |
| ECCC | GEPS5 | 4 | 0.35° x 0.35° L45 | 0.1 hPa | 22 |
| JMA | GEPS1701 | 5 | TL479/319 L100 | 0.01 hPa | 15 |
| UKMO | GloSea5-GC2.0 | 7 | 0.83° x 0.56° L85 | 85 km | 20 |
| KMA | GloSea5-GC2.0 | 3 | 0.83° x 0.56° L85 | 85 km | 20 |
| CNRM | CNRM-CM 6.0 | 15 | TL255 L91 | 0.01 hPa | 20 |
| NCEP | CFSv2 | 4 | T126 L64 | 0.02 hPa | 151 |
| HMCR | RUMS | 10 | 1.1° x 1.4° L28 | 5 hPa | 22 |
| CMA | BCC-CPS-S2Sv1 | 4 | T106 L40 | 0.5 hPa | 151 |
| BoM | POAMA-P24 | 3 x 11 | T47 L17 | 10 hPa | 30 |

40-80°N-averaged zonal-mean meridional eddy heat flux (denoted [v*T*], where the star denotes a departure from the zonal-mean, and the square brackets dictate a zonally-averaged quantity) is used as a proxy for vertically-propagating wave activity flux, since it is directly proportional to the vertical component of the Eliassen-Palm flux vector (e.g. Edmon Jr et al., 1980). 60-90°N area-averaged (polar cap) geopotential height anomalies are used as a diagnostic for the strength of the SPV; positive anomalies indicate an anomalously weak SPV (i.e., a negative NAM). Both [v*T*] and polar cap heights are weighted by cosine-latitude. Anomalies are computed with respect to the daily 00Z climatology; this is initialisation-date, lead-time dependent in the case of the S2S hindcasts. Standardised anomalies are computed in the hindcasts by dividing by the initialisation-date, lead-time dependent standard deviation of all ensemble members. In the BoM model, the 33 ensemble members comprise 3 different configurations of the model with 11 members each (Hudson et al., 2013); anomalies are thus computed with respect to the climatology for each version before forming the grand ensemble.

Empirical orthogonal function (EOF) analysis is performed on mean sea-level pressure (MSLP) anomalies in a grid box in the North Atlantic covering parts of Scandinavia and Greenland, bounded by 60-85°N, 60°W-50°E. These bounds are chosen to approximately cover a similar region to the S-G dipole of L19, as well as a consideration of the climatological exit region of the eddy-driven jet and the stationary wave pattern in the North Atlantic around 60°N. All data are weighted by the square-root of the cosine of latitude to provide equal-area weighting in the computation of the EOFs, and in all cases, the EOFs are scaled by the square-root of the eigenvalue to give unit standard deviation of the principal component (PC) timeseries. The resultant patterns are largely insensitive to the changes in the domain boundaries within ∼10°. In the hindcasts, EOFs are either analysed as (a) the "model" EOFs, or (b) the projection onto the ERA5 "observed" EOF over the same period. In the case of (a), MSLP anomalies are computed with respect to the model climate for each initialisation date over the November 1999–March 2010 period, and then EOF analysis is performed across all ensemble members at each forecast day for initialisations through the

NDJFM period. There is negligible change to the variance or pattern correlation statistics of the ERA5 EOFs over the equivalent time period as a 4 week model forecast. However, beyond 4 weeks, there are some changes due to the seasonal cycle, but the results are qualitatively similar to week 3-4. For brevity, we do not show analysis of forecasts beyond week 4.

Statistical significance of the regression analyses is assessed by bootstrap re-sampling, performed 5,000 times on random paired samples (with replacement) selected from the data for each point in space/time. If zero lies outside the 2.5th to 97.5th percentiles of these re-sampled slopes, then the regression is deemed statistically significant at the 95% confidence level.

## 3.3 S-G pattern in Reanalysis

### 3.3.1 EOF structure and characteristics

The first two EOFs, shown in Figure 3.1 as a regression with MSLP anomalies, together explain the majority (62%) of the variance in the analysis domain and are well-separated according to the criterion of North et al. (1982) (Figure 3.12 in the supporting information). There is little difference between the EOFs for the full ERA5 dataset and the much smaller S2S common period (Figure 3.13); for the sake of a larger sample size, we primarily analyse the full ERA5 dataset. The leading EOF (Figure 3.1a) explains 36% of the variance; we define it such that a positive loading is characterised by a cyclonic MSLP anomaly centred over Iceland extending across most of the analysis domain. There is also an associated remote anticyclonic MSLP anomaly across the central North Atlantic resembling the Azores high. Overall, the structure of this first EOF is similar to the NAO despite the much smaller region over which it was computed (for example, the NAO of Hurrell (1995) is computed over 20-80°N 90°W-40°E). The pattern is persistent; the PC timeseries has an autocorrelation *e*-folding timescale of 7 days. Hereafter, we refer to this first EOF as the "zonal pattern".

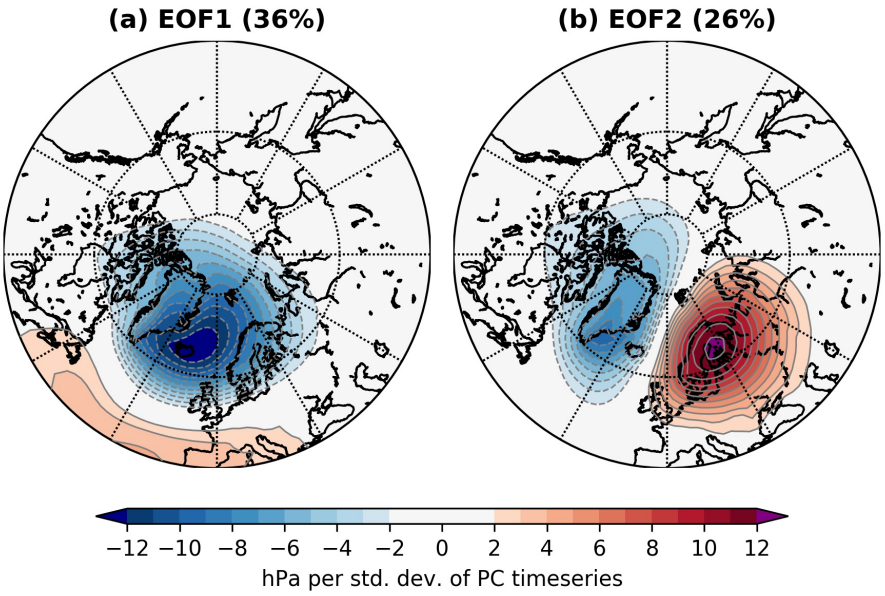


**Figure 3.1:** The first two EOFs of November–March 1980-2019 ERA5 MSLP anomalies in the region 60-85°N, 60°W-50°E, expressed as a linear regression between the principal component timeseries (scaled by the square-root of the eigenvalues) and MSLP anomalies at each grid point. The number in parentheses indicates the percentage of total variance within the analysis region explained by the EOF.

The second EOF (Figure 3.1b), explaining 26% of the variance, consists of a zonal dipole structure, defined here such that a positive loading has an anticyclonic anomaly over Scandinavia and a cyclonic anomaly extending over most of Greenland. This pattern shares strong similarities with the S-G dipole of L19, the Greenland-Scandinavia cluster of Cassou et al. (2004), and to a lesser extent the Scandinavian blocking regime (Cassou, 2008) and the Ural blocking anomaly of Peings (2019). It also closely resembles the MSLP anomalies associated with anticyclonic wave breaking and block onset near 20°E in Masato et al. (2012) (their Figure 5f). Hereafter, we refer to the principal component timeseries of the second EOF as the "S-G index", and the *positive* loading of this second EOF as the "S-G pattern". The S-G index is more transient than the zonal pattern with an autocorrelation *e*-folding timescale of 4 days, consistent with the relatively short timescale of anticyclonic wave breaking identified in L19.

### 3.3.2 Stratospheric relationship

To establish the observed relationship between the S-G pattern and the SPV, we perform lagged linear regression between the S-G index and zonal-mean eddy heat flux and polar cap geopotential height anomalies (Figure 3.2), and the amplitude anomalies of wavenumbers 1 and 2 at 60°N (Figure 3.3). Composite-based analyses (not shown) yield very similar results, confirming the suitability of the linear regression approach. At negative lags of 10-20 days, the S-G pattern is associated with a strong SPV precursor, consistent with the Greenland trough (i.e., the absence of Greenland blocking) present in the S-G pattern which is more likely during a strengthened SPV (e.g. Charlton-Perez et al., 2018). The significantly enhanced heat flux in the stratosphere during this time may be explained by sharpened potential vorticity gradients on the edge of the strengthened SPV acting as a waveguide (e.g. Scott et al., 2004). In agreement with the threshold event-based composite results of L19, on short lags, the S-G pattern is associated with a transient period of anomalously enhanced heat flux (Figure 3.2a), forming a coherent pulse from the troposphere to the upper-stratosphere on a timescale of ∼5 days. Wavenumber-2 is significantly amplified throughout the column during this time (Figure 3.3b), in agreement with the coherence between the S-G pattern and the climatological-mean eddy height field (c.f. Figure 3.4c). There is a concomitant abrupt development of a weak vortex anomaly, which persists and descends through the stratosphere over the following ∼2 months. These results further motivate analysis of the S-G pattern as a significant contributor to SPV variability across the S2S timescale, beyond just the major SSWs considered in L19.

The relationship with eddy heat flux (Figure 3.2a) remains significant and positive for lags of several weeks, with evidence of a secondary peak in the upper-troposphere lower-stratosphere at lags of ∼25-30 days during a period of wavenumber-1 amplification (wavenumber-2 is anomalously suppressed during this time). The magnitude of the weak vortex anomaly intensifies in the middle and lower stratosphere following this second peak (accordant with the relationship between *cumulative* heat flux and SPV strength (Polvani and Waugh, 2004)). There is little evidence of downward coupling of the weak vortex

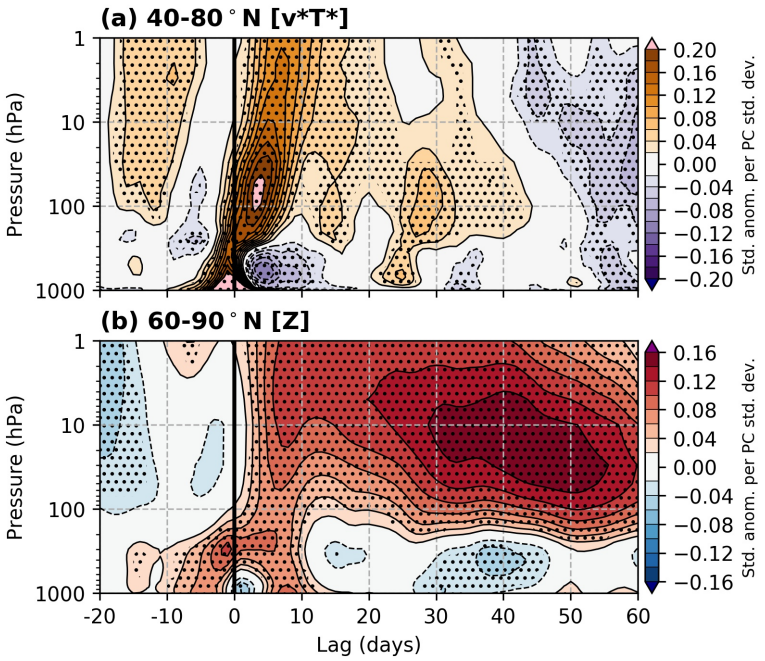


**Figure 3.2:** Lagged linear regression between the S-G index and standardised anomalies of zonally-averaged (a) 40-80°N eddy heat flux (denoted [v*T*]) and (b) 60-90°N geopotential height (denoted [Z]) for NDJFM 1980-2019 in ERA5. Stippling indicates significance at the 95% confidence level according to a bootstrap re-sampling procedure.

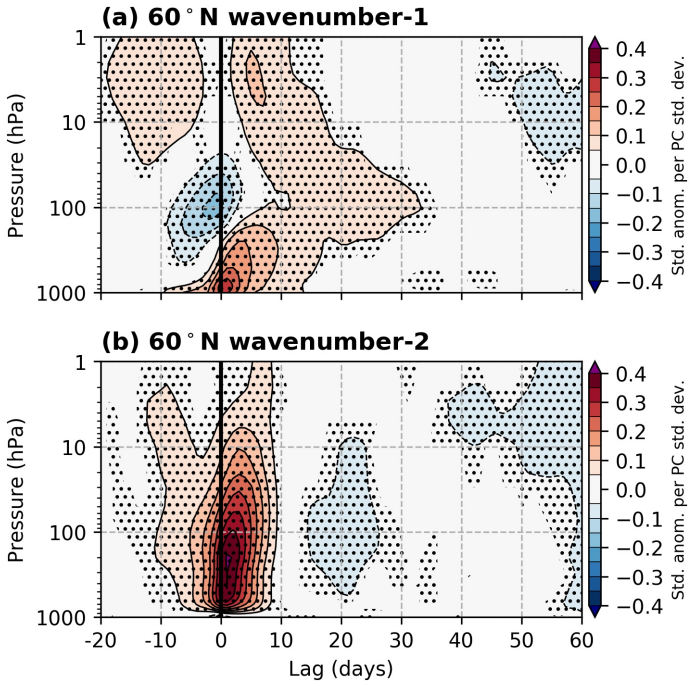


**Figure 3.3:** Lagged linear regression between the S-G index and standardised anomalies of the amplitude anomaly of (a) wavenumber-1 and (b) wavenumber-2 at 60°N for NDJFM 1980-2019 in ERA5. Stippling indicates significance at the 95% confidence level according to a bootstrap re-sampling procedure.

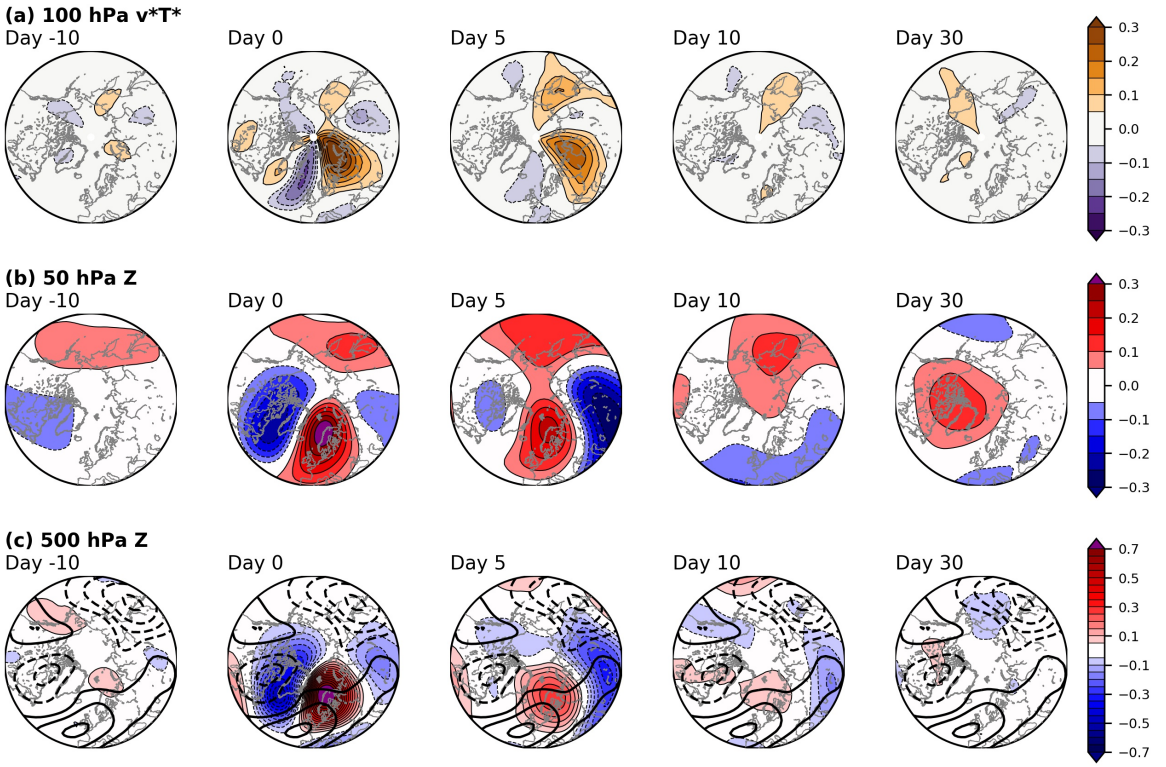


**Figure 3.4:** Linear regression at various lags between the S-G index and standardised anomalies of (a) 100 hPa meridional eddy heat flux (v*T*), (b) 50 hPa and (c) 500 hPa geopotential height (Z) for NDJFM 1980-2019 in ERA5. In (c) the thick contours show the mean 500 hPa eddy height field (50 m intervals between -200 and 200 m, excluding the 0 contour). Units are standardised anomaly per standard deviation of the PC timeseries. All contoured anomalies are significant at the 95% confidence level according to a bootstrap re-sampling procedure.

anomaly into the troposphere, although it is possible this is due to the intrinsically larger tropospheric variability which may not be captured in a lagged linear regression approach. However, downward coupling is apparent when only the S2S common period is used, despite otherwise similar results (Figure 3.14). Fully determining a cause of these differences is beyond the scope of the study.

To provide further insight into the dynamics introduced in the cross-section analysis, polar-stereographic maps of lagged regressions with standardised anomalies of (a) 100 hPa meridional eddy heat flux and (b) 50 hPa and (c) 500 hPa geopotential height are shown in Figure 3.4, alongside the mean 500 hPa eddy height field as a diagnostic of the climatological stationary waves. Eddy heat flux is anomalously amplified at 100 hPa in a sector north and east of Scandinavia (days 0 and 5), and subsequently downstream over

eastern Asia (day 5), consistent with the constructive interference between the S-G pattern and the mean eddy height field in the North Atlantic sector. The 50 hPa geopotential height anomalies form a clear wavenumber-2 pattern, reminiscent of an SPV split event, on days 0 and 5 with positive height anomalies intersecting negative height anomalies across the pole. Subsequently, the wave field decays, while positive geopotential height anomalies remain over the central Arctic indicative of a significantly weakened SPV. The aforementioned amplification of wavenumber-1 and secondary heat flux pulse at around 30 days (c.f. Figures 3.2a and 3.3a) can be seen to emanate from the northeast Pacific, downstream of an amplified tropospheric Aleutian low.

## 3.4 S-G pattern in S2S models

### 3.4.1 Representation of EOFs

In order to assess the ability of S2S models to replicate the observed patterns of variability, we repeat the same EOF analysis as performed in ERA5 but in the 10 S2S model hindcasts, where the EOF is computed across all ensemble members at each lead-time (to assess lead-time dependent biases). The first two EOFs of all S2S models have very similar structures to the equivalent observed EOFs; weekly-mean pattern correlations do not drop below 0.7 for either EOF (see Figure 3.15). The lower pattern correlations are for EOF1; visual inspection reveals this is likely due to slight shifts in the centre of the Icelandic cyclonic anomaly. For EOF2, the pattern correlations exceed 0.95. Thus, we can be confident that the S2S models replicate well the structure of the variability and that the model EOFs can be directly compared with those from ERA5. For brevity, the week 4 mean EOF patterns for all models are shown in Figure 3.5.

We quantify variability biases in the S2S models by computing the ratio between the explained variance fraction of the model EOFs with the equivalent ERA5 EOF (Figure 3.6). *Generally*, these S2S models have more than the observed variance fraction in EOF1, and

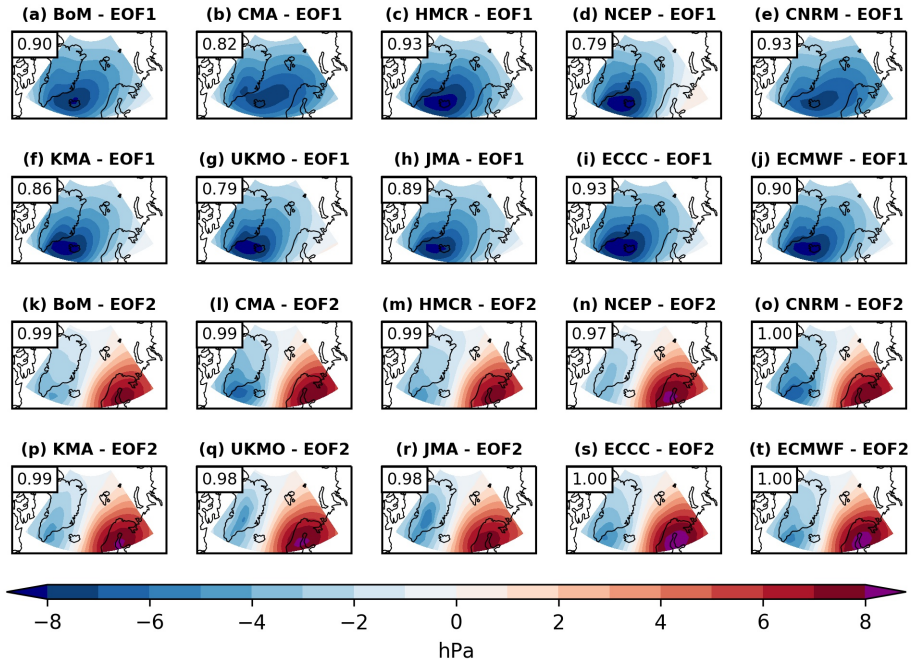


**Figure 3.5:** The average week 4 (forecast days 21–27) pattern of the (a–j) first and (k–t) second EOFs of NDJFM MSLP anomalies in 10 S2S models (multiplied by the square-root of the eigenvalue; units hPa). The models are sorted by horizontal resolution (lowest–highest). The area-weighted pattern correlation with the equivalent ERA5 EOF is also shown in the top-left corner of each panel.

correspondingly less in EOF2, in common with generic model biases in wave breaking and blocking. No model has statistically more than observed variance fraction in EOF2. The biases are especially large in weeks 3 and 4, and are largest for HMCR, CMA, and BoM, with 20-30% more than the observed variance in EOF1 and correspondingly less in EOF2. HMCR exhibits large variance biases even in the first forecast week. These three models have the lowest horizontal resolution ($>1^{\circ}$), as well as the lowest vertical resolution and lowest lid height. The statistics here indicate no dependence on ensemble size or initialisation frequency. Note that biases are small in NCEP despite it having the fourth lowest horizontal resolution and being the oldest model version used in our study – suggesting that multiple factors likely contribute to these biases.

We also considered the ratio of the *total* variance in the models with that in ERA5 (see Figure 3.16). In CMA, the total variance is close to that in ERA5, while in BoM and HMCR the total variability declines with lead-time to ∼75% of ERA5. Interpreting

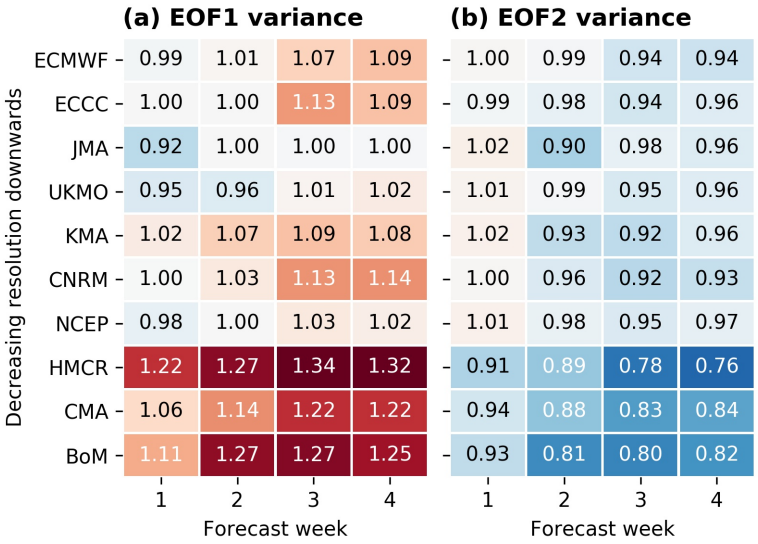


**Figure 3.6:** Weekly-mean ratio of explained variance between the model EOFs and the equivalent ERA5 EOFs for initialisations in the period NDJFM 2000–2010. The models are sorted by horizontal resolution.

these statistics in a physically meaningful sense is more challenging, but suggests that the characteristics of the biases in HMCR and BoM are different to that in CMA. We also note that other models, in particular ECCC, UKMO and CNRM, have slightly more total variability than ERA5. Despite these variance biases, the persistence of the patterns (as measured by the autocorrelation *e*-folding timescale of the PC timeseries) in all models is not significantly different to that in ERA5 for either EOF at all forecast days.

### 3.4.2 Predictability

Next, we consider the ability of the models to accurately predict the evolution of the observed EOFs. We assess this by first projecting the NDJFM 2000-2010 EOFs computed from ERA5 onto the model MSLP anomalies for each ensemble member, to generate a forecast PC timeseries. As an initial analysis of deterministic skill, Figure 3.7 shows ensemble-mean correlation skill, defined as the first day when correlation drops below 0.6, for both the zonal and S-G patterns. For all models and for both patterns, the limit of ensemble-mean correlation skill lies within 15 days. There is an indication of 1-2 days of additional skill for the zonal pattern, in agreement with its greater de-correlation timescale though this

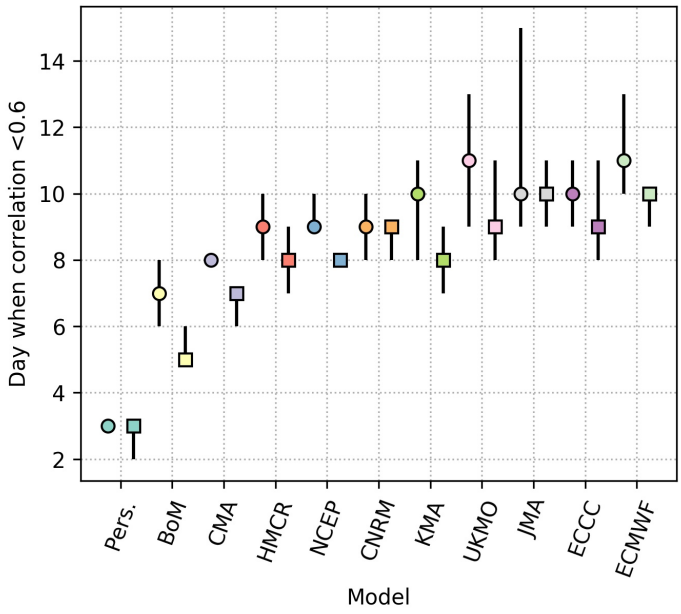


**Figure 3.7:** The first day when the ensemble-mean correlation with ERA5 drops below 0.6, for the zonal pattern (EOF1, circle markers) and S-G pattern (EOF2, square markers). Error bars indicate 95% confidence intervals from 5,000 bootstrap re-samples, where these correspond to the first day of skill below 0.6 in the 2.5 and 97.5 percentiles of the re-sampled correlations. The models are sorted left-to-right by increasing horizontal resolution. A persistence forecast is also shown for reference.

difference is only significant for CMA and NCEP. There is again evidence of resolution dependence (although minimal for models with higher resolution than CMA) and the impact of the variability biases discussed in the preceding subsection. BoM exhibits particularly poor performance, with only 5-6 days of skill for the S-G pattern and 6-8 days for the zonal pattern. In contrast, ECMWF has 9-10 days of skill in the S-G pattern and 10-13 days in the zonal pattern.

To assess the general performance of the ensemble systems on the subseasonal timescales, Figure 3.8 shows the receiver operating characteristic skill score (ROCSS) (e.g. Wilks, 2019) for a 1 $\sigma$ threshold (where the sigma threshold is set by the ERA5 EOF) for both the zonal and S-G patterns. Similar results are obtained for various positive thresholds. The ROCSS is the additional area under the ROC curve (i.e., a plot of true positive rate (TPR) versus false positive rate (FPR)) versus a non-skilled forecast (TPR = FPR). Here, a true positive is counted as $m$ ensemble members correctly predicting the S-G/zonal pattern index exceeding the threshold in the target week (and vice versa), where $m$ is varied from

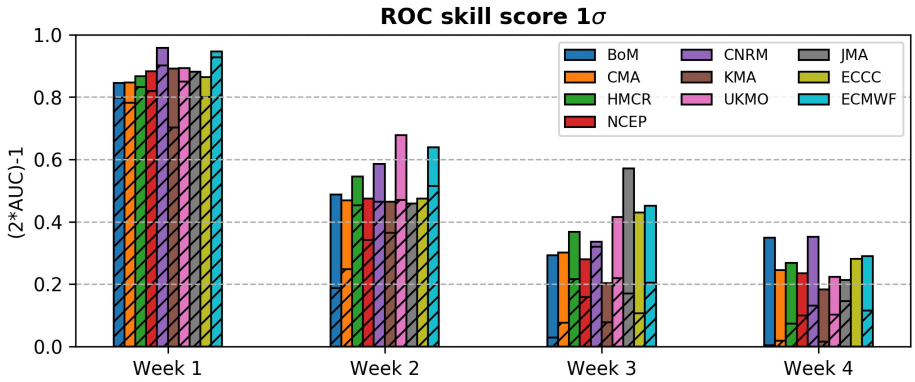


**Figure 3.8:** ROC skill score for $\geq 1\sigma$ zonal pattern (solid colours) and S-G pattern (hatched) events. The models are sorted left-to-right by increasing horizontal resolution.

0 to the full ensemble size. The ROCSS for *all* models at *all* lead-times is higher for the zonal pattern than for the S-G pattern, increasing to 50-100% larger at weeks 3 and 4, consistent with the biases introduced in Figure 3.6. Skill in the S-G pattern for BoM and CMA is appreciably lower than other models by week 2, but at weeks 3 and 4 all models have only small skill. BoM, CMA, KMA, and ECCC have negligible skill in the S-G pattern at week 4, despite comparable skill in the zonal pattern to other models. We also performed the same analysis but for the lead-time dependent 1 $\sigma$ threshold in each model, obtaining similar results.

## 3.5 S-G–Stratosphere relationship in S2S models

In this section, we assess the ability of S2S models to capture the observed relationship (established in Section 3.2) between the S-G pattern and both enhanced heat flux and a weakened SPV in the following weeks to months. First, the ERA5 EOF is projected onto the MSLP anomalies in each ensemble member, and then lagged linear regression is performed across all ensemble members and initialisation dates. Figure 3.9 shows the lagged regression between the S-G pattern on the first day of the forecasts with the subsequent 40-80°N-averaged eddy heat flux anomalies, and Figure 3.10 shows the equivalent for polar cap geopotential height anomalies. All 10 models capture the immediate relationship with enhanced eddy heat flux from the troposphere to the stratosphere within the first week, and all but the low-topped HMCR model capture the subsequently weakened SPV (c.f.

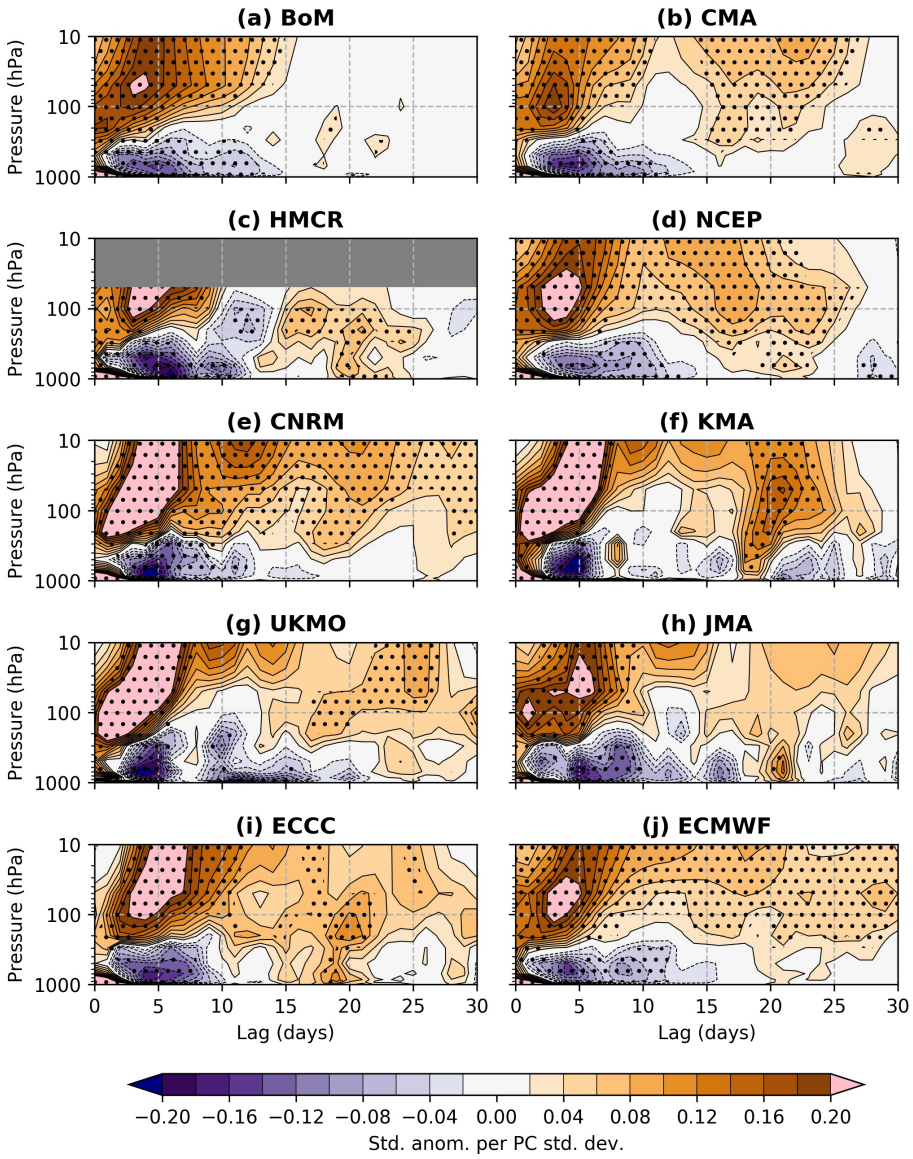


**Figure 3.9:** Linear regression between the S-G index in the first day of the forecast and standardised anomalies of 40-80°N eddy heat flux for the subsequent 30 forecast days in hindcasts from 10 S2S models in NDJFM 2000-2010. The linear regression is carried out across all ensemble members. Stippling indicates significance at the 95% confidence level according to a bootstrap re-sampling procedure. The contour scale is chosen to match that in Figure 3.2.

Figure 3.2) for the remainder of the forecast period. However, there are subtle inter-model differences and differences with the observed relationship, though we emphasise that the differences in sample sizes make direct inter-comparison of magnitudes more challenging.

BoM captures the initial pulse of enhanced heat flux (Figure 3.9a), but there is no evidence of significantly enhanced heat flux from the troposphere to the stratosphere beyond 15

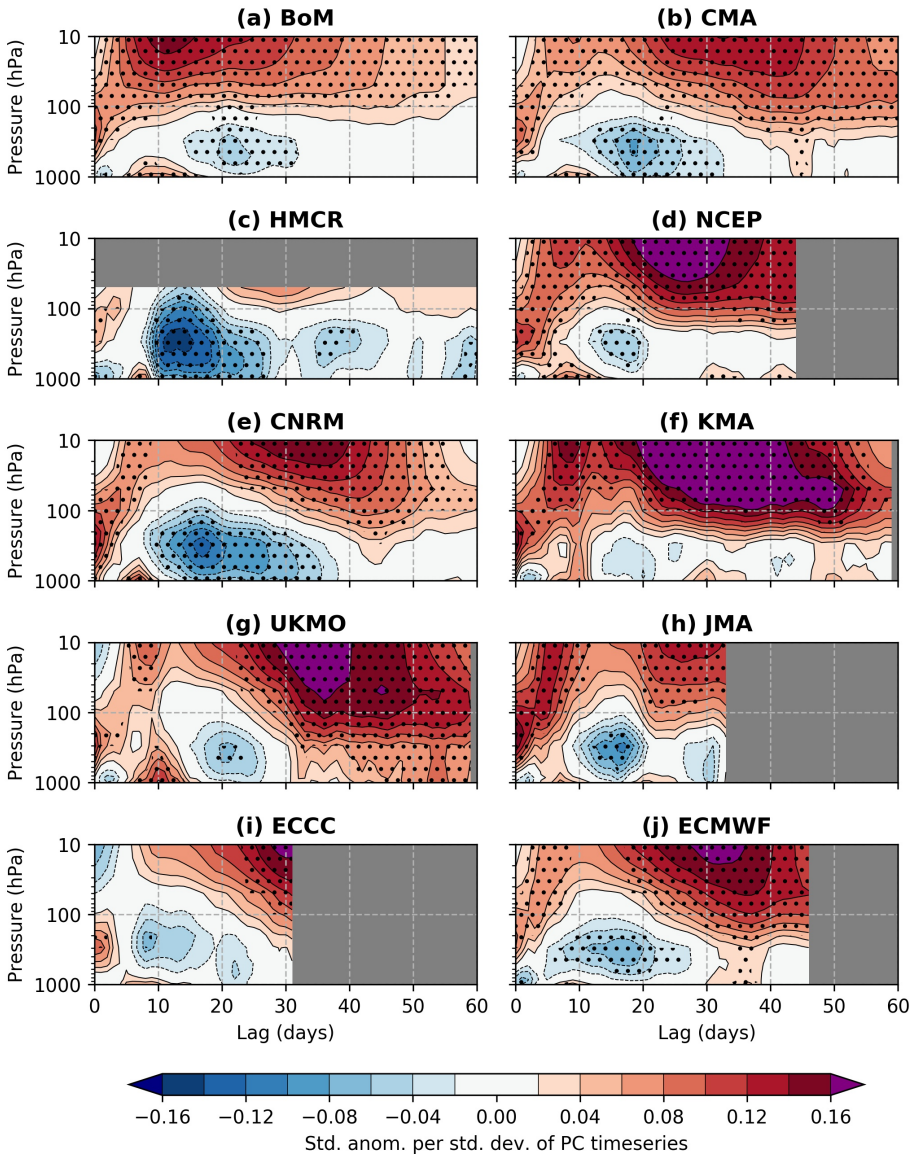


**Figure 3.10:** Linear regression between the S-G index in the first day of the forecast and standardised anomalies of 60-90°N geopotential height for the remaining forecast days in hindcasts from 10 S2S models in NDJFM 2000-2010. The linear regression is carried out across all ensemble members. Stippling indicates significance at the 95% confidence level according to a bootstrap re-sampling procedure. The contour scale is chosen to match that in Figure 3.2.

days – the second pulse (wavenumber-1) at 20-30 days is not present. Accordingly, the weak SPV anomaly (Figure 3.10a) peaks shortly after 10 days and slowly decays without evidence of amplification and downward propagation. Nevertheless, the signal for a weakened SPV over the next 60 days remains, in agreement with observations (c.f. Figure 3.2a). Overall, the structure of the heat flux evolution is closer to observations for the other models, though there is no significantly raised heat flux beyond 10 days in JMA

(Figure 3.9h) and the pattern is rather diffuse at longer lead-times in ECMWF (Figure 3.9j). Notably, the second period of significantly enhanced heat flux is captured even in the low-top HMCR model (Figure 3.9c), supportive of a mostly tropospheric-led mechanism.

Aside from BoM and HMCR, the significantly weakened SPV maximises in all models at a similar time as in observations, though with varying magnitudes. The initially significantly weakened SPV is missing from ECCC (Figure 3.10i), while the lower-stratospheric anomalies in JMA (panel (h)) around day 30 are insignificant, in contrast to other models and ERA5. The response in the troposphere is particularly varied between the models; the development of negative polar cap height anomalies seen in ERA5 is significantly present in all except KMA and ECCC, again with varying magnitudes and timings. In ERA5 this occurs most strongly at 30-40 days, while it is 10-15 days earlier in these models – consistent with a zonally-biased state or reduced blocking persistence. In UKMO, and to a lesser extent ECMWF and CMA, there is significant downward coupling into the troposphere from the stratosphere after 30 days. While ERA5 indicates some coupling (especially in the S2S common period; see Figure 3.14), the maximum in these models is both larger and earlier – occurring when ERA5 shows *negative* polar cap anomalies. We also performed the same analysis as shown in Figures 3.9 and 3.10 but at forecast day 15, as a measure of the ability of the model to internally capture the relationship once it has drifted toward its own climatological state (while retaining sufficient subsequent forecast days to assess the lagged response). The results are much weaker (shown in Figures 3.17 and 3.18); the heat flux pulse remains present, but is systematically weaker, and there is an accordingly weaker signal in the SPV strength which is not significant in KMA, ECCC or JMA (though the latter two are limited by their shorter forecast ranges).

To further this analysis, Figure 3.11 shows the weekly-mean regression coefficients across all ensemble members for eddy heat flux at (a) 300 hPa on the same day (i.e., the tropospheric wave activity associated with the S-G pattern, c.f. Figure 3.2), (b) 100 hPa at a 3-day lag, and (c) 50 hPa at a 4-day lag (where the lags correspond to maximum correlation with the

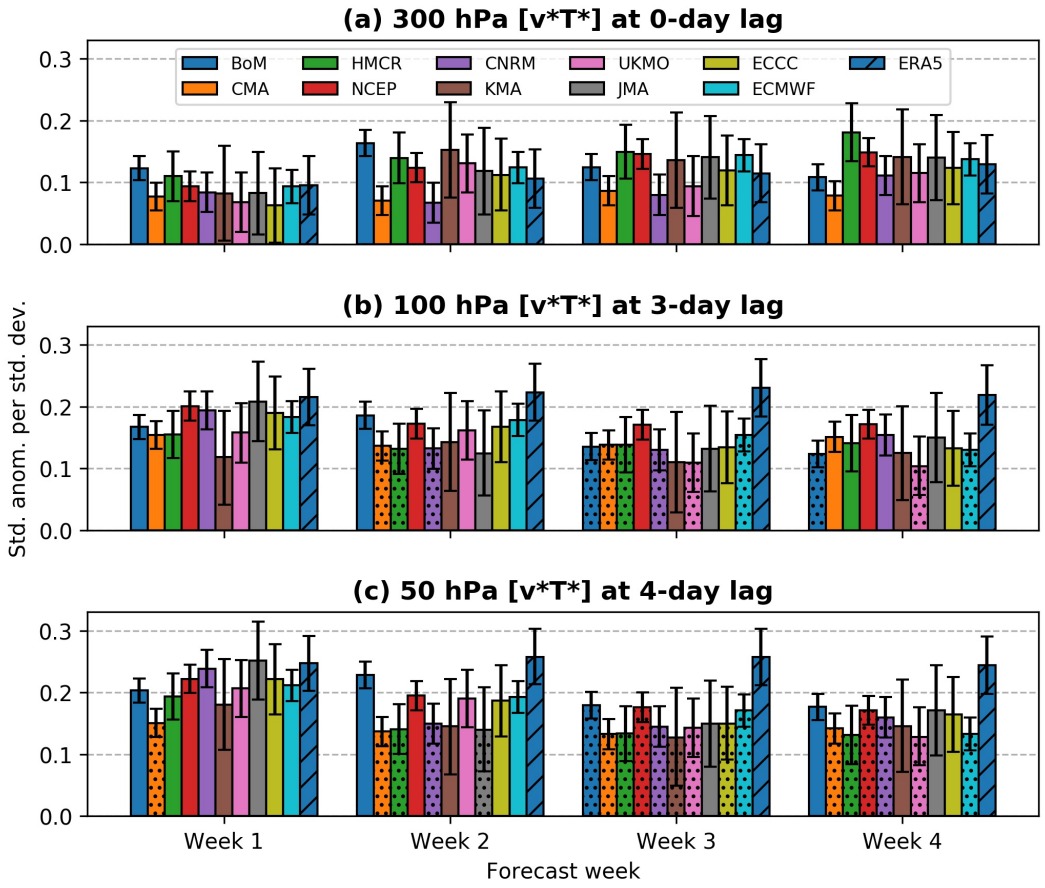


**Figure 3.11:** Weekly-mean regression coefficients between the S-G index in each ensemble member and the corresponding eddy heat flux anomalies at (a) 300 hPa on the same day, (b) 100 hPa 3 days later, and (c) 50 hPa 4 days later. The lags correspond to days with maximum correlation in ERA5. Equivalent ERA5 statistics are also shown. The models are sorted left-to-right by increasing horizontal resolution. Error bars indicate 95% confidence intervals according to a bootstrap re-sampling procedure; stippled bars indicate where these do not overlap with those of ERA5.

S-G pattern in ERA5). At 300 hPa, there is no significant difference between the regression coefficients for the models or ERA5 for any forecast week. The (statistically insignificant) increase from week 1 to week 4 may be attributable to the seasonal cycle. The agreement between the models and observations at all lead-times indicates that, in the troposphere, the influence of the S-G pattern on zonal-mean wave activity is well-represented. However the representation of the relationship with vertically-propagating wave activity in the stratosphere at 100 hPa and 50 hPa is largely lead-time dependent, with much weaker regression coefficients particularly in weeks 3 and 4 (up to 50% smaller for HMCR, UKMO and ECMWF) – though CMA is significantly weaker at all lead-times and does not decay over time. These results agree with the weaker magnitude of the SPV anomalies in Figure

3.18, providing evidence that the communication of wave activity from the troposphere to the stratosphere produced by the S-G pattern is weaker at longer lead-times, thereby generating weaker circulation anomalies. This bias may therefore further preclude subseasonal stratospheric predictability, even in the case of a well-forecast troposphere. Note that, at weeks 3 and 4, the bias is systematic across the S2S hindcasts analysed here and does not seem dependent on the resolution of the models.

## 3.6 Summary and Conclusions

In this paper, we have performed an analysis of observed (ERA5) and modelled (S2S hindcasts) wintertime tropospheric variability in a region of the northeast Atlantic associated with the exit region of the eddy-driven jet, and its relationship with vertically-propagating wave activity and the strength of the SPV. We find that the first two EOFs of MSLP anomalies describe the majority of the variability within the analysis domain. The leading mode represents a zonal pattern akin to the NAO, while the second mode depicts the S-G pattern – characterised by an anticyclonic anomaly over Scandinavia and an anomalous trough over Greenland (Figure 3.1). The S-G pattern resembles various patterns previously identified to be associated with anticyclonic wave breaking, blocking, and influences on SPV variability. We find that the S-G pattern is associated with transient amplification of wavenumber-2 and anomalously enhanced eddy heat flux into the stratosphere, with a weakened SPV following and persisting for next 2 months (Figure 3.2). The long timescale of the relationship with the SPV strength supports the importance of representing the S-G pattern – and more generally, tropospheric blocking and Rossby wave breaking – for S2S prediction.

In the 10 models from the S2S database analysed here, the structure of the two EOFs is represented well at all lead-times (Figure 3.5). However, the lowest resolution models (namely: BoM, CMA, and HMCR) exhibit a large zonal variability bias which grows with lead-time (Figure 3.6). These models have more than the observed variance fraction in the first EOF (the zonal pattern) and a proportional reduction in the second EOF (the S-G

pattern). Although the aforementioned 3 models have the largest biases, all 10 models have slightly less than the observed variance fraction in the second EOF, consistent with extensive literature on the under-representation of blocking and wave breaking in most forecast models. Our finding of the largest variability biases in BoM, CMA, and HMCR agrees well with the relative magnitudes of the biases in Rossby wave packet decay in the northeast Atlantic in Quinting and Vitart (2019) (their Figure 2), further supporting our physical interpretation of the EOFs.

We find all models have more skill in predicting the zonal pattern versus the S-G pattern, especially in weeks 3 and 4 (Figures 3.7 and 3.8). Ensemble-mean correlation skill is limited to well within 2 weeks (with a maximum of 10 days in the S-G pattern in several models, but as low as 5 days in BoM), and there are also indications of resolution dependence. Considering the link between the S-G pattern and the SPV, the timescale of correlation skill in these models (and inter-model differences) is similar to the timescale of SSW prediction (e.g. Domeisen et al., 2020b), while the very limited ROC skill in the subseasonal range is further supportive of a limitation on subseasonal stratospheric skill arising from poorly-predicted tropospheric processes. Moreover, our results indicate that the poor longer-term predictability of the S-G event preceding the February 2018 SSW (as described in L19) is not unique to that case.

The S2S models represent well the stratospheric influence of the S-G pattern when considering the initial conditions (Figures 3.9 and 3.10), though the amplitude and persistence of the SPV anomalies are weaker in lower-topped models. However, we find that the relationship between the S-G pattern and enhanced eddy heat flux into the stratosphere (and a subsequently weakened SPV) is much weaker at longer lead-times, with no clear dependence on resolution or lid height. We find evidence (Figure 3.11) that the weakening of the relationship is due to reduced attendant wave activity flux into the stratosphere (up to 50% weaker in week 4 versus ERA5 in HMCR, ECMWF and UKMO), as the tropospheric wave activity remains similar to that observed. It is possible that

lead-time dependent biases in the modelled stratospheric mean state, such as biases in the zonal winds, alter the vertical propagation or subsequent amplification of anomalous wave activity from the troposphere (perhaps arising from the interaction with the mean stationary waves (Nishii et al., 2009)). The significantly weaker regression coefficients in week 4 in the higher resolution ECMWF and UKMO models is particularly notable given the cold SPV bias in those models (Son et al., 2020).

Our results, while limited by a multi-model approach, support the importance of higher model resolution on the representation of wave breaking and blocking within the exit region of the North Atlantic eddy-driven jet. We suggest that a contribution to these biases, especially their lead-time dependent nature, arises from the decline in tropopause sharpness seen in numerical weather prediction models at longer lead-times, and its impact on Rossby wave propagation (Gray et al., 2014; Saffin et al., 2017). Future work to ascertain the relative importance of resolution and the representation of other processes (such as diabatic effects) in accurately modelling the variability in the northeast Atlantic may help address this hypothesis. Furthermore, these biases may be subsequently manifest in the downward high-latitude blocking response of the troposphere to a weakened SPV (i.e., the onset of a negative NAM) which is poorly represented in S2S models (e.g. Figure 7 in Domeisen et al. (2020c)). We note that the *negative* loading of the S-G pattern corresponds to Greenland blocking and a Scandinavian trough anomaly, and the negative loading of the zonal pattern corresponds to a negative NAO – both of which are broadly consistent with surface responses to a significantly weakened SPV or major SSW (e.g. Butler et al., 2017). Thus, while the focus of this study has been on the *positive* S-G pattern as a source of SPV weakening, the S2S model biases in variability and predictability apply to anomalies of both signs, and therefore are likely partly related to poor S2S skill in the response to SSWs.

In conclusion, the combination of zonal biases, limited sub-seasonal skill in the S-G pattern, and poor representation at longer lead-times of its subsequent impact on the SPV, is likely

to be a contributing factor to the still limited skill in predicting stratosphere-troposphere coupling on S2S timescales. A targeted approach to determining the representation within S2S models of further key tropospheric processes known to influence the SPV, such as western Pacific bomb cyclones (Attard and Lang, 2019), may illuminate additional regions where stratospheric S2S skill could be gained.

## Acknowledgements

S. H. L. acknowledges funding by the Natural Environment Research Council (NERC) via the SCENARIO Doctoral Training Partnership (NE/L002566/1). S. J. W. was supported by the National Centre for Atmospheric Science, a NERC collaborative center, under Contract R8/H12/83/001. EOF analysis was performed in Python using the package *eofs* (Dawson, 2016). This work is based on S2S data, a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP). The S2S prediction project database is available online at `https://apps.ecmwf.int/datasets/s2s`. The ERA5 reanalysis is available from the Copernicus Climate Data Store at `https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels`. We thank two anonymous reviewers for their helpful suggestions.

## 3.7 Supporting Information

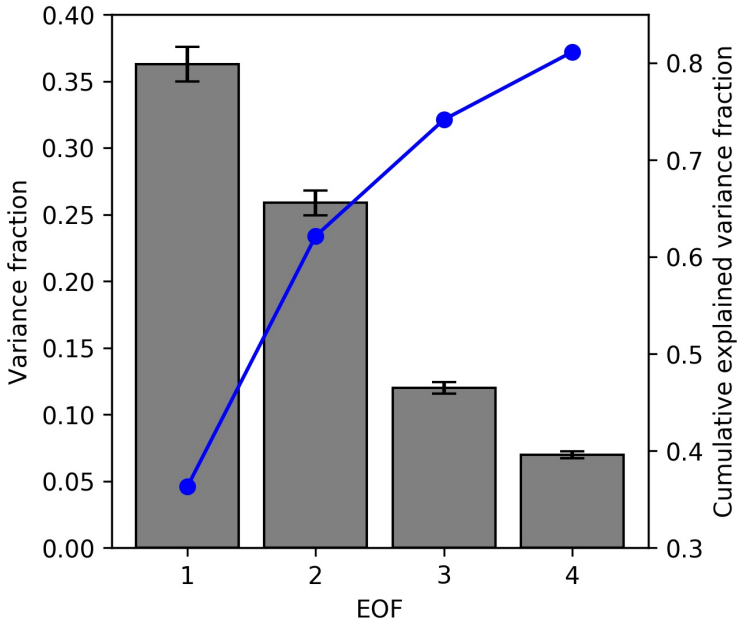


**Figure 3.12:** Explained variance fraction (bars, left-hand y-axis) and cumulative explained variance fraction (line, right-hand y-axis) for the first 4 EOFs in ERA5 November–March 1979-2019. Error bars indicate 95% confidence intervals according to North et al. (1982).

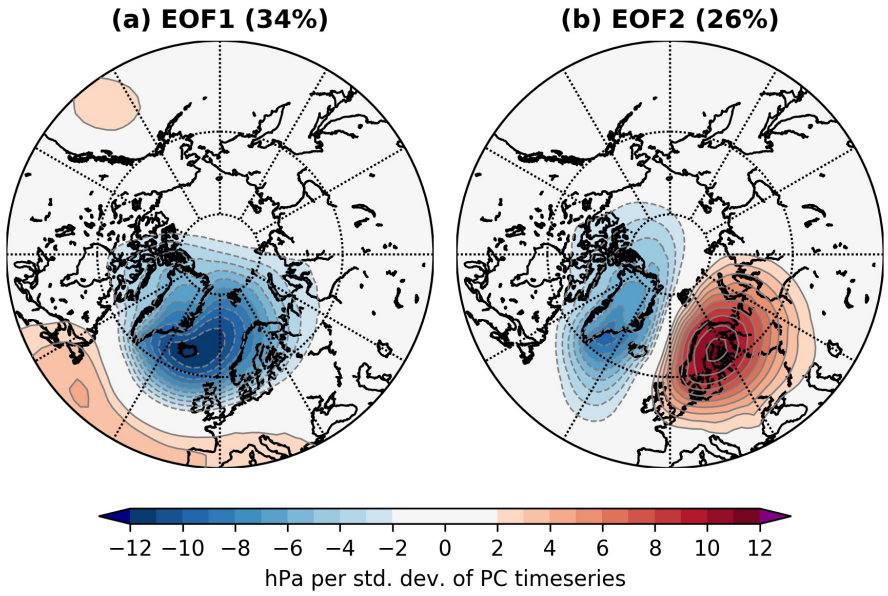


**Figure 3.13:** As in Figure 3.1 in the main text, but for the S2S common hindcast period (November–March 1999–2010).

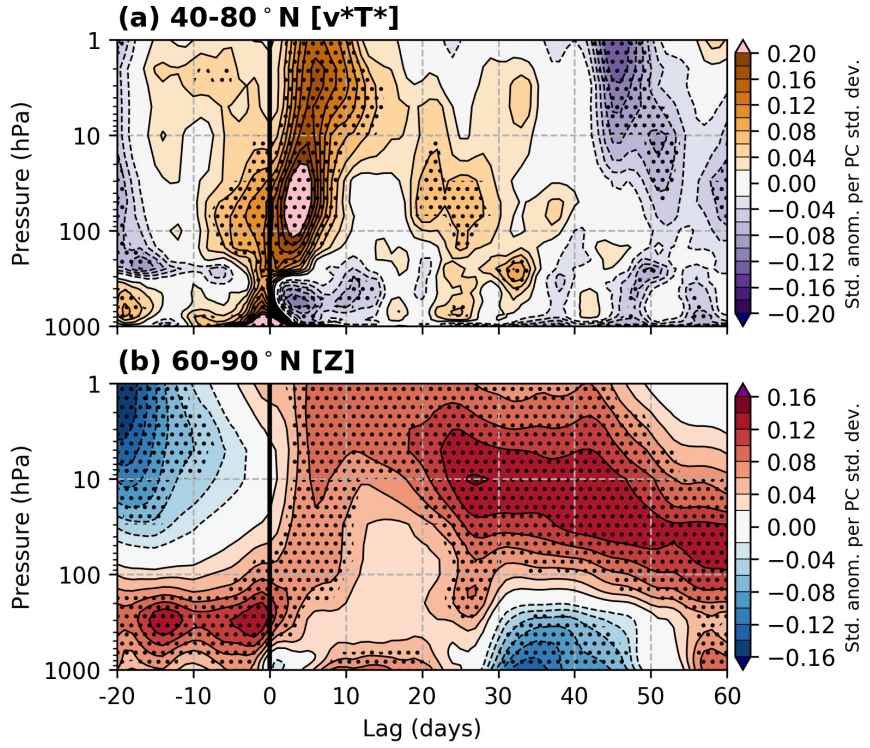


**Figure 3.14:** As in Figure 3.2 in the main text, but for November–March 1999–2010 (where anomalies are computed with respect to this period). Stippling indicates significance at the 95% confidence level according to a bootstrap re-sampling test (see main text for details).

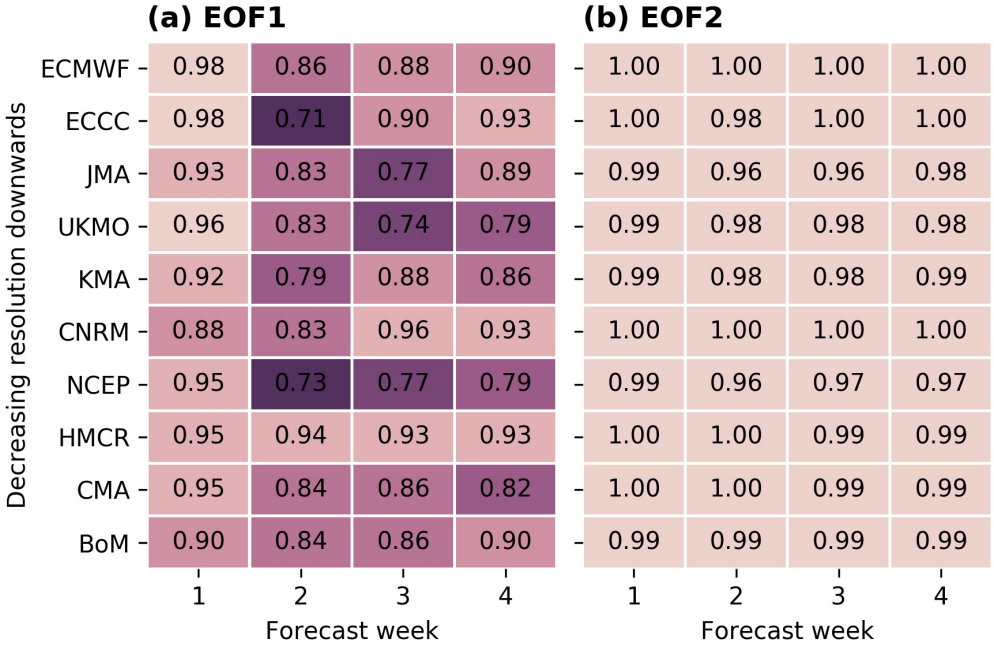


**Figure 3.15:** Weekly-mean area-weighted (cosine-latitude) Pearson's correlation coefficient between the S2S model EOFs and the equivalent ERA5 EOFs.

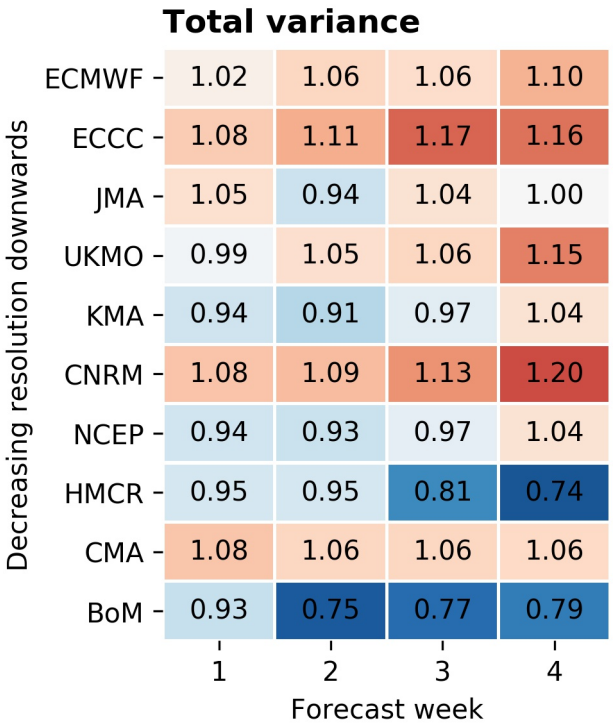


**Figure 3.16:** Ratio between the weekly-mean total variance (sum of eigenvalues) in the S2S models and the equivalent in ERA5.

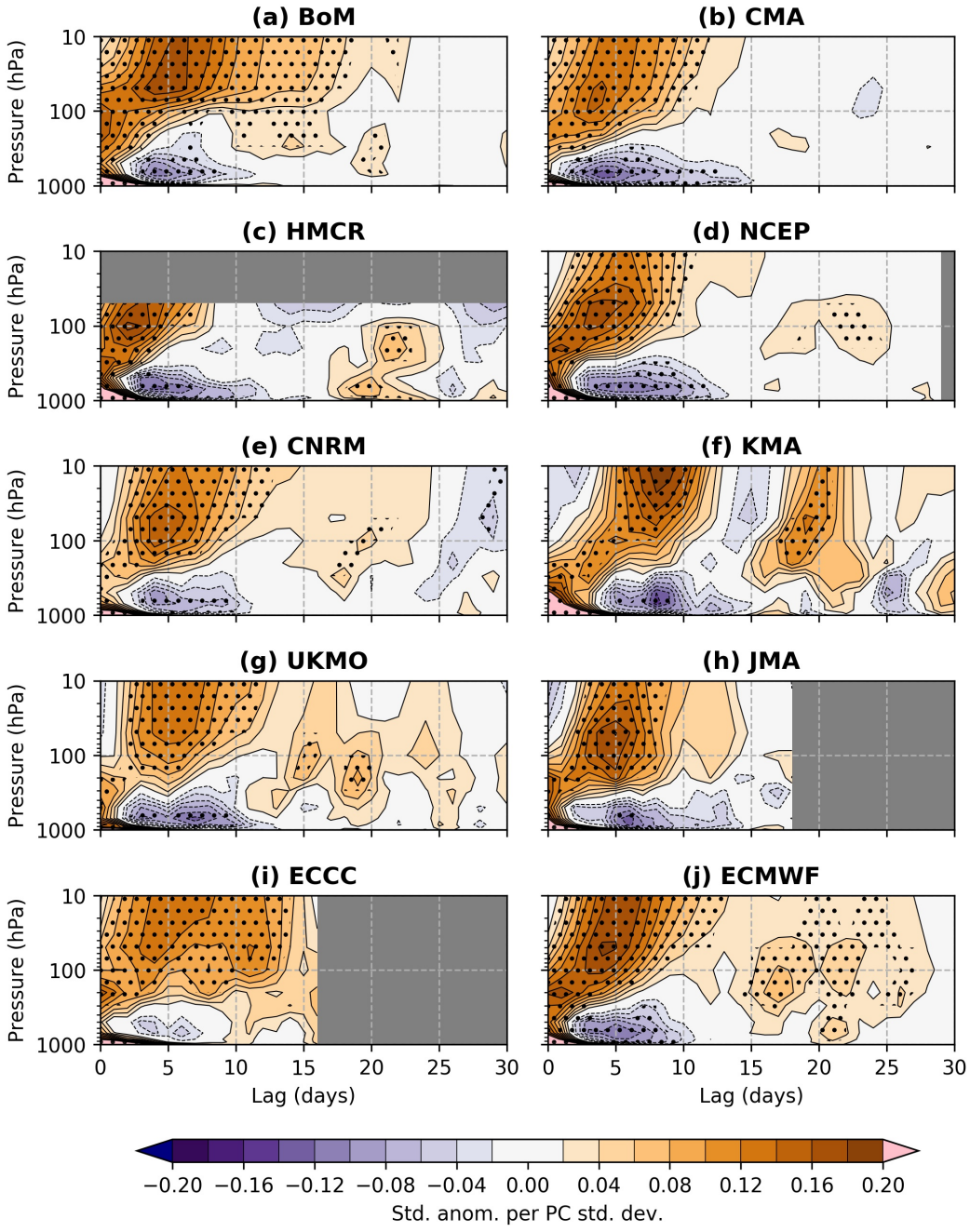


**Figure 3.17:** As in Figure 3.9 in the main text but for the S-G index on forecast day 15.

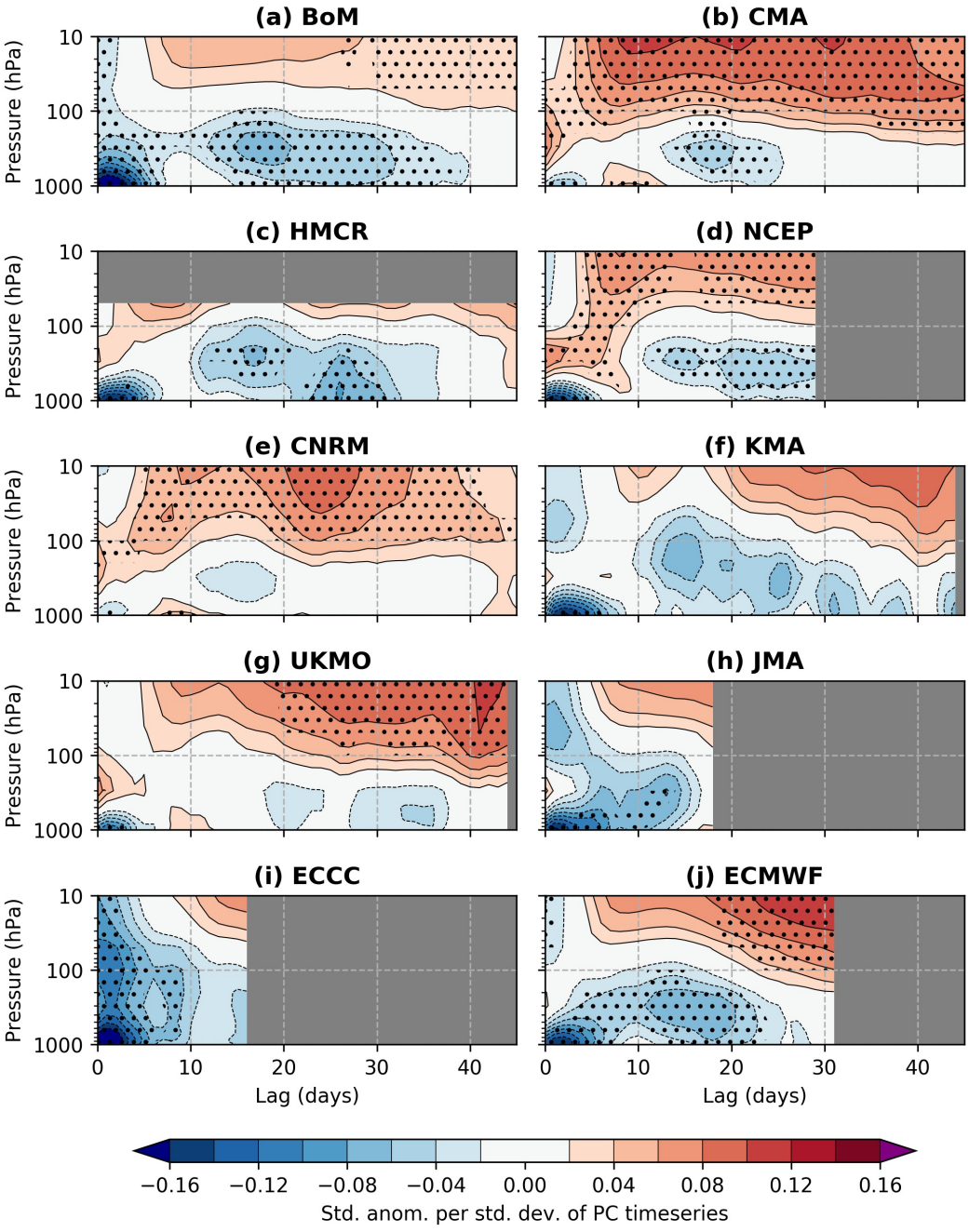


**Figure 3.18:** As in Figure 3.10 in the main text but for the S-G index on forecast day 15.

# Chapter 4

# Wintertime North American Weather Regimes and the Arctic Stratospheric Polar Vortex

This chapter has been published in *Geophysical Research Letters* with the following reference:

**Lee, S. H.**, J. C. Furtado, and A. J. Charlton-Perez, 2019: Wintertime North American Weather Regimes and the Arctic Stratospheric Polar Vortex. *Geophysical Research Letters*, **46**, 14892-14900, https://doi.org/10.1029/2019GL085592.

# Abstract

The impact of the Arctic stratospheric polar vortex on persistent weather regimes over North America is so far under-explored. Here we show the relationship between four wintertime North American weather regimes and the stratospheric vortex strength using reanalysis data. We find that the strength of the vortex significantly affects the behavior of the regimes. Whilst a regime associated with Greenland blocking is strongly favored following weak vortex events, it is not the primary regime associated with a widespread, elevated risk of extreme cold in North America. Instead, we find that the regime most strongly associated with widespread extremely cold weather does not show a strong dependency on the strength of the lower-stratospheric zonal-mean zonal winds. We also suggest that stratospheric vortex morphology may be particularly important for cold air outbreaks during this regime.


## 4.1 Introduction

The behavior of the stratospheric polar vortex (SPV) is known to influence wintertime tropospheric weather patterns on subseasonal-to-seasonal (S2S) timescales (∼15-60 days ahead) and provide a source of predictability (e.g. Kodera and Chiba, 1995; Kolstad et al., 2010; Sigmond et al., 2013; Tripathi et al., 2015b). The variability of the SPV includes strong vortex events (Tripathi et al., 2015a) and weak vortex events, including major sudden stratospheric warmings (SSWs) (e.g. Charlton and Polvani, 2007). Whilst the mean response to an SSW or weakened SPV is a negative phase of the tropospheric Northern Annular Mode (NAM) and equatorward shift of the eddy-driven jets in the troposphere in the weeks-to-months after (Baldwin and Dunkerton, 2001; Kidston et al., 2015), there is a large amount of case-by-case and regional variability (Karpechko et al., 2017; Kretschmer et al., 2018a). Weather regimes provide a helpful framework for examining stratosphere-troposphere coupling. Regimes describe the large-scale atmospheric configuration on any given day and are based on recurrent and persistent patterns in the large-scale circulation (Michelangeli et al., 1995). Because regimes exist on longer

timescales than synoptic weather patterns, they provide an opportunity for longer-range prediction, useful for the energy sector (Beerli et al., 2017; Grams et al., 2017) and for the prediction of cold weather extremes in winter (Ferranti et al., 2018b). Charlton-Perez et al. (2018) described the influence of the strength of the SPV on weather regimes in the North Atlantic, where the tropospheric response to changes in the stratospheric circulation is typically largest. Using four Atlantic wintertime regimes (following Cassou (2008)), they show the SPV strength significantly affects the occurrence and persistence of each regime, and the transition between regimes. This approach helps illuminate some of the reasons behind different tropospheric responses to stratospheric changes (including, but not limited to, SSWs) in a statistical sense.

Whilst the tropospheric response to changes in the SPV is more variable across North America than in the Euro-Atlantic sector, it has been implicated in driving recent extreme cold weather outbreaks in this region (so-called "polar vortex outbreaks"; Waugh et al. (2017)). These are among recent billion-dollar weather and climate disasters in the United States (NOAA, 2019). The North American sector is also partly influenced by Atlantic weather patterns and the NAM, which typically respond strongly to changes in the stratosphere. Kretschmer et al. (2018a) used cluster analysis in the lower stratosphere to elucidate the influence of the SPV on cold extremes in both North America and Eurasia, finding that a pattern associated with planetary wave reflection was important for anomalous cold over North America. This follows earlier work by Kodera et al. (2016), who found a Pacific blocking response to SSWs dominated by planetary wave reflection, with a downstream trough over North America. In addition, the Pacific sector tropospheric response to stratospheric perturbations is not necessarily of the same sign as in the Euro-Atlantic sector (Ambaum et al., 2001).

Although some prior work has described regimes across North America in a similar sense to the Atlantic regimes (Amini and Straus, 2019; Riddle et al., 2013; Robertson and Ghil, 1999; Straus et al., 2007; Vigaud et al., 2018), the use of regimes is not as common in

this region. The number of regimes and the westward and eastward extent of the region used to define the regimes varies between studies, capturing different aspects of Pacific and Atlantic variability. Moreover, the relationship between these regimes and changes in the stratospheric vortex has not yet been quantified.

In this article, we define four tropospheric wintertime regimes across the North American sector and describe the relationship between the regimes and the SPV. We also investigate the link between these regimes and the occurrence of extremely cold weather across North America.

## 4.2 Data and Methods

We use 00Z data from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al., 2011) for all days in December–March in the period January 1979 to December 2017 (a total of 4729 days). December to March is chosen as it encompasses the period of largest SPV variability (e.g. all observed major SSWs have occurred in these months (Butler et al., 2017)). The data are re-gridded to 2.5° horizontal resolution for computational efficiency and since we are considering only large-scale features.

We perform an empirical orthogonal function (EOF) decomposition of linearly de-trended 500 hPa geopotential height anomalies (with respect to the daily January 1979–December 2017 climatology) in the sector 180-30°W, 20-80°N (Figure 4.6). This region is chosen to include the Pacific jet exit region and include relevant North Atlantic variability. De-trending is performed to account for the climate change signal, although it does not notably alter the results (not shown). Data are weighted by the square-root of the cosine of latitude to give equal-area weighting in the covariance matrix. We retain the leading 12 modes of variability, which represent 80% of the total variance in the 500 hPa geopotential height anomaly field.

We then perform *k*-means clustering with *k*=4 using the Python package *scikit-learn* (Pedregosa et al., 2011). All days are then assigned to a regime based on their minimum Euclidean distance to the cluster centroids; we do not employ "no-regime" days (Grams et al., 2017). The resultant regimes are very similar to those found in Vigaud et al. (2018); they show these regimes are a significant representation based on the classifiability index of Michelangeli et al. (1995), so we do not repeat that calculation here. Our four regimes remain largely unchanged as a subset when five or six clusters are used, further indicating they are dominant patterns and form a concise characterization with reasonably large individual sample sizes.

The probability of regime occurrence ($p$), which we term the occupation frequency, is given by ratio of the number of days in a given regime ($n$) to the total number of days ($N$) in the sample:

$$p = \frac{n}{N} \tag{4.1}$$

We use 95% confidence intervals with a normal approximation to a binomial proportion confidence interval, given by:

$$p \pm Z\sqrt{\frac{p(1-p)}{N'}} \tag{4.2}$$

where $Z = 1.96$ from the standard normal distribution. To account for the persistence of the regimes, we employ an effective sample size $N'$, found by using the 1-day persistence probability $r_1$ (e.g. Wilks, 2011) for each regime in each vortex state,

$$N' = N\frac{1-r_1}{1+r_1} \tag{4.3}$$

We do not scale $N$ for confidence intervals on the transition probabilities, since these are independent of the preceding regime. We define the strength of the SPV to be the tercile categories of daily zonal-mean zonal wind at 100 hPa and 60°N, following Charlton-Perez et al. (2018). The 100 hPa level is chosen to represent the coupling layer between the stratosphere and troposphere and include only the effects of stratospheric perturbations which propagate into the lower stratosphere. The results are not qualitatively sensitive to

the choice of lower-stratospheric level (not shown).

Statistical significance of the composite maps is determined by bootstrap re-sampling with replacement. We construct 95% confidence intervals using 50,000 re-samples per regime over all December to March days in the period 1979–2017. Random days are selected in blocks corresponding to the observed regime 'events', to test the null hypothesis that the composites are the result of random sub-sampling of winter days. Further detail on the bootstrapping method is provided in the Supporting Information.

## 4.3 Results

### 4.3.1 Circulation regimes

Composites of mean 500 hPa geopotential height anomalies for each of the four regimes are shown in Figure 4.1. The regimes are very similar to those defined in Straus et al. (2007) (despite a slightly different domain and analysis period) so we follow their naming convention. The least frequent regime (with an occupation frequency of 20%), is the Arctic High (ArH) regime (Figure 4.1a). It is associated with anomalously high geopotential heights over Greenland and the Canadian archipelago (Greenland blocking), and lower than normal geopotential heights over the Atlantic east of the United States but no significant height anomalies in the Pacific sector. The regime resembles the negative phase of the North Atlantic Oscillation (NAO–), and its occupation frequency is equivalent to the NAO– regime in Charlton-Perez et al. (2018). It is also similar to the tropospheric anomalies associated with cluster 5 in Kretschmer et al. (2018a), which they associate with stratospheric planetary wave absorption.

The Arctic Low (ArL) regime (Figure 4.1b) is not a direct counterpart of the ArH regime and is slightly more frequent (25%). Whilst the ArL regime is associated with opposite height anomalies to the ArH regime in the vicinity of Greenland and is somewhat

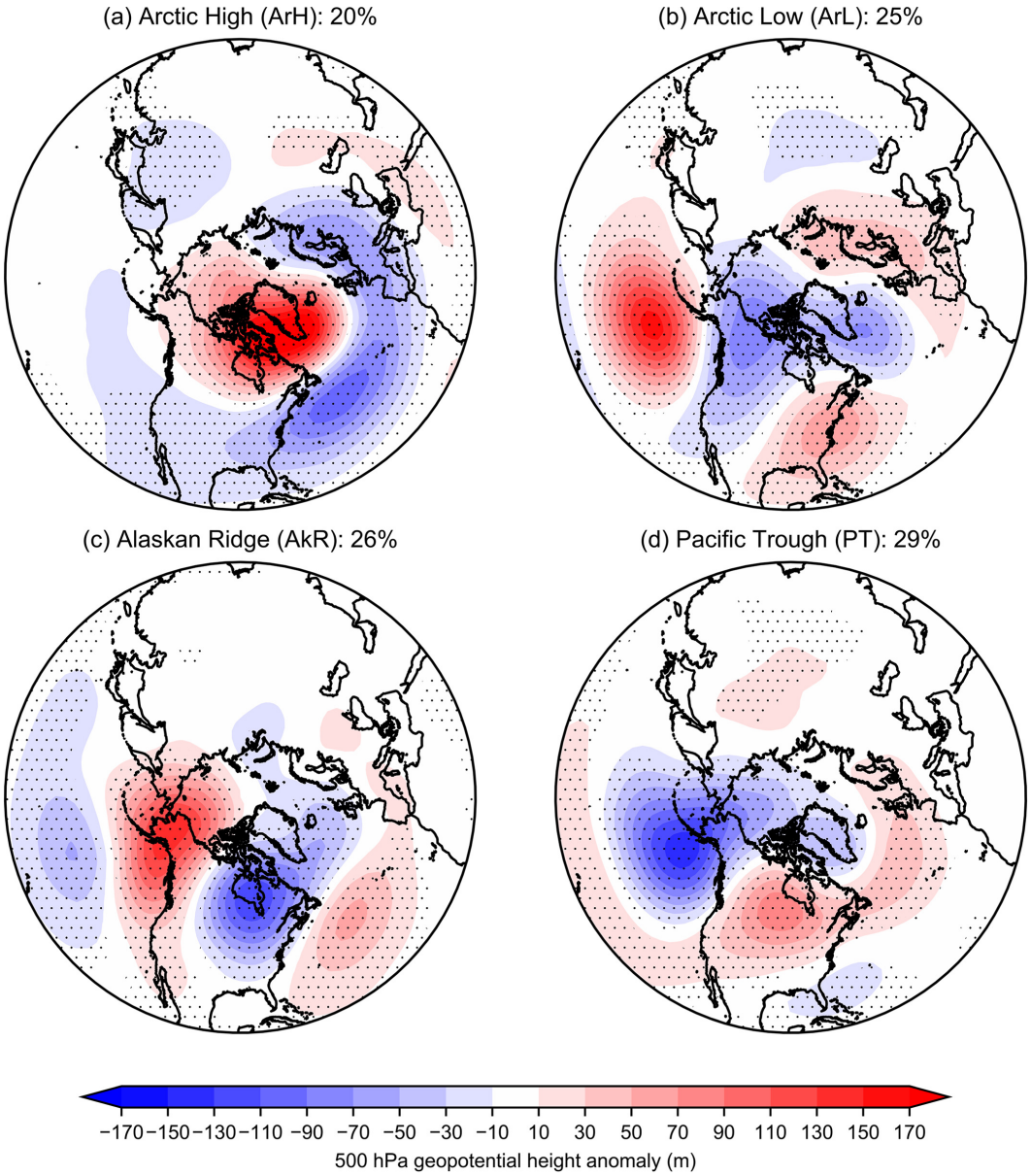


**Figure 4.1:** Composite mean 500 hPa geopotential height anomalies (meters) for each of the four regimes. Anomalies are expressed with respect to the de-trended daily January 1979–December 2017 mean. Percentages indicate the occupation frequency of the regime (the percentage of days assigned to the regime in the November–March period). Stippling indicates significance at the 95% confidence level according to a two-sided bootstrap re-sampling test.

similar to the positive NAO (NAO+), the main signature is a ridge-trough-ridge pattern extending from the Pacific across North America, which resembles the negative phase of the Pacific–North American (PNA–) pattern. The ridge anomaly in the northeast Pacific indicates this regime is associated with a weakened Aleutian low and resembles a negative North Pacific Oscillation (NPO–) (Linkin and Nigam, 2008; Rogers, 1981).

The Alaskan Ridge (AkR) regime (Figure 4.1c), occurring on 26% of days, strongly resembles the Tropical–Northern Hemisphere (TNH) pattern (Mo and Livezey, 1986) and the North American dipole (Wang et al., 2015), the latter of which was linked to the extremely cold North American winter of 2013-14. This regime is also similar to the tropospheric response to cluster 4 in Kretschmer et al. (2018a), which they associate with the reflection of planetary waves by the stratosphere. We note that the AkR and ArL regimes are closest to the patterns during "polar vortex outbreaks" over North America.

The most frequent regime (29%) is the Pacific Trough (PT) (Figure 4.1d), which consists of an anomalous trough centred near Alaska, and an anomalous ridge over continental North America. The trough is consistent with a positive phase of the NPO (NPO+) and the enhancement of the Aleutian Low associated with El Niño, whilst the pattern across North America resembles the positive PNA (PNA+).

### 4.3.2 Relationship with the stratosphere

To quantify the relationship between the stratospheric state and each regime, and by considering the long persistence of lower-stratospheric anomalies during winter (Figure 4.7), we calculate the time-lagged difference in the probability of each regime between weak and strong SPV states. We calculate this difference for the 30 days before and after each day in each regime, conditional on the SPV state at a zero-day lag (day 0) (Figure 4.2).

All but the AkR regime exhibit probability changes greater in magnitude than 0.1, which generally peak around day 0, supporting a stratospheric influence (since this is the given state on which we condition the probability, and we would expect a near-contemporaneous regime response). The ArH regime displays the greatest difference. Its occurrence probability is 0.3–0.4 greater in a contemporaneously weak vortex versus a strong vortex; this difference exceeds 0.1 for all negative lags, which is likely influenced by the long persistence of weak SPV states (and the persistence of this regime in those

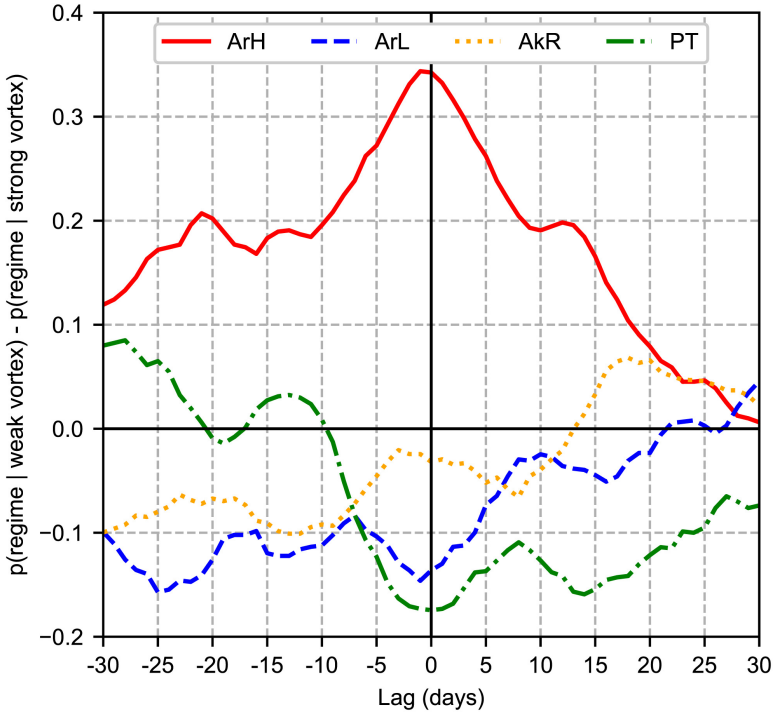


**Figure 4.2:** Difference in the occurrence probability of each regime between weak and strong stratospheric polar vortex states for -30 to +30 day lags, conditional on the vortex state at day 0.

conditions, c.f. Figure 4.3b). Moreover, for almost 20 days following a weak SPV, the probability of the ArH regime is more than 0.1 greater than following a strong SPV. Conversely, the probability of the ArL regime is around 0.1 less in the 30 days preceding a weak SPV, but this difference rapidly decays for positive lags. The PT regime becomes 0.1–0.2 less likely following a weak SPV versus a strong SPV for up to 25 days; it does not display a large change in likelihood for negative lags beyond ∼5 days.

Motivated by the preceding analysis, we next compute the probability of each regime given the SPV strength on the preceding day (Figure 4.3a). Although this is near-instantaneous, it provides a potentially useful framework for extended-range forecasting owing to the persistence and predictability of SPV strength anomalies, and the intrinsic persistence of regimes themselves. The ArH regime demonstrates the largest sensitivity to the stratospheric state, consistent with its negative NAO-like characteristics, with an approximately linear relationship with the tercile SPV strength categories. This regime is seven times more likely following weak SPV states than strong SPV states and is the most likely regime

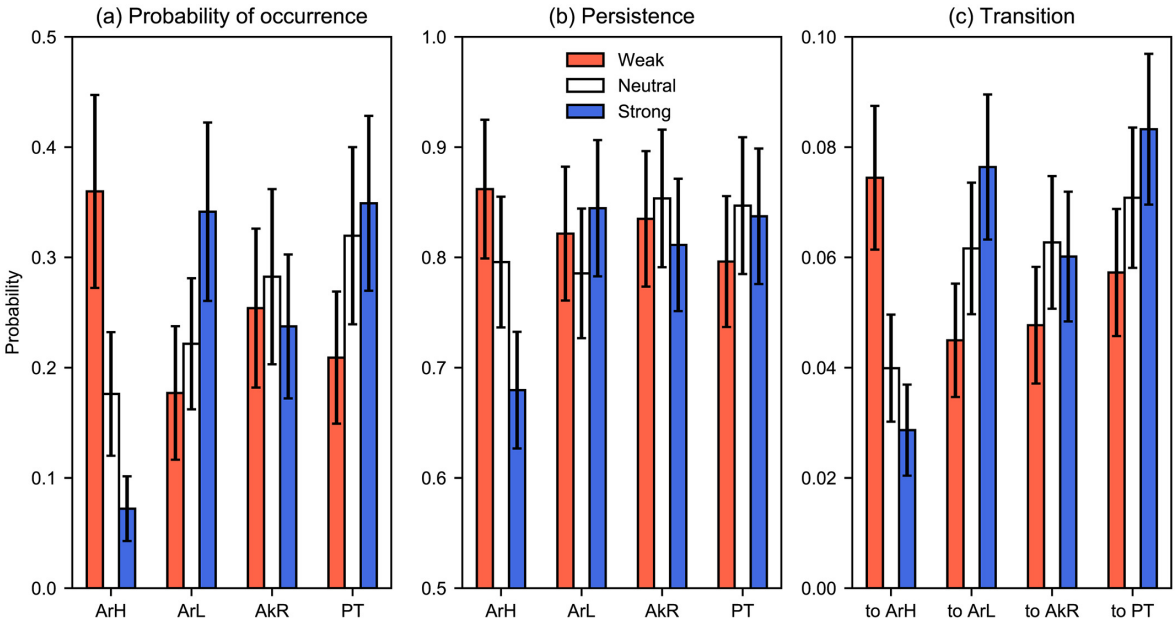


**Figure 4.3:** (a) Probability of occurrence, (b) persistence, and (c) transition of each regime given the tercile category of the stratospheric polar vortex strength on the preceding day. Error bars indicate 95% binomial proportion confidence intervals using a normal approximation (see text for details). Colors indicate the tercile category of the 100 hPa 60°N zonal-mean zonal wind based on daily January 1979–December 2017 climatology.

following a weak SPV. The likelihood of the ArL regime increases with increasing SPV strength; it is approximately twice as likely following a strong versus a weak SPV. For the AkR regime, the dependency on the antecedent SPV strength is statistically insignificant. The PT regime is most likely following neutral and strong SPV conditions, and its behavior is generally similar to the ArL regime.

To further understand vortex-dependent changes in the occurrence probabilities, we compute the probability of persisting in a given regime the following day given the SPV strength on the current day (Figure 4.3b). The persistence of the ArH regime is most strongly dependent on the antecedent SPV strength. Its persistence decreases markedly from 0.86 following a weak SPV to 0.68 following a strong SPV, the lowest persistence probability of any of the regimes for any stratospheric state. This behavior is consistent with its similarity to NAO– (c.f. Figure 3 in Charlton-Perez et al. (2018)). None of the other three regimes exhibit significant changes in persistence probability depending on the SPV strength. Similar results are found when the total duration of each regime is stratified

by the SPV strength on the day of transition into the regime (Figure 4.8), though this metric suggests enhanced duration of the PT regime during strong SPV conditions.

We also consider changes in the transitions between regimes. In Figure 4.3c we show the probability of transitioning from any other regime into a given regime the following day, given the SPV strength on the current day. Transitioning into the ArH regime is 2.5 times more likely during a weak SPV versus a strong SPV. The opposite is true for the ArL and PT regimes, but the relationship is slightly weaker, with the transitions approximately 50% more likely following a strong SPV versus a weak SPV. We also show the difference in specific regime transitions between a weak and a strong SPV in Table 4.1, but emphasize that the sample sizes are much smaller for individual transitions ($n$ = 38–90, and even smaller when categorized by SPV strength), making a robust analysis difficult.

In order to discern the association between these regimes and the middle-stratospheric polar vortex (where major SSWs are commonly defined), we show the composite-mean contemporaneous 10 hPa geopotential height anomalies in Figure 4.4. The pattern during the ArH regime resembles a weak or displaced SPV with an anomalous wavenumber-1 configuration, consisting of anomalously high (low) geopotential heights over the central Arctic (southwest North America and northwest Europe). The anomaly pattern at 10 hPa is similar to that at 500 hPa indicating an equivalent barotropic anomaly structure. The ArL pattern is mostly opposite to ArH, with a strengthened SPV indicated by anomalously low geopotential heights over the central Arctic. The Pacific ridge anomaly present in this regime at 500 hPa does not extend to 10 hPa. The AkR regime features an anomalous wavenumber-2 splitting-type pattern with ridge anomalies in the Atlantic and Pacific, and an anomalous trough over North America. The ridge anomaly over Alaska and trough anomaly over central North America are also present at 500 hPa. The trough anomaly centred near the Hudson Bay is consistent with the similarity of this regime to the "polar vortex" outbreaks driven by a distortion to the vortex.

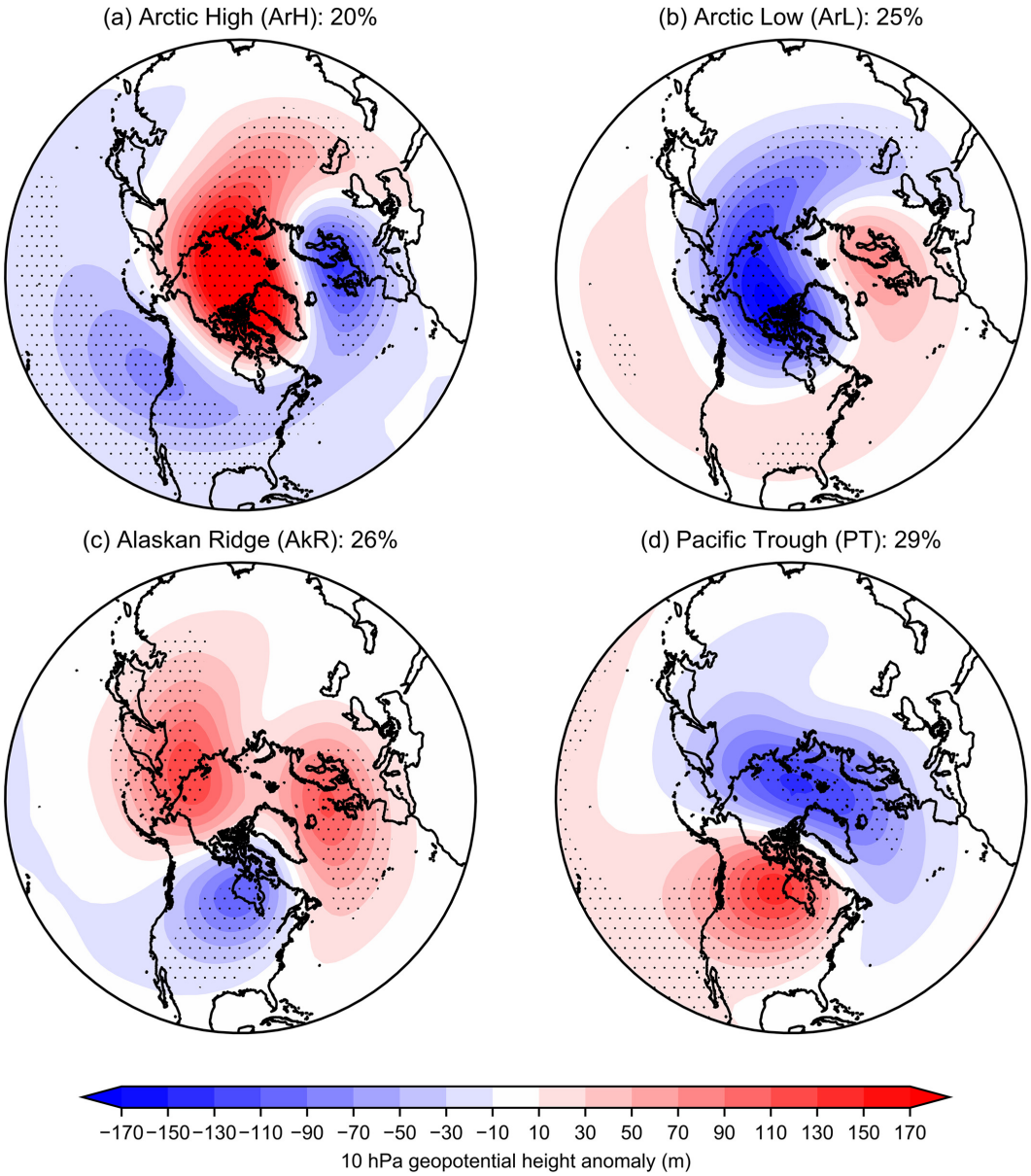


**Figure 4.4:** Composite mean 10 hPa geopotential height anomalies (meters) for days classified in each of the four regimes. Anomalies are expressed with respect to the detrended January 1979–December 2017 mean. Stippling indicates significance at the 95% confidence level according to a two-sided bootstrap re-sampling test.

Whilst the AkR regime does not have occurrence, persistence or transition preferences dependent on the antecedent zonal-mean zonal winds, the contemporaneous 10 hPa anomalies indicate significant disruption to the mid-stratospheric vortex. Therefore, this aspect of vortex variability may not be captured in the 100 hPa 60°N zonal-mean zonal wind; instead, the AkR regime may be more influenced by the morphology of the SPV. Additionally, the similarity of this regime to both the response to reflecting major SSWs described in Kodera et al. (2016) and the patterns found during SPV intensification in

Limpasuvan et al. (2005) indicates a potential relationship with stratospheric variability. The PT regime is associated with a wavenumber-1 anomaly pattern consisting of a barotropic anomalous ridge over North America and a strengthened SPV.

### 4.3.3 Relationship with cold air outbreaks

We next assess the relationship between these regimes and the occurrence of potentially dangerous cold weather outbreaks. To do this, we calculate the probability of severe cold for each regime as the number of days in each regime with normalized 2 m temperature anomalies more than 1.5 standard deviations below the daily mean (similar to the criterion of Thompson and Wallace (2001)). This calculation is performed at each grid-point, and the result is shown in Figure 4.5. Corresponding maps of composite mean 2 m temperature anomalies for each regime are shown in Figure 4.9. Despite the large differences between the likelihood, location and extent of cold weather outbreaks in these regimes, we emphasize that all four can bring cold-weather impacts to parts of the Northern Hemisphere.

Whilst the ArH regime (Figure 4.5a) is the most sensitive to the stratospheric state (c.f. Figures 4.2 and 4.3), we find that it is not the most important for widespread wintertime cold weather outbreaks across North America (though there is a significant risk of severe cold (5-10%) for all but northeastern North America during this regime). Moreover, the magnitude of the mean temperature anomalies during this regime are relatively small (Figure 4.9a). The ArH regime is instead associated with the highest risk (>20%) of severe cold only across northwest Europe, consistent with its NAO– characteristics. We find that severe cold weather outbreaks across the continental interior of North America are most likely during the AkR regime (Figure 4.5c), with chances of severe cold exceeding 20%, and mean temperature anomalies widely 5°C below normal (Figure 4.9c). The ArL regime (Figure 4.5b) is associated with a 10-15% chance of extreme cold across western North America, including Alaska, whilst in the central and east of the United States there is an absence of extreme cold during this regime. The PT regime (Figure 4.5d) features

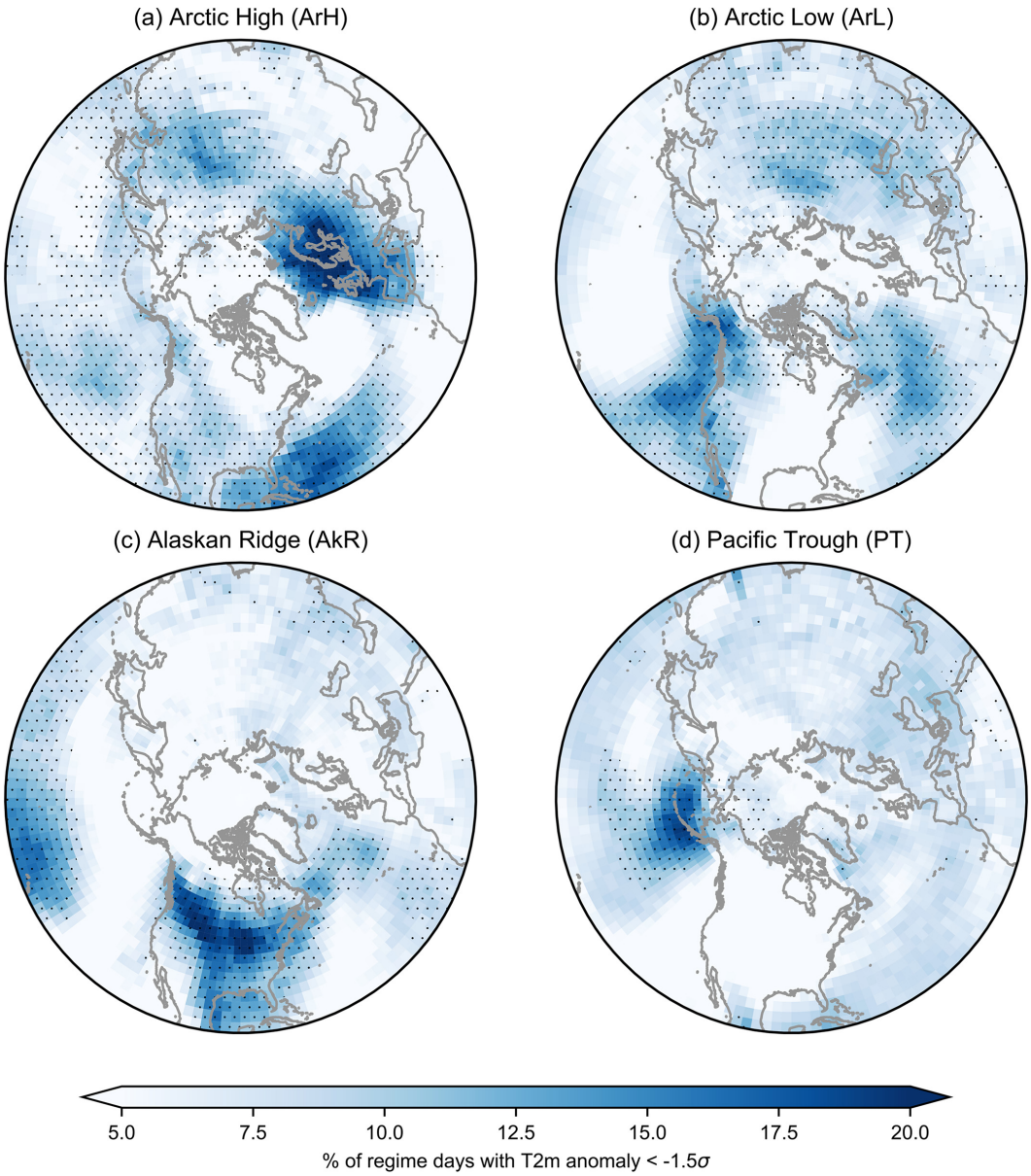


**Figure 4.5:** Percent of all days in each regime with daily standardized 2 m temperature anomalies $<$ -1.5 $\sigma$ (with respect to the linearly de-trended daily January 1979–December 2017 mean). Stippling indicates significance at the 95% confidence level according to a one-sided bootstrap re-sampling test.

an absence of extreme cold across most of North America, with mean temperatures widely more than 5°C above normal (Figure 4.9d). Extreme cold during this regime is typically confined to western Alaska and the Aleutian Islands, consistent with the western periphery of the anomalous trough. The PT regime also has the lowest overall risk of cold weather outbreaks across the Northern Hemisphere.

## 4.4 Summary and Conclusions

In this study we have shown that the behavior of three of four wintertime North American weather regimes is significantly linked to the antecedent strength of the SPV. We find that whilst the ArH regime is most sensitive to the SPV strength, it is not the most important for *widespread* extreme cold outbreaks in North America – particularly in central and northern areas where such extremes correspond to the coldest absolute temperatures. Instead, we find that the AkR regime – which does not display a significant dependence on the lower-stratospheric zonal-mean zonal wind – is associated with the greatest risk of extreme cold across most of North America. Though Figure 4.4c suggests a possible link exists with the state of the SPV, the similarity of this regime to the TNH pattern suggests that tropical forcing may also exhibit a large control on its behavior (e.g. Hartmann, 2015).

Further work should address the ability of sub-seasonal forecast models to correctly capture the downward coupling of stratospheric anomalies onto these regimes, as well as illuminating the dynamics involved, such as Rossby wave breaking (e.g. Michel and Rivière, 2011), and the impact of model biases. It should also be investigated whether Pacific phenomena on intra-seasonal (such as the Madden-Julian Oscillation (MJO)) to seasonal (e.g. the El Niño-Southern Oscillation (ENSO)) and decadal scales (e.g. the Pacific Decadal Oscillation (PDO)) interact constructively or destructively with the stratospheric influence.

## Acknowledgments

S.H.L. was funded by the Natural Environment Research Council (NERC) via the SCENARIO Doctoral Training Partnership (NE/L002566/1). J.C.F. acknowledges funding via the National Science Foundation (NSF) Atmospheric and Geospace Sciences Program (award #AGS-1657905). The ERA-Interim reanalysis dataset is available at `https://apps.ecmwf.int/datasets/`. We acknowledge two anonymous reviewers for their detailed comments which helped improve the paper.

## 4.5 Supporting Information

### Comment on bootstrapping procedure

The bootstrap re-sampling procedure used for the composite maps operates as follows. For each regime:

1. For each regime occurrence, compute the number of consecutive days in the regime, to produce "blocks" of regime days.

2. Randomise each set of consecutive days to produce random sets of numbers with the same structure as observed.

3. Select these days of data from all days in the months December to March from 1 January 1979 to 31 December 2017.

4. Repeat (2) and (3) 50,000 times.

5. At each grid-point, find the 2.5/97.5 percentiles of the re-sampled distribution. If the observed value lies below/above these, then it is classified as statistically significant.

**Table 4.1:** Difference in probabilities for specific regime transitions in a weak stratospheric vortex versus a strong stratospheric vortex, defined by the lower and upper tercile anomaly categories of 100 hPa 60°N zonal-mean zonal wind. The sample size for each transition is shown in brackets. Left-hand column indicates the starting regime (D0), further columns indicate the transition on the following day (D+1).

| D0/D+1 | ArH | ArL | AkR | PT |
|---|---|---|---|---|
| ArH | 0.18 (783) | -0.06 (38) | -0.07 (61) | -0.05 (70) |
| ArL | 0.05 (46) | -0.02 (954) | -0.01 (75) | -0.02 (86) |
| AkR | 0.03 (35) | -0.01 (90) | 0.02 (1006) | -0.04 (75) |
| PT | 0.07 (86) | -0.02 (83) | 0.00 (62) | -0.04 (1138) |

**Figure 4.6:** Map showing the region used for cluster analysis, delineated by the red polygon (180–30°W, 20–80°N)

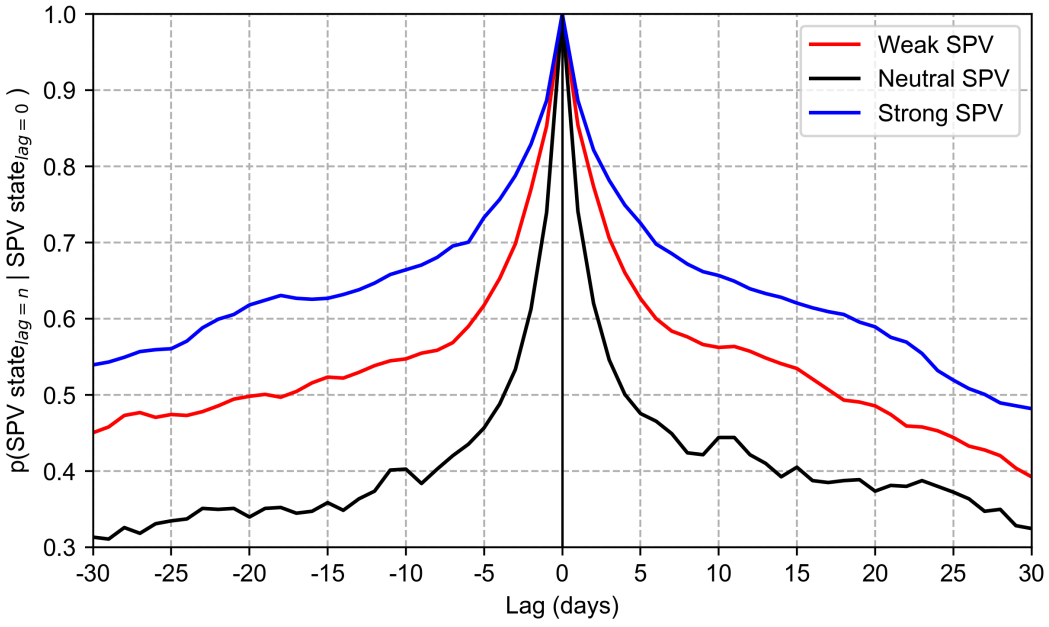


**Figure 4.7:** -30 to +30 day lagged conditional probability of tercile category of stratospheric vortex strength (100 hPa 60°N zonal-mean zonal wind) given the tercile category on day 0.

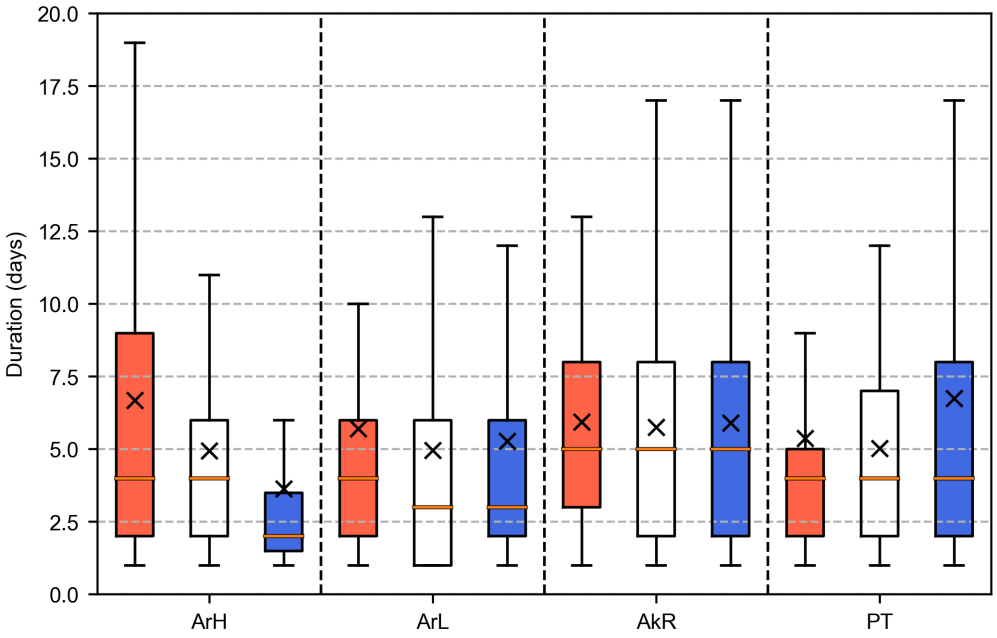


**Figure 4.8:** Boxplots of the total duration of each regime stratified by the tercile anomaly categories (red = lower tercile, white = middle tercile, blue = upper tercile) of 100 hPa 60°N zonal-mean zonal wind on the first day of the regime. Crosses indicate the mean, horizontal orange lines indicate the median. Outliers beyond 1.5 times the interquartile range are hidden.

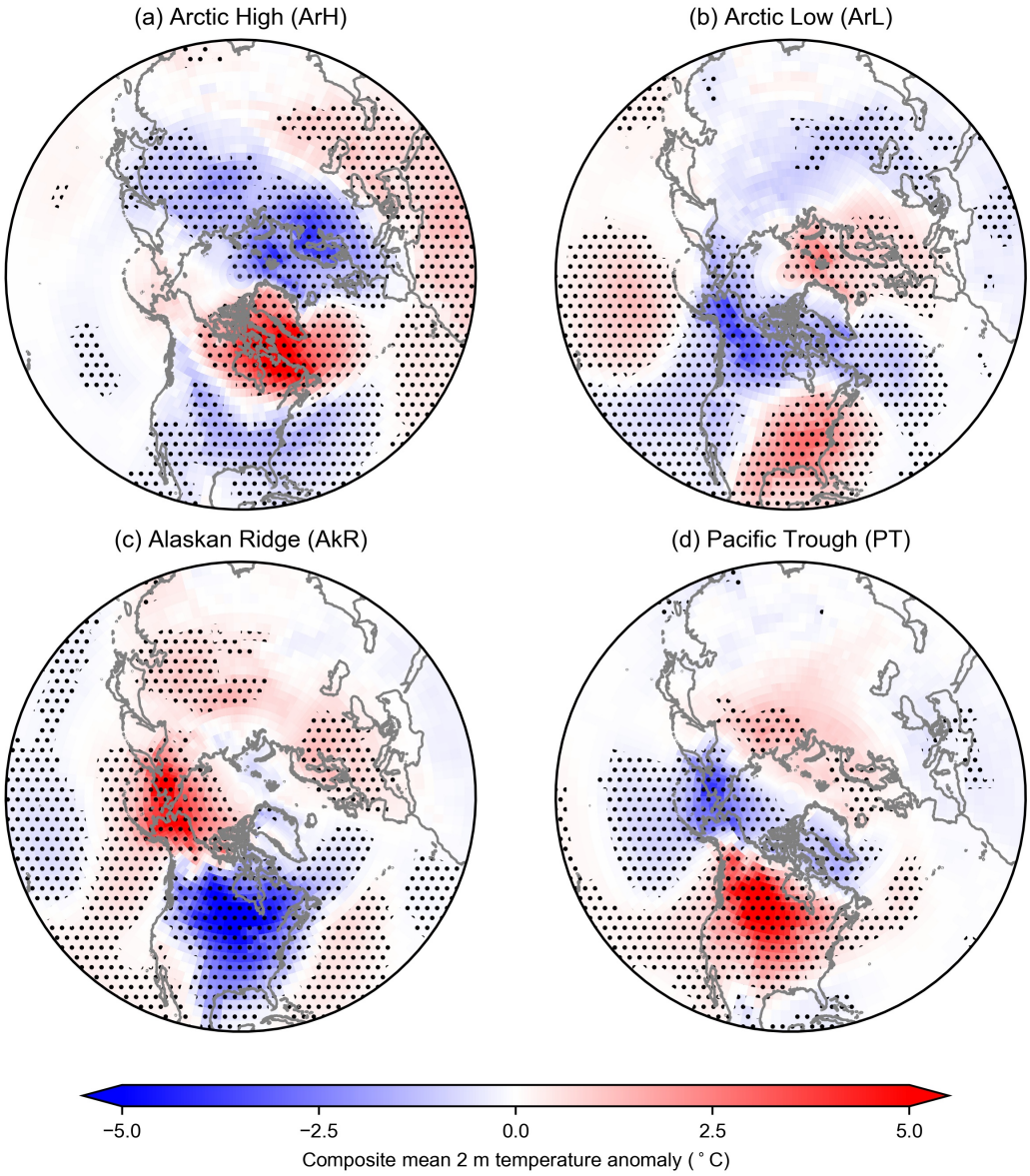


**Figure 4.9:** Composite mean 2 m temperature anomalies for each regime. Anomalies are expressed with respect to the linearly de-trended daily January 1979–December 2017 mean. Stippling indicates regions that are significant at the 95% confidence level according to a two-sided bootstrap re-sampling test (see main text for details).

# Chapter 5

# How do stratospheric perturbations influence North American weather regime predictions?

This chapter has been revised following review for *Journal of Climate* with the following reference:

**Lee, S. H.**, A. J. Charlton-Perez, S. J. Woolnough, and J. C. Furtado, 2021: How do stratospheric perturbations influence North American weather regime predictions? *Journal of Climate*, in review.

# Abstract

Observational evidence shows changes to North American weather regime occurrence depending on the strength of the lower-stratospheric polar vortex. However, it is not yet clear how this occurs or to what extent an improved stratospheric forecast would change regime predictions. Here we analyze four North American regimes at 500 hPa, constructed in principal component (PC) space. We consider both the location of the regimes in PC space and the linear regression between each PC and the lower-stratospheric zonal-mean winds, yielding a theory of which regime transitions are likely to occur due to changes in the lower stratosphere. Using a set of OpenIFS simulations, we then test the effect of relaxing the polar stratosphere to ERA-Interim on subseasonal regime predictions. The model start dates are selected based on particularly poor subseasonal regime predictions in the European Centre for Medium Range Weather Forecasts CY43R3 hindcasts. While the results show only a modest improvement to the number of accurate regime predictions, there is a substantial reduction in Euclidean distance error in PC space. The average movement of the forecasts within PC space is found to be consistent with that expected from the linear theory for moderate-to-large lower-stratospheric zonal wind perturbations. Overall, our results provide a framework for interpreting the stratospheric influence on North American regime behavior. The results can be applied to subseasonal forecasts to understand how stratospheric uncertainty may affect regime predictions, and to diagnose which regime forecast errors are likely to be related to stratospheric errors.


## 5.1 Introduction

The framework of large-scale weather regimes is now increasingly used in wintertime subseasonal-to-seasonal (S2S) prediction (∼2 weeks to 2 months ahead; White et al. (2017)), although the concept of a weather "regime" is not new (Rex, 1951). Regimes are characteristically recurrent, persistent, and quasi-stationary (e.g., Michelangeli et al., 1995) with typical timescales of weeks, well-suited to the subseasonal scale where they can manifest "windows of opportunity" for skillful extended-range forecasts (Mariotti et al.,

2020; Robertson et al., 2020).

Unlike empirical orthogonal functions (EOFs) (e.g., Hannachi et al., 2007), regimes defined through clustering methods are not bound by orthogonality or variance partitioning constraints. These regimes can therefore more closely represent the full anomalous flow configuration on a given day by benefiting from 'mode-mixing' and are accordingly easier to interpret, providing a useful way to understand extended-range ensemble forecasts. By characterizing recurrent flow configurations, weather regimes can also be used to diagnose flow-dependent predictability (Ferranti et al., 2015; Matsueda and Palmer, 2018). From an impacts perspective, regimes have been used to better understand meteorological impacts on energy demand (e.g., Grams et al., 2017; van der Wiel et al., 2019; Garrido-Perez et al., 2020), precipitation and wildfire risk (Robertson and Ghil, 1999; Robertson et al., 2020), and public health (Charlton-Perez et al., 2019; Huang et al., 2020).

A significant source of tropospheric subseasonal predictability during boreal winter is variability in the Arctic stratospheric polar vortex (e.g., Butler et al., 2019a; Domeisen et al., 2020c), including sudden stratospheric warmings (SSWs; e.g., Charlton and Polvani, 2007) and strong vortex events (e.g., Limpasuvan et al., 2005; Tripathi et al., 2015a). The downward influence of the stratosphere can be viewed as the modulation of weather regime transition and persistence. Perhaps the simplest regime framework employs the two phases of the North Atlantic Oscillation (NAO), which are similar to the Northern Annular Mode (NAM) and Arctic Oscillation (AO) patterns and strongly influenced by the stratosphere (Ambaum et al., 2001; Baldwin and Thompson, 2009; Hitchcock and Simpson, 2014). More complex regime analyses for the North Atlantic-European sector invoke four (e.g., Vautard, 1990; Cassou, 2008), six (Falkena et al., 2020) or seven (e.g., Grams et al., 2017) regimes depending on the method, focus, or purpose of the analysis.

Using four North Atlantic regimes, Charlton-Perez et al. (2018) found significant differences in the occurrence likelihood of three regimes between strong and weak lower-stratospheric

vortex states, while the probability of Scandinavian Blocking was invariant. Beerli and Grams (2019) related the stratospheric modulation of Atlantic weather regimes to whether or not the regime projected strongly onto the NAO pattern. They emphasized that regimes which do not project strongly onto the NAO provide a route for a wider variety of weather patterns following anomalous stratospheric vortex states. Subsequently, Maycock et al. (2020) analyzed the North Atlantic response to SSWs from the perspective of modulation of the three eddy-driven jet regimes, finding an increase in the occurrence and persistence of the southernmost regime (corresponding to the negative NAO). Domeisen et al. (2020a) assessed the varying degrees of stratosphere-troposphere coupling following major SSWs (e.g., Karpechko et al., 2017; White et al., 2019) by considering the regimes present during SSW onset and in the weeks afterward, suggesting the antecedent state of the troposphere may play an important role in determining subsequent downward coupling.

In recent years, the influence of the stratosphere on North American climate variability has received increased attention, likely owing to the extreme cold-air outbreaks during winter 2013–14 which accompanied disruption to the polar vortex (Yu and Zhang, 2015; Waugh et al., 2017). However, less attention has been given to explicitly viewing the impact of the stratosphere on North American weather from a regimes perspective. As North America is influenced by weather from both the Atlantic and Pacific to different degrees across the continent, a challenge with defining North American regimes is the choice of domain. Some studies (e.g., Amini and Straus, 2019; Fabiano et al., 2021) focus on upstream variability in the Pacific-North American (PNA) sector (akin to the Atlantic regimes with respect to Europe), while others focus on the continent as a whole and incorporate both Atlantic and Pacific variability. Despite some methodological differences, a growing number of studies have defined a consistent and reproducible set of four wintertime regimes in the 500 hPa geopotential height anomaly field centered over North America (e.g., Straus et al., 2007; Vigaud et al., 2018; Lee et al., 2019c; Robertson et al., 2020). The regimes capture both PNA-like and NAO-like behavior.

More specifically, Lee et al. (2019c) analyzed these four North American regimes (namely: Arctic High, Arctic Low, Alaskan Ridge, and Pacific Trough) in the context of the strength of the lower-stratospheric polar vortex in reanalysis. They found significant differences between the occurrence of three of the regimes during strong and weak stratospheric vortex states of a similar magnitude to those in Charlton-Perez et al. (2018) for the North Atlantic. The Alaskan Ridge regime did not show a relationship with the stratospheric vortex strength, but was found to be strongly linked to North American coldwaves. Lee et al. (2019c) hypothesized that tropical forcing (e.g., Wang et al., 2014) or stratospheric wave reflection (Kodera et al., 2016; Kretschmer et al., 2018a; Matthias and Kretschmer, 2020) may dominate driving the Alaskan Ridge, owing to the similarity of the regime to patterns associated with both. As a purely observation-based study, the results of Lee et al. (2019c) were non-causal and did not assess when or how changes in the stratospheric state would change regime occurrence, or whether improved stratospheric forecasts would yield better regime predictions. Addressing these points is therefore a goal of the present study.

To diagnose the downward influence of the stratosphere on the troposphere, and changes in tropospheric forecast skill arising from a correctly-predicted stratosphere, model experiments in which the stratospheric state is artificially nudged or relaxed to a different state (such as that from reanalysis) have been used. Most studies have focused on the seasonal-scale effects (Douville, 2009; Hitchcock and Simpson, 2014; Jung et al., 2010a,b). However, Kautz et al. (2020) used relaxation experiments on S2S timescales to quantify the role of the February 2018 SSW in the predictability and onset of the subsequent Eurasian cold wave. They found an increased probability of surface cold extremes in forecasts with a nudged stratosphere, but that the evolution of the lower-stratospheric NAM following the SSW – rather than simply the occurrence of the SSW – was important for more accurate tropospheric forecasts. The importance of persistent lower stratospheric anomalies in eliciting a tropospheric response is consistent with climate model studies (Maycock and Hitchcock, 2015; Runde et al., 2016) and the Polar-Night Jet Oscillation

events of Hitchcock et al. (2013).

Although SSWs and their strong vortex counterpart are typically harbingers of persistent anomalous lower-stratospheric NAM states (Baldwin and Dunkerton, 2001), they do not necessarily propagate into the lowermost stratosphere, and anomalous lower-stratospheric NAM states can occur without a typical mid-stratospheric precursor. Hence, analysis of the effect of the stratosphere on the troposphere need not only focus on such extreme mid-stratospheric circulation events. Further, the NAM in the lower stratosphere during midwinter possesses a very long timescale (over 4 weeks; Baldwin et al. (2003)), key for the S2S prediction scale. In this study, we focus entirely on subseasonal variability in the strength of the lower-stratospheric polar vortex, diagnosed through the zonal-mean zonal wind at 100 hPa and 60°N (U100). We do not explicitly consider SSWs or strong vortex events.

The overall goal of this study is to understand how changes or uncertainty in the subseasonal lower stratospheric vortex state can influence changes or uncertainty in predictions of North American weather regimes. We do this first by a statistical analysis of the regimes and their underlying EOFs in reanalysis, and then through analyzing a set of model experiments in which the stratosphere is nudged toward reanalysis. A greater understanding of the relationship between stratospheric variability and regimes will help in both the real-world understanding and interpretation of regime forecast uncertainty, and in subsequent studies of regime dynamics and predictability. It would also be a useful tool to examine how model biases affect the representation of stratosphere-troposphere coupling.

The paper is organized as follows. Section 5.2 introduces the data, methods and model experiments. Section 5.3 defines the regimes and their underlying EOFs, and the relationship between these EOFs and the lower-stratospheric polar vortex strength. Section 5.4 develops a theory of how the stratosphere may influence regime behavior. Section 5.5 presents the results of a modeling study used to test the theory. A summary and conclusion of our work

follows in Section 5.6, including implications for S2S prediction.

## 5.2 Data and Methods

### 5.2.1 Hindcasts and Reanalysis

For historical analysis and verification, we use the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis (Dee et al., 2011). Hindcasts are taken from version CY43R3 of the ECMWF extended-range prediction system (used to produce operational forecasts from July 2017–June 2018) as part of the S2S database. The hindcasts consist of an 11-member ensemble (1 unperturbed member and 10 perturbed members) initialized from ERA-Interim twice per week. The model has a resolution of Tco639[1] up to day 15 and Tco319 after day 15, and 91 vertical levels[2]. All data are sampled once per day at 0000 UTC, and re-gridded to 2.5° latitude-longitude resolution for computational efficiency and since we are only considering large-scale fields.

### 5.2.2 Regime Definitions

The definition of North American weather regimes follows that of Lee et al. (2019c), extended by 1 year. We take 500 hPa geopotential heights (Z500) in the region 180-30°W 20-80°N in all December–March days in the period 1 January 1979–31 December 2018 in ERA-Interim (4,840 days), and subtract the daily climatology over this period. (Any trends in Z500 are found to have little impact on the regimes, so detrending is not performed.) Then, data are weighted by the square-root of cosine latitude, and EOF analysis is performed, retaining the leading 12 EOFs which explain close to 80% of the variance. $k$-means clustering is then performed (Pedregosa et al., 2011) in the 12-dimensional principal component (PC) space, with $k$ set to 4. In addition to reducing the dimensionality of the clustering problem and filtering smaller-scale variability, performing the clustering in PC space produces a co-ordinate system which enables

[1] Tco = cubic octahedral spectral truncation

[2] Details of the prediction system can be found on the ECMWF website: `https://confluence.ecmwf.int/display/S2S/ECMWF+Model+Description+CY43R3`

interpretation of the regimes in terms of their comprising EOFs – linking two widely-used prediction frameworks. After generating the clusters, each day is then assigned to one of the four regimes by the minimum Euclidean distance to the cluster centroids in PC space.

For regime assignment in the hindcasts, the model Z500 climate is first subtracted, to account for systematic biases. The model climate is computed for each initialization date and lead-time over the 20-year hindcast period. Then, the daily data are projected onto the 12 EOFs, and each day is assigned to a regime based on these pseudo-PC loadings. As an additional forecast diagnostic in the model experiments, weekly-mean regimes are produced by first averaging the PCs over a 7-day period and then assigning to a regime; these are found to be largely consistent with the regime occupying the majority of days within each week (not shown).

### 5.2.3 Regime Bust Criteria

We select subseasonal regime "busts" from the ECMWF hindcasts where there is strong ensemble support ($\geq$ 7 members, or approximately two-thirds) for *one* specific incorrect regime to be dominant (i.e., present on at least 8 days) during days 14–27 (weeks 3–4). These criteria are designed to pick out cases which suggest a strong, but incorrect, subseasonal signal constraining the model analogous to a 'precise but inaccurate' forecast. As such, these forecasts may be (erroneously) interpreted as windows of opportunity, with potentially large real-world impacts. We choose only hindcasts initialized during December–February, as the seasonal cycle may affect week 3-4 forecasts initialized during March. These criteria yield 31 initialization dates. A further stipulation is applied such that the initialization dates must be separated by at least 21 days to avoid analyzing multiple instances of the same event; in these cases, the earliest initialization date is selected. This filters the number of cases to 20 (i.e., on average 1 per winter), which are listed in Table 5.1. Except for forecasts of Arctic High verifying as Alaskan Ridge, all forecast-verification combinations are included at least once (not by design).

No stratospheric error criteria are included in order to assess both to what extent poor subseasonal regime forecasts are associated with stratospheric errors and the effect of stratospheric relaxation even in cases with a relatively well-forecast stratosphere. We find that the majority of bust cases (14 of the 20 initialization dates, including 8 week 3 and 12 week 4 forecasts) feature ensemble-mean U100 error magnitudes $\geq 3$ m s$^{-1}$, approximately the upper tercile of those in the December–February hindcasts.

**Table 5.1:** North American regime busts in ECMWF CY43R3 hindcasts (HC) from December 1997–February 2017. The week 3-4 dominant (dom.) regime is that which is predicted by $\geq$7 ensemble members (64%) to be present on $\geq$8 days during days 14–27 inclusive, verified against the ERA-Interim regime which is present for $\geq$8 days during the same time period. Week 3 and week 4 regimes are the regime of the weekly-mean field with the largest ensemble support. $\varepsilon$U is the ensemble-mean error in the 100 hPa 60°N zonal-mean zonal winds averaged over each week. The data are grouped by the dominant regime prediction and then sorted by the week 4 $\varepsilon$U.

| Initialization | W3-4 dom. % [ERA] | W3 HC [ERA] | W3 $\varepsilon$U (m s$^{-1}$) | W4 HC [ERA] | W4 $\varepsilon$U (m s$^{-1}$) |
|---|---|---|---|---|---|
| **Arctic High** | | | | | |
| 21 Dec 2005 | 64 [PT] | ArH [PT] | -0.5 | ArH [PT] | 4.2 |
| 1 Feb 2009 | 64 [ArL] | ArH [ArL] | 2.5 | ArH [ArL] | 3.2 |
| 8 Feb 2010 | 73 [PT] | ArH [ArH] | 0.3 | ArH [PT] | -4.8 |
| 29 Jan 1998 | 64 [PT] | PT [PT] | -8.5 | ArH [PT] | -6.7 |
| **Arctic Low** | | | | | |
| 29 Jan 2001 | 73 [AkR] | ArL [ArL] | 6.5 | ArL [AkR] | 8.5 |
| 28 Dec 2016 | 82 [AkR] | ArL [AkR] | 2.7 | ArL [AkR] | 3.0 |
| 8 Feb 2006 | 64 [ArH] | ArL [ArH] | 4.8 | ArL [ArH] | 2.3 |
| 22 Jan 1999 | 64 [PT] | ArL [PT] | -1.5 | ArL [PT] | 1.0 |
| 19 Feb 2011 | 64 [PT] | ArL [PT] | -0.3 | ArL [PT] | -0.6 |
| 4 Dec 2011 | 64 [PT] | ArL [ArL] | 0.1 | ArL [PT] | -1.3 |
| **Alaskan Ridge** | | | | | |
| 11 Dec 2001 | 64 [ArH] | AkR [ArH] | 2.3 | AkR [PT] | 3.1 |
| 15 Feb 2017 | 64 [ArL] | AkR [ArL] | -0.6 | AkR [AkR] | 2.6 |
| 4 Dec 2003 | 73 [PT] | ArH [PT] | 0.4 | AkR [ArL] | -3.0 |
| **Pacific Trough** | | | | | |
| 12 Feb 1999 | 64 [ArH] | PT [PT] | 3.3 | PT [ArH] | 14.0 |
| 8 Jan 2010 | 64 [ArH] | PT [ArH] | 4.1 | ArH [ArH] | 8.7 |
| 25 Dec 2015 | 73 [ArH] | PT [ArH] | 7.7 | PT [ArH] | 7.7 |
| 7 Dec 2000 | 64 [ArH] | PT [ArH] | 7.3 | PT [ArH] | 2.8 |
| 18 Jan 2016 | 73 [AkR] | PT [PT] | 0.3 | PT [AkR] | 0.4 |
| 21 Dec 2014 | 73 [AkR] | AkR [AkR] | -1.7 | PT [AkR] | -2.1 |
| 25 Dec 2006 | 82 [ArL] | PT [ArL] | -5.8 | PT [ArL] | -8.7 |

### 5.2.4 OpenIFS Model

For model experiments, we use OpenIFS[3] version 43r3v1 – a research version of the ECMWF IFS (Integrated Forecast System) model CY43R3, but without data assimilation. The model is initialized from ERA-Interim and run on a linear Gaussian grid with T255 resolution, 60 vertical levels (i.e., the resolution of ERA-Interim), and a time step of 45 minutes. Output data are bi-linearly interpolated onto a 2.5° latitude-longitude grid. Each ensemble consists of an unperturbed member and 20 perturbed members, in which spread is generated by the Stochastically Perturbed Parametrization Tendencies (SPPT) and Stochastic Kinetic Energy Backscatter (SKEB) schemes (Leutbecher et al., 2017). The ensemble size is chosen as a balance between the potential gain from additional members compared with the 11-member hindcasts and computational expense. The OpenIFS runs crucially differ from the operational model in both resolution and in that there is no representation of initial condition uncertainty, so some differences between these model runs and the equivalent hindcasts are to be expected. As we are primarily considering forecasts on timescales of several weeks, the initial condition uncertainty is considered less important, and the stochastic schemes generate spread comparable to the hindcasts in the fields analyzed in this study.

For each initialization date, two sets of ensembles are produced: a control (CTR) run in which the forecast freely evolves (comparable with the equivalent hindcast, notwithstanding the model differences), and a relaxed (RLX) run in which the Arctic stratosphere is nudged toward ERA-Interim using the IFS relaxation scheme (e.g., Jung et al., 2010a). The relaxation scheme operates by applying a non-physical tendency to the model equations of the form:

$$\lambda(X_{obs} - X) \tag{5.1}$$

where $X$ is a model prognostic variable, $X_{obs}$ is the "observed" value from ERA-Interim and $\lambda$ (units: (time step)$^{-1}$) is the relaxation coefficient controlling the strength of the

[3]Specific details of the model can be found at: `https://confluence.ecmwf.int/display/OIFS/Release+notes+for+OpenIFS+43r3v1`

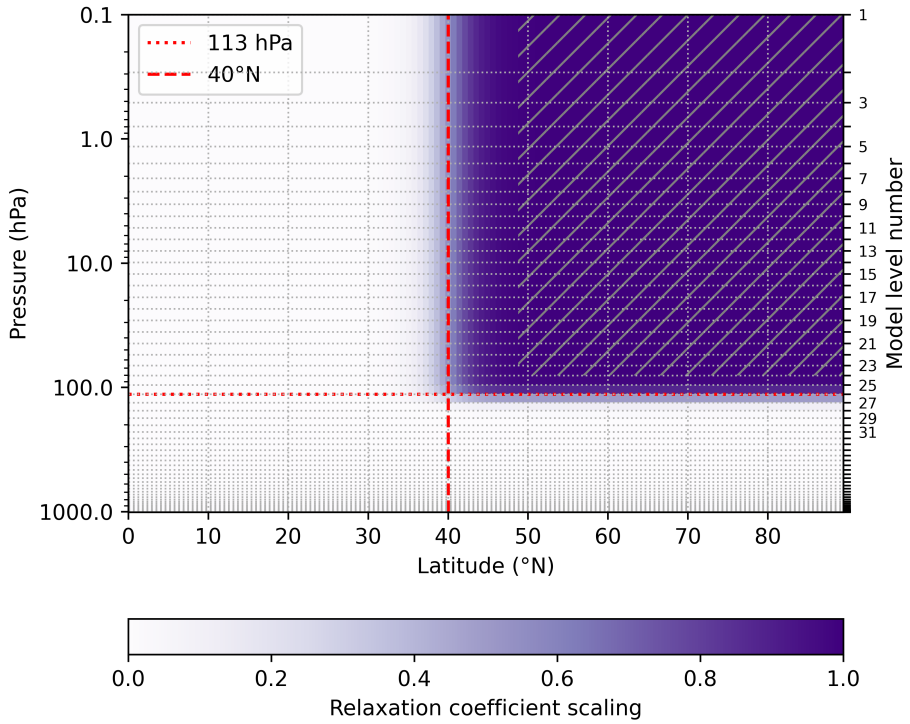


**Figure 5.1:** Vertical and latitudinal profile of the relaxation coefficient scaling (i.e., a value of 1 denotes full relaxation, here with a timescale of 12 hours), for both pressure (left-hand ordinate) and model level number (right-hand ordinate and horizontal grid lines; labeled to level 31 for clarity). The red dashed and dotted lines denote the bounds, in latitude and height respectively, where the coefficient is 0.5. The hatched area denotes the region where the scaling is at least 0.99.

forcing (following e.g., Jeuken et al. (1996); Magnusson (2017)). $X_{obs}$ at each model time-step is generated by linear interpolation between 6-hourly reanalysis files. A relaxation timescale of 12 hours is used in this study, corresponding to $\lambda = 0.0625$ (time step)$^{-1}$ given the 45-minute model time-step, which can be interpreted as nudging the model state at each time-step by 6.25% of the departure from the reanalysis. Vorticity, divergence, and temperature are relaxed in model grid-point space with an exponential taper at both the latitude and model-level boundaries.

A profile of the relaxation domain is shown in Figure 5.1. The domain boundaries are chosen to both maximize constraint of the polar lower stratosphere while allowing for a sufficiently smooth taper to avoid negative numerical effects, and to remain largely poleward and upward of the subtropical jet to reduce directly constraining the tropical upper-tropospheric waveguide. The choice of domain is also limited by the vertical level

spacing of the model in the upper troposphere-lower stratosphere. We employ a weaker stratospheric nudging than some previous studies (e.g., Jung et al., 2010a; Kautz et al., 2020), but note that the relaxation in our study extends further into the lower stratosphere. Analysis of the output fields show this relaxation strength is enough to constrain the model. Time series of the U100 forecasts from the CTR and RLX experiments and the corresponding verification from ERA-Interim are shown in Figure 5.11 in the Supplemental Material.

As the random seed used in the stochastic schemes is fixed for each ensemble member, the equivalent ensemble members in the CTR and RLX experiments differ only by the stratospheric nudging. In analyzing the OpenIFS runs, we assume the model climatology is equivalent to that of the corresponding CY43R3 hindcasts.

### 5.2.5 Significance testing

Throughout the paper, statistical significance is assessed at the 95% confidence level by bootstrap re-sampling (e.g., Wilks, 2019). Random samples (with replacement) are taken from the population and the quantity under analysis (e.g., a regression coefficient) is calculated and stored. This process is repeated 10,000 times, and then a confidence interval is constructed from the appropriate percentiles of this distribution (2.5-97.5 for two-sided 95% confidence).

## 5.3 Regimes and EOFs

The centroids of the four regimes (expressed as the Z500 field reconstructed from the sum of the centroid loading in the leading 12 EOFs), along with the percent of days assigned to each (the occupation frequency), are shown in Figure 5.2a–d. In terms of both spatial patterns and the ranking of occupation frequency, these match the regimes of Lee et al. (2019c) and so we follow their naming convention (after Straus et al., 2007): Arctic High (ArH), Arctic Low (ArL), Alaskan Ridge (AkR), and Pacific Trough (PT).

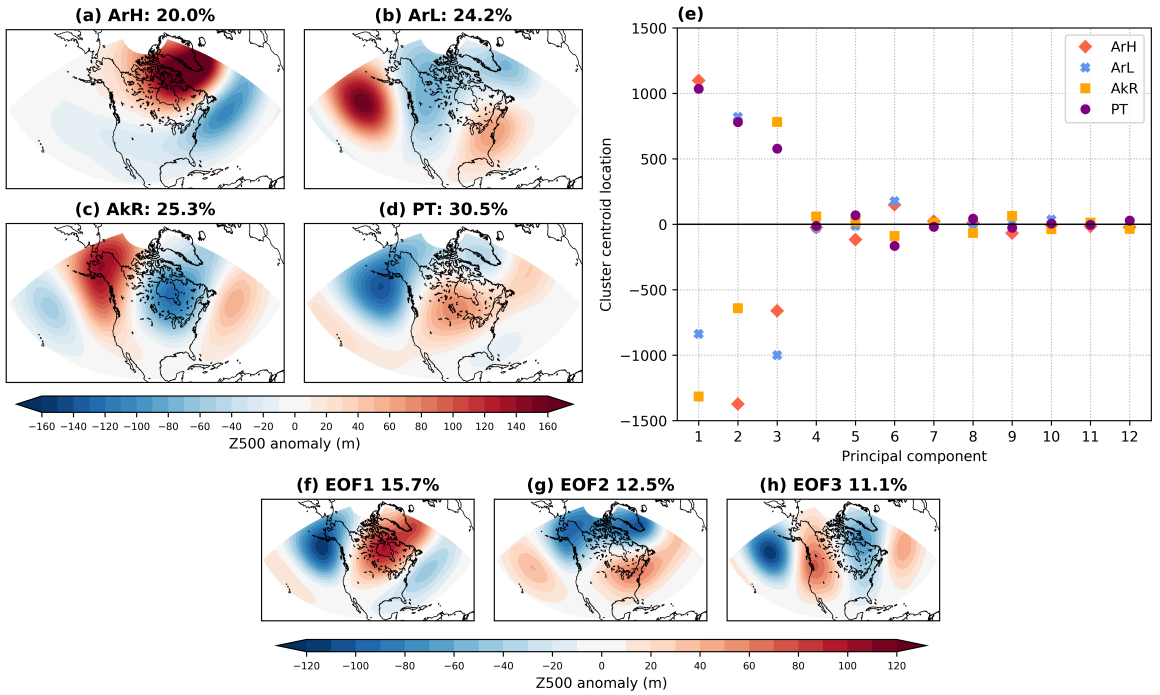


**Figure 5.2:** (a–d) Centroids of the four regimes, expressed as 500 hPa geopotential height anomalies with respect to daily 1979–2018 climatology in ERA-Interim, and the percent of days assigned to each regime in all December–March days in the period 1 January 1979–31 December 2018. (e) Co-ordinates of the regime centroids in 12-dimensional principal component space. (f–h) The leading three EOFs (multiplied by the square-root of the eigenvalue) of daily 500 hPa geopotential height anomalies in the domain 180-30°W 20-80°N, and the percent of total variability explained by each EOF.

The co-ordinates of the regime centroids in the leading 12 PCs are shown in Figure 5.2e. Only the leading three PCs have large contributions to the centroids; performing the same clustering analysis but retaining only the leading three PCs yields very similar patterns, with only 4% of days assigned to a different regime. Therefore, we now focus our analysis on these leading three EOFs.

Maps of the EOFs and the percent of the total variance explained are shown in Figure 5.2f–h. In total, these three EOFs explain close to 40% of the daily variance within the domain, and are well-separated according to the criterion of North et al. (1982). The sign of the EOFs is here defined such that a positive loading produces an anomalous trough in the northeast Pacific. EOF1 is similar to the PNA (Wallace and Gutzler, 1981) but slightly eastward-shifted, and is also similar to the pattern obtained by regressing the surface-based PNA onto upper-tropospheric height anomalies (e.g., Baldwin and Thompson, 2009).

It also bears some similarity to the Tropical-Northern Hemisphere (TNH) pattern (Mo and Livezey, 1986; Liang et al., 2017). There is also a meridional dipole in the North Atlantic in the eastern edge of the domain, reminiscent of NAO-like variability. EOF2 has a meridional dipole in Z500 anomalies, and thus some similarity to the surface NAM/AO, but with a center of action over Alaska which is not characteristic of the surface NAM (e.g., Thompson and Wallace, 1998). EOF3 is characterized by a wavenumber-2 pattern across the domain.

Comparison of these regional EOFs with the leading three EOFs for the Northern Hemisphere poleward of 20°N (Figures 5.12–5.14 in the Supplemental Material) shows a high degree of similarity in both the correlation of the PC timeseries (Pearson's correlation $r \geq 0.77$; $p < 0.05$) and spatially (area-weighted pattern correlation $\geq 0.87$ over the North American domain). We can therefore be confident that the leading three EOFs used in the clustering are regional manifestations of hemispheric variability, and that hemispheric variability is dominant in the smaller domain under consideration. The EOFs presented here – with the most NAM-like pattern in EOF2, while the leading EOF contains NAM/NAO and PNA-like characteristics – agrees well with the upper-tropospheric EOF analysis of Baldwin and Thompson (2009). For all three North American EOFs, the $e$-folding timescales of the PC timeseries are synoptic scale: i.e., 5–7 days.

To understand the relationship between regime occurrence and the lower-stratospheric vortex presented in Lee et al. (2019c), we examine the relationship between U100 and the leading EOFs which define the clusters. We perform linear regression between each PC timeseries and the contemporaneous U100 to see how changes in U100 may modulate the location of a point within the 3D-PC space and thus its regime attribution. The instantaneous relationship is used since we are considering the lower stratosphere as an upper boundary condition to the troposphere, with both a much longer memory (e.g., Baldwin et al., 2003) and greater predictability (Son et al., 2020); lagged relationships (not shown) reveal these coefficients are either effectively maximised at lag 0 or, considering

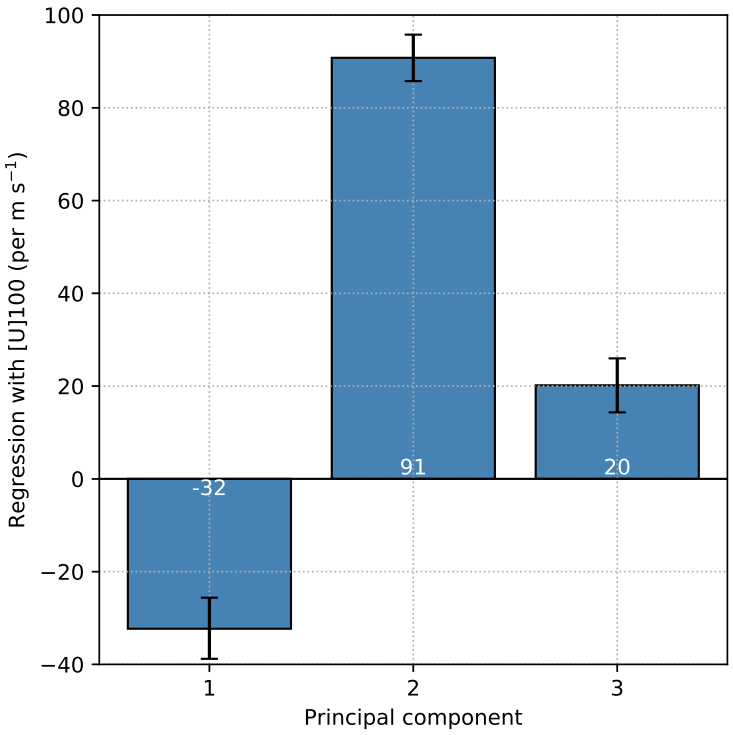


**Figure 5.3:** Linear regression coefficients between the 100 hPa 60°N zonal-mean zonal wind and the raw PC timeseries of the leading three EOFs, in all December–March days in ERA-Interim 1979-2018. Error bars indicate 95% confidence intervals obtained by bootstrapping with replacement (see Section 5.25.2.5 for details).

uncertainty, largely invariant for $\pm 7$ days (within the PC $e$-folding timescale). Some of this relationship may relate to the vertical extension of a primarily tropospheric zonal wind signature associated with these EOFs into the lower stratosphere. However, on subseasonal scales – well beyond tropospheric de-correlation timescales – this remains the component of the structure which is potentially predictable.

The regression coefficients are shown in Figure 5.3. Although the coefficients for all three EOFs are significantly different from zero, the linear relationship is 3-5 times stronger for EOF2. Similarly, the Pearson's correlations between U100 and PCs 1 and 3 are small ($r = -0.13$ and $0.10$, respectively), but moderate for PC2 ($r = 0.42$). Thus, the effect of the stratosphere in this 3D EOF space is mostly contained within EOF2, which is consistent with its annular-like spatial pattern and the height-dependent NAM results of Baldwin and Thompson (2009). The sign of the regression coefficients is such that a decrease in U100 is associated with an increase in Z500 in the vicinity of Greenland/the northern node of

the NAO, in agreement with the canonical response of the troposphere to a weakened stratospheric vortex.

## 5.4 Theory of Regime Transitions and the Stratosphere

In this section, we develop a theory of which regime transitions may be possible solely due to a stratospheric perturbation by jointly considering the linear relationship between U100 and the three PCs, and the location of the regimes within the space spanned by the three PCs. The theory can be interpreted as an idealized framework where all else is instantaneously equal and only the stratosphere is changed, retaining potential predictability arising from other tropospheric processes.

Using the regression coefficients between U100 and the PC timeseries, we define the stratospheric perturbation vector, $\boldsymbol{\beta}$. This vector represents the movement within the 3D-PC space arising from a perturbation to U100, $\Delta$U, that is explained by the linear relationship:

$$\boldsymbol{\beta} = \Delta U \begin{pmatrix} -32 \\ 91 \\ 20 \end{pmatrix} \tag{5.2}$$

$\boldsymbol{\beta}$ is not a function of the position within PC space and is thus constant for a given $\Delta$U. While the truncation to a 3D-PC space was earlier motivated by the coordinates of the regime centroids, the linear relationship between the leading three EOFs and U100 also accounts for nearly all of the linear relationship with Z500 (Figure 5.15).

The transition vector $\boldsymbol{\gamma}$ between two points (e.g., two cluster centroids) within this space is then defined as the respective distances between the co-ordinates in the three PCs:

$$\boldsymbol{\gamma} = \begin{pmatrix} \Delta PC_1 \\ \Delta PC_2 \\ \Delta PC_3 \end{pmatrix} \tag{5.3}$$

where $\Delta PC_k = PC_k(B) - PC_k(A)$ for the transition from point A to point B. Hence, inverse transitions have an equal but opposite transition vector: $\boldsymbol{\gamma}$(A, B) = $-\boldsymbol{\gamma}$(B, A).

The angle $\theta$ between $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ follows as:

$$\theta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \arccos\left(\frac{\boldsymbol{\beta} \cdot \boldsymbol{\gamma}}{\|\boldsymbol{\beta}\|\|\boldsymbol{\gamma}\|}\right) \tag{5.4}$$

where $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ denotes the Euclidean norm of a 3D vector $\mathbf{x}$.

We use this framework to model which regime transitions are possible solely with stratospheric forcing by considering whether the vectors $\boldsymbol{\beta}$ (either positive or negative) and $\boldsymbol{\gamma}$ point in a similar direction, known as "cosine similarity" (e.g., Han et al., 2012). If $\theta \geq 90^\circ$ ($\cos\theta \leq 0$), then no component of the regime transition or movement within the 3D-PC space can be explained by the linear relationship between the PCs and U100, since the contribution of $\boldsymbol{\beta}$ would be 0 (in the case of maximally dissimilar vectors, $\theta = 90^\circ$) or oppose $\boldsymbol{\gamma}$ ($\cos\theta < 0$). A smaller angle indicates the transition is more likely since the projection of $\boldsymbol{\beta}$ in the direction of $\boldsymbol{\gamma}$ is larger (as $\cos\theta$ is larger), thus requiring a smaller $\Delta$U. We focus on angles, rather than explicit distances, since the distances between regimes for any point are dependent on the initial location.

Figure 5.4 presents a 3D depiction (in the space spanned by the leading three EOFs) of $\boldsymbol{\beta}$ (both positive and negative; i.e., for a strengthening or weakening stratospheric vortex) applied to each regime centroid and the transition vector $\boldsymbol{\gamma}$ between the centroids. The regime centroids form a tetrahedron in this space. Some of the transition vectors lie closer to $\boldsymbol{\beta}$ than others owing to their relative locations within this space. For example, the positive $\boldsymbol{\beta}$ vector and the transition vector from the ArH to PT centroids are close, while the transition vectors from the AkR centroid are almost perpendicular to either sign of $\boldsymbol{\beta}$.

The angles between the centroid $\boldsymbol{\gamma}$ vectors and $\boldsymbol{\beta}$ are quantified in the protractor-like polar plots in Figure 5.5. The angles are expressed such that both positive and negative $\boldsymbol{\beta}$ are

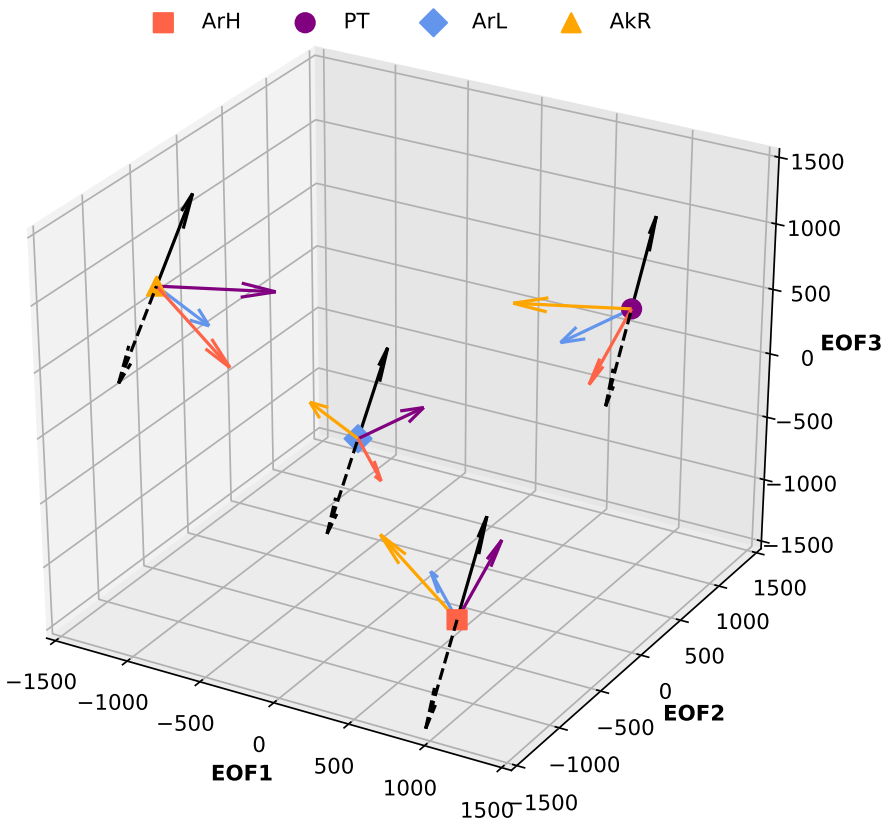


**Figure 5.4:** Visualization of the regimes in the space occupied by the leading three EOFs. Colored markers indicate the regime centroids. Colored arrows represent the transition vectors from each centroid to the other centroids, scaled to 0.25x. The black arrows show the stratospheric perturbation vector, scaled to a $\pm$10 m s$^{-1}$ perturbation (solid positive; dashed negative), which is the same at all points.

aligned with 0° (thus, the angle between each $\boldsymbol{\gamma}$ and $\boldsymbol{\beta} < 0$ is a reflection of that to $\boldsymbol{\beta} > 0$ about 90°). For a point starting at the ArH centroid (Figure 5.5a), there is substantial cosine similarity between $\boldsymbol{\beta} > 0$ and transition vectors to *all* other regimes (for all three, $\theta < 60°$). The similarity is strongest for the transition vectors to PT and ArL, which have approximately equal cosine similarity. The angles between $\boldsymbol{\beta} < 0$ and all three transition vectors are $> 90°$; thus, the linear theory does not allow a transition away from ArH given $\Delta$U$< 0$. Overall, ArH has the largest number of transition vectors with small angles/high cosine similarity. Equally, the minimum angle between either sign of $\boldsymbol{\beta}$ and any $\boldsymbol{\gamma}$ vector is between $\boldsymbol{\beta} < 0$ and transitions to ArH (Figure 5.5b–d). This is consistent with the observed probability of transitions into, and the persistence of, ArH/NAO– being the most sensitive of both the North American and North Atlantic regimes to the strength of U100 (Charlton-Perez et al., 2018; Lee et al., 2019c).

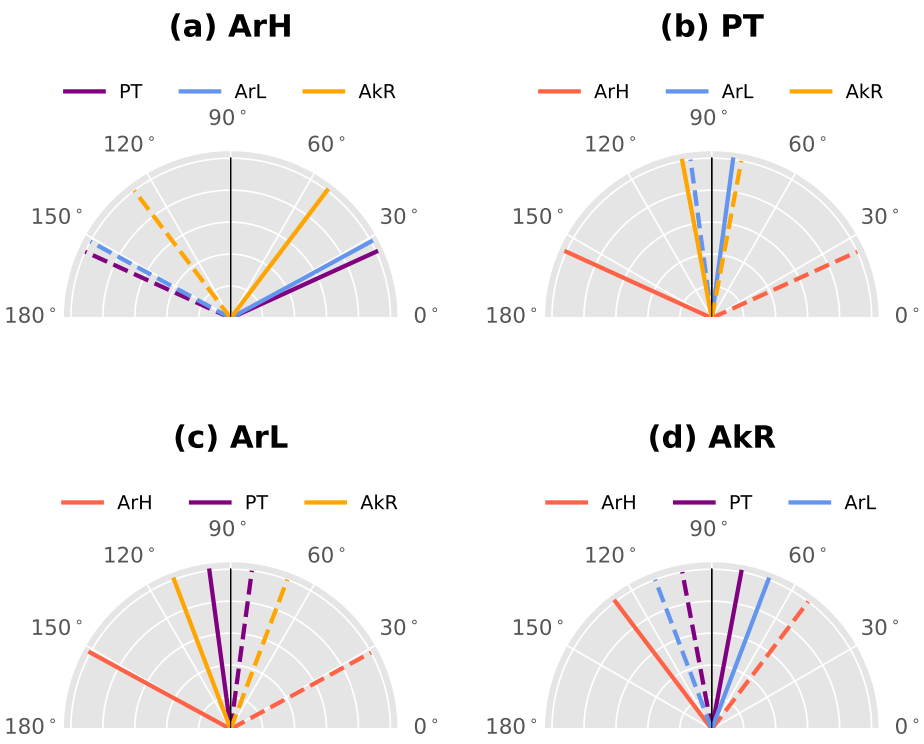


**Figure 5.5:** Polar plots showing angles between the stratospheric perturbation vector (solid positive; dashed negative) and the centroid transition vector for each of the four regimes in 3D EOF space, as visualized in Figure 5.4.

For the PT regime (Figure 5.5b), there is a small angle between the negative $\boldsymbol{\beta}$ vector and the transition vector to ArH (i.e., equal and opposite to the positive $\boldsymbol{\beta}$ and the transition from ArH to PT). While transitions are possible to both AkR with $\boldsymbol{\beta} < 0$, and to ArL with $\boldsymbol{\beta} > 0$, the angles are close to $90°$ suggesting that these are unlikely. Considering the ArL regime (Figure 5.5c), transitions to all three other regimes are possible with $\boldsymbol{\beta} < 0$. The smallest angle is to the ArH transition vector, while the angles to the PT and AkR transitions are large. No regime transitions from ArL are possible in this framework with $\Delta$U$> 0$. Lastly, the angles between the transition vectors and $\boldsymbol{\beta}$ are all relatively large for AkR (Figure 5.5c) – as previously suggested by the 3D depiction in Figure 5.4. For $\boldsymbol{\beta} < 0$, only a transition to ArH has an angle $< 90°$. Transitions to ArL and PT are possible with $\boldsymbol{\beta} > 0$, but the angles are relatively large and thus more unlikely.

We next extend our analysis beyond points initiating at the centroids and incorporate the effect of spread around the PC space spanned by each regime. First, we consider all the assigned regime days in ERA-Interim. The leading three PCs are then perturbed by $\boldsymbol{\beta}$ in the range $-30 \leq \Delta$U$\leq 30$ m s$^{-1}$, and subsequently re-assigned to a regime by minimum Euclidean distance. The maximum magnitude of $\Delta$U is chosen here to be close to the maximum observed variability in U100; the largest U100 errors in individual CY43R3 ensemble members are close to $\pm 20$ m s$^{-1}$. Note that in reality, the tropospheric response may be larger for a smaller $\Delta$U as a consequence of the linear framework.

Figure 5.6 depicts the conditional probability, for each initial regime, of either remaining in the same regime or transitioning to each of the other regimes for each $\Delta$U. Only those transition pathways with $\theta < 90^{\circ}$ occur, and the relative likelihood manifests the degree of similarity (i.e., the angle) between $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$. There are no transitions away from ArH for $\Delta$U$< 0$ (Figure 5.6a) or away from ArL for $\Delta$U$> 0$ (Figure 5.6c). For $\Delta$U$< 0$, the dominant transition for all regimes is to ArH. For $\Delta$U$> 0$, transitions from ArH to PT dominate (Figure 5.6a) while transitions to ArL dominate for AkR and PT (Figure 5.6b,d). Transitioning into AkR from any other regime is unlikely even for large $|\Delta$U$|$, while transitioning out of AkR is the least likely for any of the regimes where a transition pathway exists (despite its unique approximately equal sensitivity for either sign of $\Delta$U). Although not explicitly shown, there is also evidence of multiple transitions occurring as $|\Delta$U$|$ increases. For example, the probability of transitioning into AkR from each of the other regimes reaches a peak for $|\Delta$U$|$ between 10–20 m s$^{-1}$ before declining.

As a general diagnostic of the sensitivity of each initial regime state to a lower-stratospheric perturbation, we can consider the probability of transitioning out of the regime for $\Delta$U $= \pm 10$ m s$^{-1}$ (approximately equal to the maximum week 3–4 ensemble-mean U100 error magnitude in CY43R3 hindcasts). For $\Delta$U $= 10$ m s$^{-1}$, 58% of ArH days transition into a new regime, while only 17% of AkR days and 6% of PT days do so. For $\Delta$U $= -10$ m s$^{-1}$, the sensitivity of PT and ArL is approximately equal, with 39% of PT and 38% of ArL

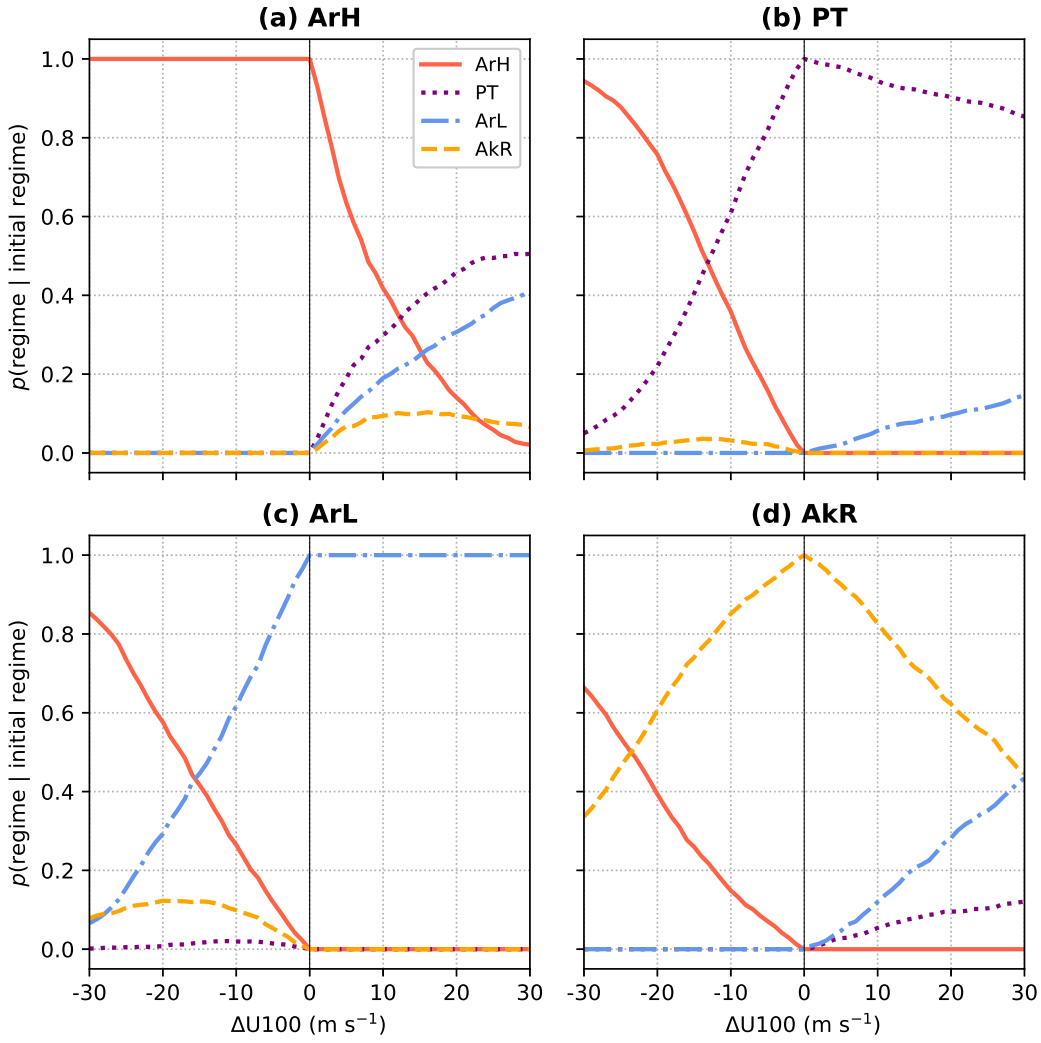


**Figure 5.6:** Given each initial regime (a–d), the conditional probability of either remaining in the same regime or transitioning to each of the other regimes, when all days assigned to each regime in ERA-Interim are perturbed by the stratospheric perturbation vector in the range $-30 \leq \Delta U \leq 30$ m s$^{-1}$.

days transitioning into a new regime. Only 15% of AkR days transition into a new regime.

Overall, the results presented in Figures 5.4–5.6 are in agreement with the observed differences in regime occurrence in strong and weak stratospheric vortex states in Lee et al. (2019c). The theory also gives results consistent with the relationship between the regimes (particularly ArH and ArL) and the concurrent NAO index (Figure 5.16), given the strong modulation of the NAO by the stratosphere. Further, the proposed framework yields insight into specific regime transitions under different vortex states which are not limited by the observational sample size. In summary:

- $\Delta$U$< 0$ moves the majority of points within PC space toward *only* ArH, consistent with this regime being the *only* one more likely under weak vortex conditions.

- $\Delta$U$> 0$ does little to changing the regime assignment for days initially assigned to ArL or PT, while these are favored transitions for initial ArH and AkR states. This is consistent with ArL and PT being more likely under strong vortex conditions.

- Very large $\Delta$U is required to shift toward and away from AkR, with a similar proportion of transitions resulting from both positive and negative perturbations. This behavior is consistent with the observed statistically equal occurrence of this regime in strong and weak vortex states.

These conclusions are highly idealized, requiring both a perfectly linear response and the sole (or dominant) change being to U100. It is also possible that $\beta$ may be sensitive to the initial position within PC space. However, the corroboration with observations suggests the potential use of this framework in interpreting the regime response to changes and uncertainty in the stratosphere on subseasonal timescales. The analysis in the next section considers whether imposing stratospheric relaxation yields a tropospheric response consistent with this simple but novel linear theory.

## 5.5 Model Experiments

In analyzing the results of the relaxation experiments, we seek to answer the following two questions:

- What is the effect of stratospheric relaxation on regime forecast accuracy in these cases?

- Regardless of the forecast accuracy, is the change in the forecast consistent with the linear theory in Section 5.4?

### 5.5.1 Regime predictions

A comparison between the weekly-mean regimes in the CTR and RLX ensembles, for weeks 3 and 4, is shown in Figure 5.7. The improvement in the total number of ensemble members with a correctly-assigned weekly-mean regime is modest: 13% in week 3 and 15% in week 4. Therefore (recalling that these cases were selected as particularly poor forecasts), the overall fraction of correctly assigned regimes remains low in the RLX experiment: 40% in week 3 and 25% in week 4. Any improvement is also case dependent. The greatest improvement in week 3 is in the 11 December 2001 case (7 more members correctly assigned to ArH), and in the 29 January 1998 case (5 more members correctly assigned to PT). The latter was a case with a very large U100 error (c.f. Table 5.1). In several cases, there is a *decrease* in the number of correctly-assigned ensemble members. Thus, constraining the stratospheric state is not enough to fix these regime bust cases – which may be unsurprising given that only a selection of these cases have large stratospheric errors, while all have largely inaccurate regime predictions. This result indicates that the stratospheric state should not be viewed as exerting simple *control* on the subseasonal tropospheric flow over North America.

Figure 5.7 also shows that there are changes to the number of ensemble members assigned to the *incorrect* regimes, regardless of whether there is a change to the number assigned to the correct regime. On a member-by-member basis, 34% and 57% of the total ensemble members in weeks 3 and 4 respectively are assigned to a different regime in the RLX experiments. Thus, by week 4, the stratospheric nudging has shifted the majority of ensemble members into a new regime – suggesting significant movement within the PC space in which the regimes are assigned. For example, in week 4 of the 11 December 2001 case, there is no increase in the number of members correctly assigned to PT, but there is a gain of 8 ensemble members assigned to AkR (with ArH and ArL losing 4 members each). While a full case-by-case analysis may yield further specific insight, it is beyond the scope of this study; we instead focus on the general results across this set of forecasts.

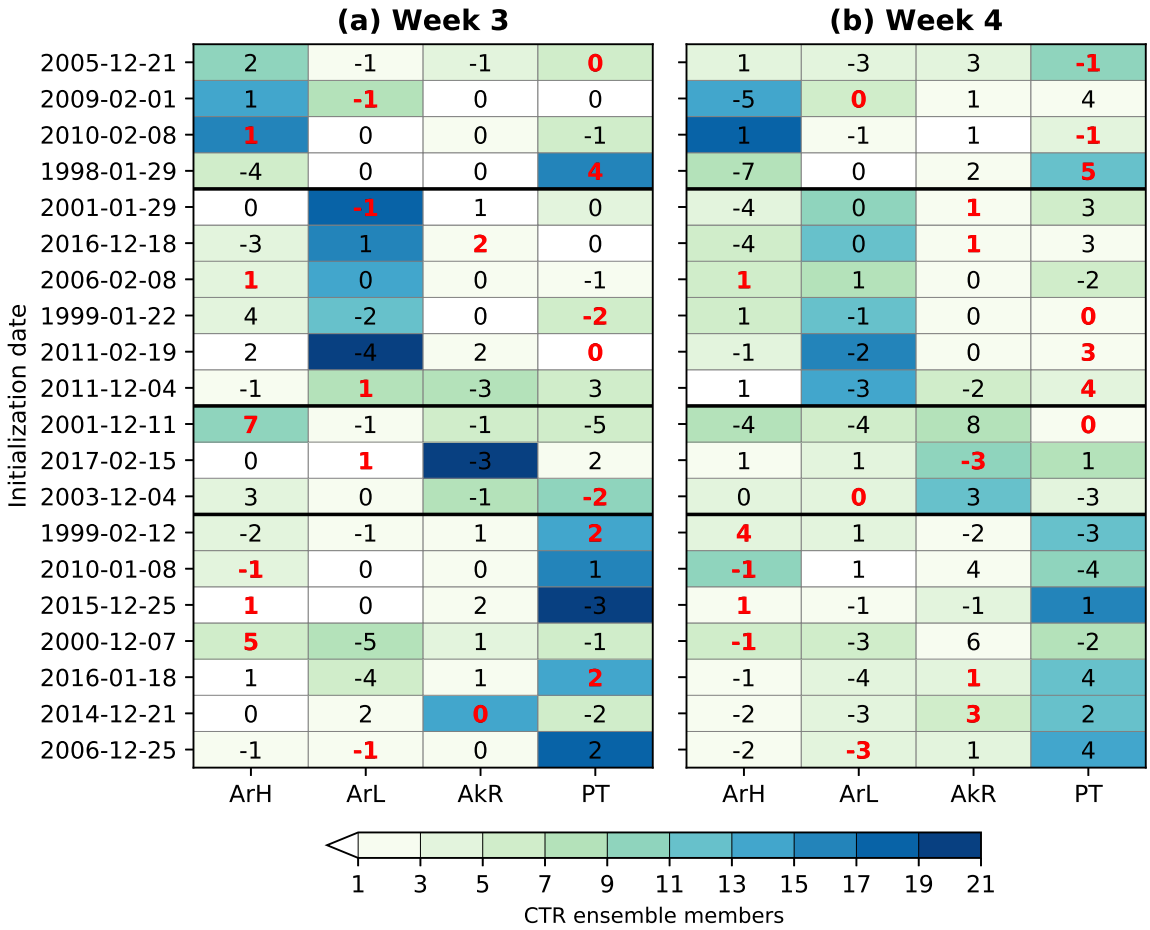


**Figure 5.7:** (values) The change in the number of ensemble members assigned to each weekly-mean regime for (a) week 3 and (b) week 4 (RLX-CTR). (colors) The number of ensemble members assigned to each regime in the CTR ensemble. The ERA-Interim regime is shown in bold red font. Grouping is as in Table 5.1.

### 5.5.2 Error reduction in PC space

Despite the small and case-dependent regime improvement, for almost all cases the mean Euclidean distance error of the ensemble in 3D-PC space is reduced (Figure 5.8a). This diagnostic is useful because it incorporates changes to forecasts which maintain the same regime attribution and is proportional to the root-mean square error (RMSE) of the Z500 field reconstructed from the leading 3 EOFs (see Supplemental Material; note that because non-normalized PCs are used, the total error on subseasonal timescales is dominated by the EOFs with the largest eigenvalues). Hence, in the space in which regimes are assigned, the RLX forecasts are almost entirely closer to the verification. The improvement is maximized in week 3 (median 14%), with only 2 cases showing an increase in error (21

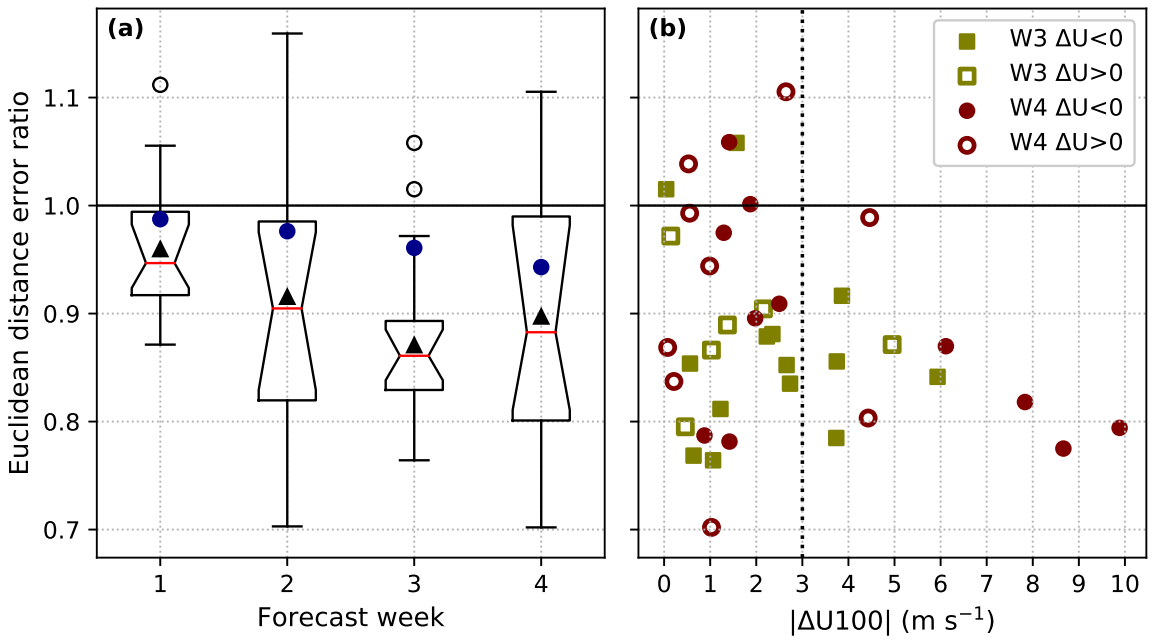


**Figure 5.8:** (a) Boxplots of the ratio between the ensemble-mean Euclidean distance error in 3D PC-space between the weekly-averaged RLX and CTR ensembles for the 20 cases. Red lines denote the median, and notches show 95% confidence intervals obtained by 10,000 bootstrap re-samples (with replacement). Black triangles denote the mean. Blue circles represent the average ratio obtained by statistically perturbing the CTR PCs by the stratospheric perturbation vector multiplied by the change in U100 between the CTR and RLX ensembles. Whiskers extend to 1.5 times the interquartile range or extremes (whichever is smaller); outliers shown as open circles. (b) Scatter plot of the week 3 (green squares) and week 4 (maroon circles) error ratio against the magnitude of the ensemble-mean change in U100 between CTR and RLX.

December 2005 and 8 February 2010, both of which had negligible week 3 U100 errors in the CTR run). The median improvement in week 4 is 12%, but with much greater spread than week 3. There was a 30% improvement in a single case (21 December 2014), while four cases show no change or increased error (7 December 2000, 11 December 2001, 8 February 2010, and 15 February 2017).

Also shown in Figure 5.8a is the mean change in Euclidean distance error obtained by perturbing the PCs of the CTR ensemble by $\beta$ multiplied by $\Delta$U between the CTR and RLX experiments. This shows that a simple statistical nudge of the PCs using the known linear relationships also yields an error reduction of on average $\sim$50% of that obtained by running the full dynamical relaxation experiment. Thus, a substantial component of the dynamical effect of imposing a different stratospheric state on

these EOFs can be explained by the observed linear relationship between the PCs and U100.

To understand whether larger stratospheric forcing yields larger error reduction, Figure 5.8b shows the case-by-case change in ensemble-mean Euclidean distance error against the magnitude of the U100 change between the CTR and RLX experiments for weeks 3 and 4. There is no immediately clear relationship, with the greatest error reduction occurring with a U100 change of only 1 m $s^{-1}$ while the largest error increase occurs with a U100 change of 2.6 m $s^{-1}$ (8 February 2010). The large relative error reduction for small $\Delta$U suggests a potential role of zonally-asymmetric corrections or other changes to the vortex which do not project strongly onto U100 (and thus fall outside the framework proposed here). However, across this set of 20 cases, for $\Delta$U exceeding 3 m $s^{-1}$, there is a systematic error reduction. We revisit this apparent threshold in the analysis below.

### 5.5.3 Movement within PC space

We now investigate whether the movement of the forecasts within 3D-PC space is consistent with what might be expected from the linear perturbation theory established in Section 5.4. For this analysis, we analyze three vectors and three different angles within PC space. Figure 5.9 shows a schematic of this approach. The vectors are defined as follows:

- CTR-ERA: the vector between the CTR forecast and the verification from ERA-Interim (i.e., the error in the CTR forecast).
- CTR-RLX: the vector between the CTR and RLX forecasts.
- CTR-STAT: the vector between the CTR forecast and the CTR forecast statistically perturbed by $\beta$ multiplied by $\Delta$U between CTR and RLX ensembles (STAT).

Then, the size of the three angles can be used to answer the following questions:

- $\theta_1 = \theta$(CTR-ERA, CTR-RLX): Does stratospheric relaxation move the CTR forecast toward the verification?

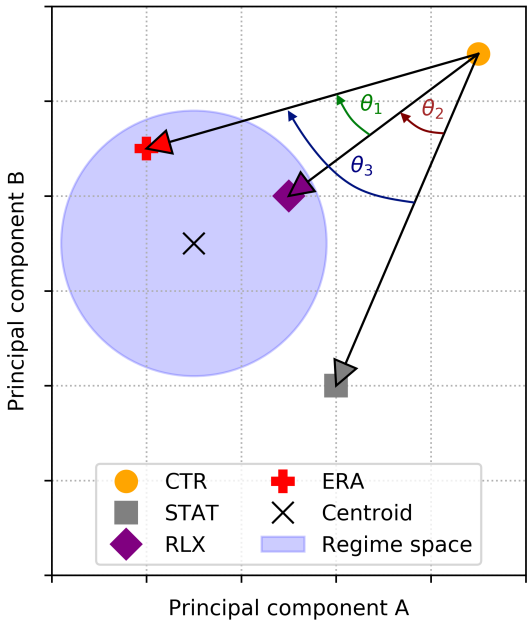


**Figure 5.9:** Schematic of the angle-based approach (here in a 2D PC space). There are three vectors: the vector from the control forecast to the ERA-I verification (CTR-ERA, red); the vector from the control forecast to the relaxed forecast (CTR-RLX, purple); and the stratospheric perturbation vector to the statistically nudged forecast (CTR-STAT, gray). $\theta_1$ denotes the angle between CTR-ERA and CTR-RLX, $\theta_2$ between CTR-RLX and CTR-STAT, and $\theta_3$ between CTR-ERA and CTR-STAT.

- $\theta_2 = \theta$(CTR-RLX, CTR-STAT): Does stratospheric relaxation move the CTR forecast in the direction expected from $\beta$?
- $\theta_3 = \theta$(CTR-ERA, CTR-STAT): Does statistical nudging by $\beta$ move the CTR forecast toward the verification?

A scatter of the week 3 and week 4 angles versus the magnitude of $\Delta$U between the CTR and RLX experiments is shown in Figure 5.10. To focus on the overall shift of the ensemble in the relaxed experiments, and since $\beta$ is defined from linear best-fit regression coefficients, we perform this analysis on the perturbations to the PCs and U100 averaged across the ensemble. Nevertheless, a similar overall story emerges when considering the results across all individual ensemble members. Figure 5.10a shows that in the majority of cases and in both weeks 3 and 4, the stratospheric relaxation generally moved the predictions toward the verification. Only 2 cases in week 3 and 6 cases in week 4 do not exhibit any similarity (i.e., $\theta > 90°$). These results are consistent with the reduction in Euclidean distance error and its relationship with the magnitude of $\Delta$U (Figure 5.8).

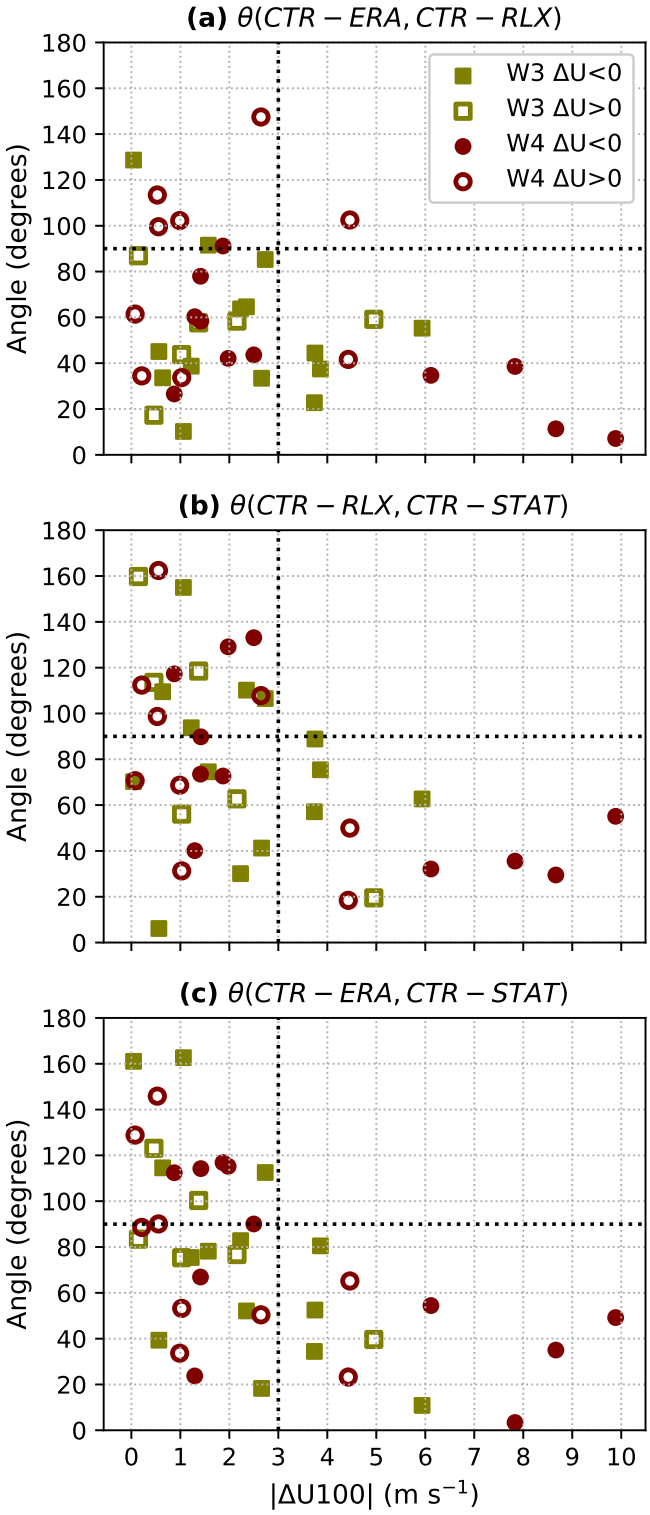


**Figure 5.10:** Scatter plots of the magnitude of the ensemble-mean weekly-mean U100 change between the CTR and RLX experiments, versus the angle between (a) CTR-ERA and CTR-RLX ($\theta_1$), (b) CTR-RLX and CTR-STAT ($\theta_2$), and (c) CTR-ERA and CTR-STAT ($\theta_3$), in 3D-PC space.

Figure 5.10b assesses whether the stratospheric perturbation vector outlined in Section 5.4 is a good representation of the effect of a dynamically-applied stratospheric perturbation. For $|\Delta U| <\sim 3$ m s$^{-1}$, the points are scattered across almost the full range of angles, indicating no clear relationship between the linear theory and the movement of these forecasts in PC space. However, although the sample is smaller, for $|\Delta U| >\sim 3$ m s$^{-1}$, the angles are systematically much smaller than $90^\circ$ – especially for week 4 forecasts which feature larger $\Delta U$. Hence, we conclude that on average, these forecasts moved in PC space in the general direction expected from the linear theory.

Finally, Figure 5.10c assesses whether the simple statistical perturbation moves the CTR forecast toward the verification without running a full dynamical experiment (c.f. Figure 5.10a). As in (b), but unlike in (a), there is no clear evidence of vector similarity for small $\Delta U$, but there is evidence of a systematic shift for $\Delta U$ exceeding $\sim 3$ m s$^{-1}$ in magnitude. As a result, for larger U100 errors the tropospheric forecast can be partially corrected statistically (as indicated by Figure 5.8a), but there is evidently additional gain from a dynamically-corrected stratosphere even for small $\Delta U$.

The 3 m s$^{-1}$ threshold is most apparent for angles involving $\beta$, though there is some suggestion for the behavior of the RLX experiment (in terms of both angles and Euclidean distance error). It is not clear why 3 m s$^{-1}$ should be a threshold; it may be related to the signal magnitude required to emerge above the typical ensemble-mean variability. Across the CY43R3 hindcasts, 3 m s$^{-1}$ is approximately two-thirds of the standard deviation of the ensemble-mean U100 in weeks 3-4 ($\sim 4.5$ m s$^{-1}$), although these are not directly comparable owing to the smaller hindcast ensemble size. As mentioned in Section 5.2, 3 m s$^{-1}$ is also approximately the upper tercile of the ensemble-mean week 3–4 U100 error magnitude in the CY43R3 hindcasts, and so errors of this magnitude are a reasonably frequent occurrence.

In week 4 (when $\Delta$U is generally largest), the magnitude of the correlations between the ensemble-mean change in the PCs and the ensemble-mean $\Delta$U from CTR to RLX (and thus the individual components of $\beta$) are maximized. These correlations are largest for EOF2 ($r = 0.60, p < 0.05$) and EOF3 ($r = 0.48, p < 0.05$) but the correlation is small and insignificant for EOF1 ($r = -0.19, p = 0.40$; although it is similar to that in ERA-Interim). Hence, across the range of cases studied here, the response of EOF1 to stratospheric perturbations is not well approximated by the linear theory. This may be due to non-linearity, or that the relationship between the EOF and U100 is not causal (recalling the similarity between the EOF and patterns related to tropical forcing). Sample size may be an issue, given that the small expected response in EOF1. There may also be limitations in the representation of stratosphere-troposphere coupling in the model, such as the overestimation of the NAO response reported by Kolstad et al. (2020) using a similar but more recent ECMWF forecast model (CY45R1). The relatively low vertical resolution employed here, particularly in the upper troposphere and lower stratosphere, may also have limited the downward coupling and forecast improvement arising from the stratosphere (Kawatani et al., 2019; Domeisen et al., 2020c).

## 5.6 Summary and Conclusions

Understanding and exploiting stratospheric variability is a key way in which the accuracy and usefulness of S2S forecasts and the fidelity of stratosphere-troposphere coupling within models can be increased. In this study, we investigated how perturbations to the strength of the lower-stratospheric polar vortex can influence North American weather regime predictions. Our novel technique involved jointly considering the linear relationship between the vortex strength and the leading EOFs that contribute to the regimes (Figure 5.3), and the relative location of the regimes within the EOF-space (Figure 5.4). We used an angle-based approach to quantify which transitions are likely to occur (using cosine similarity) for a given regime and stratospheric perturbation (Figure 5.5). These results agree with the observed changes in regime occurrence under different stratospheric vortex states reported in Lee et al. (2019c) and provide an explanation for the regime

behavior. However, both the regimes framework and EOFs are defined primarily from a mathematical, rather than physical, standpoint, and therefore the results of this work largely focus on the mathematics of regime attribution.

We then performed a set of stratospheric relaxation model experiments, selecting 20 cases from the ECMWF hindcasts in which there was strong, coherent ensemble support for an incorrect regime to dominate during weeks 3-4. The majority (14) of these cases featured large (upper tercile) U100 errors in either week 3 or 4 or both, suggesting a link to the erroneous tropospheric forecasts. We found that the stratospheric relaxation is not enough to eliminate the regime errors, but the relaxation does lead to shifts in the ensemble distribution of the regimes within each forecast indicating substantial movement within PC-space (Figure 5.7). The results also showed that there is an overall 10-20% improvement in the accuracy of the forecasts in terms of Euclidean distance error/RMSE, which is most consistent in cases where the stratospheric error is larger (Figure 5.8).

Analysis of the transition vectors between the CTR and RLX forecasts in PC-space provided insight into the effect of stratospheric relaxation in the space in which regimes are assigned. The results (Figure 5.10) illustrated that stratospheric relaxation generally moved the forecasts toward the ERA-Interim verification *and* in the direction of that expected from the linear theory, while statistically nudging the CTR ensembles by the corresponding stratospheric perturbation vector also generally moves the forecasts toward the verification. For $|\Delta \mathrm{U}| > \sim 3$ m s$^{-1}$, this effect was particularly pronounced. Therefore, our results provide evidence that – all else being equal – the average shift of an ensemble on the subseasonal timescale to a lower-stratospheric zonal wind perturbation exceeding $\sim$3 m s$^{-1}$ is broadly predictable.

Consequently, the experiments also support the proposed linear theory of which regime transitions may be possible solely because of changes to the stratospheric state (Figure 5.6). We therefore propose that this approach can be used to identify, *a priori*, the

regime forecast-verification scenarios in which lower-stratospheric errors are more likely to have played a substantial role – and thus toward understanding the overall contribution to subseasonal North American weather regime forecast accuracy. Further, the method could be used in real-time (qualitatively or quantitatively), to interpret links between regime forecast uncertainty and stratospheric forecast uncertainty. This approach is likely to be most useful 2-3 weeks before SSWs or strong vortex events, when abrupt forecast shifts (e.g., Lee et al., 2019b) are more likely due to the current predictability limit of these phenomena (Domeisen et al., 2020b). It may also be plausible to use the technique on-the-fly to linearly impose alternate regime "storylines" arising from a different stratospheric evolution without running additional dynamical forecasts.

Moreover, the dominantly linear and apparently generic response to the lower-stratospheric forcing on these timescales is somewhat similar to the long-lag response following SSWs in the model experiments of White et al. (2020). Our results also further emphasize the importance of changes to the *lower* stratospheric vortex (as in e.g., Hitchcock and Simpson (2014)); hence, predicting the stratospheric penetration depth of anomalous vortex states is likely to be key to gaining subseasonal predictive skill from the stratosphere.

The framework presented here suggests that the resultant anomalous tropospheric flow pattern (at least over North America) following an abrupt change to the lower-stratospheric vortex strength generally depends on the prior state of the troposphere, with certain prior states more likely to result in a concurrent or canonical regime transition. Such an approach is not a new idea (Gerber et al., 2009), but as a result, potential gains in subseasonal regime predictability from the stratosphere may be limited by the otherwise inherent unpredictability of the troposphere. This limitation is consistent with the regime forecasts remaining largely inaccurate even in cases where large lower-stratospheric errors were corrected.

Employing a stronger stratospheric nudging in the model experiments presented in this

paper may produce greater improvement in the regime forecasts. On the other hand, constraining the prediction too strongly would exceed a realistically-achievable level of stratospheric forecast accuracy on these scales. Model experiments with a greater horizontal and vertical resolution may also yield better results, with evidence supporting a link between increased resolution and better representation of modes of variability in S2S models (Quinting and Vitart, 2019; Lee et al., 2020a) and downward stratosphere-troposphere coupling (Kawatani et al., 2019). The 60-level model version used in the experiments performed here (limited by the resolution of ERA-Interim) is coarser than the 91-level model used operationally, suggesting there is scope for the impact of an improved stratospheric forecast to be greater in the operational model (and thus lead to more regime shifts).

Further, we have exclusively considered the effect of changes to the strength of the lower-stratospheric polar vortex defined through the zonal-mean zonal wind at 100 hPa and 60°N. A more complex analysis may incorporate the effects of wave propagation (Perlwitz and Harnik, 2003; Kodera et al., 2008), vortex morphology (Cohen et al., 2021), or the representation of ozone chemistry (e.g., Oehrlein et al., 2020). While the use of zonal-mean quantities is motivated by annular modes, the approach can mask important sub-hemispheric variability such as localized wave reflection (e.g., Matthias and Kretschmer, 2020).

A case-by-case analysis of the dynamics involved, including the interplay between stratospheric errors and other leading sources of subseasonal prediction (e.g., the Madden-Julian Oscillation, which can act together with stratospheric variability; Barnes et al., 2019; Green and Furtado, 2019) is a potentially fruitful avenue of future work. Moreover, using the proposed angular diagnostic to assess the tropospheric regime response to stratospheric perturbations across a much larger set of simulations (and in different geographic regions) will aid in understanding the robustness of the results of this study.

## Acknowledgments

S.H.L. was funded by the Natural Environment Research Council (NERC) via the SCENARIO Doctoral Training Partnership (NE/L002566/1). S.J.W was supported by the National Centre for Atmospheric Science, a NERC collaborative centre. This work is based on S2S data. S2S is a joint initiative of the World Weather Research Programme (WWRP) and the World Climate Research Programme (WCRP). The original S2S database is hosted at ECMWF as an extension of the TIGGE database. The authors thank Glenn Carver and Marcus Koehler at ECMWF for their support with OpenIFS and preparation of the initial conditions and relaxation data and two anonymous reviewers for their helpful comments on the manuscript.

## Data Availability Statement

The ERA-Interim reanalysis and ECMWF hindcasts are available from the ECMWF website: `https://apps.ecmwf.int/datasets/`. OpenIFS experiment data used in this study are available at `https://doi.org/10.5281/zenodo.4818044`. EOF and $k$-means clustering analysis were performed using the freely-available Python packages *eofs* (Dawson, 2016) and *scikit-learn* (Pedregosa et al., 2011) respectively.

## 5.7 Supplemental Material

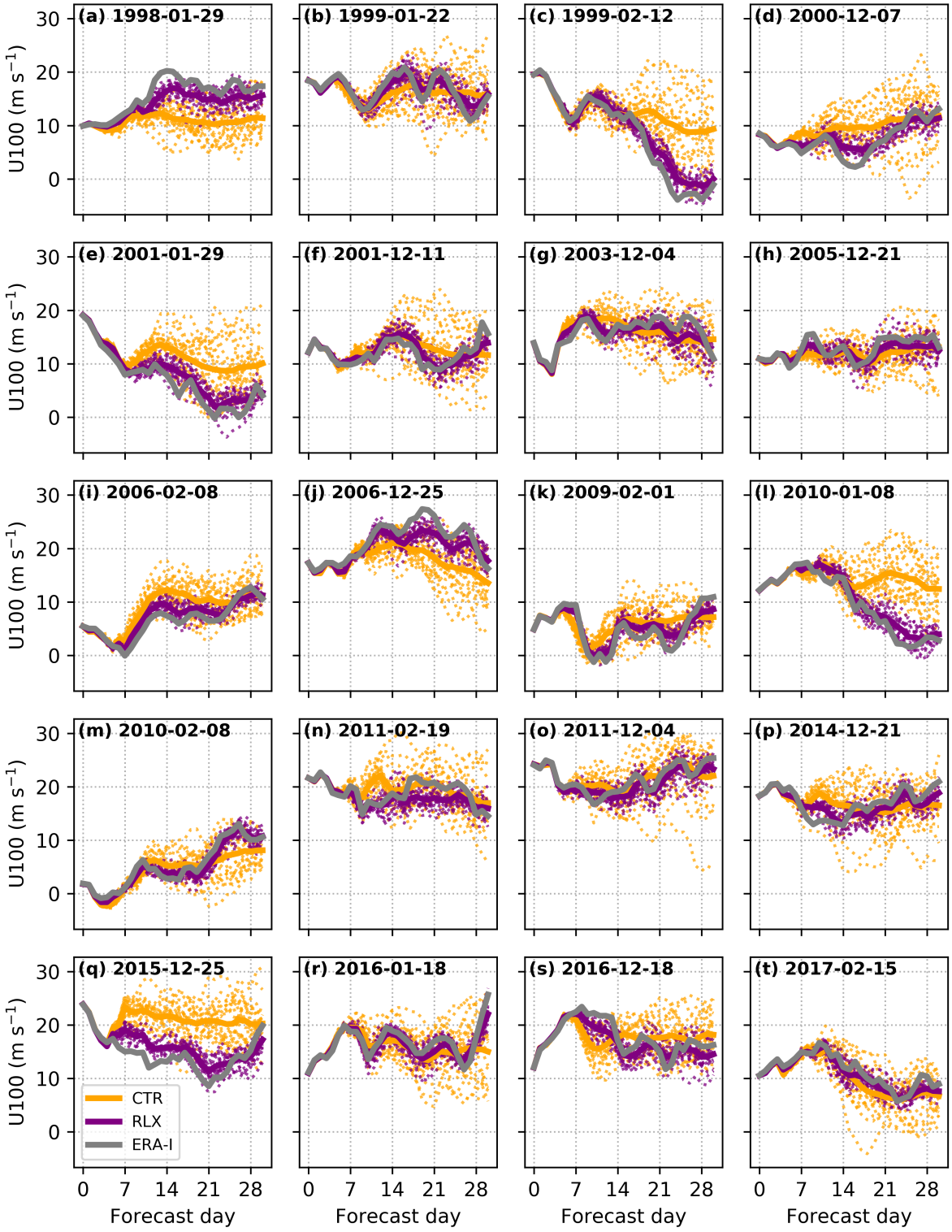


**Figure 5.11:** Time series of 100 hPa 60°N zonal-mean zonal wind from the 20 OpenIFS experiments showing: (orange) control (CTR) forecasts, (purple) relaxed (RLX) forecasts, and (gray) ERA-Interim verification. Solid lines denote the ensemble mean.

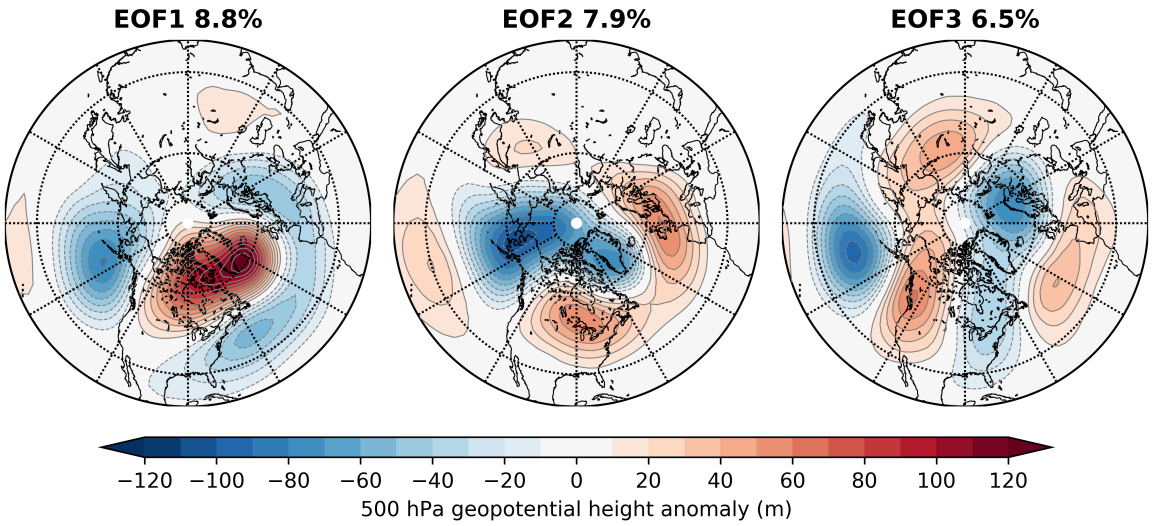


**Figure 5.12:** The leading 3 EOFs (multiplied by the square-root of the eigenvalue; units m) of 500 hPa geopotential height anomalies poleward of 20°N in 00Z ERA-Interim data for December–March from 1 January 1979–31 December 2018. The percent of total variance explained by each EOF is also shown.

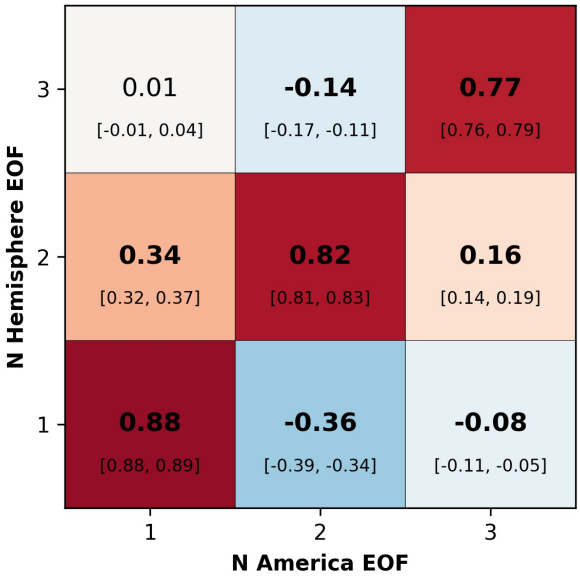


**Figure 5.13:** Pearson's correlation coefficients between the PC timeseries of the leading three North American EOFs and the leading three Northern Hemisphere EOFs, in 00Z ERA-Interim data for December–March from 1 January 1979–31 December 2018. Bold indicates the correlation is significantly different from zero at the 95% confidence level, according to 10,000 bootstrap re-samples with replacement (confidence intervals shown in square brackets).

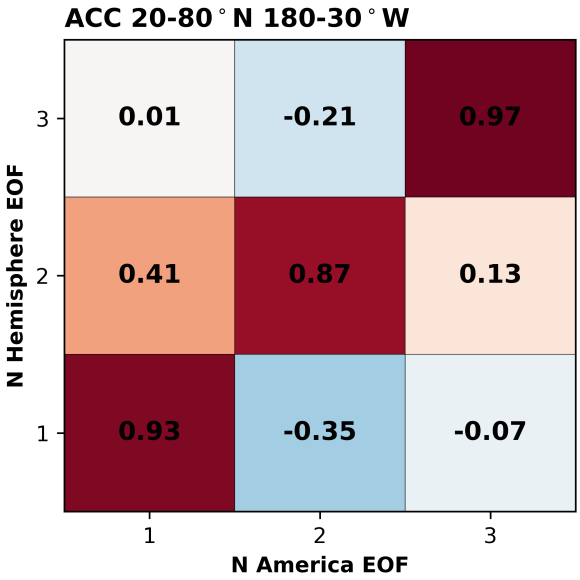


**Figure 5.14:** Cosine-latitude weighted pattern correlation coefficients between the North American and Northern Hemisphere EOFs over the North American EOF domain 20-80°N, 180-30°W.

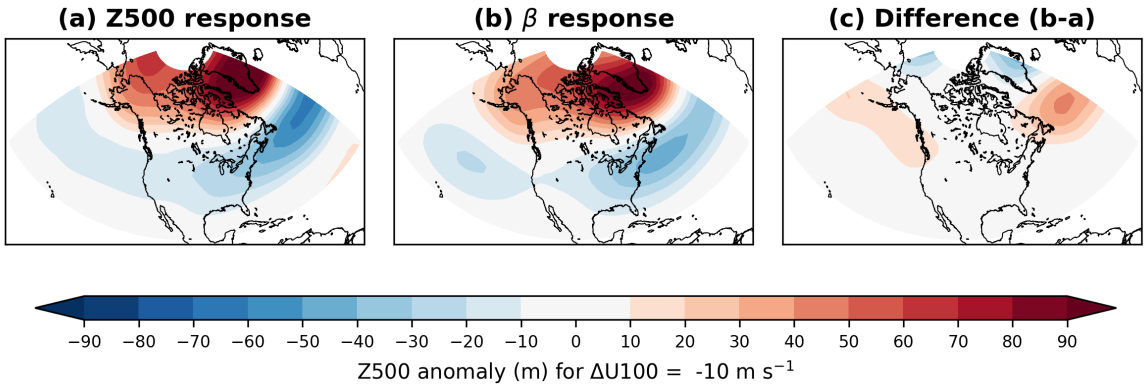


**Figure 5.15:** (a) The linear response of Z500 anomalies to a U100 perturbation of -10 m $s^{-1}$ as inferred from linear regression at each grid point between U100 and Z500 anomalies for December–March in ERA-Interim 1 January 1979–31 December 2018. (b) The response of Z500 inferred from $\beta$ (the linear regression coefficients of the leading three EOFs with U100) computed as the field reconstructed from the leading three EOFs multiplied by the components of $\beta$. (c) The residual linear response not explained by $\beta$, shown as (b) - (a).

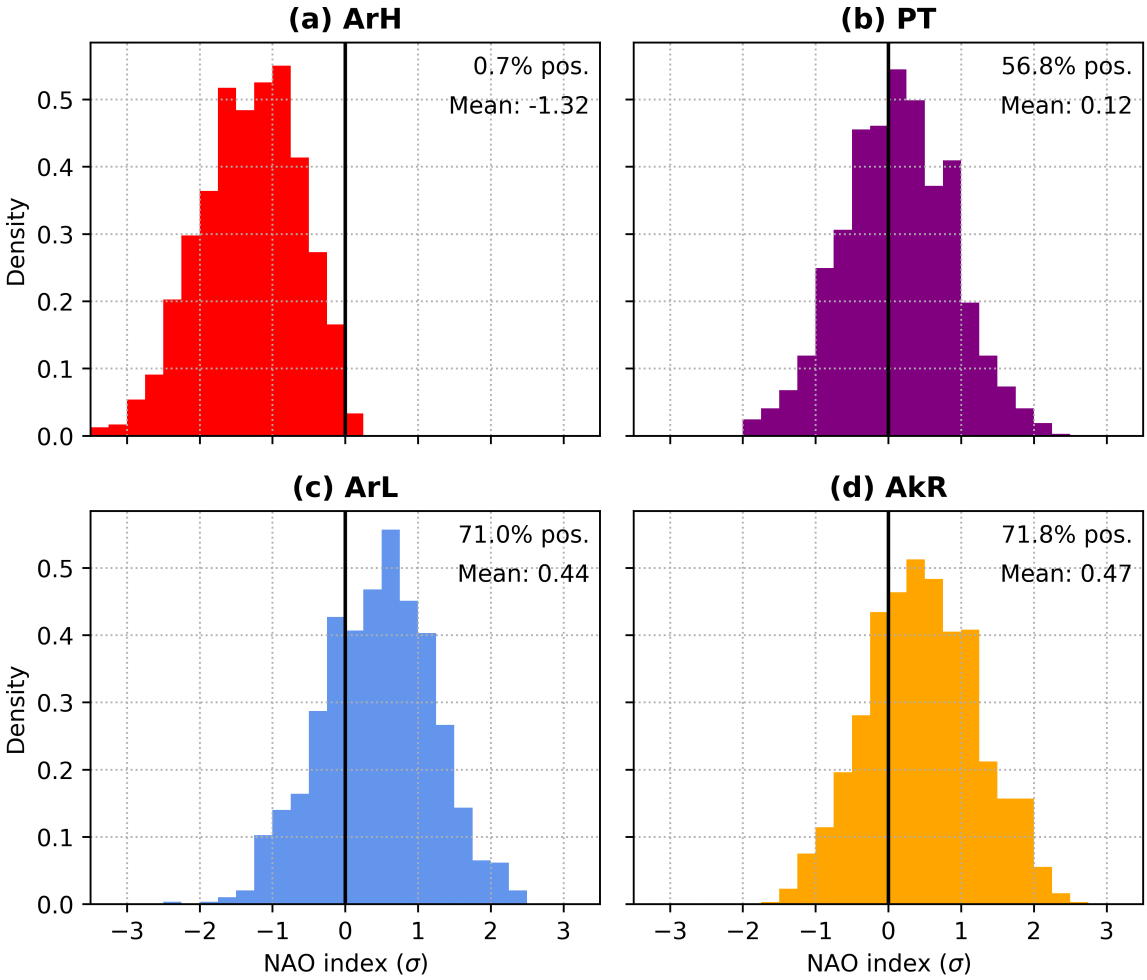


**Figure 5.16:** Histograms of the standardized NAO index (defined as the leading EOF of daily DJFM Z500 anomalies in the domain 20-80°N, 90°W-40°E over 1 January 1979–31 December 2018) for all days assigned to each regime. Also shown are the % of days with a positive NAO and the mean NAO value for each regime.

**Derivation of the relationship between the RMSE of a full field (or that reconstructed from a subset of EOFs) and the Euclidean distance error in PC space**

Let $i, j$ be grid points in latitude and longitude respectively, then $N = i \times j$ is the total number of grid points in the domain. Let $f_n$ be the forecast and $o_n$ be the observation at the $n$th grid point.

Then, the (non area-weighted) domain RMSE is:

$$RMSE = \left\{ \frac{1}{N} \sum_{n=1}^{N} (f_n - o_n)^2 \right\}^{1/2} \tag{5.5}$$

In $K$-dimensional principal component (PC) space, the Euclidean distance error is:

$$d = \left\{ \sum_{k=1}^{K} (p_k(f) - p_k(o))^2 \right\}^{1/2} \tag{5.6}$$

where $p_k(f) - p_k(o)$ are the $k$th PC scores of the forecast and observation respectively.

The physical field, and thus its error, can be reconstructed using the PC scores of the forecast and observations and the EOFs. Denoting these with a hat:

$$\hat{f_n} - \hat{o_n} = \sum_{k=1}^{K} [p_k(f) - p_k(o)] \mathbf{e}_{n,k} = \sum_{k=1}^{K} p_k(\epsilon) \mathbf{e}_{n,k} \tag{5.7}$$

where $p_k(\epsilon)$ is the error in the PC scores and $\mathbf{e}_k$ is the $k$th EOF.

Substituting into the equation for RMSE,

$$RMSE = \left\{ \frac{1}{N} \sum_{n=1}^{N} \left[ \sum_{k=1}^{K} p_k(\epsilon) \mathbf{e}_{n,k} \right]^2 \right\}^{1/2} \tag{5.8}$$

Applying the square of a summation, this becomes:

$$RMSE = \left\{ \frac{1}{N} \sum_{n=1}^{N} \left[ \sum_{k=1}^{K} p_k(\epsilon)^2 \mathbf{e}_{n,k}^2 + 2 \sum_{k=1}^{K} \sum_{l=1}^{k-1} p_k(\epsilon) \mathbf{e}_{n,k} p_l(\epsilon) \mathbf{e}_{n,l} \right] \right\}^{1/2} \tag{5.9}$$

Then, due to the orthonormality of the EOFs, this simplifies to:

$$RMSE = \left\{ \frac{1}{N} \sum_{k=1}^{K} p_k(\epsilon)^2 \right\}^{1/2} \tag{5.10}$$

$$\Rightarrow RMSE = \frac{d}{N^{1/2}} \tag{5.11}$$

Hence, the constant of proportionality is the square-root of the dimensions $N$ of the original dataset.

# Chapter 6

# Conclusions

The downward propagation of stratospheric circulation anomalies into the troposphere has now long been recognised (Kodera et al., 1990; Baldwin and Dunkerton, 1999), although the direction of causality has been only confirmed for around a decade (e.g. Gerber et al., 2009; Hitchcock and Simpson, 2014). In that time, the understanding of stratosphere-troposphere coupling has grown, owing in part to both the natural increase in sample size with time and the development of 'high-top' models which can fully represent stratospheric variability (Charlton-Perez et al., 2013). Much attention is now being devoted to the role of stratospheric variability in Northern Hemisphere subseasonal prediction (Domeisen et al., 2020b,c) as part of the international S2S Prediction Project.

This thesis set out to further our understanding of subseasonal stratosphere-troposphere coupling and its applications within S2S forecasting. Both 'upward' (i.e., processes driving stratospheric variability and their model representation) and 'downward' (i.e., the influence of the stratosphere on the troposphere) aspects were considered. Although these have been analysed separately here, both aspects form part of a coupled two-way system.

The main outcomes of the thesis and how they answer the three key questions outlined in Chapter 1 are summarised in the following section. The implications and limitations of the findings are discussed alongside potential avenues of further work.

## 6.1 Summary and Discussion

### 6.1.1 What tropospheric processes influence stratospheric variability on subseasonal timescales?

The abrupt predictability onset of the February 2018 SSW (lead-times <12 days; Figure 2.1), which was followed by a high-impact coldwave in Europe, was investigated in Chapter 2. Previous and subsequent studies have shown that a more accurate model representation of the SSW (in operational forecasts and relaxation experiments) was associated with a better representation of the anomalous tropospheric weather patterns in the following weeks (Karpechko et al., 2018; Kautz et al., 2020; Knight et al., 2021), motivating a greater understanding of the mechanisms driving the SSW and limiting its predictability. S2S model forecasts were increasingly confident of a stronger-than-normal SPV during February 2018, before transitioning to predictions of a major SSW within a few days (see also Butler et al. (2020)). Such a rapid change-of-state can be hugely problematic for subseasonal prediction because of the different tropospheric weather conditions expected in opposing vortex states, particularly when there is little or no signal within prediction systems for the change to occur.

The results in Chapter 2 showed that the predictability onset of the 2018 SSW was linked to synoptic-scale anticyclonic wave breaking in the northeast Atlantic (Figure 2.6). This tropospheric precursor was diagnosed as the "Scandinavia-Greenland (S-G) dipole" through correlation analysis between the spread in S2S model forecast MSLP and SPV strength during the week before the SSW (Figure 2.3). The wave breaking was associated with driving enhanced vertically-propagating wave activity into the stratosphere and subsequent vortex weakening. It was also related to the downstream amplification of Ural blocking, which had been previously shown as a precursor to the 2018 event (Karpechko et al., 2018), and consistent with anomalous ridging in the Scandinavia-Urals region as an SSW precursor (e.g. Martius et al., 2009; Kolstad and Charlton-Perez, 2011; Peings, 2019).

Furthermore, the S-G dipole and the subsequent stratospheric evolution were not unique to 2018: a composite event-based analysis showed that previous similarly strong S-G dipoles were on average followed by a weakened SPV (Figure 2.11). Additionally, slightly more than one-third (35%) of major SSWs between 1979–2017 were preceded within 15 days by a similar evolution, emphasising both the role of North Atlantic anticyclonic wave breaking in SSW onset (albeit limited by the observational sample size) and the wide range of tropospheric precursors to SSWs. Although the S-G dipole was defined from the forecast evolution in 2018, its link with previous SSWs was used to support its role in driving the 2018 event. However, such diagnostics cannot confirm the ultimate driver of the 2018 SSW; to do so could involve localised tropospheric nudging or potential vorticity inversion experiments (e.g. O'Neill et al., 2017).

The abrupt vortex weakening in 2018 highlights the impact that synoptic-scale processes – like anticyclonic wave breaking – can have on a much larger spatio-temporal scale, and that errors in their prediction can rapidly grow upscale, manifesting the mixed-scale nature of subseasonal prediction. The ensemble correlation analysis technique used in Chapter 2 was able to capture the transient wave-breaking precursor, including the track of an extratropical cyclone. In contrast, the more commonly-used time-averaged analysis (such as in Karpechko et al. (2018)) is only able to capture the longer-lived quasi-stationary blocking anticyclone. The S-G dipole/wave breaking provides a mechanism for both the onset of blocking and the *sudden* changes to the stratospheric state, versus the slower evolution of a block. Similar case studies may illuminate further transient processes important for driving stratospheric variability. A recent example is the occurrence of an intense Siberian anticyclone and North Pacific bomb cyclones during the week before the January 2021 SSW, a configuration which – like the S-G dipole – constructively interferes with the stationary wave field and may thus serve as a transient source of enhanced vertically-propagating wave activity.

Further, Chapter 2 provides evidence that, whilst the stratospheric NAM may be on average

related to the cumulative eddy heat flux over the preceding weeks to months (Polvani and Waugh, 2004; Hinssen and Ambaum, 2010), transient events can have significant effects. In 2018, a transient event was necessary to initiate a stratospheric regime transition in what appears to have been a sensitive point between two 'attractors' (i.e., opposing vortex states). The response of the stratosphere to transient pulses of wave activity is significantly governed by the state of the stratosphere itself; favourable stratospheric states can arise through 'pre-conditioning' (Albers and Birner, 2014; Lawrence and Manney, 2020). Dependence on the stratospheric state was not explicitly explored in this thesis, but likely explains why not every S-G dipole event is quickly followed by a major SSW (as was the case in 2018). The rapid transition in 2018 may be evidence of sufficient "tuning" of the SPV (Albers and Birner, 2014) such that simply the additional vertically-propagating wave activity supplied by the S-G wave breaking event was sufficient to cause the observed explosive wavenumber-2 amplitude growth. In a similar sense, wave activity arising from S-G wave breaking may itself contribute to pre-conditioning the SPV to subsequent wave activity.

### 6.1.2 How well do current S2S forecast models represent these tropospheric processes, and is there evidence that model biases limit stratospheric skill?

The poor predictability of the 2018 event and known related model biases in blocking and wave breaking in northwest Europe (e.g. Ferranti et al., 2015; Quinting and Vitart, 2019) motivated quantifying model representation of the S-G dipole, explored in Chapter 3. This chapter defined the "Scandinavia-Greenland (S-G) pattern" as the second-leading mode of MSLP variability in the northeast Atlantic (Figure 3.1), a more generalised form of the S-G dipole from Chapter 2 (i.e., not dependent on the evolution in 2018 for its definition). The S-G pattern has some similarity to previously-defined patterns/regimes, most notably the Greenland–Scandinavia cluster of Cassou et al. (2004), but places greater emphasis on the Greenland centre-of-action (versus the more common emphasis on the Scandinavian ridge) and emphasises transience, rather than persistence (i.e., the pattern has a 4-day *e*-folding timescale). Lagged linear regression analysis (Figures 3.2 and 3.3) showed a

strong relationship with transient wavenumber-2 amplification, enhanced eddy heat flux into the stratosphere, and a significantly weakened SPV over subsequent weeks (also consistent with the evolution in 2018). This analysis moved beyond the composite-based analysis of Chapter 2, and confirmed the S-G pattern as a contributor to subseasonal SPV variability. The representation and predictability of the S-G pattern was then analysed in hindcasts from 10 S2S models.

These results showed significant biases in S-G pattern variability (Figure 3.6): considering the leading two EOFs, S2S models systematically favoured too much zonal (NAO-like) variability relative to ERA5 reanalysis and accordingly too little S-G variability. This bias was shown to amplify with lead-time and was particularly notable in weeks 3 and 4, again for all models. S-G pattern predictability (Figures 3.7 and 3.8) was also found to be lower than for the zonal pattern, particularly in weeks 3–4 for all models. Deterministic predictability (measured through correlation of the ensemble mean exceeding 0.6) ranged from 5–10 days, consistent with the timescale of deterministic SSW predictability (and indeed, predictions of the 2018 SSW) in the respective prediction systems (e.g. Domeisen et al., 2020b). Evidence of a (primarily horizontal) resolution dependence was found: the variability biases and poorest skill were largest in the three lowest-resolution models, although this is not itself conclusive of the role of resolution as many inter-model differences are present. However, the apparent role of resolution in modelling the S-G pattern agrees well with the effect of resolution on the representation of blocking and wave breaking in climate models (e.g. Schiemann et al., 2017). Good agreement was found between the biases in the EOFs and the biases in Rossby wave packet decay in northwest Europe reported in Quinting and Vitart (2019), supporting the physical interpretation of the EOFs.

Further, a decline in the relationship between the S-G pattern and a weakened SPV at subseasonal lead-times was found, owing to biases in the associated upward wave propagation/zonal-mean eddy heat flux (Figure 3.11). Consequently, even accurate representation and predictions of the tropospheric S-G pattern may be insufficient for

better subseasonal SPV predictions in the presence of the stratospheric biases. The physical reasons for the heat flux bias were not apparent; S2S data availability constraints precluded a detailed analysis, and it did not appear related to model resolution or mean-state biases in the stratospheric zonal winds. However, damped heat flux (suppressed vertical wave propagation) suggests a wave structure which is insufficiently westward-tilting with height, and thus the heat flux bias may be related to biases in the mean-state planetary wave structure with which the S-G pattern interacts.

Taken together, these results suggest that the poor predictions in 2018 were not a unique event but a manifestation of a problematic pattern within subseasonal forecast models. It was not, however, *directly* determined if poor S-G pattern predictions limited the predictability of SSWs other than that in 2018. To the best of the author's knowledge, this was the first study to specifically link biases in Rossby wave breaking and blocking in northwest Europe to potential subseasonal impacts through the SPV, suggesting the bias has longer-term impacts by propagating beyond its own timescale (i.e., in addition to any concurrent, direct effects on European weather predictability). Thus, improving the model representation of this pattern (which may be achieved by rectifying similar biases in blocking and wave breaking) and its influence on the SPV will likely increase subseasonal stratospheric, and resultant tropospheric, forecast skill. Interestingly, the three models with a horizontal resolution coarser than $1^{\circ}$ exhibited the largest biases in the S-G pattern; testing the representation of the pattern within the same model at different resolutions could confirm this apparent threshold.

### 6.1.3 How do circulation anomalies in the stratosphere influence subseasonal tropospheric weather patterns?

The influence of the stratosphere on the troposphere was analysed in Chapter 4, using the framework of weather regimes over North America. The approach was motivated by a previous study of the North Atlantic regimes (Charlton-Perez et al., 2018) and evidence of a link between SPV variability and subseasonal North American weather (Kodera et al.,

2016; Kretschmer et al., 2018a). Following a similar approach to earlier studies (Straus et al., 2007; Vigaud et al., 2018), four 500 hPa geopotential height regimes were defined using *k*-means clustering in principal component space: Arctic High, Arctic Low, Alaskan Ridge, and Pacific Trough (Figure 4.1). The domain of the clustering was centred on North America (rather than upstream, as in e.g. Fabiano et al. (2021)) to include both Pacific and Atlantic atmospheric variability, and to focus on the continent itself for the purpose of relating regimes to societal impacts.

The relationship with the stratosphere was quantified through probabilistic analysis of regime occurrence, persistence, and transition frequencies (Figure 4.3) stratified by the tercile strength of the lower-stratospheric polar vortex (where circulation anomalies are both most persistent and most likely to influence the troposphere). While only the near-instantaneous relationship between SPV strength and the regimes was defined, the long persistence of lower-stratospheric anomalies provides a natural extension to subseasonal scales. For three of the four regimes, significant differences in behaviour were found depending on the strength of the SPV: the Arctic High regime was found to be around seven times more likely when the lower-stratospheric vortex was weak versus strong, while the Arctic Low and Pacific Trough regimes were found to be up to twice as likely when the vortex was strong. The Alaskan Ridge regime, which did not show an apparent sensitivity to the stratospheric vortex strength, was found to be the regime most strongly associated with severe cold over North America (Figure 4.5). However, the Alaskan Ridge regime resembles the tropospheric response to stratospheric wave reflection, which has been linked to North American coldwaves (Kretschmer et al., 2018a; Matthias and Kretschmer, 2020) and may at least partly explain the absence of a signal in the zonal-mean zonal wind.

Although non-causal by design, these results suggested a significant downward impact of the SPV on North American weather. Therefore, accurate stratospheric predictions *may* improve subseasonal regime predictability (subsequently considered in Chapter 5). By

accounting for the likelihood of the mediating regime, weak SPV states are unlikely to be the primary subseasonal predictor for severe North American cold (unlike in Eurasia; Kolstad et al. (2010)). However, the regime-independent probability of cold weather over North America remains larger when the vortex is weak versus strong. A recent study by Huang et al. (2021) also found that the relationship between a weakened SPV and severe cold air outbreaks is stronger over Europe and East Asia compared with North America. The observed regime probabilities also provide a set of benchmark statistics against which the stratosphere-troposphere coupling in models can be tested. Further, while there are probabilistic changes, various different weather regimes are possible in both weak and strong vortex states. That is, the difference in the probability of an *individual* regime in *different* vortex states exceeds the difference in the probability between *different* regimes in the *same* vortex state, in agreement with the variety of tropospheric weather patterns following extreme stratospheric events (e.g. Karpechko et al., 2017).

The analysis of the four North Atlantic regimes by Charlton-Perez et al. (2018) reported two regimes that were more likely when the SPV was strong (Atlantic Ridge and NAO+), one that was more likely when the SPV was weak (NAO–), and one which was equiprobable (Scandinavian Blocking). This 2-1-0 pattern was also found for the North American regimes, despite only similarity between the NAO– and Arctic High regimes. It is not clear why the same probabilistic differences should also apply to the North American regimes, but it may relate to the behaviour of the underlying EOFs and the position of the regimes in EOF-space (see Chapter 5).

The observations-based regimes work of Chapter 4 was extended in Chapter 5 by seeking to understand how changes to the strength of the lower-stratospheric vortex influence the regimes. First, regression analysis was performed between the lower-stratospheric zonal-mean zonal winds and the leading three EOFs which define the regimes (Figure 5.3), providing a continuous framework rather than the binary regime classification. The regression coefficients were used to construct a 3D vector representing the linear response

in 3D-PC space to a lower-stratospheric zonal wind perturbation. A joint consideration of this vector with the location of the regime clusters in 3D-PC space yielded a theory (a statistical model) of which regime transitions are likely to occur due to changes in the strength of the lower-stratospheric vortex (Figures 5.4 and 5.5). These results were consistent with the observed differences in regime occurrence found in Chapter 4 (including the lack of a significant difference in Alaskan Ridge regime occurrence during different SPV states), providing a novel view of regime modulation by the stratosphere which incorporates the antecedent state of the troposphere.

The effect of dynamically imposing a stratospheric change was tested in a set of relaxation experiments using the ECMWF OpenIFS model, where the stratosphere was nudged to reanalysis. North American forecast accuracy (assessed by Z500 RMSE; Figure 5.8) was improved overall in weeks 3–4, confirming the potential benefits of a more accurate SPV forecast, but this improvement translated to only a modest improvement in the regime predictions. The lack of a substantial improvement to the regime forecasts suggests that a correct stratospheric forecast is insufficient to gain a correct North American regime forecast, although it does contribute to improving the forecast.

Selecting forecast start dates which possessed strong ensemble support for an incorrect regime regardless of the stratospheric error may have limited the potential improvement from dynamically correcting the stratosphere. Nonetheless, the average ensemble response in 3D-PC space (i.e., the vector between the control and nudged simulations; Figure 5.10) was consistent with the linear 3D regression vector framework, particularly for moderate-to-large ($\geq 3$ m s$^{-1}$) 100 hPa zonal wind perturbations (which may correspond to a change of vortex state).

Thus (in the absence of other changes), the relative impact of stratospheric changes on subseasonal North American regime predictions is potentially predictable through the effect of the linear 3D vector. The vector framework suggests that the regime state following

a stratospheric perturbation depends on the *prior* regime state; a paradigm in which the tropospheric response is determined by the troposphere (Gerber et al., 2009; Garfinkel et al., 2013; Maycock et al., 2020). The results presented in Chapter 5 may be particularly useful in interpreting scenarios where the SPV forecast shifts to a different extreme within a short time-frame, such as in the forecasts analysed in Chapter 2 before the 2018 SSW.

The conclusions are, however, limited by the sample size of the model runs, and would benefit from analysis of a larger set. The relatively low (60-level) vertical resolution employed in the experiments may have precluded a full tropospheric realisation of the corrected stratosphere. It is also important to note that, as the stratospheric state is influenced by the tropospheric state through vertically propagating waves, then the extent to which stratospheric nudging can be used to infer downward causality as an 'independent' source of skill is somewhat limited.

## 6.2 Future work

### 6.2.1 The Scandinavia-Greenland Pattern and tropospheric drivers of stratospheric variability

While previous cases of anticyclonic wave breaking/amplified S-G pattern were followed by major SSWs, the rapid vortex split within a few days of the wave break in 2018 was more unusual. Indeed, for the SPV to be so sensitive to such a single tropospheric event suggests both a particularly chaotic state and that other favourable factors were present (both in the troposphere and stratosphere). Rao et al. (2020) and Butler et al. (2020) show other large-scale patterns favourable for SSWs were present in the lead-up to the February 2018 SSW, including the highly-amplified MJO phase 6–7 during late January-early February. Multiple regression, or conditional composites (notwithstanding the sample size limitations in the historical record) may yield insight into when and why certain S-G events have an apparently much greater impact on the SPV. Although the wider tropospheric state may play a role, determining the stratospheric state which is best

conditioned for a large-amplitude S-G response is likely important.

The role of the MJO in the February 2018 SSW has been subsequently explored by Statnaia et al. (2020) and Knight et al. (2021), who related it to tropospheric Rossby wave trains which amplified planetary wavenumber-2 (associated with the vortex split). Statnaia et al. (2020) also supported the role of Atlantic anticyclonic wave breaking introduced in Chapter 2. Hence, the planetary wave field was likely sufficiently pre-conditioned such that effect of localised amplification via the S-G pattern was maximised (as the S-G pattern interacts constructively with wavenumber-2). It is possible that extratropical wave trains induced by the MJO or other intraseasonal tropical convection partly contributed to the wave breaking pattern in the North Atlantic. Such an evolution would be consistent with the lagged response of the Scandinavian Blocking regime following MJO phase 6 in Cassou (2008), though days with an amplified S-G pattern are not necessarily assigned to the Scandinavian Blocking regime.

A link to tropical convection could provide a clear route by which subseasonal S-G pattern predictability could be increased. Recently, Wang and Tan (2020) suggested a similar "Scandinavian Pattern" can be preceded (though not always) by convection in the North Atlantic by as much as 2 weeks, although these convection-preceded events were not found to precede significant SPV disruption. Otherwise, the S-G pattern may be ultimately dominated by synoptic variability which may fundamentally limit its predictability. However, in the case where specific timing of S-G wave breaking is not important for the stratospheric evolution (i.e., where the S-G pattern is simply a contributor to accumulated stratospheric heat flux), then lower intrinsic predictability may not be too consequential. In either case, a larger ensemble size which specifically targets spread in this region on subseasonal timescales (and addresses existing biases) could allow for sufficient probabilistic prediction of extreme scenarios with sufficient lead-time.

This thesis has focused on the positive S-G pattern as a precursor to SPV weakening.

The *negative* S-G pattern (characterised by an anomalous ridge over Greenland and trough over Scandinavia) has not been explicitly explored although many of the results, such as variability biases, are not sign-dependent. It was noted in the conclusions of Chapter 3 that the *negative* S-G pattern may be likely *following* SSWs, or alternatively, before strong vortex events through suppression of the stationary wave field. Its difference from the NAO may contribute to the spectrum of regime states in the region following SSWs (e.g. Beerli and Grams, 2019). Thus, while the analysis focused on the positive S-G pattern as a precursor to vortex weakening, the results may extend to the representation and predictability of the downward influence of the stratosphere in the northeast Atlantic. To that end, a subsequent study by Afargan-Gerstman et al. (2020) defined a Z500 anomaly dipole very similar to the negative S-G pattern (using a two-box method similar to that in Chapter 2). They linked its occurrence to marine cold air outbreaks, and found it was more likely in the four weeks following SSWs – confirming that investigating the downward impact of the stratosphere on the S-G pattern would be a useful avenue of future research with real-world impact.

### 6.2.2 North American weather regimes and the tropospheric response to the stratosphere

Although Chapter 5 suggested that even a perfect stratospheric forecast was insufficient to produce accurate subseasonal North American regime forecasts, this result may be partly an effect of the 'bust' cases being chosen independently of the stratospheric error. A much greater improvement may occur for cases where the dominant forecast error lies in the stratosphere. Hence, the sources of the particularly poor predictions in these cases must lie elsewhere, and are as yet undetermined. Other subseasonal tropospheric signals must be capable of destructively interfering with that from the stratosphere; an obvious candidate on subseasonal scales is the MJO, and so understanding the joint influence of the MJO and SPV is crucial to a more complete view of potential North American S2S regime predictability. A joint SPV-MJO approach has been used (Barnes et al., 2019; Green and Furtado, 2019) but not from the standpoint of North American regimes. Knight et al. (2021) used relaxation experiments to show a component of the negative NAO during the

2018 'Beast from the East' cold-wave arose from tropical forcing, while its amplification and persistence was only obtained with the SSW. This case exemplifies the interplay between subseasonal phenomena, which in 2018 seem to have constructively interfered with one another to produce an historic weather event.

Currently, there is no study which fully assesses whether better stratospheric forecasts are associated with better North American regime forecasts, or to what extent this may be captured by different S2S models (including whether S2S models correctly represent the observed regime occurrence statistics outlined in Chapter 4). The results of Chapter 5 provide a good foundation for this avenue of future work. One may readily pose the question: is the linear 3D vector ($\beta$) the same in the model as in observations? If not, it may provide an explanation for any biases that may exist in regime occurrence during different SPV states. Analysis in this vein may further contribute to identifying 'windows of opportunity' from the stratosphere. In addition, work to better understand to what extent stratospheric wave reflection events contribute to the occurrence of the Alaskan Ridge regime (building on Matthias and Kretschmer (2020)) is ongoing.

Having established the relationship between regimes and the SPV strength, it may be possible – given a suitably large ensemble – to 'pair' or cluster the regime forecasts with the SPV strength in each ensemble member. This approach may extract more physical meaning from ensemble spread or run-to-run changes in forecast uncertainty, yielding a conditional probability of the regime given the stratospheric state. As a thought experiment, consider a subseasonal ensemble forecast where 40% of ensemble members have a weak vortex and 60% have a strong vortex[1], and the ensemble regime attribution is:

- 40% Arctic High
- 30% Arctic Low
- 20% Pacific Trough

[1]This may seem like an extreme bifurcation, but consider forecasts before the 2018 SSW shown in Figure 2.2.

- 10% Alaskan Ridge

In isolation, this ensemble distribution suggests a strong signal for the Arctic High regime (twice the climatological likelihood of 20%). But, in this example, there is large uncertainty in the SPV strength, and so by pairing these with the stratosphere we may find (in a highly idealised state):

- 40% weak SPV of which 100% predict Arctic High.
- 60% strong SPV of which 50% Arctic Low, 33% Pacific Trough, 17% Alaskan Ridge.

Thus, the dominance of Arctic High in this example arises purely because *all* the weak SPV ensemble members have transitioned into that regime – while the strong SPV members, which dominate the ensemble, are split into *three different* regimes. The results presented in Chapters 4 and 5 support interpreting the ensemble spread as arising due to uncertainty in the SPV state because the regime likelihood is consistent with the vortex strength. Hence, by viewing the tropospheric forecast uncertainty through the lens of the stratosphere (motivated by the usually much larger predictability of the extratropical stratosphere), the favoured outcome may be seen as *not* Arctic High *unless* the vortex were to weaken (which could occur, for example, due to S-G wave breaking). Otherwise, considering the regime statistics presented in this thesis, the substantially increased probability of the Arctic High appears inconsistent with the stratospheric forecast favouring a stronger vortex. This apparent discord could be incorrectly interpreted as an effect of other tropospheric signals overwhelming that from the stratosphere. Uncertainty otherwise present in the troposphere then contributes to the spread among the remaining regimes with a predicted strong vortex.

In this way, the subseasonal tropospheric predictability arising from the stratosphere, which was not immediately evident, is exploited in a 'storyline'-like approach. As with accounting for the cumulative effects of large-scale dynamical uncertainty in quantifying regional climate change (Zappa and Shepherd, 2017; Shepherd et al., 2018), alternative, dynamically-consistent scenarios are presented, but not directly assigned probabilities. Some preliminary analysis suggests there are cases where such an approach is useful (e.g.

during the onset of the 2018 SSW and in some of the cases analysed in Chapter 5), which warrants further investigation. However, a significant challenge in taking this further is the small S2S hindcast ensemble sizes which generally preclude sub-setting, while the larger operational ensembles are limited by the lifespan of individual model versions. Larger S2S ensembles are therefore likely to be a rewarding use of computational resources, while a multi-model approach may also be a useful method of effectively increasing ensemble size.

Nevertheless, viewing the subseasonal evolution from the standpoint of dynamical processes opens the door to hierarchically clustering forecasts based first on known sources of predictability and then user-relevant outcomes. These can be then interpreted using known relationships. The "tubing" method (Atger, 1999), formerly used by ECMWF to classify deviations from the medium-range ensemble mean (or 'most likely' scenario) in a meteorologically meaningful sense, may provide a useful basis.

## 6.3 Concluding Remarks

This thesis demonstrated that polar vortex variability can be generated via Scandinavia-Greenland anticyclonic wave breaking and that this process was important in the onset of the February 2018 SSW, with predictability of the wave breaking limiting the predictability of the SSW. Moreover, current S2S forecast models struggle to represent and predict the Scandinavia-Greenland pattern and its influence on the SPV, suggesting that improvements in the modelling the pattern will improve stratospheric forecast skill.

Using the framework of weather regimes, it was shown that changes in the strength of the lower-stratospheric vortex impact the likelihood of different tropospheric weather regimes over North America. This set of statistics helps to quantify the stratospheric influence on tropospheric weather in a way useful for S2S prediction. A model for understanding how the stratosphere modulates North American regimes was developed and tested, providing an explanation for why some tropospheric regime states are more sensitive to a stratospheric perturbation than others. Nudging the stratosphere to observations

– mimicking an accurate stratospheric forecast – produced more accurate tropospheric forecasts over North America, but was insufficient to eliminate the erroneous regime forecasts.

There emerges a view that potential gains in subseasonal tropospheric predictability arising from the stratosphere may be severely limited by the troposphere itself – which may also affect vertically-propagating wave activity, yielding erroneous stratospheric predictions (with subsequent surface impacts, as was the case in 2018). This view was exemplified in the results of the stratospheric relaxation experiments: to an extent, these experiments resemble scenarios where the stratosphere is in a very predictable state (such as following major SSWs), yet in the cases studied here it did not translate to comparable tropospheric windows of opportunity. As a consequence, a question which follows is: does the stratosphere provide subseasonal windows of opportunity *on its own*? At least for North America, the results presented here suggest not. Thus, there should be a focus on either quantifying the remaining sources of uncertainty when the stratospheric uncertainty is low (e.g. low ensemble spread), or ascertaining how the stratosphere *can* produce a window of opportunity together with the rest of the Earth system.

To be able to confidently exploit the source of extended-range predictability present in the lower stratosphere therefore may now require giving greater attention to tropospheric processes, rather than the stratosphere itself (in the case of the now predominant high-top models). Process-based methods to unpack either conflicting or harmonious subseasonal signals, which can utilise the findings of this thesis, provide one route to extract potentially useful information otherwise swamped by inherent tropospheric chaos on subseasonal scales.

# Bibliography


Afargan-Gerstman, H., and D. I. Domeisen, 2020: Pacific modulation of the North Atlantic storm track response to sudden stratospheric warming events. *Geophysical Research Letters*, **47 (2)**, e2019GL085 007, doi:10.1029/2019GL085007.

Afargan-Gerstman, H., I. Polkova, L. Papritz, P. Ruggieri, M. P. King, P. J. Athanasiadis, J. Baehr, and D. I. Domeisen, 2020: Stratospheric influence on North Atlantic marine cold air outbreaks following sudden stratospheric warming events. *Weather and Climate Dynamics*, **1 (2)**, 541–553, doi:10.5194/wcd-1-541-2020.

Albers, J. R., and T. Birner, 2014: Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings. *Journal of the Atmospheric Sciences*, **71 (11)**, 4028–4054, doi:10.1175/JAS-D-14-0026.1.

Ambaum, M. H., and B. J. Hoskins, 2002: The NAO troposphere–stratosphere connection. *Journal of Climate*, **15 (14)**, 1969–1978, doi:10.1175/1520-0442(2002)015⟨1969:TNTSC⟩2.0.CO;2.

Ambaum, M. H., B. J. Hoskins, and D. B. Stephenson, 2001: Arctic oscillation or North Atlantic oscillation? *Journal of Climate*, **14 (16)**, 3495–3507, doi:10.1175/1520-0442(2001)014⟨3495:AOONAO⟩2.0.CO;2.

Amini, S., and D. M. Straus, 2019: Control of storminess over the Pacific and North America by circulation regimes. *Climate Dynamics*, **52 (7)**, 4749–4770, doi:10.1007/s00382-018-4409-7.

Andrews, D., J. Holton, and C. Leovy, 1987: *Middle Atmosphere Dynamics*. Academic Press, 489 pp.

Anstey, J. A., and Coauthors, 2013: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution. *Journal of Geophysical Research: Atmospheres*, **118 (10)**, 3956–3971, doi:10.1002/jgrd.50231.

Atger, F., 1999: Tubing: An alternative to clustering for the classification of ensemble forecasts. *Weather and Forecasting*, **14 (5)**, 741–757, doi:10.1175/1520-0434(1999)014⟨0741:TAATCF⟩2.0.CO;2.

Attard, H. E., and A. L. Lang, 2019: Tropospheric–stratospheric coupling following tropospheric blocking and extratropical cyclones. *Monthly Weather Review*, **147 (5)**, 1781–1804, doi:https://doi.org/10.1175/MWR-D-18-0335.1.

Baldwin, M. P., and T. J. Dunkerton, 1999: Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research: Atmospheres*, **104 (D24)**, 30 937–30 946, doi:10.1029/1999JD900445.

Baldwin, M. P., and T. J. Dunkerton, 2001: Stratospheric harbingers of anomalous weather regimes. *Science*, **294 (5542)**, 581–584, doi:10.1126/science.1063315.

Baldwin, M. P., D. B. Stephenson, D. W. Thompson, T. J. Dunkerton, A. J. Charlton, and A. O'Neill, 2003: Stratospheric memory and skill of extended-range weather forecasts. *Science*, **301 (5633)**, 636–640, doi:10.1126/science.1087143.

Baldwin, M. P., and D. W. Thompson, 2009: A critical comparison of stratosphere–troposphere coupling indices. *Quarterly Journal of the Royal Meteorological Society*, **135 (644)**, 1661–1672, doi:10.1002/qj.479.

Baldwin, M. P., and Coauthors, 2021: Sudden stratospheric warmings. *Reviews of Geophysics*, **59 (1)**, e2020RG000 708, doi:10.1029/2020RG000708.

Bao, M., X. Tan, D. L. Hartmann, and P. Ceppi, 2017: Classifying the tropospheric precursor patterns of sudden stratospheric warmings. *Geophysical Research Letters*, **44 (15)**, 8011–8016, doi:10.1002/2017GL074611.

Barnes, E. A., S. M. Samarasinghe, I. Ebert-Uphoff, and J. C. Furtado, 2019: Tropospheric and stratospheric causal pathways between the MJO and NAO. *Journal of Geophysical Research: Atmospheres*, **124 (16)**, 9356–9371, doi:10.1029/2019JD031024.

Barriopedro, D., and N. Calvo, 2014: On the relationship between ENSO, stratospheric sudden warmings, and blocking. *Journal of Climate*, **27 (12)**, 4704–4720, doi:10.1175/JCLI-D-13-00770.1.

Beerli, R., and C. M. Grams, 2019: Stratospheric modulation of the large-scale circulation in the Atlantic–European region and its implications for surface weather events. *Quarterly Journal of the Royal Meteorological Society*, **145 (725)**, 3732–3750, doi:10.1002/qj.3653.

Beerli, R., H. Wernli, and C. M. Grams, 2017: Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society*, **143 (709)**, 3025–3036, doi:10.1002/qj.3158.

Berckmans, J., T. Woollings, M.-E. Demory, P.-L. Vidale, and M. Roberts, 2013: Atmospheric blocking in a high resolution climate model: Influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, **14 (1)**, 34–40, doi:10.1002/asl2.412.

Birner, T., and J. R. Albers, 2017: Sudden stratospheric warmings and anomalous upward wave activity flux. *Sola*, **13 (Special Edition)**, 8–12, doi:10.2151/sola.13A-002.

Black, R. X., 2002: Stratospheric forcing of surface climate in the Arctic Oscillation. *Journal of Climate*, **15 (3)**, 268–277, doi:10.1175/1520-0442(2002)015⟨0268:SFOSCI⟩2.0.CO;2.

Bloomfield, H. C., D. J. Brayshaw, P. L. Gonzalez, and A. Charlton-Perez, 2021: Pattern-based conditioning enhances sub-seasonal prediction skill of European national energy variables. *Meteorological Applications*, **28 (4)**, e2018, doi:doi.org/10.1002/met.2018.

Butchart, N., 2014: The Brewer-Dobson circulation. *Reviews of Geophysics*, **52 (2)**, 157–184, doi:10.1002/2013RG000448.

Butler, A., and Coauthors, 2019a: Sub-seasonal predictability and the stratosphere. *Sub-seasonal to seasonal prediction*, Elsevier, 223–241, doi:10.1016/B978-0-12-811714-9.00011-5.

Butler, A. H., A. Charlton-Perez, D. I. Domeisen, I. R. Simpson, and J. Sjoberg, 2019b: Predictability of Northern Hemisphere final stratospheric warmings and their surface impacts. *Geophysical Research Letters*, **46 (17-18)**, 10 578–10 588, doi:10.1029/2019GL083346.

Butler, A. H., Z. D. Lawrence, S. H. Lee, S. P. Lillo, and C. S. Long, 2020: Differences between the 2018 and 2019 stratospheric polar vortex split events. *Quarterly Journal of the Royal Meteorological Society*, **146 (732)**, 3503–3521, doi:10.1002/qj.3858.

Butler, A. H., D. J. Seidel, S. C. Hardiman, N. Butchart, T. Birner, and A. Match, 2015: Defining sudden stratospheric warmings. *Bulletin of the American Meteorological Society*, **96 (11)**, 1913–1928, doi:10.1175/BAMS-D-13-00173.1.

Butler, A. H., J. P. Sjoberg, D. J. Seidel, and K. H. Rosenlof, 2017: A sudden stratospheric warming compendium. *Earth System Science Data*, **9 (1)**, 63–76, doi:10.5194/essd-9-63-2017.

Cassou, C., 2008: Intraseasonal interaction between the Madden–Julian oscillation and the North Atlantic Oscillation. *Nature*, **455 (7212)**, 523–527, doi:10.1038/nature07286.

Cassou, C., L. Terray, J. W. Hurrell, and C. Deser, 2004: North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing. *Journal of Climate*, **17 (5)**, 1055–1068, doi:10.1175/1520-0442(2004)017⟨1055:NAWCRS⟩2.0.CO;2.

Charlton, A. J., A. O'neill, P. Berrisford, and W. Lahoz, 2005: Can the dynamical impact of the stratosphere on the troposphere be described by large-scale adjustment to the stratospheric PV distribution? *Quarterly Journal of the Royal Meteorological Society*, **131 (606)**, 525–543, doi:10.1256/qj.03.222.

Charlton, A. J., and L. M. Polvani, 2007: A new look at stratospheric sudden warmings. Part I: Climatology and modeling benchmarks. *Journal of Climate*, **20 (3)**, 449–469, doi:10.1175/JCLI3996.1.

Charlton-Perez, A. J., R. W. Aldridge, C. M. Grams, and R. Lee, 2019: Winter pressures on the UK health system dominated by the Greenland Blocking weather regime. *Weather and Climate Extremes*, **25**, 100 218, doi:10.1016/j.wace.2019.100218.

Charlton-Perez, A. J., L. Ferranti, and R. W. Lee, 2018: The influence of the stratospheric state on North Atlantic weather regimes. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 1140–1151, doi:10.1002/qj.3280.

Charlton-Perez, A. J., and Coauthors, 2013: On the lack of stratospheric dynamical variability in low-top versions of the CMIP5 models. *Journal of Geophysical Research: Atmospheres*, **118 (6)**, 2494–2505, doi:10.1002/jgrd.50125.

Charney, J., and P. Drazin, 1961: Propagation of planetary-scale disturbances from the lower into the upper atmosphere. *Journal of Geophysical Research*, **66**, 83, doi:10.1029/JZ066i001p00083.

Chevallier, M., F. Massonnet, H. Goessling, V. Guémas, and T. Jung, 2019: The role of sea ice in sub-seasonal predictability. *Sub-seasonal to seasonal prediction*, Elsevier, 201–221, doi:10.1016/B978-0-12-811714-9.00010-3.

Choi, H., B.-M. Kim, and W. Choi, 2019: Type classification of sudden stratospheric warming based on pre-and postwarming periods. *Journal of Climate*, **32 (8)**, 2349–2367, doi:10.1175/JCLI-D-18-0223.1.

Cohen, J., L. Agel, M. Barlow, C. I. Garfinkel, and I. White, 2021: Linking Arctic variability and change with extreme winter weather in the United States. *Science*, **373 (6559)**, 1116–1121, doi:10.1126/science.abi9167.

Cohen, J., and J. Jones, 2011: Tropospheric precursors and stratospheric warmings. *Journal of Climate*, **24 (24)**, 6562–6572, doi:10.1175/2011JCLI4160.1.

Colucci, S. J., and M. E. Kelleher, 2015: Diagnostic comparison of tropospheric blocking events with and without sudden stratospheric warming. *Journal of the Atmospheric Sciences*, **72 (6)**, 2227–2240, doi:10.1175/JAS-D-14-0160.1.

Coy, L., S. Eckermann, and K. Hoppel, 2009: Planetary wave breaking and tropospheric forcing as seen in the stratospheric sudden warming of 2006. *Journal of the Atmospheric Sciences*, **66 (2)**, 495–507, doi:10.1175/2008JAS2784.1.

Coy, L., and S. Pawson, 2015: The major stratospheric sudden warming of January 2013: Analyses and forecasts in the GEOS-5 data assimilation system. *Monthly Weather Review*, **143 (2)**, 491–510, doi:10.1175/MWR-D-14-00023.1.

Davies, P. A., and Coauthors, 2021: The wet and stormy UK winter of 2019/2020. *Weather*, doi:10.1002/wea.3955.

Davini, P., S. Corti, F. D'Andrea, G. Rivière, and J. von Hardenberg, 2017: Improved winter European atmospheric blocking frequencies in high-resolution global climate simulations. *Journal of Advances in Modeling Earth Systems*, **9 (7)**, 2615–2634, doi:10.1002/2017MS001082.

Dawson, A., 2016: eofs: A library for EOF analysis of meteorological, oceanographic, and climate data. *Journal of Open Research Software*, **4 (1)**, doi:10.5334/jors.122.

de la Cámara, A., T. Birner, and J. R. Albers, 2019: Are sudden stratospheric warmings preceded by anomalous tropospheric wave activity? *Journal of Climate*, **32 (21)**, 7173–7189, doi:10.1175/JCLI-D-19-0269.1.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137 (656)**, 553–597, doi:10.1002/qj.828.

Dirmeyer, P. A., P. Gentine, M. B. Ek, and G. Balsamo, 2019: Land surface processes relevant to sub-seasonal to seasonal (S2S) prediction. *Sub-Seasonal to Seasonal Prediction*, Elsevier, 165–181, doi:10.1016/B978-0-12-811714-9.00008-5.

Domeisen, D. I., and A. H. Butler, 2020: Stratospheric drivers of extreme events at the Earth's surface. *Communications Earth & Environment*, **1 (1)**, 1–8, doi:10.1038/s43247-020-00060-z.

Domeisen, D. I., C. M. Grams, and L. Papritz, 2020a: The role of North Atlantic–European weather regimes in the surface impact of sudden stratospheric warming events. *Weather and Climate Dynamics*, **1 (2)**, 373–388, doi:10.5194/wcd-1-373-2020.

Domeisen, D. I., and Coauthors, 2020b: The role of the stratosphere in subseasonal to seasonal prediction: 1. Predictability of the stratosphere. *Journal of Geophysical Research: Atmospheres*, **125 (2)**, e2019JD030 920, doi:10.1029/2019JD030920.

Domeisen, D. I., and Coauthors, 2020c: The role of the stratosphere in subseasonal to seasonal prediction: 2. Predictability arising from stratosphere-troposphere coupling. *Journal of Geophysical Research: Atmospheres*, **125 (2)**, e2019JD030 923, doi:10.1029/2019JD030923.

Douville, H., 2009: Stratospheric polar vortex influence on Northern Hemisphere winter climate variability. *Geophysical Research Letters*, **36 (18)**, doi:10.1029/2009GL039334.

Edmon Jr, H., B. Hoskins, and M. McIntyre, 1980: Eliassen-Palm cross sections for the troposphere. *Journal of Atmospheric Sciences*, **37 (12)**, 2600–2616, doi:10.1175/1520-0469(1980)037⟨2600:EPCSFT⟩2.0.CO;2.

Fabiano, F., V. L. Meccia, P. Davini, P. Ghinassi, and S. Corti, 2021: A regime view of future atmospheric circulation changes in northern mid-latitudes. *Weather and Climate Dynamics*, **2 (1)**, 163–180, doi:10.5194/wcd-2-163-2021.

Falkena, S. K., J. de Wiljes, A. Weisheimer, and T. G. Shepherd, 2020: Revisiting the identification of wintertime atmospheric circulation regimes in the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, **146 (731)**, 2801–2814, doi:10.1002/qj.3818.

Feng, P.-N., H. Lin, J. Derome, and T. M. Merlis, 2021: Forecast Skill of the NAO in the Subseasonal-to-Seasonal Prediction Models. *Journal of Climate*, **34 (12)**, 4757–4769, doi:10.1175/JCLI-D-20-0430.1.

Ferranti, L., S. Corti, and M. Janousek, 2015: Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, **141 (688)**, 916–924, doi:10.1002/qj.2411.

Ferranti, L., L. Magnusson, F. Vitart, and D. Richardson, 2018a: A new product to flag

up the risk of cold spells in Europe weeks ahead. *ECMWF Newsletter*, **(158)**, 1–8, doi:10.21957/k2rlf88oe1.

Ferranti, L., L. Magnusson, F. Vitart, and D. S. Richardson, 2018b: How far in advance can we predict changes in large-scale flow leading to severe cold conditions over Europe? *Quarterly Journal of the Royal Meteorological Society*, **144 (715)**, 1788–1802, doi: 10.1002/qj.3341.

Frame, T. H., J. Methven, N. M. Roberts, and H. A. Titley, 2015: Predictability of frontal waves and cyclones. *Weather and Forecasting*, **30 (5)**, 1291–1302, doi: 10.1175/WAF-D-15-0039.1.

Furtado, J. C., J. Cohen, E. J. Becker, and D. C. Collins, 2021: Evaluating the relationship between sudden stratospheric warmings and tropospheric weather regimes in the NMME phase-2 models. *Climate Dynamics*, **56 (7)**, 2321–2338, doi:10.1007/s00382-020-05591-x.

Garfinkel, C. I., J. J. Benedict, and E. D. Maloney, 2014: Impact of the MJO on the boreal winter extratropical circulation. *Geophysical Research Letters*, **41 (16)**, 6055–6062, doi: 10.1002/2014GL061094.

Garfinkel, C. I., S. B. Feldstein, D. W. Waugh, C. Yoo, and S. Lee, 2012: Observed connection between stratospheric sudden warmings and the Madden-Julian Oscillation. *Geophysical Research Letters*, **39 (18)**, doi:10.1029/2012GL053144.

Garfinkel, C. I., D. L. Hartmann, and F. Sassi, 2010: Tropospheric precursors of anomalous northern hemisphere stratospheric polar vortices. *Journal of Climate*, **23 (12)**, 3282–3299, doi:10.1175/2010JCLI3010.1.

Garfinkel, C. I., D. W. Waugh, and E. P. Gerber, 2013: The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere. *Journal of Climate*, **26 (6)**, 2077–2095, doi:10.1175/JCLI-D-12-00301.1.

Garfinkel, C. I., I. White, E. P. Gerber, M. Jucker, and M. Erez, 2020: The building blocks

of Northern Hemisphere wintertime stationary waves. *Journal of Climate*, **33 (13)**, 5611–5633, doi:10.1175/JCLI-D-19-0181.1.

Garrido-Perez, J. M., C. Ordóñez, D. Barriopedro, R. García-Herrera, and D. Paredes, 2020: Impact of weather regimes on wind power variability in western Europe. *Appl. Energy*, **264**, 114 731, doi:10.1016/j.apenergy.2020.114731.

Gensini, V. A., D. Gold, J. T. Allen, and B. S. Barrett, 2019: Extended US tornado outbreak during late May 2019: A forecast of opportunity. *Geophysical Research Letters*, **46 (16)**, 10 150–10 158, doi:10.1029/2019GL084470.

Gerber, E., C. Orbe, and L. M. Polvani, 2009: Stratospheric influence on the tropospheric circulation revealed by idealized ensemble forecasts. *Geophysical Research Letters*, **36 (24)**, doi:10.1029/2009GL040913.

Goss, M., E. A. Lindgren, A. Sheshadri, and N. S. Diffenbaugh, 2021: The Atlantic jet response to stratospheric events: a regime perspective. *Journal of Geophysical Research: Atmospheres*, **126 (7)**, e2020JD033 358, doi:10.1029/2020JD033358.

Grams, C. M., R. Beerli, S. Pfenninger, I. Staffell, and H. Wernli, 2017: Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nat. Climate Change*, **7 (8)**, 557–562, doi:10.1038/nclimate3338.

Grams, C. M., L. Ferranti, and L. Magnusson, 2020: How to make use of weather regimes in extended range predictions for Europe. *ECMWF Newsletter*, **165**, 14–19.

Gray, S. L., C. Dunning, J. Methven, G. Masato, and J. M. Chagnon, 2014: Systematic model forecast error in Rossby wave structure. *Geophysical Research Letters*, **41 (8)**, 2979–2987, doi:10.1002/2014GL059282.

Green, M. R., and J. C. Furtado, 2019: Evaluating the joint influence of the Madden-Julian oscillation and the stratospheric polar vortex on weather patterns in the Northern hemisphere. *Journal of Geophysical Research: Atmospheres*, **124 (22)**, 11 693–11 709, doi:10.1029/2019JD030771.

Greening, K., and A. Hodgson, 2019: Atmospheric analysis of the cold late February and early March 2018 over the UK. *Weather*, **74 (3)**, 79–85, doi:10.1002/wea.3467.

Hall, R. J., D. M. Mitchell, W. J. Seviour, and C. J. Wright, 2020: Tracking the stratosphere-to-surface impact of Sudden Stratospheric Warmings. *Journal of Geophysical Research: Atmospheres*, e2020JD033881, doi:10.1029/2020JD033881.

Han, J., M. Kamber, J. Pei, and Coauthors, 2012: Getting to know your data. *Data Mining*, Elsevier Amsterdam, Netherlands, 39–82, doi:10.1016/B978-0-12-381479-1.00002-2.

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.*, **27 (9)**, 1119–1152, doi:10.1002/joc.1499.

Hardiman, S. C., N. J. Dunstone, A. A. Scaife, D. M. Smith, J. R. Knight, P. Davies, M. Claus, and R. J. Greatbatch, 2020: Predictability of European winter 2019/20: Indian Ocean dipole impacts on the NAO. *Atmospheric Science Letters*, **21 (12)**, e1005, doi:10.1002/asl.1005.

Hartley, D. E., J. T. Villarin, R. X. Black, and C. A. Davis, 1998: A new perspective on the dynamical link between the stratosphere and troposphere. *Nature*, **391 (6666)**, 471–474, doi:10.1038/35112.

Hartmann, D. L., 2015: Pacific sea surface temperature and the winter of 2014. *Geophysical Research Letters*, **42 (6)**, 1894–1902, doi:10.1002/2015GL063083.

Haynes, P., M. McIntyre, T. Shepherd, C. Marks, and K. P. Shine, 1991: On the "downward control" of extratropical diabatic circulations by eddy-induced mean zonal forces. *Journal of Atmospheric Sciences*, **48 (4)**, 651–678, doi:10.1175/1520-0469(1991)048⟨0651:OTCOED⟩2.0.CO;2.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi:10.1002/qj.3803.

Hinssen, Y. B., and M. H. Ambaum, 2010: Relation between the 100-hPa heat flux and stratospheric potential vorticity. *Journal of the atmospheric sciences*, **67 (12)**, 4017–4027, doi:10.1175/2010JAS3569.1.

Hitchcock, P., T. G. Shepherd, and G. L. Manney, 2013: Statistical characterization of Arctic polar-night jet oscillation events. *Journal of Climate*, **26 (6)**, 2096–2116, doi:10.1175/JCLI-D-12-00202.1.

Hitchcock, P., and I. R. Simpson, 2014: The downward influence of stratospheric sudden warmings. *Journal of the Atmospheric Sciences*, **71 (10)**, 3856–3876, doi:10.1175/JAS-D-14-0012.1.

Hitchcock, P., and I. R. Simpson, 2016: Quantifying eddy feedbacks and forcings in the tropospheric response to stratospheric sudden warmings. *Journal of the Atmospheric Sciences*, **73 (9)**, 3641–3657, doi:10.1175/JAS-D-16-0056.1.

Huang, J., P. Hitchcock, A. C. Maycock, C. M. McKenna, and W. Tian, 2021: Northern hemisphere cold air outbreaks are more likely to be severe during weak polar vortex conditions. *Communications Earth & Environment*, **2 (1)**, 1–11, doi:10.1038/s43247-021-00215-6.

Huang, W. T. K., A. Charlton-Perez, R. W. Lee, R. Neal, C. Sarran, and T. Sun, 2020: Weather regimes and patterns associated with temperature-related excess mortality in the UK: a pathway to sub-seasonal risk forecasting. *Environmental Research Letters*, **15**, doi:10.1088/1748-9326/abcbba.

Hudson, D., A. G. Marshall, Y. Yin, O. Alves, and H. H. Hendon, 2013: Improving intraseasonal prediction with a new ensemble generation strategy. *Monthly Weather Review*, **141 (12)**, 4429–4449, doi:10.1175/MWR-D-13-00059.1.

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269 (5224)**, 676–679, doi:10.1126/science.269.5224.676.

Jeuken, A., P. Siegmund, L. Heijboer, J. Feichter, and L. Bengtsson, 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *Journal of Geophysical Research: Atmospheres*, **101 (D12)**, 16 939–16 950, doi:10.1029/96JD01218.

Jia, L., and Coauthors, 2017: Seasonal prediction skill of northern extratropical surface temperature driven by the stratosphere. *Journal of Climate*, **30 (12)**, 4463–4475, doi:10.1175/JCLI-D-16-0475.1.

Julian, P. R., and K. B. Labitzke, 1965: A study of atmospheric energetics during the january–february 1963 stratospheric warming. *Journal of Atmospheric Sciences*, **22 (6)**, 597–610, doi:10.1175/1520-0469(1965)022⟨0597:ASOAED⟩2.0.CO;2.

Jung, T., M. Miller, and T. Palmer, 2010a: Diagnosing the origin of extended-range forecast errors. *Monthly Weather Review*, **138 (6)**, 2434–2446, doi:10.1175/2010MWR3255.1.

Jung, T., T. Palmer, M. Rodwell, and S. Serrar, 2010b: Understanding the anomalously cold European winter of 2005/06 using relaxation experiments. *Monthly Weather Review*, **138 (8)**, 3157–3174, doi:10.1175/2010MWR3258.1.

Karpechko, A. Y., 2018: Predictability of sudden stratospheric warmings in the ECMWF extended-range forecast system. *Monthly Weather Review*, **146 (4)**, 1063–1075, doi:10.1175/MWR-D-17-0317.1.

Karpechko, A. Y., A. Charlton-Perez, M. Balmaseda, N. Tyrrell, and F. Vitart, 2018: Predicting sudden stratospheric warming 2018 and its climate impacts with a multimodel ensemble. *Geophysical Research Letters*, **45 (24)**, 13–538, doi:10.1029/2018GL081091.

Karpechko, A. Y., P. Hitchcock, D. H. Peters, and A. Schneidereit, 2017: Predictability of downward propagation of major sudden stratospheric warmings. *Quarterly Journal of the Royal Meteorological Society*, **143 (704)**, 1459–1470, doi:10.1002/qj.3017.

Kautz, L.-A., I. Polichtchouk, T. Birner, H. Garny, and J. G. Pinto, 2020: Enhanced extended-range predictability of the 2018 late-winter Eurasian cold spell due to the strato-

sphere. *Quarterly Journal of the Royal Meteorological Society*, **146 (727)**, 1040–1055, doi:10.1002/qj.3724.

Kawatani, Y., K. Hamilton, L. J. Gray, S. M. Osprey, S. Watanabe, and Y. Yamashita, 2019: The effects of a well-resolved stratosphere on the simulated boreal winter circulation in a climate model. *Journal of the Atmospheric Sciences*, **76 (5)**, 1203–1226, doi:10.1175/JAS-D-18-0206.1.

Kidston, J., A. A. Scaife, S. C. Hardiman, D. M. Mitchell, N. Butchart, M. P. Baldwin, and L. J. Gray, 2015: Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience*, **8 (6)**, 433–440, doi:10.1038/ngeo2424.

Kim, H., J. H. Richter, and Z. Martin, 2019: Insignificant QBO-MJO prediction skill relationship in the SubX and S2S subseasonal reforecasts. *Journal of Geophysical Research: Atmospheres*, **124 (23)**, 12 655–12 666, doi:10.1029/2019JD031416.

Knight, J., and Coauthors, 2021: Predictability of European Winters 2017/2018 and 2018/2019: Contrasting influences from the Tropics and stratosphere. *Atmospheric Science Letters*, **22 (1)**, e1009, doi:10.1002/asl.1009.

Kodera, K., and M. Chiba, 1995: Tropospheric circulation changes associated with stratospheric sudden warmings: A case study. *Journal of Geophysical Research: Atmospheres*, **100 (D6)**, 11 055–11 068, doi:10.1029/95JD00771.

Kodera, K., H. Mukougawa, and A. Fujii, 2013: Influence of the vertical and zonal propagation of stratospheric planetary waves on tropospheric blockings. *Journal of Geophysical Research: Atmospheres*, **118 (15)**, 8333–8345, doi:10.1002/jgrd.50650.

Kodera, K., H. Mukougawa, and S. Itoh, 2008: Tropospheric impact of reflected planetary waves from the stratosphere. *Geophysical Research Letters*, **35 (16)**, doi:10.1029/2008GL034575.

Kodera, K., H. Mukougawa, P. Maury, M. Ueda, and C. Claud, 2016: Absorbing and reflecting sudden stratospheric warming events and their relationship with tropospheric

circulation. *Journal of Geophysical Research: Atmospheres*, **121 (1)**, 80–94, doi:10.1002/2015JD023359.

Kodera, K., K. Yamazaki, M. Chiba, and K. Shibata, 1990: Downward propagation of upper stratospheric mean zonal wind perturbation to the troposphere. *Geophysical Research Letters*, **17 (9)**, 1263–1266, doi:10.1029/GL017i009p01263.

Kolstad, E. W., T. Breiteig, and A. A. Scaife, 2010: The association between stratospheric weak polar vortex events and cold air outbreaks in the Northern Hemisphere. *Quarterly Journal of the Royal Meteorological Society*, **136 (649)**, 886–893, doi:10.1002/qj.620.

Kolstad, E. W., and A. J. Charlton-Perez, 2011: Observed and simulated precursors of stratospheric polar vortex anomalies in the Northern Hemisphere. *Climate Dynamics*, **37 (7)**, 1443–1456, doi:10.1007/s00382-010-0919-7.

Kolstad, E. W., C. O. Wulff, D. I. Domeisen, and T. Woollings, 2020: Tracing North Atlantic Oscillation forecast errors to stratospheric origins. *Journal of Climate*, **33 (21)**, 9145–9157, doi:10.1175/JCLI-D-20-0270.1.

Kretschmer, M., J. Cohen, V. Matthias, J. Runge, and D. Coumou, 2018a: The different stratospheric influence on cold-extremes in Eurasia and North America. *npj Climate and Atmospheric Science*, **1 (1)**, 1–10, doi:10.1038/s41612-018-0054-4.

Kretschmer, M., D. Coumou, L. Agel, M. Barlow, E. Tziperman, and J. Cohen, 2018b: More-persistent weak stratospheric polar vortex states linked to cold extremes. *Bulletin of the American Meteorological Society*, **99 (1)**, 49–60, doi:10.1175/BAMS-D-16-0259.1.

Krüger, K., B. Naujokat, and K. Labitzke, 2005: The unusual midwinter warming in the Southern Hemisphere stratosphere 2002: A comparison to Northern Hemisphere phenomena. *Journal of Atmospheric Sciences*, **62 (3)**, 603–613, doi:10.1175/JAS-3316.1.

Kushner, P. J., and L. M. Polvani, 2004: Stratosphere–troposphere coupling in a relatively simple AGCM: The role of eddies. *Journal of Climate*, **17 (3)**, 629–639, doi:10.1175/1520-0442(2004)017⟨0629:SCIARS⟩2.0.CO;2.

Lang, A. L., K. Pegion, and E. A. Barnes, 2020: Introduction to special collection: "Bridging weather and climate: Subseasonal-to-seasonal (S2S) prediction". *Journal of Geophysical Research: Atmospheres*, **125 (4)**, e2019JD031 833, doi:10.1029/2019JD031833.

Lawrence, Z. D., and G. L. Manney, 2020: Does the Arctic stratospheric polar vortex exhibit signs of preconditioning prior to sudden stratospheric warmings? *Journal of the Atmospheric Sciences*, **77 (2)**, 611–632, doi:10.1175/JAS-D-19-0168.1.

Lawrence, Z. D., J. Perlwitz, A. H. Butler, G. L. Manney, P. A. Newman, S. H. Lee, and E. R. Nash, 2020: The remarkably strong Arctic stratospheric polar vortex of winter 2020: Links to record-breaking Arctic oscillation and ozone loss. *Journal of Geophysical Research: Atmospheres*, **125 (22)**, e2020JD033 271, doi:10.1029/2020JD033271.

Lee, R. W., S. J. Woolnough, A. J. Charlton-Perez, and F. Vitart, 2019a: ENSO modulation of MJO teleconnections to the North Atlantic and Europe. *Geophysical Research Letters*, **46 (22)**, 13 535–13 545, doi:10.1029/2019GL084683.

Lee, S. H., A. Charlton-Perez, J. Furtado, and S. Woolnough, 2019b: Abrupt stratospheric vortex weakening associated with North Atlantic anticyclonic wave breaking. *Journal of Geophysical Research: Atmospheres*, **124 (15)**, 8563–8575, doi:10.1029/2019JD030940.

Lee, S. H., A. J. Charlton-Perez, J. C. Furtado, and S. J. Woolnough, 2020a: Representation of the Scandinavia–Greenland pattern and its relationship with the polar vortex in S2S forecast models. *Quarterly Journal of the Royal Meteorological Society*, **146 (733)**, 4083–4098, doi:10.1002/qj.3892.

Lee, S. H., J. C. Furtado, and A. J. Charlton-Perez, 2019c: Wintertime North American weather regimes and the Arctic stratospheric polar vortex. *Geophysical Research Letters*, **46 (24)**, 14 892–14 900, doi:10.1029/2019GL085592.

Lee, S. H., Z. D. Lawrence, A. H. Butler, and A. Y. Karpechko, 2020b: Seasonal forecasts of the exceptional Northern Hemisphere winter of 2020. *Geophysical Research Letters*, **47 (21)**, e2020GL090 328, doi:10.1029/2020GL090328.

Leutbecher, M., and Coauthors, 2017: Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society*, **143 (707)**, 2315–2339, doi:10.1002/qj.3094.

Liang, Y.-C., J.-Y. Yu, E. S. Saltzman, and F. Wang, 2017: Linking the tropical Northern Hemisphere pattern to the Pacific warm blob and Atlantic cold blob. *Journal of Climate*, **30 (22)**, 9041–9057, doi:10.1175/JCLI-D-17-0149.1.

Lim, Y., S.-W. Son, and D. Kim, 2018: MJO prediction skill of the subseasonal-to-seasonal prediction models. *Journal of Climate*, **31 (10)**, 4075–4094, doi:10.1175/JCLI-D-17-0545.1.

Limpasuvan, V., D. L. Hartmann, D. W. Thompson, K. Jeev, and Y. L. Yung, 2005: Stratosphere-troposphere evolution during polar vortex intensification. *Journal of Geophysical Research: Atmospheres*, **110 (D24)**, doi:10.1029/2005JD006302.

Lin, H., G. Brunet, and J. S. Fontecilla, 2010: Impact of the Madden-Julian Oscillation on the intraseasonal forecast skill of the North Atlantic Oscillation. *Geophysical Research Letters*, **37 (19)**, doi:10.1029/2010GL044315.

Linkin, M. E., and S. Nigam, 2008: The North Pacific Oscillation–west Pacific teleconnection pattern: Mature-phase structure and winter impacts. *Journal of Climate*, **21 (9)**, 1979–1997, doi:10.1175/2007JCLI2048.1.

Lorenz, D. J., and E. T. DeWeaver, 2007: Tropopause height and zonal wind response to global warming in the IPCC scenario integrations. *Journal of Geophysical Research: Atmospheres*, **112 (D10)**, doi:10.1029/2006JD008087.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21 (3)**, 289–307, doi:10.3402/tellusa.v21i3.10086.

Madden, R. A., and P. R. Julian, 1972: Description of global-scale circulation cells in the tropics with a 40–50 day period. *Journal of Atmospheric Sciences*, **29 (6)**, 1109–1123, doi:10.1175/1520-0469(1972)029⟨1109:DOGSCC⟩2.0.CO;2.

Maddison, J., S. Gray, O. Martínez-Alvarado, and K. Williams, 2019: Upstream cyclone influence on the predictability of block onsets over the Euro-Atlantic region. *Monthly Weather Review*, **147 (4)**, 1277–1296, doi:10.1175/MWR-D-18-0226.1.

Magnusson, L., 2017: Diagnostic methods for understanding the origin of forecast errors. *Quarterly Journal of the Royal Meteorological Society*, **143 (706)**, 2129–2142, doi:10.1002/qj.3072.

Mariotti, A., and Coauthors, 2020: Windows of opportunity for skillful forecasts subseasonal to seasonal and beyond. *Bulletin of the American Meteorological Society*, **101 (5)**, E608–E625, doi:10.1175/BAMS-D-18-0326.1.

Martius, O., L. Polvani, and H. Davies, 2009: Blocking precursors to stratospheric sudden warming events. *Geophysical Research Letters*, **36 (14)**, doi:10.1029/2009GL038776.

Masato, G., B. Hoskins, and T. J. Woollings, 2012: Wave-breaking characteristics of mid-latitude blocking. *Quarterly Journal of the Royal Meteorological Society*, **138 (666)**, 1285–1296, doi:10.1002/qj.990.

Matsueda, M., 2009: Blocking predictability in operational medium-range ensemble forecasts. *Scientific Online Letters on the Atmosphere,*, **5**, 113–116, doi:10.2151/sola.2009-029.

Matsueda, M., and T. Palmer, 2018: Estimates of flow-dependent predictability of wintertime Euro-Atlantic weather regimes in medium-range forecasts. *Quarterly Journal of the Royal Meteorological Society*, **144 (713)**, 1012–1027, doi:10.1002/qj.3265.

Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter Northern Hemisphere. *Journal of Atmospheric Sciences*, **27 (6)**, 871–883, doi:10.1175/1520-0469(1970)027⟨0871:VPOSPW⟩2.0.CO;2.

Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *Journal of Atmospheric Sciences*, **28 (8)**, 1479–1494, doi:10.1175/1520-0469(1971)028⟨1479:ADMOTS⟩2.0.CO;2.

Matthewman, N. J., J. G. Esler, A. J. Charlton-Perez, and L. M. Polvani, 2009: A new look at stratospheric sudden warmings. Part III: Polar vortex evolution and vertical structure. *Journal of Climate*, **22 (6)**, 1566–1585, doi:10.1175/2008JCLI2365.1.

Matthias, V., and M. Kretschmer, 2020: The influence of stratospheric wave reflection on North American cold spells. *Monthly Weather Review*, **148 (4)**, 1675–1690, doi:10.1175/MWR-D-19-0339.1.

Maycock, A. C., and P. Hitchcock, 2015: Do split and displacement sudden stratospheric warmings have different annular mode signatures? *Geophysical Research Letters*, **42 (24)**, 10–943, doi:10.1002/2015GL066754.

Maycock, A. C., G. I. Masukwedza, P. Hitchcock, and I. R. Simpson, 2020: A regime perspective on the North Atlantic eddy-driven jet response to sudden stratospheric warmings. *Journal of Climate*, **33 (9)**, 3901–3917, doi:10.1175/JCLI-D-19-0702.1.

McIntyre, M. E., and T. Palmer, 1983: Breaking planetary waves in the stratosphere. *Nature*, **305 (5935)**, 593–600, doi:10.1038/305593a0.

Michel, C., and G. Rivière, 2011: The link between Rossby wave breakings and weather regime transitions. *Journal of the Atmospheric Sciences*, **68 (8)**, 1730–1748, doi:10.1175/2011JAS3635.1.

Michelangeli, P.-A., R. Vautard, and B. Legras, 1995: Weather regimes: Recurrence and quasi stationarity. *Journal of the Atmospheric Sciences*, **52 (8)**, 1237–1256, doi:10.1175/1520-0469(1995)052⟨1237:WRRAQS⟩2.0.CO;2.

Mitchell, D. M., L. J. Gray, J. Anstey, M. P. Baldwin, and A. J. Charlton-Perez, 2013: The influence of stratospheric vortex displacements and splits on surface climate. *Journal of Climate*, **26 (8)**, 2668–2682, doi:10.1175/JCLI-D-12-00030.1.

Mo, K. C., and R. E. Livezey, 1986: Tropical-extratropical geopotential height teleconnections during the Northern Hemisphere winter. *Monthly Weather Review*, **114 (12)**, 2488–2515, doi:10.1175/1520-0493(1986)114⟨2488:TEGHTD⟩2.0.CO;2.

Nishii, K., H. Nakamura, and T. Miyasaka, 2009: Modulations in the planetary wave field induced by upward-propagating Rossby wave packets prior to stratospheric sudden warming events: A case-study. *Quarterly Journal of the Royal Meteorological Society*, **135 (638)**, 39–52, doi:10.1002/qj.359.

NOAA, 2019: NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters. URL https://www.ncdc.noaa.gov/billions/.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Monthly Weather Review*, **110 (7)**, 699–706, doi:10.1175/1520-0493(1982)110⟨0699:SEITEO⟩2.0.CO;2.

Oehrlein, J., G. Chiodo, and L. M. Polvani, 2020: The effect of interactive ozone chemistry on weak and strong stratospheric polar vortex events. *Atmos. Chem. Phys.*, **20 (17)**, 10 531–10 544, doi:10.5194/acp-20-10531-2020.

O'Neill, A., C. Oatley, A. J. Charlton-Perez, D. Mitchell, and T. Jung, 2017: Vortex splitting on a planetary scale in the stratosphere by cyclogenesis on a subplanetary scale in the troposphere. *Quarterly Journal of the Royal Meteorological Society*, **143 (703)**, 691–705, doi:10.1002/qj.2957.

Pedregosa, F., and Coauthors, 2011: Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.*, **12**, 2825–2830.

Peings, Y., 2019: Ural blocking as a driver of early-winter stratospheric warmings. *Geophysical Research Letters*, **46 (10)**, 5460–5468, doi:10.1029/2019GL082097.

Perlwitz, J., and N. Harnik, 2003: Observational evidence of a stratospheric influence on the troposphere by planetary wave reflection. *Journal of Climate*, **16 (18)**, 3011–3026, doi:10.1175/1520-0442(2003)016⟨3011:OEOASI⟩2.0.CO;2.

Perlwitz, J., and N. Harnik, 2004: Downward coupling between the stratosphere and troposphere: The relative roles of wave and zonal mean processes. *Journal of Climate*, **17 (24)**, 4902–4909, doi:10.1175/JCLI-3247.1.

Polvani, L. M., and D. W. Waugh, 2004: Upward wave activity flux as a precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *Journal of Climate*, **17 (18)**, 3548–3554, doi:10.1175/1520-0442(2004)017⟨3548:UWAFAA⟩2.0.CO;2.

Quinting, J., and F. Vitart, 2019: Representation of synoptic-scale Rossby wave packets and blocking in the S2S prediction project database. *Geophysical Research Letters*, **46 (2)**, 1070–1078, doi:10.1029/2018GL081381.

Quiroz, R. S., 1986: The association of stratospheric warmings with tropospheric blocking. *Journal of Geophysical Research: Atmospheres*, **91 (D4)**, 5277–5285, doi:10.1029/JD091iD04p05277.

Rao, J., and C. I. Garfinkel, 2021: The Strong Stratospheric Polar Vortex in March 2020 in Sub-Seasonal to Seasonal Models: Implications for Empirical Prediction of the Low Arctic Total Ozone Extreme. *Journal of Geophysical Research: Atmospheres*, **126 (9)**, e2020JD034 190, doi:10.1029/2020JD034190.

Rao, J., C. I. Garfinkel, and I. P. White, 2020: Predicting the downward and surface influence of the February 2018 and January 2019 sudden stratospheric warming events in subseasonal to seasonal (S2S) models. *Journal of Geophysical Research: Atmospheres*, **125 (2)**, doi:10.1029/2019JD031919.

Rex, D. F., 1951: The effect of Atlantic blocking action upon European climate. *Tellus*, **3 (2)**, 100–112, doi:10.3402/tellusa.v3i2.8617.

Riddle, E. E., M. B. Stoner, N. C. Johnson, M. L. L'Heureux, D. C. Collins, and S. B. Feldstein, 2013: The impact of the MJO on clusters of wintertime circulation anomalies over the North American region. *Climate Dynamics*, **40 (7-8)**, 1749–1766, doi:10.1007/s00382-012-1493-y.

Robertson, A. W., S. J. Camargo, A. Sobel, F. Vitart, and S. Wang, 2018: Summary of workshop on sub-seasonal to seasonal predictability of extreme weather and climate. *npj Climate and Atmospheric Science*, **1 (20178)**, doi:10.1038/s41612-017-0009-1.

Robertson, A. W., and M. Ghil, 1999: Large-scale weather regimes and local climate over the western United States. *Journal of Climate*, **12 (6)**, 1796–1813, doi:10.1175/1520-0442(1999)012⟨1796:LSWRAL⟩2.0.CO;2.

Robertson, A. W., N. Vigaud, J. Yuan, and M. K. Tippett, 2020: Toward Identifying Subseasonal Forecasts of Opportunity Using North American Weather Regimes. *Monthly Weather Review*, **148 (5)**, 1861–1875, doi:10.1175/MWR-D-19-0285.1.

Rogers, J. C., 1981: The North Pacific Oscillation. *Journal of Climatology*, **1 (1)**, 39–57, doi:10.1002/joc.3370010106.

Runde, T., M. Dameris, H. Garny, and D. Kinnison, 2016: Classification of stratospheric extreme events according to their downward propagation to the troposphere. *Geophysical Research Letters*, **43 (12)**, 6665–6672, doi:10.1002/2016GL069569.

Saffin, L., S. Gray, J. Methven, and K. Williams, 2017: Processes maintaining tropopause sharpness in numerical models. *Journal of Geophysical Research: Atmospheres*, **122 (18)**, 9611–9627, doi:10.1002/2017JD026879.

Saravanan, R., and P. Chang, 2019: Midlatitude mesoscale ocean-atmosphere interaction and its relevance to s2s prediction. *Sub-Seasonal to Seasonal Prediction*, Elsevier, 183–200, doi:10.1016/B978-0-12-811714-9.00009-7.

Scaife, A., and Coauthors, 2016: Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters*, **17 (1)**, 51–56, doi:10.1002/asl.598.

Schiemann, R., and Coauthors, 2017: The resolution sensitivity of Northern Hemisphere blocking in four 25-km atmospheric global circulation models. *Journal of Climate*, **30 (1)**, 337–358, doi:10.1175/JCLI-D-16-0100.1.

Scott, R., D. Dritschel, L. M. Polvani, and D. Waugh, 2004: Enhancement of Rossby wave breaking by steep potential vorticity gradients in the winter stratosphere. *Journal of Atmospheric Sciences*, **61 (8)**, 904–918, doi:10.1175/1520-0469(2004)061⟨0904:EORWBB⟩2.0.CO;2.

Scott, R., and L. M. Polvani, 2004: Stratospheric control of upward wave flux near the tropopause. *Geophysical Research Letters*, **31 (2)**, doi:10.1029/2003GL017965.

Shepherd, T. G., and Coauthors, 2018: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Climatic change*, **151 (3)**, 555–571, doi:10.1007/s10584-018-2317-9.

Sigmond, M., J. Scinocca, V. Kharin, and T. Shepherd, 2013: Enhanced seasonal forecast skill following stratospheric sudden warmings. *Nature Geoscience*, **6 (2)**, 98–102, doi:10.1038/ngeo1698.

Simpson, I. R., P. Hitchcock, T. G. Shepherd, and J. F. Scinocca, 2011: Stratospheric variability and tropospheric annular-mode timescales. *Geophysical Research Letters*, **38 (20)**, doi:10.1029/2011GL049304.

Smy, L., and R. Scott, 2009: The influence of stratospheric potential vorticity on baroclinic instability. *Quarterly Journal of the Royal Meteorological Society*, **135 (644)**, 1673–1683, doi:10.1002/qj.484.

Son, S.-W., H. Kim, K. Song, S.-W. Kim, P. Martineau, Y.-K. Hyun, and Y. Kim, 2020: Extratropical prediction skill of the Subseasonal-to-Seasonal (S2S) prediction models. *Journal of Geophysical Research: Atmospheres*, **125 (4)**, e2019JD031 273, doi:10.1029/2019JD031273.

Statnaia, I. A., A. Y. Karpechko, and H. J. Järvinen, 2020: Mechanisms and predictability of Sudden Stratospheric Warming in winter 2018. *Weather and Climate Dynamics*, **1 (2)**, 657–674, doi:10.5194/wcd-1-657-2020.

Straus, D. M., S. Corti, and F. Molteni, 2007: Circulation regimes: Chaotic variability versus SST-forced predictability. *Journal of Climate*, **20 (10)**, 2251–2272, doi:10.1175/JCLI4070.1.

Taguchi, M., 2014: Predictability of major stratospheric sudden warmings of the vortex split type: Case study of the 2002 southern event and the 2009 and 1989 northern events. *Journal of Atmospheric Sciences*, **71 (8)**, 2886–2904, doi:10.1175/JAS-D-13-078.1.

Taguchi, M., 2018: Comparison of subseasonal-to-seasonal model forecasts for major stratospheric sudden warmings. *Journal of Geophysical Research: Atmospheres*, **123 (18)**, 10 231–10 247, doi:10.1029/2018JD028755.

Tamarin, T., and Y. Kaspi, 2017: The poleward shift of storm tracks under global warming: A Lagrangian perspective. *Geophysical Research Letters*, **44 (20)**, 10–666, doi:10.1002/2017GL073633.

Thompson, D. W., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters*, **25 (9)**, 1297–1300, doi:10.1029/98GL00950.

Thompson, D. W., and J. M. Wallace, 2000: Annular modes in the extratropical circulation. Part I: Month-to-month variability. *Journal of Climate*, **13 (5)**, 1000–1016, doi:10.1175/1520-0442(2000)013⟨1000:AMITEC⟩2.0.CO;2.

Thompson, D. W., and J. M. Wallace, 2001: Regional climate impacts of the Northern Hemisphere annular mode. *Science*, **293 (5527)**, 85–89, doi:10.1126/science.1058958.

Tripathi, O. P., A. Charlton-Perez, M. Sigmond, and F. Vitart, 2015a: Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. *Environmental Research Letters*, **10 (10)**, 104 007, doi:10.1088/1748-9326/10/10/104007.

Tripathi, O. P., and Coauthors, 2015b: 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*, **141 (689)**, 987–1003, doi:10.1002/qj.2432.

Tripathi, O. P., and Coauthors, 2016: Examining the predictability of the stratospheric sudden warming of january 2013 using multiple nwp systems. *Monthly Weather Review*, **144 (5)**, 1935–1960, doi:10.1175/MWR-D-15-0010.1.

van der Wiel, K., H. C. Bloomfield, R. W. Lee, L. P. Stoop, R. Blackport, J. A. Screen, and F. M. Selten, 2019: The influence of weather regimes on European renewable energy production and demand. *Environmental Research Letters*, **14 (9)**, 094 010, doi:10.1088/1748-9326/ab38d3.

Vautard, R., 1990: Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. *Monthly Weather Review*, **118 (10)**, 2056–2081, doi:10.1175/1520-0493(1990)118⟨2056:MWROTN⟩2.0.CO;2.

Vigaud, N., A. W. Robertson, and M. K. Tippett, 2018: Predictability of recurrent weather regimes over North America during winter from submonthly reforecasts. *Monthly Weather Review*, **146 (8)**, 2559–2577, doi:10.1175/MWR-D-18-0058.1.

Vitart, F., 2017: Madden—Julian Oscillation prediction and teleconnections in the S2S database. *Quarterly Journal of the Royal Meteorological Society*, **143 (706)**, 2210–2220, doi:10.1002/qj.3079.

Vitart, F., and A. W. Robertson, 2018: The sub-seasonal to seasonal prediction project (s2s) and the prediction of extreme events. *npj Climate and Atmospheric Science*, **1 (1)**, 1–7, doi:10.1038/s41612-018-0013-0.

Vitart, F., A. W. Robertson, and D. L. Anderson, 2012: Subseasonal to seasonal prediction project: Bridging the gap between weather and climate. *Bulletin of the World Meteorological Organization*, **61 (2)**, 23.

Vitart, F., and Coauthors, 2017: The subseasonal to seasonal (S2S) prediction project database. *Bulletin of the American Meteorological Society*, **98 (1)**, 163–173, doi:10.1175/BAMS-D-16-0017.1.

Wallace, J. M., and D. S. Gutzler, 1981: Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly Weather Review*, **109 (4)**, 784–812, doi:10.1175/1520-0493(1981)109⟨0784:TITGHF⟩2.0.CO;2.

Wang, M., and B. Tan, 2020: Two types of the Scandinavian pattern: their formation mechanisms and climate impacts. *Journal of Climate*, **33 (7)**, 2645–2661, doi:10.1175/JCLI-D-19-0447.1.

Wang, S.-Y., L. Hipps, R. R. Gillies, and J.-H. Yoon, 2014: Probable causes of the abnormal ridge accompanying the 2013–2014 California drought: ENSO precursor and

anthropogenic warming footprint. *Geophysical Research Letters*, **41 (9)**, 3220–3226, doi:10.1002/2014GL059748.

Wang, S.-Y. S., W.-R. Huang, and J.-H. Yoon, 2015: The North American winter 'dipole' and extremes activity: A CMIP5 assessment. *Atmospheric Science Letters*, **16 (3)**, 338–345, doi:10.1002/asl2.565.

Waugh, D. W., A. H. Sobel, and L. M. Polvani, 2017: What is the polar vortex and how does it influence weather? *Bulletin of the American Meteorological Society*, **98 (1)**, 37–44, doi:10.1175/BAMS-D-15-00212.1.

White, C. J., and Coauthors, 2017: Potential applications of subseasonal-to-seasonal (S2S) predictions. *Meteorological Applications*, **24 (3)**, 315–325, doi:10.1002/met.1654.

White, I., C. I. Garfinkel, E. P. Gerber, M. Jucker, V. Aquila, and L. D. Oman, 2019: The downward influence of sudden stratospheric warmings: Association with tropospheric precursors. *Journal of Climate*, **32 (1)**, 85–108, doi:10.1175/JCLI-D-18-0053.1.

White, I. P., C. I. Garfinkel, E. P. Gerber, M. Jucker, P. Hitchcock, and J. Rao, 2020: The generic nature of the tropospheric response to sudden stratospheric warmings. *Journal of Climate*, **33 (13)**, 5589–5610.

Wilks, D., 2011: *Statistical Methods in the Atmospheric Sciences*. Elsevier, 818 pp.

Wilks, D., 2019: *Statistical Methods in the Atmospheric Sciences*. Elsevier, 818 pp.

Wittman, M. A., A. J. Charlton, and L. M. Polvani, 2007: The effect of lower stratospheric shear on baroclinic instability. *Journal of the atmospheric sciences*, **64 (2)**, 479–496, doi:10.1175/JAS3828.1.

WMO, 2018: WWRP/WCRP Sub-seasonal to Seasonal Prediction Project (S2S) Phase II Proposal (November 2018–December 2023). URL http://s2sprediction.net/file/documents_reports/P2_Pro_201811.pdf, 33 pp.

Woolnough, S. J., 2019: The Madden-Julian Oscillation. *Sub-Seasonal to Seasonal Prediction*, Elsevier, 93–117, doi:10.1016/B978-0-12-811714-9.00005-X.

Wulff, C. O., and D. I. Domeisen, 2019: Higher subseasonal predictability of extreme hot European summer temperatures as compared to average summers. *Geophysical Research Letters*, **46 (20)**, 11 520–11 529, doi:10.1029/2019GL084314.

Yu, B., and X. Zhang, 2015: A physical analysis of the severe 2013/2014 cold winter in North America. *Journal of Geophysical Research: Atmospheres*, **120 (19)**, 10–149, doi:10.1002/2015JD023116.

Zappa, G., and T. G. Shepherd, 2017: Storylines of atmospheric circulation change for European regional climate impact assessment. *Journal of Climate*, **30 (16)**, 6561–6577, doi:10.1175/JCLI-D-16-0807.1.