The influence of the stratospheric state on North Atlantic weather regimes

Article

Accepted Version

The influence of the Stratospheric state on North Atlantic Weather Regimes

Andrew J. Charlton-Perez,[a]*, Laura Ferranti[b], Robert W. Lee[a,c]

[a]*Dept. of Meteorology, Univ. of Reading, Reading, UK*
[b]*European Centre for Medium Range Weather Forecasts, Reading, UK*
[c]*NCAS-Climate, Univ. of Reading, Reading, UK*

*Correspondence to: Lyle Building, Univ. of Reading, Whiteknights, Reading, Berks, RG6 6BB, a.j.charlton-perez@reading.ac.uk

Stratosphere-troposphere coupling is often viewed from the perspective of the annular modes and their dynamics. Despite the obvious benefits of this approach, recent work has emphasised the greater tropospheric sensitivity to stratospheric variability in the Atlantic basin than in the Pacific basin. In this study, a new approach to understanding stratosphere-troposphere coupling is proposed, with a focus on the influence of the stratospheric state on North Atlantic weather regimes (during extended winter, November to March). The influence of the strength of the lower stratospheric vortex on four commonly used tropospheric weather regimes is quantified. The negative phase of the North Atlantic Oscillation is most sensitive to the stratospheric state, occurring on 33% of days following weak vortex conditions but on only 5% of days following strong vortex conditions. An opposite and slightly weaker sensitivity is found for the positive phase of the North Atlantic Oscillation and the Atlantic Ridge regime. For the North Atlantic Oscillation regimes, stratospheric conditions change both the probability of remaining in each regime and the probability of transitioning to that regime from others. A logistic regression model is developed to further quantify the sensitivity of tropospheric weather regimes to the lower stratospheric state. The logistic regression model predicts an increase of 40-60% in the probability of transition to the negative phase of the North Atlantic Oscillation for a one standard deviation reduction in the strength of the stratospheric vortex. Similarly it predicts a 10-30% increase in the probability of transition to the positive phase of the North Atlantic Oscillation for a one standard deviation increase in the strength of the stratospheric vortex. The stratosphere-troposphere coupling in the European Centre for Medium Range Weather Forecasts, Integrated Forecasting System model is found to be consistent with the re-analysis data by fitting the same logistic regression model.

Key Words: STRATOSPHERE-TROPOSPHERE COUPLING, WEATHER REGIMES

Received . . .

1. Introduction and Motivation

It is now widely accepted that variability in the winter stratosphere is coupled to low-frequency North Atlantic variability on medium-range, sub-seasonal and seasonal timescales during the winter season (e.g. Tripathi *et al.* 2015b; Sigmond *et al.* 2013; Scaife *et al.* 2016). Stratospheric variability, either Stratospheric Sudden Warming (SSW) events or periods in which the vortex is anomalously strong, typically result in anomalous conditions throughout the lower stratosphere on sub-seasonal timescales (between 30 and 60 days) (Limpasuvan *et al.* 2004, 2005). During this time period, the annular mode index in the lower stratosphere and troposphere is, on average, shifted toward negative values during SSW events and positive values during strong vortex events (Baldwin and Dunkerton 2001; Beerli *et al.* 2017). Negative values of the annular mode index are associated with a weakened polar night jet in the lower stratosphere and a southward shifted eddy-driven jet in the troposphere. Quantifying and understanding the strength and nature of stratosphere-troposphere coupling has important implications for sub-seasonal prediction for a range of different economic sectors, including renewable energy generation (Grams *et al.* 2017; Zubiate *et al.* 2017).

Although this general behaviour has been known for some time, understanding of the mechanisms by which this connection occurs and the extent to which these mechanisms are captured by numerical models remain elusive. A number of recent reviews (Kidston

et al. 2015; Tripathi *et al.* 2015a) discuss in detail the different mechanisms that have been proposed to explain this effect, with a strong emphasis on the influence of lower stratospheric conditions on upper tropospheric eddy momentum fluxes. Careful consideration of the momentum budget during SSW events (Hitchcock and Simpson 2016) suggests that it is the planetary scale tropospheric eddies which are directly influenced by the lower stratospheric state, with resulting changes to the synoptic-scale eddies simply a consequence of eddy-jet feedbacks in the troposphere,but there remains debate in the literature about the details of the mechanism.

Disentangling the mechanisms which underlie stratosphere-troposphere coupling is hampered by, on one-hand, the limited observational record of coupling events which is available and on the other by the challenge of designing experiments which can cleanly distinguish between different coupling mechanisms. One approach to dealing with the limited observational record is to examine stratosphere-troposphere coupling in numerical models. The approach of Hitchcock and Simpson (2014) which uses artificial stratospheric damping to simulate the same SSW event a large number of times demonstrates that the response to SSW events is strongly concentrated in the North Atlantic sector (see their Fig. 11) and the significant amount of variation in the strength of the tropospheric response that is dependent solely upon antecedent tropospheric conditions (see their Fig. 12). To date, the most comprehensive attempt to design experiments which distinguish between different proposed mechanisms for stratosphere-troposphere coupling are those described by Garfinkel *et al.* (2013). Garfinkel et al. also highlight the importance of tropospheric processes in determining stratosphere-troposphere coupling, focussing their attention on arguments associated with tropospheric jet persistence and its links to eddy-jet feedbacks. Garfinkel et al. propose a significant non-linearity in the sensitivity of different tropospheric states to stratospheric variability. In their idealised model, tropospheric jets centred at 40° latitude were more than three times more sensitive (in terms of the shift in the tropospheric jet position) to stratospheric perturbations than jets centred at 30° or 50° latitude.

In both of these cases, and in other studies including the original investigations of Baldwin and Dunkerton (1999), event-to-event variability in stratosphere-troposphere coupling is apparent. The key hypothesis of our study, linked to previous work, is that much of the event-to-event variability in stratosphere-troposphere coupling is the result of tropospheric low-frequency variability rather than simply a function of the strength and longevity of the SSW or strong vortex event. To investigate this hypothesis an alternative and complimentary approach to understanding stratosphere-troposphere coupling is adopted, through the use of North-Atlantic weather regimes.

The use of weather regimes to describe low-frequency tropospheric variability is reviewed extensively in Hannachi *et al.* (2017). While there is some debate about the uniqueness and relevance of regime analysis in the North Atlantic, a large number of authors have found this approach to be a useful way to describe low frequency variability and its connection to external drivers. In this context, external refers to other parts of the climate system which vary independently of the extra-tropical North Atlantic atmosphere. Most relevant to this work is the approach taken by Cassou *et al.* (2004) and Cassou (2008) whose studies establish how North Atlantic Sea-Surface Temperatures and the Madden-Julian Oscillation influence the frequency of occurrence of different weather regimes in the North Atlantic and the likelihood of transitions between these regimes. Our study adopts a similar approach, but in this case applied to the influence of the strength of the polar vortex in the lower stratosphere as a proxy for stratospheric variability. A number of studies, most recently Karpechko (2015), have shown that the strength of the polar vortex close to the tropopause has the strongest connection to tropospheric variability (perhaps as this is also the region where stratospheric anomalies are longest lived).

A recent study by Beerli *et al.* (2017) began to examine the response of Atlantic weather regimes to changes in the strength of the stratospheric vortex. In this study, we seek to expand on this approach using a simple model of regime behaviour in the troposphere in order to answer the following two questions:

- Which North Atlantic weather regimes are influenced by the stratospheric polar vortex?
- By how much does the probability of occurrence of North Atlantic weather regimes change in response to typical variations of the stratospheric polar vortex?

In answering these questions, it is necessary to first introduce a simple model of North Atlantic weather regimes (henceforth weather regimes) in order to understand how the stratospheric state modifies their behaviour. Once the simple model is established, a logistic regression model is used to quantify the size of the influence of the stratospheric state on the frequency of occurrence of weather regimes. In the first part of the paper, the model of stratospheric influence on weather regimes is developed using ERA-Interim re-analysis data. As in previous studies, the limited observational record makes quantifying some parts of the stratospheric influence on weather regimes difficult and so in the second part of the paper we also make use of a large set of simulations from the ECMWF Integrated Forecasting System model to be able to further quantify this influence. We first show that the model faithfully reproduces the stratospheric influence on weather regimes determined from the re-analysis data. As shown by Dawson *et al.* (2012) it is important to use a relatively high resolution for this purpose, since climate resolution simulations are unable to faithfully reproduce the same weather regimes observed in re-analysis data. We use a model configuration (Ferranti *et al.* 2015) which has been previously shown to faithfully reproduce weather regime behaviour in the North Atlantic for this purpose.

2. Methodology

In all analysis in this study, data from the extended winter season (NDJFM) is considered.

2.1. Regime identification

The methodology used to derive weather regimes in both the ERA-Interim re-analysis data (Dee *et al.* 2011) and the IFS re-forecasts is similar to that used by Ferranti *et al.* (2015) and previously by Michelangeli *et al.* (1995) and Cassou (2008). Briefly, regimes are calculated as 'k-means' clusters of the phase space spanned by the first 14 leading Empirical Orthogonal Functions (EOFs) of the 500hPa geopotential height field over the North Atlantic domain (90°W-30°E; 20-80°N) for the cold season (November to March). For the re-analysis data, the winter seasons from 1979-1980 to 2016-2017 are used, giving a total of 38 winter seasons. The choice of 14 EOFs and 4 clusters is standard in the literature, but the details of the clustering are insensitive to the particular choices of these parameters (not shown).

This analysis produces the four familiar cluster patterns shown in Fig. 1 which are henceforth labelled as the Positive NAO (NAO+), Negative NAO (NAO-), Scandinavian Blocking (BL) and Atlantic Ridge (AR) regimes. Three of the weather regimes describe the three preferred North Atlantic jet stream locations (Woollings *et al.* 2010b), namely, NAO-, NAO+ and AR broadly correspond to southern, central and northern jet-states respectively. Recent work also links the BL regime to a fourth mixed jet state and highlights the complexity of comparing regimes from this kind of clustering approach and measures of the jet latitude (Madonna *et al.* 2017). Other studies have emphasised the correlation between jet shifting and strengthening which can be observed when analysing the first and second EOFs of the Atlantic zonal wind (Sparrow *et al.* 2009; Sheshadri and Plumb 2017).

The four regimes are used in the ECMWF medium-range clustering products (Ferranti and Corti 2011) to provide additional information about the ensemble in terms of large-scale circulation and to allow an objective verification of the regime transitions.

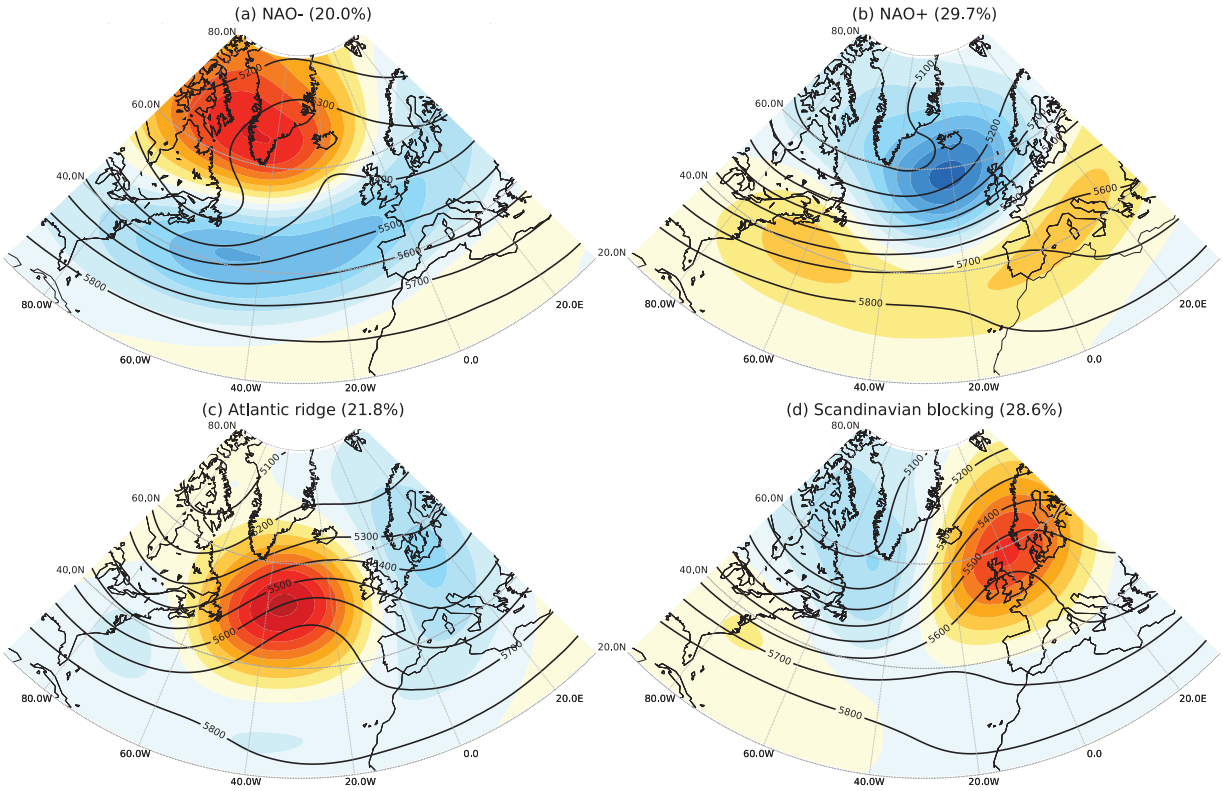

Figure 1. Cluster centroids for the four regimes used in the study. Colours show 500hPa geopotential height anomalies for the centroid for the (a) NAO- (b) NAO+ (c) Atlantic Ridge and (d) Scandinavian Blocking regimes. For reference, the average geopotential height field is plotted for each regime in black contours. The percentage of days in which each regime occurs is shown at the top of each panel, this may not add up to 100% due to rounding

In both the re-analysis and model datasets each day is assigned to one of these four regimes based on the minimum Euclidian distance between the regime centroid and the daily 500hPa geopotential height anomaly as described in Ferranti *et al.* (2015). In some cases, the projection of an individual daily pattern onto all four of the centroids can be relatively weak and some authors have chosen to include a seperate 'no regime' classification in order to remove these cases from subsequent analysis (e.g. Grams *et al.* 2017). In our analysis we choose not to take this approach, but do comment on the strength of the projection onto each cluster centroid for particular cases throughout the manuscript.

The model dataset used is a set of ensemble reforecasts from the ECMWF Integrated Forecast System (IFS, cycle 40R1, operational in 2014 with an enhanced ensemble size of 11 ensemble members). The data used covers 20 cold seasons from November 1994 to February 2014. All reforecasts are made with the same version of the model, thereby avoiding any discontinuities related to changes in model configuration. The first 14 days of each hindcast is discarded, so that the tropospheric evolution in each hindcast can be thought of as an independent sample of the influence of the stratospheric state on weather regime frequency.

2.2. Stratospheric polar vortex strength

Throughout the manuscript we characterise the strength of the stratospheric polar vortex by calculating the daily-mean, zonal-mean, zonal wind anomaly at 100hPa and 60° N. While other, more complex, measures might also be used to calculate the strength of the polar vortex in the lower stratosphere, for example the annular mode index, the correlation between these measures and a simple diagnostic of the zonal mean zonal wind is very high. In the first-order Markov models which are developed, the daily state of the polar vortex is assigned to one of three tercile categories, where the thresholds for these categories are calculated from the full set of daily zonal mean zonal wind anomaly in either the re-analysis data or model. For the ERA-Interim data, the threshold for the weak vortex state is -2.30 ms^{-1} and for the strong vortex state is 2.19 ms^{-1}.

3. Results

3.1. Regime frequency changes in response to the stratosphere

As an initial exploration of the influence of stratospheric vortex strength on weather regimes in the North Atlantic, changes to the frequency of occurrence, persistence and transition between regimes are first estimated. The frequency of weather regime occurrence

($\hat{p}$) is a widely used metric (e.g. Cassou 2008) and is estimated here simply by dividing the number of times a given regime occurs (a) by the total number of days sampled in each subset (n).

$$\hat{p} = \frac{a}{n} \tag{1}$$

A standard 95 % Wald confidence interval for the frequency is calculated since the sample size is large:

$$\hat{p} \pm 1.96\sqrt{\frac{1}{n}\hat{p}(1-\hat{p})} \tag{2}$$

To estimate the influence of the stratospheric vortex strength on the frequency of occurrence of weather regimes, this metric is calculated for three sub-sets of the data based on the state of the polar vortex (weak, neutral and strong) on the previous day. The frequency of occurrence estimates are shown in Fig. 2(a). This frequency is written as (where R_i represents one of the four weather regimes):

$$\hat{p}(R_i(t+1)|U_{100}(t)). \tag{3}$$

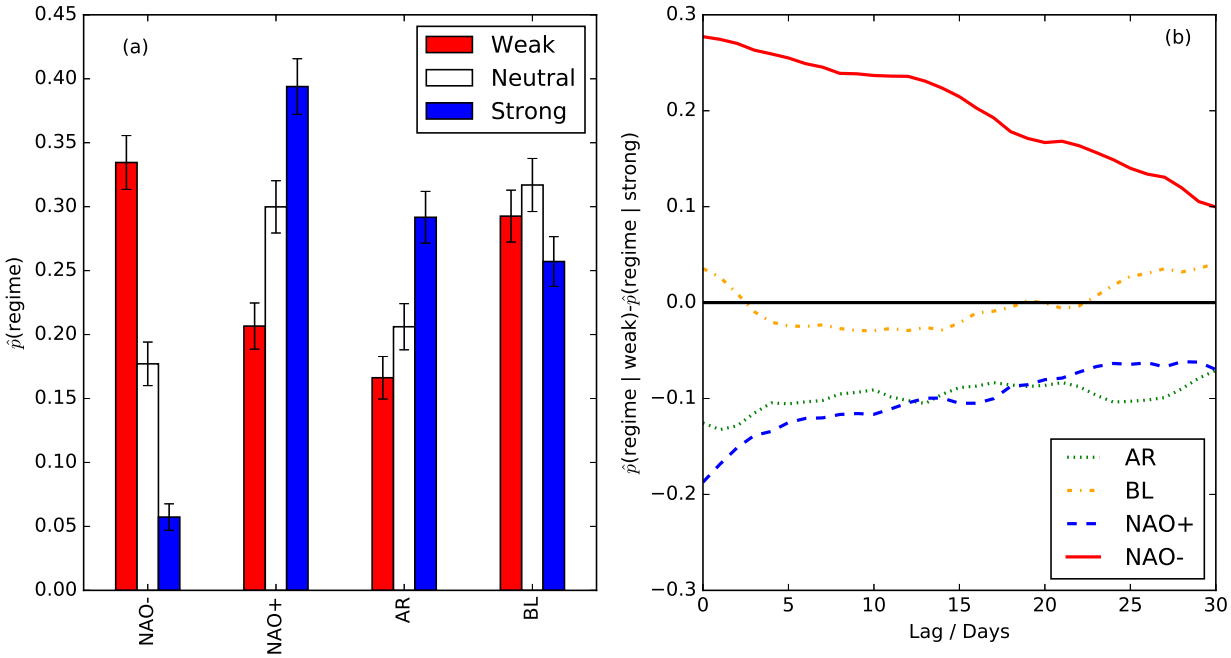

Figure 2. (a) Frequency of occurrence of North Atlantic weather regimes immediately following weak stratospheric vortex conditions (left, red bar), neutral stratospheric vortex conditions (central, white bar) and strong stratospheric vortex conditions (blue, right bar). The 95% confidence interval (see text for details) is shown at the end of each bar. (b) Difference between frequency of occurrence of each regime in weak and strong stratospheric vortex conditions as a function of the lag between the stratospheric vortex conditions and the tropospheric regime. For example, a lag of 20 indicates that the frequency of occurrence of each tropospheric regime is conditional upon the stratospheric state twenty days before. Each line shows a different tropospheric regime, NAO- (red, solid), NAO+(blue, dashed), AR (green, dotted) and BL (orange, dot-dash).

For neutral conditions, BL is the most frequently observed regime, followed by NAO+. Stratospheric vortex conditions significantly shift weather regime frequency as expected. Following weak stratospheric vortex conditions, NAO- occurs almost twice as frequently as during neutral stratospheric vortex conditions with consequent reductions in the occurrence frequency of both AR and NAO+. Following strong stratospheric vortex conditions, the opposite sensitivity is found, large reductions in NAO- frequency and increases in both AR and NAO+. Changes to BL frequency are smaller and not significant. Previous studies have highlighted the links between Greenland Blocking and stratospheric vortex conditions (Woollings *et al.* 2010a; Davini *et al.* 2014), reflected here in the sensitivity of the NAO- regime.

Since the lower stratosphere has a long memory during winter, the qualitative sensitivity of tropospheric weather regimes to antecedent stratospheric conditions can be reproduced for a large range of lags between the variable chosen to represent the stratosphere and the tropospheric weather regime. Fig. 2(b) shows the difference in occurrence frequency between weak and strong stratospheric vortex states as a function of lag between the stratosphere and troposphere. A lag of twenty days in this figure indicates the estimated frequency of occurrence of each tropospheric regime is conditional upon the stratospheric vortex state twenty days previously. As expected, the difference between occurrence frequency of the NAO-, NAO+ and AR regimes declines as the lag between the stratospheric vortex state and tropospheric weather regime increases. However, a large difference remains for the NAO-, NAO+ and AR remains even at 30 days lag, consistent with the sensitivity to the stratospheric state shown in Fig. 2(a). For the remainder of the manuscript we therefore continue to consider weather regime sensitivity to the stratospheric vortex state on the previous day.

The regime sensitivity to the antecedent stratospheric vortex state could be related either to changes to the probability of transition to the AR, NAO+ and NAO- regimes or changes to probability of persistence of the same regimes. These quantities are estimated by further sub-setting the observed regime data used in Fig. 2 adding a condition related to the weather regime on the previous day in addition to the stratospheric vortex state. For example, the NAO- regime days in the weak vortex subset are further subset into two sets

in which the weather regime on the previous day was NAO- and cases where the weather regime on the previous day was any other regime. The number of days in which the previous regime was NAO- is used to estimate the persistence probability for the NAO-/weak vortex case.

$$\hat{p}(R_i(t+1)|R_i(t)\ \&\ U_{100}(t)) \tag{4}$$

The number of days in which the previous regime was not NAO- is used to estimate the transition probability for the NAO-/weak vortex case.

$$\hat{p}(R_i(t+1)|R_{j\neq i}(t)\ \&\ U_{100}(t)) \tag{5}$$

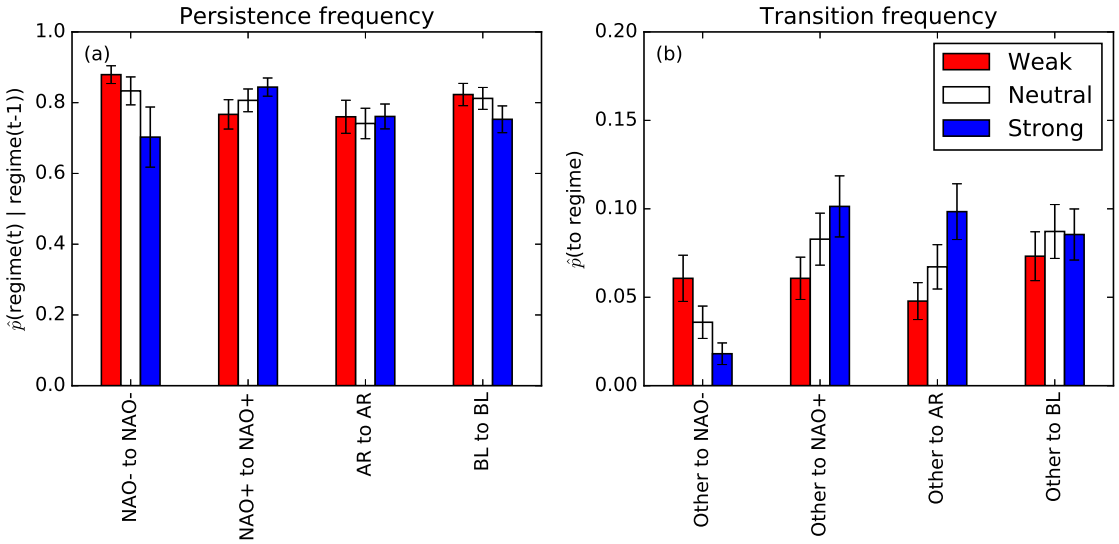

Figure 3. (a) Persistence and (b) transition frequency for weather regimes immediately following weak stratospheric vortex conditions (left, red bars), neutral stratospheric vortex conditions (central, white bars) and strong stratospheric vortex conditions (right, blue bars). The 95% confidence interval (see text for details) is shown at the end of each bar.

Figure 3 shows the persistence and transition probabilities for each of the four weather regimes following, weak, neutral and strong stratospheric vortex conditions. The changes to persistence and transition probabilities are consistent with what might be expected given the changes in regime frequency shown in Fig. 2, at least for the two NAO regimes. For antecedent weak vortex conditions, the NAO- regime is significantly more persistent and a transition to this regime is more likely when compared to antecedent strong vortex conditions. The NAO+ regime has an opposite sensitivity, with significantly enhanced persistence and transition frequency for antecedent strong vortex conditions compared to weak vortex conditions. The AR regime has the same sensitivity as the NAO+ regime in it's transition frequency but no significant sensitivity in it's persistence.

3.2. A simple Markov model

It is helpful to the subsequent discussion to cast the weather regime coupling to the stratosphere revealed in the previous section in terms of a simple Markov model. An example for the NAO- regime is shown in Fig. 4. Fig. 4 shows the North Atlantic state as being either NAO- or any one of the other regimes (other state in the figure). The arrows indicate persistence probability for NAO- (1-α) and the transition probabilities between the NAO- state and other states (α&β), which can be estimated from the analysis in the previous section.

Describing North Atlantic regimes in this way assumes that state tomorrow depends only upon the state today. The model is shown on the right hand side of Fig. 5 and is defined by the transition matrix, P, which in this case is:

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix} = \begin{bmatrix} 0.83 & 0.17 \\ 0.04 & 0.96 \end{bmatrix} \tag{6}$$

After N-steps of the model, the probability of being in either the NAO- or other state can be calculated from the N-th power of P. In the limit of large n, the computed probability represents the steady state probability of the two states. For the case with a neutral stratospheric vortex above:

$$\lim_{N\to\infty} P^N = \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix} = \begin{bmatrix} 0.19 & 0.81 \\ 0.19 & 0.81 \end{bmatrix} \tag{7}$$

The model described by P represents the observed frequency of occurrence of the NAO- state when the stratospheric vortex is neutral very well ($\hat{p} = 0.19$). To describe the impact of changes to the stratospheric state, an alternate Markov model that reflects the increased persistence and likelihood of transition into the NAO- regime, is shown on the left hand side of Fig. 5.

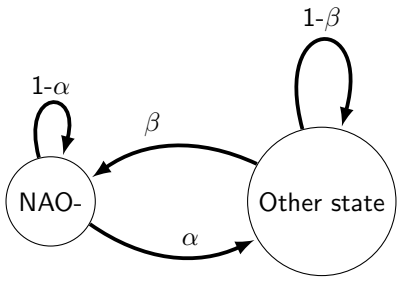

Figure 4. Model of transitions between the NAO- and other states for the Neutral case. Circles indicate the weather regime. Arrows indicate transitions between states with the labels indicating the probability of following each path.

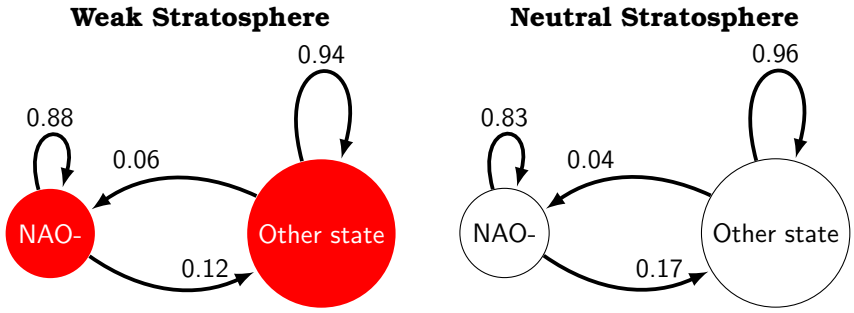

Figure 5. Model of transitions between the NAO- and other states for the weak and neutral stratospheric regimes. Circles indicate the weather regime. Arrows indicate transitions between states with the labels indicating the probability of following each path.

In this case, the steady state probability (for the weak vortex state) is:

$$\lim_{N\to\infty} P^N = \begin{bmatrix} 0.33 & 0.67 \\ 0.33 & 0.67 \end{bmatrix} \tag{8}$$

For a positive recurrent Markov chain like the one described above the mean recurrence time (the expected time taken to return to a given state) is given by the reciprocal of the steady state probability. For the two models discussed above, the recurrence time for the NAO- state is approximately five days when the stratospheric vortex is in a neutral state ($\hat{p}$=0.19, 1/$\hat{p}$=5.3), three days when the stratospheric vortex is in a weak state ($\hat{p}$=0.33, 1/$\hat{p}$=3.0) and sixteen days when the stratospheric vortex is in a strong state ($\hat{p}$= 0.063, 1/$\hat{p}$=15.9)).

The simple Markov chain model introduced here is an over-simplification of the real regime behaviour and it's connection to the stratosphere, but nonetheless helps to illustrate why diagnosing stratosphere-troposphere coupling is difficult with the limited data record available. Typically, daily composites following SSW or strong vortex events are used to diagnose the strength and extent of stratosphere-troposphere coupling in re-analysis datasets and model integrations. The analysis above suggests that for composites of 30-100 events, daily estimates of the strength of the coupling will be strongly influenced by noise from the underlying tropospheric regime dynamics. A simple estimate of the size of this effect can be made numerically by simulating a large number of 60-day Markov chains with the same properties as the weak vortex chain above and constructing daily composites of the NAO- frequency.

For a large enough sample size, the daily average frequency of the NAO- regime should be 0.33 for each of the days in the composite. Variations in the daily NAO- frequency related to insufficient sampling can be quantified by calculating the daily standard deviation. In a toy experiment in which this process was repeated 100 times for a sample of 30 weak vortex events, the average daily standard deviation is 0.08. Even for a sample size of 100 events, the average daily standard deviation is 0.05. This result implies that composites of stratosphere-troposphere coupling events (e.g. Baldwin and Dunkerton 2001) likely contain significant sampling noise on the daily timescale and should only be used with some caution.

3.3. Quantifying stratosphere-troposphere coupling

In order to quantify coupling between the stratosphere and tropospheric weather regimes and compare coupling in different datasets and models it is necessary to develop a more comprehensive statistical model that can account for the effect of changes to the strength of the stratospheric vortex and the likelihood of transition between different weather regimes.

For this purpose, a logistic regression model is used. Logistic models are standard tools used in many branches of applied sciences where binomial observations are encountered (Peng *et al.* 2002). In the case of the regime data, this naturally forms a set of binomial variables ($R_i(t)$) for each regime where $R_i(t)$=1 indicates the weather regime on day t is regime i. Here the i refers to each of the different weather regimes (NAO-, NAO+, AR, BL) in turn. The logistic model predicts the probability (π_i) that a given set of antecedent conditions today (the current weather regime and strength of the zonal mean zonal wind at 100hPa) will lead to a particular weather regime tomorrow. In order to deal with the non-linear nature of problems with binary response variables, a logistic regression model applies the logit transform to the dependent variable (in this case the probability of occurrence of a weather regime), where:

$$\text{logit}(O_i) = \ln\left(\frac{\pi_i}{1-\pi_i}\right) \tag{9}$$

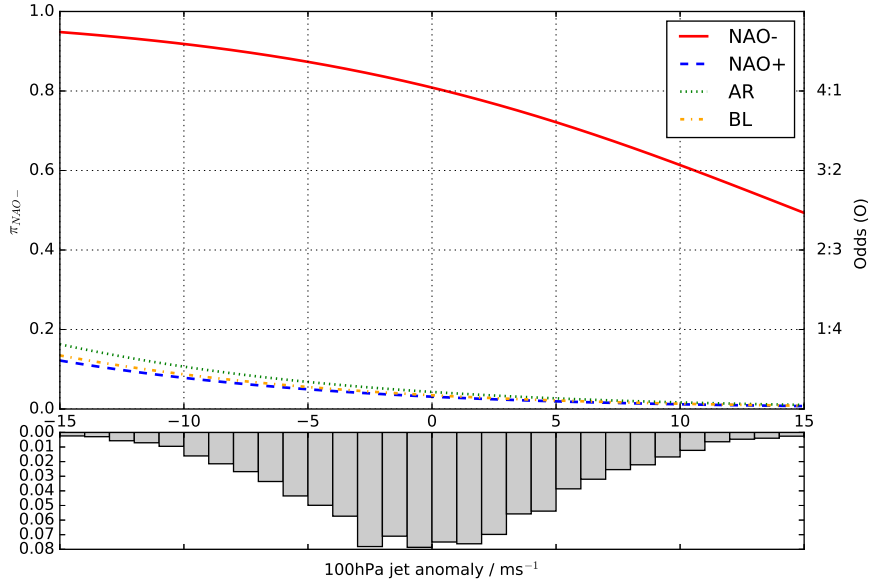

Figure 6. Model prediction for the NAO- regime for the logistic model fit to ERA-Interim data. Lines show the variation in the probability of occurrence of the NAO- regime as a function of the stratospheric wind anomaly. Solid, red line shows the probability when the previous state is also NAO-. The other coloured lines show the probability when the previous state is NAO+ (dashed, blue line), AR (dotted, green line) and BL (dash-dotted, orange line). A histogram of the stratospheric wind anomaly is shown below the x-axis.

The regression model is then very similar to a standard Ordinary Least Squares regression:

$$\text{logit}(O_i(t{+}1)) = \alpha U_{100}(t) + \sum_{j \neq i} \beta_j R_j(t) + c. \tag{10}$$

In the above expression, $R_j(t)$ is a binary indicator variable for each of the weather regimes other than the one which the model predicts. The model is fitted using each of the four weather regimes in turn. The parameters of the model (α, β_j, c) are then used to diagnose the strength of stratosphere-troposphere coupling for each regime. Note that this model is linear in the log-odds of the predicted regime but not the probability. (Typically, parameters of the model are expressed as an odds-ratio, which for the β_j parameters is the ratio between the odds when the prior regime does and does not occur.

The nature of the data means that this model cannot simply be fit in the standard manner using a maximum likelihood estimate. Since daily data for each winter season is used, it is strongly autocorrelated within each season (breaking the required assumption of independence) but has little autocorrelation between seasons. Therefore, we make use of an alternative quasi-likelihood method, the Generalized Estimating Equations (Ziegler and Vens 2010) which explicitly allows for this correlation structure when fitting the model. The model is implemented and fit using the statsmodels (Seabold and Perktold 2010) and pandas libraries (McKinney 2010), freely available for python.

An example of the diagnosis of coupling which can be gained from the model is shown in Fig. 6. The red line in this figure shows the predicted probability of occurrence of the NAO- regime as a function of the stratospheric wind anomaly ($\pi_{\text{NAO}-}$), when the existing regime is also NAO-. As expected, the model shows a large increase in probability of the NAO- regime when stratospheric winds are weak and a reduced probability when stratospheric winds are strong. The other coloured lines show the predicted occurrence probability of the NAO- regime when the existing regime is NAO+, AR or BL. There are small differences between the three antecedent regimes, but the largest difference is between regime persistence (the red line) and regime transition. For example, when stratospheric wind anomalies are zero, the probability of transition from any of the three other regimes to NAO- is small (around 0.04). Therefore the model captures a similar picture of the coupling as that shown in Fig. 4 and Fig. 5.

Comparison of the behaviour for all four regimes can be made by fitting the model with each of the regimes as the dependent variable in turn. Central estimates for the model parameters (α, β_j, c) are shown in Table 1. Stratosphere-troposphere coupling for each of the regimes is described by the odds ratio for the zonal wind predictor. For the NAO- case, the odds ratio $(\text{O}_{(\text{NAO}-|\text{U}_{100}=5)}/\text{O}_{(\text{NAO}-|\text{U}_{100}=0)})$ is 0.61 for a $5ms^{-1}$ change in the stratospheric zonal wind anomaly (approximately one standard deviation). Comparing with the red line in Fig. 6 helps to interpret the odds ratio. In Fig. 6 the probability of occurrence of the NAO- state for a wind anomaly of $0ms^{-1}$ and a prior NAO- state (red line) is approximately 0.8, or odds of 4:1. The diagnosed odds ratio of 0.61 indicates that the odds of the NAO- state for a wind anomaly of $5ms^{-1}$ should be 61% of 4:1 which are odds of 2.4:1 or a probability of approximately 0.7 as shown by the red line.

The models for the NAO+ and AR regimes have an odds ratio greater than one for the stratospheric wind anomaly as expected, indicating increased probability of occurrence for positive wind anomalies. However, the coupling strength for these regimes is much smaller than for the NAO- regime, a 16% increase in the odds of NAO+ and AR regimes for a $5ms^{-1}$ change in the stratospheric wind. Furthermore, the BL regime does not show significant sensitivity to the stratospheric wind anomaly.

These results are consistent with the idealised modelling study of Garfinkel *et al.* (2013). The NAO- regime (typically associated with a jet located around 40°N (Madonna *et al.* 2017) is more sensitive to stratospheric perturbations than regimes associated with jets located further poleward (NAO+ and AR).

Table 1. Central estimates of model parameters for ERA-Interim data. Where an estimated parameter is not significant at a p-value of 0.05 it is written in italic. The intercept is shown as the probability of remaining in the dependent weather regime when the stratospheric wind anomaly is zero (it can be transformed back into the model parameter through the logit transform). The parameter for the stratospheric zonal mean zonal wind anomaly is shown as the odds ratio ($O_{(i|U_{100}=5)}/O_{(i|U_{100}=0)}$) of the change for a $5ms^{-1}$ (an approximately 1 s.d.) change in wind anomaly. Parameters for weather regimes are shown as the odds ratio ($O_{(i|j=j)}/O_{(i|j\neq j)}$). An odds ratio of one indicates no change in the odds for a given model predictor. Odds ratios greater than one indicate increases in the odds of the dependent weather regime. Odds ratios less than one indicate decreases in the odds of the dependent weather regime.

	NAO-(t+1)	NAO+(t+1)	AR(t+1)	BL(t+1)
Intercept	0.81	0.81	0.80	0.75
Stratospheric zonal wind(t)	0.61	1.16	1.16	*0.93*
NAO-(t)		0.01	0.02	0.02
NAO+(t)	0.01		0.03	0.02
AR(t)	0.01	0.02		0.03
BL(t)	0.01	0.02	0.03	

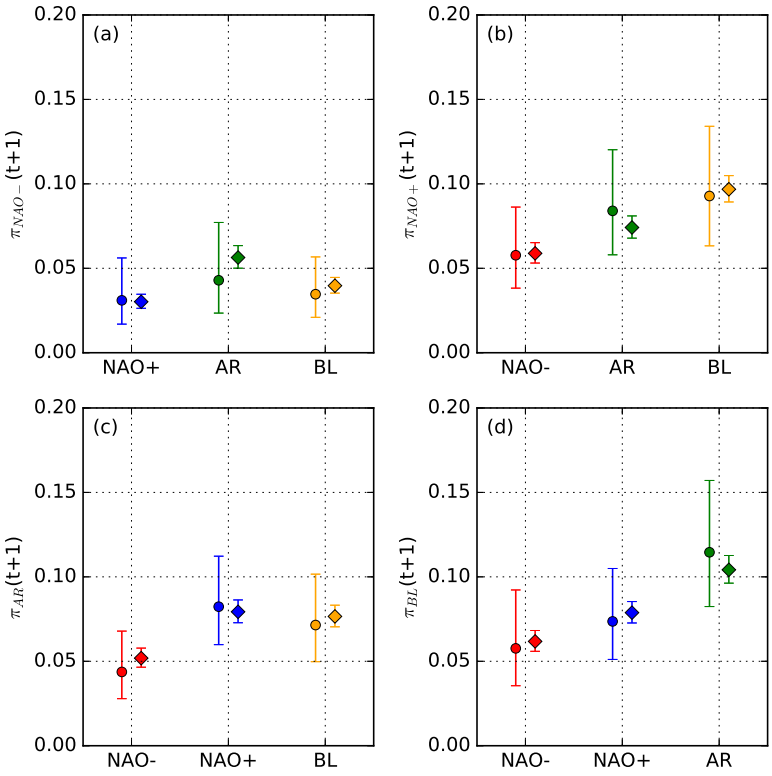

Figure 7. Model estimated transition probabilities for the (a) NAO-, (b) NAO+, (c) AR and (d) BL regimes from each of the other regimes. Filled dots show the estimate from the ERA-Interim dataset and diamonds show the estimate for the ECMWF IFS dataset. Whiskers on each estimate show the 95% confidence interval.

As the odds ratios for regime transitions are so small, reflecting the low-frequency memory of the weather regimes as described by the simple Markov model, it is helpful to calculate and compare changes to the occurrence probability for a specific stratospheric state (π_i). Figure 7 shows model estimates of the transition probabilities for each of the four regimes. The transition probabilities are largest for the NAO+ and BL regimes, reflecting the higher occurrence of these regimes in the dataset, and are smallest for the NAO- regime. The transition probabilities are consistent with the right panel of Fig. 3 giving confidence in the fit of the model. For the ERA-Interim data, although, for example in the NAO- case, the transition probability from AR to NAO- is larger than NAO+ to NAO-, the confidence intervals for the two estimates overlap. When considering transitions to all four regimes, similar arguments hold, it appears that some regime transitions are preferred above others but this cannot be conclusively stated given the limited re-analysis data record available. We will return to this point when discussing the ECMWF IFS dataset below.

The model can also be used to diagnose how stratospheric wind anomalies influence transition probabilities for each regime. For example for the NAO- regime, the model estimate of the impact of an approximately one standard deviation change in the stratospheric wind anomaly can be calculated from:

$$\Delta p(NAO-(t+1)) = p(NAO-(t+1)|R_j(t)=1 \& U_{100}(t)=-5) - p(NAO-(t+1)|R_j(t)=1 \& U_{100}(t)=0) \qquad (11)$$

This quantity is shown in Fig. 8(a). Equivalent quantities for the NAO+ and AR regimes (although this time representing a difference in probability between a positive 5ms^{-1} and zero lower stratospheric wind anomaly) are shown in Fig. 8(b) and (c).

The changes in transition probability associated with a one standard deviation zonal mean zonal wind anomaly for the NAO- regime are much larger than for the NAO+ regime as expected from Table 1. The average sensitivity of the transition probability to stratospheric wind anomalies is consistent with the simple analysis shown in Fig. 3. Changes of the transition probability are between 0.015 and 0.025 for transitions to the NAO- regime, a fractional change of 40-60% while changes of the transition probability for transitions to

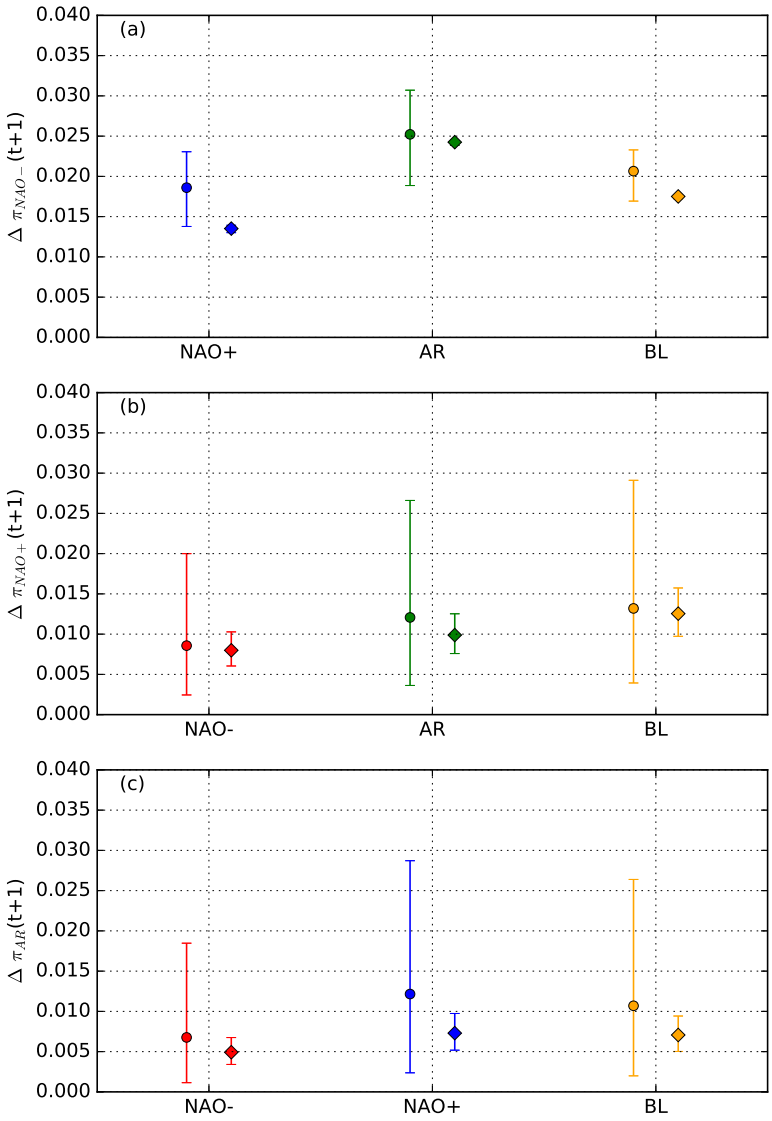

Figure 8. Model estimated impact of an approximately one standard deviation ($5ms^{-1}$) change to the stratospheric zonal wind anomaly on the transition probabilities for the (a)NAO-, (b) NAO+ and (c) AR regimes from each of the other regimes. Filled dots show the estimate from the ERA-Interim dataset and diamonds show the estimate for the ECMWF IFS dataset. Whiskers on each estimate show the 95% confidence interval. For the NAO- regime the estimated change is from 0 to $-5ms^{-1}$ and for the NAO+ regime the estimated change is from 0 to $+5ms^{-1}$ in order that the changes have the same sign.

the NAO+ regime are between 0.005 and 0.015, a fractional change of 10-30%. Very similar changes occur for transition to the AR regime (Fig. 8(c))

3.4. Stratosphere-troposphere coupling in the ECMWF model

The same statistical model used to diagnose stratosphere-troposphere coupling in the ERA-Interim dataset can also be fit to the larger ECMWF IFS reforecast dataset. The purpose of this analysis is two-fold, firstly it provides a means of comparing coupling in the ECMWF model to the re-analysis in order to understand if it is consistent with that diagnosed in the re-analysis. Secondly, if the coupling is broadly consistent, then the much larger ECMWF IFS dataset can reveal differences in the transition probabilities between different regimes which cannot be consistently diagnosed in the smaller ERA-Interim dataset.

An identical logistic model is fit to the reforecast dataset, using the same Generalized Estimating Equations fitting method. Clusters in the IFS dataset are the individual ensemble members of each forecast, we assume that there is autocorrelation between consecutive days of each forecast but no correlation between each forecast member.

Table 2 shows central estimates of the model parameters, analogously to the parameters shown in 1. All model parameters are consistent between the two datasets, with overlapping confidence intervals (not shown). This indicates that both the weather regime behaviour and stratosphere-troposphere coupling in the IFS model is representative of the re-analysis and so can be used as a useful proxy for the real atmosphere.

Model estimates of the transition probabilities (π_i) in the IFS model are shown as diamonds in Fig. 7. Central estimates for the transition probabilities are quite similar for the ERA-Interim and IFS datasets, with the same preference for particular regime transitions for all four model fits. The much larger dataset available for the IFS model, means that confidence estimates for the transition probabilities are much smaller, giving confidence that the differences in transition probabilities between regimes, while small, are robust. For the NAO- regime (Fig. 7(a)), the transition probability from the AR regime is largest, approximately double the transition probability from the NAO+ regime. For the NAO+ regime (Fig. 7(b)), there is a clear ordering with the largest transition probability

Table 2. Central estimates of model parameters for ECMWF IFS model data. Where an estimated parameter is not significant at a p-value of 0.05 it is written in italic. The intercept is shown as the probability of remaining in the dependent weather regime when the stratospheric wind anomaly is zero (it can be transformed back into the model parameter through the logit transform). The parameter for the stratospheric zonal mean zonal wind anomaly is shown as the odds ratio ($O_{(i|U_{100}=5)}/O_{(i|U_{100}=0)}$) of the change for a $5ms^{-1}$ (an approximately 1 s.d.) change in wind anomaly. Parameters for weather regimes are shown as the odds ratio ($O_{(i|j=j)}/O_{(i|j\neq j)}$). An odds ratio of one indicates no change in the odds for a given model predictor. Odds ratios greater than one indicate increases in the odds of the dependent weather regime. Odds ratios less than one indicate decreases in the odds of the dependent weather regime.

	NAO-(t+1)	NAO+(t+1)	AR(t+1)	BL(t+1)
Intercept	0.81	0.81	0.78	0.76
Stratospheric zonal wind(t)	0.68	1.15	1.10	*1.00*
NAO-(t)		0.01	0.03	0.02
NAO+(t)	0.01		0.03	0.02
AR(t)	0.01	0.02		0.03
BL(t)	0.01	0.03	0.03	

from the BL regime, followed by the AR and then the NAO- regime. For the AR regime, (Fig. 7(c)) the transition probability from the NAO- regime is smaller than the transition probability from the NAO+ or BL regimes. For the BL regime (Fig. 7(d)), there is a clear ordering in transition probability with highest probability for the AR regime, followed by the NAO+ and NAO- regime.

Finally, it is useful to compare the sensitivity of regime transitions to the stratospheric state in the IFS model with the re-analysis. The diamonds in Fig. 8 show the change in the transition probability due to a 5ms^{-1} change in the strength of the stratospheric jet. In all cases, the change in transition probability is consistent with the estimates from the ERA-Interim analysis, but at the lower end of the confidence interval. For the NAO- regime, there is a clear seperation between the change in transition probability for the three possible precursor regimes. Transitions from AR to NAO- regime are most affected by stratospheric changes ($\Delta p \approx 0.025$) followed by transitions from the BL regime to NAO- and then transitions from the NAO+ to NAO- regime. For the NAO+ and AR regimes, even with the larger IFS dataset, similar significant differences in the impact of stratospheric winds on transition probabilities are not present.

4. Discussion and Conclusions

In this study, the influence of stratospheric conditions on the North Atlantic troposphere is analysed by considering changes to weather regime frequency and transition. A clear sensitivity of the NAO-, NAO+ and AR regimes to stratospheric conditions is demonstrated, consistent with that shown by Beerli (2017). The simple model of regime frequency as a one-step Markov process which was developed in section 3.1, makes it possible to demonstrate that this change in regime frequency is the result of changes in persistence probability for the NAO-, NAO+ and AR regimes and the probability of transition into the same three regimes.

With this model in mind, a more complex logistic regression model can be fit to both the ERA-Interim re-analysis and the ECMWF IFS model. An important primary conclusion of this analysis is that the coupling between the stratosphere and North Atlantic in the IFS model is consistent with the re-analysis. By fitting a logistic regression model it is possible to be confident about this result, because it makes use of the full dataset in each case producing model fit parameters which are highly significant. A standard approach to comparing stratosphere-troposphere coupling in models by using composites of extreme events in the stratosphere would use a much smaller fraction of the data and would be hampered by the kinds of sampling uncertainties discussed in section 3.2.

The logistic model shows that there is significantly non-linearity in the response of the troposphere to stratospheric anomalies, consistent with Garfinkel *et al.* (2013). The NAO- regime is the most sensitive to changes in the stratospheric state, with smaller but still significant sensitivity for the AR and NAO+ regimes. The BL regime is not sensitive to the stratospheric state in either the ERA-Interim and ECMWF IFS datasets. Furthermore, for the transitions to the NAO- regime, the antecedent North Atlantic regime is an important determinant of the North Atlantic response. For example, transitions between the AR and NAO- are enhanced to a greater extent than transitions between the NAO+ and NAO- for the same change in stratospheric winds. These two results are important both for forecasters seeking to make use of stratosphere-troposphere coupling when diagnosing model responses for individual SSW or strong vortex events and for model developers.

Understanding the dynamical reasons why the AR regime is more sensitive to stratospheric changes is beyond the scope of this study, but recent work (e.g. Michel and Rivière 2011; Swenson and Straus 2017) which links regime transitions to changes in Rossby Wave Breaking (RWB) points to a possible direction for further study. In particular, transitions to NAO- states are often preceded by enhancement of cyclonic RWB over North America and Greenland, accompanied by weak zonal wind anomalies and enhanced horizontal wind shear on the poleward side of the jet. Similar easterly zonal wind anomalies would be produced in the lower stratosphere and upper stratosphere on the poleward side of the jet during SSW events.

Acknowledgement

We thank Remo Beerli, Christian Grams and Heini Wernli for useful dicussions and comparison of our analysis with their forthcoming work. RWL was funded by NERC grant NE/P006787/1, the InterDec project funded by Belmont Forum and JPI-Climate Collaborative Research Action on Climate Predictability and Inter-Regional Linkages. We thank two reviewers for their helpful and constructive reviews of our initial submission.

References

Baldwin MP, Dunkerton TJ. 1999. Propagation of the Arctic Oscillation from the stratosphere to the troposphere. *Journal of Geophysical Research: Atmospheres* **104**(D24): 30 937–30 946.

Baldwin MP, Dunkerton TJ. 2001. Stratospheric harbingers of anomalous weather regimes. *Science* **294**(5542): 581–584.

Beerli R. 2017. Sources of sub-seasonal predictability for energy industry-relevant weather events. *qjrms* (ETH No. 24488).

Beerli R, Wernli H, Grams CM. 2017. Does the lower stratosphere provide predictability for month-ahead wind electricity generation in Europe? *Quarterly Journal of the Royal Meteorological Society* **In press**.

Cassou C. 2008. Intraseasonal interaction between the Madden-Julian Oscillation and the North Atlantic Oscillation. *Nature* **455**(7212): 523.

Cassou C, Terray L, Hurrell JW, Deser C. 2004. North Atlantic winter climate regimes: Spatial asymmetry, stationarity with time, and oceanic forcing. *Journal of Climate* **17**(5): 1055–1068.

Davini P, Cagnazzo C, Anstey J. 2014. A blocking view of the stratosphere-troposphere coupling. *Journal of Geophysical Research: Atmospheres* **119**(19).

Dawson A, Palmer T, Corti S. 2012. Simulating regime structures in weather and climate prediction models. *Geophysical Research Letters* **39**(21).

Dee DP, Uppala S, Simmons A, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P, Bechtold P, Belijaars A, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer A, Haimberger L, Healey S, Hersbach H, Holm E, Isaksen L, Kallberg P, Kohler M, Matricardi M, McNally A, Monge-Sanz B, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thepaut JN, Vitart F. 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society* **137**(656): 553–597.

Ferranti L, Corti S. 2011. New clustering products. *ECMWF Newsletter* **127**: 6–11.

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

Garfinkel CI, Waugh DW, Gerber EP. 2013. The effect of tropospheric jet latitude on coupling between the stratospheric polar vortex and the troposphere. *Journal of Climate* **26**(6): 2077–2095.

Grams CM, Beerli R, Pfenninger S, Staffell I, Wernli H. 2017. Balancing europe?s wind-power output through spatial deployment informed by weather regimes. *Nature climate change* **7**(8): 557.

Hannachi A, Straus DM, Franzke CL, Corti S, Woollings T. 2017. Low Frequency Nonlinearity and Regime Behavior in the Northern Hemisphere Extra-Tropical Atmosphere. *Reviews of Geophysics* .

Hitchcock P, Simpson IR. 2014. The downward influence of stratospheric sudden warmings. *Journal of the Atmospheric Sciences* **71**(10): 3856–3876.

Hitchcock P, Simpson IR. 2016. Quantifying Eddy Feedbacks and Forcings in the Tropospheric Response to Stratospheric Sudden Warmings. *Journal of the Atmospheric Sciences* **73**(9): 3641–3657.

Karpechko AY. 2015. Improvements in statistical forecasts of monthly and two-monthly surface air temperatures using a stratospheric predictor. *Quarterly Journal of the Royal Meteorological Society* **141**(691): 2444–2456.

Kidston J, Scaife AA, Hardiman SC, Mitchell DM, Butchart N, Baldwin MP, Gray LJ. 2015. Stratospheric influence on tropospheric jet streams, storm tracks and surface weather. *Nature Geoscience* **8**(6): 433.

Limpasuvan V, Hartmann DL, Thompson DW, Jeev K, Yung YL. 2005. Stratosphere-troposphere evolution during polar vortex intensification. *Journal of Geophysical Research: Atmospheres* **110**(D24).

Limpasuvan V, Thompson DW, Hartmann DL. 2004. The life cycle of the northern hemisphere sudden stratospheric warmings. *Journal of Climate* **17**(13): 2584–2596.

Madonna E, Li C, Grams CM, Woollings T. 2017. The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Quarterly Journal of the Royal Meteorological Society* .

McKinney W. 2010. Data structures for statistical computing in python. In: *Proceedings of the 9th Python in Science Conference*, van der Walt S, Millman J (eds). pp. 51 – 56.

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

Michelangeli PA, Vautard R, Legras B. 1995. Weather regimes: Recurrence and quasi stationarity. *Journal of the Atmospheric Sciences* **52**(8): 1237–1256.

Peng CYJ, Lee KL, Ingersoll GM. 2002. An introduction to logistic regression analysis and reporting. *The Journal of Educational Research* **96**(1): 3–14.

Scaife A, Karpechko AY, Baldwin M, Brookshaw A, Butler A, Eade R, Gordon M, MacLachlan C, Martin N, Dunstone N. 2016. Seasonal winter forecasts and the stratosphere. *Atmospheric Science Letters* **17**(1): 51–56.

Seabold J, Perktold J. 2010. Statsmodels: Econometric and statistical modeling with python. In: *Proceedings of the 9th Python in Science Conference*.

Sheshadri A, Plumb RA. 2017. Propagating annular modes: Empirical orthogonal functions, principal oscillation patterns, and time scales. *Journal of the Atmospheric Sciences* **74**(5): 1345–1361.

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

Sparrow S, Blackburn M, Haigh JD. 2009. Annular variability and eddy–zonal flow interactions in a simplified atmospheric gcm. part i: Characterization of high-and low-frequency behavior. *Journal of the Atmospheric Sciences* **66**(10): 3075–3094.

Swenson ET, Straus DM. 2017. Rossby Wave Breaking and Transient Eddy Forcing during Euro-Atlantic Circulation Regimes. *Journal of the Atmospheric Sciences* **74**(6): 1735–1755.

Tripathi OP, Baldwin M, Charlton-Perez A, Charron M, Eckermann SD, Gerber E, Harrison RG, Jackson DR, Kim BM, Kuroda Y, Lang A, Mahmood S, Mizuta R, Roff G, Sigmond M, Son SW. 2015a. 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.

Tripathi OP, Charlton-Perez A, Sigmond M, Vitart F. 2015b. Enhanced long-range forecast skill in boreal winter following stratospheric strong vortex conditions. *Environmental Research Letters* **10**(10): 104 007.

Woollings T, Charlton-Perez A, Ineson S, Marshall A, Masato G. 2010a. Associations between stratospheric variability and tropospheric blocking. *Journal of Geophysical Research: Atmospheres* **115**(D6).

Woollings T, Hannachi A, Hoskins B. 2010b. Variability of the North Atlantic eddy-driven jet stream. *Quarterly Journal of the Royal Meteorological Society* **136**(649): 856–868.

Ziegler A, Vens M. 2010. Generalized estimating equations. *Methods of Information in Medicine* **49**(5): 421–425.

Zubiate L, McDermott F, Sweeney C, O'Malley M. 2017. Spatial variability in winter nao?wind speed relationships in western europe linked to concomitant states of the east atlantic and scandinavian patterns. *Quarterly Journal of the Royal Meteorological Society* **143**(702): 552–562, doi:10.1002/qj.2943, URL http://dx.doi.org/10.1002/qj.2943.