Response to Reviewers

We would like to thank both reviewers for their time and effort to review our manuscript. We here list the major changes that we made in the manuscript during the revision:

- Role of the lower stratospheric persistence in the downward response: We have now included an extended discussion and analysis of the lower stratospheric persistence in the Introduction, Results, and Discussion sections of our manuscript.
- Role of weather regimes: We have rewritten the manuscript to omit causal statements and emphasise more that regimes at the onset might be an indicator for the potential surface impact of an SSW.
- Statistical testing: We have carefully re-written and significantly extended the section about statistical testing. In particular, the hypotheses have now been clearly formulated.
- Case studies: We have removed the case studies at the suggestion of the reviewer.
- Persistence of the tropospheric weather regimes and causality: We have clarified the relevant timescales and persistence of the regimes in the new version of the manuscript. We have also clarified the definition of the weather regimes and their evolution. In particular, we are indeed not suggesting that the response 6-8 weeks after the SSW onset is a direct cause of the weather regime at lag 0.
- We have limited the Figure showing the surface temperature response to lags 0 - 25 days allowing us to include the composite for all SSWs for better comparability.

Please find below the detailed responses (in *cursive*) to the reviewers' comments and suggestions. All line indications refer to the new (annotated) manuscript at the end of the reviewer response.

Reviewer 1:

General comments: The paper investigates response in the troposphere in the North Atlantic/European (NAE) region following stratospheric sudden warming (SSW) events in reanalysis data. The study finds that Greenland blocking and Atlantic trough (AT) are more likely weather regimes weeks after SSWs. In addition, the study investigates the role of tropospheric weather patterns during SSW onset in the subsequent tropospheric response. It is found that it is mostly for SSWs with European blocking at their onset that the canonical response of cold surface extremes over Europe is observed weeks following SSWs. In contrast, for SSWs with AT at the onset, mild conditions over NAE region after SSWs are observed. The remaining tropospheric flow patterns at the onset of SSWs were not associated with clear surface response following SSWs.

Given the large case to case variability of surface responses following SSWs, this study is a step in the right direction in trying to further understand when an SSW is likely to be followed by surface extremes. The paper shows that not all SSWs are followed by anomalous tropospheric weather patterns, therefore suggesting that caution must be exercised when generalising results from composite analysis involving all SSWs. Such knowledge is important for subseasonal to seasonal predictability when trying to assess if the downward impact from SSW is to be expected.

I commend the authors on statistical rigour and strongly encourage to carry out similar analysis (in the future) in the context of sub seasonal to seasonal prediction models where the robustness of the results to sampling uncertainty and the impact on predictability can both be assessed. I recommend this paper for publications and have only very minor comment detailed below.

We would like to thank the reviewer for their positive evaluation of our manuscript. We fully agree that extending the analysis to S2S prediction models will be very worthwhile, and we have commented on this in the new manuscript at the end of the Discussion section (lines 311 - 319).

Specific comment:

It would be helpful if the authors could put their study in the context of previously published studies that assessed which SSWs give stronger response. In particular the persistence and amplitude of lower stratospheric anomalies following SSWs are known to affect surface response (Hitchcock et al, 2013, JClim; Kodera et al., 2016, JGR; Runde et al., 2016, GRL; Karpechko et al, 2017; QJRMS; Polichtchouk et al, 2018, JAS). For example, is there any evidence to suggest that SSWs that had European Blocking at the onset also have larger and longer lasting anomalies in the lower stratosphere than the other cases?

We agree that the persistence of lower stratospheric anomalies is a very important issue, and we have included a more thorough discussion of this matter and the relevant citations in the manuscript. Our analysis reveals that 2 out of 5 SSW events with the GL regime at the onset have a persistent lower stratospheric temperature response, while 4 out of 6 EuBl cases have a persistent lower stratospheric temperature response. This goes in the right direction by indicating that the shorter (longer) persistence in the lower stratosphere for the SSWs associated with GL (EuBl) may add some support to the persistence of the tropospheric response, but the statistics are too small to provide a clear result. We have included a more extensive discussion of the above references and lower stratospheric persistence into the manuscript (Introduction: lines 44 - 55, Results: lines 232-245, Discussion: lines 303 - 310). For Polichtchouk et al 2018, JAS: After discussion with Inna Polichtchouk we decided that her 2018 GRL paper might be most relevant here, hence we have now included a reference to her 2018 GRL paper instead.

Technical comment:

P7, L180: "NAM), the number" → "NAM), and the number"

Thanks for spotting this, this has been corrected (line 202).

Reviewer 2

Summary

The study asks whether the North Atlantic European (NAE) weather regime present at the onset of an SSW has a bearing on the subsequent evolution of the tropospheric state. The topic fits within the scope of WCD and a study on this topic fits more broadly into a body of literature that has investigated possible factors to explain why some SSWs appear to couple to the troposphere and others do not. The study uses ERA-Interim reanalysis data (26 SSWs) and views these through the lens of 7 NAE weather regime types introduced in earlier work by one of the co-authors. The authors conclude that European Blocking at the time of SSW onset favours Greenland Blocking in the subsequent weeks, while Greenland blocking at onset favours an Atlantic Trough following the SSW. One major limitation of the study is the small sample of observed SSWs, which are then subdivided across the 7 regimes. This leaves only small samples for each subcategory. Earlier studies provide a cautionary tale about interpreting small subsets of SSWs (e.g., Mitchell et al. (2013) and Maycock and Hitchcock (2015)), and the authors fall foul to some of these issues. The authors undertake bootstrap analyses to test for the significance of results, but this is mainly comparing to samples drawn from non-SSW periods. If the purpose is to test whether knowledge of the NAE state at the onset of an SSW can provide additional knowledge over and above knowledge of an SSW, the null hypothesis should be either that the tropospheric state following SSWs with a given day-0 regime is not distinguishable from that for all SSWs and/or that it is not distinguishable from SSWs with a different day-0 regime. This requires calculating differences (and their significance) between the regime subsets. The authors also make no attempt to rule out other confounding factors that might affect their interpretation of the role of NAE regimes. For example, studies have found a relationship between the amplitude of lower stratospheric anomalies around the onset and the subsequent tropospheric NAM response. This was also pointed out by reviewer 1, but I think it is hugely important for the interpretation of the present results. The manner of presentation implicitly assumes the differences are a consequence of the day-0 regime, but since no other factors are tested for or displayed it is impossible to determine whether this is the case. This is especially pertinent given the small sample sizes being dealt with. Overall, while the topic itself is potentially interesting, I found the manuscript disappointing both in terms of setting out the motivation for why/how the NAE state could have a long-lasting impact on the subsequent response and in terms of weaknesses in the analysis that I did not feel support the conclusions for the added value of knowing an SSW has occurred AND the day-0 NAE regime as compared to simply knowing anSSW has occurred. I therefore recommend to reject the manuscript in its current form. Recommendation: Reject.

We thank the reviewer for the thorough evaluation of the manuscript. We will answer the specific comments below.

Major comments

1) Hypotheses and statistical tests.

a) Your statistical test in Fig. 2 and 3 asks whether the SSW periods are different from non-SSW periods (climatology). This is fine for Figs 2a and 3a, where you ask about the overall signal of SSWs compared to no-SSWs, but what you are asking in Fig 2b-d is whether knowledge of the day-0 NAE regime provides extra information over and above the general knowledge of an SSW. Step 1) is there an SSW? Step 2) if yes, what is the NAE regime? Therefore, to my mind the relevant test is whether panels (b-d) are different from each other and/or different from (a). The same applies to Fig. 3. See e.g., important lessons from a parallel case on whether split vs. displacement SSWs show different coupling. Mitchell et al. (2013) performed similar analysis to that here for the NAM, but instead stratifying events based on split and displacement types (rather than on NAE type); importantly they neglected to test the significance of their differences, which was later done by Maycock and Hitchcock (2015) who estimated that the difference is not significant. You could do something similar here constructing a bootstrap distribution of the difference between two sets of N SSW samples.

We fully agree that the number of SSWs in the observational record of the satellite era is very small. Thus, we have put significant effort into the design of the statistical tests, as also commended by Reviewer 1, to take sampling uncertainty into account and obtain meaningful results. We recognize that in the manuscript this procedure may not have been explained in sufficient detail. We have expanded on this in the revised version of the manuscript (section 2.2). Furthermore, we clarify the procedure again in the following.

The overarching question we address in this study is whether after SSWs we can detect robust tropospheric geopotential height anomalies and whether these anomalies are significantly different from situations without an SSW. Hence, the relevant null hypothesis is that the tropospheric geopotential height anomalies after SSWs are indistinguishable from geopotential height anomalies commonly occurring in the absence of an SSW. The testing procedure follows two important steps:

1. *First, we assess the* ***robustness*** *of the samples by performing a Monte Carlo resampling. For that purpose, the dripping paint plots are re-computed by resampling the original samples 100 times with repetitions. This yields confidence intervals of the dripping paint plots, estimating the uncertainty inherent in each sample. Due to the small sample size, these confidence intervals are relatively large.*

2. *Second, we compute 1000 random samples of the same size as the original sample but for random periods with the same weather regime(s) at the central date but no SSW occurring within* ± 60 *days, yielding estimates of the distributions of geopotential height anomalies occurring in the absence of SSWs. Testing for **significance** is done by comparing the confidence intervals and distributions obtained from the random samples.*

Following this procedure, we thus show that the anomalies in the ALL composite are not robust at the 10 % level but at the 25 % (Fig. A4a in the manuscript), indicative of the large variability in the tropospheric response. Yet, the anomalies observed between 10-20 days after the SSW are statistically significantly different from non-SSW periods at the 10 % (Fig. 3a in the manuscript). Now, we can ask the question whether there are subsamples of all SSWs that show a more robust response that is also statistically different from periods with no SSW. We select these subsamples according to the weather regime present at the time of the SSW. Note that the null hypothesis here is still the same as for the ALL sample, namely that no significantly different anomalies occur. Thus, we find that particularly robust and significant anomalies occur if the SSW is timed with European blocking, for example.

The reviewer suggests to test for the null hypothesis that the geopotential height anomalies between samples are identical. This null hypothesis may be appropriate if we were interested in the question whether geopotential height anomalies between the samples are different from each other. While this may be an interesting alternate path of investigation, it is not the question we pose here. For the reviewer's interest, we have nevertheless done mutual tests of the Greenland blocking, European Blocking, and the cyclonic regimes samples (Figure 1). This reveals that the geopotential height anomalies in the European blocking case, for example, are significantly different from the Greenland blocking case. The difference is less significant with respect to the cyclonic regimes samples (especially the anomalies before day 20).

We would like to stress that mutually testing the individual samples for difference does not a priori tell anything about whether the flow evolution in the presence of an SSW is different from that in the absence of an SSW. Similar differences between samples may in principle also arise from samples obtained from random days with a given weather regime irrespective of whether an SSW occurred or not.

We hope that with these additional clarifications, we are able to convince the reviewer that our testing procedure is appropriate for the questions addressed in the study. In the revised manuscript, we have more carefully explained the statistical testing procedure in subsection 2.2 and discuss the appropriate null hypothesis.

b) Fig. 3: These dripping paint diagrams are notoriously sensitive to sampling uncertainty and for such small sample sizes I strongly question their representativity. Charlton and Polvani (2007) stated in relation to their assessment of the impacts of split and displacements (p.462, Section 6) “We started our analysis by first constructing time - height composites of the NAM index for the

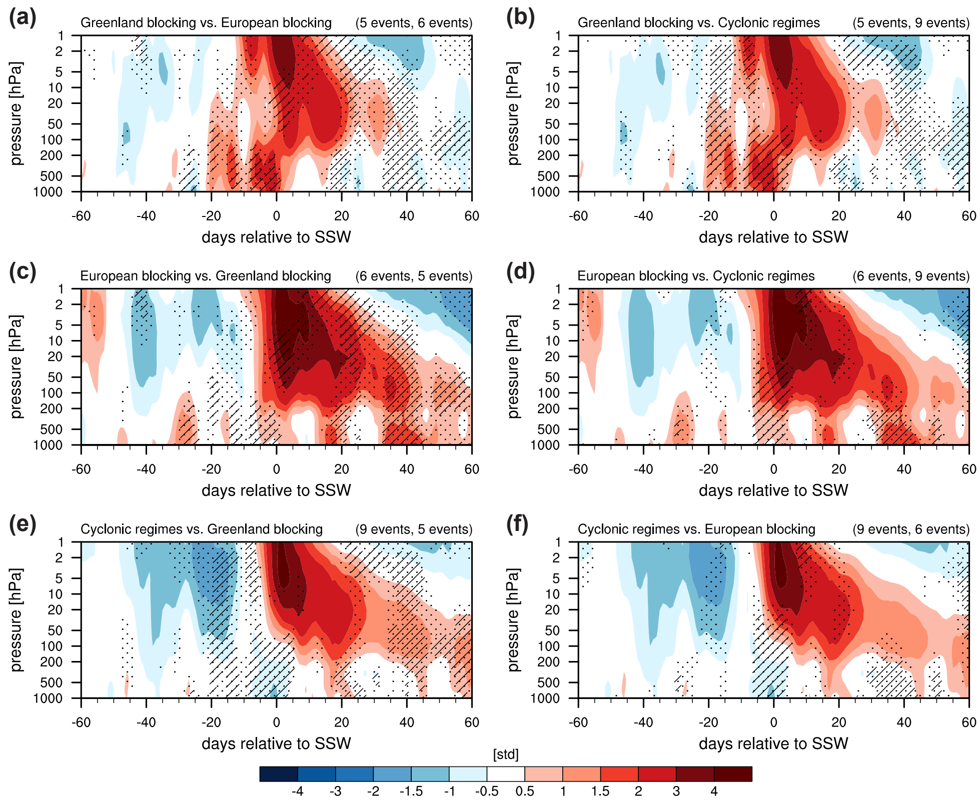

Figure 1: Standardized geopotential height anomalies as in Fig. 3 in the manuscript for (a, b) Greenland blocking cases, (c, d) European blocking cases, and (e, f) cyclonic regimes. Stippling and hatching indicates statistically significantly different anomalies from the other two samples (as indicated in the title of each panel) obtained from an overlap of the confidence intervals by less than 25 % and 10 %, respectively.

two types of SSW. However, the structure of the NAM index for the two types of SSW was found to be extremely sensitive, particularly in the troposphere: the size and timing of the composite NAM index anomalies following the events could be substantially altered by adding or removing even a single event. Hence, composite time-height NAM plots could not be used to examine differences in tropospheric impact between the vortex splits and vortex displacements." They made this point in relation to splits and displacements which have bigger sample sizes than those considered here. A similar point was also made by Maycock and Hitchcock (see e.g., their Fig. 3). Charlton and Polvani (2007) instead use the integrated NAM index to assess differences between splits and displacements. You could try an approach along these lines instead.

Thank you for this comment. Yes, we strongly agree that the small sample size requires careful testing. As already detailed in the response to the previous comment, we assess the robustness of the composites by a Monte Carlo resampling procedure (see answer to question 1a). Based on these we indeed find geopotential height anomalies that are significantly different from zero, that is, they are robust.

This is shown for each sample in Fig. A4 of the manuscript. For instance, the positive anomalies around day 0 in the Greenland blocking composites or those between 10-20 and around 40 days in the European blocking composites are robust at the 10 % level. In the ALL composite, in contrast, positive anomalies around day 10-20 are robust at the 25 % but not the 10 % level. This shows the large case-to-case variability of the tropospheric flow in the aftermath of an SSW. Importantly, this variability is reduced during specific times in the European blocking and Greenland blocking subsets of SSWs.

c) No attempt is made to rule out other possible associations than the NAE regime at day-0. For example, what if there is an indirect relationship to some other factor, such as the amplitude and persistence of the stratospheric anomalies themselves (e.g., Karpechko et al., 2017). There is some hint in Fig. 3 that the character of the stratospheric anomalies is different for these particular subsets of events; might that not be important? The sample sizes available here are very limiting in being able to say what is going on. To my mind, other more effective studies on related topics of downward coupling have combined reanalysis and model results (Karpechko et al., 2017; Maycock and Hitchcock, 2015). Reviewer 1 talks about following this up with a study on S2S models. If the authors do plan this, my recommendation would be to combine the current results with such a model study.

(see also response to Reviewer 1, specific comment) We agree that the persistence of lower stratospheric anomalies is a very important issue, and we have included a more thorough discussion of this matter and the relevant citations in the manuscript (Introduction: lines 44 - 55, Results: lines 232-245, Discussion: lines 303 - 310). In addition we have performed an analysis which reveals that 2 out of 5 SSW events with the GL regime at the onset have a persistent lower stratospheric temperature response, while 4 out of 6 EuBl cases have a persistent lower stratospheric temperature response. This goes in the right direction by indicating that the shorter (longer) persistence in the lower stratosphere for the SSWs associated with GL (EuBl) may add some support to the persistence of the tropospheric response, but the statistics are too small to provide a clear result.

Also, we fully agree that extending the analysis to S2S prediction models will be extremely worthwhile. However, it is currently unknown to what extent complex prediction models are indeed able to represent the variety of tropospheric responses to stratospheric forcing. Although simplified models indeed show a role of the troposphere in the downward impact, as the reviewer points out, this has not been sufficiently tested in more complex models beyond the canonical response and selected case studies. From a preliminary analysis of S2S prediction model data we anticipate large biases and a very complex role of the representation of stratosphere - troposphere coupling in prediction models. This will be complex to disentangle, and we will therefore not be able to cover the analysis of the model data in this study. We have now added a more comprehensive section on this in the Discussion section of the paper (lines 311-319).

d) L193-196 "The immediate positive geopotential height anomalies and the weak tropospheric response in the aftermath of the event are archetypal for SSWs with GL at the onset and not the result of cancellations in the composites. Indeed, they are also evident for individual SSWs, such as the SSW on 8-Dec-1987 that exhibited a dominant GL regime for an extended period around the SSW onset (Figs. 4a and 6a)." I find Figures 4 and 6 completely uninformative. You have chosen examples to support your proposed hypotheses, but the key information is what comes from the behaviour across all events, as shown in Fig 2 and 3. To give just one example, you could have chosen instead the GL event on 9 Feb 2010 which shows the GL regime for 3 weeks after the onset. Presumably this event is not "archetypal" but it is one of your 5 cases. I would also argue you cannot conclude something is "archetypal" when you have only 5 events. I suggest removing these arguably cherry picked case studies and providing more comprehensive evidence for a detectible difference between the subsets discussed would improve the manuscript. This also applies to the discussion of the 2018/2019 events, which I found too cursory and descriptive to provide any real insight.

As requested by the reviewer we have removed the case studies. To answer the reviewer's question about the 2010 SSW event, we have here included the dripping paint plot for the 2010 SSW event (Figure 2). The event again supports the general structure of the events that are dominated by GL at the onset of the SSW, with a strong tropospheric signal already before the onset, and the tropospheric response generally limited to the weeks just after the SSW event and including an AT response. We agree that the event exemplifies variability between events in the same sub-category, but we would also like to emphasize that this event as well as all others in the GL category indeed show a very different response to e.g. the EuBl cases.

2) Existence of plausible mechanism(s). L53-57: "Given the large tropospheric internal variability and the influence of other remote effects mentioned above, it appears plausible that also the tropospheric state at the time of occurrence of an SSW plays an important role in shaping the characteristics of its downward impact. For example, tropospheric jet characteristics have been suggested to affect the downward impact of SSW events (Chan and Plumb, 2009; Garfinkel et al., 2013)." I appreciate the goal of the study is not to explain but rather to diagnose, but this point is central to the whole premise of the study. However, the studies cited here are highly idealised and explore a much wider range of basic states than is plausible for the real world in idealised models that do not produce the type of NAE regime behaviour described here. I therefore do not agree this is supporting evidence for the proposed hypothesis. Indeed, no mechanism or theory is provided to justify why, or in what way, the tropospheric NAE state at day-0 would influence the subsequent NAE state up to +60 days. If that is the motivation to pursue this analysis, then some hypothesis for a mechanism is needed to explain an effect that extends far beyond the characteristic decorrelation timescale of the NAE circulation. It appears the proposal is for a vague mechanism related to internal tropospheric dynamics. However, this seems to defy the premise of

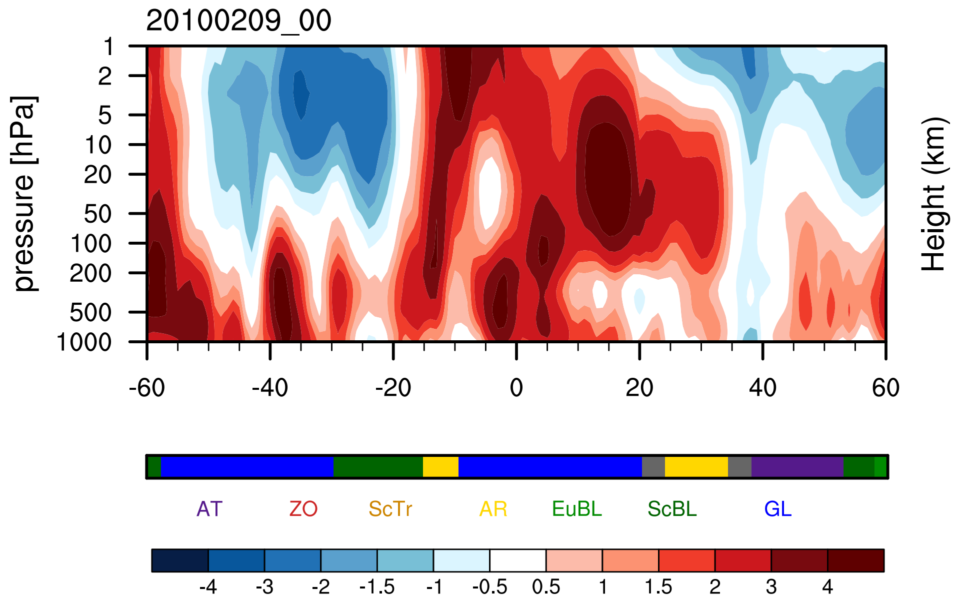

Figure 2: As Figure 4 in the original manuscript but for the SSW on Feb 9, 2010.

why SSWs are useful for predictability in the first place, which is that their intrinsic timescale is much longer than the 'memory' of the tropospheric circulation. To make a more convincing case for this, more discussion is needed around the persistence characteristics of the regimes themselves and the canonical transitions amongst the regimes to put the behaviour following SSWs into context.

We concur that our initial formulation was misleading as it implied the regime at the time of the SSW to determine the tropospheric flow several weeks later. Our intention is rather to highlight that the regime at SSW onset might be indicative of the subsequent SSW response for a limited set of SSW events. Indeed our results point in that direction for the GL and EuBL regimes. With our rewriting we are careful to only convey the observation of the diagnostic study and not to imply causality. In that sense we do also not suggest that the regime at SSW onset itself would have a deterministic response up to 60 days later which are indeed different timescales. It is rather that the SSW modulates the likelihood of some regimes at these subseasonal time scales. Furthermore, we have clarified the relevant timescales and persistence of the regimes vs. the SSW response directly in the introduction of the new version of the manuscript (lines 67 - 81). We have also removed the references to the idealized downward impact studies as suggested by the reviewer.

3) Timescales. Related to 2), a more careful description of the relevant timescales is needed in the introduction. Since the downward influence of SSWs may last for up to 6-8 weeks are the authors proposing that the NAE regime on day 0 bears some relevance for the response in week

6? Or are the authors talking about the downward coupling over a shorter period following the onset, e.g. in week 1? This does not become clear until one gets into the results, so some explicit statements on timescales in the abstract and introduction would clarify this and this should tie into the discussion of mechanisms.

(see also answer to query 2 above) We have clarified the relevant timescales in the introduction section (lines 67 - 81). We are indeed not suggesting that the response 6-8 weeks after the SSW onset is a direct consequence of the the weather regime at lag 0. We have formulated this more carefully throughout the revised manuscript by emphasising that rather the stratospheric state might modulate conditions that favor the occurrence or transition of some regimes.

4) Dataset. Why is ERA-Interim used and not a longer reanalysis like JRA-55 which contains more SSWs (41 compared to 26 in Butler et al (2017))? For rare events, the benefits of increased sample size can outweigh other uncertainties in the pre-satellite era (Hitchcock, 2019).

Hitchcock, P.: On the value of reanalyses prior to 1979 for dynamical studies of stratosphere-troposphere coupling, Atmos. Chem. Phys., 19, 2749?2764, https://doi.org/10.5194/acp-19-2749-2019, 2019.

Thank you for this comment. We had considered including pre-satellite data for the original manuscript, but we were worried that the atmospheric state would be much less constrained and rely more on the model itself, hence we decided against using additional SSWs with a poorer representation of the atmospheric circulation.

5) Other studies. The introduction ignores important information on past efforts (and their degree of success) in identifying stratospheric factors that may influence downward coupling, e.g.:

Charlton, A.J. and L.M. Polvani, 2007: A New Look at Stratospheric Sudden Warmings. Part I: Climatology and Modeling Benchmarks. J. Climate, 20, 449? 469,https://doi.org/10.1175/JCLI3996.1
This reference has been included in the Introduction.

Karpechko, A.Y., Hitchcock, P., Peters, D.H.W. and Schneidereit, A. (2017), Predictability of downward propagation of major sudden stratospheric warmings. Q.J.R. Meteorol. Soc., 143: 1459-1470. doi:10.1002/qj.3017
This reference had already been included in our original manuscript. The discussion on this reference has been extended in the Introduction.

Mitchell, D.M., L.J. Gray, J. Anstey, M.P. Baldwin, and A.J. Charlton-Perez, 2013: The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate. J. Climate, 26, 2668?2682, https://doi.org/10.1175/JCLI-D-12-00030.1

This reference has been included in the Introduction.

Maycock, A. C., and Hitchcock, P. (2015), Do split and displacement sudden stratospheric warmings have different annular mode signatures?, Geophys. Res. Lett., 42, 10,943-10,951, doi:10.1002/2015GL066754.
This reference has been included in the Introduction.

Nakagawa, K. I., and Yamazaki, K. (2006), What kind of stratospheric sudden warming propagates to the troposphere? Geophys. Res. Lett., 33, L04801, doi:10.1029/2005GL024784.
This reference has been included in the Introduction.

Runde, T., Dameris, M., Garny, H., and Kinnison, D. E. (2016), Classification of stratospheric extreme events according to their downward propagation to the troposphere, Geophys. Res. Lett., 43, 6665? 6672, doi:10.1002/2016GL069569.
This reference has been included in the Introduction.

Seviour, W. J. M., Gray, L. J., and Mitchell, D. M. (2016), Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models, J.Geophys.Res. Atmos., 121, 1400-1413, doi:10.1002/2015JD024178.
This reference has been included in the Introduction.

All of the above references, in addition to further references to clarify the answers to the reviewer questions as well as papers published in the meantime, have been included in the new version of the manuscript.

Specific comments

L6: following the weeks after an SSW → in the weeks following an SSW
This has been corrected (line 8).

L45-49 you need to add e.g., to these reference lists as they are highly selective L54 remove 'also'
This section has been re-written.

L55 occurrence of an SSW also plays
This sentence has been removed.

L65-67 "Prior work based on this extended regime definition revealed important differences in the surface weather response to the state of the stratosphere, which remain hidden using the canonical four NAE regimes (Papritz and Grams, 2018; Beerli and Grams, 2019)." It seems this needs expanding as this is important justification for the current approach of using seven regimes rather

than four. What specifically is missed? Also what did Papritz and Grams, 2018 and Beerli and Grams, 2019 show in relation to the two questions investigated here? Did they analyse similar things? What did they find?
Papritz and Grams (2018) investigated how weather regimes modulate the storm track and by that the occurrence of marine cold air outbreaks over the north-eastern North Atlantic. They further showed that the winter mean strength of the polar vortex is related to frequency changes of the weather regimes. They did, however, not consider instantaneous anomalies in the strength of the polar vortex or SSWs. We removed the reference when rewriting the introduction.

L113-115 These statements are not visible from Figure 1 without some specific information on frequencies given in the text or in a table. Also in Figure 2a I see a peak in AR between -20 to -10 days but I cannot see a clear higher frequency for ScTr compared to all the other states. Rather the peak in EuBL immediately before onset seems to be a clearer feature for all events.
We have reformulated and clarified this paragraph (1st paragraph of section 3) so that Figures 1 and 2 are looked at together by the reader. We have in addition included a table (Table 1) in order to clarify the occurrence of weather regimes at the onset of the SSW events.

L138-139: All the subsets you are dealing with are small sample sizes. The no regime cases can provide a useful null hypothesis, i.e. what anomalous regime frequencies can apparently arise without a specific regime being identified at day 0?
We do not agree that the no regime category would provide a suitable null hypothesis for the hypotheses addressed in this study. First, the no regime category arises due to a weak projection of the geopotential height anomalies in the Atlantic European sector on one particular cluster centroid. This does not necessarily imply that the geopotential height signal is particularly weak. Instead, this may just reflect a progression between regimes. Second, we aim to test whether the regime succession is different in the presence of an SSW compared to the case where no SSW took place. Thus, the null hypothesis is that the regime progression in the aftermath of an SSW is the same as in cases with the same initial regime but no SSW. We have clarified this in section 2.2.

Fig. 2 and A2: The choice to use 5-day running means and to test for significance on that basis has implications related to the intrinsic persistence characteristics of each regime. But these timescales differ – e.g., from Fig. 1 it appears the canonical persistence timescale of GL is longer than, say, EuBL. It needs to be mentioned how the authors have accounted for the intrinsic persistence of each regime in choosing the smoothing window. Also, do you need to account for autocorrelation in your statistical tests?
We have now described the regime persistence for the different regimes in more detail. The 5 day running mean arises from the minimum duration of a regime. We now state this also in the main text (lines 144-150). See also answer to your point 2 above.

I find Figure A2 more informative than Figure 2 since what you wish to highlight is the anomalous frequencies associated with particular subsets of data not the absolute frequencies. This is more clearly seen in Figure A2. For example, it becomes clear that the significant anomalies in AT at lags -35 to -15 is because the frequency is anomalously low (i.e. a negative anomaly). I suggest switching them and putting Fig. A2 in the main text and Fig. 2 in the Appendix.

Thank you for this comment, we have switched Figures 2 and A2 and adapted the text accordingly.

L163-164 "Given the strong influence of the tropospheric state at the time of the onset of an SSW on the weather regime frequencies in the subsequent days" I don't agree you have demonstrated this in Section 3. See major comment 1.

We have removed this statement.

L170-171 These are weaker thresholds than one would typically associate with "robust" and "highly robust"

Yes, we agree and we have removed the terms robust and highly robust. We have also clarified this in the text and we now explicitly state the level, i.e., 25 % or 10 % (lines 203-209).

L 179-187 and Fig. A3: Why are the tropospheric Z anomalies so weak? Is it a matter of plotting (e.g., contour intervals)? The SSW compendium NAM composite for the same set of events in ERA-Interim (see Fig. 1 below) looks quite different from your Z anomalies (Butler et al., 2017).

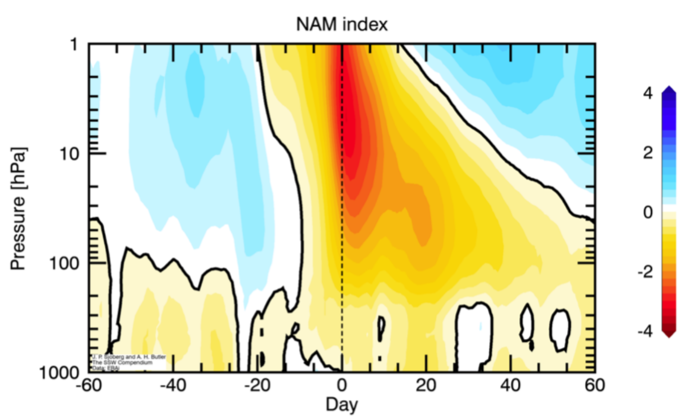

Fig. 1. SSW compendium composite NAM anomaly for SSWs in ERA-Interim (Butler et al., 2017).

Comparing to the NAM figure from the SSW compendium (copied again here), intervals of 0.25 stddev are used to show the downward influence, and shading starts at zero. In our figure we use intervals of 0.5 stddev, and shading starts at 0.5. Hence, in the NAM compendium figure, if using the same colorbar that we use here, only the downward influence around a lag of 20 days would be shaded, as is the case in our figure, with a few additional places close to the surface, e.g. around lags of -35 to -30 and +40 days. These figures are therefore very comparable.

Fig 5: Is any statistical testing applied to the anomalies? The caption does not mention it.
To keep the rather small panels as easy to read as possible, we do not show statistical significance here (this is now Figure 4).

Typographical

Figures – the two shades of green for EuBL and ScBL are hard to differentiate
We improved Figures 1, 2 and A2 in order to better differentiate the green shades.

The role of North Atlantic-European weather regimes in the surface impact of sudden stratospheric warming events

Daniela I.V. Domeisen[1,*], Christian M. Grams[2,*], and Lukas Papritz[1,*]

[1]Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland
[2]Institute of Meteorology and Climate Research (IMK-TRO), Department Troposphere Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[*]Equally contributing authors

Correspondence: Daniela I.V. Domeisen (daniela.domeisen@env.ethz.ch)

Abstract. Sudden stratospheric warming (SSW) events can significantly impact tropospheric weather for a period of several weeks, in particular in the North Atlantic European **(NAE) region**. **While the stratospheric forcing often projects onto the North Atlantic Oscillation (NAO)**, **the** tropospheric response **to SSW events**, if any, **is highly variable** and it remains an open question what determines the existence, location, timing, and strength of the downward impact. We here explore **how the variable tropospheric response to SSW events in the NAE region can be characterised in terms of a refined set of seven weather regimes and if the tropospheric flow in the North Atlantic region at the onset of SSW events is an indicator of the subsequent downward impact. The weather regime analysis reveals** the Greenland blocking (GL) and Atlantic Trough (AT) regimes as the most frequent large-scale patterns **in the weeks following an SSW.** While the GL regime is dominated by high pressure over Greenland, AT is dominated by a southeastward shifted storm track in the North Atlantic. **The flow evolution associated with GL and the associated cold conditions over Europe in the weeks following an SSW occur most frequently if a blocking situation over western Europe and the North Sea (European Blocking) prevailed at the time of the SSW onset.** In contrast, **an AT regime associated with** mild conditions is more likely if GL occurs already at SSW onset. **For** the remaining tropospheric flow regimes during SSW onset **we cannot identify a dominant flow evolution**. **Although it remains unclear what causes these relationships,** the results **suggest that specific tropospheric states during the onset of the SSW are an indicator of the subsequent tropospheric flow evolution in the aftermath of an SSW**, which could provide crucial guidance for subseasonal prediction.

1 Introduction

Sudden stratospheric warming events can have a significant impact on the tropospheric large-scale circulation and hence on surface weather (Baldwin and Dunkerton, 2001). However, a robust detection and quantification of the downward impact of SSWs remains challenging. First of all, the number of SSWs in the record of satellite-era reanalysis is small (26 events from

1979 - 2019), while the case-to-case variability **in terms of their tropospheric impact** is large. Second, the internal variability of the troposphere itself is high, such that it can mask a stratospheric influence. Predicting if, when, and where a downward impact from SSW events will occur is therefore not straightforward, yet a better prediction of the type and timing of a downward impact would significantly benefit a wide range of users.

The **tropospheric impact of SSW events** is communicated by a range of mechanisms including synoptic and planetary-scale waves (e.g. Song and Robinson, 2004; Domeisen et al., 2013; Hitchcock and Simpson, 2014; Smith and Scott, 2016). In particular, after SSW events the North Atlantic-European region (NAE) tends to exhibit more persistent states of the negative phase of the North Atlantic Oscillation (NAO-, Domeisen, 2019), as well as more frequent transitions towards NAO- and fewer away from NAO- (Charlton-Perez et al., 2018). NAO- is associated with enhanced meridional air mass exchanges, in particular, more cold air outbreaks in Northern Europe but fewer over the Nordic Seas (Kolstad et al., 2010; Kretschmer et al., 2018b; Papritz and Grams, 2018; Huang and Tian, 2019), as well as **increased** precipitation in Southern Europe (Butler et al., 2017; Ayarzagüena et al., 2018). The Pacific sector tends to be less strongly affected in the aftermath of SSW events (Greatbatch et al., 2012; Butler et al., 2017), though the occurrence of wave reflection in the stratosphere can be associated with Pacific blocking (Kodera et al., 2016) and cold spells over North America (Kretschmer et al., 2018a). Given the preferred occurrence and the increased persistence of certain surface signatures after SSW events as compared to climatology, medium- to long-range predictability over Europe has been suggested to increase after SSW events (Sigmond et al., 2013; Domeisen et al., 2015; Karpechko, 2015; Butler et al., 2016; Scaife et al., 2016; Jia et al., 2017; Beerli et al., 2017; Butler et al., 2019; Domeisen et al., 2020a), **although SSW events themselves are often not predictable beyond deterministic lead times (Taguchi, 2014, 2016; Domeisen et al., 2020b).**

Despite the preferred occurrence of the negative phase of the NAO, the downward influence of an SSW event on the evolution of the tropospheric flow can be highly variable between events. This issue is further complicated by the fact that there exists a range of different metrics for characterizing the downward impact, with each definition yielding a different set and number of SSW events with apparent surface impacts. **In particular, the occurrence and type of downward impact has been investigated with respect to the SSW geometry (Charlton and Polvani, 2007; Mitchell et al., 2013; Maycock and Hitchcock, 2015; Seviour et al., 2016), though no statistically robust differences with respect to wave geometry emerge in the tropospheric response. More promising pathways include SSW precursors (Nakagawa and Yamazaki, 2006; White et al., 2019), the evolution of the stratosphere - troposphere system following the SSW (Kodera et al., 2016), and in particular the persistence of the lower stratospheric response (Hitchcock et al., 2013a; Karpechko et al., 2017; Runde et al., 2016; Polichtchouk et al., 2018) after the SSW event. These studies use indices for the downward effect that are based on exclusively stratospheric or a combination of stratospheric and tropospheric indicators. In this study we will investigate purely tropospheric indicators of downward impact for SSW events.** Definitions of a downward impact using tropospheric indicators **are generally based** on large-scale circulation indices such as the NAO (Charlton-Perez et al., 2018; Domeisen, 2019) or tropospheric jet location (Garfinkel et al., 2013; Afargan-Gerstman and Domeisen, 2020; Maycock et al., 2020).

While a causal downward link from the stratosphere after SSW events has been confirmed **in idealized experiments** (e.g. Gerber et al., 2009), remote forcings can affect both the stratosphere and the troposphere, and thereby either mask or strengthen the downward response from the stratosphere. Indeed, a range of tropical remote connections can impact the NAE region through both a tropospheric and a stratospheric pathway (Attard et al., 2019), such as the Quasi-Biennial Oscillation (QBO) (Gray et al., 2018; Andrews et al., 2019), the MJO (Garfinkel et al., 2014; Barnes et al., 2019), and the El Niño Southern Oscillation (ENSO) (Jiménez-Esteve and Domeisen, 2018; Domeisen et al., 2019), in addition to extratropical tropospheric forcing in the North Pacific (Honda and Nakamura, 2001; Sun and Tan, 2013; Drouard et al., 2013), Arctic sea ice (Sun et al., 2015), and snow cover in Eurasia (Cohen et al., 2014). It therefore has to be kept in mind that the stratosphere is often only one possible forcing of the troposphere. In addition, it has recently been suggested that the precursors to SSW events with a downward influence differ from those without such a tropospheric impact in terms of strength and location (Domeisen, 2019; Zhang et al., 2019), in particular with respect to forcing over Eurasia (White et al., 2019; Tyrrell et al., 2019; Peings, 2019).

Given the large variability of the tropospheric flow evolution following SSW events and the influence of other remote factors mentioned above, the prediction of the SSW response in the troposphere is difficult for an individual event, despite the general shift towards NAO negative conditions in a statistical sense. The goal of this study is **to investigate if tropospheric flow regimes in the NAE region help to understand the variability of the SSW response in the observational record. More specifically, we here address the question if the tropospheric flow evolution in the NAE region after an SSW is statistically different from that without an SSW using seven weather regimes in the NAE region. Weather regimes are quasi-stationary, recurrent, and persistent patterns of the large-scale extratropical circulation (e.g. Michelangeli et al., 1995). While many studies showed that there are preferred transitions between different regimes, internal tropospheric variability is high and a regime onset often occurs on short timescales (e.g. Vautard, 1990; Michel and Rivière, 2011). Therefore predictability due to regimes arises from regime persistence on time scales of several days rather than typical regime sequences over several weeks. However, recent work revealed important shifts of regime occurrence and transition probabilities between regimes on subseasonal time scales of several weeks dependent on the external forcing such as the stratospheric polar vortex state (Charlton-Perez et al., 2018; Papritz and Grams, 2018; Beerli and Grams, 2019). This motivates the study at hand aiming at investigating if the variability in the tropospheric flow evolution following SSW events can be characterised in terms of the weather regime at the time of SSW onset.**

2 Data and Methods

2.1 Data and Classifications

ERA-interim reanalysis (Dee et al., 2011) from 1979 to present is used for all figures. The SSW central dates are identified as the day of reversal of the zonal mean zonal winds to easterly at 10 hPa and 60°N in midwinter (Nov - Mar) for ERA-Interim (1979 - 2019), yielding 26 SSW events for the period 1979-2019. For the period 1979 to 2013, these reversal dates are given in Table 2 in Butler et al. (2017). The SSW central dates for the remaining years are 12-Feb-2018 and 02-Jan-2019.

The tropospheric flow over the NAE region is described in terms of quasi-stationary large-scale flow patterns, given by seven year-round weather regimes defined in Grams et al. (2017) based on six-hourly data for the period 1979-2015. We use this weather regime classification to stratify SSWs according to the large-scale tropospheric flow conditions at their onset. To do so, we select for each SSW the dominant weather regime that is active during at least one 6-hourly time step throughout the **onset** day of the SSW **(see Table 1)**. As for the canonical seasonal definition using four regimes (e.g. Michelangeli et al., 1995; Michel and Rivière, 2011; Ferranti et al., 2015; Charlton-Perez et al., 2018), the mean patterns of the seven regimes are based on a k-means clustering in the phase space spanned by the leading seven empirical orthogonal functions (EOFs; explaining 76% of the variance) of 10-day low-pass filtered 500hPa geopotential height anomalies. In addition, **the normalized projection (weather regime index I_{WR})** following Michel and Rivière (2011) **for each of the seven regimes** is employed for defining objective weather regime life cycles and for a filtering of time steps without a clear regime structure ("no regime" category). **In essence, an active life cycle requires an I_{WR} above a certain threshold for at least 5 consecutive days (minimum duration for an active regime life cycle) and a continuous increase/decrease during the onset/decay phases (see methods of Grams et al., 2017, for details). As different life cycles can be active simultaneously, in particular during the onset and decay phases, individual days are attributed to a specific regime life cycle only if I_{WR} is also the maximum of all I_{WR}.** The life cycle definition allows for a continuous extension of the weather regime attribution to more recent data without repeating the EOF analysis and clustering (here done for the years 2016-2019).

Three of the seven regimes are dominated by a cyclonic 500 hPa geopotential height anomaly ("cyclonic regimes"; cf. Figs. A1a-c): the Atlantic Trough (AT) regime with cyclonic activity shifted towards western Europe, the Zonal regime (ZO), and the Scandinavian Trough (ScTr) regime. The remaining four regimes are dominated by a positive geopotential height anomaly and are referred to as "blocked regimes" (Figs. A1d-g): Atlantic Ridge (AR), European Blocking (EuBL), Scandinavian Blocking (ScBL), and Greenland Blocking (GL).

A potential modulation of the frequency of occurrence of the seven regimes can be understood in terms of the link between the respective regimes and the NAO (Beerli and Grams, 2019, their Figs. 2,6) and the link between the stratospheric polar vortex and the NAO (Charlton-Perez et al., 2018). Since ZO and ScTr project onto NAO+, they are suppressed after a weakening of the stratospheric polar vortex, and vice versa after a strengthening. In contrast, GL strongly projects onto NAO- and is enhanced following a weak stratospheric polar vortex, while it is suppressed in the aftermath of a strong vortex. EuBL and AT do not project strongly onto either NAO phase and are, thus, only weakly modulated by the strength of the stratospheric polar vortex.

2.2 Statistical testing

The rare occurrence of SSWs (26 events during 1979-2019) and the subsequent stratification according to tropospheric flow conditions requires careful statistical testing to extract significant and robust results that are distinct from sampling uncertainty. **The overarching questions we address in this study are whether after SSWs the tropospheric flow evolution is different from situations without an SSW and to what extent this depends on the tropospheric state at the time of the SSW. To investigate the latter, we will consider subsamples of all SSWs. Hence, in all cases the relevant null hypothesis is that**

Table 1. Weather Regime at the onset of SSW events in the observational record (1979-2019)

dominant weather regime (lag 0)		SSW central date
Atlantic Trough	(AT)	14.03.1988, 24.01.2009, 24.03.2010
Zonal Regime	(ZO)	21.02.1989, 15.12.1998, 22.02.2008, 12.02.2018
Scandinavian Trough	(ScTr)	26.02.1999, 20.03.2000
Atlantic Ridge	(AR)	04.12.1981, 23.01.1987, 02.01.2019
European Blocking	(EuBL)	22.02.1979, 24.02.1984, 01.01.1985, 11.02.2001, 21.01.2006, 07.01.2013
Scandinavian Blocking	(ScBL)	-
Greenland Blocking	(GL)	04.03.1981, 08.12.1987, 30.12.2001, 24.02.2007, 09.02.2010
no regime at onset	(no)	29.02.1980, 18.01.2003, 05.01.2004

the flow evolution after SSWs is indistinguishable from that occurring in the absence of an SSW. The testing procedure, thus, comprises the following two steps:

1. **First, we assess the** *robustness* **of the samples by performing a Monte Carlo resampling. For that purpose, we resample the original samples 100 times with repetitions. The number of random samples is chosen according to the maximum number of possible combinations with repetitions of the smallest subset of SSW events that will be considered in this study ($N = 5$ events corresponding to 126 independent combinations). This yields confidence intervals, estimating the uncertainty inherent in each sample. Due to the small sample size, these confidence intervals are relatively large.**

2. **Second, we compute 1000 random samples of the same size as the original sample but for random periods with the same weather regime at the central date but no SSW occurring within ± 60 days, yielding estimates of the distributions in the absence of SSWs. Prescribing the same weather regime at the central date for the random samples filters out signals which might result from regime persistence or preferred regime transitions independent of external forcings. Testing for** *significance* **is done by comparing the confidence intervals and distributions obtained from the random samples for overlap.**

Applying this method to geopotential height anomalies, we consider anomalies as robust if the width of the confidence interval is smaller than the amplitude of the anomaly. In addition, the sample mean is significant at, e.g., the 10% level, if the confidence intervals overlap by less than 10% with the Monte Carlo distribution. A similar procedure is applied to test significance of lagged weather regime occurrence.

3 Weather regimes during SSW events

As a first step, we evaluate the sequence of weather regimes from 60 days before to 60 days after an SSW for all 26 SSW cases during 1979-2019 (Fig. 1**, cf. Tab. 1**). **This figure suggests a preferred occurrence of AT (purple) and GL (blue)**

during the weeks after an SSW compared to the weeks before. This is further emphasized by the 5-day running mean

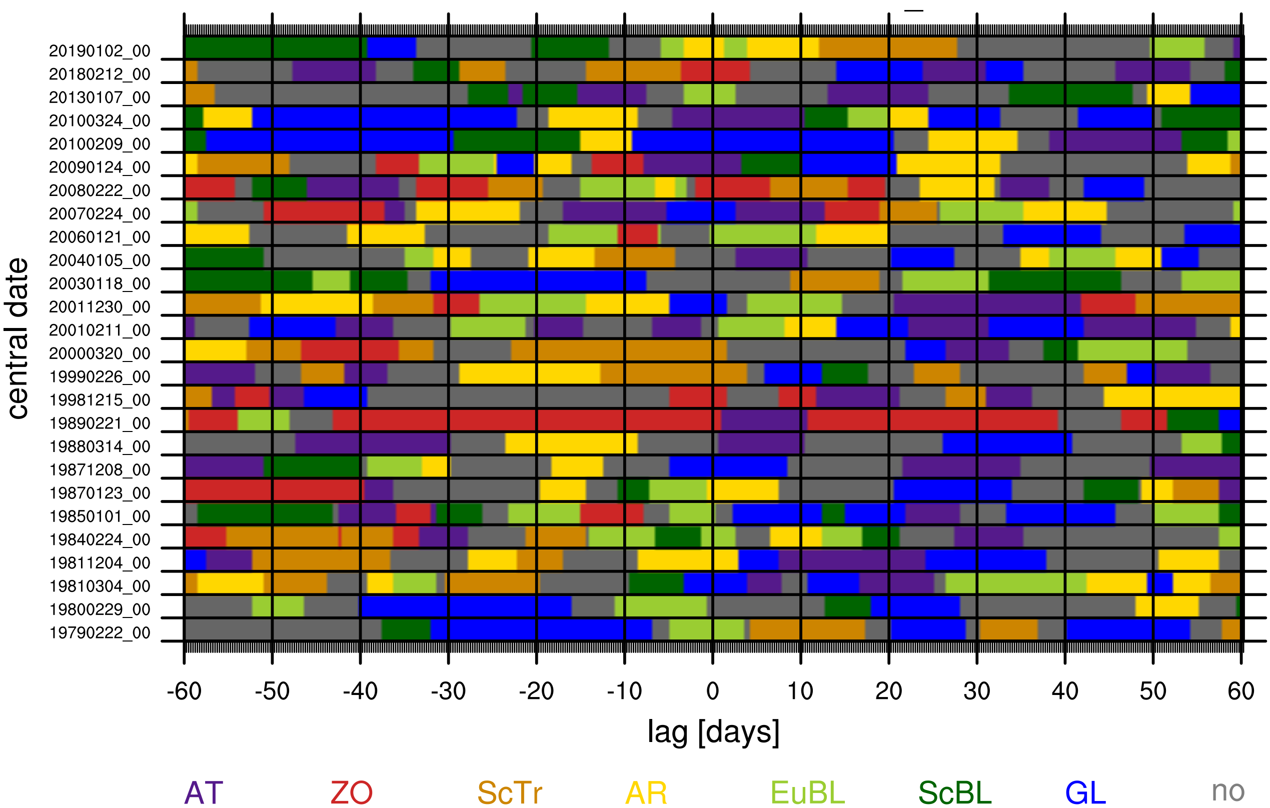

Figure 1. The sequence of the dominant weather regimes (colors indicated in legend) for -60 to +60 days with respect to the onset of each of the 26 SSW events (lag 0) between 1979 and 2019. The central dates of the SSW events are indicated on the left.

of the anomalous frequency of weather regimes around SSW events, which provides a more complete overview over the modulation of regime frequencies after SSWs (Figure 2). **We show the 5-day running mean frequency anomaly to account for the 5-day minimum duration of an active regime life cycle. Different from the testing procedure outlined in Section 2.2, we here consider the distribution of lagged 5-day mean frequencies by selecting for each day in the original sample a random day $\pm$15 days around the original day of year but from a different winter.** In addition, the random day must exhibit the same weather regime as the original day to replicate potential regime-dependence. We then compute the mean lagged weather regime frequency for each random sample as for the original sample and test for significance at the 10% level (bold). **For reference, we show the absolute frequencies of weather regimes in Fig. A2.**

GL and EuBL are the most prominent regimes at the onset of SSW events with a relative 5-day mean frequency of more than 18% and 20%, respectively (Figure A2a). The frequency of EuBL is significantly enhanced from 5 days prior until the onset of the SSW (Figure 2a), in agreement with Woollings et al. (2010) and Nishii et al. (2011). The cyclonic regimes ZO, ScTr, as well as the blocked regimes AR and ScBL tend to be suppressed at the time of SSW events. **This is consistent with the**

strong projection of the ZO and ScTr regimes onto NAO+, which also tends to be suppressed after SSW events (Charlton-Perez et al., 2018). **On the other hand, AR (yellow, significant peak around lag -20 to -10) and the related ScTr (orange, significant range at lag -10) regimes are more frequent in the period 1-3 weeks before the SSW onset. The prominence of AR around 15 days before the onset of an SSW event agrees** with the suggested precursor role of blocking over the Atlantic before SSW events (Martius et al., 2009). After the SSW onset, AT frequencies are significantly enhanced, peaking at more than 23% after 7 days **(Figure A2a) corresponding to a frequency anomaly of 15% for the same lag (Figure 2a).** Thereafter, GL (lag 12 to 42 days) and AT (lag 17-35 days) are the most likely weather regimes **with enhanced frequency anomalies of up to 15%** (Fig. A2a), but **in absolute terms** frequencies for both are only around 20-25% and none of the two clearly dominates (Figure 2a). This behaviour obscures the potential tropospheric impact of an SSW in a composite as AT and GL trigger contrasting large-scale weather conditions (rather mild and windy for AT, cold and calm for GL) for large parts of Europe (Beerli and Grams, 2019).

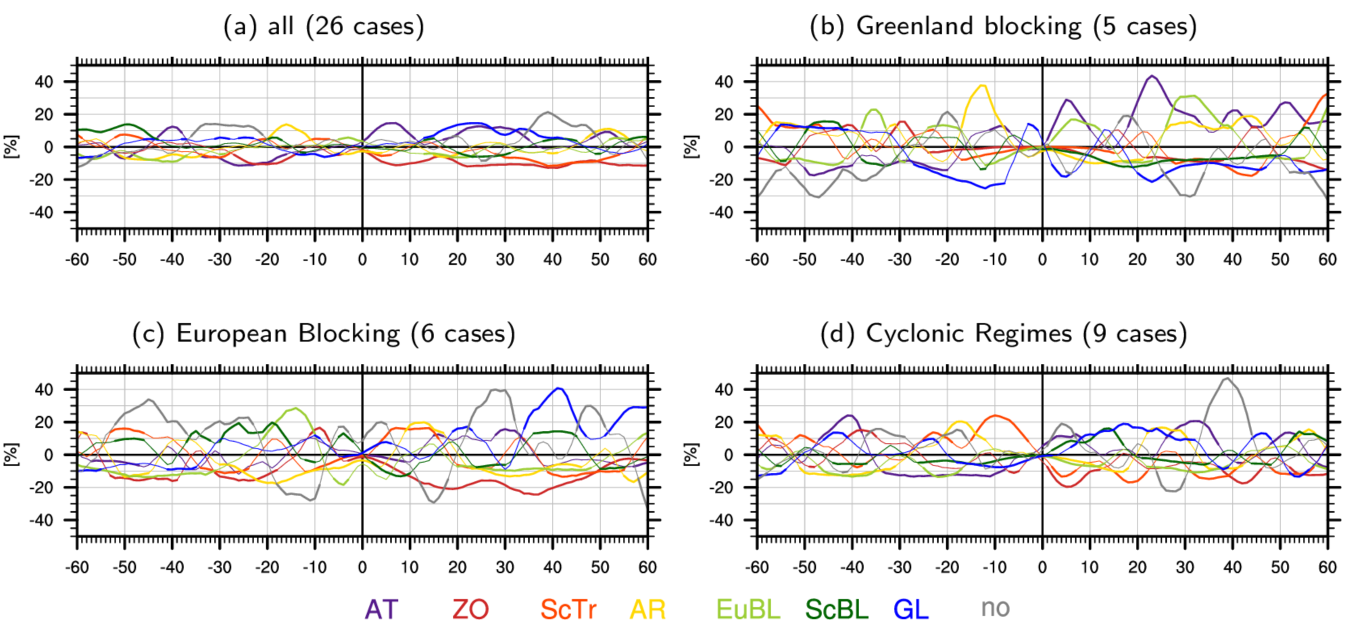

Figure 2. 5-day running mean of the *anomalous* frequency of weather regimes centred on the onset of the SSW event (lag 0) relative to the mean of the climatological distribution for (a) all SSW events and (b-d) conditional on the dominant weather regime at lag 0: (b) Greenland blocking, (c) European blocking, and (d) cyclonic regimes (ZO, AT, and ScTr). The 5-day mean frequencies are computed from 6-hourly weather regime data from lag -60 days to lag 60 days. Note the different y-axis in (a). Note that anomalous frequencies at lag 0 in c, d are - by construction - close to zero as the same regime is prescribed for computing the mean from the 1000 Monte-Carlo samples. The bold parts of the lines indicate significant deviations from climatology (see text for details).

We now sub-divide the 26 SSW events with respect to the weather regime that dominates during the SSW onset: GL (5 cases), EuBL (6 cases), and the cyclonic regimes (ZO, ScTr, AT; 9 cases). The remaining 6 cases either have no clear regime signature (no-regime, 3 events) or are associated with AR (3 events) at their onset. Because of the small sample size, we do not consider these cases here. For the GL subset (Fig. 2b / A2b), all other regimes are subsequently suppressed except for AT and EuBL. The frequency of GL itself drops immediately after the SSW to below 10% around a lag of 20 days – far below its climatological mean frequency. AT, and to a lesser degree also EuBL, become significantly more frequent immediately after the SSW until about a lag of 10 days, reaching absolute frequencies of 35% and 20%, respectively **(Fig. A2b)**. After a period with no clear regime assignment, AT becomes the dominant regime starting at lag 18 days **with anomalous frequencies above 40% (Fig. 2b)**, peaking above 50% absolute frequency about 23 days after the SSW and remaining significantly enhanced until a lag of 33 days (Fig. A2b). From lag 25 days until lag 40 days, EuBL becomes significantly enhanced peaking at 40% absolute frequency around lag 30 days.

For the EuBL subset (Fig. 2c / A2c), the subsequent regime frequencies are quite different to GL at the onset of an SSW. First, the frequencies of ScTr and AR are significantly enhanced directly after the SSW, with peaks at 20% and 30% absolute frequency at lag 10 days. This is then followed by a period of preferred occurrence of GL (around lag 20 days) and AT (around lag 30 days) with a frequency of about 25% each. The dominance of GL from lag 35 to 45 days (50% peak frequency) is particularly striking. At the same time, also ScBL is enhanced with a frequency of 20%, while all other regimes are suppressed.

Cyclonic regimes at the time of the SSW (Fig. 2d / A2d) exhibit a less prominent regime frequency modulation after an SSW compared to the EuBL and GL subsets. Still, GL (lag 10-35 days, lag 45 days), AR (lag 20-30 days), and AT (lag 25-35 days) are significantly enhanced, but absolute frequencies barely exceed 20%. **Note that this corresponds to significantly increased frequencies of 10-20% for these regimes in the considered time windows. However,** most often no single regime dominates after an SSW event with a cyclonic regime at lag 0, hinting at cases with a "missing" response after the SSW event.

Despite the large tropospheric variability in the aftermath of SSW events, the investigation of lagged regime frequencies reveals that (1) the AT and GL regimes are more likely to follow an SSW (as compared to other weather regimes) and (2) that this subsequent modulation is sensitive to the tropospheric flow regime at the onset of the SSW. The dominance of EuBL and GL at the time of the SSW onset **hints at a significantly more likely GL response (after EuBL at lag 0) vs. AT (after GL at lag 0) after an SSW, respectively. Thus the stratospheric impact on the evolution of the tropospheric flow in the NAE region and hence the associated surface weather may be connected to the presence of a particular tropospheric regime at the onset of the SSW.**

4 Temporal Evolution of the Downward Impact

We focus in the following on the modulation of the stratosphere-troposphere coupling for the previously discussed sets of SSWs. For that purpose, we evaluate the temporal evolution of standardized geopotential height anomalies averaged over the NAE sector (-80°E to 40°E / 60°N to 90°N) by compositing a given set of SSW events. Using the full hemisphere, that is, the full longitude range instead of the here used sectorial view over the North Atlantic, yields the same qualitative results due to the strong imprint of the anomalies induced by the SSW in the NAE sector (Fig. A3).

Compositing all SSW events (Fig. 3a) yields the classical dripping paint plot of Baldwin and Dunkerton (2001, their Fig. 2). Qualitative differences to the figure from Baldwin and Dunkerton (2001) are due to the different variable (geopotential height in our study vs NAM) **and** the number of events (26 in our study vs 18) for a different time period (1979-2019 in our study vs 1958-1999). When compositing all SSW events, the downward impact between 10 to 60 days after the SSW onset is robust **at the 25 % but not the 10 % level** (see Fig. A4a). Together with the relatively weak amplitude of the anomalies, this reflects the large case-to-case variability in the tropospheric impact of SSWs. **A more robust anomaly (10 % level)** can only be observed at around 15 days after an SSW. **This anomaly around a lag of 15 days is unlikely to be obtained from a random sampling as evident from the less than 10 % overlap between the confidence and random distributions (Fig. 3a). This suggests that in the immediate aftermath of an SSW (lag 15 - 25 days), indeed positive geopotential height anomalies over the NAE sector are significantly more likely than in the absence of an SSW.**

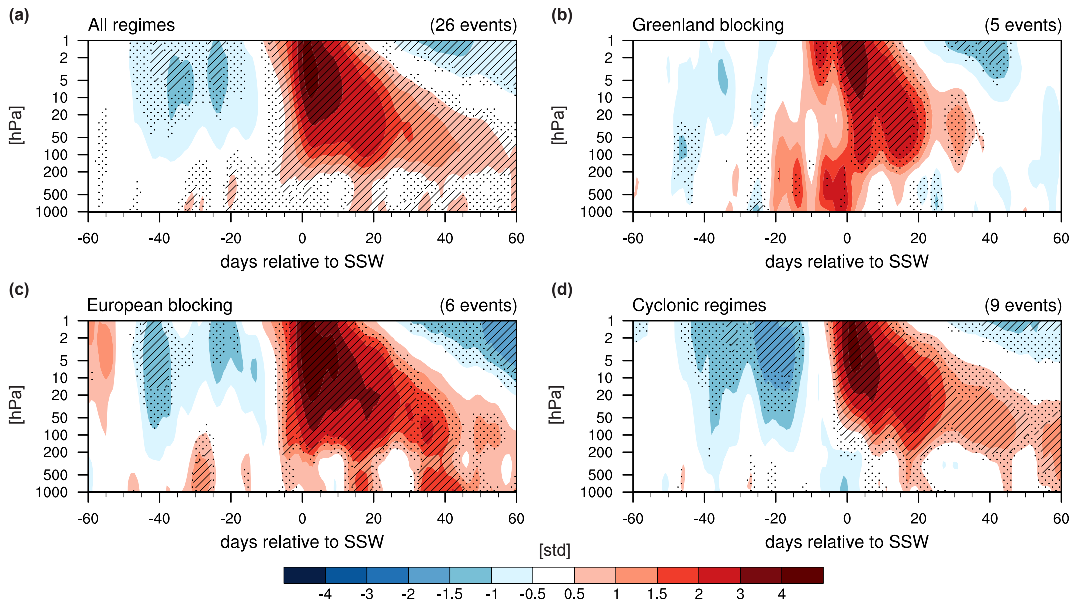

Figure 3. Standardized geopotential height anomalies for the sector -80°E to 40°E / 60°N to 90°N for (a) all SSW events, and (b - d) subdivided by the weather regime that is dominant at the onset of the SSW event as indicated in the panel titles. Hatching (stippling) indicates that the confidence intervals and the random distributions overlap by less than 25% (10%).

SSW events that occur during GL (Fig. 3b) are associated with an immediate, strongly positive anomaly in the troposphere. Consistent with Fig. A2b, when GL is present at the onset of the SSW, GL or AR are often already present before the SSW event, which is likely the cause of the positive tropospheric geopotential height anomalies several days prior to the event. Notably, there are no significant and robust (cf. Figs. 3b and A4b) anomalies after 10 days of the onset of the SSW except for a weak negative geopotential height anomaly after 20 days, indicating a cyclonic flow regime in the NAE region. This is consistent with the significantly enhanced likelihood for the occurrence of the AT regime at this lag (Fig. A2b). **Note that both the immediate positive geopotential height anomalies and the weak tropospheric anomalies in the aftermath of the event are not the result of cancellations in the composites but are rather typical across cases.**

For EuBL at the onset of the SSW event, only a weakly significant positive anomaly can be observed at the time of the SSW, but highly robust, significant, and strongly positive geopotential height anomalies are present in the troposphere at lags of 15 - 20 and 30 - 55 days after the SSW event (Fig. 3c). These positive anomalies are consistent with the finding that first AR and then GL are much more likely in the aftermath of an SSW with EuBL at lag 0 (compare to Fig. A2c). Furthermore, comparing to the panel for all SSW events (Fig. 3a) indicates that the EuBL cases dominate the perceived downward response in the canonical response for SSW events.

During cyclonic regimes at the onset of the SSW, there is no substantial tropospheric anomaly in the NAE region at the time of the SSW, but a positive albeit weak anomaly can be observed around days 15 - 20 after the SSW event (Fig. 3d). **This anomaly is not robust at the 10 % level, but it is significantly different from a random sample at least at the 25 % level.** Several SSWs with a cyclonic regime at the onset are followed by GL at a longer lag (Fig. A2d), thus, likely causing these anomalies. Still, the GL frequencies only reach 25% at most and also other regimes occur more often albeit with low frequencies around 25%. These findings and the small amplitude of the anomalies suggest that the variability in the tropospheric flow evolution after SSWs is large after a cyclonic regime at lag 0, which is also confirmed by the inspection of individual cases (not shown).

While not the focus of this study, the question arises whether other factors might contribute to the differing surface evolution in the aftermath of the SSW event. In fact, it is certainly the case that tropospheric variability at SSW onset alone is not responsible for the response to SSWs. In particular, differing values of the amplitude and persistence of the lower stratospheric anomaly can be observed in Fig. 3 between the different composites. In general, the events with EuBL at the onset also tend to have a longer stratospheric persistence, although a long (but slightly weaker) persistence can also be observed for cyclonic regimes at the onset of the SSW. The five SSW events associated with GL have a more short-lived lower stratospheric response. As events with a persistent lower stratospheric response are often associated with so-called polar jet oscillation (PJO) events, a comparison with Table 1 in Karpechko et al. (2017) reveals that 2 out of 5 SSW events with a GL regime (and, respectively, 4 out of 6 EuBl events) at the onset have a persistent lower stratospheric temperature response characterized by a polar jet oscillation (PJO) event (Kuroda and Kodera, 2004; Hitchcock et al., 2013b). While this is not a clear result, it indicates that the shorter (longer) persistence in the lower stratosphere for the SSWs associated with GL (EuBL) may add some support to the persistence of the tropospheric response, but the statistics are too small to provide a clear result. Similarly, 4 out of 6 EuBL events are split events (rather than displacements), while 2 out of 5 GL events are split events.

5 Impact on Surface Weather

Since each weather regime is associated with characteristic surface weather, the modulation of regime successions in the aftermath of an SSW by the tropospheric state at the time of **an SSW might contribute** to the marked variability in the surface impact. Hence, we here consider spatial composites of 2m temperature anomalies and anomalies of 500 hPa geopotential height

(Z500') for the three groups of SSW events discussed in the previous sections **(Fig. 4a-c) and for all SSW events (Fig. 4d) for days 0 to 25 after the SSW (cf. Fig. A5 for days 25 - 50)** .

During SSWs dominated by GL at the onset, initially strong warm anomalies prevail over Greenland and the Canadian Archipelago, whereas western Russia and Scandinavia are anomalously cold, consistent with the anomalous ridge over Greenland and the low geopotential height anomalies over Scandinavia (Fig. 4a). **With the subsequent progression of weather regimes - typically towards the cyclonic AT regime or EuBL - mild conditions are established throughout central Europe from a lag of 20 days onwards. This is in stark contrast to the negative NAO and the associated cold conditions that are commonly expected as the canonical response to SSWs over Europe (Butler et al., 2017; Kolstad et al., 2010; Domeisen et al., 2020a).**

For SSWs that are dominated by EuBL at the onset, cold anomalies prevail over Northern Europe, albeit also extending over large parts of central Europe (Fig. 4b). **They peak at -4 K to -6 K around lags beyond 20 days, which corresponds well with the occurrence of the GL regime. Note that cold anomalies in the composite for all SSWs are much weaker (cf. Fig. 4d) reaching -1 K to -3 K in Central Europe. The retrogression of initial positive Z500' over the eastern North Atlantic to Greenland along with a strengthening of negative Z500' over the southeastern North Atlantic around lag 15-25 days is striking. Furthermore, GL is associated with warm anomalies over Greenland and North America.**

Finally, as expected by the varied regime succession for the SSWs with cyclonic regimes at their onset, composite temperature and Z500' are weaker for these events (Fig. 4c). **Thus, the canonical response of surface temperature (i.e., the composite for all SSWs, Figure 4d) is the result of averaging over – in important regions opposing – temperature anomalies for SSWs with GL, EuBL, or a cyclonic regime at the onset.**

6 Summary and Discussion

This study **aimed to shed light on the large case-to-case variability of the tropospheric response to SSW events and their associated surface impacts, as well as** the dependence on the tropospheric weather regime at the onset of the SSW. To that end, we have exploited in a statistical framework the observational record of the satellite era (1979 - 2019) as represented in the ERA-Interim reanalysis. Our conclusions are as follows:

1. In the aftermath of an SSW event, the tropospheric flow in the NAE region exhibits an evolution that is unlikely to occur in the absence of an SSW. Specifically, positive geopotential height anomalies related to Greenland blocking are statistically more likely to occur after the onset of the SSW than in the absence of an SSW. This is consistent with the expected (canonical) **negative NAO** response of the troposphere to SSWs (e.g. Charlton-Perez et al., 2018).

2. The significant and robust positive geopotential height anomalies found in the period 10-60 days after SSWs are predominantly the result of SSWs with European blocking at their onset. This is manifest **for this subset of events** in a transition from EuBL to GL that then dominates at lags of 15-20 and 30-55 days after the SSW onset, which is statisti-

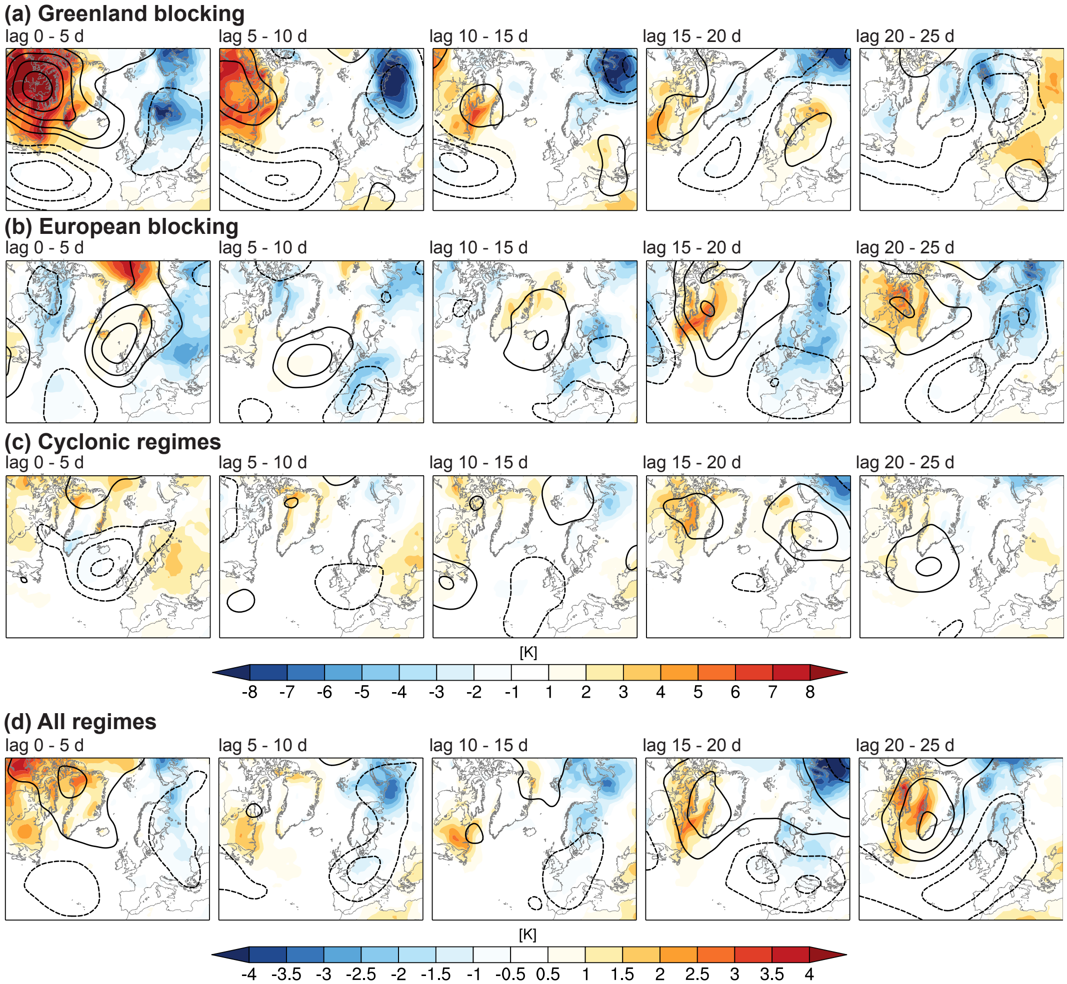

Figure 4. Surface impact for SSWs with (a) Greenland blocking, (b) European blocking, and (c) cyclonic regimes at the onset, as well as (d) for all SSWs. Shading indicates composite 2m temperature anomalies, black contours correspond to geopotential height anomalies at 500 hPa at intervals of (a-c) 50 gpm and (d) 25 gpm. Negative values are dashed. Fields are averaged over 5 days between lags 0 to 25 days. Note the different scales for temperature in (a-c) and (d).

cally significantly different from the natural progression from EuBL to GL. For other tropospheric regimes at SSW onset the tropospheric response is weaker and **less robust and significant.**

3. For **Greenland blocking at the SSW onset**, a weak preference for cyclonic flow regimes around 20-30 days after the SSW is apparent, with an **opposite** surface response in the aftermath of the SSW as compared to SSW onsets dominated by EuBL.

4. SSWs that occur during cyclonic weather regimes exhibit a considerably weaker and less significant response with a modestly enhanced likelihood for GL.

Depending on the tropospheric weather regime at SSW onset different surface signatures result. Specifically, the canonical signature in 2m temperature, e.g., cold conditions prevailing over much of northern Europe, occurs for the EuBL cases. In contrast, mild temperatures in large parts of Europe are found for SSWs with GL at their onset. It is important to distinguish these cases, since although EuBL and GL frequently occur at the onset of SSW events, they lead to a different subsequent **evolution and different associated surface temperatures. In particular, the most common transitions are from EuBL to GL and from GL to AT around 3-4 weeks after the SSW, respectively,** along with their contrasting large-scale weather impacts (Beerli and Grams, 2019). These findings corroborate that the presence of either a EuBL or GL regime at SSW onset will allow us to disentangle the difference in **surface weather**, and hence to determine if and when a "downward impact" of the SSW is expected. This is highly relevant for subseasonal forecasting.

While these findings are limited by the small sample size of available SSW events, the rigorous statistical testing for significance and robustness performed here suggests that the large case-to-case variability in the tropospheric response to SSWs **can be described in terms of NAE weather regimes and depends on the regime** at the onset of the SSW. Our findings confirm that while the stratosphere does not represent the sole forcing of the tropospheric state, for many events it may be able to nudge the tropospheric flow into a particular direction by suppressing some weather regimes and by **favoring others, as found in Charlton-Perez et al. (2018). We here in addition** show that the susceptibility of the troposphere to the stratospheric nudging depends on the tropospheric state at the time of the SSW. **Other factors that can modulate the tropospheric response are the persistence of the temperature anomaly in the lower stratosphere (Hitchcock et al., 2013a; Karpechko et al., 2017; Runde et al., 2016; Polichtchouk et al., 2018), as well as upstream effects in the eastern North Pacific (Afargan-Gerstman and Domeisen, 2020). An analysis of differences in the lower stratospheric persistence for the here considered weather regimes did not yield conclusive results, which warrants further studies. In particular, a model analysis to quantify the respective contributions to the tropospheric impact of different remote factors in comparison to the role of local North Atlantic variability might shed further light onto the complex role of stratosphere - troposphere coupling in surface weather.**

However, it is currently not sufficiently known to what extent complex prediction models are able to represent the diversity of tropospheric responses to stratospheric forcing, as this has not been sufficiently tested in models beyond the canonical response and selected case studies. From a preliminary analysis of S2S prediction model data we anticipate large biases and a very complex role of the representation of stratosphere - troposphere coupling in prediction models that will be difficult to disentangle. Hence, while state-of-the-art subseasonal prediction systems are often unable to forecast **at the time of occurrence of the SSW event** if a surface response is to be expected, our findings suggests that the presence or absence – and in fact the timing – of a surface impact following SSW events might **in some cases be** predictable based on the weather regime at the onset of the SSW event. This could significantly improve the subseasonal prediction of tropospheric winter weather over Europe.

Data availability. The ERA-interim reanalysis data (Dee et al., 2011) is available from ECMWF at https://apps.ecmwf.int/datasets/data/interim-full-daily/.

Appendix A

Author contributions. The authors together initiated and designed the study and all authors contributed to data analysis, discussion of results, and writing.

Competing interests. The authors declare no competing interests.

Acknowledgements. Support from the Swiss National Science Foundation through project PP00P2_170523 to DD is gratefully acknowledged. The contribution of CMG was supported by the Helmholtz Association as part of the Young Investigator Group "SPREADOUT" (grant VH-NG-1243). Data analysis and visualisation were performed using the NCAR Command Language (UCAR/NCAR/CISL/VETS, 2014).

References

Afargan-Gerstman, H. and Domeisen, D. I. V.: Pacific Modulation of the North Atlantic Storm Track Response to Sudden Stratospheric Warming Events, Geophysical Research Letters, 47, 18, https://doi.org/https://doi.org/10.1029/2019GL085007, 2020.

Andrews, M. B., Knight, J. R., Scaife, A. A., and Lu, Y.: Observed and simulated teleconnections between the stratospheric Quasi-Biennial Oscillation and Northern Hemisphere winter atmospheric circulation, Journal of Geophysical Research – Atmospheres, https://doi.org/10.1029/2018JD029368, 2019.

Attard, H. E., Lang, A. L., Attard, H. E., and Lang, A. L.: The Impact of Tropospheric and Stratospheric Tropical Variability on the Location, Frequency, and Duration of Cool-Season Extratropical Synoptic Events, Monthly Weather Review, 147, 519–542, 2019.

Ayarzagüena, B., Barriopedro, D., Perez, J. M. G., Abalos, M., de la Camara, A., Herrera, R. G., Calvo, N., and Ordóñez, C.: Stratospheric Connection to the Abrupt End of the 2016/2017 Iberian Drought, Geophysical Research Letters, 45, 12,639–12,646, 2018.

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

Barnes, E. A., Samarasinghe, S. M., Uphoff, I. E., and Furtado, J. C.: Tropospheric and Stratospheric Causal Pathways Between the MJO and NAO, Journal of Geophysical Research-Atmospheres, 124, 9356–9371, 2019.

Beerli, R. and Grams, C. M.: 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, 0, https://doi.org/10.1002/qj.3653, 2019.

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

Butler, A., Charlton-Perez, A., Domeisen, D. I. V., Garfinkel, C., Gerber, E. P., Hitchcock, P., Karpechko, A. Y., Maycock, A. C., Sigmond, M., Simpson, I., and Son, S.-W.: Sub-seasonal Predictability and the Stratosphere, in: Sub-Seasonal to Seasonal Prediction, pp. 223–241, Elsevier, https://doi.org/10.1016/B978-0-12-811714-9.00011-5, 2019.

Butler, A. H., Arribas, A., Athanassiadou, M., Baehr, J., Calvo, N., Charlton-Perez, A., Déqué, M., Domeisen, D. I. V., Fröhlich, K., Hendon, H., Imada, Y., Ishii, M., Iza, M., Karpechko, A. Y., Kumar, A., MacLachlan, C., Merryfield, W. J., Müller, W. A., O'Neill, A., Scaife, A. A., Scinocca, J., Sigmond, M., Stockdale, T. N., and Yasuda, T.: The Climate-system Historical Forecast Project: do stratosphere-resolving models make better seasonal climate predictions in boreal winter?, Quarterly Journal of the Royal Meteorological Society, 142, 1413–1427, 2016.

Butler, A. H., Sjoberg, J. P., Seidel, D. J., and Rosenlof, K. H.: A sudden stratospheric warming compendium, Earth System Science Data, 9, 63–76, 2017.

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

Charlton-Perez, A. J., Ferranti, L., and Lee, R. W.: The influence of the stratospheric state on North Atlantic weather regimes, Quarterly Journal of the Royal Meteorological Society, 144, 1140–1151, 2018.

Cohen, J., Furtado, J. C., Jones, J., Barlow, M., Whittleston, D., and Entekhabi, D.: Linking Siberian Snow Cover to Precursors of Stratospheric Variability, Journal of Climate, 27, 5422–5432, 2014.

Dee, D., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A., Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hersbach, H., Holm, E., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A., Monge-Sanz, B., Morcrette, J., Park, B., Peubey, C., Rosnay,

P. d., Tavolato, C., Thepaut, J., and Vitart, F.: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Q. J. Royal Met. Soc., 137, 553–597, 2011.

Domeisen, D. I. V.: Estimating the Frequency of Sudden Stratospheric Warming Events from Surface Observations of the North Atlantic Oscillation, Journal of Geophysical Research-Atmospheres, 124, https://doi.org/10.1029/2018JD030077, 2019.

Domeisen, D. I. V., Sun, L., and Chen, G.: The role of synoptic eddies in the tropospheric response to stratospheric variability, Geophys. Res. Lett., 40, 4933–4937, https://doi.org/10.1002/grl.50943, 2013.

Domeisen, D. I. V., Butler, A. H., Fröhlich, K., Bittner, M., Müller, W. A., and Baehr, J.: Seasonal Predictability over Europe Arising from El Niño and Stratospheric Variability in the MPI-ESM Seasonal Prediction System, Journal of Climate, 28, 256–271, https://doi.org/10.1175/JCLI-D-14-00207.1, 2015.

Domeisen, D. I. V., Garfinkel, C. I., and Butler, A. H.: The Teleconnection of El Niño Southern Oscillation to the Stratosphere, Reviews of Geophysics, 57, https://doi.org/10.1029/2018RG000596, 2019.

Domeisen, D. I. V., Butler, A. H., Charlton-Perez, A. J., Ayarzaguena, B., Baldwin, M. P., Dunn Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A. Y., Kim, H., Knight, J., Lang, A. L., Lim, E.-P., Marshall, A., Roff, G., Schwartz, C., Simpson, I. R., Son, S.-W., and Taguchi, M.: The role of the stratosphere in subseasonal to seasonal prediction. Part 2: Predictability arising from stratosphere - troposphere coupling, Journal of Geophysical Research-Atmospheres, https://doi.org/10.1029/2019JD030923, 2020a.

Domeisen, D. I. V., Butler, A. H., Charlton-Perez, A. J., Ayarzaguena, B., Baldwin, M. P., Dunn Sigouin, E., Furtado, J. C., Garfinkel, C. I., Hitchcock, P., Karpechko, A. Y., Kim, H., Knight, J., Lang, A. L., Lim, E.-P., Marshall, A., Roff, G., Schwartz, C., Simpson, I. R., Son, S.-W., and Taguchi, M.: The role of the stratosphere in subseasonal to seasonal prediction. 1. Predictability of the stratosphere, Journal of Geophysical Research-Atmospheres, https://doi.org/10.1029/2019JD030920, 2020b.

Drouard, M., Rivière, G., Arbogast, P., Drouard, M., Rivière, G., and Arbogast, P.: The North Atlantic Oscillation Response to Large-Scale Atmospheric Anomalies in the Northeastern Pacific, dx.doi.org, 70, 2854–2874, 2013.

Ferranti, L., Corti, S., and Janousek, M.: Flow-Dependent Verification of the ECMWF Ensemble over the Euro-Atlantic Sector, Quarterly Journal of the Royal Meteorological Society, 141, 916–924, https://doi.org/10.1002/qj.2411, 2015.

Garfinkel, C. I., Waugh, D. W., and Gerber, E. P.: The Effect of Tropospheric Jet Latitude on Coupling between the Stratospheric Polar Vortex and the Troposphere, Journal of Climate, 26, 2077–2095, 2013.

Garfinkel, C. I., Benedict, J. J., and Maloney, E. D.: Impact of the MJO on the boreal winter extratropical circulation, Geophysical Research Letters, 41, 6055–6062, 2014.

Gerber, E. P., Orbe, C., and Polvani, L. M.: Stratospheric influence on the tropospheric circulation revealed by idealized ensemble forecasts, Geophysical Research Letters, 36, L24 801, 2009.

Grams, C. M., Beerli, R., Pfenninger, S., Staffell, I., and Wernli, H.: Balancing Europe's wind-power output through spatial deployment informed by weather regimes, Nature Climate Change, 7, nclimate3338–562, 2017.

Gray, L. J., Anstey, J. A., Kawatani, Y., Lu, H., Osprey, S., and Schenzinger, V.: Surface impacts of the Quasi Biennial Oscillation, Atmospheric Chemistry And Physics, 18, 8227–8247, 2018.

Greatbatch, R. J., Gollan, G., Jung, T., and Kunz, T.: Factors influencing Northern Hemisphere winter mean atmospheric circulation anomalies during the period 1960/61 to 2001/02, Quarterly Journal of the Royal Meteorological Society, 138, 1970–1982, 2012.

Hitchcock, P. and Simpson, I. R.: The Downward Influence of Stratospheric Sudden Warmings, J. Atmos. Sci., 71, 3856–3876, 2014.

Hitchcock, P., Shepherd, T. G., and Manney, G. L.: Statistical Characterization of Arctic Polar-Night Jet Oscillation Events, Journal of Climate, 26, 2096–2116, 2013a.

Hitchcock, P., Shepherd, T. G., Taguchi, M., Yoden, S., and Noguchi, S.: Lower-stratospheric Radiative Damping and Polar-night Jet Oscillation Events, J. Atmos. Sci., 70, 1391–1408, 2013b.

Honda, M. and Nakamura, H.: Interannual Seesaw between the Aleutian and Icelandic Lows. Part II: Its Significance in the Interannual Variability over the Wintertime Northern Hemisphere, Journal of Climate, 14, 4512–4529, 2001.

Huang, J. and Tian, W.: Eurasian Cold Air Outbreaks under Different Arctic Stratospheric Polar Vortex Strength, Journal of the Atmospheric Sciences, https://doi.org/10.1175/JAS-D-18-0285.1, 2019.

Jia, L., Yang, X., Vecchi, G., Gudgel, R., Delworth, T., Fueglistaler, S., Lin, P., Scaife, A. A., Underwood, S., and Lin, S.-J.: Seasonal Prediction Skill of Northern Extratropical Surface Temperature Driven by the Stratosphere, Journal of Climate, 30, 4463–4475, 2017.

Jiménez-Esteve, B. and Domeisen, D. I. V.: The Tropospheric Pathway of the ENSO-North Atlantic Teleconnection, Journal of Climate, 31, 4563–4584, https://doi.org/10.1175/JCLI-D-17-0716.1, 2018.

Karpechko, A. Y.: Improvements in statistical forecasts of monthly and two-monthly surface air temperatures using a stratospheric predictor, Quarterly Journal of the Royal Meteorological Society, 141, 2444–2456, 2015.

Karpechko, A. Y., Hitchcock, P., Peters, D. H. W., and Schneidereit, A.: Predictability of downward propagation of major sudden stratospheric warmings, Q. J. Royal Met. Soc., 104, 30 937, 2017.

Kodera, K., Mukougawa, H., Maury, P., Ueda, M., and Claud, C.: Absorbing and reflecting sudden stratospheric warming events and their relationship with tropospheric circulation, Journal of Geophysical Research-Atmospheres, 121, 80–94, 2016.

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

Kretschmer, M., Cohen, J., Matthias, V., Runge, J., and Coumou, D.: The Different Stratospheric Influence on Cold-Extremes in Eurasia and North America, npj Climate and Atmospheric Science, 1, 44, https://doi.org/10.1038/s41612-018-0054-4, 2018a.

Kretschmer, M., Coumou, D., Agel, L., Barlow, M., Tziperman, E., and Cohen, J.: More-Persistent Weak Stratospheric Polar Vortex States Linked to Cold Extremes, Bulletin of the American Meteorological Society, 99, 49–60, 2018b.

Kuroda, Y. and Kodera, K.: Role of the Polar-night Jet Oscillation on the formation of the Arctic Oscillation in the Northern Hemisphere winter, Journal of Geophysical Research, 109, D11 112, 2004.

Martius, O., Polvani, L., and Davies, H.: Blocking precursors to stratospheric sudden warming events, Geophys. Res. Lett., 36, L14 806, 2009.

Maycock, A. C. and Hitchcock, P.: Do split and displacement sudden stratospheric warmings have different annular mode signatures?, Geophysical Research Letters, 42, 10 943–10 951, 2015.

Maycock, A. C., Masukwedza, G. I. T., Hitchcock, P., and Simpson, I. R.: A Regime Perspective on the North Atlantic Eddy-Driven Jet Response to Sudden Stratospheric Warmings, Journal of Climate, 33, 3901–3917, 2020.

Michel, C. and Rivière, G.: The Link between Rossby Wave Breakings and Weather Regime Transitions, Journal of the Atmospheric Sciences, 68, 1730–1748, https://doi.org/10.1175/2011JAS3635.1, 2011.

Michelangeli, P.-A., Vautard, R., and Legras, B.: Weather Regimes: Recurrence and Quasi Stationarity, Journal of the Atmospheric Sciences, 52, 1237–1256, https://doi.org/10.1175/1520-0469(1995)052<1237:WRRAQS>2.0.CO;2, 1995.

Mitchell, D. M., Gray, L. J., Anstey, J., Baldwin, M. P., and Charlton-Perez, A. J.: The Influence of Stratospheric Vortex Displacements and Splits on Surface Climate, Journal of Climate, 26, 2668–2682, 2013.

Nakagawa, K. I. and Yamazaki, K.: What kind of stratospheric sudden warming propagates to the troposphere?, Geophysical Research Letters, 33, 30 937–4, 2006.

Nishii, K., Nakamura, H., and Orsolini, Y. J.: Geographical Dependence Observed in Blocking High Influence on the Stratospheric Variability through Enhancement and Suppression of Upward Planetary-Wave Propagation, J. Climate, 24, 6408–6423, 2011.

Papritz, L. and Grams, C.: Linking low-frequency large-scale circulation patterns to cold air outbreak formation in the north-eastern North Atlantic, Geophys. Res. Lett., 45, 2542–2553, https://doi.org/10.1002/2017GL076921, 2018.

Peings, Y.: Ural Blocking as a driver of early winter stratospheric warmings, Geophysical Research Letters, pp. 2019GL082 097–18, 2019.

Polichtchouk, I., Shepherd, T. G., and Byrne, N. J.: Impact of Parametrized Nonorographic Gravity Wave Drag on Stratosphere-Troposphere Coupling in the Northern and Southern Hemispheres, Geophysical Research Letters, 45, 8612–8618, 2018.

Runde, T., Dameris, M., Garny, H., and Kinnison, D. E.: Classification of stratospheric extreme events according to their downward propagation to the troposphere, Geophysical Research Letters, 43, 6665–6672, 2016.

Scaife, A. A., Karpechko, A. Y., Baldwin, M. P., Brookshaw, A., Butler, A. H., Eade, R., Gordon, M., MacLachlan, C., Martin, N., Dunstone, N., and Smith, D.: Seasonal winter forecasts and the stratosphere, Atmospheric Science Letters, 17, 51–56, 2016.

Seviour, W. J. M., Gray, L. J., and Mitchell, D. M.: Stratospheric polar vortex splits and displacements in the high-top CMIP5 climate models, Journal of Geophysical Research-Atmospheres, pp. 1–14, 2016.

Sigmond, M., Scinocca, J. F., Kharin, V. V., and Shepherd, T. G.: Enhanced seasonal forecast skill following stratospheric sudden warmings, Nature Geoscience, 6, 1–5, 2013.

Smith, K. L. and Scott, R. K.: The role of planetary waves in the tropospheric jet response to stratospheric cooling, Geophysical Research Letters, 43, 2904–2911, 2016.

Song, Y. and Robinson, W. A.: Dynamical Mechanisms for Stratospheric Influences on the Troposphere, Journal of the Atmospheric Sciences, 61, 1711–1725, 2004.

Sun, J. and Tan, B.: Mechanism of the wintertime Aleutian Low–Icelandic Low seesaw, Geophysical Research Letters, 40, 4103–4108, 2013.

Sun, L., Deser, C., and Tomas, R. A.: Mechanisms of Stratospheric and Tropospheric Circulation Response to Projected Arctic Sea Ice Loss*, Journal of Climate, 28, 7824–7845, 2015.

Taguchi, M.: 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, J. Atmos. Sci., 71, 2886–2904, 2014.

Taguchi, M.: Connection of predictability of major stratospheric sudden warmings to polar vortex geometry, Atmospheric Science Letters, 17, 33–38, 2016.

Tyrrell, N. L., Karpechko, A. Y., Uotila, P., and Vihma, T.: Atmospheric Circulation Response to Anomalous Siberian Forcing in October 2016 and its Long-Range Predictability, Geophysical Research Letters, 104, 30,937–11, 2019.

UCAR/NCAR/CISL/VETS: The NCAR Command Language (Version 6.1.2) [Software], Boulder, Colorado, https://doi.org/http://dx.doi.org/10.5065/D6WD3XH5, http://www.ncl.ucar.edu/, 2014.

Vautard, R.: Multiple Weather Regimes over the North Atlantic: Analysis of Precursors and Successors, Monthly Weather Review, 118, 2056–2081, https://doi.org/10.1175/1520-0493(1990)118<2056:MWROTN>2.0.CO;2, 1990.

White, I., Garfinkel, C. I., Gerber, E. P., Jucker, M., Aquila, V., and Oman, L. D.: The Downward Influence of Sudden Stratospheric Warmings: Association with Tropospheric Precursors, Journal of Climate, 32, 85–108, 2019.

Woollings, T., Charlton-Perez, A., Ineson, S., Marshall, A. G., and Masato, G.: Associations between stratospheric variability and tropospheric blocking, J. Geophys. Res.: Atmospheres, 115, D06 108, 2010.

Zhang, R., Tian, W., Zhang, J., Huang, J., Xie, F., and Xu, M.: The Corresponding Tropospheric Environments during Downward-extending and Non-downward-extending Events of Stratospheric Northern Annular Mode Anomalies, Journal of Climate, pp. JCLI–D–18–0574.1, 2019.

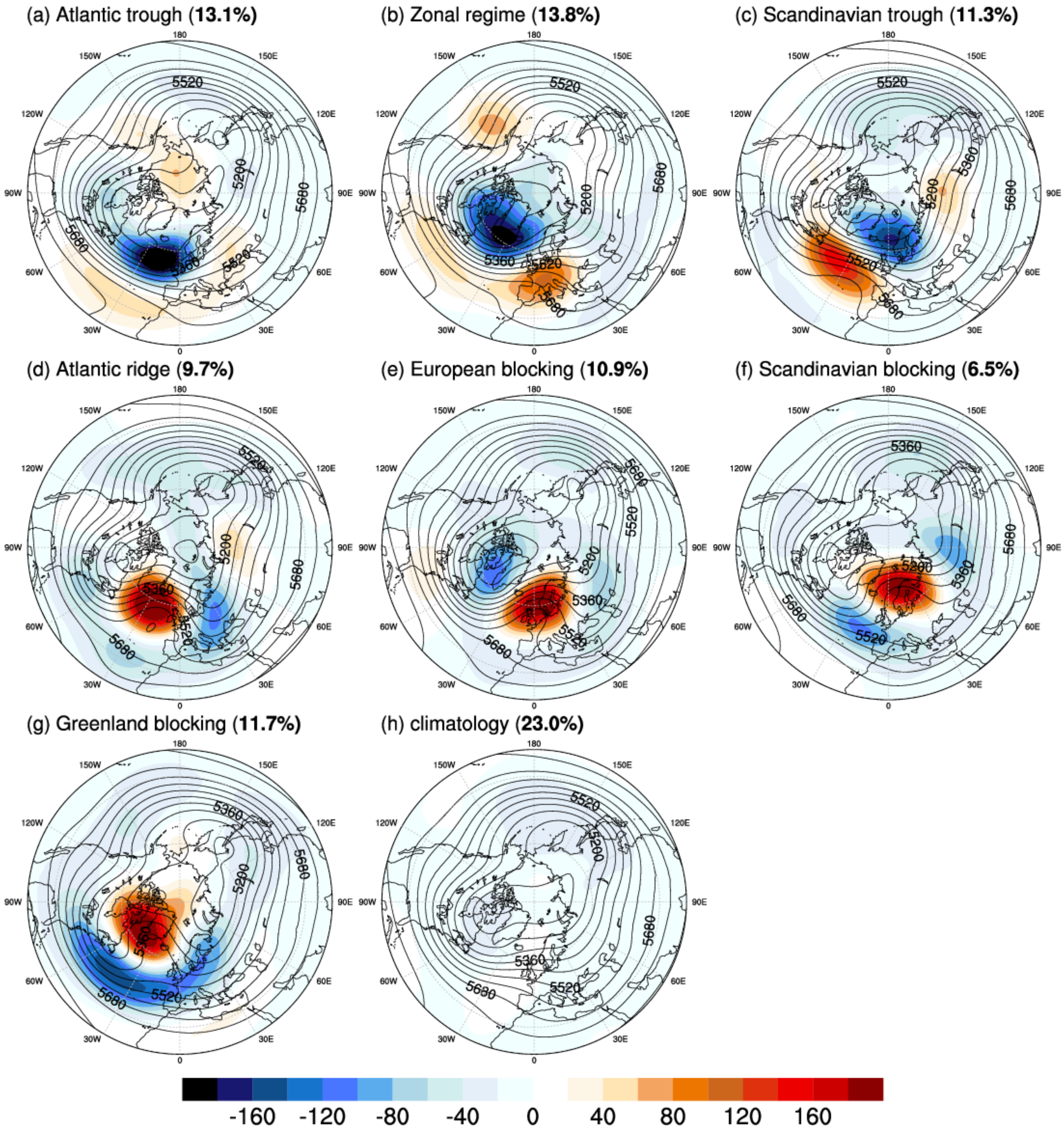

Figure A1. Composite mean 10-day low-pass filtered 500 hPa geopotential height anomaly (shading, every 20 gpm), and mean absolute 500 hPa geopotential height (black contours, every 20 gpm) for all winter days in ERA-Interim (DJF, 1979-2015) attributed to one of the 7 weather regimes (a-g) and the climatological mean (h). Regime name and relative frequency (in percent) are indicated in the sub-figure captions.

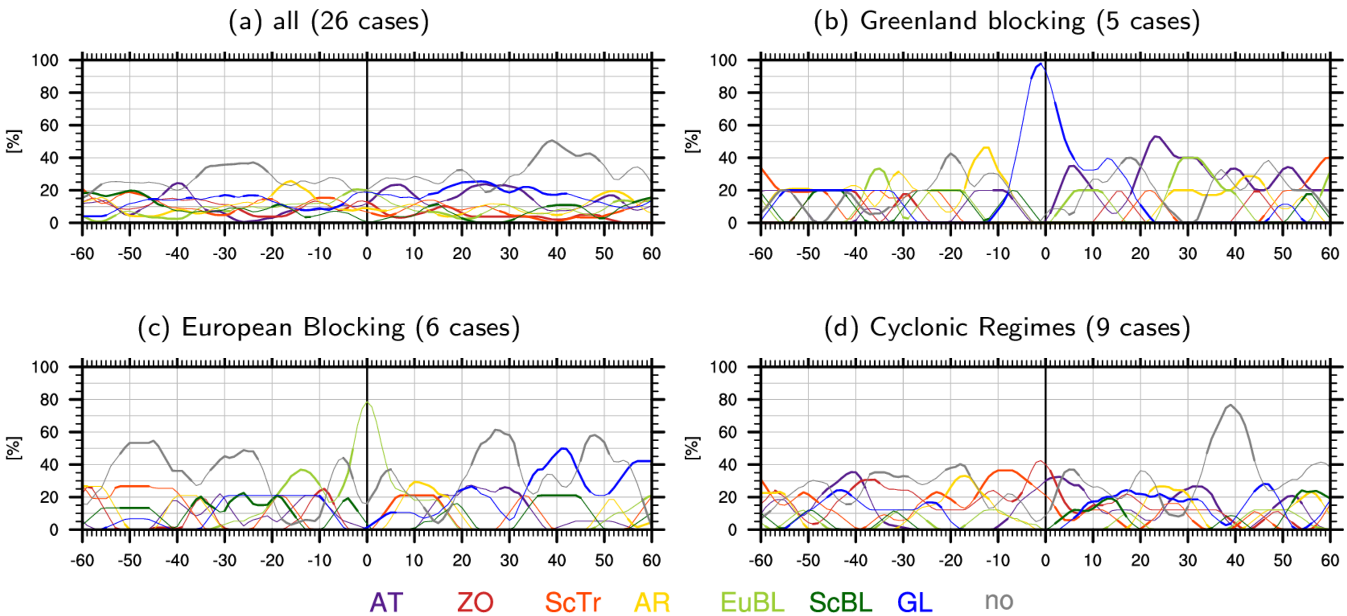

Figure A2. **As Figure 2 but for the 5-day running mean** ***absolute*** **frequency of weather regimes centred on the onset of the SSW event (lag 0) for (a) all SSW events and (b-d) conditional on the dominant weather regime at lag 0: (b) Greenland blocking, (c) European blocking, and (d) cyclonic regimes (ZO, AT, and ScTr). The 5-day mean frequencies are computed from 6-hourly weather regime data from lag -60 days to lag 60 days. Note the different y-axis in (a). The bold parts of the lines indicate significant deviations from climatology (see text for details).**

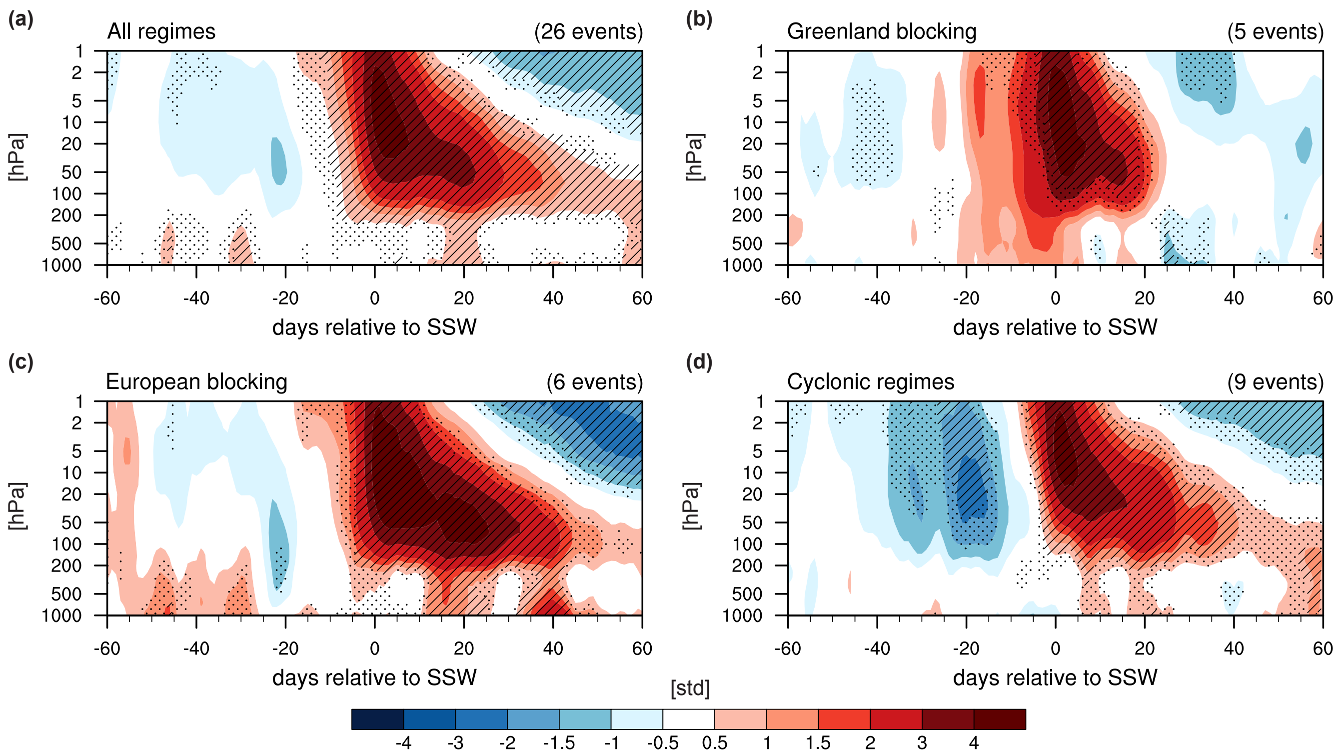

Figure A3. As Figure 3 but for the full longitude range, i.e. for the polar cap poleward of 60°N.

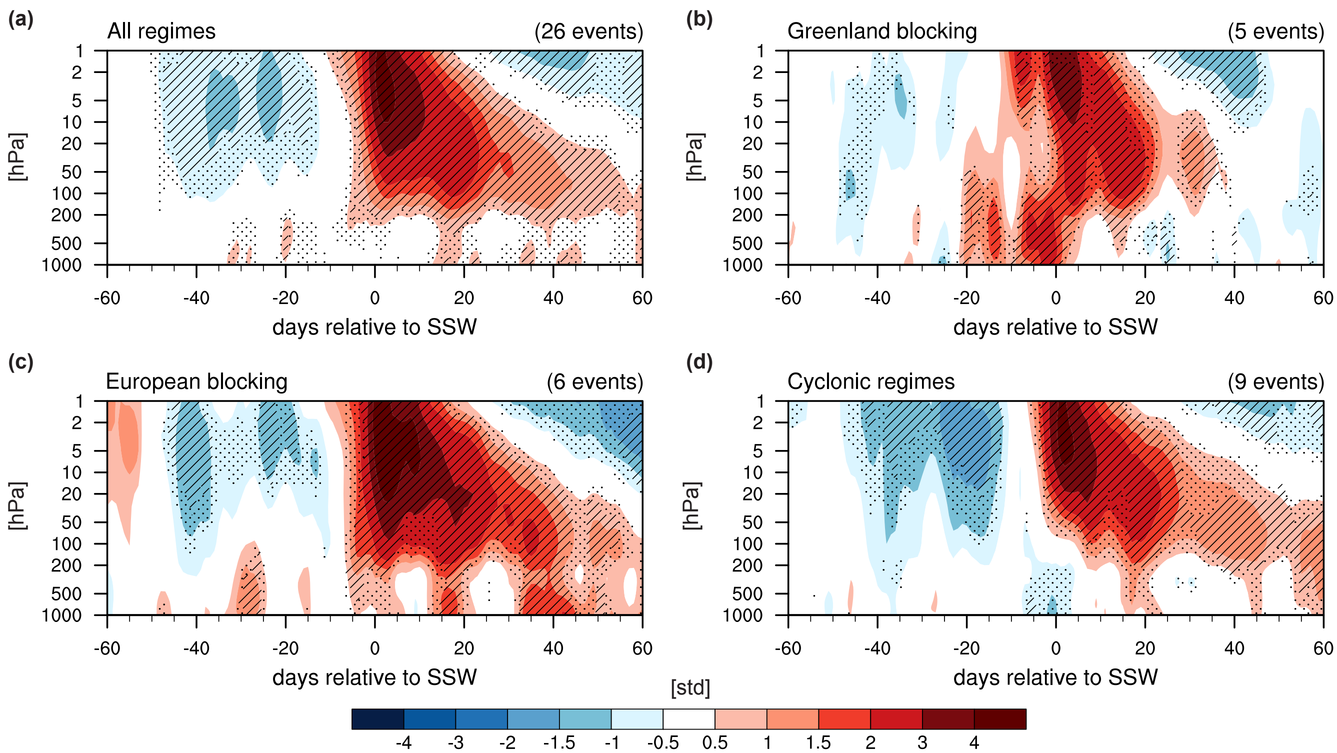

Figure A4. Standardized geopotential height anomalies for the sector -80°E to 40°E / 60°N to 90°N for (a) all SSW events, and (b-d) subdivided by the weather regime that is dominant at the onset of the SSW as indicated by the titles of the panels. Robustness is assessed using confidence intervals by resampling the SSW events 100 times with repetition. **If the magnitude of the anomaly exceeds the interquartile or the 10^{th}-90^{th} percentile ranges the anomaly is highlighted by stippling or hatching, respectively. See Section 2.2 for details.**

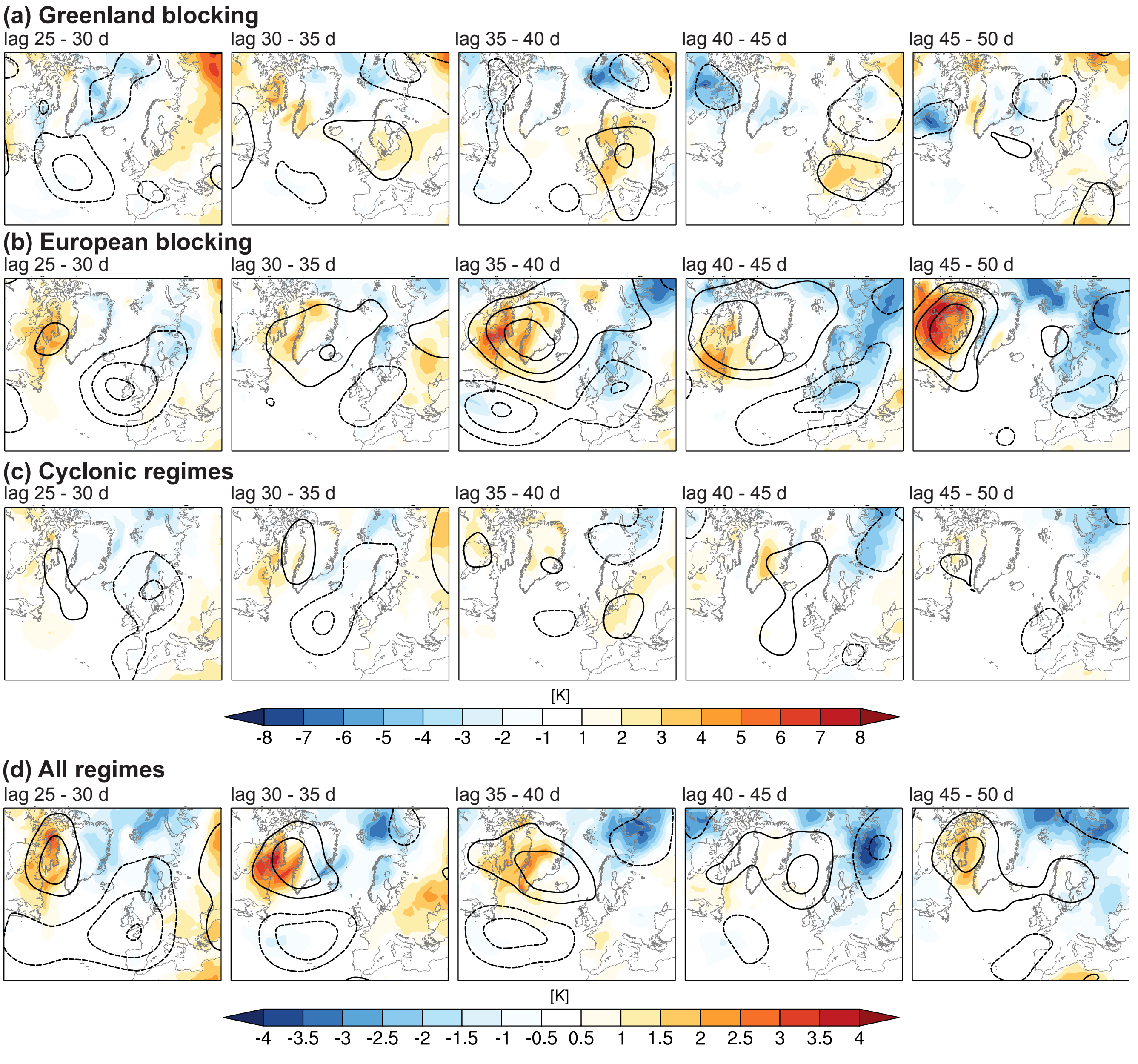

Figure A5. As Fig. 4 but for days 25 - 50.