

Manuscript Title: Stratospheric influence on the winter North Atlantic storm track in subseasonal reforecasts

Authors: H. Afargan-Gerstman, D. Büeler, C. Ole Wulff, M. Sprenger, D. I. V. Domeisen

Recommendation: Major revisions

Overall Opinion

This is my second review of the manuscript. The prior version of the manuscript lacked two key elements that I thought made it unsuitable for publication: (1) The study focused on only one model, whereas the database has several other models with hindcasts that could be explored for the same phenomenon; and (2) the paper lacked any dynamical insight into *why* the results were the way they were. The authors responded that doing the analysis for more than one model would be too laborious and that the ECMWF model was a well-trusted subseasonal model. I *kind of* agree with this thought, though I think the paper would be strengthened with more than one model analyzed for the study. But, I won't hold up publication of the study based just on that issue. For the second point, the authors expanded their analysis to look at the nature of downward propagation and the tropospheric circulation after a strong or weak vortex event. Specifically, the authors contrast between ensemble members which correctly predict the anomaly of the North Atlantic cyclone frequency following a strong/weak polar vortex event and those that do not. However, I don't think the methodology used actually addresses the dynamical interpretation of the results that the authors intend.

So, taken together, I would consider the paper ready for publication but only after **major revisions**, particularly to the dynamical interpretation portion of the paper.

Major Revisions

1. **Sampling Issues and Figure 4.** The authors continue to note that sample size is a concern for many of their results, and I agree with this point. However, one area this is not addressed enough is in **Figure 4**. In particular, the sample size between reanalysis and reforecasts for the SSW and strong vortex events is about a factor of 10 different, which makes comparison of the probability distribution functions (PDFs) very difficult (in fact, I question the reproducibility and representation of a distribution of a variable with only 14 samples). For example, the authors could repeatedly sample 14 random cases (with replacement) from the reforecasts to make a comparative PDF with reanalysis. I think in its current form, it is hard to argue about statistically significant differences in these distributions, whether comparing the cases or comparing reanalysis to reforecasts. The authors may want to pursue alternate strategies to strengthen this argument in the paper.
2. **Dynamical Interpretation.** I like the efforts that the authors made in trying to bring some dynamical insight into their model evaluation study of Atlantic storm track changes due to stratospheric polar vortex variability. However, I am not convinced that the analyses shown actually accomplish this effort. In particular, the authors use mean sea level pressure (MSLP) to represent the tropospheric circulation changes after "successful" and "unsuccessful" forecasts. However, storm tracks are defined in this study using MSLP. So, it is a bit circular to argue that differences in MSLP ("the tropospheric circulation") are the

leading reason why there are changes in storm tracks (which are determined by MSLP). I like the use of lower tropospheric winds and even looking at the lower stratosphere (Z100). But, the authors should reconsider how they measure the tropospheric circulation and consider other variables for that other than MSLP.

Minor Revisions

1. **Lines 14-18.** This sentence is long and confusing to understand. Please revise.
2. **Lines 74-75.** The acronym "ERA5" already contains the word "reanalysis" in it. So, it is redundant to say "ERA5 reanalysis."
3. **Line 228.** Please move the comma from after "ERA5" to after "reforecasts."
4. **Lines 313.** There is no need to define "MSLP" again here. Also, I think it is unnecessary to introduce another acronym into the paper for the 100 hPa geopotential height anomalies. Instead, the authors can just use the already-defined acronym for 100 hPa geopotential height anomalies (Z100). Can you just write "Z100 anomalies?"
5. **Lines 338-340.** This sentence structure (with the parentheses) is no longer favored in journal articles for readability and understanding. Please rephrase as two sentences or in another way. Same comment for **Lines 411-412.**
6. **Line 353.** It looks like there is a missing figure reference here ("??").
7. **Lines 362-363.** I am unclear what "larger natural variability" means, particularly in reference to model reforecasts. What does "natural variability" in a simulated atmosphere mean?
8. **Line 364.** "...found to b associated with a..." → "...found to be associated with a..."
9. **Lines 445-448.** The authors previously mentioned that they were going to have a Github site to make their data publicly accessible. This site is not listed here - please add the information for completion.