**MWR-D-22-0318, Biernat et al., Responses to Reviewers**

The authors thank the reviewers for their feedback on the revised manuscript. Our responses (red) to the reviewer comments (black) are given below. We provide line numbers corresponding to the latest manuscript to indicate where changes were made. At the end of the current document, we indicate one very minor grammatical correction that was made in the latest manuscript.

**Responses to Reviewer #1**

**Major comment:**

1) I disagree with the authors’ response to not have a hypothesis. There already is a hypothesis implied in this study stated by the authors’ “expectations” (lines 129-131; 462-466). The purpose of this study, stated in the abstract, is the “increase understanding of processes that influence the forecast skill....” Therefore, this is a process study, and in order to improve an understanding of processes, there should be an a priori expectation of what processes may be important. It seems from what they’ve stated and based on previous literature, latent heating can lead to large forecast errors in midlatitude cyclones. So perhaps the hypothesis is that you do not expect Arctic Cyclones to be physically different from midlatitude cyclones, and thus you hypothesize

that low forecast error will be associated with Arctic Cyclones with relatively high latent heating. Using rationale like this would hardly add much length to the manuscript. This is almost stated as such on lines 129-131 already.

**Response:**

The reviewer has good points for the inclusion of a hypothesis. We agree with the reviewer, and so we have added the following text to L149–157 of the latest manuscript:

As previously discussed, forecast errors related to baroclinic processes and latent heating are expected to contribute to forecast errors in the synoptic-scale flow over the Arctic and to forecast errors in ACs. It is further expected that more vigorous baroclinic processes and latent heating may be associated with greater forecast errors. It is hypothesized that there are more vigorous baroclinic processes and latent heating over the Arctic and in the vicinity of ACs during low-skill periods compared to high-skill periods, and that there are relatively vigorous baroclinic processes and latent heating in the vicinity of ACs during low-skill periods that are characterized by low forecast skill of intensity.

Minor Comments:

1) 56: What exactly is a forecast “bust”?

**Response:**

A forecast “bust” is defined by Yamagami and Matsueda (2021) as a very poor forecast that occurs when the anomaly correlation coefficient (ACC) of 500-hPa geopotential height over the Arctic at a 144-h forecast lead time is less than the climatological 10th percentile value of ACC and when the root mean square error (RMSE) of 500-hPa geopotential height over the Arctic at the 144-h forecast lead time is greater than the climatological 90th percentile value of RMSE.

We added some of this discussion as a new sentence on L56–58 of the latest manuscript. Because of this new sentence, the first sentence of the corresponding paragraph from the previous version of the manuscript has been split into the two sentences found on L55–56 and on L58–61 of the latest manuscript. Also, since we now define “root mean square error” as “RMSE” on L57 of the revised manuscript, we replaced “root mean square error (RMSE)” with “RMSE” on L192 of the revised manuscript.

2) 176: In the Figure 1 caption, there are not any names of nations on the plot. Perhaps

the authors mean “select regions.”

**Response:**

This is a good point. We replaced “nations” with “select regions” in the Fig. 1 caption on L189 of the latest manuscript.

3) 185: resolution → spacing

**Response:**

We replaced “resolution” with “spacing” on L198 of the latest manuscript. For consistency, we also replaced “resolution” with “spacing” on L238 of the latest manuscript and on L331 of the latest manuscript.

4) 521-527: Does Figure 8a suggest that low-skill periods resemble the positive phase of

the Arctic Oscillation? If so, please add to text.

**Response:**

In the past, we investigated the Arctic Oscillation (AO) during low-skill periods and high-skill periods using the normalized daily AO index from the Climate Prediction Center (CPC) (referenced below). Although Fig. 8a shows relatively low 500-hPa geopotential height over the Arctic during low-skill periods, which may resemble the positive phase of the AO, we found that the mean AO index during low-skill periods is slightly negative and not statistically significantly different from the mean AO index during the 2007–2017 climatological period (not shown). Given that there are no statistically significant results for the AO index, we chose to not discuss the AO index in the manuscript.

Reference:

CPC, 2021: Normalized daily AO indices. CPC. [Available online

at ftp://ftp.cpc.ncep.noaa.gov/cwlinks/norm.daily.ao.index.b500101.current.ascii.]

5) 555-573: It is insightful that the low-skill and high-skill ACs have statistically distinguishable

environments. It would be useful if the authors offered at least one sentence

on their interpretation as to the low-skill ACs not being statistically distinguishable

from climatology while the high-skill ones are distinguishable.

**Response:**

We added the following discussion to L582–589 of the latest manuscript:

It is expected that the intensity of ACs may influence the magnitude of baroclinic processes and latent heating in the vicinity of ACs, and vice versa, such that statistically significantly stronger and weaker ACs may tend to be located in regions of statistically significantly more and less, respectively, vigorous baroclinic processes and latent heating. There being statistical significance in the mean value of intensity for only ACs during high-skill periods (Fig. 9a) is thus likely consistent with there being statistical significance in the mean values of all quantities shown in Figs. 9b–d and Figs. 10a–d for only ACs during high-skill periods.

6) 666, 815: What result would have been needed for the authors to not consider the

evidence “weak?” Doesn’t the statistical significance test just determine yes or no?

**Response:**

While the statistical significance test does determine whether or not there is statistical significance in the mean intensity RMSE for a given forecast lead time, we want there to be statistically significantly higher mean intensity RMSE of ACs during low-skill periods compared to high-skill periods for the majority of the seven forecast lead times (i.e., at least four of the seven forecast lead times) for us to consider the evidence that ACs are associated with lower forecast skill of intensity during low-skill periods compared to high-skill periods to not be weak. Having a statistically significantly higher mean intensity RMSE of ACs during low-skill periods compared to high-skill periods for the majority of the seven forecast lead times, as opposed to three of the seven forecast lead times, as found in our study, would indicate that ACs are more consistently associated with lower forecast skill of intensity during low-skill periods compared to high-skill periods.

Other change:

Changed “midlatitudes cyclones” to “midlatitude cyclones” on L131–132 of the latest manuscript, as “midlatitude” should not be plural in this instance.