

Overall Thoughts

This is my second review of the manuscript. The authors did a great job acknowledging and addressing the various comments from the two reviews of the original manuscript. In particular, the authors have streamlined their discussion of the results, discussed further the summer results, and clarified some of their methodological choices. Altogether, the manuscript is stronger and clearer and remains scientifically interesting. I still think that the manuscript has the potential to be significant in the seasonal forecasting community and in research involving Arctic and mid-latitude weather connections across timescales. I still think there are a couple of areas that need some slight improvement, but this will not ultimately curtail the publication of the manuscript. Hence, I recommend that the paper be **accepted for publication pending minor revisions**.

Minor Revisions

1. **ACC Skill Score Discussion.** I am not going to belabor or get into a “back-and-forth” with the authors about the use of ACC as forecast skill. I am well aware that the community has adopted this metric to represent a form of forecast skill. If not already read and/or of interest, I encourage the authors to read Murphy and Epstein (1989) to understand the context of ACC as a skill metric.

With that, I appreciate the additional context provided by the authors on comparing ACC “high skill” to the relatively large RMSE errors (**Lines 475-477**). This type of caveat reflects the spirit of my original comment – a high ACC score does not necessarily mean high forecast skill for a particular variable when compared to other skill metrics. Admittedly, for forecasting purposes, magnitudes of features could be considered more important than just pattern correlation, particularly when forecasting extremes on subseasonal-to-seasonal timescales. So, thank you for adding this text! There are a couple of other areas in the manuscript where this concept could be emphasized as well.

- a. **Abstract:** In lieu of the caveat above, the statement that there is a “significant increase in seasonal forecast skill” of Z500 isn’t quite accurate. The improvement is shown only for ACC, not for all skill metrics; e.g., RMSE remains essentially unchanged on average.
 - b. **Line 621:** The improvement in “forecast skill” is only shown for ACC, so this would be an improvement of the ACC but necessarily of the total Z500 “forecast skill.”
2. **Field Significance.** The authors acknowledge the gridpoint-based hypothesis testing when looking for significant differences in ACC scores. However, they have not mentioned performing *field significance/false discovery rate* tests on the Z500 field in response to this concern. Have the authors investigated these tests? If so, what is the

resulting map? The conclusions may or may not change, but this added analysis would be important for completeness.

Grammatical / Typographical Error Corrections

1. **Lines 5 and 99.** Please capitalize “Northern Hemisphere.”
2. **Line 89.** “The ensemble of 50 atmospheric initial conditions are based on ERA5...” → “The ensemble of 50 atmospheric initial conditions is based on ERA5...”
3. **Line 91.** “an ensemble of 11 conditions *are* used...” → “an ensemble of 11 conditions *is* used...”
4. **Line 96.** “...the generation of the ocean and sea ice initial conditions *do not* make use of...” → “...the generation of the ocean and sea ice initial conditions *does not* make use of...”
5. **Line 165.** The parameter name “rn_csrd” should be “rn_csrdg” based on Table 1 (or conversely, Table 1’s labeling is incorrect).
6. **Line 200.** “The increased SIC in the *Chuckhi* and Bering seas in DJF *are not* significant...” → “The increased SIC in the *Chukchi* and Bering Seas in DJF *is not* significant...”
7. **Line 328.** Double period after “absent.”
8. **Line 335.** “The linear correlation ranges between –0.5 *to* –0.7 in the two seasons...” → “The linear correlation ranges between –0.5 *and* –0.7 in the two seasons...”
9. **Line 421.** Remove “is” after “SIT.”
10. **Line 423.** “An asymmetric impact associated with irreversible processes such as ridging *are* also...” → “An asymmetric impact associated with irreversible processes such as ridging *is* also...”
11. **Line 452.** “*Greenland sea*” → “*Greenland Sea*”
12. **Line 471.** Eliminate “the” between “surprising” and “there.”
13. **Line 503.** “strenghtening” → “strengthening”
14. **Line 560.** “heatfluxes” → “heat fluxes”

15. **Line 602.** Missing the closing parenthesis after “S9.”

Reference

Murphy, A. H., and E. S. Epstein, 1989: Skill scores and correlation coefficients in model verification. *Monthly Weather Review*, **117**, 572–582, [https://doi.org/10.1175/1520-0493\(1989\)117%3C0572:SSACCI%3E2.0.CO;2](https://doi.org/10.1175/1520-0493(1989)117%3C0572:SSACCI%3E2.0.CO;2).