

## ***Interactive comment on “Seasonal statistical-dynamical prediction of the North Atlantic Oscillation by probabilistic post-processing” by André Düsterhus***

### **Anonymous Referee #1**

Received and published: 21 October 2019

This study presents a novel empirical postprocessing method for combining dynamical seasonal weather forecasts with previously identified statistical predictors in order to achieve enhanced predictive skill. The proposed method involves constructing bootstrapped pdfs for individual ensemble member predictions. These are then combined and normalised to give a pdf. The statistical predictors are treated in the same manner. The normalised model and predictor pdfs are then combined by multiplication and normalised again to produce a final forecast pdf. Results for the NAO show improved performance of the combined pdf, particularly compared to the model pdf alone. Results for atmospheric fields are patchy, likely reflecting the teleconnections involved.

C1

I congratulate the author on an interesting methodology and a concisely written manuscript. Although empirical in nature, the proposed post-processing method is clearly statistically motivated and attempts to capture and quantify relevant uncertainty. However, I would like to see a few details clarified, specifically:

Lines 56-60 – Please describe briefly the bootstrap procedure. It is not necessary to explain the block bootstrap proposed by Wang et al. (2014), but it is important to know what is being bootstrapped. I assume it is the daily pressure fields?

Section 3.3 - The Earth Mover’s Distance is an interesting method for evaluating the quality of the forecasts, but in addition it would be useful to see an evaluation based on a conventional strictly proper scoring rule, such as the CRPS mentioned in the discussion.

Section 3.4 - I think “meteorological variable fields” in line 97 and “atmospheric field” in line 100 refers to the individual ensemble member fields? Please clarify.

Section 3.4 - I am unclear what exactly is taking place in Equation 5. Are these the  $e_i(v)$  and  $M(v)$  for the NAO forecast then used to weight entire fields for the atmospheric variables, or are these computed individually for each variable, or at each grid box for each variable? Please clarify.

Minor points and typos:

Lines 23-24: “The paper” -> “This paper”, “bases on selection” -> “bases selection”, “on the basis of statistical predictors” -> on “statistical predictors”

Line 33: “at the hand of” -> “for”

Section 2: Were the grid boxes latitude weighted before computing the EOFs?

Lines 66-71 - A table summarising the definitions of the predictors might be helpful here.

Line 81: “clearer signal as” -> “clearer signal than”

C2

Lines 190-2: This doesn't make sense to me, sharper predictions should have less uncertainty. Do you mean variability between predictions? Please clarify

Lines 206-7: There are many good reasons not to use ACC (e.g., not a proper scoring rule), but the ACC does NOT assume normality. Some tests of ACC assume normality, but the statistic itself does not.

---

Interactive comment on Nonlin. Processes Geophys. Discuss., <https://doi.org/10.5194/npg-2019-50>, 2019.