This study analyzes the impact of the initial state on the forecast skill at week 3 (days 19-25). The flow dependent predictability being an inherent feature of the atmosphere-ocean dynamics, affects generally any forecast range and it has been extensively studied. At the extended ranges, when the forecast skill can be generally modest, knowing which flow conditions can lead to a more accurate forecast is particularly relevant. At sub-seasonal and seasonal time scale the higher predictability is typically associated with the distribution of tropical heat anomalies (e.g. ENSO and MJO) or the anomalies in the stratosphere. In this paper the focus is on the forecast skill conditioned by to the flow configuration over the Atlantic sector at initial time. The well-known 4 Euro-Atlantic regimes are used to stratify the initial conditions. The authors show that the forecast initiated with the two phases of NAO are more skillful at week 3 than the forecasts initiated in other flow configuration. This result is consistent with findings from previous work done on flow dependent verification at medium range.

The analysis is limited to the boreal extended winter and it uses 20 years of reforecast data from two sub-seasonal forecast, namely the new version of CNRM and the ECMWF operational system. The skill evaluation is based on the 2m temperature over the Northern Hemisphere. Although the flow dependent skill assessment is based on anomaly correlation, a probabilistic skill estimate is also included.

The study is well presented with an appropriate number of figures and a valuable discussion of the results.

Specific comments:

L 80-85 In the computation of daily forecast anomalies, the model daily climate has been computed as a function of lead-time? Please explain what has been done.

L155 It is hard to believe that at week3 and 4 the effect of the initialization shock is still evident.

L165 For a “naïve forecast” is intended the climatology? Please explain
Comparing Fig.1 and Fig.2 it is clear that the CRPSS is rather more informative than the ACC. In fact, a positive ACC, even if statistically significant, does not translate into a skillful forecast. Why not using the CRPSS as well as ACC for fig.4 and fig.5?

L205-210 The case of 2009-2010 was characterized by a very persistent negative NAO. Considering that the NAO negative typically has a longer life cycle than the NAO positive (Dawson 2012), a forecasts made by persistence could be equally skillful. On the contrary, for the forecast initiated in NAO positive the persistence forecast is expected to be less accurate because the NAO positive life cycle is shorter. By adding an extra bar in Fig.4 indicating the number of forecasts based on persistence with ACC>0, we could see if the good persistence forecast are actually associated with the NAO negative cases.

Results from Fig.6 could indicate that the NAO negative is more persistent in the forecast than the NAO positive.

Showing the typical residence time for the NAO positive and for the NAO negative separately using the regimes daily attribution would help to interpret the results. Fig. 9 indicates that for strong NAO conditions the residence time can be 20-30 days, could it be that this is mainly associated with the strong NAO negative phase?

Pag.12 For both systems, the forecast initiated with NAO+ indicate higher ACC already by day 5 and beyond day 15 the ACC difference increases.

L 220 A skillful forecast is a forecast that performs better than climatology or than persisting the anomalies of previous 5-8 days. Some of the 68 forecasts that have an ACC>0 in both systems might have a skill close to climatology. Please add some discussion on this.

L280 By considering the very long period of era5 (1950-2017) we might introduce some important decadal variations. Please comment.