This paper investigates the behavior of North Atlantic ARs in reanalysis and ensembles of global simulations in historical and future simulations. The manuscript is in good shape. It is well-written and has a clear structure. I don't have major concerns for the paper, but some minor suggestions and comments for discussion:

We thank the reviewer for a thorough review of our manuscript and his/her valuable suggestions that really help to improve the manuscript.

L193: Guan and Waliser (2015) does not provide a thorough review of AR detection methods, instead, GW15 explained their AR detection algorithm in detail. Since GW15's algorithm is not adopted in this paper, it is better to cite a different paper as the "overview". For example:

Shields et al. (2018): Atmospheric River Tracking Method Intercomparison Project (ARTMIP): project goals and experimental design, Geosci. Model Dev., 11, 2455-2474, https://doi.org/10.5194/gmd-11-2455-2018, 2018.

Thank you for this suggestions. We will cite Shields et al., 2018 and remove Guan and Waliser (2015).

Figure 4: I wonder how the duration distribution would change if the duration threshold to tuned to 6 hours or 12 hours. How will the duration (including those excluded short-lived events (<18hrs)) change in the future climate?

We have have recalculated this exemplary for the MPI-ESM realization. The main effect is, as expected, an increase of about 55% in the total of detected ARs. The left hand figure below compares the distributions calculated with a threshold set to 18 hours (orange) and 12 hours (blue). Generally, both distributions are similar and indicate most ARs in the short term fraction (with >50% ARs 30 hours or shorter). The right hand figure displays the relative change for the RCP8.5 scenario [%, i.e. future minus historical period *100 / historical period)]. In both distributions the tendency towards extra long (48 hours or higher) ARs can be seen. However, these extended ARs remain typically rare (<10%) in both historical and future climate.

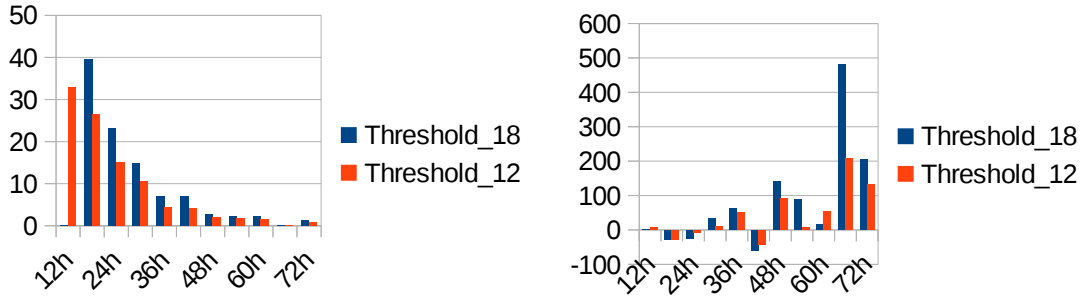

Figure 5a: instead of showing the actual number of AR days, it might be helpful to show the percentage of AR days to the number of days of the period.

We agree. We have adopted figure 5a accordingly. We have also changed Fig. 8 in the same way which displays the change in AR days in a future warmer climate. (See the new figure 5 in our reply to RC3)

Table 3 and L367: Is it possible that the dynamical field is more active in the future simulation as well?

Table 3 shows that the number of AR increased in the future runs: it makes sense if the historical AR threshold is applied in the future run – however, here the future ARs are detected with IVT thresholds calculated from future simulation, so the higher moisture load is reflected in the future IVT threshold to

some extent. Therefore, I am curious if the change in dynamical filed also contribute to the increase in future AR number.

Yes, it is likely that both dynamical as well the thermodynamical changes contribute to the increase in AR number. Hence, it could be that more ARs are initiated in the source region (outside our model domain) or that the pathway from the source area to Europe changed. This we can not analyze from our results but this question was addressed in global CMIP5 models (which are more suited for this question) . Lavers et al., 2013 argue for mostly thermodynamical changes (moisture increase) for the future increase in ARs. Gao et al., (2016), likewise found thermodynamics playing a dominant role but pointed out potential dynamical changes in the seasonal distribution of ARs.

L450: I am not sure if I fully understand the methodology. Are the origins of ARs being categorized by the southernmost latitude? If so, isn't it more related to the curvature of AR's shape than the actual origin?

No it's not the southernmost. We will reformulate this description. First, for every AR that is detected over land we check the time and the latitude when the AR was registered for the first time at 10°W. Hence if we have an AR entering a land grid cell we "look" in our "archived" when this AR it was registered for the first time. For the post-statistical analysis we then summarize by regions south of 45° degree between 45-60° and >60° to obtain the results.

L563: please rewrite, grammar problem.

"The regional climate model is further used to investigate the dynamics of ARs in present and future climate. For this in an ensemble of in total 34 simulations was carried to downscale global climate model scenarios from the CMIP5 suite with a coarse resolution between 1.4 – 3°. The models were used to downscale three greenhouse gas scenarios (RCP2.6, RCP4.5, and RCP8.5). "

Done. The paragraph reads now:

The regional climate model is further used to investigate ARs in the present and future climate. For this an ensemble of global CMIP5 climate simulations (1.4° - 3° resolution) was downscaled to reach 0.22° resolution. In total 34 simulations were carried out with the regional model following greenhouse gas scenarios RCP2.6, RCP4.5, and RCP8.5. For this in an ensemble of in total 34 simulations was carried to downscale global climate model scenarios from the CMIP5 suite with a coarse resolution between 1.4 – 3°. The models were used to downscale three greenhouse gas scenarios (RCP2.6, RCP4.5, and RCP8.5).