Referee Comment #3

General comments:

The paper is well structured, and the algorithm adopted is easy to understand thanks to the illustrative figures. Although several approaches have been developed to identify ARs, their innovation relies on overcoming the RCMs' limitations where most of the runs are focused over land, and this precludes capturing the long way over the ocean. The success of this approach will allow the use of RCMs to provide more accurate precipitation amounts than GCMs and to perform less computationally costly simulations such as online aerosol runs to understand ARs mechanisms. Then, I found this work a valuable advance to analyze the impacts of the AR's landfalling.

Under these arguments, I recommend accepting this work after addressing a minor revision detailed below.

We sincerely appreciate your thoughtful and positive feedback on our research. Your comments on the algorithm, as well as on the innovation in addressing RCMs limitations, are of great value for our team. Your recommendation to accept the work after addressing a minor revision is encouraging. We have addressed your suggested revisions in the specific comments section to ensure the quality of the manuscript. Thank you for considering our work a valuable advance in the field.

Specific comments:

Introduction

In line 55, the authors mention the lack of research about the impact of aerosols on ARs but they did not discuss the challenges nor mention previous works such as Counterbalancing influences of aerosols and Greenhouse gases on atmospheric Rivers by Baek and Lora.

Thank you for the remark. We are going to include this discussion in the Introduction section.

Methods

How can the AIRA be sure that is detecting an AR and not the branch of a low system with a bigger enough radius? Does $\Delta\Theta < 25$ guarantees this fact? Maybe introducing SLP values will avoid this concern.

Thank you for your question. Yes, the maximum direction difference allows us to ensure that moisture transport is taking place in the direction of the AR, distinguishing it from the branch of a low system. With respect to the slp suggestion, we kindly appreciate it. However, although it may provide some extra information which would also allow us to distinguish low systems from ARs, the currently presented method (AIRA) is sufficiently capable of performing this distinction.

In Table 1 the authors show the imposed parameters. To demonstrate the robustness of the approach some discussions about the sensitivity of these parameters are needed. For instance, how many percentages of ARs increase/decrease if the IVT threshold is modified?

Thank you for your comment. This remark was also mentioned by the other referees. Following your suggestions and those of the other two referees, we have performed an analysis of the sensitivity to the IVT threshold given a fixed minimum duration and the sensitivity to the duration threshold given a fixed IVT threshold. The results are exposed in the tables below and include the variation in the number of ARs in each simulation, the number of common ARs events, the percentage of common AR time steps and the mean intensity and mean duration of the identified ARs.

Table: Sensitivity analysis to the IVT threshold, given a fixed minimum duration, of the number of ARs identified in the three simulations, the number of common AR events, the percentage of common AR time-steps and the mean intensity and duration of the ARs of the three simulations.

T = 10 h	Γ **(kg m^{-1} s^{-1})**								
	200	225	250	275	**300**	325	350	375	400
ARs BASE (#)	194	212	230	236	**244**	245	252	244	238
ARs ARI (#)	166	195	210	217	**248**	254	247	230	234
ARs ARCI (#)	173	205	222	232	**250**	244	243	230	233
ARs COM (#)	39	54	63	73	**80**	92	94	86	91
COM time-steps (%)	24.79	28.51	32.11	33.65	**37.16**	40.54	38.91	38.11	40.65
$\overline{IVT}$ **BASE (kg m^{-1} s^{-1})**	344.42	380.58	407.61	435.15	**469.20**	495.67	523.24	549.25	579.26
$\overline{IVT}$ **ARI (kg m^{-1} s^{-1})**	345.14	373.10	407.43	440.88	**465.47**	491.42	520.35	551.15	589.44
$\overline{IVT}$ **ARCI (kg m^{-1} s^{-1})**	347.59	377.23	404.54	434.99	**459.18**	490.43	517.50	550.55	574.19
$\overline{d}$ **BASE (h)**	53.17	50.73	47.54	45.36	**42.55**	40.44	40.11	37.75	36.35
$\overline{d}$ **ARI (h)**	56.76	52.44	51.24	47.47	**43.13**	41.26	39.00	37.69	36.46
$\overline{d}$ **ARCI (h)**	56.61	53.45	48.71	46.72	**43.79**	43.10	41.82	40.03	36.76

Table: Sensitivity analysis to the minimum duration threshold, given a fixed IVT threshold, of the number of ARs identified in the three simulations, the number of common AR events and the percentage of common AR time-steps.

IVT = 300 kg m^{-1} s^{-1}	**T (h)**								
	4	6	8	**10**	12	14	16	20	24
ARs BASE (#)	267	262	253	**244**	233	222	209	183	162
ARs ARI (#)	261	259	254	**248**	232	225	212	193	170
ARs ARCI (#)	267	261	254	**250**	233	226	212	198	171
ARs COM (#)	86	85	84	**80**	74	69	65	58	50
COM time-steps (%)	35.73	35.83	35.94	**37.16**	36.19	36.12	35.97	36.41	36.75

On one hand, a lower IVT threshold results in a decrease in the number of ARs but also in an increase of their duration, because two very close in time events could be identified as a single but longer event. On the other hand, increasing the IVT threshold over 300 kg m-1 s-1 reduces the mean duration of the ARs but has little impact on the number of ARs itself. For instance, the selection of an IVT threshold of 400 kg m-1 s-1 would have resulted in a decrease in the number of ARs in BASE, ARI and ARCI of 2.5%, 5.6% and 6.8%, respectively.

With respect to the sensitivity of the duration threshold, the results turned as expected. The higher the minimum duration imposed, the lower the number of ARs identified that meet this condition. Furthermore, we also wanted to remark that the selected parameter (T=10h), gives rise to the highest percentage of common AR time steps, with 80 common events that have allowed us to perform our comparison study.

To better contextualize your methodology, I missed a discussion comparing the AIRA approach with other methodologies of other tracking approaches, For instance, a review can be found in: Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project Goals and Experimental Design by Ruth et al.

Thank you for your comment. Other referees also noted this issue. In the revised manuscript, we are going to put AIRA in the ARTMIP context and classification, including its main differences with the IDL ARDT (Ramos et al., 2016) and Brands ARDT (Brands et al., 2017) algorithms, which are the most similar to AIRA and also detect ARs over the Iberian Peninsula. As a preliminary observation, the main contrast is that both algorithms make use of spatial tracking, while AIRA never uses it, as it is intended to perform also in regions close to the domain edges. This is indeed the case in our study, with the detection lines located very near the limits of the spatial domain.

Results

Following the previous comment, some validation against observations (e.g. satellite images) and/or using the ARs inventory/catalogs is needed to be the coherence of your approach with the ARs already identified along the bibliography.

In line 196 the authors mention. "It was found that most of the ARs identified by AIRA also matched those identified by global-scale algorithms, as reported by Brands et al. (2017)." How many coincidences did you find? Did you find more 'real' ARs in BASE or in ARCI? Do you think that some discrepancies may be due to a different approach or the use of an RCM instead of a GCM?

By the time this research was conducted, there was a website mentioned in Brands et al. (2017) with their *Atmospheric Rivers Archive* available: http://www.meteo.unican.es/atmospheric-rivers. This catalogue documented all ARs detected by their algorithm using ERA-20C data and we compared our results with it (see figure below for some qualitative examples). Unfortunately, the page was shutdown. To answer your questions, we have contacted the authors and they have provided us a database with all the information through a Zenodo repository (https://doi.org/10.5281/zenodo.8010794), although the representation tool is not available anymore. In the revised manuscript, we are going to assess the coincidences to the fullest extent possible to answer your first two questions.

With respect to your last comment, we think that the discrepancies could be mainly due to differences in the methodology approach, like the IVT threshold or the shorter detection line used by the Brands ARDT to study W Iberia region. In addition, we have considered the same line to study ARs on the southwest of the IP, while Brands ARDT employed a different line for S Iberia. Furthermore, aerosol effects can cause spatial deviations, as seen in this research, potentially pushing ARs out of the study area and lowering the number of coincidences. A more detailed discussion of these differences will be included in the revised manuscript.

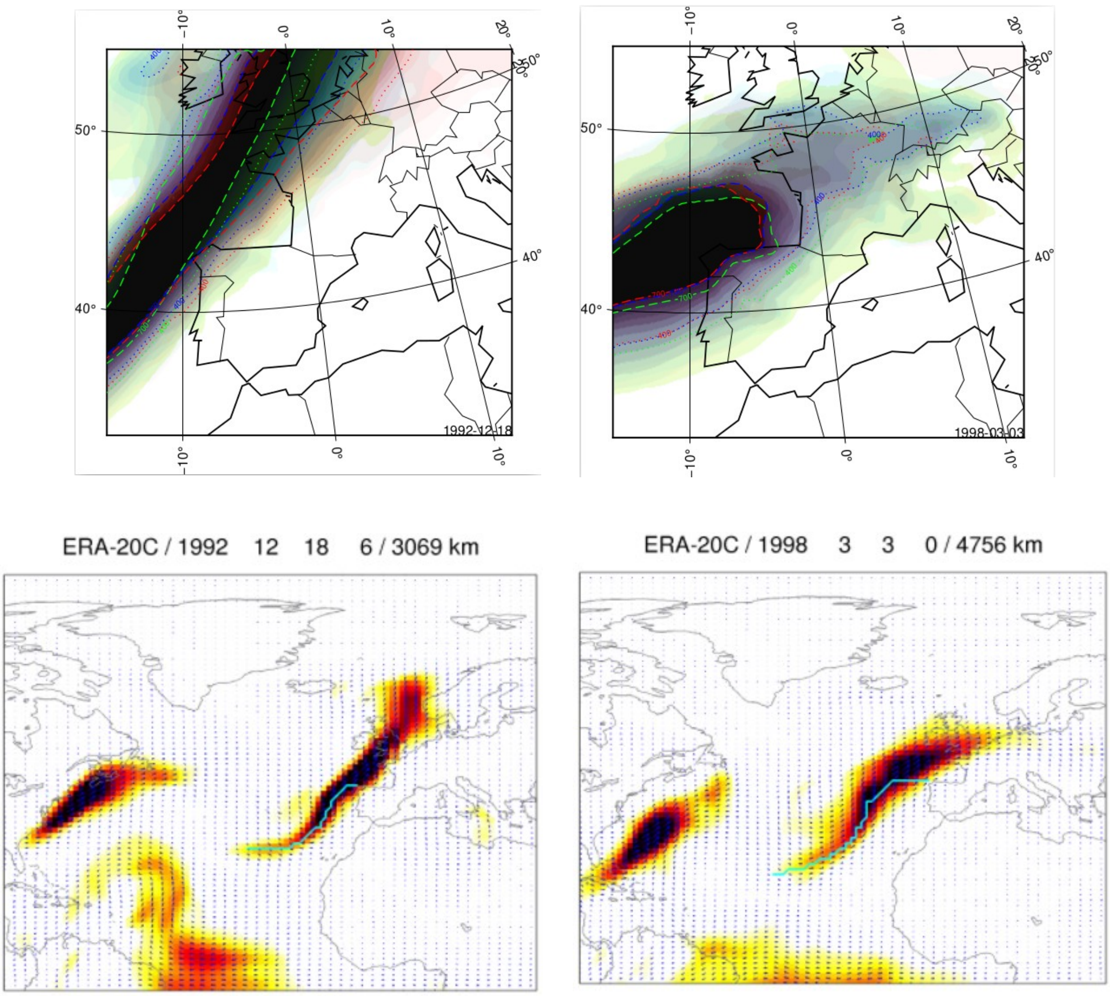

Figure: ARs identified the 1992-12-18 (left) and 1998-03-03 (right) by AIRA (top) and Brands ARDT (bottom, Brands et al., 2017). In the top images, green, red and blue contours/shades represent the ARs of the BASE, ARI and ARCI simulations, respectively.

In Line 224 the authors assert that the ARs explain the 30% of the precipitation, it is not clear what area did you use to obtain this value, and the Fig. 5 shows strong spatial variability to perform a spatial average.

ARs don't explain the same percentage of total precipitation in every cell of the domain, as can be seen in Fig. 5. We have stated that "In all three simulations, it is apparent that the maximum percentage of total precipitation attributable to the presence of ARs is close to 30% and occurs along the western Iberian coast, which is the impact zone of the ARs". This means that ARs could explain up to a 30% of the precipitation of a given location/grid cell and we have also shown the locations in which this maximum takes place. Nevertheless, 30% constitues a maximum, thus the percentage of total precipitation related to ARs decreases in the rest of the study domain, especially in the points located far from the impact zone of the ARs. Furthermore, these percentages were calculated for the whole period, but as you have mentioned in the last paragraph of your specific comments, these percentages will have temporal variability, due to the interannual variability of ARs, thus changing the precipitation.

Furthermore, how accurate is the precipitation during these events? Is ARCI or BASE more representative of the observed precipitation?

This is a very interesting question. Let's use the observed precipitation of the Iberia database. Considering the AR-related precipitation of each simulation and the observed precipitation of those same days (different from simulation to simulation), we can plot the Taylor diagram shown below. The correlation coefficient of the three simulations is higher than 0.85 and ARCI presents a lower standard deviation than the rest. The three simulations represent quite accurately their related observed precipitation.

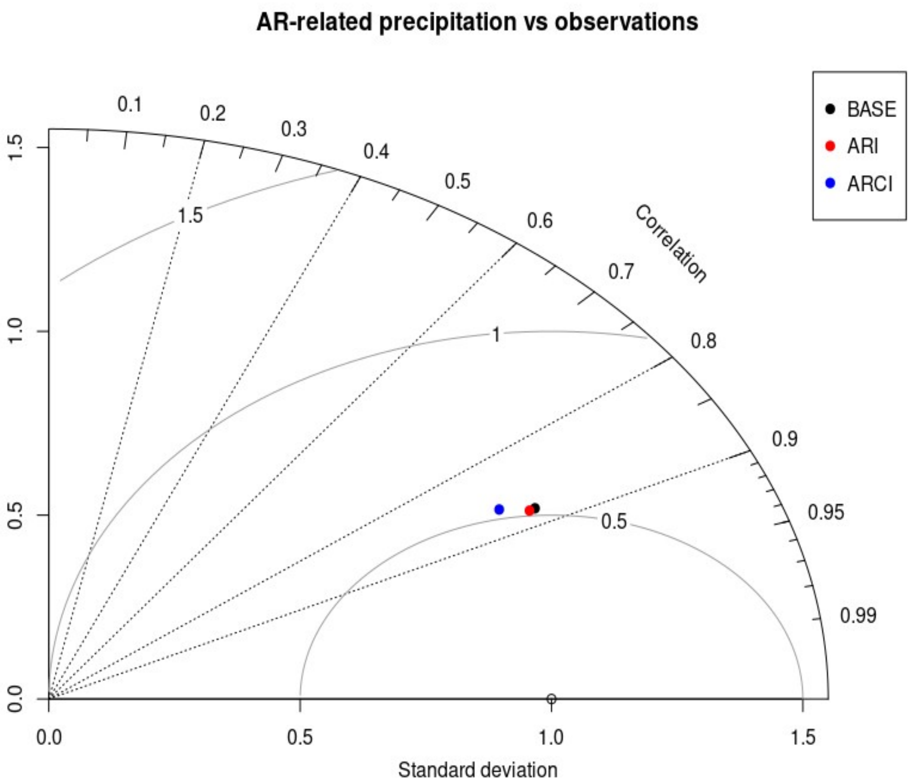

Figure: Normalised Taylor diagram of the AR-related precipitation in BASE (black), ARI (red) and ARCI (blue) with respect to the observed precipitation of the set of AR days of each experiment.

In Line 232. Only 37 % of the coincidence of ARs between ARCI and BASE looks like a few percent. When the simulations are described there isn't any mention of nudging or re-initialization of initial conditions has been mentioned. What percentage of these discrepancies could be due to different treatments of aerosols or due to internal variability of the simulations?

Thank you for your question. We have already mentioned some plausible explanations to this few percentage in the manuscript: "Only 37% of the time steps with ARs coincide among all three simulations concerning the BASE total. This low percentage could be attributed to weak events and the temporal limitations of the identified ARs, where the IVT threshold is exceeded in some simulations but not in others. Furthermore, aerosol effects can cause spatial deviations, as seen in the following sections, potentially pushing ARs out of the study area, decreasing the time steps with AR on the detection lines in some experiments, and thus lowering the coincidence percentage."

Although nudging was applied to the outer domain, neither nudging nor re-initialization of initial conditions have been used in the target (inner) domain. We were interested in allowing the model to run "freely" in this domain once the initial conditions had been stablished, in order to see how the different aerosol treatments affected the simulations. Nudging can reduce the internal variability of the model but it would have prevented us from obtaining the desired conditions in the simulations. Another referee requested to include a more detailed explanation of the simulations design, so we will address these comments there.

To determine the exact percentage of the discrepancies that could be due to the internal variability of the model, a deeper and more complex study should be done. It would require repeating the simulations to address their variability. This is an interesting question that falls out of the scope of this research.

When sea salt and dust clusters are analyzed (Fig. 7 and 10) It will be interesting to see mean ARs trajectories for each cluster (for instance superimposed with dotted lines).

Thank you very much for your comment. Following also the suggestion of Referee #2, we have added the trajectories of each AR belonging to a cluster and their mean trajectory, but instead of performing this approach to all the clusters and representing it on Fig. 7 and 10, we have focused on the most relevant clusters (discussed in the manuscript), and we have added the representation of the trajectories to Fig. 9 and 12, where the thickness fields are shown. For instance, you can find the resulting representation of the ARI clusters 2-3 in the figure below this paragraph. Each thin arrow represents an AR. It is located on its mean latitude with its mean direction and the length of the arrow is proportional to its mean intensity. The thicker arrow represents the mean characteristics of the ARs belonging to the cluster.

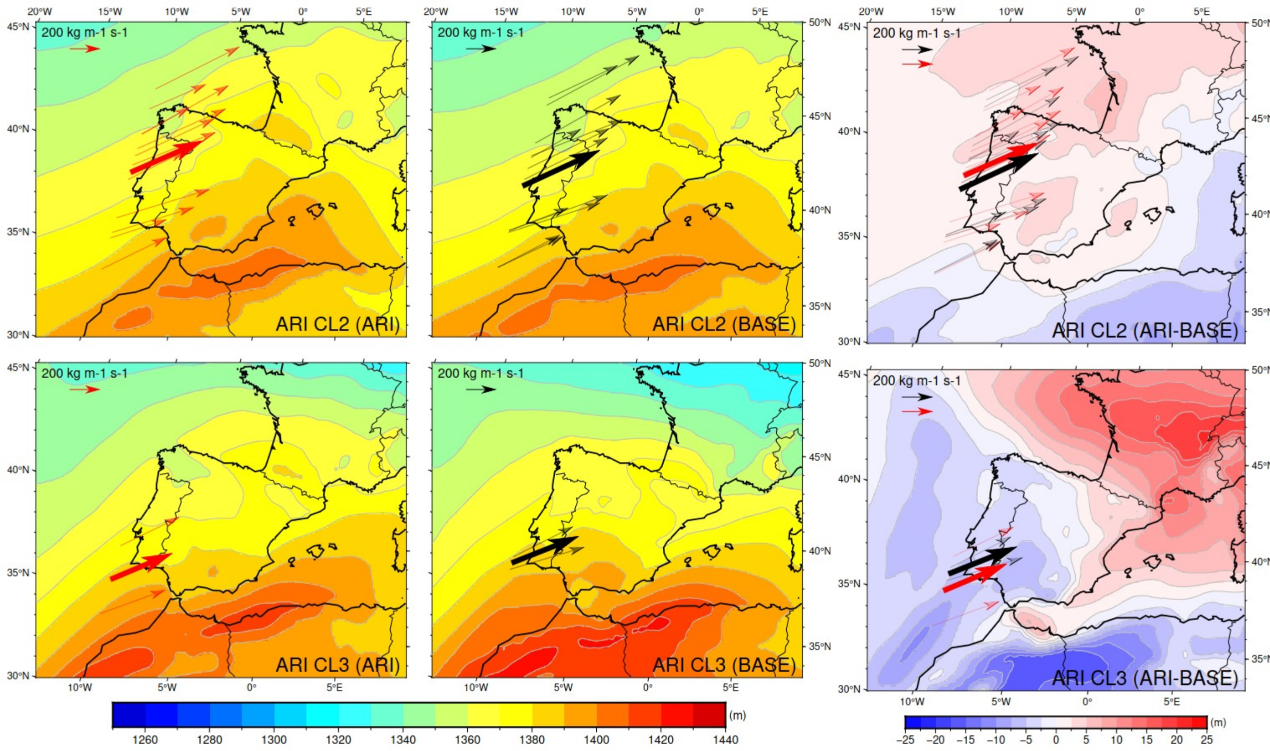

Figure: ARI and BASE mean thickness fields of the atmospheric layer between 1,000 and 850 hPa of the common AR events belonging to clusters 2 and 3 in the ARI simulation and ARI-BASE thickness differences. The same time steps are included in the representations of both experiments.

In the analysis of the differences to better understand the thermodynamics and dynamics changes, it will be illustrative to analyze whether the IVT changes are more due to IWV or winds.

In the analysis of the differences, we have found that direct, semi-direct and indirect aerosol effects play an important role in ARs behaviour and characteristics. These effects were translated into temperature differences that give rise to changes in the thermodynamic properties of the cloud droplets, as discussed along the manuscript. These thermodynamic/temperature changes trigger the thickness field differences and thus are the origin of the dynamic changes (weakening-strengthning of the thickness field gradient and winds).

Throughout the work, I missed more analysis about the impacts of ARs on precipitation. I understand that may be the scope of future work.

Thank you for your remark. You are right, a more in-depth study about the impacts of aerosols on the precipitation related to ARs is intended as the main topic of future works. However, we strongly appreciate your precipitation-related comments above and we have added some calculations/representations to the manuscript, especially to the case studies, with the aim of making this work more complete.

For the case studies will be interesting to show the spatial distributions of the precipitation (accumulated during the whole event and/or hourly) for the three simulations; BASE, ARI, and ARCI. These will provide some insights about how the intensity and trajectory of ARs impact on the precipitation distributions.

As said right above, we have included this to the case studies, following your valuable suggestion. The new results show that the precipitation distributions of the involved days are quite different from simulation to simulation, being the ARCI distribution the most similar to the observed precipitation.

Furthermore, the authors found around 30% of ARs impact precipitation but this percentage will have spatial and temporal variability. For instance, as ARs have an interannual variability also their impact on precipitation will be significant.

Thank you for the remark. We have calculated the percentage of the total accumulated precipitation that could be related to the presence of ARs in the whole period. However, due to the interannual variability of ARs, this general percentage is supposed to change from year to year if we perform the calculation yearly. In addition, the spatial distribution may also be dependent of this interannual variability.

Finally, it will be interesting a further understand the low impact on precipitation of the ARs over Galicia, Is it less frequency of ARs, more precipitation due to cold fronts, or orographic arguments?

This was already slightly discussed in the manuscript: "In Galicia, located in the northwest region of the IP, this percentage is slightly lower owing to the greater amount of precipitation that is not associated with ARs". In fact, ARs discharge significantly more precipitation over the Galicia region than over the rest of the study domain. However, the precipitation related with other phenomena, like cold fronts, is even greater thus deriving in a lower percentage of AR-related precipitation in the area.