

Response to Reviewer 1

December 2019

1 General comment

The manuscript analyses the quality of precipitation reforecasts over the Mediterranean region performed with the Arpege model and a 10-member ensemble consisting of different convection parametrization schemes. The statistical data analysis is of very high technical quality. However the manuscript is very long, is sometimes a bit too technical, a large list of results from different well elaborated tools - one would wish to see a bit more physical interpretations in places - and one might possibly also see in addition to the observations some other reference, in particular the ERA5. Therefore I suggested major revisions, but as you will see these are not really major and should be very straightforward to do as it consist of some shortening as specified, a few sentences on physical interpretation and if possible I and likely the community would like to see on at least a few plots the results from the ERA5 (at least for forecast step 12 or so) which should shed more light on the quality/interpretation of the Arpege reforecasts (though they are somewhat penalized by initialising with a low-resolution ERA-Interim).

RESPONSE: Thanks to the referee for his general comment. First we agree with the fact that the manuscript could be improved and shortened. The revised manuscript has been shortened with more interpretation of the results. Concerning ERA5, if the suggestion is to use ERA5 as a bench-marking dataset of our 30-year reforecast with its 10-member ensemble, it is not directly feasible in our sense. Indeed our study focuses on 24h precipitation fields whereas ERA5 10-member ensemble only provides 18h forecasts for which the observation density is too low. Our intention was not precisely to emphasize the relative quality of the PEARP system in comparison with other ensemble systems. It has been already done even though on much shorter dataset. It is more about to understand the inner behaviour of the PEARP error model representation in term of predictability in the case of HPEs.

Another technical limitation to undertake such a comparison comes, as stated by the referee, from the differences in resolution of the two systems, 10 kms for PEARP and around 25 kms for ERA5. Finally, the differences in the initialization, ERA-interim with a 75kms resolution and ERA5-25kms in the case of ERA5, could also be a sensitive factor of deviation in the comparison of the two systems.

2 Major comments

- add on a few plots at least the results with the ERA5 -remove Tables 7-9 or alternatively if you prefer Tables 5-6 -remove Figures 11 and 12 and corresponding text on page 25

RESPONSE: The answer to ERA5 suggestion has been developed in the major comment answer (see previous section). We agreed with the reviewer suggestion to remove some of the tables which were too technical and the text has been simplified leading to an overall reduced length of the manuscript. Some too technical parts have also been removed.

CHANGE: In the conclusion section (extensively modified, at the suggestion of referee number 2), a statement concerning ERA5 is added: “The model physics perturbation technique should then play a greater role in the control of the ensemble dispersion. In this perspective, the novel reanalysis ERA5 would be interesting to use, in particular its perturbed members, to improve the uncertainty from initial conditions in the reforecast.”

Technical description of k-sample Anderson-Darling test has been removed. Lines from [18 367] to [18 368]. Eq. (9) and (10) are removed.

Tables 5 and 6 are removed. Text from [19 379] to [20 394] is replaced by:

“The tests are performed for the comparison of each pair of PDFs combined from the four deep convection families and from the three clusters classification. For the A-component, PCMT physics distributions depart significantly from B85 schemes at all lead times, while B85 mod and CAPE perform as B85, meaning that the modified versions of B85 weakly affect physics behaviour (not shown).

With respect to the S-component distributions, k -sample AD tests show significant differences between B85 and PCMT physics for LT12, but not for the longest lead times (not shown). At LT34, we observe a convergence of the physics scheme towards a homogeneous distribution, meaning that the differences between physics are negligible.

The test applied to the location component does not reveal significant differences between the PDFs. We suppose that the limited dimensions of the domain employed in this study, as well as its irregular shape, may lead to a less coherent estimation of the location, resulting in a degradation of the score significance. Since the L-component result is not informative about HPEs, it is ignored hereafter.”

In order to simplify the text and facilitate the interpretation of results, Tables 7,8,9, and 10 have been replaced by a new figure attached to the document. Text from [21 410] to [22 424] is replaced by:

“The results of the error diagnostic *err* for the the A-component are shown in Fig. (new figure)(a). Errors increase with lead time. We note

that the negative errors are always at least twice as large as the positive ones. Forecasted averaged rainfall amounts are almost always underestimated. PCMT produces overall better A-component statistics, except for cluster 3 at LT34. It is interesting to observe that the weakest errors are associated with cluster 3, which is the most extreme one. Since cluster 3 collects a large number of precipitation events impacting the Cévennes chain, we may suppose that the domain averaged rainfall amounts are more predictable in situations of precipitation driven by the orography. Concerning the S-component evaluation (see Fig. (new figure)(b)), structures of rainfall patterns are better forecasted for heavy rainfall events (clusters 3 and 5), than for the remaining classes of events. In contrast to the A-component, the S-component exhibits the highest err_+ for B85 scheme for most of the cases (majority of + sign in Fig. (new figure)(b)), whereas this trend is not systematic for PCMT physics. PCMT globally performs better than B85, except for cluster 2. As with the amplitude A, the S-component gets worse for longer lead times, resulting in a shift to larger err_- for both B85 and PCMT physics (more - sign for LT34 in Fig. (new figure)(a) and (b)). The lowest values of S-component are achieved for cluster 5. Cluster 5 HPEs are known to have specific regional properties whose influence on S-component results should be studied with further diagnostics.”

The caption of the new figure is: “Dumbbell plot of integrated error diagnostics computed using eq. 13. Colours refer to B85 (orange) and PCMT (blue) deep convection parametrization schemes. Results are stratified on the basis of the clusters and lead times. Symbols denote whether positive or negative errors dominate. These signs are defined using the following definition: $-$ (bold) if $\frac{err_-}{err_+} \geq 2$; $-$ if $1.1 \leq \frac{err_-}{err_+} < 2$; $\approx$ if $0.9 < \frac{err_-}{err_+} < 1.1$; $+$ if $0.5 < \frac{err_-}{err_+} \leq 0.9$; $+$ (bold) if $\frac{err_-}{err_+} \leq 0.5$.”

We agree also with the reviewer that Fig. 12 and its corresponding comments were not necessary for the results analysis whole comprehension and hence they have been removed.

3 Minor comments

- Abstract l4 ”flash-flood”→”flash floods” l6 ”hindcast”→”hindcasts” -l27 :”a quasi-stationary synoptic system to slow the convective system”, odd sentence, the convective systems are not ”slowed” by the convective system

RESPONSE: [Thanks to the referee. The text has been corrected](#)

CHANGE: “a quasi-stationary convective system that persists over the threat area”

- l45 ”weather warning triggering”→”triggering of weather warnings”

RESPONSE: [The text has be corrected](#)

CHANGE: As suggested by the reviewer

- -186 "forecasts on the basis of the region of the domain ..." ????? rewrite whole sentence

RESPONSE: Thanks to the referee for suggesting a reformulation of the sentence. The text has be corrected

CHANGE: "In particular we focus on the quality of the spatial forecasts corresponding to days of intense precipitation over the French Mediterranean region"

- -191 "In detail, section 2.1 .." → "section 2.1 .."

RESPONSE: The text has be corrected

CHANGE: as suggested by the reviewer

- -195 ", and furtherly based on ..." delete

RESPONSE: The text has be corrected

CHANGE: As suggested by the reviewer

- "In ECUME oceanic fluxes are maximized"? what do you mean by that

RESPONSE: thanks to the referee for requiring a clarification. The sentence is modified to clarify the discrepancy between *ECUME* and *ECUME_{mod}*

CHANGE: "In *ECUME_{mod}* evaporation fluxes above sea surfaces are enhanced

- 1135 "realized" → "provided" -1232 "both in" delete -page 12 caption Figure 4 "the the" delete "the" -1312 "higher lead time ranges" → "longer lead times" -1349 "at first lead times" → "at short lead times"

RESPONSE: Thanks to the referee for suggesting some corrections. The text has be corrected

CHANGE: As suggested by the reviewer

- -1356 "while CAPE shows a behaviour different from". Can you explain why and add to the text?

RESPONSE: Thanks to the referee for suggesting a specification about the interpretation of the behaviour of CAPE. The distribution of CAPE physics for L-component actually exhibits a different behaviour compared to the other ones. This effect may be related to the closure used in the deep convection parametrization scheme. CAPE implements the same scheme as B85, but it uses a closure based on CAPE. This discrepancy affects the spatial distribution of precipitation.

CHANGE: "The use of a closure based on CAPE, rather than on the convergence of humidity seems to modulate the location of precipitation produced by this deep convection parametrization scheme. Moreover, at LT34 CAPE is characterized by a lower number of strong location errors, compared to the other physics."

- -1375 "standardized" → "normalized" -1503 "flat and large objects" → "flat and larger"
 RESPONSE: Thanks to the referee for suggesting some corrections. The text has be corrected
 CHANGE: As suggested by the reviewer
- -1504-5 please rewrite this sentence
 RESPONSE: Thanks to the referee for this comment. The sentence is not clear, indeed. The text has be corrected
 CHANGE: "As in grid-point rainfall verification, all the SAL components get worse as a function of lead time"
- 506 "negative-skewed" → "negatively-skewed"
 RESPONSE: The text has be corrected
 CHANGE: As suggested by the reviewer
- -1508 " .. component behaviour .. to the clusters" rewrite this sentence
 REPONSE: Thanks to the referee for this comment. The sentence is not clear, indeed. The text has be corrected
 CHANGE: "Concerning the regional disparities in model behaviour, the SAL diagnostics divided according to the clusters show interesting results"
- -1520-524 delete as with material Figures 11,12
 RESPONSE: Thanks to the referee for suggesting a shortening of this article. As already stated in response of the major comment, we deleted Fig. 12 and the corresponding text
 CHANGE: Remove Fig. 11. Remove text from [25 448] to [25 463]. Remove text [25 465]: "even when its scaled object mass R^* is close to O_2 "
- -1525 "On another hand" → "Furthermore," -1515 "ranges higher" → "ranges longer"
 RESPONSE: The text has be corrected
 CHANGE: As suggested by the reviewer

References