EVALUATION SUMMARY: I would like to thank to the authors for their hard work to update their paper. Generally, I am happy with the answers to my questions, but I think there are still two issues, which needs some more attention:

1. *There is no particular analysis why BARRA is expected to provide additional information with respect to its global counterparts and with respect to dynamical downscaling. Only generalities are mentioned here and I would be interested in the particular aspects of BARRA at that regard.*
2. *The answer for the grid-point storm question is not satisfactory in my opinion. Please don't try to convince the readers that it is normal in an NWP model to have grid-point storms. This is a numerical problem, which must be avoided. It is coming from the fact that the convection scheme (and the parameterisation schemes in general) is not suited to the resolution used for the reanalysis. Normally the convection scheme should have adapted to have the proper match!*

So overall, I am still not fully happy with the paper, but due to the very positive attitude of the other three reviewers I don't want to block the paper from publication. Therefore, I suggest, to accept the paper provided a proper thought is given to these two issues. See below my original comments (in italics) the answer of the authors and my latest feedback (red italics).

The main question for a regional reanalysis is to clearly demonstrate whether the use of such system is justified, which means that more value can be added to the global reanalysis then it would be the case with a pure dynamical downscaling. For this question one has to understand the additional information brought into the regional system in terms of more precise dynamical and physical description of the atmosphere, but also in terms of additional and advanced use of observations. I miss a summary of this kind from the manuscript though some of these aspects are highlighted here and there in the paper.

[AR] The introduction has reviewed several papers on the usefulness of regional reanalyses over dynamical downscaling, underpinning efforts around the various regional reanalysis projects internationally.

My point here was not a general assessment of the value of regional reanalysis with respect to dynamical downscaling, but a particular one which analyses the merit of BARRA in that regard.

The comparisons of short-ranged O-A (observation – analysis) and O-B (observation – background) statistics in Table 1 (now moved to Table S1 of the Supplementary Material) showed that, with O-A being consistently better O-B for various observational types, an analysis within the BARRA-R system yields a more accurate short-ranged forecast than simply using the background, where the background from the previous analysis is, by extension, better than pure dynamical downscaling from the very first cycle.

I think, the fact that O-A is better than O-B does not show that the reanalysis is better than dynamical downscaling (it was also admitted by the authors answering to another question of my original review). In case of dynamical downscaling there is a higher resolution dynamics and physics of the model and the surface characteristics are described in more details (I mean on dynamical downscaling that a model is used to downscale the lower resolution information possibly also taking into account a better surface description).

2.3. The existence of the "grid-point storms" is embarrassing since such numerical problems should not happen in a reanalysis, where a robust and properly (thoroughly) tested NWP system should be used. Normally, the reanalysis should not be run if such problems are not yet solved. There is a need for a thorough explanation how this could happen and how this deficiency compromises the validity of the reanalysis results.

[AR] The Unified Model is sufficiently robust to be useful for many operational meteorological centres in Australia, UK, India, Singapore, Korea, South Africa and New Zealand. The issue of "grid- point storms" is also not unique to UM but for instance, also occurs in the widely-used Weather Research and Forecasting (WRF) model from NCAR. When the convective (sub-grid) parameterization scheme in non-convective resolving models does not stabilize the air column, meteorological events can develop at the smallest resolvable scales in the model, producing unrealistically strong vertical velocities and precipitation (Scinocca and McFarlane, 2004; Williamson, 2013). The resulting "grid-point storms" occur more readily in models with higher horizontal resolutions (Williamson, 2013). The issue becomes unavoidable for BARRA-R as it aims to be sufficiently higher resolution than global reanalyses but could not be sufficiently high resolution (< 2 km) (and computationally prohibitive) to resolve convection explicitly without the need for a convective parameterization scheme.

Further, we do not think that the wet biases in BARRA-R over the tropics and New Zealand are entirely due to grid point storms. Additional analyses have been made to identify the location of

precipitation excess in the tropics and New Zealand (Figure S3 in Supplementary Material). We found that the higher precipitation in BARRA-R are concentrated at high or sharp topographical regions in PNG, Indonesia, Sumatra and small Indonesian Islands, and west coast and Southern Alps of New Zealand. At these locations, GPCC (gauge analysis) and TMPA would underestimate the precipitation. With these considerations, the actual levels of bias observed in BARRA-R are not entirely clear.

I am still NOT convinced at all that the grid-point storms are unavoidable details of a numerical model. These are really numerical artefacts, which should be avoided! Regarding the answer of the authors:

- *It is not an argument that other models (e.g. WRF) and other centres (UK, India, Singapore, Korea, South Africa, New Zealand) have the same problem. This is not an answer to the question!*
- *As the authors properly mention this problem is coming from the discrepancy between the convection scheme and the non-convective resolving model. It is well-known that in the so called grey resolution zone (typically around 3-7km resolution range) adequate convection scheme should be used. The occurrence of the grid-point storms indicate that the applied convection scheme is not suited to that resolution!*

I think the only way to circumvent this issue in the article is (i) admit this problem (which is already the case in the manuscript), (ii) properly explain its origin, (iii) warn the users particularly if they would like to have a local evaluation and (iv) convince the readers/users that this problem does not have a significant impact on the climate quality of the reanalysis. But, please don't use such arguments that it is also apparent in other models and centres!!

2.8. page 2, line 25: please give reference for the Copernicus reanalysis

[AR] Agreed. I have added a reference to Ridal et al. (2017).

Ridal et al (2017) is a reference to UERRA and not to Copernicus reanalysis (ERA5). Use for instance Hersbach and Dee (2016).

2.35. page 16, line13-14: it is important to get an overview in this paper about the relative merits between reanalysis and downscaling, since this gives justification for having reanalysis instead of simple downscaling. Therefore, some information about this issue should be provided at an early part of this paper.

[AR] The introduction has reviewed several papers on the usefulness of regional reanalyses over dynamical downscaling, underpinning efforts around the various regional reanalysis projects internationally.

Again, I mean this particularly for BARRA and not in general!

The comparisons of short-ranged O-A (observation – analysis) and O-B (observation – background) statistics in Table 1 (now moved to Table S1 of the Supplementary Material) showed that, with O-A being consistently better O-B for various observational types, an analysis within the BARRA-R system yields a more accurate short-ranged forecast than simply using the background, where the background from the previous analysis is, by extension, better than pure dynamical downscaling from the very first cycle.

See my feedback above for the same issue!

Two small additional issues: please use ERA5 without hyphen and I think one has to use short-range instead of short-ranged.