The reviewed manuscript corresponds to the authors response to the mayor revisions of the article: “Analysis of the streamflow extremes and long term water balance in Liguria Region of Italy using a cloud permitting grid spacing reanalysis dataset” by Silvestro et al.

General comments

- The authors present an approach (modelling chain) to use a long-term and high spatial resolution (4km) regional climate model with further bias correction to force a hydrological model over a large region in Italy. Calibration of the hydrological model was performed when possible and then parameters were transferred to ungauged basins. On gauged basins the model seems to represent relatively well streamflow; however, I see a significant limitation in the approach used to transfer these parameters to ungauged basins, as these are purely empirical, and the average of the calibrated parameters were used in ungauged. A GEV distribution was fitted to the simulated Annual Discharge Maxima (ADM) at all gauged sites and compared with observed streamflow, for which I see a fair agreement as fitted GEV 95% confidence interval sometimes does not cover observations (3 out of 6 in my opinion). I think the authors underestimate the role of the hydrological model in the overall analysis and discussion through the manuscript; more emphasis should be given to it, as this is largely driving the simulated hydrological response of these basins. The Ratio(t) is defined to assess the performance of a regionalization approach in characterizing streamflow for several return periods (T). Honestly I have somewhat a hard time understanding the usefulness of this regionalization approach, as it does not seems to work very well. I encourage the authors to demonstrate that the Ratio(T) is being well represented, as I see problems in small and large scale basins (Figure 11). The analysis of the effect of downscaling precipitation on the streamflow should be revised (see specific comments). The final analysis of the mass balance components (runoff ratio) seems very useful and maybe it should be more deeply explored and the main focus of the manuscript.
- I acknowledge that the authors responded to most of the comments from the previous revision; however, there are still significant issues with the manuscript in my opinion.
- For latter revisions I would encourage adding more details to the author’s response letter, specifically where the changes are located in the manuscript (page and line), that way it would be much easier for us to go through the revisions.
- I believe that the English of the manuscript needs to be further improved.

Specific Comments:

1 Introduction:

- Page 4, Line 22-23: I agree in the use of high-resolution RCM to reproduce small-scale rainfall events, but I think you need a reference to this or even better if you can show that this is the case with the original dataset in your study region.

2 Material and Methods:

Study area and case study:

- Figure 1: To help the description of the figure, use (a), (b) and (c) to refer to each of the plot. And explain each of them as an inset of the others. I'm not sure about the utility of including the Curve Number, if you want to keep this you need to reference the source of this information. Also for the legend use "Curve Number" instead of C.N., try to avoid acronyms, same for DEM, etc.
- Page 5, line 12: refer to table 1.
- Page 6, line 5: What about the raingage distribution vs elevation, is orographic precipitation well represented by these stations? Discuss.
- Page 6, line 5-7: be more specific, how low?
- Page 6, line 7-8: what data was used for calibration/validation?, I assume streamflow discharge, but it is unclear.

Downscaling the rainfall

- Pag 10, line 10: Define Lr and tr**.**
- Pag 10, line 10-12: add reference.

The hydrological model

- This section needs better organization, use subsections if needed to explain: model inputs, calibration/validation and model structure (parameters, spatial resolution, etc.)
- How is evapotranspiration being calculated?
- pag 13, line 15-17: Calibration procedure is not very clear. What do you mean by 11 sections, what's a "section" (basin?). What ground stations measurements were interpolated? I assume you calibrate and validate the model using the downscalled reanalysis data, please clarify.
- At what temporal scale did you calculated NS and REHF, hourly or daily?
- What range did you use for calibration and why?, explain. How did you come up with the 500 mm and 1 (Vxmax and Rf) for all the basins, wild guess?
- Pag. 14, line 16-17: This point is crucial in the manuscript. I disagree that taking average parameters values from calibration into ungauged basins is a satisfactory approach, especially as this are highly empirical values, and this is critical to analyze the results. I would encourage the authors to demonstrate that the impact of taking average parameters in ungauged basins is minimal. For this purpose you can use the average parameter values in the model of a gauged basin and demonstrate that these parameters have little impact on the streamflow performance. I think a significant assumption is being made here, which is not easy to support given the empirical basis of the model.
- For the regional analysis it is necessary to know how much of your study region is covered by gauged and ungauged basins, this can help discussing the impact of the assumptions in model's parameters.
- Table 5: need to include the periods available for discharge, and ADM.

3 Results

- Figure 2 and 3: I would recommend showing the difference in precipitation as a percentage with respect to the observed precipitation, or as (mm), so it is easier to compare with the other maps which are in mm.
- Pag 17, line 17: not sure what you mean by "are largely confirmed"?. Please change the units of the difference map so the comparison is easier. And do the same when you analyze this difference.
- Figure 4: Include coefficient of determination of correlation and mean bias.
- Figure 5 and 6: Values presented in this figures are good for the analysis; however, I think you already have enough figures regarding the performance of EXPRESS-Hydro (Fig 2 to 6) I would try to merge or remove some, maybe by adding more information to Table 3 you can remove one or two plots without losing much information for the analysis.
- As ADM are a key aspect of your analysis, I would also show the performance of the bias-corrected precipitation time-series in representing peak precipitation, you could look at this by comparing the probability distribution functions of simulated and observed precipitation. I think this is very important as your peak flows are driven by the peak precipitation events, not by the precipitation volume. This will help you in the discussion of the ADM model's performance.
- Pag 18, line 21-24: why did you chose these 4 basins and nor others?
- You should try to merge figure 7, 8 and 9.
- Why did you choose these 4 basins to contrast the analysis?
- Explain what is that you are testing with Kolmorogov-smirnov test, and this should be in your methods too. I know you are talking about ADM distribution but it should be clear from in your manuscript.

Section 3.3

- Figure 10: are you using data from every grid point in your model to construct this curve or selected basin outlet?, this need to be clarified in the text. Avoid shortening the words in the figure; you have enough room in the plot, change 'Calib' to calibration, 'simul' to simulations, etc. This applies to all plots/tables.
- Can you comment or interpret the step-changes in the observed growth curve of figure 10 in terms of the dominating hydrological processes in the region, and why you don't see that in the simulated values?
- It seems to me that the relatively better performance in the regional growth curve, both calibrated and total area, could be the results of errors compensation as figure 7 to 9 show that observed distribution of ADM lies outside of the confidence boundary in some cases (Bisagno, Magra, Argentina; 3 of 6). This needs more discussion in the manuscript. If this approach is meant to be used in ungauged basins, the model needs to prove that it works in the calibrated basins. This can be a major problem in the proposed modelling chain. If the authors can prove that the model in calibrated basins can represent ADM, this will help a lot in supporting the proposed modelling chain.

- Figure 11 shows that small basins are underestimated (as discussed in the manuscript), and it also shows an overestimation for the Ratio(T) for large basins, how can the authors explain this?. Also I'm not convinced that the B.C. Ratio(T) shows an improvement in Figure(11), as small basins that used to have a good performance (near 1) now they are overestimating (Ratio>1). Overall I'm doubtful about the usefulness of this Ratio(T) as it does not seem to produce good results. I think the authors need to show that this index, which tries to show that the regional curve is suitable for all basins, works.
- Figure 12 has little discussion in the manuscript and the comparison that the authors do (line 9-11, page 22) is hard to see from the figure.
- Page 21, line 16-18: Then why not excluding these from the analysis?

Section 3.4

This analysis should be changed. In order to fairly compare the effect of downscaling in the mean annual streamflow, which is what I assumed you are comparing (clarify), you should show results from the hydrological model calibrated using the non-downscaled precipitation (unless you did so, but it is not clear from the text) versus the calibrated hydrological model using downscaled precipitation. If you want to show this analysis you should also include the effect of downscaling on Nash-Sutcliffe, REHF and bias against observations.