Dear reviewers and dear Editor,

We uploaded a new version of the paper. We tried to insert most of the requests arose during the reviewing process that, in our opinion, helped to improve the paper. Various modifications were done along all the manuscript but we report the main structure changes:

- "Material and methods" chapter was lengthened since we move there the description of the methods as requested
- More details about hydrological model, study area and characteristics of the basins were added (also new tables and new versions of the figures are now present)
- Graphs and tables regarding precipitation analysis were added to deepen the analysis as requested
- Some figures were merged as requested
- The discussion of results of hydrological analysis is now extended, we inserted comments and discussion arose during the review process together with new version of some figures.
- A new section to show how rainfall downscaling affects simulations/results was added
- We revised the language of all the sections; if editor and reviewers agree that the paper is now improved regarding the scientific point of view but it still needs improvement regarding the language, we are open to submit the manuscript to a language revision service.

We hope the paper is now suitable for publication.

Answers to the reviewer comments are already available on web but we report them in the following and we briefly described how we modified the manuscript.

The authors.

Reviewer1

Introduction

- Weak, include more about issues with peakflow estimation and uncertainties, previous studies in the region, hydrological modelling, downscaling and bias correction. Much of the text in other sections should be moved to the introduction (detailed below)
- Page3, line 13-16: I would move it to the methodology section.

We improved the introduction following the reviewer suggestions, also including more material about previous studies in the region

Study Site and Case Study

Need a better description of the study basins: how many are they and where? Include geomorphological characterization (area, slope, etc), vegetation cover and soil description; are these rivers regulated by dams or do they have extraction, hydrological regime? How is the climatology of this region(precipitation: snow and rain, and temperature?), you should give more details about this. If there are other studies in these basins they should be included and mentioned.

We expanded the description of the study area and its climatology in the new version of the manuscript
We added more information in Table 1 for all the points/sections where discharge data were available.

- Page 4, line 11-12: "The response time ..." add Reference.
Added the new text (Maidment, 1992; Giannoni et al., 2005)

- Figure1: Add to the figure a North Arrow, scale bar, color bar for topography and its source. I would zoom in to the area of interest and add an inset plot showing the regional location of the study site within Europe.
We modified the figure as requested in the new version

- How many stations from the OMIRL network do you use in the study? What about their quality control, any issues that should be mentioned in this regard? Any snowfall measurements?

- Page 4, line 22: Mention about "historical validated data", what do you mean by that, validated against what exactly? Is there a reference showing the validation?
- Need a better description of the driving data used in the study. Number of stations, a better map showing their location with respect to the basins, available period, gaps in the data if any and variables being recorded (more details).
- Similar for the annual maxima time series, need to include an official identifier number (if available) or name of the hydrometric stations. How reliable are these records? This is particularly important when you are assessing peakflow.
We collected the available information and improve the description in the text, adding some tables

EXPRESS-Hydro reanalysis

- Page 5, line 6-20: Rewrite, move to introduction and add references, this is not part of the methodology. Here you should only describe the EXPRESS-Hydro reanalysis.
These changes will be done in the new version.

- Avoid the use of superlative like "very high", instead state the resolution.
Agreed, we will replace in the text with "4 km grid spacing"

- Remove Figure 2, not relevant.
Agreed, figure was removed.

- Add reference for ERA-Interim and WRF.
Added:
- Page 6, line 1-17: This is more of a discussion about Pieri et al, I would remove it or synthesize as part of the introduction section. This section should be about the reanalysis' technical features, pros and cons and its suitability for the area of study. What variables are you using from EXPRESS-Hydro? Temporal and spatial resolution? Available period? Etc.
Agreed, this section was shortened in the new version of the Manuscript, meanwhile it is worth to clarify that Continuum uses from EXPRESS-Hydro the following variables (at 4 km grid spacing and 3 hours time resolution): 2m temperature, 10 m wind, rain depth, downward short wave flux at ground surface, 2m air relative humidity; this is partially done in a previous answer (first comment of this section)

Bias correction of rainfall fields (B.C.)

Replace "Pieri et al (2015) reanalysis" by "EXPRESS-Hydro reanalysis" (everywhere in the text)

The expression was replaced in the text according to the reviewer request.

Page 6, line 22-23: no need to include the link for the dataset again.

The link was removed from the text.

Why don't you use the observed dataset as opposed to the corrected EXPRESS-Hydro reanalysis?

A sentence was added in Section 2.1 in order to specify that. Given the fact that the extreme events that occur in the region of study are due to short-duration precipitation events, it was not possible to employ directly these data from the Atlante Climatico dataset, so they were used only for the bias-correction of the EXPRESS-Hydro dataset.

You should not give additional information about EXPRESS-Hydro or observed dataset here; those should be listed in their respective sections.

The information about the meteorological datasets was moved in the previous Sections.

Page 8, line 17-18: that's part of the results section

This sentence was intended as a feature-by-design of the particular bias-correction procedure that was employed. The sentence was modified in order to be more clear.

Did you perform any correction to other weather variables used in the hydrological model, such as air temperature?

A sentence was added in the text in order to clarify this.

Is there any snowfall in the study basins, was this bias-correction also applied to snowfall? Please clarify

The snowfall in the region of study is present, but not in significant quantity. The contribution of snowfall to the extreme flood events can be considered basically negligible. Furthermore, no snowfall gauges were available in the observed dataset, so the bias-correction was applied on the precipitation.

Is there a particular reason you performed the CDF correction at a monthly scale, how different are the precipitation patterns between months? Maybe with an annual or seasonal correction is enough?

As explained in Section 2.2, also with specific reference to Figures numbered as 3-7, strong seasonality is present in both observed (as confirmed by the Atlante Climatico Liguria) and EXPRESS-Hydro datasets. The differences are not negligible also in terms of the single months patterns, thus an annual or seasonal correction would not have been sufficient. A specific sentence about the monthly spatial patterns was added in Section 2.2 in order to specify this.

You need to better argue the use of rainfall station interpolation to bias-correct the reanalysis, spatial interpolation can introduce significant errors in you rainfall estimations, particularly if your monitoring network in sparse, the rainfall regime is very heterogonous or has significant local topographical controls. If any of these apply, then I'm not sure how robust the spatial interpolation is, which can lead to significant problems in the hydrological model performance. You said that regression with other variable, such as elevation, were tested did not show significant correlation, but is your monitoring network dense enough to argue that? In other words, how is the elevation distribution of your monitoring network? If most of them are at lower elevations then no correlation should be expected.

The interpolation was employed only at monthly scale, in order to correct the monthly cumulate of the EXPRESS-Hydro dataset. Given the fact that the events of interest are of short duration and spatially concentrated, the error introduced by the interpolation should not affect the structure of such events. The possible distortion introduced by the interpolation can be considered acceptable considering that the raingauge network used for the interpolation itself is sufficiently dense and spatially homogeneous to avoid large areas without sensors. Furthermore, the distribution of the sensors along the different elevations is uniform enough to assure the significance of the correlation analysis. A sentence was added to the text in order to clarify the interpolation issues.

Precipitation analysis
- Page 13, line 13-20: Apparently, these are not your results should not be in the result section
In this part of the text we compare our analysis with the one done by Pieri et al..We can better clarified this point and even reduce these comments

- Figure 3 to 7: are these your results or from Pieri et al?. If there are not yours you should remove and just reference them. Otherwise you can combine them and only show the

difference map for annual and seasonal scales. Note that the mean daily errors are between -3 and 3 mm/d, which for annual rainfall mean an average error of +-1095 mm/yr, which is quite significant and will have a significant impact in your simulations, this should be reconsidered.

These figures are part of our work, we will better clarified this point in the text. As shown in some part of the region errors are quite large. For this reason we applied a rainfall bias correction. Still it is worth to mention that errors are definitely lower than those corresponding to coarser grid spacing dynamically downscaled (e.g. 12 km) reanalysis.

- Is Figure 3 using the Bias-corrected EXPRESS-Hydro reanalysis? unclear
Without B.C. We better clarified this in the caption.

Page 14, line 12-17: do this analysis at a basin-scale as that's the relevant scale of the study. I'm not convinced about that EXPRESS-Hydro "reproduces quite well" the observed precipitation. Need further analysis. How are the extreme precipitation events represented by the reanalysis? This is critical to properly represent peakflows. What about temperature?

We made the analysis on 4 tests basins along the study area just to give an idea of the possible variability of the results when going to basin scale.
Concerning the other questions, Pieri et al. (2015) assessed the EXPRESS-HYDRO results capability to correctly reproduce the statistics of intense precipitation events both over Europe CORDEX domain at large and over the Great Alpine Region (GAR). The following figure shows the map of the number of days with rainfall rate larger than 10 mm for E-OBS and for EXPRESS-HYDRO runs. According to E-OBS, frequent heavy precipitation days are essentially localized over the Alps, western Norway, Scotland, and Portugal. This spatial distribution is well captured by EXPRESS-HYDRO runs.
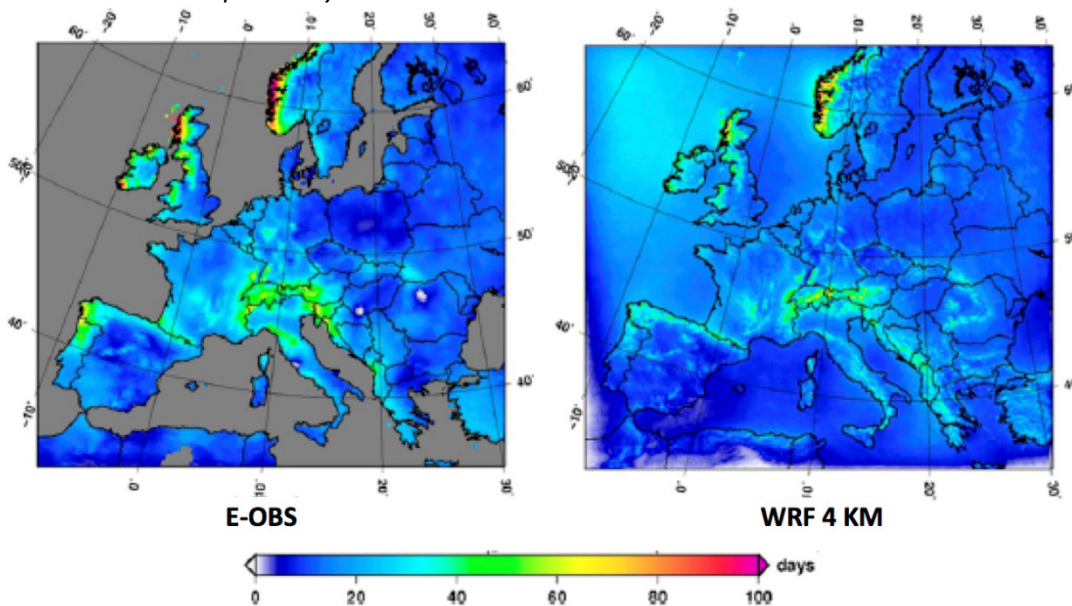

Figure 1: Number of days with heavy precipitation (>10 mm) for the period 1979–98 (modified from Pieri et al. 2015).

Pieri et al. (2015) investigated the distribution of the probabilities of exceedance of precipitation thresholds for the daily rainfall rate on the GAR domain, for all seasons, for the 12 and 4 km grid spacing model results and for EURO4M-APGD observational dataset. The high-resolution run at its original 4 km grid spacing agrees well with observations in most seasons but underestimates the probabilities of intense precipitation (at about > 100mm/day) when aggregated at 12 km, except in winter when the aggregated high-resolution run is close to the observations.
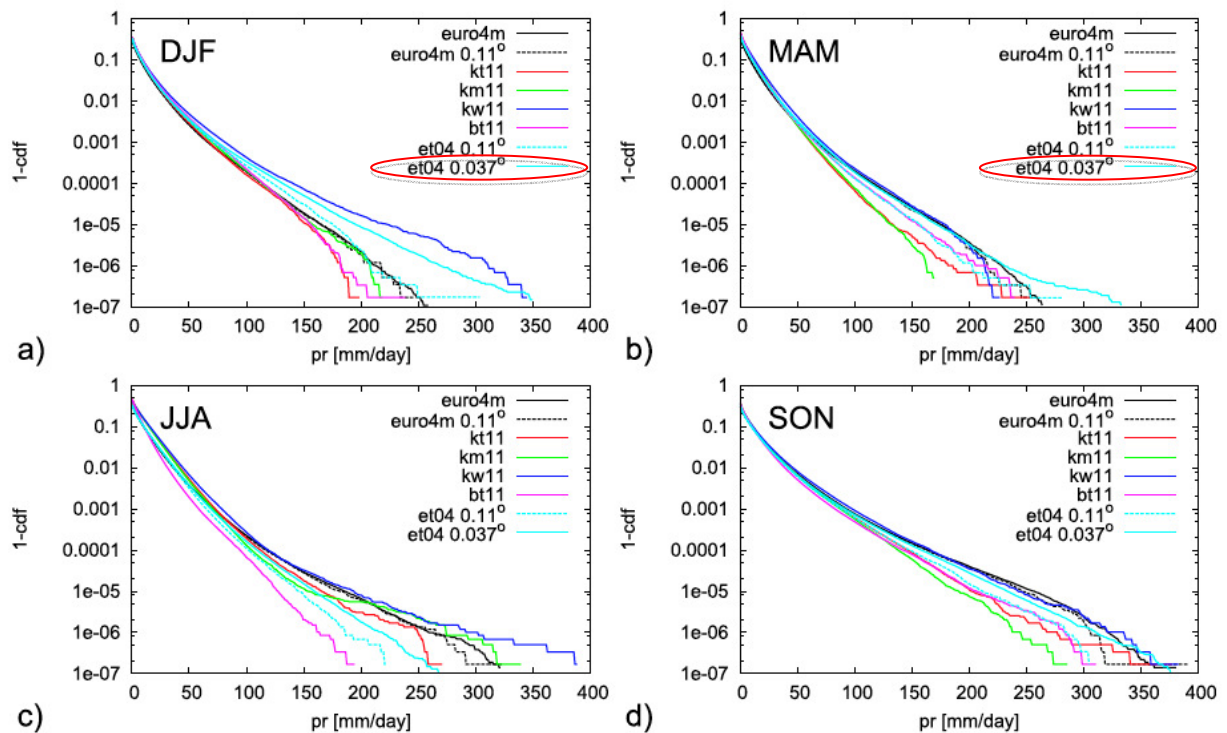
FIG. 7. Exceedance probability for 1979–98 over the GAR, for different WRF simulations and for EURO4M-APGD, for seasons (a) DJF, (b) March–May (MAM), (c) JJA, and (d) September–November (SON). The exceedance probability is computed as 1 minus the precipitation cdf.

Can you evaluate the performance of the spatio-temporal downscaling?, this seems rather critical when assessing peakflow that may be generated from intense hourly rainfall-runoff events.

The testing and application of RainFARM in the study area was done in many works, some already cited in the text (Rebora et al. 2006; Silvestro et al. 2012). As stated by reviewer this is a potentially critical point and we discussed the effects on the very small basins in the text and in the following answers to the comments. We will also improve comments on text.

*To assess the impact of downscaling we made the hydrological simulation with BC rainfall but **without** applying the downscaling, we estimated the ratio of Qmean (mean of 30 years ADM) without and with Downscaling for each pixel.*

Results are plot versus drainage area (see figure below). It is quite evident the impact of downscaling, the impact generally increases when drainage area decreases. This graph also help to answer to one of the comment below regarding the underestimation of quantiles for very small basins: the downscaling is crucial to face this issue but probably the adopted configuration is not sufficient when drainage area drops below certain values.

The hydrological model: Continuum

- Page 10, line 11-12: Reference
We removed this sentence.

- What's the soil surface temperature used for in the model?
As mentioned in the text (page 10, line 7) LST is an output of the model

- How is evapotranspiration being simulated by the model? Does the model require air temperature?
Yes, in the revised text we added the following sentence:" Continuum needs as input the following variables: rainfall, air temperature, short-wave solar radiation, wind velocity, air relative humidity."

- Does the model simulate snowmelt and accumulation?
The model simulates snow dynamics, we added a sentence in paragraph 2.5 : "Snow melting and accumulation is simulated with simple equations forced with air temperature and solar radiation(Maidment, 1992) as described in Silvestro et al. (2015)."

- Include the units of the parameters.
Done in the new version

- This section needs clarification regarding the parameterization approach. There are 6 parameters requiring calibration, did you use the same values for all your basins and landcover? How was the calibration performed, automatically or manually? What was the period used for calibration? Did you calibrate using hourly streamflow measurements? What are the final parameter values for the calibrated and uncalibrated parameters?
We added some more information and a new table (Table 2) with final parameter values as requested. "NS and REHF were combined in a multi-objective function to carry out the calibration using the approach proposed in Madsen (2000). Calibration was done analyzing the parameter space using a brute force approach on 2011-2013 period in order to find the parameter set that optimize the multi-objective function. Curve Number map is derived by the CORINE Land cover (http://www.sinanet.isprambiente.it/it/progetti/corine-land-cover-1).The final setting is similar to the one described in Davolio et al. (2017)."

- Can you describe the meaning of the REHF index
We added the following sentence in section 2.5: "While NS aims to assess the general reproduction of streamflow, REHF score has the aim of giving high weight to high flows leading the calibration to better reproduce the flood events."

- Page 12, line 18-19: basins without data for calibration/validation should be removed from the analysis. The empirical nature of the model does not support any parameter transferability; therefore, assuming average values from calibration at other basins should not be used.
We know that this is not the optimal condition of work, it is clear that it would be better to have a very large number of calibrated basin, but this is also quite common in real application.
*We only partially gree with this comment for at least two reasons: i)*the Continuum hydrological model (and its predecessor DRiFt, that was designed only for simulating floods at event scale) demonstrated to well work in the study environment *(Giannoni et al., 2000, 2005; Gabellani et al., 2008; Silvestro et al 2013, 2015, Cenci et al 2016…) giving reasonable and reliable results also for uncalibrated sections when used for flood forecast aims (Regione Marche, 2016). This is valid when the basins have similar characteristics especially regarding the surface response to intense rainfall events and main genesis of rainfall-runoff process, as a consequence parameters have often similar values moving from a catchment to another. We agree with reviewer that if we would use Ligurian average parameters in other environments, with completely different characteristics (example flat basins,) results could be affected. ii)Even the benchmark regional analysis (Boni et al., 2007) is done with the model DRiFt which was not possible to calibrate in each place of the target region. The hydro model is used indeed to have flow time series in un-gauged basins*
We stressed the problem of calibration in the text but we believe it would be interesting to keep the analysis over the entire region.

- Page 13, line 1-3: clarify, did you run the model for the entire period and used the final conditions from year 2008 as initial conditions for the year 1979?
No, we modified the sentence in order to better clarify:" In order to have reduced warm up impacts on 1979 simulation a first run was done starting from a predefined initial condition and the state variables simulated on the 31st of December of every year (from 1979 to 2008) were used averaged to estimate a reasonable initial condition for 1st January 1979 to be used in the final simulation.

- What meteorological variables does the model use?

We added the following sentence in section 2.5: "Continuum needs as input the following variables: rainfall, air temperature, short-wave solar radiation, wind velocity, air relative humidity."

Distribution of the annual discharge maxima
- I think that Fig. 9 to 11 show that model representation of peakflow is somewhat weak, and often the simulations without B.C. show better results (see fig 9 Bisagno La Presa), which is what the Kolmogorov-Smirnov test show as well (only 60% pass the test). This problem can be due to the problems in representing annual rainfall (fig 3-7) and hydrological model!.
We added comments regarding the results; yes, as already mentioned in the text, the fact that results are not optimal in some sections are probably due to a combination of different causes: i) sometimes and somewhere the rainfall reanalysis is probably poorly representative of real rainfall and B.C. does not correct it enough ii) hydrological model setting could be not optimal in some basins iii) observed peak simulated and peaks are often referred to different time periods, this fact together to the hydro-climatic regime typical of the study region (flash flood regime with high variability of ADM) could have a certain impact on final results.
The study is not aiming to show that the applied chain reproduces perfectly observations and benchmark (Boni et al., 2007 for regiona analysis) but the overarching goal is to combine, for the first time of the best of authors knowledge in a complex orography area such as Liguria, a cloud-permitting grid spacing reanalysis together with a state of the art hydrological modelto understand what is possible to do with such kind of chain, and thus offering a methodological framework for future similar analysis.
In any case we still believe that the results are interesting and not so weak even in respect to the hydrometerological characteristics of the study region (frequent occurrence of flash floods).

- Figure 12 should be replaced by a table.
Agreed, done in the new version of the manuscript

- Page 17, line 1-23: This is methodology not results, move to the methodology section.
*Agreed, done in the new version of the manuscript*

- I think you need to show that the model works well representing ADM at a basin-scale to perform the regional analysis. So far, the simulations at only 9 basins pass the statistical test (out of 15 from fig 12), but then the regional analysis is performed using all of them? Are those basins without streamflow records also included? (they should not). The relatively good agreement from the regional frequency distribution function analysis (fig 13) could be due to a compensatory effect.

We agree only partially with this approach (in brief: not considering streamflow series in un-calibrated sections) and we tried to explain our point of view (which is also based on previous publications). It is, in our opinion and based on our scientific experience, indeed interesting to look at the results in a distributed perspective using modeled time series to reproduce streamflow where no observations are available (being conscious of the possible uncertainty and errors of simulated time series). This is an approach similar to the one done in a previous scientific papers, in particular the one that describe our benchmark: Boni et. al (2007). The benefit of using a large number of time series is often larger than the uncertainty introduced by not optimal parameter set, especially when ADM are normalized and used to build a regional curve.

The following figure is generated using time series from the 9 calibrated basins (Thyrrenian side) and it is compared with the one generated with all sections with drainage area > 15 km^2. In the new version of the paper we added the figure.
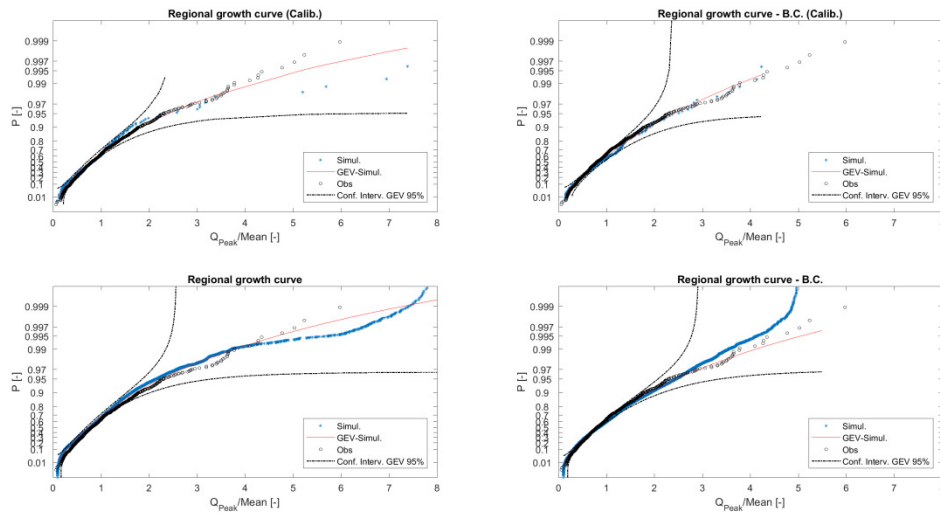
Moreover it is worth to highlight that in many cases simulated ADM distributions have similar shapes to observed ones and suffer of a sort of bias (probably driven by errors in rainfall or hydro-model), in other cases the simulated ADM distribution is only partially out of the confidence interval, but hydro-climatology seems to be discretely reproduced. This is a probable motivation because regional curve is quite well represented. For example, in the La Presa case the evident Bias is presumably lead by one of the effects presented in section 3.3, we report the text: "……it is possible that EXPRESS-Hydro well reproduces the events at small time and spatial scales (3-6 hours, 10-100 km2) in that part of the region but generally underestimates monthly cumulates, in this case B.C. could lead to streamflow overestimation….". When considering the Normalized values they contribute to good results in building regional curve.

Regional Analysis of the annual discharge maxima

- Page 17, line 1-15: This should be in the methodology section.

Agreed, we moved this part on a new paragraph (2.7) in section 2, in the new version of the manuscript.

- I suggest expanding about the usefulness of the regional ADM curve, I could see the potential when dealing with prediction in ungauged basins, otherwise I'm not sure what's the purpose of computing this regional curve. The relatively good agreement of this curve against observations showed in figure 13 could be due to some compensatory effect between basins; need to expand on this as well.

We inserted some more comments about the usefulness of the regional curve in the new version of the paper. The relatively good agreement could be also due the aforementioned motivation: even if in some basins the fit is not optimal the simulated ADM reproduce discretely the typical hydrological regime of the study region.

- The authors argue that for small scale basins (<50 km2) the simulated ADM curve (Q(T) model)underestimates the regional curve, and attribute this to problems in the reanalysis precipitation quality. I think this analysis is wrong and the fact that the Q(t)model underestimates Q(T)reg (i.e.Ratio(T) <1) only suggests that the regional analysis is not representative of small basins, which could be due to several factors not address in the discussion. (1) If the number of points used to develop the regional curve comes in majority from larger basins, then I wouldn't expect the regional model to represent small scales

basins; it is unclear how many large or small scale basins (or grid points) where use to construct the regional curve, this could help the discussion.(2) The role that the hydrological model is playing in representing peakflows is not sufficiently explored. From the Kolmgorov-Smirnov test, only 9 out of 15 basins past the test (60%), which is not sufficient in my opinion. Problems with model parameterization and process representation can probably explain most of the ADM mismatch for the scenario with bias-corrected rainfall. The representation of peak rainfall events by the reanalysis is unclear, and those are the eventsthat produce peakflows, assuming an entirely rainfall-driven streamflow, which is also unclear from the manuscript. (3) If the EXPRESS-Hydro reanalysis cannot represent the "small-scale" rainfall events, then what was the purpose behind the spatio-temporal downscaling? I think these issues should be better explored and described in the text, particularly regarding uncertainties with hydrological model.

We added some considerations based on the comments of the reviewer: they are potentially good reasons and clear explanations for underestimation in very small basins, thus we will improve the discussion of the results in this respect.

As a general comment we would like to highlight again that we are exploring the potentialities of the presented modeling chain to reproduce ADM in an environment made by small basins, in complex topography areas, especially testing the usage of a high resolution (cloud permitting) reanalysis. To the best of our knowledge this is one of the first works of this kind done in the considered study area and we have not the claim to solve all the possible issues and problems in a unique work, or to define a definitive methodology that can be applied in any case. On the other side the fact that in some cases the system cannot reproduce in optimal way observations (..or better we should say their estimation..) cannot be considered a bad result or weak point, simply the analysis shows that in some cases the system does not work optimally; in many others the results are good, 9 out of 15 basins are working well also accounting that we are working only with models and only observations on very coarse time scale are used in the system (for Bias correction).

We improved discussion and motivation but we do not believe that our analysis is totally unuseful. The model was calibrated and we have a consolidated experience in applying it in the studied environment (we already cited various works), so it appears unlikely that a possible not perfect parameterization leads to such a general underestimation on small basins (more probable a more uniform distribution between over and underestimation); anyway the model spatial resolution could play a role since the representativeness of morphology degrades for small drainage areas with general smoothing effect that affects results; we put a threshold for the analysis (basins with Area<15 km^2 are neglected) but the degradation effect is clearly continuous from larger to smaller drainage areas. But there is surely another issue to account for: the used time resolution (1 hour after downscaling) could be not sufficient in some cases for these basins, various works (example Silvestro et al., 2016; Rebora et al., 2013) and the experience prove that rainfall time pattern on scales lower than 1 hour have significant impact on high flows for such small basins. We already commented this fact in the text (pgg 18 lines 16-18 and also in section 4: Discussion and conclusions). The combination of these effects could lead to a flow underestimation (that reflects on the index flow which is estimated as average of the 30 year series of ADM) when drainage area decreases. This effect is compensated in the centre of the region where B.C. seems to lead to a general overestimation (see also comments in the text) of quantiles. Figure 7 also seems to support this motivation.

The regional curve is built using simulated time series for cells with upstream area>15 km^2 (pgg 13 lines 6-10), so it should not be affected by preponderance of large basins.

As we discussed in the text, probably the results demonstrate that very small basins should be treated with a different setting of the system and a dedicated work (even, as mentioned in the text, trying to exploiting more deeply the downscaling potentialities) , we would expand the discussion adding all these considerations. (See also response and figure to comment regarding the downscaling)

Water Balance and Runoff coefficient

I am not sure about the point of calculating what the authors refer to the "runoff coefficient". Iwould suggest looking at basin or sub-basin scale runoff ratio (runoff volume/precipitation) as aproxy to the water balance, that way you can avoid storage problems. I don't understand why ifall the relevant mass fluxes are being simulated by the model, the authors don't calculate themass balance directly. I would re-focus this section to a basin-scale mass balance analysis.

The "runoff coefficient" is exactly the runoff ratio and we changed the name in the text as suggested. We did the analysis of runoff ratio because, as mentioned in the text, we have a term of comparison from Hydrologic Annual Survey. For the Analysis at pixel scale we used the first formulation showed in section 3.4 because rain and evt are the standard output and input available for common hydrological models (also for Continuum), while other components (input and output of each cell) are not saved as output. It's clear that the mass balance at single cell scale is closed. Long term runoff ratio can be estimated in both ways (function of P and Q, or function of P and Evt). We better clarified these points in the text.

Table 2 shows model streamflow bias, this should be part of a calibration-validation section. No need to show runoff coefficient for the scenario without B.C. as this will clearly be worst than the scenario with B.C.

We better clarified the sense of these results in the text. Runoff coefficient (..or runoff ratio) are obtained using Express HYDRO as input, so they cannot be considered part of calibration (which is done using observations on a period where they are continuously available) but are part of the hydrological reanalysis. We still believe interesting to keep both results (with and without B.C.) because they show the impact of rainfall B.C. but also of the other variables (there are already comments about this on page 22 lines 11-17); anyway if the editor and reviewer believe that results without B.C. are unuseful we can remove them

1) While I am not a native English speaker, I need to point out that the paper is poorly structured and the quality of the English is low. This is a very serious concern that has to be solved even before discussing the technical details. Many sentences are quite long and with poor grammar. In each section, the length of the paragraphs varies too much (from a few lines to an entire page). I suggest the authors to have a native English speaker proofread the paper.
We reviewed the general quality of language and writing, moreover we are open to ask to native English speaker to check the manuscript, if editor and reviewer retain it necessary..

2) The paper structure needs to be improved. The methodology presents results of the hydrologic model calibration; there is a section on the datasets, but new datasets seem to appear in the Results (or maybe they are the same, but named in a different way). There are too many figures that can be merged (see below).
Agreed, we improved the description and naming of data and give more detail about hydrological model. Some parts of the paper were restructured by moving some sections in different parts of the manuscript. Some figures was merged.

3) The authors need to describe better what is the novelty of their analyses. The idea of applying a cascade-based approach to issue hydrologic predictions has been used in other studies to evaluate climate change impacts (in this case, climate model outputs are used instead of reanalysis) and improve hydrometeorological forecasts (in this case, outputs of Numerical Weather Prediction models are used instead of reanalysis). The Introduction of the paper does not discuss previous applications that follow the same strategy and does not highlight what are the main contributions of this study (in my opinion, it is the analysis of the simulated distribution of flood extremes). In addition, the authors should clearly state which new analyses and simulations have been performed in the paper –I guess these are bias correction, statistical downscaling and hydrologic modeling– or conduced in other studies –the dynamical downscaling of ERAINT.

It is true that the similar modeling cascade were already applied, but to the best of our knowledge this is one of the first works of this kind done in the considered study area in which the environment is constituted by small basins with flash flood hydrological regime. Moreover the usage of such high resolution reanalysis is quite new. We highlighted these points in the text as suggested trying to better clarifying the novelty elements of the work.

4) The authors need to present more details on the hydrologic model, including: how the model has been parameterized, calibrated, and validated with observed data; which soil and vegetation have been used (maybe show a map as well?); which observed hydrometeorological forcings have been adopted and how they have been interpolated in space; and show examples of simulated and observed hydrographs in the calibration and validation periods.
*Agreed, we improve both sectios 2.1 and 2.5. Moreover we have a recent published paper to be cited where we used a similar setting of the model on the same study area for different purpose (Davolio et al., 2017). We did not devote too much space regarding the calibration of the model because it was partially faced in other works and it is not the core of the paper but it is a "functional" activity for the other analysis. To summarize the performance of the model we inserted a table with the skill scores values. In figure 1 we added a Curve Number map in order to give information about soil use and vegetation.*

5) A discussion on the performances of the statistical downscaling algorithm is missing.
The testing and application of RainFARM in the study area for forecast purposes was done in many works, some already cited in the text (Rebora et al. 2006; Silvestro et al. 2012). In this work it is used, as mentioned

in the text, to generate a possible downscaled rainfall scenario from 4km-3hours to 1km-1hour resolution. Effectively we did not test the benefit of using or not the downscaling.

*To assess the impact of downscaling we made the hydrological simulation with BC rainfall but **without** applying the downscaling, we estimated the ratio of Qmean without and with Downscaling for each pixel. Results are plot versus drainage area (see figure below). It is quite evident the impact of downscaling, the impact generally increases when drainage area decreases. The downscaling is quite important even in this application especially to deal with small basins.*
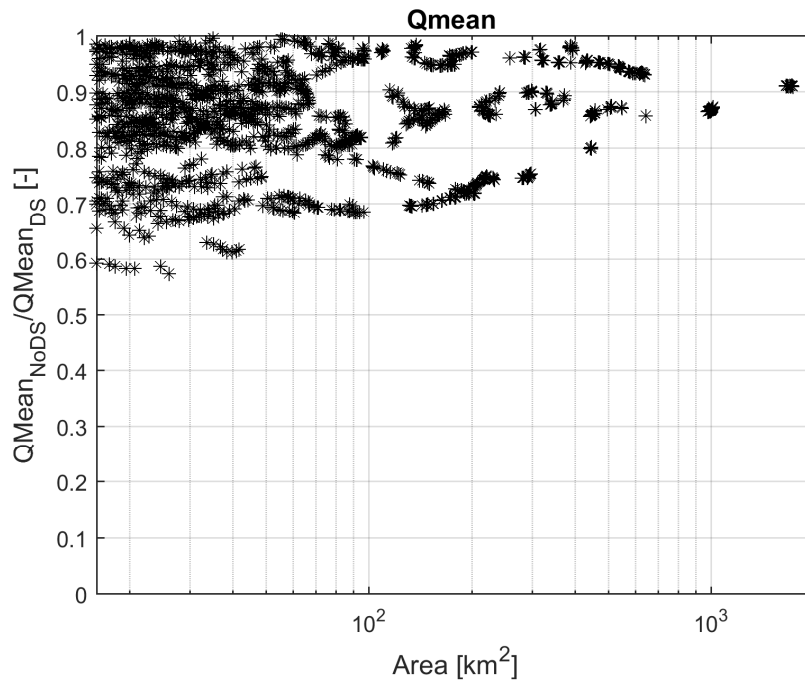
6) The relatively long discussion of the comparison of the WRF simulations with an alternative rain gauge network to the one used by Pieri et al. (2015) should be shortened. By the way, are the two rain gauge networks completely different?

We can better clarify this point in the text. The raingauge data set used in Pieri et al. (2015) is a subset of the one used in our work.

Figures.
We modified the figures following the reviewer suggestions.

References

Cenci L., Laiolo P., Gabellani S., Campo L., Silvestro F., Delogu F., Boni G., and Rudari R., "Assimilation of H-SAF Soil Moisture Products for Flash Flood Early Warning Systems. Case Study: Mediterranean Catchments", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 9, no. 12, pp. 5634-5646, Dec. 2016. DOI: 10.1109/JSTARS.2016.2598475

Davolio, S., F. Silvestro, and T. Gastaldo, 0: Impact of rainfall assimilation on high-resolution hydro-meteorological forecasts over Liguria (Italy). J. Hydrometeor., 18, 2659-2680,https://doi.org/10.1175/JHM-D-17-0073.1

Giannoni, F., Roth., G., and Rudari, R.: A Semi – Distributed Rainfall – Runoff Model Based on a Geomorphologic Approach. Physics and Chemistry of the Earth, 25/7-8, 665-671, 2000.

Giannoni, F., Roth, G. Rudari, R.: A procedure for drainage network identification from geomorphology and its application to the prediction of the hydrologic response, Advances in Water Resources, 28(6), 567-581, 2005, doi:10.1016/j.advwatres.2004.11.013.

Regione Marche. -Regionalizzazione delle portate massime annuali al colmo di piena per la stima dei tempi di ritorno delle grandezze idrologiche. http://www.regione.marche.it/Regione-Utile/Protezione-Civile/Progetti-e-Pubblicazioni/Studi-Meteo-Idro#Studi-Idrologici-e-Idraulici. Last access date:19/10/2017

Rebora, N., Molini, L., Casella, E., Comellas, A., Fiori, E., Pignone, F., Siccardi, F., Silvestro, F., Tanelli, S., Parodi, A.: Extreme Rainfall in the Mediterranean: What Can We Learn from Observations? J. Hydrometeorol., 14, 906-922, 2013.

Silvestro, F., Rebora, N., Giannoni, F., Cavallo, A., Ferraris, L., (2016) The flash flood of the Bisagno Creek on 9th October 2014: an "unfortunate" combination of spatial and temporal scales. Journal of Hydrology, 541, Part A, Pages 50–62, doi:10.1016/j.jhydrol.2015.08.004