Author response to the Referees

Title: Improving runoff estimates from regional climate models: a performance analysis in Spain

Authors: D. González-Zeas, L. Garrote, A. Iglesias and A. Sordo-Ward

Article Type: Research Paper

Dear Editor,

Please find enclosed the response to all comments made by the two Referees to the manuscript "Improving runoff estimates from regional climate models: a performance analysis in Spain" to be considered for publication in Hydrology and Earth System Science.

We gratefully acknowledge the helpful comments that have contributed to the improvement of our paper. We have made the suggested revisions by the two Referees to improve on the presentation of methods and analysis of performance; the changes are addressed in comments 1 to 26**.**

Thank you very much for your consideration.

Sincerely,

Dunia González

Department of Hydraulic and Energy Engineering
Universidad Politécnica de Madrid, Spain

The comments of the Referees are numbered **1 to 26**. Following each comment, **R** indicates the response of the authors.

Referee 1 General comment

1. General comment. A good example of how to condense a very large data set from 10 different RCMs into few meaningful analysis tests is provided in this work an in-depth insight into the applicability of RCM for hydrological and water management evaluations which is really a hot topic in today's hydrology debate. This study underlines the value of the direct surface runoff simulated by RCMs, suggesting that, despite the need of bias correction, the accuracy of the results may be comparable to those obtained by other studies using the climatic model output as an input to finer resolution water balance models (standard procedure). Besides this, 5 aridity index based on climate simulations are also adopted as a long-term mean water balance estimator. The novelty of this work is mainly in the use of these two approaches adopted to test the accuracy of different RCM in predicting discharge in large rivers basins. Natural discharge data from the entire mainland territory of Spain are compared with different RCM output downscaled to 2.5 min resolution (about 4x4 km) using different (but very simple) interpolation methods thus evaluating the ability of each method to reproduce the observed behaviour. The interpolation alternatives consider the use of the direct runoff from the RCMs and the mean annual calculated using different aridity indices and are evaluated in terms of their bias and index of agreement with observations thus providing a convincing picture of the performance of the different RCMs, interpolation methods/grids and aridity index formulation. In this sense, this work could help the scientific community in further advancement towards a better use of climate scenarios in hydrological applications. A few improvements in the presentation of methods, application and results will probably enhance the evaluation of this paper by the community concerned with the hydrological impacts of climate change at river basin scale.

R: We have made the suggested revisions by the two Referees to improve on the presentation of methods and analysis of performance; specific changes are addressed below and in the response to Referee 2.

Referee 1 Specific Comments

2. Complex downscaling methods. The problem of bias when dealing with RCM output is well introduced in section 3.4, though the presentation of more popular techniques of bias correction does not give proper credit to the statistical downscaling

methodology and related literature works which represent one of the more efficient way to operate bias corrections on the whole empirical frequency distribution of the interested variable (i.e. quantile mapping method, applied by e.g. Wilby et al. already referenced papers; Deque, Global and Planetary Change, 2007; Bardossy and Pegram, WRR, 2011).

R: We have modified Section 3.4 to recognise the value of statistical downscaling techniques. To that end we have expanded Section 3.4 including discussion of the use of quantile mapping methods (Wilby et al., 2000; Déqué, 2007; Bardossy and Pegram, 2011). A practical limitation of the quantile mapping method is data availability. Here we analyse behaviour of discharge over a very large territory (about 500,000 km^2) and the data limitations over this territory are the main reason for selecting simple interpolation methods. We use two interpolation methods to evaluate the uncertainty derived from the approach. In our study, observed monthly runoff is available at the basin scale, and only annual runoff is available in the grid cells. Therefore the alternatives that minimize bias with respect to the observed values provide only annual values of runoff.

We have included in the revised reference list the following references:

Bárdossy, A. and Pegram, G.: Downscaling precipitation using regional climate models and circulation patterns towards hydrology, Water Resour. Res., 47, W04505, doi:10.1029/2010WR009689, 2011.

Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenica scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16-26, 2007.

Wilby, R. L., Hay, L. E., Gutowski, W. J. Jr., Arritt, R. W., Takle, E. S., Pan, Z., Leavesley, G. H., and Clark, M. P.: Hydrological responses to dynamically and statistically downscaled climate model output, Geophys. Res.Lett. 27, 1199-1202, 2000.

3. Performance. The bias correction method presented in section 3.4 is applied to the monthly series in section 4.3. Comparison against observation was performed between monthly frequency distributions while the reader would have found a comparison between time series more significant. Therefore it would be interesting to report the performance of the applied bias correction using the NS to the time series before sorting operation. Moreover, the comparison between frequency distributions themselves (Fig. 6) is only qualitative or based on overall statistical indicators such as the NS which is not common for distributional comparisons. Comparison between frequency distributions is usually performed with specific statistical test exploring both the average and extreme behaviour of the variables (see e.g. Portoghese et al., NHESS, 2011) while the adopted performance test are more commonly used to analyze the

fitting between time series. Furthermore the authors should put emphasis on distributional differences since it is likely that major effect of climate change will concern extreme events (floods/droughts) and climate variability in general rather than the average behaviours.

R: We have modified the performance analysis and included the methods for comparing frequency distributions as suggested by the Referee. The completely new revised section, including new tables and figures, is included below.

Time series analysis is a challenge when considering that the runoff simulations of the RCM do not correspond with an actual calendar date, but corresponds to a potential realisation of a weather pattern on a given decade, therefore year-to-year comparisons may be misleading. Therefore we first compare monthly climate patterns using the Kolmogorov-Smirnov (K-S) test − the corrected monthly distributions of runoff with respect to the observed monthly distributions of runoff −, second we represent the annual hydrological cycle by using regime curves (30 year average monthly values), and finally we use the NS coefficient of efficiency to evaluate the hydrological cycle. This analysis of distributional differences is relevant to climate change since a major effect of climate change will concern extreme events (floods/droughts) and climate variability in general rather than the average behaviours.

We have included a non-parametric test for equality of CDF of two populations instead of NS indicator, in order to evaluate on what extend the corrected monthly frequency distributions are representative of the observed data. We have applied a Kolmogorov-Smirnov (K-S) test. We have added in the section 3.4 a description of the K-S test.
The Fig. 6 has been changed in order to show the results obtained for the 10 RCM simulations instead only the mean according to the Referee 2 suggestion. Now the Fig. 6 shows the monthly distribution of runoff in Spain simulated by the 10 RCMs and their mean. The results are shown for the direct runoff, bias corrected runoff by UNH/GRDC dataset and bias corrected runoff by Schreiber formula.
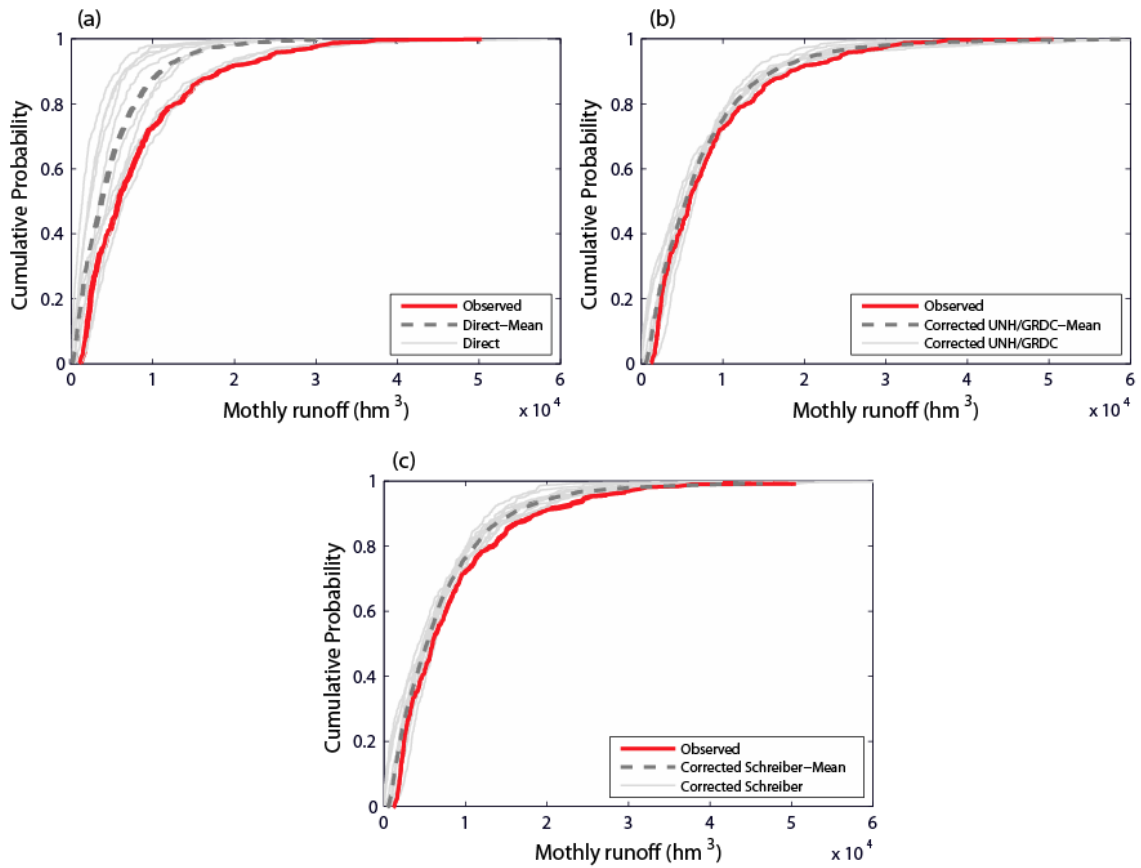
Fig. 6. Monthly distribution of runoff in Spain. The figures show the cumulative probability of the mean runoff of every month of the entire period of analysis (1961-1990). The comparison is made between the observed data and (a) direct runoff, (b) bias corrected runoff by UNH/GRDC dataset and (c) bias corrected runoff by Schreiber. Red line is observed monthly distribution of runoff, grey lines are monthly distributions of runoff simulated by the 10 RCMs and dotted line is the mean of the simulations.

Table 4 (in section 4.3) outlines the results of the K-S test corresponding to the Fig. 6, for the 10 RCM simulations and their mean.

Table 4. Test results (p-values) of the K-S test for the difference between observed and (a) direct runoff, (b) runoff bias-corrected by UNH/GRDC and (c) runoff bias-corrected by Schreiber formula in Spain, at 95% confidence level. Bold characters are used to remark the passed test.

Alternatives	DMI1	DMI2	DMI3	ETH	GKSS	ICTP	KNMI	MPI	SMHI	UCM	Mean
Direct Runoff	1,9E-26	1,5E-27	6,0E-21	8,1E-48	0,04	1,0E-04	2,3E-03	1,0E-14	3,4E-11	9,0E-09	1,4E-07
UNH/GRDC	0,03	**0,07**	0,01	5,6E-09	**0,48**	3,1E-03	3,7E-05	**0,10**	3,4E-11	3,4E-09	**0,12**
Schreiber	1,8E-03	0,02	1,8E-03	2,3E-08	**0,20**	2,3E-03	5,3E-04	**0,10**	2,0E-11	9,0E-09	**0,06**

We have included a new Section 3.5 in order to represent the annual hydrological cycle using regime curves and the NS coefficient of efficiency to compare the 30-year

5

average monthly values. Studies like those by Wolock and McCabe (1999) and Sperna Weiland et al. (2010) have also used the NS to represent the ability of the simulations to reproduce the mean monthly observed data.

We have included in the revised reference list the following references:

Wolock, D.M. and McCabe, G.J.: Explaining spatial variability in mean annual runoff in the conterminous United States, Clim. Res., 11, 149-159, 1999.

Sperna Weiland, F.C., van Beek, L.P.H., Kwadijk, J.C.J. and Bierkens, M.F.P.: The ability of a GCM-forced hydrological model to reproduce global discharge variability, Hydrol. Earth Syst. Sci., 14, 1595-1621, 2010.

Finally we further analyse the mean monthly runoff in Spain (regime curves) for the 10 RCMs (Fig. 7), and calculate the NS coefficient of efficiency to represent the performance of the applied bias. Figure 8 has been changed and now it shows the NS efficiency coefficients of the 30-year average monthly runoff for Spain and its 14 river basin districts in order to take into account the spatial performance.
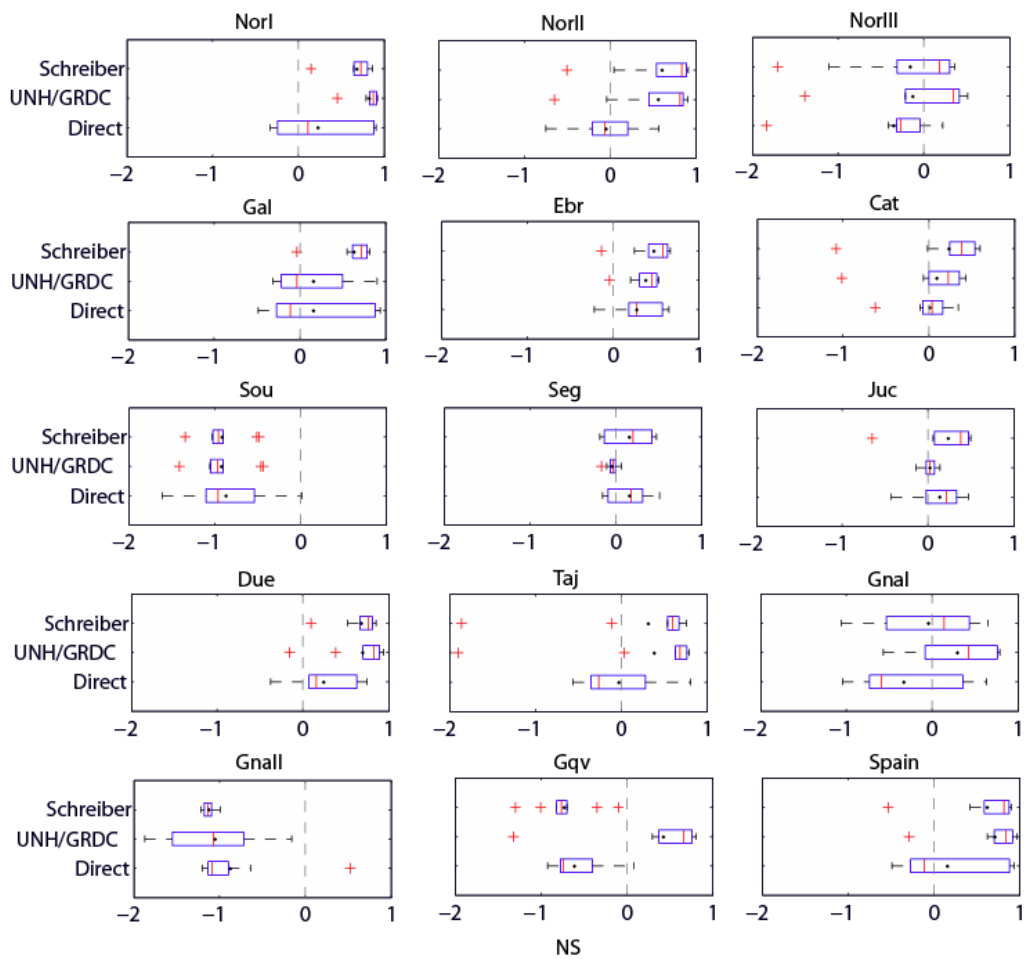
Fig. 8. Nash-Sutcliffe (NS) efficiency coefficients of the 30-year average monthly mean of direct and corrected surface runoff for Spain and its 14 river basin districts. The boxplots show the NS obtained by the 10 RCM simulations, the lines extend up to 1.5 times the interquartile range to the right and left of the box, the box extends from the 25th percentile to the 75th percentile, the line within the box indicates the median simulations, the black point shows the mean of the simulations and red crosses outside the box indicate the outliers. The river basin district code is defined in Table 1.

Referee 1 Technical corrections

4. On page 177, lines 16 and 17, the term layer referred to data records corresponding to river section measurements is not properly correct. May be changed into runoff dataset.
R: Revised.

5. On page 178, line 21, the term desegregations is wrong. Should be changed in disaggregation. This term is used somewhere else through the text. Please check.
R: Revised.

6. On page 183, line 28, the term relation may be ambiguous. Better using ratio.
R: Revised.

7. On page 183, line 28, and in all other cases (please check) the term evaporation should be evapotranspiration according to the commonly used notation and adopted formula reported in Table 2. According to the same water balance terms, the adjective current on page 184, line 1, and in all other cases (please check), should be changed into actual.
R: Revised.

Referee 2 General comment

8. General comment. The paper by González-Zeas et al. focuses on the assessment and evaluation of monthly runoff time series derived from the bias-corrected output of 10 regional climate model simulations, comparing the results of a variety of simple interpolation methods against 'observed values' (which in fact were estimated through applying the hydrological model SMPA, due to a lack of gauging data reflecting natural flow conditions) for Spain's mainland, comprising 338 sub-basins. For further comparison to the application of direct surface runoff products from RCM, five functional descriptions of the aridity index and the UNH/GRDC global runoff data set is used. The rationale of this study is to reveal the applicability of direct and processed RCM outputs of runoff series for large scale areas, for which no reliable calibration of distributed hydrological model is feasible. The value of this paper for the scientific community can be seen in the demonstration that reasonable runoff estimates from RCM data seem achievable, even in applying only rather simple techniques. Overall, the paper is well written, generally well structured and relatively concise in its description of applied methods and results. however, I think that in some parts it is a little bit too simplistic and straightforward so that certain elements could and should be addressed in more detail (while others could be left out with no harm), to actually explain the limitations and inaccuracies and specific conditions under which the proposed approach(es) can be beneficial.

R: We have made the suggested revisions by the two Referees to improve on the presentation of methods and analysis of performance; specific changes are addressed below and in the response to Referee 1.

Referee 2 Specific comments

9. Abstract. I do have a problem with the first sentence to start with, because this is clearly nothing you address any further in your paper – monthly time series of present state runoff estimates are by itself not at all important to assess the impact of climate change on (future) water availability: The methods applied to obtain good estimates from RCM are (as you claim)! But once bias-correction is involved, you carefully need to argue whether you trust that BC-terms/factors can still be valid for future climate, especially when considering the loss of (physical) data consistency through bias-correcting only one or few variables from an RCM (but all this is a much more critical debate then the one you should raise here). I am fully in line with you that your proposed methodology proves valuable and helpful when trying to provide estimates for runoff conditions in current state climate and thus value-adding RCM outputs for an application in water resources management (models) (as explained later on).

R: We have modified the abstract; the new sentence is" An important step to assess water availability is to have monthly time series representative of the current situation".

10. Introduction. You provide a very comprehensive overview on current issues and problems in large scale hydrology and explain reasonably well the potential and limits of RCM data with regard to relevant literature. Again, I am not so sure whether you should try and make that link to CC impact analysis.

R: The following paragraph has been added: "Our study focuses on providing estimates for runoff conditions in current situation using the RCM outputs; nonetheless if the RCM runoff is to be used in CC impact analysis, the first step is to analyze how the RCM simulations reproduce the current situation and therefore our methodology focuses on the control scenario as a first step before linking to CC impact analysis."

The link to CC impact analysis has not been raised here, however, we also agree with you that the BC techniques may not be applicable in modified climate.

11. Chapter 2.2. Can you please comment on the performance of the calibrated SIMPA model over Spain, i.e., what is the quality of the reference you are using? Since it has been calibrated (validate?), I assume such information should be given to the reader without a need to find the Spanish reference.

R: The SIMPA model has been calibrated over 100 control points of Spain, using gauging stations where stream flows are measured in natural regime. The model shows a good fit at all control points and these results are available in the White Paper Book of Water in Spain (MARM, 2000). The monthly time series estimated by the SIMPA model are also being used for the National Water Master Plan of the Water Framework Directive and therefore are the best global estimate available in Spain.

In order to provide more detail information about the SIMPA model we have added two references to the manuscript related to model calibration:

MARM, Ministry of Environment: White Paper Book of Water in Spain, 2000. http://www.mma.es/rec hid/libro b/index.htm.
Álvarez, J., Sánchez, A. and Quintas, L.: SIMPA, a GRASS based tool for Hydrological Studies, Proceedings of the FOSS/GRASS users Conference, Bangkok, Thailand, 12-14 September, 2004.

12. Chapter 2.3. You have using 10 RCMs (eight models, 1 in three members), all driven by the same GCM. Can you please comment on why you made that choice and why you didn't consider to use a set of different RCMs driven by different GCMs? It

may well be that this has been considered to be outside the scope of your paper, but I think that it is quite crucial to understand your intention behind using 10 RCM model outputs, especially since you seem to average them later on, which is not really valid.

R: The modeling chain including socio economic assumptions, climate models, regional downscaling and bias correction provides a basis for a variety of uncertainties. A major source of uncertainty in climate change impacts on water availability is derived from the representation of runoff in the regional models. Here we focus on the concrete issue of downscaling runoff and bias correction. An uncertainty study is out of the scope of this research, and may include, different assumptions on greenhouse gas concentrations defined by socio-economic projections, different GCMs that represent the climate system, and a range of downscaling methods, among others. However, conclusion obtained by Quintana Seguí et al. (2010), shows that the downscaling and bias-correction of the RCM has proved very useful when analyzing the impact with a single GCM given that it is possible to get a range of variation of the uncertainties.

Regarding your comment about to average the 10 RCM simulations, we agree with you, therefore we have changed the Fig. 6 and Fig. 8 in order to represent the behaviour of the 10 RCM results.

The reference mentioned is:

Quintana Seguí, P., Ribes, A., Martin, E., Habets, F. and Boé, J.: Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins, J. Hydrol., 383(1-2), 111-124, 2010.

13. Chapter 3.1. a) The fact that you actually considered the impact of the two available coordinate systems is good and thoughtful. b) The interpolation due to the scale mismatch between RCM and hydrological units is obviously necessary, yet the schemes you apply are really the most simple ones available. I trust that this is ok for the purpose of your study, but I think you should indicate the vast availability of more sophisticated methods (e.g. of purely higher statistical order or even mass/energy-conserving approaches considering the underlying topography that comes into play when you turn fluxes into flows) by ways of quoting relevant literature.

R: We have expanded section 3.1 including relevant literature related with the more sophisticated methods, according to your suggestion. The new references included are:

Daly, C., Gibson, W.P., Taylor, G.H., Johnson, G.L. and Pasteris, P.: A knowledge-based approach to the statistical mapping of climate, Clim. Res., 22, 99-113, 2002.
Hijmas, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25, 1965-1978, 2005.

Thornton, P.E., Running, S.W. and White, M.A.: Generating surfaces of daily meteorology variables over large regions of complex terrain, J. Hydrol., 190, 214-251, 1997.

14. Chapter 3.2.: a) Did you ever look at PET in the RCM output data? If no, why not (not available to you?), if so, how does it compare to your Hargreaves estimate (Eq. 3)? If there should be a mismatch (which I would definitely expect), what does that imply on your runoff estimates with regard to closing the water balance?

R: Some studies suggest that potential evaporation provided by climate-model simulations are strongly biased and not suitable for direct use in hydrological models (Teutschbein et al., 2011). Because the limitations of the PET in the RCMs, the PRUDENCE project does not make readily available outputs of PET. However, for computing runoff using the five functional descriptions of the aridity index, it is necessary to compute PET. We have used the Hargreaves method because it uses information that is provided by the RCM simulations and it was also the method based on temperature with highest rating to calculate PET reported by Jensen et al. (1990) and the only method based on temperature recommended by Shuttleworth (1993).

Teutschbein, C., Wetterhall, F., Seibert, J., 2011. Evaluation of Different Downscaling Techniques for Hydrological Climate-Change Impact Studies at the Catchment Scale, Climate Dynamics, 37:2087–2105, doi 10.1007/s00382-010-0979-8
Jensen, M.E., Burman, R.D. and Allen, R.G.: Evapotranspiration and irrigation water requeriments, ASCE Manuals and Rep. on Eng.Pr., 70, 360 pp, Am.Soc. of Civ. Eng. Reston, Va., 1990.
Shuttlewort, W.J.: Evaporation in Handbook of Hydrology, edited by D.R. Maidment, p. 418, McGraw-Hill, New York.

15. Chapter 4.1. Is there a reasonable explanation why O-D outperforms the other interpolation schemes (Fig 4a and b)? If yes, please do explain.

R: We consider that the O-D method outperforms the other interpolation schemes because of the influence of the size of the basins. Given that in general the study basins are small, using an interpolation method does not give a greater precision by including values around in the interpolation, because of the spatial correlations, so the Direct method performs better. In addition, the use of the CRU data suppose a resampling of the data simulated by the original simulations of the RCMs, therefore, the original coordinate systems provide the better results.

See also Response to comment num. 3 of Referee 1.

16. As you mention, the deviances among the four alternatives are markedly small (probably not significant?) – are they arbitrary?

R: Even if the differences among the four alternatives are markedly small, we have seen that they are systematic.

17. Chapter 4.2 +4.3 With regard to Eq. 6 and in the light of the monthly time series you are providing and evaluating later on, have some doubts that this simple approach is feasible here. In applying only an annual correction factor, which you then superimpose on each monthly runoff value, you are accounting only for the bias in runoff volume. In ranking according the best performance later on, your approach totally neglects than certain RCMs may perform bad in terms of a bias in volume (which is fairly easy to correct), but are capable to represent the annual hydrological cycle (which would be much more difficult to compensate, if the seasonality is not preserved…). This becomes evident in Figure 7, for which you claim that the 'corrected RCM series adequately represent the seasonal cycle' (on page 191, line 19). This has not much to do with your correction method, doest it? Please comment.

R: We correct bias on annual time series and at the same time maintaining the seasonality given by each RCM simulations, in view of the functional forms based on the aridity index which only provides annual values of runoff. In the case of UNH/GRDC dataset, mean monthly runoff and mean annual runoff are available. Nonetheless, we have found that mean monthly runoff from the UNH/GRDC are not able to reproduce the observed seasonality. Using the annual corrections factors from the alternatives that minimize bias with respect to observed values, we have obtained 10 bias-corrected monthly time series with runoff volume equal to that obtained with the Schreiber formula and 10 bias-corrected monthly time series with runoff volume equal to that provided by the UNH/GRDC dataset. Each of these bias-corrected time series maintain the seasonality simulated by their respective RCMs.

The results shown in Fig. 7 highlight the capability of the 30 year average monthly bias-corrected values obtained by each RCM to reproduce the observed annual hydrological cycle. However, the capability to reproduce the seasonality is not related to our correction method, it is related with the ability of each RCM simulation to reproduce the observed seasonality. In this context, considering that some of the RCM simulations better represent the observed seasonality than others, we agree with you that it is important to evaluate the behaviour of each RCM rather than their average. For this reason, we have changed the Fig. 6 and Fig. 8 in order to represent the results of the 10 RCM simulations individually.

18. Page 190, line 26: the bias in RCMs is not 'inherent' per se, please avoid this term.

R: This has been eliminated.

19. Fig. 6 and following explanation on page 191, line 23: You cannot draw a serious conclusion from averaging runoff from 10 different RCMs

R: According to your suggestion, we have changed Fig. 6 in order to present the results obtained by the 10 RCM simulations.
See also Response to comment num. 3 of Referee 1.

20. Page 192, line 1-2: The NS values you indicate are referring to what? A time series of monthly runoff values or the cumulative probability distributions? In the latter case, I don´t think that NS is the appropriate objective function and you should find a more appropriate measure.

R: According to your suggestion as well as the suggestion of Referee 1, we have used a statistical indicator for distributional comparisons.
See also Response to comment num. 3 of Referee 1.

21. Page 192, line 11f. and Fig. 8: Again, I don´t think you should show the average of 10 RCMs, but much rather I advise to use a boxplot, indicating all 10, but also the mean and median of the 10.

R: The Fig. 8 has been changed according to your suggestion. Fig. 8 shows a boxplot of the 10 RCM results, indicating the mean and median of the 10 RCM simulations. Please, see response to comment 3.

22. Figure 8: Can you also show the NS values when referring to the lines given in Figure 7? Does that make a difference?

R: We have used K-S test to compare the monthly distribution of runoff and according to the suggestion of Referee 1, we have used the NS values to evaluate the performance of the 30-year average monthly values in Spain and its 14 River Basin Districts. Please, see response to comment 3.

23. Chapter 5: a) Page 193, line 12-18: The link that is made here is not clear to me. You claim that the simple disaggregation scheme O-D has been applied in earlier studies, but proved not very efficient due to the absence of a bias-correction step. I think that these processing steps are to be seen independently, there is no causal link-and the impact of bias-correction clearly dominates the choice of interpolation scheme, no?

R: We completely agree with this comment, so we have re-written the paragraph in order to avoid confusion.

24. b) Page 193, lines 26, 27: contents repeated

R: The repeated contents have been eliminated.

25. c) In your conclusion, it would be helpful to add some more concluding explanation to the description of performances. What do we actually learn from your findings, if in some cases the RCM data, sometimes the Schreiber and mostly the UNH/GRDC data provide best results?

R: We have modified the conclusions in order to present a better explanation about our findings and the following paragraph has also been added: "In general, the findings of this study conclude that, for using direct surface runoff outputs from RCMs, the O-D interpolation method at basin scale and the bias-correction with annual factors given by the UNH/GRDC dataset or Schreiber's climate formula alternatives are necessary in order to obtain simulated monthly runoff that most closely approximate to the observed values".

26. d) The regional differences in your results are quite interesting and I think you present some logical arguments (e.g. catchment size) to explain the findings. Still, I am not convinced that there isn`t more in this. Is there any geographical context and some clear causes for these obvious differences? Could you provide an answer on why Schreiber´s formula works best in certain basins while it fails in others?

R: The results show some regional differences which may be caused by: 1) the deficiency of the RCMs to simulate the climatic variables in arid and semi-arid regions, 2) the UNH/GRDC dataset has been calibrated in the Spanish gauging stations that belong to the RBDs that give the better performance results by the UNH/GRDC alternative, 3) the inability of the climate formulas to capture the impact of rapid precipitation events over arid and semi-arid regions and 4) the basins size.

In order to represent the regional differences spatially, following we may add a figure indicating the RBDs where the UNH/GRDC and Schreiber formula perform better. If the Referee thinks this figure is interesting, we will include it in the revised manuscript.
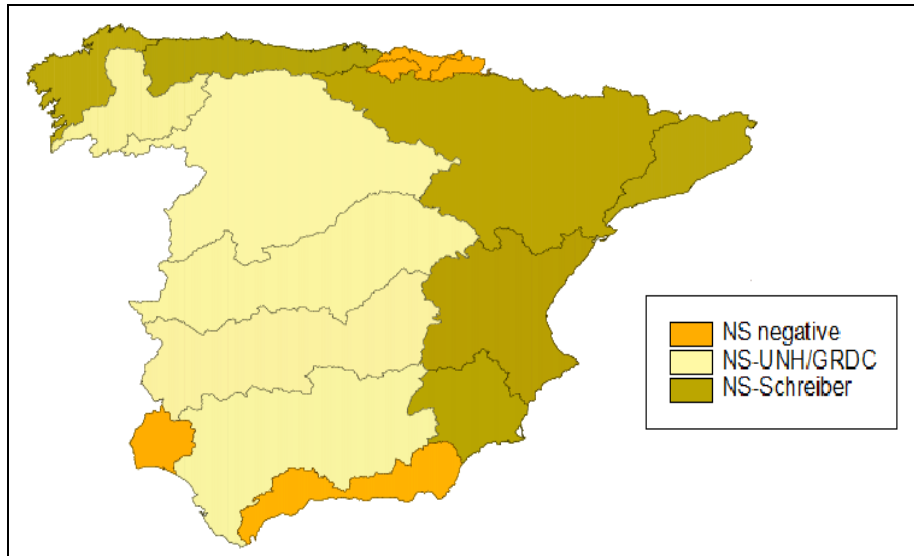
Fig. Nash-Sutcliffe coefficients for the regimes curves (30-year average monthly mean runoff) in the 14 RBDs of Spain. The figure shows the RBDs where UNH/GRDC dataset and Schreiber formula perform better and the RBDs with negative values of NS.

In wet regions, like in the north of Spain, where the precipitation always exceeds the potential evaporation, any error in the precipitation translates to approximately the same absolute error in runoff (which will result in higher relative error since runoff is always less than the precipitation). In semidry regions, the runoff-generation processes are highly nonlinear; therefore the errors in precipitation translate into even greater errors in the runoff. In arid regions, the water balance calculation does not produce any runoff (in some cases due to the inability of the climate formulas to capture the impact of rapid rain events over small scales), therefore the runoff estimate is virtually insensitive to the precipitation inputs. So, it is necessary to improve precipitation estimates in arid and semiarid regions, where slight changes in precipitation can result in dramatic changes in the runoff, due to the nonlinearity of the runoff generation processes. In the case of Spain, Guadalquivir and South water district basins provide the worst results with the Schreiber formula and it's because of the geographic influence, given that these river basin districts are located in the South of Spain with semiarid and arid climatologic characteristics. On the other hand, Schreiber's formula performs better in the RBDs where UNH/GRDC dataset has not been calibrated.