The authors are grateful to Reviewer #1 for sharing his concerns with the authors of this submission. Your concerns were carefully taken into consideration for the manuscript revision, leading to a more consistent and scientifically sound manuscript. Major changes and additions were performed in the entire manuscript. In the pages below, we elaborate with the major and minor concerns that led you to the decision for major revision. A point by point reply is given below.

1. Does the paper address relevant scientific questions within the scope of HESS? YES, it explores the response of European river basins to climate change

2. Does the paper present novel concepts, ideas, tools, or data? YES, although projections for hydrological impacts of climate change in itself are not new. Projections with this particular model JULES and based on this set of CORDEX simulations are new. It uses comprehensively bias corrected data following new methods (itself described elsewhere). The focus on low flows and droughts as presented here, is also relatively unexplored.

3. Are substantial conclusions reached? YES

4. Are the scientific methods and assumptions valid and clearly outlined? YES

5. Are the results sufficient to support the interpretations and conclusions? YES minor: twice (p7281 L1 and p7292 L1) a statement is made on floods which to my opinion cannot be derived from the present analysis. I suggest to simply omit these.

REPLY: The flood related unsupported statements were removed from the manuscript.

6. Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? YES

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? YES, extensively

8. Does the title clearly reflect the contents of the paper? The title mentions water stress which is a function of both availability and demand. This is not directly analyzed in the paper. Also hydrological model biases are not analysed, only the effects of forcing biases. I suggest to change the title to something like: "High-end climate change impacts on European runoff and low flows: exploring the effects of forcing biases"

REPLY: Following the reviewer's indication, the title was changed to better correspond with the topic of the manuscript. The new title is "High-end climate change impact on European runoff and low flows. Exploring the effects of forcing biases".

9. Does the abstract provide a concise and complete summary? YES

10. Is the overall presentation well-structured and clear? Overall, the paper is well structured and clear. However, the introduction is too long and its structure is not always clear: â˘A ´c p 7268-7269 (para 1 and 2) are OK â˘A ´c para 3, 5 and 7 are bit long but generally OK â˘A ´c p7270-7271 (para 4): the discussion on added stresses of population growth and human activities is not relevant in present context; a single statement reflecting its significance here or in the discussions section (4.1) suffices â˘A ´c para 9, p7273, is superfluous after para 8 â˘A ´c p7273-7274 (para 10): the discussion of GHMs/LSMs is not too relevant in the present context â˘A ´c p7274-7275 (para 11): the JULES discussion can be omitted here and partially merged with section 2.2 Section 2 is OK except that I would put the present 2.5 – Bias correction directly following the present 2.1 – climate/forcing data. Section 3, 4 and 5 are OK

REPLY: The reviewer's suggestions were carefully taken into account into the revised version of the Introduction. The Introduction now is shorter and more focused on the topics of this study as redundant information have been removed. Specific changes made per paragraph are:

- Paragraph 4 was deleted as its content (added stresses of population growth and human activities on climate change impact) is not relevant with this study. A reference on the significance of this topic is made in Paragraph 1 of the Introduction.
- Paragraph 5 on multi-model assessments was significantly shortened and moved to the end of Paragraph 3
- Paragraph 9 was removed from the manuscript as it was repeating the information of Paragraph 8.
- As the content of Paragraph 10 (discussion on other LSMs) was not very relevant to the context of this study, only a short definition of the LSMs has been kept of this paragraph. This is now merged into a paragraph where the JULES model is briefly mentioned.
- The discussion about JULES in Paragraph 11 was removed from the Introduction section and was moved to the Data and Methods section (merged with Section 2.2.)
- The research objectives in the last paragraph of the Introduction have been reformulated.

In Section 2 we followed the reviewer's indication of moving Section 2.5 (Bias Correction) right after Section 2.1.

11. Is the language fluent and precise? Yes, the quality of english is generally high and precise, with a few minor exceptions

12. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

I miss a brief description of the forcing data: how large are biases? While the five models are similar in projected temperature change (due to the time slice strategy, centering on ïA¿Dˇ T =2 and 4K resp.) there is no indication of how their precipitation changes. What is the CV over the 5 models here before/after bias correction? Can we have 5 maps with ïA¿Dˇ P? Could be part of section 2.1 or a new starting subsection in 3 or an expansion of 3.6. For other variables biases could be presented in supplementary material.

REPLY: Following the reviewer's indication, two Figures describing the effect of bias correction on the forcing variables, have been added to the Electronic Supplementary Material. Figure S1 shows the effect of bias correcting against the WFDEI dataset and Figure S2 against the E-OBS dataset. Results are shown for precipitation and temperature as the rest of the forcing variables were not bias adjusted. The absolute differences between bias corrected and raw input (bc-raw) are shown for all the participating GCMs and for their ensemble mean. The cv between the ensemble members before and after bias correction has also been calculated. In each sub-figure, the spatial average of each illustrated map is noted in each sub-figure.

These were used to also understand the differences between the two observational datasets. A relative comment deduced from these figures has been added to Section 3.6.: "From Figures S1 and S2 of the ESM (showing the effect of bias correction on the forcing variables of precipitation and temperature) it can be deduced that that E-OBS corrected precipitation has lower values than precipitation adjusted against the WFDEI dataset. This explains the lower runoff produced by the E-OBS bias adjusted dataset, as it is reasonable for the differences in precipitation to reflect on the output of the hydrological model. "

The two figures are shown below:

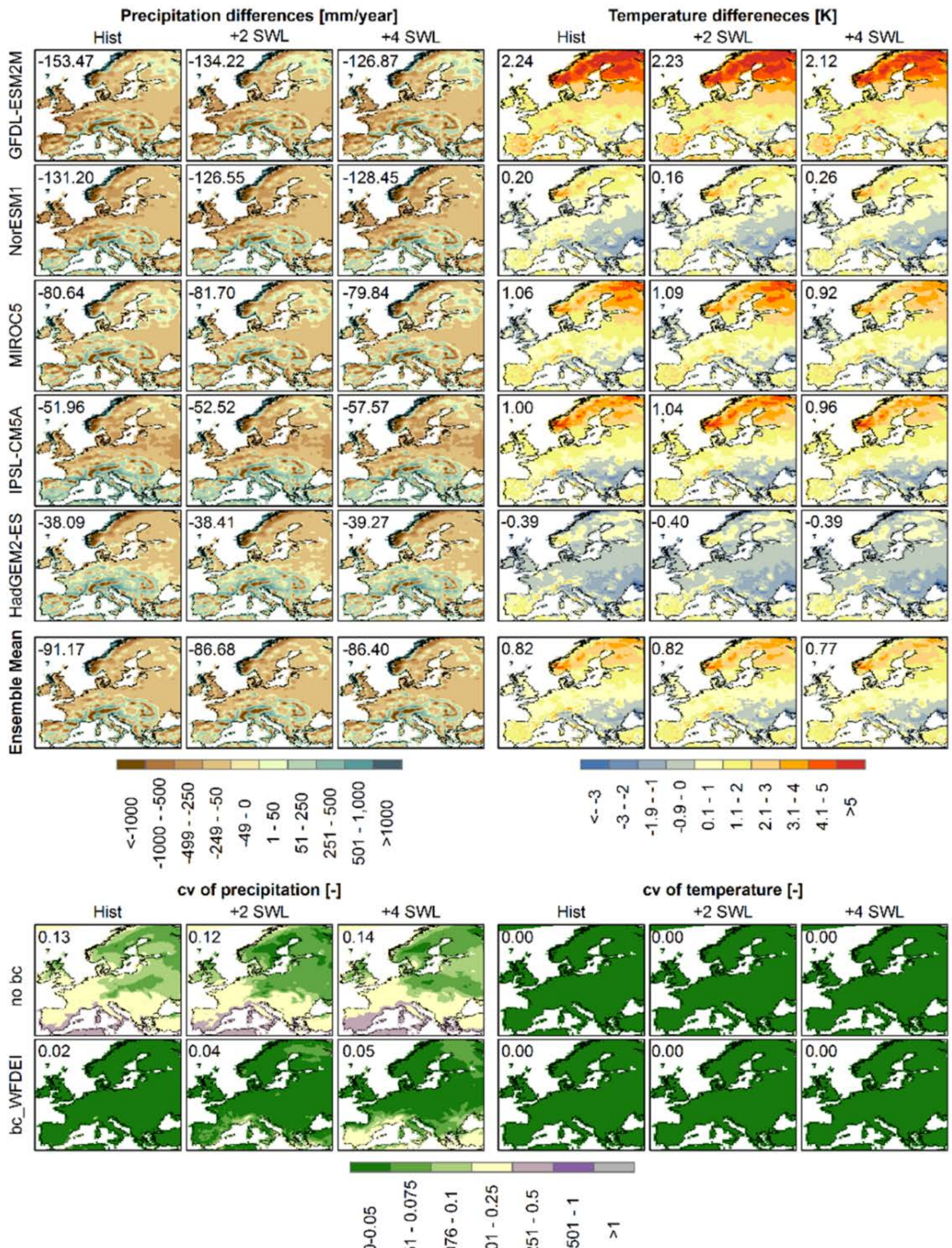


Figure S1. Absolute differences between Euro-CORDEX data bias adjusted against the WFDEI dataset and raw Euro-CORDEX data, for the variables of precipitation (right block) and temperature (left block). Differences are calculated from the historical (1976-2005), +2 SWL and +4 SWL time-slice averages, for all dynamical downscaled GCMs and their ensemble mean. Bottom block: Coefficient of variation between the ensemble members, for raw and bias corrected against the WFDEI dataset precipitation and temperature forcing variables, for the historical, +2 SWL and +4 SWL time-slices. The average value for the pan-European area is shown in each sub-figure.

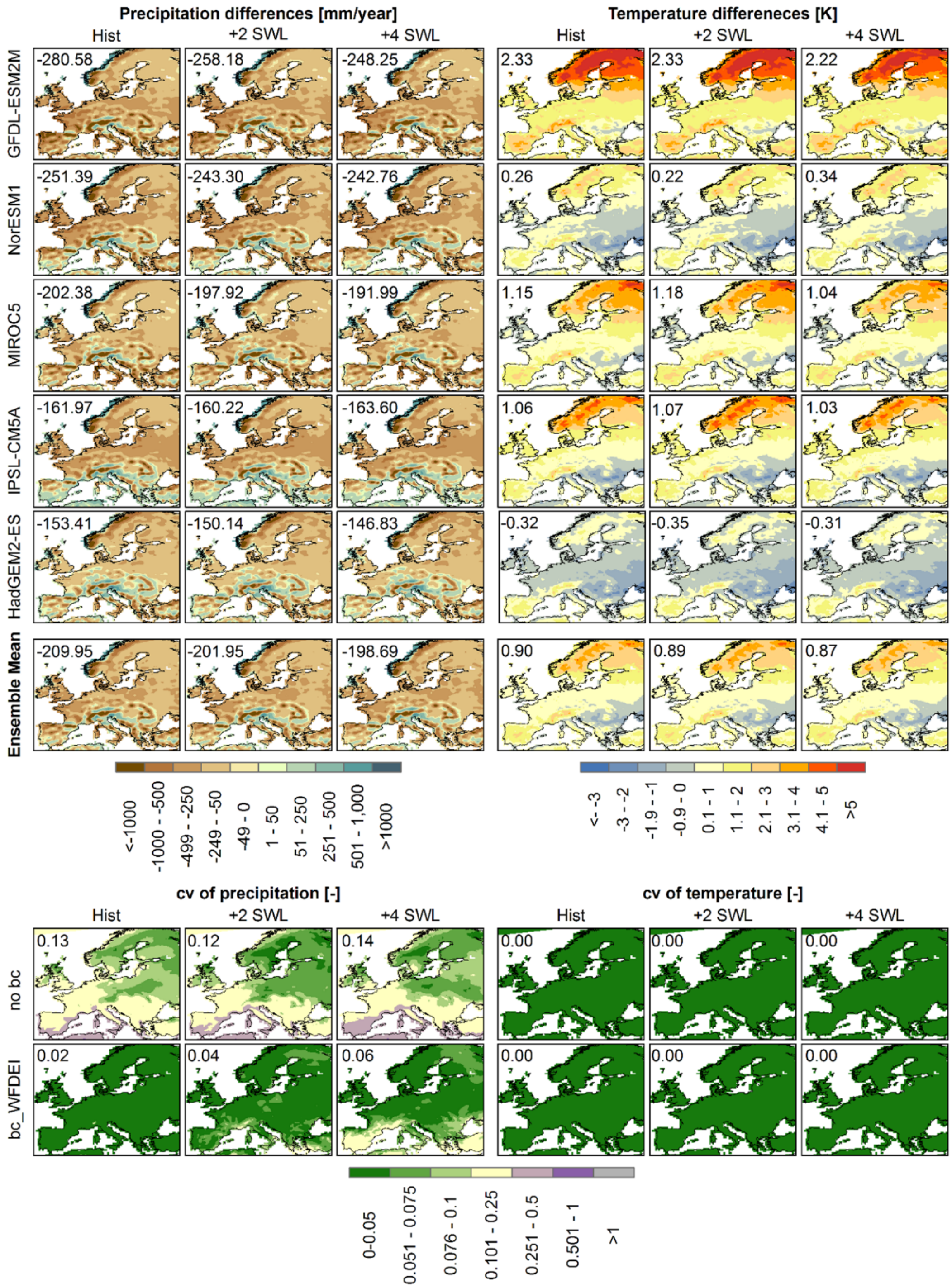


Figure S2. Absolute differences between Euro-CORDEX data bias adjusted against the E-OBS dataset and raw Euro-CORDEX data, for the variables of precipitation (right block) and temperature (left block). Differences are calculated from the historical (1976-2005), +2 SWL and +4 SWL time-slice averages, for all dynamical downscaled GCMs and their ensemble mean. Bottom block: Coefficient of variation between the ensemble members, for raw and bias corrected against the E-OBS dataset precipitation and temperature forcing variables, for the historical, +2 SWL and +4 SWL time-slices. The average value for the pan-European area is shown in each sub-figure.

In section 2.2 I miss a paragraph on the hydrological performance of JULES over Europe from previous studies. How well does it perform wrt discharge (average, high and especially low flows)? And then in the discussion 4.1 what does that imply for the results of the present paper?

REPLY: A piece on the hydrological performance of JULES over Europe has been added to the description of the model section in Section 2 (Data & Methods).
"Other studies give insight into the hydrological performance of JULES specifically. Blyth et al. (2011) extensively evaluated the JULES model for its ability to capture observed fluxes of water and carbon. Concerning discharge, their findings suggest that for the European region seasonality is captured well by the model. For temperate regions (like most of central Europe) to model exhibited a tendency towards underestimating river flows due to overestimation of evapotranspiration. Prudhomme et al. (2011) assessed JULES' ability in simulating past hydrological events over Europe. In general terms the model was found to capture the timing of major drought events and periods with no large-scale droughts present were also well reproduced. The model showed a positive drought duration bias, more profoundly present in northwest Spain and East Germany-Czech Repuplic. Prudhomme et al. (2011) argue that this feature is related to overestimation of evaporation by the model. For regions where droughts tend to last longer, JULES exhibited a better ability of reproducing the drought events' characteristics. Gudmundsson et al. (2012) compared nine large scale hydrological models, and their ensemble mean, based on their skill in simulating the interannual variability of observed runoff percentiles in Europe. According to the overall performance (accounting for all examined percentiles and evaluation metrics), JULES was ranked third best out of the 10 models, after the multi-model ensemble mean and the GWAVA model. For low and moderately low flows, expressed as 5th and 25th percentile respectively, JULES is also in the top three models regarding the representation of interannual variability in runoff. In the study of Gudmundsson et al. (2012b), where an ensemble of hydrological models is evaluated for their ability to capture seasonal runoff climatology in three different hydroclimatic regime classes in Europe, JULES exhibits a good performance, comparable to that of the best performing multi-model ensemble mean. In other studies employing multi-model ensembles, focusing on the whole European region (Gudmundsson and Seneviratne, 2015) or a single basin in Europe (Harding et al., 2014; Weedon et al., 2015) JULES' simulations also correspond with these of the other models."

There is redundancy between fig 2 and 4, and 3 and 5 respectively. Can be reduced in discussion with the technical editor perhaps? E.g., adding perhaps one column each in figures 2 and 3 with the CV of absolute change only.

REPLY: Following the reviewer's indication Figures 2 -5 have been restructured. Figure 2 was merged with Figure 4 and Figure 3 with Figure 5. Moreover, the CV for the projected period was substituted with the CV of the absolute differences. Finally, a sub-figure showing model agreement towards a wetter change in the projected time-slice has been added to the two new Figures. The new version of the Figures is shown below.

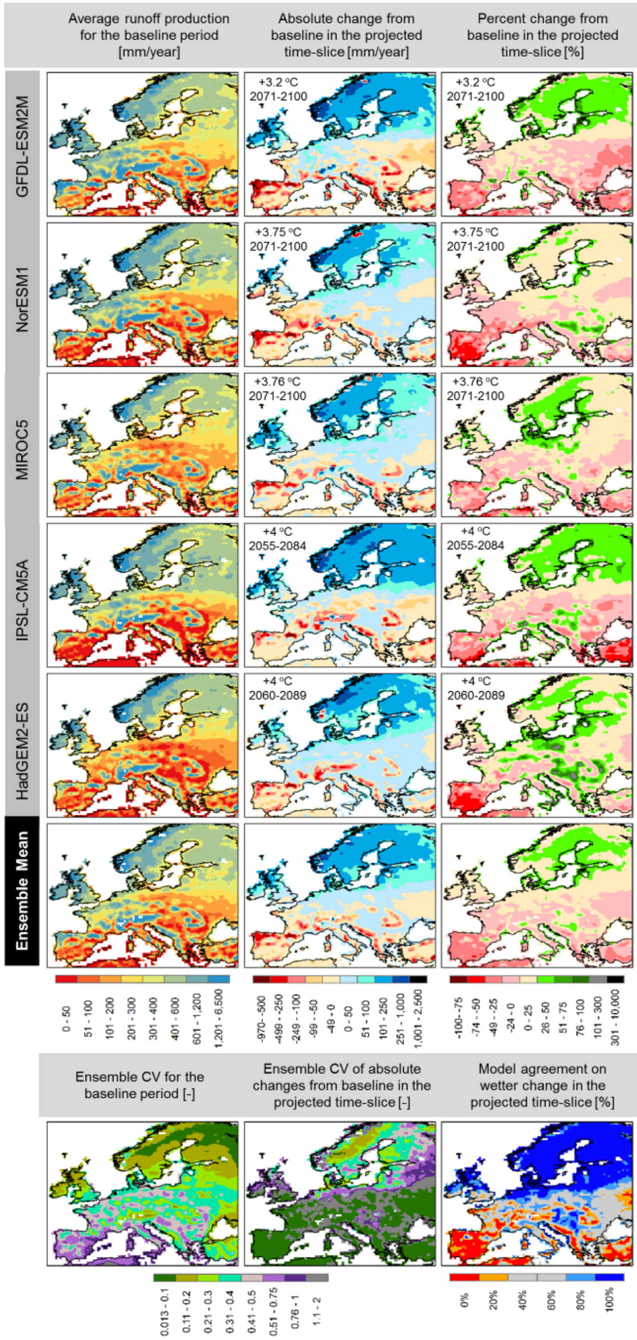


Figure 1. Average runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. Runoff production averaged over the baseline period (1976-2005) (left column), absolute change in runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

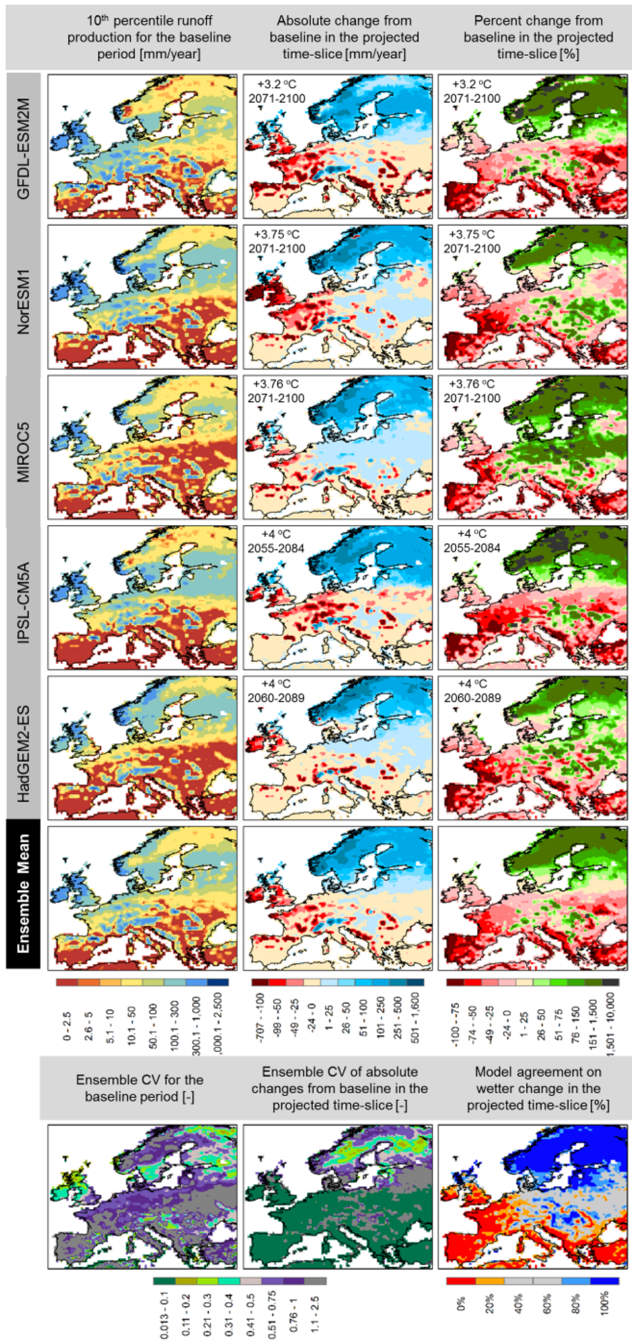


Figure 2. 10th percentile of runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. 10th percentile runoff production derived on an annual basis and averaged over the baseline period (1976–2005), absolute change in 10th percentile runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

Figure 9 second block is wrong: should be Guadiana and Elbe instead of a repetition of Rhine and Danube

REPLY: This mistake has been eliminated in this revised version of the manuscript.

14. Are the number and quality of references appropriate? YES, though the number of refs is on the high side

15. Is the amount and quality of supplementary material appropriate? Since the paper has a (perhaps secondary) focus on the effect of bias correction and even the use of different reference sets in these, I would like to see more information on initial biases of the 5 models with respect to the 2 ref sets. For precip in the paper itself, for the other variables in the supplementary material. Also the change signal for the forcing data, at least for precip, should be presented, e.g. in maps

REPLY: The added Figures S1 and S2 tackle the issue of the differences between the observational datasets and initial biases. Apart from these two figures, more additions have been made to the ESM to support our findings.

In the supplement pdf the reviewer asks: "why span ECS range if you only look at 2 and 4 degree warming periods?"

REPLY: The following was added to Section 2.1 (description of climate data):

"Using the SWL concept constitutes the results independent of the timing that the warming occurs. Although by definition of the SWL, the models reach the same level of warming in their time-slices, the different model sensitivity reflects on the evolution of temperature in the time-slice, as more sensitive models are expected to have higher rates of changes in the period before and after a specific SWL is achieved compared to the less sensitive models. Moreover, considering models of different ECS is important to express the range of other than temperature forcing variables produced by the GCMs (eg. radiation). "

The authors are grateful to Reviewer #2 for sharing his concerns with the authors of this submission. Your concerns were carefully taken into consideration for the manuscript revision, leading to a more consistent and scientifically sound manuscript. In the following pages we present a point by point reply to your general and specific/technical comments.

General comments:

1. The title of the manuscript does not reflect the general topic of the manuscript. Also within the main body of text, the authors mix up the definition of water stress (which is a function of demand versus supply) with the definition of hydrological drought/low flows. I suggest the authors to stick to the definition of average and low flows/runoff and hydrological drought throughout the manuscript.

REPLY: Following the reviewer's indication, the title was changed to better correspond with the topic of the manuscript. The new title is "High-end climate change impact on European runoff and low flows. Exploring the effects of forcing biases". Moreover, references on water stress were eliminated from the manuscript.

2. The introduction section is too long and consists of redundant information. I takes too long for the reader reaches the main goals of the manuscript. Please remove redundant text and restructure this section please, specifically related to: a. Flooding (not studied in this paper: can be removed, also in the discussion section) b. Comparison of all the GCMs/LSMs: keep it short and focus on results for the EU continent. Do not present everything here, might be more appropriate in the discussion section. c. Do not mix up the concepts of drought and water stress: focus on drought also when referring to literature in this section. d. The model JULES is now explained both in the introduction and in the methods section. I would suggest to replace the majority of this piece of text to the methods section. Only briefly mention JULES in one/two sentences in the introduction. e. Use consistent namings: e.g. when referring to the climate change scenarios +2/+4 degrees global warming. f. Please rephrase the research goals. The goals in itself are fine but they could be defined more precisely.

REPLY: The reviewer's suggestions were carefully taken into account into the revised version of the Introduction. The Introduction now is shorter and more focused on the topics of this study as redundant information have been removed. Specific changes made per paragraph are:

- Paragraph 4 was deleted as its content (added stresses of population growth and human activities on climate change impact) is not relevant with this study. A reference on the significance of this topic is made in Paragraph 1 of the Introduction.
- Paragraph 5 on multi-model assessments was significantly shortened and moved to the end of Paragraph 3
- Paragraph 9 was removed from the manuscript as it was repeating the information of Paragraph 8.
- As the content of Paragraph 10 (discussion on other LSMs) was not very relevant to the context of this study, only a short definition of the LSMs has been kept of this paragraph. This is now merged into a paragraph where the JULES model is briefly mentioned.
- The discussion about JULES in Paragraph 11 was removed from the Introduction section and was moved to the Data and Methods section (merged with Section 2.2.)

- The research objectives in the last paragraph of the Introduction have been reformulated as follows:
- “The research objectives set by this study are the following:
  i) To identify changes posed on the hydrological cycle (mean state and lower extremes) at +4 oC global warming compared to a baseline situation, and relative to the target of 2 ºC warming.
  ii) To analyse the effect of bias correction on projected hydrological simulations. To achieve this, both raw and bias corrected Euro-CORDEX data were used as input forcing in the impact model.
  iii) To assess the effect of the observational dataset used for bias correction.
  iv) To identify climate change induced changes in drought climatology at the basin scale.”

3. One of the goals is to evaluate the average runoff and low flows under a +2 and +4 degrees global warming scenario. I would suggest therefore to add to table 1 the years in which a +2 degree global warming is reached for each of the GCMs.

REPLY: The time-slices centering at +2 degrees of warming are added to Table 1. The new version of Table 1 is:

Table 1. Euro-CORDEX climate scenarios used to force JULES.

	GCM	+2 SWL time-slice	Exceeded warming level (ºC) in the +2 SWL time-slice	+4 SWL time-slice	Exceeded warming level (ºC) in the +4 SWL time-slice	Equilibrium Climate Sensitivity (K)
1	GFDL-ESM2M	2040-2069	2	2071-2100	3.2	2.44
2	NorESM1	2036-2065	2	2071-2100	3.75	2.80
3	MIROC5	2037-2066	2	2071-2100	3.76	2.72
4	IPSL-CM5A	2018-2047	2	2055-2084	4	4.13
5	HadGEM2-ES	2024-2053	2	2060-2089	4	4.59

4. Only 2 of 5 models reach a 4 degree global warming before 2100, can we really speak of a 4 degree global warming scenario then? And is it fair to compare the output of these GCM modelling results with each other or to estimate and ensemble-mean value? Please elaborate.

REPLY: The main reason for not having models capturing +4 degrees, was data availability. For the models that did not reach +4 SWL, data were not available after 2100 so as to extend the analysis. For the 5 available models, the average exceeding warming level is 3.74 ^{0}C. We considered that this is close enough to +4 to examine the models as an ensemble at +4 SWL. In the manuscript it is stated that the term “+ 4 SWL time-slice” will be adopted for all models for reasons of consistency. The authors believe that it is fair to compare the output and derive ensemble mean of the 5 models, as it is very typical for climate change assessment studies to extract ensemble means using time-slices (eg 2071-2100), without considering how different the temperature projections are for each ensemble member in that period.

5. The +2 and + 4 degrees refer to a 'global warming condition' whereas this study looks specifically to the European conditions. Could you elaborate a bit more on the temperature differences (and differences in precipitation accordingly) between the GCMs for the European continent when using the +2/+4 global warming scenario time slices? How could these differences influence your analysis/results?

REPLY: In the Introduction section there is a comment on the way +2 degrees global warming reflects on Europe: "The effect of a 2 ^{o}C global warming for the European climate was examined by Vautard et al. (2014). The study revealed that warming in Europe is projected to be higher than the global average of 2 ^{o}C".

To elaborate more on this, the following piece of text was added:

"Temperature increases of up to 3 ^{o}C were found for the winter season over north-western Europe and for the summer months over sourthern Europe. Heavy precipitation was found to increase over the whole continent for all seasons, with the exception of southern Europe during summer."

And on the effect of +4 ^{o}C in Europe: "The +4 ^{o}C global warming scenario is also translated in more intense temperature increases in Europe, especially for the summer season (World Bank, 2014)."

6. In this study, only the JULES model is being used for hydrological simulations. I would suggest the authors to elaborate a bit more on the performance of JULES compared to other models, both in the baseline situation and given the future simulations. Moreover, it would be good to show/discuss how well the JULES model matches observational data, preferably with a focus on the pan-European continent.

REPLY: A piece on the hydrological performance of JULES over Europe has been added to the description of the model section in Section 2 (Data & Methods).

"Other studies give insight into the hydrological performance of JULES specifically. Blyth et al. (2011) extensively evaluated the JULES model for its ability to capture observed fluxes of water and carbon. Concerning discharge, their findings suggest that for the European region seasonality is captured well by the model. For temperate regions (like most of central Europe) to model exhibited a tendency towards underestimating river flows due to overestimation of evapotranspiration. Prudhomme et al. (2011) assessed JULES' ability in simulating past hydrological events over Europe. In general terms the model was found to capture the timing of major drought events and periods with no large-scale droughts present were also well reproduced. The model showed a positive drought duration bias, more profoundly present in northwest Spain and East Germany-Czech Repuplic. Prudhomme et al. (2011) argue that this feature is related to overestimation of evaporation by the model. For regions where droughts tend to last longer, JULES exhibited a better ability of reproducing the drought events' characteristics. Gudmundsson et al. (2012) compared nine large scale hydrological models, and their ensemble mean, based on their skill in simulating the interannual variability of observed runoff percentiles in Europe. According to the overall performance (accounting for all examined percentiles

and evaluation metrics), JULES was ranked third best out of the 10 models, after the multi-model ensemble mean and the GWAVA model. For low and moderately low flows, expressed as 5th and 25th percentile respectively, JULES is also in the top three models regarding the representation of interannual variability in runoff. In the study of Gudmundsson et al. (2012b), where an ensemble of hydrological models is evaluated for their ability to capture seasonal runoff climatology in three different hydroclimatic regime classes in Europe, JULES exhibits a good performance, comparable to that of the best performing multi-model ensemble mean. In other studies employing multi-model ensembles, focusing on the whole European region (Gudmundsson and Seneviratne, 2015) or a single basin in Europe (Harding et al., 2014; Weedon et al., 2015) JULES' simulations also correspond with these of the other models."

7. The authors used two hydrological indicators to identify changing climate trends, the average and 10th percentile of runoff production. Reading the manuscript, it did not became clear to me however how the authors applied these indicators. Did they use monthly or yearly values? And if they used monthly values, did they used a variable threshold approach to estimate the 10th percentile values? Or did they use a fixed 10th value over all months? Please clarify.

REPLY: Average and 10th percentile runoff production were deduced from monthly runoff data. For the analysis of the gridded results with the SWL time-slice approach, each indicator was computed from the monthly values of all years in the time-slice.
For the analysis of basin averaged runoff regime, the two hydrologic indicators were calculated per year, for all the years of the simulation. This resulted in time-series of basin aggregated average and 10th percentile runoff production, spanning from 1971 to 2100. Clarifications for the derivation of the hydrological indicators have been added to the manuscript:
"The two hydrological indicators were deduced from monthly runoff data. For the analysis of the gridded results at pan-European scale with the SWL time-slice approach, each indicator was computed from the monthly values of all years in the time-slice. For the analysis of basin aggregated runoff regime, the two hydrologic indicators were calculated per year, for all the years of the simulation. This resulted in time-series of basin aggregated average and 10th percentile runoff production, spanning from 1971 to 2100."
Note that a different approach was used in defining the drought threshold (described in current Section 2.5, which was 2.4 in the first manuscript). For this, daily values of discharge were used, and a daily varying threshold computed based on the values of the historical period was applied to the 1971-2100 discharge time-series.

8a. With respect to the examination of drought climatology the authors explain that they 'counted the number of days per year that extreme lows in flow occur'. First, using a 10th percentile value is not yet really an extreme low (I would say extreme lows would be using a 5th percentile or 1st percentile).

REPLY: The limit set for defining low flows was defined considering its relevance for drought formation. The 10th percentile value has been used in numerous studies as a limit for identifying drought conditions (Hannaford et al., 2011; Prudhomme et al., 2011, 2014; Roudier et al., 2015; Stahl, 2001). As the use of the word "extreme" mostly refers to lower percentiles, we substituted this word in the manuscript with "particular": "changes in drought climatology, i.e. the number of days per year that particular lows in flow occur."

8b. Secondly, I'm wondering whether the authors used any buffering methods (defining minimum/maximum inter-event times or minimum length of dry conditions in order to be considered a drought) to estimate the drought climatology, see for example Tallaksen et al. (1997) and Sung and Chung (2014). Thirdly, could the authors argue why they used a 'total number of days per year with extreme lows', rather than a 'total/max number of consecutive days with extreme lows', which might be a more appropriate indicator for drought climatology? Finally, Wander et al. (2015) argue that – when evaluating drought conditions under climate change- it is better to use a transient variable threshold approach as (aquatic and terrestrial) ecosystems are able to adapt to changing drought conditions. I would suggest the authors to at least discuss the use of this transient variable threshold here or in the discussion section.

REPLY: We did not use a method to buffer the discharge time-series before comparing them with the daily varying drought threshold. The authors acknowledge the different methods that the reviewer refers to, although we have made different choices for our analysis. Comments on the different options (not followed in this study) for drought identification are added to the discussion section:
"Our analysis of drought climatology at the basin scale was based on the total number of days under a predefined daily varying drought threshold. We did not employ any buffering criterion for the days under threshold to be accounted for in the total sum (as discussed for example by (Sung and Chung, (2014) and (Tallaksen et al., (1997)). The use of such a criterion would have decreased the calculated dry days. However, as the interpretation of the results of this study is mostly oriented in identifying trends of change rather than absolute numbers describing the future regime, the lack of a buffering criterion is not supposed to notably affect the extracted conclusions. (Wanders et al., (2015) employed a transient variable threshold for the assessment of the drought conditions under climate change, considering a gradual adaptation of the ecosystem on the altered hydrological regime. This is an interesting alternative, especially for climate change mitigation and adaptation studies. In our study we aimed to identify global warming induced changes in the future hydrological state without considering adaptation, thus the same historically derived threshold was applied to the whole length of the simulated runoff time-series."

9. Please mention that the results presented in section 3.1 refer to the use of the Euro-CORDEX data without bias-correction.

REPLY: "raw" was added before the "Euro-Cordex forcing data" to indicate that these data have not undergone any bias adjustment. This was also added to the caption of the relative figures.

10. Section 3 is quite wordy about (significant) positive and negative trends to describe the changes in average and 10th percentile runoff values towards future conditions. However, only few statements are actually backed up with numbers/statistics. I would suggest to execute some extra statistical analyses to give your results some extra body. E.g. the trend observations could easily be backed up with a simple regression analysis giving a 'number' to the trend (coefficient) and a feeling of (in)significance of the *result (R-squared and significance level of the estimated coefficient in the regression). Outcomes could be mentioned briefly in text – e.g. between brackets- and in the figures.*

REPLY: Following the reviewer's suggestion, a regression analysis was performed with the annual time-series of basin aggregated runoff production and total number of drought days per year. The analysis employed a linear regression model for trend estimation and the trend significance was tested at the 0.05 significance level. Relevant information have been added to Section 2.4: "The trend of the annual time-series was investigated employing a linear regression analysis to estimate the sign and the average rate of the trend. The significance of the trend was tested at the 95% confidence interval via a Student-t test."
Trend and p values have been added to the relevant figures and tables of summary statistics of the regression analysis have been added to the ESM (Table S1,S2,S3).

11. I miss in the discussion section a piece of text elaborating on the use of JULES, the performance of JULES and the potential use of other GHMs/LSMs.

REPLY: The piece of text added in Section 2, gives information on JULES' hydrological performance, with focus on the European region. It also provides information on the performance of JULES compared to other GHMs/LSMs. At the discussion section, a piece of text about the use of other GCMs and the uncertainty stemming from the use of a single impact model has been added:
"In our study only one impact model (JULES) was used. Hagemann et al. (2013) argue that impact model induced uncertainty in future hydrological simulations is larger than that of the GCMS for some regions of the land surface and suggest using multi-impact model ensembles to deal with this issue. However useful conclusions can be drawn also from studies employing a single GHM/LSM. Examples of such single model climate change impact assessments performed recently are the studies of Schneider et al. (2013) and Laizé et al. (2013) with the WaterGAP GHM, the studies of Arnell and Gosling (2013), Gosling and Arnell (2013) and Arnell et al. (2013) with the GHM MacPDM and of Hanasaki et al. (2010) using the H08 LSM."

Specific/Technical comments:

12. 'GFDL and NorESM1 exhibiting generally wetter patterns' (p 7281, line 8-9). Looking at figure 2 I would say that NorESM1 is also generating relatively dry patterns for southern Europe.

REPLY: The reviewer's observation is correct. The sentence was intended to refer to the wetter patterns GFDL and NorESM1 show in the northern regions but it was confusingly structured. This sentence has now changed accordingly: *"GFDL and NorESM1 exhibiting generally wetter patterns for northern Europe and Scandinavian Peninsula, and with IPSL describing drier patterns, especially for southern Europe."*

13. 'all models agree' (p 7281, line 11 & 20): I would suggest to add a sub-figure to figure 4 & 5 that shows the modelling agreement with the ensemble mean in terms of estimated change (+/-) in average/10th percentile runoff.

REPLY: As the reviewer suggested, a sub-figure showing model agreement has been added to the revised version of the figures. The model agreement here is defined as agreement towards a wetter change in the projected time-slice. Thus 100% agreement means that all the five models project more runoff than in the baseline period while 0% agreement means that all five models show a drier response in the projected time-slice. The revised Figures are shown below:

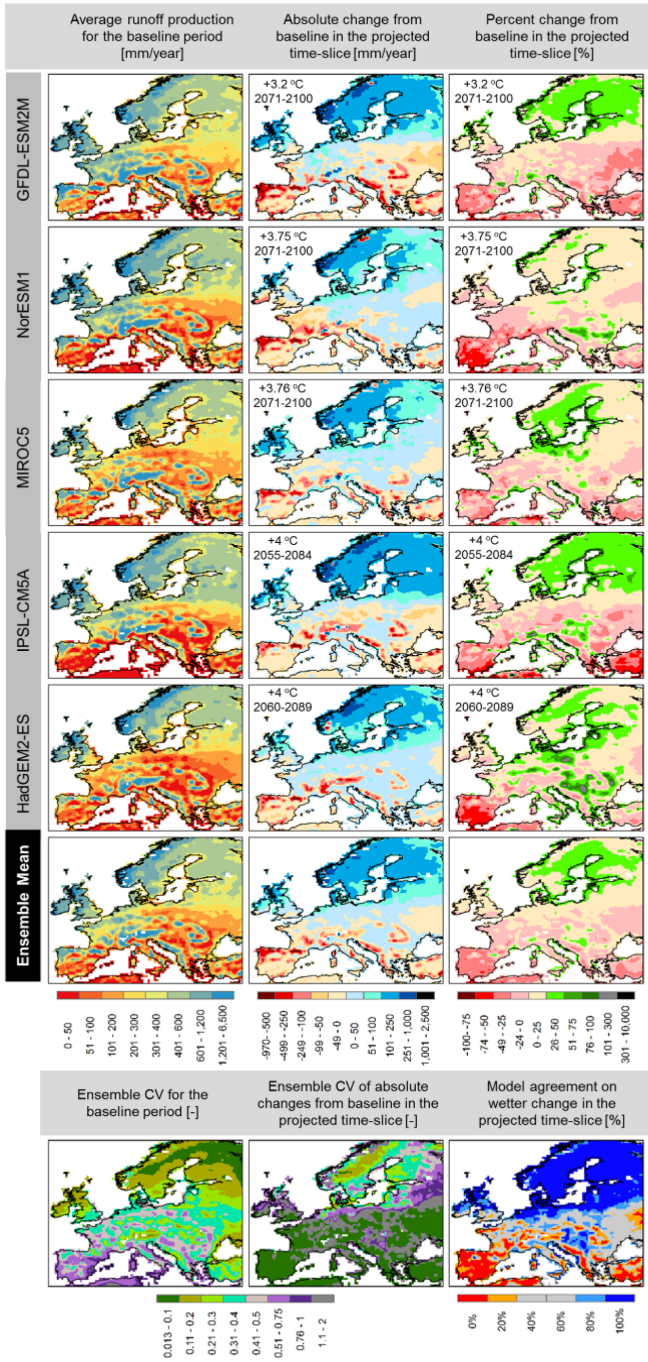


Figure 2. Average runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. Runoff production averaged over the baseline period (1976-2005) (left column), absolute change in runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

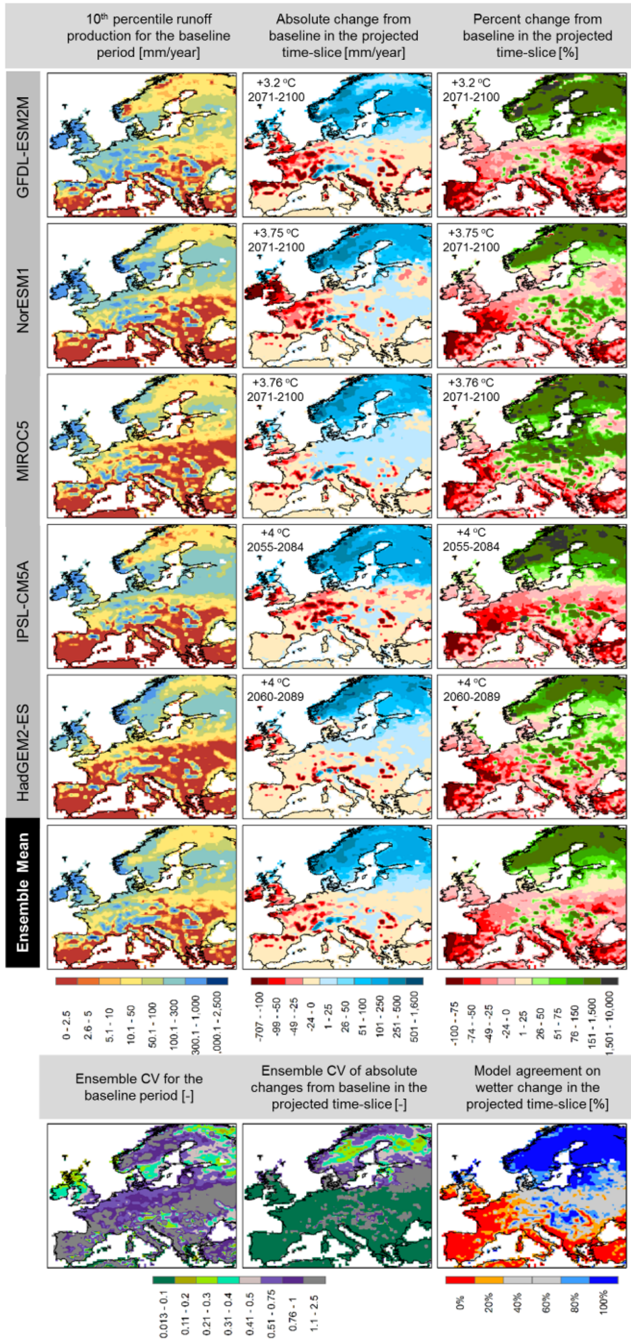


Figure 3. 10th percentile of runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. 10th percentile runoff production derived on an annual basis and averaged over the baseline period (1976–2005), absolute change in 10th percentile runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

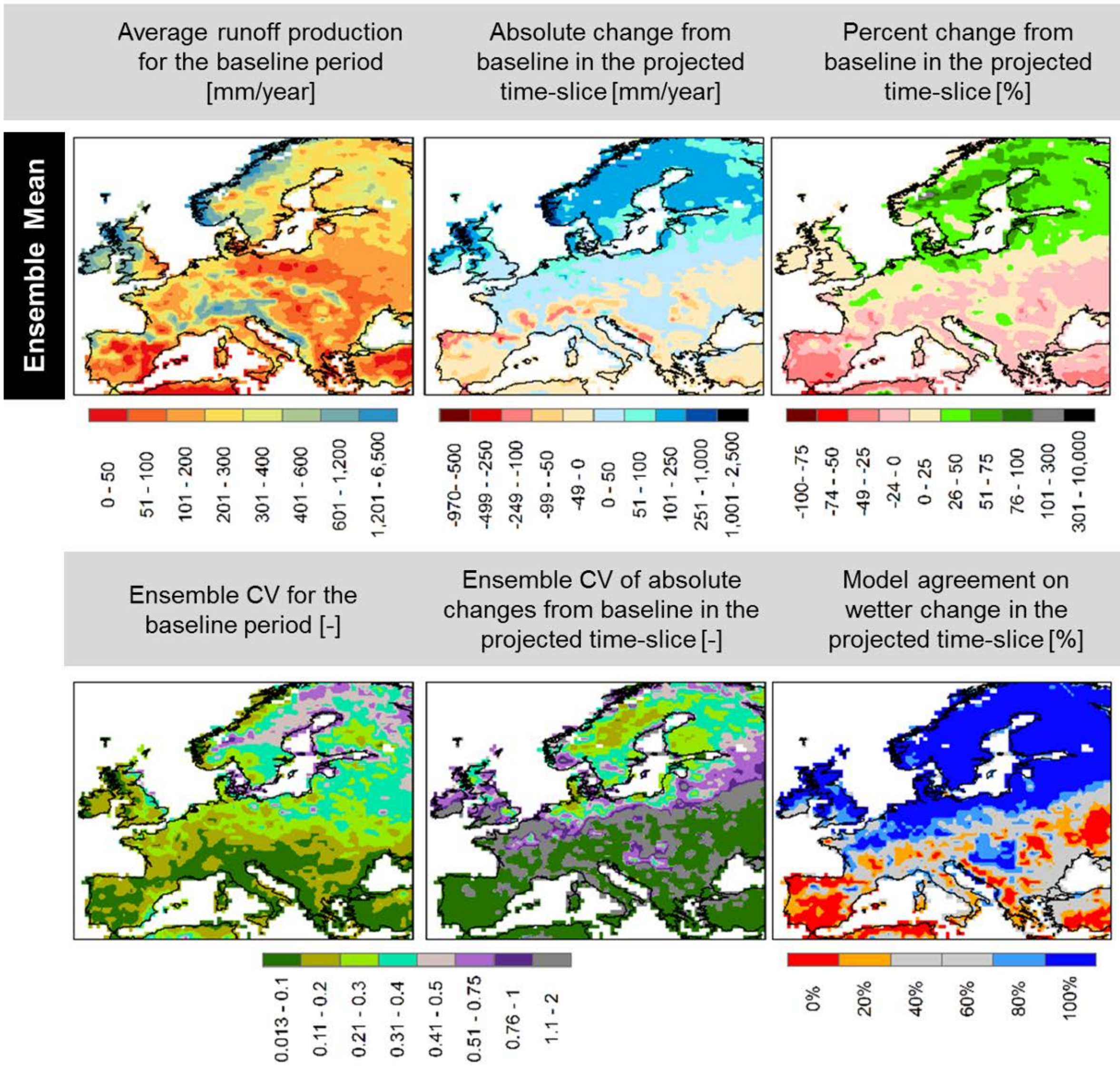


Figure 4. Ensemble mean of average runoff production from Euro-CORDEX data bias adjusted against the WFDEI dataset. Top row: Runoff production averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice. Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

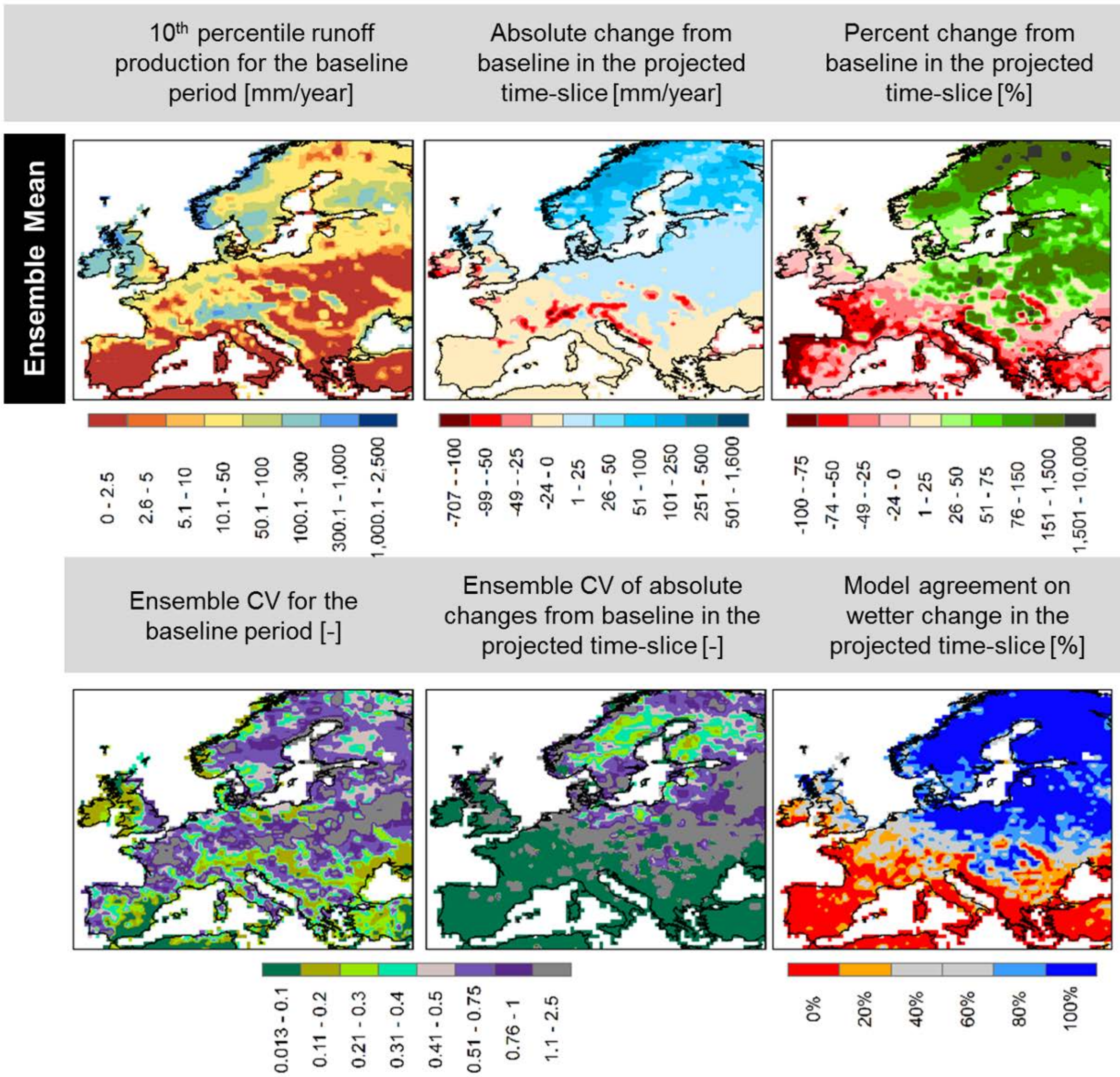


Figure 5. Ensemble mean of 10th percentile runoff production from Euro-CORDEX data bias adjusted against the WFDEI dataset. Top row: 10th percentile runoff production derived on an annual basis averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice. Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

14. 'with MIROC5 being the only ensemble member that expands this wetter climate down to central Europe': Looking at figure 3 I would argue that HadGEM2-ES is also showing some significant expansion of this wetter climate down to central Europe.

REPLY: The reviewer's observation is correct. MICOC5 is distinguished in the fact that has a stronger spatial coherence in the expansion of the wetter changes towards the south and also higher values of change. In the revised version this sentence has been modified to include the HadGEM2 model: "…with MIROC5 and HadGEM2 being the two ensemble member that expand this wetter climate down to central Europe".

15. 'Thus averaging … projected changes' (p 7282, line 2-3): Isn't this always the case with taking an average ensemble-mean?

REPLY: It is. In the revised version this sentence is again used but as an observation rather than as a finding: "Concerning the ensemble mean, smoothing of the projected changes due to averaging has revealed clear patterns of change…".

16. 'making it easier to identify clear patterns of change' (p 7282, line 3): Is this really the case? I would argue that an ensemble-mean might be useful but that it could also create pseudo-results (the average value is not per se the true value namely), therefore it is important to consider the full spread of GCM-forced outcomes as plausible results (unless you have information on the reliability of the different GCM-forcings).

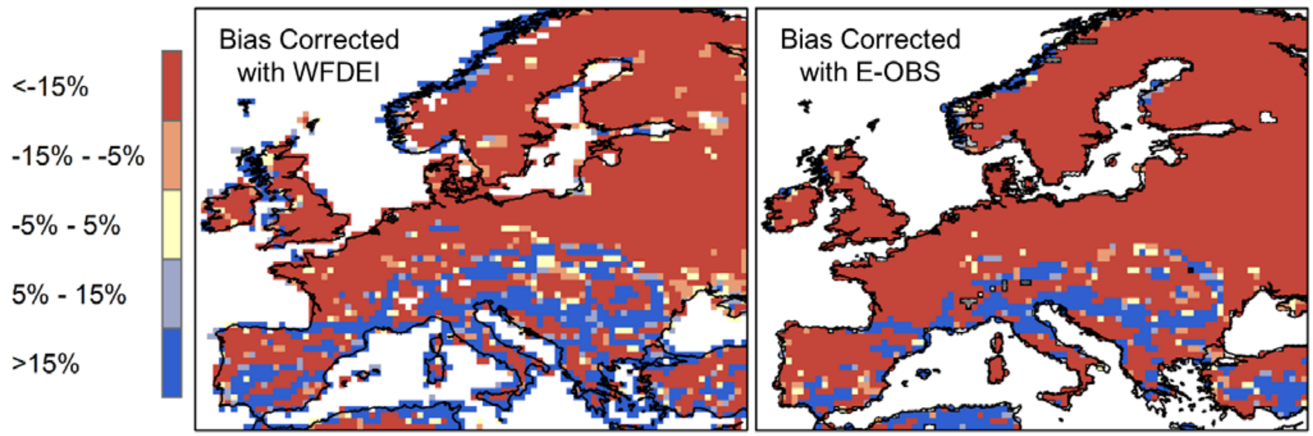
REPLY: The reviewer's comment is very legitimate. The sentence has been changed to reflect that ensemble-mean results can be misleading if they are not interpreted in conjunction with the ensemble members' agreement. Here we tackle the model agreement issue by examining both the cv of the changes (gives us information on "how much the models agree") and the agreement towards wetter of drier response ("what they models agree on"). The changed sentence is the following: "Concerning the ensemble mean, smoothing of the projected changes due to averaging has revealed clear patterns of change, which however have to be interpreted considering the full spread of the GCM-forced outcomes and the agreement between them in order to avoid misguided conclusions."

17. 'For 10th percentile … part of Europe' (p 7282, line 16-19): Could you argue why this is the case?

REPLY: This could be explained by the cv calculation formula (standard deviation/mean). For 10th percentile runoff both the mean and standard deviation are lower than for average runoff. If the mean of 10th percentile runoff decreases more than the standard deviation does, the computed cv for 10th percentile runoff will be lower compared to the cv of average runoff.

18. 'bias adjustment of the forcing data resulted in a drier hydrological response from the JULES model' (p 7282, line 24-25): Could you support this statement with some numbers? E.g. xx % of the total pan-European land surface area shows a drier output using the bias-corrected forcing data compared to using the non-bias corrected forcing data, xx % shows insignificant change, xx% shows a wetter output.

REPLY: Following the reviewer's suggestion, a Figure has been added to the ESM (Figure S3), showing the difference between bias corrected and non-bias corrected output in the baseline period. Changes between -5% and 5% have been classified as insignificant, >5% as wetter output and <-5% as drier output. The percent of pan-European area each change category occupies, along with the average value of the change (in absolute and percent terms) have been calculated. Additions based on these findings have been made in the manuscript: "Bias adjustment of the forcing data resulted in a drier ensemble mean runoff for the baseline period for 70.40% of the pan-European land surface, in comparison to 26.01% of the land area that had a wetter response after bias adjustment. The remaining 3.59% of the European area had changes that were classified as insignificant". Figure S3 is shown below:



	Bias Corrected with WFDEI			Bias Corrected with E-OBS		
	Drier output	Wetter output	Insignificant change	Drier output	Wetter output	Insignificant change
Percent of pan-European land area	70.40%	26.01%	3.59%	83.62%	14.67%	1.70%
Average percent change	-44.15%	148.77%	-0.53%	-56.10%	215.33%	-0.87%
Average absolute change [mm/year]	-231.44	159.37	-2.96	-285.70	131.97	1.49

Figure S3. The effect of bias correction on the ensemble mean of average runoff production for the baseline period (1976-2005). Figures: Relative difference between the ensemble means of bias corrected (left:with WFDEI, right:with E-OBS) and raw forcing data. Differences between -5% and 5% are classified as insignificant, differences <-5% as drier output and differences >5% as wetter output after bias correction. Table: percent of land area that falls into each category of change and average of the changes.

19. 'with increases in runoff in northern Europe getting more pronounced in the runs after bias correction (p 7282-7283, line 28-1): does this hold both for the absolute and percentage change? Or only for the percentage change? And should in that case the difference be considered as a results of the baseline values becoming reduced in magnitude?

REPLY: This applies only to percent change. The reviewer's comment is right that this is probably due to baseline values getting smaller after bias correction. In order to avoid confusion, this statement was replaced with the following in the revised manuscript: "Projected changes from bias adjusted data exhibit very similar patterns and magnitudes with the raw data derived changes".

20. 'sign change' (p 7283, line 2): how about significance of the values? Doesn't it just all fall under 'insignificant change'?

REPLY: Changes are in the lower region (from the class:-25% to 0 to the class 0 to 25%). However the effect that bias correction had on the baseline mean on the same areas (Figure S3 in ESM), was not classified as insignificant (apart from very few gridboxes).

21. 'bias correction has … model agreement' (page 7283, line 3-4): That makes sense as we bias-correct all forcing data-sets using the same WFDEI data-set. Point for discussion should be whether this is actually a desired outcome (all outputs merging towards 'one single line'). How big is the confidence in the WFDEI data-set, for example?

REPLY: This is indeed a very interesting topic for discussion as in most analyses, one has to assume that the "observations" are correct (and here the WFDEI dataset serves as observations of historical precipitation and temperature). We introduce this issue in our study by comparing the effect of two different observational datasets (WFDEI and E-OBS) on projected runoff output from bias corrected data. The subject of how much the observational datasets can be trusted and how they affect uncertainty in simulations (Biemans et al., 2009), relates to model sensitivity to input forcing. In the now running phase of ISIMIP2, historical validation simulations were performed with three different observational datasets, in attempt to study this model sensitivity and quantify the resulting uncertainty in simulations.

22. 'For the baseline …southern Europe' (p 7283, line 6-8): Could you explain/clarify this?

REPLY: The cv calculated for the baseline period has in general lower values than the cv calculated for the +4 SWL projected period. This makes sense as the models were bias corrected against years that include the baseline period. This sentence however has been removed as it does not apply to the new version of the manuscript (now we do not present the cv for the projected period but the cv of the differences between projected and baseline periods).

23. 'basin average runoff production' (p 7283, line 20): Please clarify how you got these averaged values. Did you first averaged all runoff values and afterwards took the average and 10th percentile? Or did you basin-averaged all average (temporal) and 10th percentile values?

REPLY: From the gridded runoff output, basin aggregated time-series were produced (first spatial averaging). Based on these, the average and the 10^{th} percentile were calculated per year, resulting in annual time-series of average and 10^{th} percentile runoff production.

24. 'A common observation … input forcing' (p 7283,26-27): These decreases are really large. Did you find any corresponding decreases in the literature, or did you check the values with observed time-series (from gauges). Could you somehow explain these large decreases?

REPLY: These decreases are reasonable based on the new Figure S3 of the ESM (average of negative changes after bias corrections is around 44% of 231.44 mm/year). In current Figure 10 (previously Figure 12) there is comparison between observed and modelled historical runoff production. The decreases after bias correction are probably that large because the raw data produce too high values of runoff. Bias corrected data result in modelled runoff that is closer to observations compared to non-bias adjusted data.

25. Figure 9 is not correct. The figure shows twice the results for the Danube and Rhine whilst the results for the Elbe and Guadiana are missing.

REPLY: This mistake has been eliminated in this revised version of the manuscript. The correct figure is shown below (the trend and p value from the regression analysis are also added):

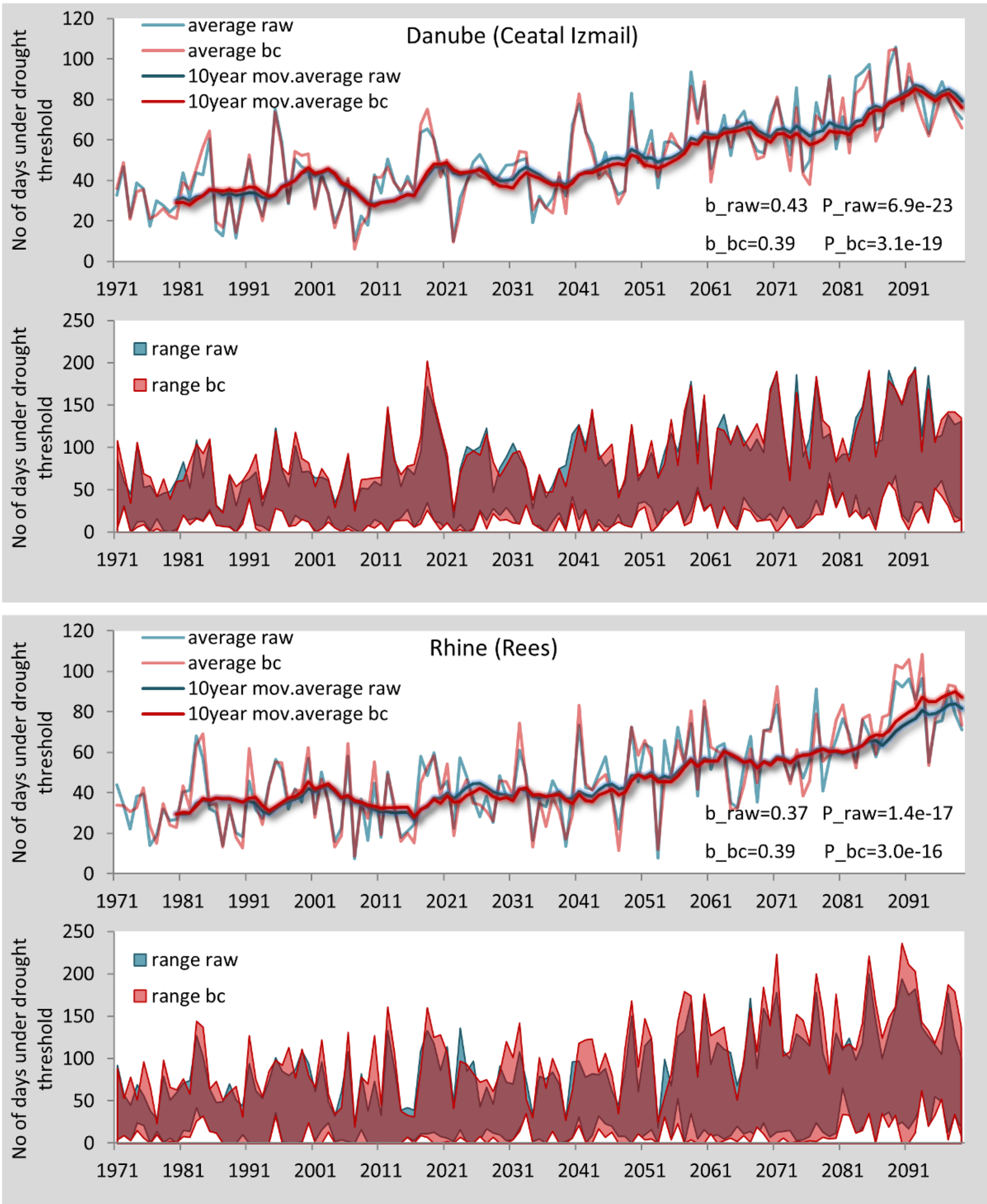


Figure 7. Number of days under drought threshold per year for raw and bias adjusted Euro-CORDEX data. Ensemble mean and 10-year moving average of the ensemble mean (top), ensemble range (bottom).

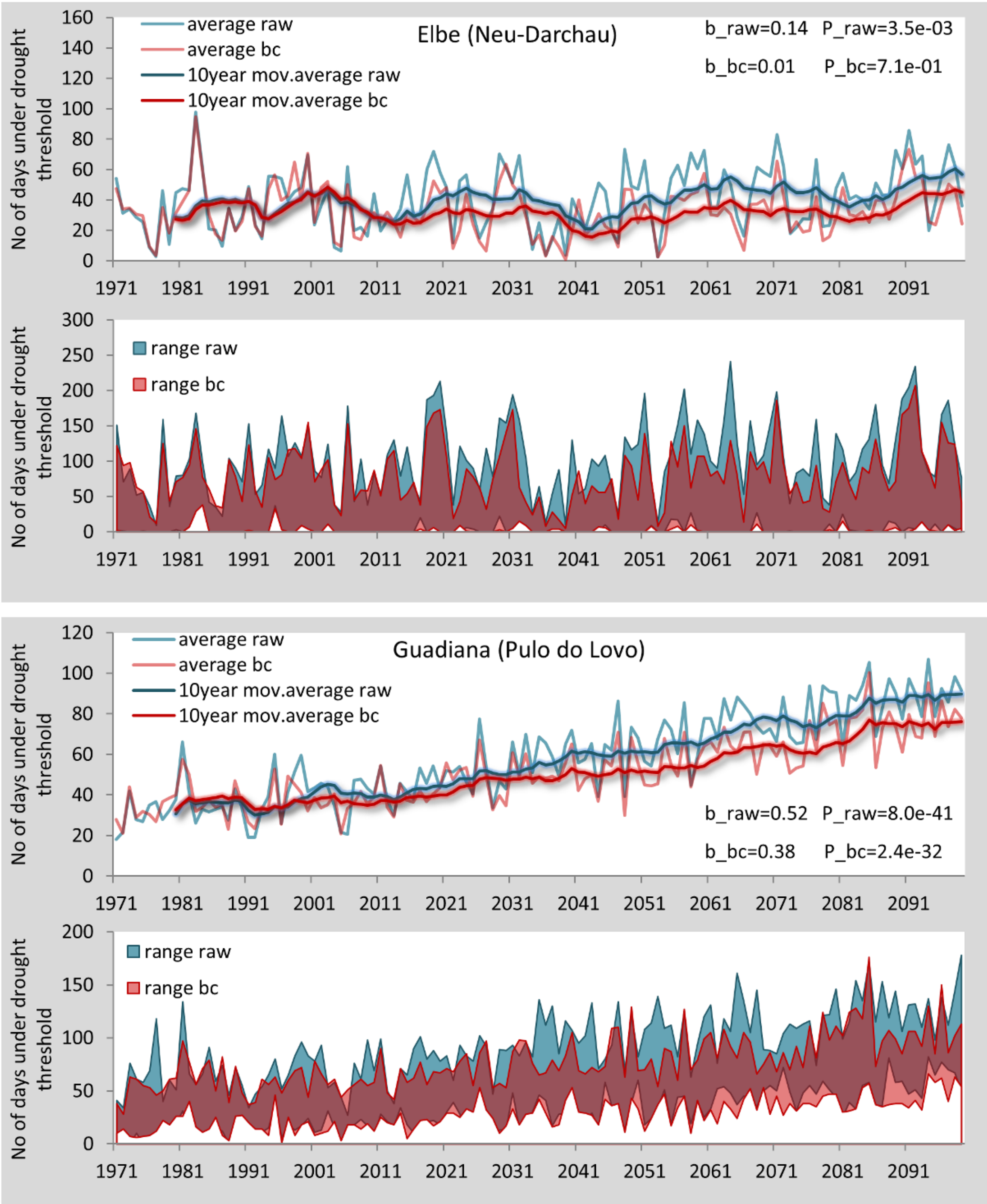


Figure 7 (continued)

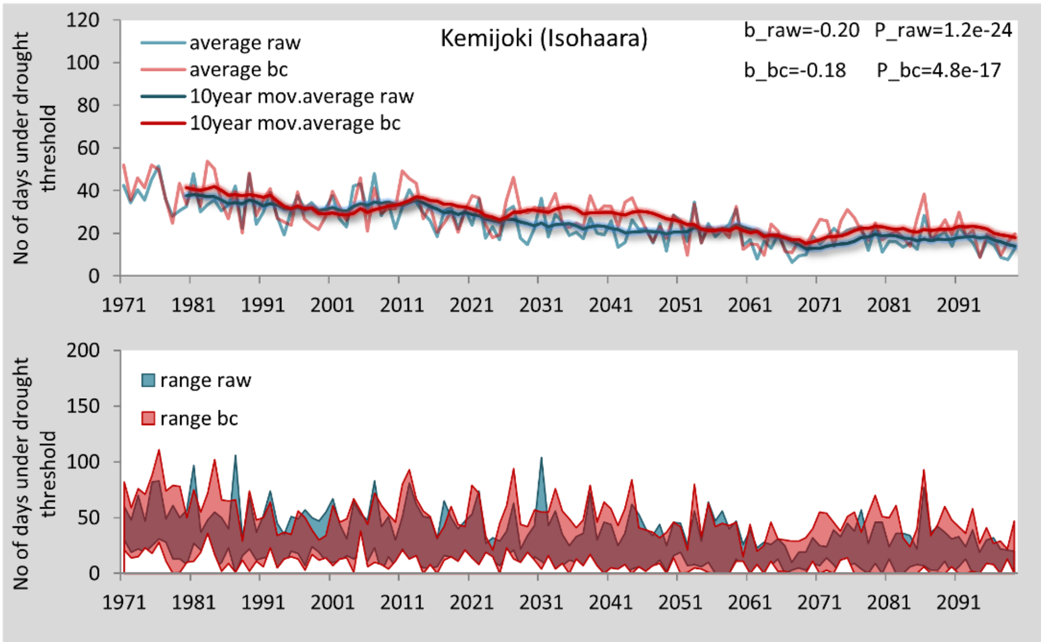


Figure 7 (continued)

26. 'the effect of climate warming is far more pronounced for the low flows', 'significantly' (p 7285, line 16-17): Is this really significantly? Did you tested this? Please use some statistical methods to support these statements.

REPLY: As this is not directly checked in this part, the statement was removed from the manuscript. However, having statistically checked the trend of the time-series from 1971 to 2100, we have a good approximation of the trend between the baseline and the +4 SWL time-slices. The scope of presenting Figure 8 (previously Figure 10) is to see how the basin aggregated runoff values change in between that previously estimated trend, i.e. check the values at +2 compared to +4 SWL. At the same time it allows us to compare the runoff values calculated by the different ensemble members and see how these deviate from the ensemble mean.

27. 'there is a significant decrease from 0 to +2 C' (7285, line 23): How did you estimated the values for 0 degrees Celsius warming as they are, following figure 10, not equal to the baseline values. Please, clarify this in the methods section.

REPLY: We did not estimated values for 0 degrees Celcius. 0 degrees correspond to the reference period of the pre-industrial state (1861-1880), set for the calculation of the SWLs. This means that by the baseline period (1976-2005) there has already been a change in temperature. In Figure 8 (previously Figure 10) we allocated the basin aggregated runoff values of the baseline period to their corresponding temperature change from pre-industrial state (this temperature change ranges from 0.3-0.5 ^{o}C between the models).

28. 'probably due to its very low values of 10th percentile runoff' (p 7286, line 7-8): In how many of the GCM runs you reach zero flow? And how reliable is zero flow for this river? Please mention this in the text.

REPLY: For all the GCMs, for the +4 SWL projected time-slice, bias corrected data give values of around 0.001 mm/year for Guadiana while for the baseline time-slice, basin aggregated 10th percentile runoff varies between 0.0015 and 0.0033 mm/year. Raw data give higher values for the baseline period (0.004-0.73 mm/year) but project values of the same order of magnitude as the bias corrected data. Only NOAA and MIROC give values higher than the other models (0.009 and 0.004 mm/year respectively). The other three give values of around 0.001 mm/year. Thus the absolute projected values for raw and bias corrected data are very close. However, the calculated percent change from baseline gives higher values for raw data, as the raw baseline values are higher than the bias corrected ones. The close to zero values of 10th percentile runoff indicate that the river exhibits intermittent flow regime. This is relevant for this particular river, as it is located in a semi-arid region and intermittent flows typically characterize its hydrological regime (Filipe et al., 2002; Collares-Pereira et al., 2000; Pires et al., 1999). Given the changes that are projected for the Iberian Peninsula at +4 SWL, it is expected that the intermittent flow regime in Guadiana might intensify.
A piece concerning this topic was added to the discussion section of the revised manuscript.

29. 'E-OBS corrected data … the observed values' (p 7287, line 11-14): Could you think of an explanation for this observation? Are there any differences between E-OBS and WFDEI that might clarify this result?

REPLY: The two newly added Figures in the ESM (Figure S1 and S2), describe the effect of bias correction on the forcing variables and help us understand the differences between the two observational datasets. (Figure S1 shows the effect of bias correcting against the WFDEI dataset and Figure S2 against the E-OBS dataset. Results are shown for precipitation and temperature as the rest of the forcing variables were not bias adjusted. The absolute differences between bias corrected and raw input (bc-raw) are shown for all the participating GCMs and for their ensemble mean. The cv between the ensemble members before and after bias correction has also been calculated. In each sub-figure, the spatial average of each illustrated map is noted in each sub-figure.)

The precipitation differences after bias correction with E-OBS are quite larger than the differences after bias correction with WFDEI (for ensemble mean the values are approximately 2.3 times bigger), meaning that E-OBS corrected precipitation has lower values than the WFDEI adjusted precipitation. It is thus reasonable that this difference in precipitation will also reflect on the output of the hydrological model, producing lower runoff values when forced with the E-OBS bias adjusted dataset.

30. Section 3.6 focuses on the basin averaged average runoff whilst I think it is (more) interesting to show also the results for the 10th percentile runoff and using not the basin-averaged numbers. Optionally figures could be placed in a supplementary.

REPLY: A Figure of the same analysis described in section 3.6 but for 10th percentile runoff has been added to the ESM (Figure S6). The following piece of text was added to Section 3.6:

"From the application of the same analysis on 10th percentile runoff production (Figure S6 of the Supplementary Material), it is deduced that for the low flows the E-OBS corrected data again produce lower values of runoff compared to WFDEI. In this case, however, even the raw forced output (which is wetter than the bias corrected) underestimates the observed 10th percentile runoff values. Regarding the percent projected changes, results from bias corrected data produce smaller values compared to the raw data while E-OBS adjusted data result in decreased changes compared to output from WFDEI adjusted forcing."

31. Could you elaborate a bit more on how the differences between the two bias corrected data-sets and their hydrological output develops, comparing the difference in input data (precipitation, temperature) with the order of magnitude differences in hydrological outputs (local and routed runoff)?

REPLY: As mentioned in the response to comment 29, the precipitation differences after bias correction with E-OBS are approximately 2.3 times larger than the differences after bias correction with WFDEI (for the ensemble mean, comparing spatially aggregated differences). For the historical period, the ensemble mean of precipitation differences for WFDEI bias corrected results is -91.17 mm/year compared to -209.95 mm/year for EOBS bias corrected results.
Using the values of Table 2 we can make a rough comparison of the effect bias correction has on input precipitation and output runoff. For WFDEI forced runs, the average of the difference between bias corrected and raw forced output for the five basins, for the historical period is around -135 mm/year. The respective number for EOBS adjusted data is -197 mm/year. Runoff differences after bias correction with E-OBS are about 1.5 times smaller than the differences after bias correction with WFDEI. Comparing this with the respective number characterizing precipitation differences (2.3) we could argue that the effect of the different observational dataset is weakened in the hydrological model output. Another comment could be that the effect of bias correction on input precipitation and output runoff are of the same order of magnitude (-91.17 mm/year vs -135 mm/year for WFDEI and -209.95 mm/year vs 197 mm/year for E-OBS).

32. 'it is … model agreement' (p 7287, line 14-15): Incomplete sentence

33. 'changed' (p 7288, line 9): change

34. 'it is … climate change' (p 7289, line 1-2): please leave out as this is only deduced and not studied.

REPLY: The flood related unsupported statement was removed from the manuscript.

35. 'It should be … average-state' (p 7289, line 4-5): Please back-up with some references.

REPLY: Similar results with increased model spread expressed as cv for low flows compared to average state flows were found by Koirala et al., (2014).

36. 'of' (p 7289, line 15): delete

37. 'remarked' (p 7291, line 25): remarkable/significant

REPLY: “Remarked” was changed to “remarkable”

38. ‘thus’ (p 7292, line 1): and

39. ‘are expected’ (p 7292, line 2): please remove floods. Moreover I wouldn’t say this so frankly, replace with: ‘could be expected’

REPLY: The statement about floods was removed and the replacement the reviewer suggested was done.

40. ‘two degrees warming’ (p 7292, line 12); +2 SWL (p 7292, line 14): please be consistent in naming

REPLY: Here we wanted to make a statement on the effect that “two more degrees of warming” (reaching +4 SWL) has on the change already documented at +2 SWL. To clarify this the last sentence was modified to the following:
“For the rest of the European region where trends are not clear or ensemble members do not agree towards the change, the effect of the further warming from +2 SWL to +4SWL, does not seem to severely affect the hydrological state, which is however already significantly altered at +2 SWL compared to pre-industrial.”

41. Table 1: (1) All RCMs are RCA4 - column could be deleted. (2) Is it of interest to present the equilibrium climate sensitivity?

REPLY: The column containing the RCM was deleted as suggested by the reviewer. We believe that the ECS is of interest as it gives some information on the behavior of the GCM.

42. Table 2 and 3: Absolute and percentage change to what? Please clarify in table or heading.

REPLY: Absolute and percent change refer to changes from baseline in the projected time-slice. Relative clarifications have been added in the Tables.

43. Please merge figure 2 & 4 (3 & 5): add one/two rows to figure 2 and 3 to show the ensemble-mean changes and the modelling agreement.

REPLY: Following the reviewer’s indication Figures 2 -5 have been restructured. Figure 2 was merged with Figure 4 (and Figure 3 with Figure 5). Moreover, the CV for the projected period was substituted with the CV of the absolute differences. Finally, a sub-figure showing model agreement towards a wetter change in the projected time-slice has been added to the two new Figures.

High-end climate change impact on European ~~water availability and stress~~runoff and low flows. Exploring the ~~presence~~ effects of forcing biases~~ and issues on adjustments~~.

L. V. Papadimitriou[1], A. G. Koutroulis[1], M. G. Grillakis[1] and I. K. Tsanis[1,2]

[1]{Technical University of Crete, School of Environmental Engineering, Chania, Greece}

[2]{ McMaster University, Department of Civil Engineering, Hamilton, ON, Canada}

Correspondence to: I. K. Tsanis (tsanis@hydromech.gr)

Abstract

Climate models project a much more substantial warming than the 2°C target under the more probable emission scenarios, making higher end scenarios increasingly plausible. Freshwater availability under such conditions is a key issue of concern. In this study, an ensemble of Euro-CORDEX projections under RCP8.5 is used to assess the mean and low hydrological states under +4 ºC of global warming for the European region. Five major European catchments were analyzed in terms of future drought climatology and the impact of +2 ºC versus +4 ºC global warming was investigated. The effect of bias correction of the climate model outputs and the observations used for this adjustment was also quantified. Projections indicate an intensification of the water cycle at higher levels of warming. Even for areas where the average state may not considerably be affected, low flows are expected to reduce leading to changes in the number of dry days and thus drought climatology. The identified increasing or decreasing runoff trends are substantially intensified when moving from the +2 to the +4 degrees of global warming. Bias correction resulted in an improved representation of the historical hydrology. It is also found that the selection of the observational dataset for

the application of the bias correction has an impact on the projected signal that could be of the same order of magnitude to the selection of the ~~RCM~~GCM.

1 Introduction

Global CO_2 emission rates ~~keep~~ have been following high-end climate change pathways leading to a future global temperature that is likely to surpass the target limit of 2ºC, despite the recent hiatus (England et al., 2015), and reach levels of +4 ºC and higher at the end of the 21st century. By that time, the seasonality of river discharge is expected to get more pronounced for one-third of the global land surface, which translates to increased high flows and decreased low flows (Van Vliet et al., 2013). By the mid-century, the hydrological regime is projected to change considerably for a significant part of the global land surface (Arnell and Gosling, 2013). The effect that global warming can have on water resources raises serious concerns on future water availability, especially under the pressure of the growing global population and the consequent increased food production needs. It is projected that the number of people coping with significantly reduced water availability will increase by 15% globally due to climate change, while the percentage of the global population living under conditions of absolute water scarcity is also projected to increase~~d is also projected to be the percentage of the global population living under conditions of absolute water scarcity~~ (Schewe et al., 2014).

In this framework, the future hydrological state needs to be assessed~~, in terms of both freshwater availability and water stresses~~. The runoff production is the component of the hydrological cycle most representative to describe freshwater availability ~~these aspects~~, as it expresses the amount of available water after the evapotranspiration and infiltration losses and before any stream formation process intervenes. Furthermore, ensembles of mean annual and seasonal runoff can provide information about the climate change impact on river flows (Döll and Schmied, 2012). Studies have shown that changes in runoff are not linearly correlated with changes in global mean temperature (Arnell and Gosling, 2013), neither are meteorological with hydrological droughts (van Huijgevoort et al., 2013), concluding that for climate change impact assessments it is fundamental to use an impact model to translate the precipitation derived signal into runoff.

A substantial number of large scale climate change impact studies that have been performed recently examine the future hydrological state analyzing projections of runoff or river flow. Fung et al. (2011) compared the projected future water availability under +2 ºC and +4 ºC of global warming, forcing the MacPDM Global Hydrological Model (GHM) with 22 GCMs from the CMIP3 experiment. Arnell & Gosling (2013) performed a global assessment of the climate driven changes in runoff based hydrologic indicators in mid-21st century, using multiple scenarios derived from the CMIP3 experiment. Schneider et al. (2013) focused on the impacts of climate change for the European river flows, using data from three bias corrected GCM scenarios. Van Vliet et al. (2013) performed a global assessment of future river discharge and temperature under two climate change scenarios, forcing a GHM with an ensemble of bias corrected GCM output. They found that the combination of lower low flows with increased river water temperature can lead to water quality and ecosystem degradation in south-eastern United States, Europe, eastern China, southern Africa and southern Australia. An investigation of the future trends in flood risk at the global scale was performed by Dankers et al. (2014) and for the European region by Alfieri et al. (2015). ~~The results of the latter study indicate that future flood hazard is mostly affected by the increased frequency of discharge extremes, rather than the absolute increase of discharge values.~~ Betts et al. (2015) performed a global assessment of the impact posed on river flows and terrestrial ecosystems by climate and land use changes described by four RCPs.~~ The study showed that, for all the climate scenarios, global warming in conjunction with elevated CO_2 concentrations result in augmented river outflows by the end of the 21st century.~~ Various multi-model hydrological simulations have been also performed, in an attempt to quantify the climate change analysis' uncertainty resulting from the impact model (Hagemann et al., 2013; van Huijgevoort et al., 2013; Dankers et al., 2014).

~~Other studies assess climate change impacts under the adding stresses of population growth and human activities. Hanasaki et al. (2013) used different combinations of socio-economic and emission scenarios to capture the effects of a wide range of climate change scenarios on future water availability. Future climate was described by 3 GCMS and the water cycle was simulated in conjunction with water demand. Their results report increased water scarcity by the end of the 21st century, even for the lower-end emission and water use changes scenario, mainly due to the increasing population in developing countries and to general changes posed by global warming on the hydrological regime. Arnell & Lloyd-Hughes (2014) examined the effects of different degrees of climate change and future population scenarios on global water resources, using a large ensemble of 19 CMIP5 climate models. Their study underlined the importance of quantifying and accounting for the adaptation and mitigation challenges when assessing climate change impacts. Haddeland et al. (2014) investigated the impact of climate~~

~~change and human interventions on global water resources. According to the study findings, the stress posed by human activities is similar, and in some cases more intense, than a respective stress caused by a +2 ºC global warming scenario. Hejazi et al. (2014) analysed the cross-sectorial changes of water needs under climate change. Results indicate that while water scarcity is an issue that many regions (mainly Middle East and India) will have to tackle with, for other regions (e.g. USA and Canada) climate change is likely to alleviate water shortage problems.~~

~~Multi-model assessments have been performed in an attempt to quantify the climate change analysis' uncertainty resulting from the impact model. Hagemann et al. (2013), performed future hydrological simulations with 3 GCMs and 8 GHMs. According to their findings, impact model induced uncertainty is larger than that of the GCMs for some regions of the land surface. With a global multi-model experiment van Huijgevoort et al. (2013) investigated the impact models' agreement on the effect of global warming on global drought in runoff and concluded that there are significant differences in the simulated drought event duration and spatial extent between the models. Comparing simulations from 9 GHMs and LSMs, Dankers et al. (2014) found that the impact models and the driving GCMs exhibit consistency in the projected patterns of flood risk at the large scale, but significant disagreements are found at the local scale, which may regard even the sign of the change. This increased basin scale uncertainty is an important issue to cope with when studying climate change adaptation locally (Dankers et al., 2013).~~

Currently, global mean temperature has increased 0.85 ºC relative to pre-industrial and already 18% of the moderate daily precipitation extremes is attributed to this warming. At +2 ºC the fraction of the global warming driven precipitation extremes is projected to rise up to 40% (Fischer and Knutti, 2015). The effect of a 2 ºC global warming for the European climate was examined by Vautard et al. (2014). The study revealed that warming in Europe is projected to be higher than the global average of 2 ºC. Temperature increases of up to 3 ºC were found for the winter season over north-western Europe and for the summer months over sourthern Europe. Heavy precipitation was found to increase over the whole continent for all seasons, with the exception of southern Europe during summer. Prospects of limiting the warming to the +2 ºC ~~is~~ target have become vanishingly small (Sanford et al., 2014) at the same time that many experts believe that we are on the +4°C path (Betts et al., 2011, 2015).

The +4 °C global warming scenario is also translated in more intense temperature increases in Europe, especially for the summer season (World Bank, 2014).

Formatted: Superscript

Significant climate change induced alterations are projected for the flow regime in Europe, with the most pronounced changes in magnitude projected for the Mediterranean region and the northern part of the continent (Schneider et al., 2013). Moreover, considering that southern Europe is identified as a possible hotspot where the fraction of land under drought will increase substantially (Prudhomme et al., 2014), along with global temperature rise exceeding +2 °C, concerns for future water availability in Europe are raising. Prolonged water deficits during long-term droughts surpass the resilience of the hydrological systems and are a significant threat to water resources security in Europe (Parry et al., 2012). In the Euro-Mediterranean regions the severity of droughts has increased during the past 50 years, as a consequence of greater atmospheric evaporative demand resulting from temperature rise (Vicente-Serrano et al., 2014). Besides southern European areas, north-western and central-eastern regions appear more drought prone than the rest of Europe (Bonaccorso et al., 2013). Streamflow projections indicate more severe and persistent droughts in many parts of Europe due to climate change, except for northern and north-eastern parts of the continent. The opposite is projected for the middle and northern parts with a highly significant signal of reduced droughts that may be reversed due to intensive water use (Forzieri et al., 2014). Consequently European cropland affected by droughts is projected to increase 7-fold (up to 700,000 km^2/year) at about +3°C of global warming (Ciscar et al., 2014) compared to the situation of the last decades. Similarly, under the same warming level, European population affected by droughts is expected to increase by a factor of seven, overcoming the 150 million/year.

GCM outputs, used as input in impact models to assess the effects of climate change, feature systematic errors and biases. To deal with these, several bias correction techniques have been developed to statistically adjust the GCM output against observations. This process adds another level of uncertainty in the chain of climate to impact modelling that has to be investigated and communicated to the impact research communities. Ehret et al. (2012) acknowledge the fact that inherent climate models' biases render them unsuitable for direct use in climate change impact assessments but express scepticism towards adopting bias correction as a standard undisputed procedure. They argue that bias adjustment hides rather than reduces the uncertainty, as the narrowing of the uncertainty range is not supported by any physical explanation. Teutschbein & Seibert (2012) also accept the need for bias correction but raise awareness towards the increased uncertainty derived from adding this step to the modelling chain. Ehret et al. (2012) introduce the issue of how “correct” is the dataset used as baseline for the bias adjustment. Haerter et al. (2011) underline that the statistical adjustments applied to GCM data with bias correction are bounded to the timescale selected for the adjustment and might have adverse effects on the statistics of another timescale. Haerter et al. (2011) also accentuate that one significant assumption is made when present day based bias correction methods are applied to climate scenario simulations; that of the bias stationarity throughout the future decades. Teng et al. (2015) argue that errors in bias corrected precipitation are inherited and augmented in modelled runoff.

~~Until climate modelling development manages to overcome the biases included in model outputs, GCM data are not an adequate forcing for GHMs. Bias correction is a helpful tool to deal with this problem, as its application improves the representation of both mean flow and seasonality and is thus fundamental for climate change analysis (Harding et al., 2014).~~

The major tools for the investigation of large scale hydrological changes due to climate change are Global Hydrological Models (GHMs) and/or Land Surface Models (LSMs). According to the classification proposed by Haddeland et al. (2011), the models that solve the water balance are considered as GHMs and the models that solve both the water and energy balance are categorized as LSMs. The LSM JULES (Joint UK Land Environment Simulator-~~(~~Best et al., 2011~~)~~) has been implemented for many recent climate change impact and model inter-comparison studies (Hagemann et al. 2013; Davie et al. 2013; Dankers et al. 2014; Prudhomme et al. 2014; Harding et al. 2014).

~~GHMs describe the lateral transfer of water and are focused on water resources (Haddeland et al., 2011) while LSMs focus on flux exchanges mainly at the vertical direction, simulating the energy, water and carbon exchanges between the land surface and the atmosphere (Zulkafli et al., 2013), as they were originally developed to provide the lower boundary for climate models. It should be noted however that for some models their classification in one of the two categories cannot be definitive, and they have been reported in the literature both as GHMs and as LSMs. According to the classification proposed by Haddeland et al. (2011), the models that solve the water balance are considered as GHMs and the models that solve both the water and energy balance are categorized as LSMs. Several references on Global Models and their applications in water related modelling applications follow. WaterGAP (Alcamo et al., 2003)~~

is a GHM well applicable for simulating the effects of climate change on water availability and irrigation demands (Döll et al., 2003; Verzano, 2009), which has been used to simulate inter-sectorial water uses under socio-economic development (Flörke et al., 2013) and to assess changes in flow regimes in Europe due to climate change (Schneider et al., 2013) and the resulting ecological risk for rivers (Laizé et al., 2013). The GHM MacPDM (Arnell, 1999) has been used in various recent studies to estimate climate change effects on global water scarcity (Gosling and Arnell, 2013) and global river flow regimes (Arnell and Gosling, 2013) and for assessing the effects of climate policy on the impacts of climate change (Arnell et al., 2013). The LSM H08 (Hanasaki et al., 2008) has been used in global applications for the estimation of river flows and sources of virtual water used for agriculture and livestock products (Hanasaki et al., 2010).

The LSM JULES (Joint UK Land Environment Simulator) has been implemented for many recent climate change impact and model inter- comparison studies (Hagemann et al. 2013; Davie et al. 2013; Dankers et al. 2014; Prudhomme et al. 2014; Harding et al. 2014). Furthermore, JULES has been used in many recent studies as a tool for evaluating the exchange of water, energy and carbon fluxes between the land surface and the atmosphere. Van den Hoof et al. (2013) assessed JULES' performance in simulating evaporative flux (and its partitions) and carbon flux in temperate Europe and evaluated an adapted version of the model for its suitability for use in climate change studies, based on the extreme summer of 2003. Marthews et al. (2012) implemented JULES in tropical forests of Andes-Amazon to simulate all components of carbon balance and study possible flux variations between sites of different altitude. Zulkafli et al. (2013) implemented JULES in a humid tropical mountain basin of the Peruvian Andes-Amazon. MacKellar et al. (2013) evaluated JULES, implemented in a region of Southern Africa, concerning its ability to simulate the catchment streamflow, testing both the PDM and the TOPMODEL runoff generation schemes. In the study of Bakopoulou et al. (2012), the sensitivity of the JULES outputs to the soil parameters of the model at a point scale was estimated. Dadson et al. (2010) sought to quantify the feedback between wetland inundation and heat and moisture fluxes in the Niger inland delta by adding an overbank flow parameterization into JULES. Burke et al. (2013) used JULES to simulate retrospectively the pan-arctic changes in permafrost and Dankers et al. (2011) assessed JULES' performance in simulating the distribution of surface permafrost in large scale catchments. In a study by Jiménez et al. (2013) soil moisture modelled with JULES is evaluated against satellite soil moisture observations.

The scope of this work is to assess future water availability and identify ~~water stress~~drought conditions in the European region under high-end scenarios of climate change. Transient hydrological simulations for the period 1971 to 2100 were performed by forcing the JULES model with five Euro-CORDEX (Coordinated Downscaling Experiment over Europe) climate projections. Water availability is described by the output of runoff production. In our analysis the model results are mainly interpreted statistically, aiming to express the changes found in the projected future periods with respect to the historical baseline state rather than describing future regimes with absolute numbers. The ~~aspects~~ research objectives ~~that are examined here include~~set by this study are the following:

i) To identify ~~C~~changes posed on the hydrological cycle (mean state and lower extremes) at +4 ºC global warming compared to a baseline situation, and relative to the target of 2 ºC warming.

ii) To analyse ~~T~~the effect of bias correction on projected hydrological simulations. To achieve this, ~~B~~ both raw and bias corrected Euro-CORDEX data were used as input forcing in the impact model.

iii) To assess~~T~~ the effect of the observational dataset used for bias correction.

iv) To identify climate change induced changes in ~~D~~drought climatology~~, along with climate change induced changes,~~ at the basin scale.

2 Data & Methods

Hydrological simulations were performed with the JULES Land Surface Model driven by Euro-CORDEX climate scenarios. To warm-up the model, 10 spin-up cycles from 1955 to 1960 were run. A daily time-step was employed for all the model runs. JULES was setup at the spatial resolution of the forcing Euro-CORDEX data which was 0.44 degrees. The model output was regridded to match a 0.5x0.5 degree grid.

Brief descriptions of the climate data and the impact model are included in the following sections.

2.1 Climate data

Projections from five Euro-CORDEX experiments under Representative Concentration Pathway RCP8.5 scenario were used as input to JULES. The climate models were selected so as to cover the range of model sensitivity, as expressed by the index of Equilibrium climate sensitivity (ECS) which spans from 2.1 to 4.7 K for the CMIP5 ensemble (Andrews et al., 2012). ECS is a useful metric of the response of a climate model, in terms of air temperature change, to a doubling of the atmospheric CO_2 concentration (Andrews et al., 2012). Another factor for selecting the participating climate models was the availability of GCM downscaled at the spatial resolution of 0.44 degrees.

Historical and projected time-slices comprise of 30-years of simulations, for which one time-slice average is extracted. The historical or baseline time-slice covers the period from 1976 to 2005. The projected time-slice varies between the models. The definition for determining the projected time-slice here is to take the 30-year average of the slice centered on the year where the +4 (or +2) Specific Warming Level (SWL) is exceeded. The reference period for the calculation of the SWL is the pre-industrial state and specifically the period from 1861 to 1880. For three of the selected scenarios the +4 SWL is achieved outside the temporal extend of this study, thus the last 30 year period available is considered instead (2071-2100). The SWL exceeded during that period for the models that reach +4 after 2100 is shown in Table 1~~Table 1~~. For reasons of consistency in terminology the time-slice of all models describing the greater SWL achieved will be referred to as +4 SWL time-slice.

Using the SWL concept constitutes the results independent of the timing that the warming occurs. Although by definition of the SWL, the models reach the same level of warming in their time-slices, the different model sensitivity reflects on the evolution of temperature in the time-slice, as more sensitive models are expected to have higher rates of changes in the period before and after a specific SWL is achieved compared to the less sensitive models. Moreover, considering models of different ECS is important to express the range of other than temperature forcing variables produced by the GCMs (eg. radiation).

Formatted: Font color: Auto

The five scenarios along with information on the time-slices extracted for our analysis and the corresponding exceeded warming levels and ECS indices are shown in Table 1~~Table 1~~. Two widely used observational datasets were used to adjust the biases of the RCMs precipitation and temperature data. The first dataset was a hybrid dataset created by the Inter-Sectoral Impact Model Integration and Intercomparison Project ISI-MIP (Warszawski et al., 2014) that consists of the WFD (Weedon et al., 2010) and WFDEI.GPCC. (Weedon et al., 2014) datasets. Additionally, the station data based European Climate Assessment & Dataset (ECA&D) and the ENSEMBLES Observations gridded dataset (E-OBS v10; Haylock et al. 2008) was also used for the bias adjustment of the aforementioned climate variables.

2.2 Bias correction method

In the present study the multi-segment bias correction (MSBC) method is used to correct the precipitation and temperature data for their biases. A detailed description of the method can be found in Grillakis et al. (2013). This bias correction methodology has the ability to better transfer the observed precipitation statistics to the raw GCM data. The method utilizes multiple discrete segments on the cumulative density function (CDF) to fit multiple theoretical distributions, as opposed to the commonly used single transfer function at the entire CDF space. Pragmatically, the method eliminates to a large extent the bias in mean precipitation, while significantly reducing the bias of the higher quantile of the precipitation CDF associated with extreme precipitation events.

~~2.2~~2.3 The JULES land surface model

JULES is a physically based land surface model that was established in 2006. It is comprised of two parts: the Met Office Surface Exchange Scheme (MOSES; Cox et al. 1998) and the Top-down Representation of Interactive Foliage and Flora Including Dynamics (TRIFFID; Cox 2001) component. MOSES is an energy and water balance model which is JULES' forerunner, and TRIFFID is a dynamic global vegetation model (Best et al., 2011; Clark et al., 2011; Cox, 2001). In our model application for this study we do not examine vegetation dynamics thus we are focusing on the MOSES component of JULES.

The meteorological forcing data required for running JULES are: downward shortwave and longwave radiation, precipitation rate, air temperature, wind-speed, air pressure and specific humidity (Best et al., 2011).

JULES has a modular structure, which makes it a flexible modelling platform, as there is the potential of replacing modules or introducing new modules within the model. The physics modules that comprise JULES include the following themes: surface exchange of energy fluxes, snow cover, surface hydrology, soil moisture and temperature, plant physiology, soil carbon and dynamic vegetation ~~dynamics~~ (Best et al., 2011), with the latter being disabled for this application.

In JULES, each gridbox is represented with a number of surface types, each one represented by a tile. JULES recognises nine surface types (Best et al., 2011), of which five are vegetation surface types (broadleaf trees, needleleaf trees, C3 (temperate) grasses, C4 (tropical) grasses and shrubs) and four are non-vegetated surface types (urban, inland water, bare soil and ice). A full energy balance equation including constituents of radiation, sensible heat, latent heat, canopy heat and ground surface heat fluxes is calculated separately for each tile and the average energy balance for the gridbox is found by weighting the values from each tile (Pryor et al., 2012).

In JULES the default soil configuration consists of four soil layers of thicknesses 0.1 m, 0.25 m, 0.65 m and 2.0 m. This configuration however can be altered by the user. The fluxes of soil moisture between each soil layer are described by Darcy's law and a form of Richards' equation (Richards, 1931) governs the soil hydrology. Runoff production is governed by two processes: infiltration excess surface runoff and drainage through the bottom of the soil column, a process calculated as a Darcian flux assuming zero gradient of matric potential (Best et al., 2011). There is also the option of representing soil moisture heterogeneity. In that case total surface runoff also includes saturation excess runoff. The model allows for two approaches to introduce sub-grid scale heterogeneity into the soil moisture: 1) use of TOPMODEL (Beven and Kirkby, 1979), where heterogeneity is taken into account throughout the soil column, or 2) use of PDM (Moore, 1985), which represents heterogeneity in the top soil layer only (Best et al., 2011). Calculation of potential evaporation follows the Penman-Monteith approach (Penman, 1948). Water held at the plant canopy evaporates at the potential rate while restrictions of canopy resistance and soil moisture are applied for the simulation of evaporation from soil and plant transpiration from potential evaporation.

JULES simulates fluxes at the vertical direction only. For hydrological applications this means that the model calculates runoff production in each gridbox which needs to be routed to estimate streamflow. The standard version of the JULES model until very recently (February 2015) did not account for a routing mechanism. To overcome this model limitation, we use a conceptual lumped routing approach based on triangular filtering in order to delay runoff response. This is applied after discriminating the gridboxes that contribute to runoff production of a specific basin from the gridded model output. Determination of gridboxes upstream of the gauging station location is implemented using the TRIP river routing scheme (Oki and Sud, 1998).

JULES has been used in many recent studies as a tool for evaluating the exchange of water, energy and carbon fluxes between the land surface and the atmosphere. Van den Hoof et al. (2013) assessed JULES' performance in simulating evaporative flux (and its partitions) and carbon flux in temperate Europe. Marthews et al. (2012) implemented JULES in tropical forests of Andes-Amazon to simulate all components of carbon balance and study possible flux variations between sites of different altitude. Zulkafli et al. (2013) implemented JULES in a humid tropical mountain basin of the Peruvian Andes-Amazon. MacKellar et al. (2013) evaluated JULES, implemented in a region of Southern Africa, concerning its ability to simulate the catchment streamflow. In the study of Bakopoulou et al. (2012), the sensitivity of the JULES outputs to the soil parameters of the model at a point scale was estimated. Dadson et al. (2010) sought to quantify the feedback between wetland inundation and heat and moisture fluxes in the Niger inland delta by adding an overbank flow parameterization into JULES. Burke et al. (2013) used JULES to simulate retrospectively the pan-arctic changes in permafrost and Dankers et al. (2011) assessed JULES' performance in simulating the distribution of surface permafrost in large scale catchments. In a study by Jiménez et al. (2013) soil moisture modelled with JULES is evaluated against satellite soil moisture observations.

Other studies give insight into the hydrological performance of JULES specifically. (Blyth et al., (2011) extensively evaluated the JULES model for its ability to capture observed fluxes of water and carbon. Concerning discharge, their findings suggest that for the European region seasonality is captured well by the model. For temperate regions (like most of central Europe) to model exhibited a tendency towards underestimating river flows due to overestimation of evapotranspiration. (Prudhomme et al., (2011) assessed JULES' ability in simulating past hydrological events over Europe. In general terms the model was found to capture the timing of major drought events and periods with no large-scale droughts present were also well reproduced. The model showed a positive drought duration bias, more profoundly present in northwest Spain and East Germany-Czech Republic. Prudhomme et al. (2011) argue that this feature is related to overestimation of evaporation by the model. For regions where droughts tend to last longer, JULES exhibited a better ability of reproducing the drought events' characteristics. (Gudmundsson et al., (2012) compared nine large scale hydrological models,

Formatted: Font color: Auto

Formatted: Font color: Auto

Formatted: Font: Not Bold

Formatted: Font: Not Bold, Not Highlight

Formatted: Font: Not Bold

and their ensemble mean, based on their skill in simulating the interannual variability of observed runoff percentiles in Europe. According to the overall performance (accounting for all examined percentiles and evaluation metrics), JULES was ranked third best out of the 10 models, after the multi-model ensemble mean and the GWAVA model. For low and moderately low flows, expressed as 5th and 25th percentile respectively, JULES is also in the top three models regarding the representation of interannual variability in runoff. In the study of ~~(~~Gudmundsson et al.~~,~~ (2012b), where an ensemble of hydrological models is evaluated for their ability to capture seasonal runoff climatology in three different hydroclimatic regime classes in Europe, JULES exhibits a good performance, comparable to that of the best performing multi-model ensemble mean. In other studies employing multi-model ensembles, focusing on the whole European region (Gudmundsson and Seneviratne, 2015) or a single basin in Europe (Harding et al., 2014; Weedon et al., 2015) JULES' simulations also correspond with these of the other models.

Formatted: Superscript

Formatted: Superscript

Formatted: Font color: Auto

~~2.3~~2.4 Identifying changing climate trends

For the assessment of the impact of the +4 ^{o}C warming relative to pre-industrial, the projected time-slices are compared to the baseline period in terms of both absolute and percent change. This is done for each ensemble member individually in order to check the variability of the projected changes and also for the ensemble mean. Two hydrologic indicators are tested, the average and the 10th percentile of runoff production.

Average runoff production is a good and widely used indicator of mean hydrological state of a region. The 10th percentile runoff is considered as a representative indicator of the low flow regime (Prudhomme et al., 2011). Consistent low flows (relative to the mean state) are connected with the formation of hydrological drought conditions. Thus the assessment of the changes in low flows could reveal trends towards more intense or/and often extreme lows in the future hydrological cycle. The impact of high-end climate scenarios on average and 10th percentile runoff is presented both as gridded results at the pan-European scale and aggregated at the basin scale for five major European river basins.

Formatted: Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

The two hydrological indicators were deduced from monthly runoff data. For the analysis of the gridded results at pan-European scale with the SWL time-slice approach, each indicator was computed from the monthly values of all years in the time-slice. For the analysis of basin aggregated runoff regime, the two hydrologic indicators were calculated per year, for all the

Formatted: Font: (Default) Times New Roman, Font color: Black

Formatted: Font: (Default) Times New Roman, Font color: Black years of the simulation. This resulted in time-series of basin aggregated average and 10th percentile runoff production, spanning from 1971 to 2100. The trend of the annual time-series was investigated employing a linear regression analysis to estimate the sign and the average rate of the trend. The significance of the trend was tested at the 95% confidence interval via a Student-t test.

The Europe study domain along with information on the catchments tested and their corresponding gauging stations are shown in Figure 1.

~~2.4~~2.5 Examination of drought climatology

Another aspect of our low flow analysis is to assess ~~to~~ changes in drought climatology, i.e. the number of days per year that ~~extreme~~ particular lows in flow occur. This is here done at the basin scale, following the threshold level method to identify days of discharge deficiencies. The threshold level method is a widely used tool for drought identification applications (Fleig et al., 2006; Vrochidou et al., 2013). According to this method, drought conditions are characterized as the periods during which discharge falls below a pre-defined threshold level. In our application, the threshold is varying daily and is established as in Prudhomme et al. (2011): for each Julian day k, the 10^{th} percentile of a 31-day window discharge centering at day k is derived, from data of all the years of the baseline period (1976-2005). The daily modelled time-series for the whole period simulated (1971-2100) is compared to the daily varying drought limit, and the number of days that fall below the threshold is summed up on an annual basis. The drought threshold is derived from the flows of the baseline period and is applied to both historical and projected flows, in order to capture the climate change induced changes in drought climatology. The regression analysis described in section 2.4 was also applied to the time-series of total drought days per year.

~~2.5 Bias correction method~~

~~In the present study the multi-segment bias correction (MSBC) method is used to correct the precipitation data for its biases. A detailed description of the method can be found in Grillakis et al. (2013). This bias correction methodology has the ability to better transfer the observed precipitation statistics to the raw GCM data. The method utilizes multiple discrete segments on the cumulative density function (CDF) to fit multiple theoretical distributions, as opposed~~

Formatted: Font: (Default) Times New Roman, Font color: Black, Not Superscript/ Subscript

Formatted: Font: (Default) Times New Roman, Font color: Black

Formatted: Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers to the commonly used single transfer function at the entire CDF space. Pragmatically, the method eliminates to a large extent the bias in mean precipitation, while significantly reducing the bias of the higher quantile of the precipitation CDF associated with extreme precipitation events.

3 Results

3.1 Hydrological simulation at Pan-European scale with raw Euro-CORDEX forcing data

Figure 2Figure 2 shows the average runoff production estimated by JULES forced with the five participating dynamical downscaled GCMs, for each model separately and for the ensemble mean. Measures of model agreement (coefficient of variation between the ensemble members and model agreement on a wetter change in the projected time-slice) are also shown in Figure 2. The change in runoff in the +4 SWL°C projected time-slice with respect to the baseline period is expressed as both absolute and percent relative difference. It is interesting to observe the variations between the models for the historical time-slice, with the low climate sensitivity GFDL and NorESM1 exhibiting generally wetter patterns , especially for northern Europe and Scandinavian Peninsula, and with IPSL describing drier patterns, especially for southern Europe. Concerning the overall agreement of the ensemble members in the baseline period the coefficient of variation is below 0.5 for most of the European region (Figure 2, bottom), indicating a good agreement of the models. In more detail, the coefficient of variation is lower for the Scandinavian region and is reduced towards the lower latitudes.

For the projected time-slice, all models agree in a general pattern of increased runoff production in northern Europe and a small part in central Europe Europe and decreased runoff production in Spain, Greecethe Mediterranean region and ,parts of Italy.. Especially for the negative trends shown in southern Europe it is important that though small in absolute terms they increase in magnitude when expressed as a percentage, meaning that small negative changes can pose severe stress in regions where water availability is already an issue.

Concerning the ensemble mean, smoothing of the projected changes due to averaging has revealed clear patterns of change, which however have to be interpreted considering the full spread of the GCM-forced outcomes and the agreement between them in order to avoid misguided conclusions. Less extreme values are encountered in the ensemble mean of

Field Code Changed

projected changes in runoff, compared to the change projected by each ensemble member individually (Figure 2). Especially for percent change a clear trend of runoff increase is revealed in northern Europe and decrease in southern Europe, with a mixed pattern for central Europe. Four or five out of the five ensemble members agree on the wetter response in the northern regions and the drier response in the southern part of Europe. The smaller cv value (cv<0.1) for the southern regions indicates that the models agree more on the value of the change compared to the changes in the Scandinavian region (0.11<cv<0.75). For central Europe there are areas of reduced agreement, with two models showing a change different in sign than the other three of the ensemble. For the same areas cv has values greater than 1, marking a large spread between the values of the five ensemble members.

Figure 3 has the same features as Figure 2 but concerns the 10^{th} percentile runoff production instead of the average. The 10^{th} percentile limit is used to describe low flows that are related to the creation of hydrological drought conditions.Even more alarming trends are deduced from Figure 3, which shows the changes in 10^{th} percentile runoff production at +4°C compared to baseline. For 10^{th} percentile runoff, model agreement in the baseline period is notably reduced compared to agreement for average runoff, with the coefficient of variation for most regions exceeding 0.5 while it exceeds the unity for a large part of Europe. For the +4 SWL projected time-slice, The 10^{th} percentile limit is used to describe low flows that are related to the creation of hydrological drought conditions. Aaccording to Figure 3Figure 3, all models agree in relative decreases in runoff production in western and southern Europe which are specifically pronounced in the western Iberian and Balkan Peninsulas. Another common trend between the models is the significant increase in runoff production in the Scandinavian Peninsula, with MIROC5 and HadGEM2 being the twoonly ensemble member that expands this wetter climate down to central Europe.

Figure 4 and Figure 5 illustrate the changes in ensemble mean behaviour in the +4 °C time-slice for average and 10^{th} percentile runoff respectively along with the coefficient of variation between the ensemble members, which serves as a measure of model agreement. As can be observed in Figure 4, less extreme values are encountered in the ensemble mean of projected changes in runoff, compared to the change projected by each ensemble member individually (Figure 2). Thus averaging has smoothed out the projected changes, making it easier to identify clear patterns of change. Especially for percent change a clear trend of runoff increase is revealed in northern Europe and decrease in southern Europe, with a mixed pattern for

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

central Europe, all between the range of -50% to 50%. In contrastRegarding the ensemble mean changes, percent change in 10^{th} percentile runoff for the ensemble mean (Figure 5Figure 3) shows more significant reductions (up to 100%) compared to average runoff (for which changes range between -50% and 50%)., It is thus deduced that the changes in low flows are more pronounced than the changes in the mean, a conclusion that points towards the overall intensification of the water cycle. The decreasing trend in 10^{th} percentile runoffwith this trend coversing most of most of thethe west and south European area (with 80% to 100% agreement on the sign of the change) while all models agree in an increase in 10^{th} percentile runoff in the Scandinavian region.

Field Code Changed

Formatted: Superscript

Formatted: Superscript

It is thus deduced that the changes in low flows are more pronounced than the changes in the mean, a conclusion that points towards the overall intensification of the water cycle.

Formatted: Strikethrough

Concerning the ensemble members' agreement, for average runoff (Figure 4) the coefficient of variation is below 0.5 for most of the European region, which indicates a good agreement of the models. The coefficient of variation is lower for the Scandinavian region and is reduced towards the lower latitudes. No significant variations can be observed between the coefficient of variation of the baseline and the projected period. For 10^{th} percentile runoff, model agreement is notably reduced, with the coefficient of variation for most regions exceeding 0.5 while it exceeds the unity for a large part of Europe.

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Strikethrough

Formatted: Highlight

3.2 Hydrological simulation at Pan-European scale with bias adjusted Euro-CORDEX forcing data

The ensemble mean of average runoff derived from the five participating downscaled GCMs, whose temperature and precipitation were bias adjusted according to the WFDEI dataset is presented in Figure 6Figure 4. From the ensemble mean runoff for the baseline period it is clear that Bbiias adjustment of the forcing data resulted in a drier ensemble mean runoff for the baseline period for 70.40% of the pan-European land surface, in comparison to 26.01% of the land area that had a wetter response after bias adjustmenthydrological response from the JULES model. The remaining 3.59% of the European area had changes that were classified as insignificant (see ESM for details). The spatial pattern is similar to that of the raw data ensemble but the magnitude is fairly reduced. PFor the projected changes from, data bias adjusted data exhibit very similar patterns and magnitudes with the raw data derived

Field Code Changed

changesment has a small but notable effect on the magnitude of the absolute and percent change, with increases in runoff in northern Europe getting more pronounced in the runs after bias correction. For some regions in central Europe, where a small negative change is reported by the raw data run, a sign change of the projected difference is documented after bias correction. Lastly, bias correction has a strong positive effect on model agreement as it can be documented from the low values of the coefficient of determination all over Europe, with the exception of the Scandinavian Peninsula where model disagreement appears increased after bias correction.

For the baseline period model agreement is stronger compared to the projected period, especially for southern Europe.

In Figure 7Figure 5, the effect of bias correction on the representation of the 10th percentile runoff is shown. As in Figure 6, a decrease in the historical ensemble mean is observed over the whole European region. Some hotspots of pronounced negative changes in western Europe have been eliminated and replaced with milder projected absolute changes. There are areas where sign change is observed (central and central-west Europe) however it is difficult to interpret this result and correlate it with bias correction as these are also the areas where models show the lowest agreement (coefficient of variation exceeding one and agreement towards wetter change 40%-60%). Although the coefficient of variation is for the baseline period is considerably reduced compared to the raw data runs, there are still areas of high model uncertainty in the representation of lower flows.

Field Code Changed

3.3 Basin averaged runoff regime

In Figure 6Figure 86, annual time-series of basin averaged runoff production (average and 10th percentile) for five European basins are shown. These cover the whole length of historical and projected years simulated (1971-2100) in an attempt to identify general trends in average and low runoff, calculating 10-year moving averages from the ensemble mean. Results in 6 Figure 6 include both raw and bias adjusted output, thus an assessment of the effect of the bias correction on the basin scale hydrology can be made. A common observation for all the basins is that runoff decreases considerably for bias adjustedment input forcing.

Formatted: Caption

For Danube, Rhine and Guadiana, significantly important slight negative trends are identified for average runoff (-0.24 mm/year and -0.35 mm/year respectively for raw output, -0.11

mm/year and -0.31 mm/year respectively for bias adjusted output) which are more pronounced for the 10^{th} percentile runoff. For Rhine, the identified trends in average runoff production of both raw and bias corrected forcing are not statistically significant. In contrast, the 10^{th} percentile runoff production in Rhine exhibits statistically significant decreasing trends, for both raw (-0.74 mm/year) and bias corrected (-0.50 mm/year) outputs. For Elbe, raw output gives an insignificant trend in average runoff and a slight decreasing trend for 10^{th} percentile runoff.a Bias corrected data result in a small but statistically significant increasing trend (0.18 mm/year) in annual average runoff while for 10^{th} percentile runoff the trend is decreasing (-0.06 mm/year, statistically significant).clear trend cannot be identified while FFor Kemijoki average and low flows, of raw and bias adjusted forcing, are allboth exhibiting statistically significant increasing trends.

Basin scale average annual runoff production for raw and bias adjusted Euro-CORDEX data as well as the +4°C absolute and percent change for each ensemble member and ensemble mean is included in Table 2. Similar information but for low flows (10^{th} percentile) are presented in the following Table 3. In Tables S1 and S2 of the ESM, the results of the linear regression applied to the average and 10^{th} percentile runoff time-series for the estimation of the trend and its significance can be found.

3.4 Drought climatology at basin scale

Figure 9Figure 7 shows the results of the drought threshold level method analysis for the five study basins, for raw and bias corrected output. For each year, the number of days under the historical drought threshold has been counted. This allows a comparison of the tendency towards the formation of drought conditions between the historical period and the projected period. As this is a statistically oriented interpretation of our data, we can see that the differences between raw and bias corrected time-series are very small, especially compared to the difference in the magnitude of their absolute values. For Danube, Rhine and Guadiana a strongeclear rising trends (all statistically significant) werecan be identified in the 10 year moving average time-series of ensemble mean of days under threshold per year. Before bias correction these were 0.43, 0.37 and 0.52 days/year for the three basins respectively and changed to 0.39, 0.39 and 0.38 days/year respectively after bias correction. For Elbe, non-bias corrected data give a slight but statistically significant increasing trend (0.14 days/year) in

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

contrast to bias corrected output that shows a statistically insignificant trend. ~~and f~~For Kemijoki~~ a~~ strong decreasing (statistically significant) trends are found~~is observed~~ for both for raw (-0.20 days/year) and bias corrected (-0.18 days/year) data. Table S3 of the ESM, tabulates the results of the linear regression applied to time-series of ensemble mean of days under threshold per year for the estimation of the time-series' trend and its significance.

~~For Elbe a sign in the trend cannot be identified. For Danube, Rhine and Kemijoki the raw and bias corrected moving averages almost completely coincide. For Elbe and Guadiana the moving averages of the raw data exhibit a slightly more intense upward trend. These are the two basins where also the range of the raw and bias corrected data vary the most.~~

3.5 Impacts of 4ºC warming relative to 2ºC warming

Figure 8 ~~Figure 10~~ shows the basin average runoff production for raw and bias corrected Euro-CORDEX data with respect to the corresponding SWL in degrees Celsius. This analysis considers the runoff values corresponding to the +2 ºC and +4 ºC SWLs, the latter ranging from 3.2 to 4 between the GCMs, and also the SWL achieved by each participating GCM in the baseline period (0.3-0.5 ºC). It is thus allowing us to examine the changes in basin runoff as temperature increases and to compare the effect of different SWLs.

Formatted: Not Highlight

Formatted: Not Highlight

Formatted: Not Highlight

Comparing the annual average runoff production for raw and bias corrected input forcing it is clear that bias corrected output exhibits a considerably reduced range, which translates in increased model agreement for the basins of Danube, Rhine, Elbe and Guadiana. In Kemijoki basin the bias adjusted output has a greater range than the raw output. Concerning the range of the low flows, an increase in model agreement for the bias corrected forcing is observed for all basins.

Examining the changes in annual average runoff, a slight decreasing trend can be identified for Danube and a slight increasing trend for Elbe while for Rhine there is not a clear trend present. In contrast, Guadiana and Kemijoki exhibit strong decreasing and increasing trends respectively. The falling trend in Guadiana is marginally intensified between +2 and +4 SWL compared to 0 to +2 SWL. The rising trend in Kemijoki does not have evident differences between +2 and +4 ºC.

~~The effect of climate warming is far more pronounced for the low flows.~~ According to the results in Figure 8 ~~Figure 10~~ the 10th percentile runoff in Danube and Rhine ~~significantly~~ decreases as SWLs increase while the opposite trend is observed for the low flows in Kemijoki. For Elbe the raw results show an intense decreasing trend up to +2 SWL which continues more moderately until +4 SWL, in contrast with ~~there is not a clear sign in trend for~~ the bias corrected output that shows milder changes with temperature increase~~,~~ ~~in contrast with the raw results that show an intense decreasing trend up to 2 SWL which continues more moderately until 4 SWL~~. For Guadiana it is difficult to observe a trend in the bias corrected low percentile runoff as the values are already very low. For the raw output however there is an ~~significant~~ abrupt decrease from 0 to +2 °C which continues with a milder trend up to +4 °C.

Figure 9~~Figure 11~~ illustrates the correlation between the percent projected change in annual average and 10th percentile runoff production from bias corrected and raw forcing, for the +2 and +4 SWLs.

Concerning the effect of bias adjustment it can be observed that regardless the significant differences in magnitude between runoff from raw and bias corrected data discussed before, the projected change in average flow by the two forcings almost coincide for the +2 SWL. For the +4 SWL the GCM range has increased for Kemijoki after bias adjustment while for the rest of the basins raw and bias corrected data result in very similar levels of same percent change. For the projected change in 10th percentile runoff, the larger spreading of the values in Figure 9~~Figure 11~~ (right column) shows that the GCM uncertainty on this field is higher. Guadiana is the only basin where bias corrected data result in an improvement in GCM agreement, probably due to its very low values of 10th percentile runoff. Kemijoki is not included in the 10th percentile scatterplots as its projected increase far exceeds the 100% limit selected. For the rest of the basins, the effect of the bias correction on the change of the 10th percentile runoff is not constant. For Guadiana and Elbe bias adjustment mostly increases percent change while for Rhine and Danube percent change is in general terms decreased after bias correction.

Comparing the difference on percent projected change in average annual runoff from +2 to +4 SWL it can be observed that temperature increase results in a slight decline in percent change for basins with small absolute values of change, causing sign changes for Danube and Rhine, and it intensifies the negative and positive changes of Guadiana and Kemijoki respectively.

For the 10th percentile runoff there is a similar response to temperature increase. For Elbe there is positive percent change at +2 SWL which falls below zero at +4 SWL while for Danube, Rhine and Guadiana the already declining projected changes present are further intensified.

3.6 Effect of observational datasets for bias correction on the output of the hydrological model

The aspect of the impact posed by the observational dataset used for bias correction to the results of the hydrological simulations is introduced in this part of our analysis. Additional model runs performed with bias adjusted Euro-CORDEX precipitation and temperature, corrected against the E-OBS (instead of the WFDEI) dataset participate in a comprehensive comparison between all the outputs used in this study. The results are illustrated in ~~Figure 12~~ Figure 10. Three different sets of outputs are compared: one driven by raw downscaled and two driven by Euro-CORDEX data bias corrected against two different datasets. The comparison considers both the mean and range of the ensembles and results are presented as basin aggregates. The first part of the comparison concerns the long-term annual average for the period 1976 to 2005 (~~Figure 12~~Figure 10, top row) and apart from the model results includes values corresponding to observations, derived from GRDC discharge measurements. Observations can serve as a baseline for this comparison, allowing us to evaluate which configuration can better simulate "true" water budget numbers and the effect of bias correction with respect to this baseline.

For all basins the raw data result in overestimates of runoff production which is though significantly reduced after bias correction. E-OBS corrected data however produce values lower than the observations (with the exception of Guadiana) while the WFDEI-corrected data produce the best simulation in terms of approximating the observed values. From Figures S1 and S2 of the ESM (showing the effect of bias correction on the forcing variables of precipitation and temperature) it can be deduced that that E-OBS corrected precipitation has lower values than precipitation adjusted against the WFDEI dataset. This explains the lower runoff produced by the E-OBS bias adjusted dataset, as it is reasonable for the differences in precipitation to reflect on the output of the hydrological model. As already has been revealed in previous stages of this analysis, it is again clear the positive impact that bias adjustment has on the increase of model agreement. The only exception is Kemijoki basin due to its high latitude position (coefficient of variation was increased after bias correction for the high latitude areas).

Changes in annual average runoff production at the +4 SWL appear to be more intensified compared to the +2 SWL (Figure 10Figure 12, middle and bottom). Although for percent change the differences of the distinctive configurations are less pronounced, variations can be observed between the two bias corrected data driven simulations. It is also interesting that the effect of bias correction on reducing the uncertainty is not that strong when looking the results from the more statistical perspective of percent projected change. The improvements in model agreement after bias adjustment however are still significant pronounced for all basins except for Rhine.

From the application of the same analysis on 10th percentile runoff production (Figure S6 of the ESM), it is deduced that for the low flows the E-OBS corrected data again produce lower values of runoff compared to WFDEI. In this case, however, even the raw forced output (which is wetter than the bias corrected) underestimates the observed 10th percentile runoff values. Regarding the percent projected changes, results from bias corrected data produce smaller values compared to the raw data while E-OBS adjusted data result in decreased changes compared to output from WFDEI adjusted forcing.

4 Discussion

4.1 Hydrological response to +4 °C global warming

In our analysis we investigated the effects of climate change on the European hydrological resources, extracting time periods that correspond to an increase of 4 °C of the global temperature, rather than using pre-defined time-slices. The same approach was followed by Vautard et al. (2013), stating that reduced GCM induced uncertainty is achieved with this method and thus the regional patterns of change in the variables of study are strengthened.

In our study only one impact model (JULES) was used. (Hagemann et al., (2013) argue that impact model induced uncertainty in future hydrological simulations is larger than that of the GCMS for some regions of the land surface and suggest using multi-impact model ensembles to deal with this issue. However useful conclusions can be drawn also from studies employing

Formatted: Superscript

Formatted: Superscript

Field Code Changed

a single GHM/LSM. Examples of such single model climate change impact assessments performed recently are the studies of Schneider et al. (2013) and ~~(~~Laizé et al.~~,~~ ~~(~~2013) with the WaterGAP GHM, the studies of ~~(~~Arnell and Gosling~~,~~ (2013), ~~(~~Gosling and Arnell ~~,(~~2013) and ~~(~~Arnell et al.~~,~~ (2013) with the GHM MacPDM and of ~~(~~Hanasaki et al. ~~,~~ (2010) using the H08 LSM.

Field Code Changed

Field Code Changed

The findings of the study regarding the climate changed induced alterations of the mean hydrological state in Europe show decreasing trends for southern Europe, including the Mediterranean region, and strong increasing trends for northern and north-eastern Europe. These follow the same patterns as identified by previous studies. Schneider et al. (2013) found that the most pronounced changes in the magnitude of European river flows are projected for the Mediterranean region and the northern part of the continent. Hagemann et al. (2013) reported positive changes in projected runoff for the high latitudes and negative changes for southern Europe. For central Europe the projected changes are smaller (mostly in the range of -25% to 25%) and thus more easily obscured by GCM and bias correction uncertainty. Arnell & Lloyd-Hughes (2014) report that the main source of uncertainty in the projected climate impact stems from the GCMs, with a range of uncertainty for the CMIP5 ensemble that is similar to that of older climate model experiments.

The projected relative changes found for 10th percentile runoff are far more pronounced than the changes in average, even for the regions where changes in average-state annual runoff were negligible. This finding implies that seasonality in runoff is likely to intensify under climate change and is in accordance with the results of Fung et al. (2011) and Van Vliet et al. (2013) who also reported pronounced seasonality in their projected simulations. This may translate to increased dry spells and thus elevated drought risks in the future. Under the light of these findings (~~With the~~ mean-state runoff changing slightly~~,~~ and ~~the~~ low-state changing significantly)~~, it is deduced that high flows are also to be considerably affected by climate change.~~, m~~M~~ore extreme hydrological ~~events~~ droughts are ~~hence~~ expected in the future ~~under the light of these findings, concerning both extreme lows (droughts) and highs (floods)~~. It should be noted however that projections of low flow bear higher uncertainty compared to average-state, as indicated by the higher values of the coefficient of variation. Similar results of increased model spread expressed as cv for low flows compared to average state flows were found by ~~(~~Koirala et al., (2014).

Specifically for the Guadiana River, the close to zero values of 10th percentile runoff encountered even in the historical period indicate that the river exhibits intermittent flow regime. This is relevant for this particular river, as it is located in a semi-arid region and intermittent flows typically characterize its hydrological regime (Collares-Pereira et al., 2000; Filipe et al., 2002; Pires et al., 1999). Given the changes that are projected for the Iberian Peninsula at +4 SWL, it is expected that the intermittent flow regime in Guadiana might intensify.

Concerning the effects of a +4 °C temperature increase on the European hydrological regime compared to a +2 °C increase, significant alterations posed by the +2 degrees of global warming are identified for south Europe and northern and north-eastern Europe, where the respective decreasing and rising trends are intensified. Fung et al. (2011) also found that changes in mean annual runoff identified at +2 are intensified at +4. More specifically, their study reports that regions where decreasing runoff trends have been found become even drier and, in contrast, areas where runoff is projected to increase are getting wetter. For most of the river basins examined by Fung et al. (2011), water stress is increased at +4 compared to +2, with the exception of a few basins where an increase in rainfall is projected thus decreasing water stress. In our study, the basins located at central Europe (Danube, Rhine and Elbe) do not exhibit significant changes in their annual average runoff values due to temperature increase from +2 to +4. For 10th percentile runoff, however, a temperature increase of +4 °C from the pre-industrial baseline results in an aggravation of the lowering trends that are already significantly affecting the low runoff regime at +2 °C.

Our analysis of drought climatology at the basin scale was based on the total number of days under a predefined daily varying drought threshold. We did not employ any buffering criterion for the days under threshold to be accounted for in the total sum (as discussed for example by (Sung and Chung, (2014) and (Tallaksen et al., (1997)). The use of such a criterion would have decreased the calculated dry days. However, as the interpretation of the results of this study is mostly oriented in identifying trends of change rather than absolute numbers describing the future regime, the lack of a buffering criterion is not supposed to notably affect the extracted conclusions. (Wanders et al., (2015) employed a transient variable threshold for the assessment of the drought conditions under climate change, considering a gradual adaptation of the ecosystem on the altered hydrological regime. This is an interesting alternative, especially for climate change mitigation and adaptation studies. In our study we aimed to identify global warming induced changes in the future hydrological state without considering adaptation, thus the same historically derived threshold was applied to the whole length of the simulated runoff time-series.

From the analysis performed on drought climatology, increased number of days per year under the historically defined drought threshold are found for the basins of Danube, Rhine and Guadiana. Our results correspond with the findings of previous studies about drought regime under climate change. Giuntoli et al. (2015), investigating future high and low flow regimes at the global scale, using multiple impact models and climate scenarios, found increased number of low flow days in Southern Europe. In the study of Wanders & Van Lanen (2015) the impact of climate change on the hydrological drought regime of different climate regions was assessed, using a conceptual hydrological model forced with 3 GCMs. The study findings describe a decrease in the frequency of drought events in the future, which however does not point towards drought alleviation. In contrast, it relates to increased drought event duration and deficit volume. These effects are more pronounced for the arid climates that already face problems of water availability.

4.2 The effect of bias correction

As proposed by Ehret et al. (2012), both raw and bias corrected data driven simulations are presented in our study, in order to comprehensively assess the effect of bias correction on our results. In four of the five study basins, raw data driven simulated runoff overestimates the corresponding observed values. After bias correction, the modelled results represent more accurately the past hydrological regime. Similar improvements in the bias corrected output have been reported by Hagemann et al. (2011), Muerth et al. (2013) and Harding et al. (2014).

For some regions, the sign of the projected change in runoff shifted after bias correction. This finding was also encountered in the study of Hagemann et al. (2011). Hagemann et al. (2011) underline that these changes in the climate signal reveal another uncertainty aspect of the GCM to GHM modelling procedure, that is inherent to the GCM but becomes apparent after the bias adjustment of the climate model output. Teng et al. (2015) argue that signal changes are produced by bias correction errors in higher percentiles' precipitation, thus adding another factor to the uncertainty of the runoff projections.

Although the absolute values of raw and bias corrected simulations differ significantly, this does not apply to the projected relative changes. Liu et al. (2014) also found that raw and bias corrected data resulted in similar estimations of relative changes for a series a variables, including ET and runoff. The study of Muerth et al. (2013) investigates the effect of bias adjustment on hydrological simulations and their climate change induced alterations. Concerning the relative changes between baseline and future time-slices, it is reported that bias correction does not influence notably the hydrologic indicators, apart from the one describing flow seasonality.

Chen et al. (2011) identify three uncertainty components in bias correction applications: the uncertainty of: the different GCM, the variable emission scenarios and that of the decade used for bias adjustment. From a comparison of the latter uncertainty source with the two former, concluded that the choice of correction decade has the smallest contribution to total uncertainty. In this paper we address another uncertainty source; that of the dataset used for correction. It was found that the WFDEI-bias corrected simulation captured better the past hydrological regime compared to the E-OBS-bias corrected configuration. The differences between the two simulations abate when results are expressed as percent change but still their variation are of the same magnitude as that between raw and bias corrected data. This implies that the selection of the observational dataset used for bias correction is not a trivial step of the modelling procedure and it should be treated as an extra factor that causes the uncertainty window of the projected hydrologic conditions to further open

5 Conclusions

In this paper, the future mean- and low- hydrological states under +4 ºC of global warming were assessed for the European region, using the novel dataset of the Euro-CORDEX climate projections. An analysis of the changes in future drought climatology was performed for five major European basins and the impact of +2 ºC versus +4 ºC global warming was estimated. Concurrently, the effect of bias correction of the climate model outputs on the projected climate was also evaluated.

The concluding remarks of this study are summarised below:

Projections show an intensification of the water cycle at +4 SWL, as even for areas where the average state is not considerably affected, there are remarkable~~ed~~ projected decreases of low flows. With the exception of the Scandinavian Peninsula and some small areas in central Europe, 10th percentile runoff production is projected to reduce all over Europe. This favours the formation of extreme hydrological events, thus more droughts ~~and floods~~ compared to the current state could be~~are~~ expected in the future due to the warming climate.

Drought climatology is projected to change to more dry days per year for the Danube, Rhine and Guadiana basins. Thus these areas are projected to experience more usual and more intense drought events in the future.

For the areas where clear decreasing or increasing runoff trends are identified in the projections, these changes are considerably intensified when moving from the +2 SWL to the +4 SWL. Decreasing trends apply to southern Europe, including the Mediterranean region, while strong increasing trends are projected for northern and north-eastern Europe. For the rest of the European region where trends are not clear or ensemble members do not agree towards the change, the effect of the ~~the~~ further ~~plus two degrees~~ warming from +2 SWL to +4SWL, does not seem to severely affect the hydrological state, which is however already significantly altered at +2 SWL compared to pre-industrial.

Bias correction results in an improved representation of the historical hydrological conditions. However, raw and bias corrected simulations exhibit minor variations for results of statistical interpretation (in our study: percent change, number of days under drought threshold).

The dataset used for bias correction can affect the quality of the projections in absolute terms to a great extent. The comparison performed here showed that the WFDEI-corrected dataset produces simulations that capture better the past observed hydrologic state compared to the E-OBS-corrected dataset and should thus be preferred for bias correction applications over Europe. The selection of the “correct” dataset is an added uncertainty to the climate impact modelling chain, with magnitude similar to that of the bias correction procedure itself.

Acknowledgements

The research leading to these results has received funding from HELIX project of the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 603864. We acknowledge the World Climate Research Programme's Working Group on Regional Climate, and the Working Group on Coupled Modelling, former coordinating body of CORDEX and responsible panel for CMIP5. We also thank the climate modelling groups (listed in Table 1~~Table 1~~ of this paper) for producing and making available their model output. We also acknowledge the Earth System Grid Federation infrastructure an international effort led by the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison, the European Network for Earth System Modelling and other partners in the Global Organisation for Earth System Science Portals (GO-ESSP). Finally, we acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu).

References

Alfieri, L., Burek, P., Feyen, L. and Forzieri, G.: Global warming increases the frequency of river floods in Europe, Hydrol. Earth Syst. Sci. Discuss., 12, 1119–1152, doi:doi:10.5194/hessd-12-1119-2015, 2015.

Andrews, T., Gregory, J. M., Webb, M. J. and Taylor, K. E.: Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models, Geophys. Res. Lett., 39(9), 1–7, doi:10.1029/2012GL051607, 2012.

Arnell, N. W. and Gosling, S. N.: The impacts of climate change on river flow regimes at the global scale, J. Hydrol., 486, 351–364, doi:10.1016/j.jhydrol.2013.02.010, 2013.

Arnell, N. W. and Lloyd-Hughes, B.: The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios, Clim. Change, 122(1-2), 127–140, doi:10.1007/s10584-013-0948-4, 2014.

Arnell, N. W., Lowe, J. A., Brown, S., Gosling, S. N., Gottschalk, P., Hinkel, J., Lloyd-Hughes, B., Nicholls, R. J., Osborn, T. J., Osborne, T. M., Rose, G. A., Smith, P. and Warren, R. F.: A global assessment of the effects of climate policy on the impacts of climate change, Nat. Clim. Chang., 3(5), 512–519, doi:10.1038/nclimate1793, 2013.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. . L. H., Ménard, C. B., Edwards, J. M., Hendry, M. a., Porson, a., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B. and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes, Geosci. Model Dev., 4(3), 677–699, doi:10.5194/gmd-4-677-2011, 2011.

Betts, R. a, Collins, M., Hemming, D. L., Jones, C. D., Lowe, J. a and Sanderson, M. G.: When could global warming reach 4°C?, Philos. Trans. A. Math. Phys. Eng. Sci., 369(1934), 67–84, doi:10.1098/rsta.2010.0292, 2011.

Betts, R. A., Golding, N., Gonzalez, P., Gornall, J., Kahana, R., Kay, G., Mitchell, L. and Wiltshire, A.: Climate and land use change impacts on global terrestrial ecosystems and river flows in the HadGEM2-ES Earth System Model using the Representative Concentration Pathways, Biogeosciences Discuss., 10(4), 6171–6223, doi:10.5194/bg-12-1317-2015, 2015.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrol. Sci. Bull., 24(1), 43–69, doi:10.1080/02626667909491834, 1979.

Blyth, E., Clark, D. B., Ellis, R., Huntingford, C., Los, S., Pryor, M., Best, M. and Sitch, S.: A comprehensive set of benchmark tests for a land surface model of simultaneous fluxes of water and carbon at both the global and seasonal scale, Geosci. Model Dev., 4(2), 255–269, doi:10.5194/gmd-4-255-2011, 2011.

Bonaccorso, B., Peres, D. J., Cancelliere, A. and Rossi, G.: Large Scale Probabilistic Drought Characterization Over Europe, Water Resour. Manag., 27(6), 1675–1692, doi:10.1007/s11269-012-0177-z, 2013.

Chen, C., Haerter, J. O., Hagemann, S. and Piani, C.: On the contribution of statistical bias correction to the uncertainty in the projected hydrological cycle, Geophys. Res. Lett., 38(20), 1–6, doi:10.1029/2011GL049318, 2011.

Ciscar, J.-C., Feyen, L., Soria, A., Lavalle, C., Raes, F., Perry, M., Nemry, F., Demirel, H., Rozsai, M., Dosio, A. and others: Climate Impacts in Europe-The JRC PESETA II project, 2014.

Clark, D. B., Mercado, L. M., Sitch, S., Jones, C. D., Gedney, N., Best, M. J., Pryor, M., Rooney, G. G., Essery, R. L. H., Blyth, E., Boucher, O., Harding, R. J., Huntingford, C. and Cox, P. M.: The Joint UK Land Environment Simulator (JULES), model description – Part 2: Carbon fluxes and vegetation dynamics, Geosci. Model Dev., 4(3), 701–722, doi:10.5194/gmd-4-701-2011, 2011.

Collares-Pereira, M. J., Cowx, I. G., Ribeiro, F., Rodrigues, J. A. and Rogado, L.: Threats imposed by water resource development schemes on the conservation of endangered fish species in the Guadiana River basin in Portugal, Fish. Manag. Ecol., 7(1-2), 167–178, doi:10.1046/j.1365-2400.2000.00202.x, 2000.

Cox, P. ., Huntingford, C. and Harding, R. .: A canopy conductance and photosynthesis model for use in a GCM land surface scheme, J. Hydrol., 212-213, 79–94, doi:10.1016/S0022-1694(98)00203-0, 1998.

Cox, P. M.: Description of the " TRIFFID " Dynamic Global Vegetation Model, 2001.

Dankers, R., Arnell, N. W., Clark, D. B., Falloon, P. D., Fekete, B. M., Gosling, S. N., Heinke, J., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: First look at changes in flood hazard in the Inter-Sectoral Impact Model Intercomparison Project ensemble., Proc. Natl. Acad. Sci. U. S. A., 1–5, doi:10.1073/pnas.1302078110, 2013.

Döll, P. and Schmied, H. M.: How is the impact of climate change on river flow regimes related to the impact on mean annual runoff? A global-scale analysis, Environ. Res. Lett., 7(1), 014037, doi:10.1088/1748-9326/7/1/014037, 2012.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K. and Liebert, J.: HESS Opinions "should we apply bias correction to global and regional climate model data?," Hydrol. Earth Syst. Sci., 16(9), 3391–3404, doi:10.5194/hess-16-3391-2012, 2012.

England, M. H., Kajtar, J. B. and Maher, N.: Robust warming projections despite the recent hiatus, Nat. Clim. Chang., 5(5), 394–396, doi:10.1038/nclimate2575, 2015.

Filipe, A. F., Cowx, I. G. and Collares-Pereira, M. J.: Spatial modelling of freshwater fish in semi-arid river systems: A tool for conservation, River Res. Appl., 18(2), 123–136, doi:10.1002/rra.638, 2002.

Fischer, E. M. and Knutti, R.: Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes, Nat. Clim. Chang., (April), 1–6, doi:10.1038/nclimate2617, 2015.

Fleig, A. K., Tallaksen, L. M., Hisdal, H. and Demuth, S.: Sciences A global evaluation of streamflow drought characteristics, , (2002), 535–552, 2006.

Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F. and Bianchi, a.: Ensemble projections of future streamflow droughts in Europe, Hydrol. Earth Syst. Sci., 18(1), 85–108, doi:10.5194/hess-18-85-2014, 2014.

Fung, F., Lopez, A. and New, M.: Water availability in +2°C and +4°C worlds., Philos. Trans. A. Math. Phys. Eng. Sci., 369(1934), 99–116, doi:10.1098/rsta.2010.0293, 2011.

Giuntoli, I., Vidal, J., Prudhomme, C. and Hannah, D. M.: Future hydrological extremes : the uncertainty from multiple global climate and global hydrological models, Earth Syst. Dyn., 1–30, doi:10.5194/esdd-6-1-2015, 2015.

Gosling, S. N. and Arnell, N. W.: A global assessment of the impact of climate change on water scarcity, Clim. Change, doi:10.1007/s10584-013-0853-x, 2013.

Grillakis, M. G., Koutroulis, A. G. and Tsanis, I. K.: Multisegment statistical bias correction of daily GCM precipitation output, J. Geophys. Res. Atmos., 118(8), 3150–3162, doi:10.1002/jgrd.50323, 2013.

Gudmundsson, L. and Seneviratne, S. I.: Towards observation-based gridded runoff estimates for Europe, , 2859–2879, doi:10.5194/hess-19-2859-2015, 2015.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., Clark, D. B., Dumont, E., Hagemann, S., Bertrand, N., Gerten, D., Heinke, J., Hanasaki, N., Voss, F. and Koirala, S.: Comparing large-scale hydrological model simulations to observed runoff percentiles in Europe, J. Hydrometeorol., 13, doi:10.1175/JHM-D-11-083.1, 2012a.

Gudmundsson, L., Wagener, T., Tallaksen, L. M. and Engeland, K.: Evaluation of nine large-scale hydrological models with respect to the seasonal runoff climatology in Europe, , 48(October), 1–20, doi:10.1029/2011WR010911, 2012b.

Haddeland, I., Clark, D. B., Franssen, W., Ludwig, F., Voß, F., Arnell, N. W., Bertrand, N., Best, M., Folwell, S., Gerten, D., Gomes, S., Gosling, S. N., Hagemann, S., Hanasaki, N., Harding, R., Heinke, J., Kabat, P., Koirala, S., Oki, T., Polcher, J., Stacke, T., Viterbo, P., Weedon, G. P. and Yeh, P.: Multimodel Estimate of the Global Terrestrial Water Balance: Setup and First Results, J. Hydrometeorol., 12(5), 869–884, doi:10.1175/2011JHM1324.1, 2011.

Haerter, J. O., Hagemann, S., Moseley, C. and Piani, C.: Climate model bias correction and the role of timescales, Hydrol. Earth Syst. Sci., 15(3), 1065–1079, doi:10.5194/hess-15-1065-2011, 2011.

Hagemann, S., Chen, C., Haerter, J. O., Heinke, J., Gerten, D. and Piani, C.: Impact of a Statistical Bias Correction on the Projected Hydrological Changes Obtained from Three GCMs and Two Hydrology Models, J. Hydrometeorol., 12(4), 556–578, doi:10.1175/2011JHM1336.1, 2011.

Hagemann, S., Chen, C., Clark, D. B., Folwell, S., Gosling, S. N., Haddeland, I., Hanasaki, N., Heinke, J., Ludwig, F., Voss, F. and Wiltshire, a. J.: Climate change impact on available water resources obtained using multiple global climate and hydrology models, Earth Syst. Dyn., 4(1), 129–144, doi:10.5194/esd-4-129-2013, 2013.

Hanasaki, N., Inuzuka, T., Kanae, S. and Oki, T.: An estimation of global virtual water flow and sources of water withdrawal for major crops and livestock products using a global hydrological model, J. Hydrol., 384(3-4), 232–244, doi:10.1016/j.jhydrol.2009.09.028, 2010.

Harding, R. J., Weedon, G. P., van Lanen, H. a. J. and Clark, D. B.: The future for Global Water Assessment, J. Hydrol., 518, 186–193, doi:10.1016/j.jhydrol.2014.05.014, 2014.

Haylock, M. R., Hofstra, N., Klein Tank, a. M. G., Klok, E. J., Jones, P. D. and New, M.: A European daily high-resolution gridded dataset of surface temperature and precipitation, J. Geophys. Res, 113(December 2007), D20119, doi:10.1029/2008JD1020, 2008.

van Huijgevoort, M. H. J., Hazenberg, P., van Lanen, H. a. J., Teuling, a. J., Clark, D. B., Folwell, S., Gosling, S. N., Hanasaki, N., Heinke, J., Koirala, S., Stacke, T., Voss, F., Sheffield, J. and Uijlenhoet, R.: Global Multimodel Analysis of Drought in Runoff for the Second Half of the Twentieth Century, J. Hydrometeorol., 14(5), 1535–1552, doi:10.1175/JHM-D-12-0186.1, 2013.

Koirala, S., Hirabayashi, Y., Mahendran, R. and Kanae, S.: Global assessment of agreement among streamflow projections using CMIP5 model outputs, Environ. Res. Lett., 9(6), 064017, doi:10.1088/1748-9326/9/6/064017, 2014.

Laizé, C. L. R., Acreman, M. C., Schneider, C., Dunbar, M. J., Houghton-Carr, H. A., Florke, M. and Hannah, D. M.: Projected flow alteration and ecological risk for pan-european rivers, River Res. Appl., 2013.

Liu, M., Rajagopalan, K., Chung, S. H., Jiang, X., Harrison, J., Nergui, T., Guenther, a., Miller, C., Reyes, J., Tague, C., Choate, J., Salathé, E. P., Stöckle, C. O. and Adam, J. C.: What is the importance of climate model bias when projecting the impacts of climate change on land surface processes?, Biogeosciences, 11(10), 2601–2622, doi:10.5194/bg-11-2601-2014, 2014.

Moore, R. J.: The probability-distributed principle and runoff production at point and basin scales, Hydrol. Sci. J., 30(2), 273–297, doi:10.1080/02626668509490989, 1985.

Muerth, M. J., Gauvin St-Denis, B., Ricard, S., Velázquez, J. a., Schmid, J., Minville, M., Caya, D., Chaumont, D., Ludwig, R. and Turcotte, R.: On the need for bias correction in regional climate scenarios to assess climate change impacts on river runoff, Hydrol. Earth Syst. Sci., 17(3), 1189–1204, doi:10.5194/hess-17-1189-2013, 2013.

Oki, T. and Sud, Y. C.: Design of Total Runoff Integrating Pathways (TRIP)— A Global River Channel Network, , 2(1), 7–22, 1998.

Parry, S., Hannaford, J., Lloyd-Hughes, B. and Prudhomme, C.: Multi-year droughts in Europe: analysis of development and causes, Hydrol. Res., 43(5), 689–706, 2012.

Penman, H. L.: Natural evaporation from open water, bare soil and grass, in Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 193, pp. 120–145., 1948.

Pires, A., Cowx, I. and Coelho, M.: Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal, J. Fish Biol., 54, 235–249, doi:10.1006/jfbi.1998.0860, 1999.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S. and Voss, F.: How Well Do Large-Scale Models Reproduce Regional Hydrological Extremes in Europe?, J. Hydrometeorol., 12(6), 1181–1204, doi:10.1175/2011JHM1387.1, 2011.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y. and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment., Proc. Natl. Acad. Sci. U. S. A., 111(9), 3262–7, doi:10.1073/pnas.1222473110, 2014.

Pryor, M., Clark, D., Harris, P. and Hendry, M.: Joint UK Land Environment Simulator (JULES) Version 3.2 User Manual, 2012.

Richards, L. A.: Capillary conduction of liquids through porous mediums, J. Appl. Phys., 1(5), 318–333, 1931.

Sanford, T., Frumhoff, P. C., Luers, A. and Gulledge, J.: The climate policy narrative for a dangerously warming world, Nat. Clim. Chang., 4(3), 164–166, 2014.

Schewe, J., Heinke, J., Gerten, D., Haddeland, I., Arnell, N. W., Clark, D. B., Dankers, R., Eisner, S., Fekete, B. M., Colón-González, F. J., Gosling, S. N., Kim, H., Liu, X., Masaki, Y., Portmann, F. T., Satoh, Y., Stacke, T., Tang, Q., Wada, Y., Wisser, D., Albrecht, T., Frieler, K., Piontek, F., Warszawski, L. and Kabat, P.: Multimodel assessment of water scarcity under climate change., Proc. Natl. Acad. Sci. U. S. A., 111(9), 3245–3250, doi:10.1073/pnas.1222460110, 2014.

Schneider, C., Laizé, C. L. R., Acreman, M. C. and Flörke, M.: How will climate change modify river flow regimes in Europe?, Hydrol. Earth Syst. Sci., 17(1), 325–339, doi:10.5194/hess-17-325-2013, 2013.

Sung, J. H. and Chung, E.-S.: Development of streamflow drought severity–duration–frequency curves using the threshold level method, Hydrol. Earth Syst. Sci., 18(9), 3341–3351, doi:10.5194/hess-18-3341-2014, 2014.

Tallaksen, L., Madsen, H. and Clausen, B.: On the definition and modelling of streamflow drought duration and deficit volume, Hydrol. Sci. J., 42(1), 15–33, doi:10.1080/02626669709492003, 1997.

Teng, J., Potter, N. J., Chiew, F. H. S., Zhang, L., Wang, B., Vaze, J. and Evans, J. P.: How does bias correction of regional climate model precipitation affect modelled runoff?, Hydrol. Earth Syst. Sci., 19(2), 711–728, doi:10.5194/hess-19-711-2015, 2015.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456-457, 12–29, doi:10.1016/j.jhydrol.2012.05.052, 2012.

Vautard, R., Gobiet, A., Jacob, D., Belda, M., Colette, A., Déqué, M., Fernández, J., García-Díez, M., Goergen, K., Güttler, I., Halenka, T., Karacostas, T., Katragkou, E., Keuler, K., Kotlarski, S., Mayer, S., van Meijgaard, E., Nikulin, G., Patarčić, M., Scinocca, J., Sobolowski, S., Suklitsch, M., Teichmann, C., Warrach-Sagi, K., Wulfmeyer, V. and Yiou, P.: The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project, Clim. Dyn., 41(9-10), 2555–2575, doi:10.1007/s00382-013-1714-z, 2013.

Vautard, R., Gobiet, A., Sobolowski, S., Kjellström, E., Stegehuis, A., Watkiss, P., Mendlik, T., Landgren, O., Nikulin, G., Teichmann, C. and Jacob, D.: The European climate under a 2 °C global warming, Environ. Res. Lett., 9(3), 034006, doi:10.1088/1748-9326/9/3/034006, 2014.

Vicente-Serrano, S. M., Lopez-Moreno, J.-I., Beguería, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., García-Ruiz, J. M., Azorin-Molina, C., Morán-Tejeda, E., Revuelto, J., Trigo, R., Coelho, F. and Espejo, F.: Evidence of increasing drought severity caused by temperature rise in southern Europe, Environ. Res. Lett., 9(4), 044001, doi:10.1088/1748-9326/9/4/044001, 2014.

Van Vliet, M. T. H., Franssen, W. H. P., Yearsley, J. R., Ludwig, F., Haddeland, I., Lettenmaier, D. P. and Kabat, P.: Global river discharge and water temperature under climate change, Glob. Environ. Chang., 23(2), 450–464, doi:10.1016/j.gloenvcha.2012.11.002, 2013.

Vrochidou, a. E. K., Tsanis, I. K., Grillakis, M. G. and Koutroulis, a. G.: The impact of climate change on hydrometeorological droughts at a basin scale, J. Hydrol., 476, 290–301, doi:10.1016/j.jhydrol.2012.10.046, 2013.

Wanders, N. and Van Lanen, H. a. J.: Future discharge drought across climate regions around the world modelled with a synthetic hydrological modelling approach forced by three general circulation models, Nat. Hazards Earth Syst. Sci., 15(3), 487–504, doi:10.5194/nhess-15-487-2015, 2015.

Wanders, N., Wada, Y. and Van Lanen, H. A. J.: Global hydrological droughts in the 21st century under a changing hydrological regime, Earth Syst. Dyn., 6(1), 1–15, doi:10.5194/esd-6-1-2015, 2015.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O. and Schewe, J.: The Inter-

Sectoral Impact Model Intercomparison Project (ISI--MIP): Project framework, Proc. Natl. Acad. Sci., 111(9), 3228–3232, 2014.

Weedon, G., P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J. and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, Water Resour. Res., 50((9)), 7505–7514, doi:10.1002/2014WR015638.Received, 2014.

Weedon, G. P., Gomes, S., Viterbo, P., Österle, H., Adam, J. C., Bellouin, N., Boucher, O. and Best, M.: The WATCH forcing data 1958--2001: A meteorological forcing dataset for land surface and hydrological models, Watch. Ed. Watch Tech. Rep., 22, 2010.

Weedon, G. P., Prudhomme, C., Crooks, S., Ellis, R. J., Folwell, S. S. and Best, M. J.: Evaluating the Performance of Hydrological Models via Cross-Spectral Analysis: Case Study of the Thames Basin, United Kingdom, J. Hydrometeorol., 16(1), 214–231, doi:10.1175/JHM-D-14-0021.1, 2015.

Table 1. Euro-CORDEX climate scenarios used to force JULES.

	GCM	+2 SWL time-slice	Exceeded warming level (°C) in the +2 SWL time-slice	+4 SWL time-slice	Exceeded warming level (°C) in the +4 SWL time-slice	Equilibrium Climate Sensitivity (K)
1	GFDL-ESM2M	2040-2069	2	2071-2100	3.2	2.44
2	NorESM1	2036-2065	2	2071-2100	3.75	2.80
3	MIROC5	2037-2066	2	2071-2100	3.76	2.72
4	IPSL-CM5A	2018-2047	2	2055-2084	4	4.13
5	HadGEM2-ES	2024-2053	2	2060-2089	4	4.59

Formatted Table

	RCM	GCM
1	RCA4	GFDL-ESM2M
2	RCA4	NorESM1
3	RCA4	MIROC5
4	RCA4	IPSL-CM5A
5	RCA4	HadGEM2-ES

1

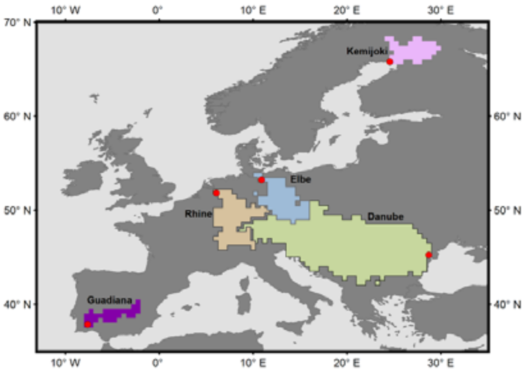
Table 2. Basin's annual average runoff production for raw and bias adjusted Euro-CORDEX data.

Basin's Annual Average Runoff Production [mm/year]												
	Raw						**Bias Corrected**					
	Historical average 1976-2005						**Historical average 1976-2005**					
Danube	462.05	362.35	383.78	304.02	266.21	355.68	219.37	249.80	201.95	226.70	229.00	225.36
Rhine	794.21	845.83	616.94	710.16	495.99	692.63	426.67	503.68	415.00	439.11	470.29	450.95
Elbe	371.88	356.72	219.68	337.42	174.41	292.02	148.70	203.39	135.98	174.79	202.12	173.00
Guadiana	166.13	71.44	116.14	46.60	81.51	96.36	93.14	96.42	90.06	79.22	89.82	89.73
Kemijoki	428.17	482.28	427.95	418.03	507.48	452.78	174.68	327.78	197.30	238.28	450.70	277.75
RCM-GCM	**RCA4-GFDL-ESM2M +3.2 (2071-2100)**	**RCA4-NorESM1 +3.75 (2071-2100)**	**RCA4-MIROC5 +3.76 (2071-2100)**	**RCA4-IPSL-CM5A +4 (2055-2084)**	**RCA4-HadGEM2-ES +4 (2060-2089)**	**MEAN**	**RCA4-GFDL-ESM2M +3.2 (2071-2100)**	**RCA4-NorESM1 +3.75 (2071-2100)**	**RCA4-MIROC5 +3.76 (2071-2100)**	**RCA4-IPSL-CM5A +4 (2055-2084)**	**RCA4-HadGEM2-ES +4 (2060-2089)**	**MEAN**
	Absolute change from baseline in the projected time-slice						**Absolute change from baseline in the projected time-slice**					
Danube	-54.57	3.36	-13.20	-42.04	-14.96	-24.28	-11.83	-1.38	3.61	-30.04	-11.48	-10.22
Rhine	59.95	-19.81	-13.23	-39.31	-20.14	-6.51	53.83	-5.91	6.09	-44.17	-21.73	-2.37
Elbe	2.05	33.91	30.00	-28.39	19.05	11.32	22.81	33.28	31.55	-5.57	25.71	21.55
Guadiana	-55.70	-37.02	-17.16	-14.09	-46.16	-34.03	-26.23	-48.81	-10.37	-28.52	-45.23	-31.83
Kemijoki	146.86	67.46	67.48	174.94	108.26	113.00	149.69	97.38	89.71	179.15	119.97	127.18
	Percent change from baseline in the projected time-slice						**Percent change from baseline in the projected time-slice**					
Danube	-11.81	0.93	-3.44	-13.83	-5.62	-6.83	-5.39	-0.55	1.79	-13.25	-5.01	-4.54
Rhine	7.55	-2.34	-2.14	-5.54	-4.06	-0.94	12.62	-1.17	1.47	-10.06	-4.62	-0.53
Elbe	0.55	9.51	13.66	-8.42	10.92	3.88	15.34	16.36	23.20	-3.19	12.72	12.46
Guadiana	-33.53	-51.82	-14.78	-30.24	-56.63	-35.31	-28.16	-50.63	-11.51	-36.00	-50.35	-35.47
Kemijoki	34.30	13.99	15.77	41.85	21.33	24.96	85.69	29.71	45.47	75.19	26.62	45.79

Formatted: English (United States)

Table 3. Basin's 10th percentile of runoff production, derived on an annual basis, for raw and bias adjusted Euro-CORDEX data.

Basin's 10th percentile on annual basis [mm/year]												
	Raw						Bias Corrected					
	Historical average 1976-2005						Historical average 1976-2005					
Danube	146.63	96.81	80.55	79.71	58.69	92.48	31.49	41.73	28.54	30.32	37.94	34.00
Rhine	250.22	258.37	162.58	200.59	109.23	196.20	98.23	120.41	93.24	101.58	107.68	104.23
Elbe	118.79	99.15	29.98	98.30	28.95	75.04	10.22	20.08	11.23	16.75	22.14	16.08
Guadiana	0.74	0.00	0.12	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.00	0.00
Kemijoki	0.80	4.50	1.10	1.47	10.79	3.73	0.25	5.91	0.53	1.00	11.60	3.86
RCM-GCM	**RCA4-GFDL-ESM2M +3.2 (2071-2100)**	**RCA4-NorESM1 +3.75 (2071-2100)**	**RCA4-MIROC 5 +3.76 (2071-2100)**	**RCA4-IPSL-CM5A +4 (2055-2084)**	**RCA4-HadGEM2-ES +4 (2060-2089)**	**MEAN**	**RCA4-GFDL-ESM2M +3.2 (2071-2100)**	**RCA4-NorESM1 +3.75 (2071-2100)**	**RCA4-MIROC 5 +3.76 (2071-2100)**	**RCA4-IPSL-CM5A +4 (2055-2084)**	**RCA4-HadGEM2-ES +4 (2060-2089)**	**MEAN**
	Absolute change from baseline in the projected time-slice						**Absolute change** from baseline in the projected time-slice					
Danube	-53.89	-23.89	-18.83	-38.22	-27.41	-32.45	-18.03	-15.89	-9.68	-22.28	-24.37	-18.05
Rhine	-89.38	-87.03	-20.39	-103.94	-43.25	-68.80	-31.43	-49.93	-19.49	-69.92	-52.57	-44.67
Elbe	-29.14	-21.01	1.21	-44.80	-9.96	-20.74	-2.03	-2.73	-0.91	-8.90	-8.52	-4.62
Guadiana	-0.73	0.00	-0.11	0.00	0.00	-0.17	0.00	0.00	0.00	0.00	0.00	0.00
Kemijoki	16.77	53.16	36.71	56.80	72.44	47.18	3.24	3.12	5.05	22.55	16.79	10.15
	Percent change from baseline in the projected time-slice						**Percent change** from baseline in the projected time-slice					
Danube	-36.75	-24.68	-23.38	-47.95	-46.71	-35.09	-57.26	-38.07	-33.90	-73.50	-64.22	-53.08
Rhine	-35.72	-33.68	-12.54	-51.82	-39.59	-35.07	-32.00	-41.46	-20.91	-68.83	-48.82	-42.86
Elbe	-24.53	-21.19	4.04	-45.57	-34.41	-27.64	-19.86	-13.58	-8.11	-53.15	-38.47	-28.71
Guadiana	-98.67	-73.37	-96.24	-26.22	-76.38	-98.01	-48.53	-50.67	-65.42	-32.31	-56.63	-53.36
Kemijoki	2088.40	1181.25	3328.72	3877.01	671.51	1264.16	1283.66	52.88	946.08	2265.11	144.71	263.09



Station Info						
River	Station	GRDC Station No	Latitude (dec. degrees)	Longitude (dec. degrees)	Observed cathment area [km²]	Modelled cathment area [km²]
Danube	Ceatal Izmail	6742900	45.22	28.72	807000	815063
Rhine	Rees	6335020	51.75	6.40	159300	177339
Elbe	Neu-Darchau	6340110	53.23	10.89	131950	133886
Guadiana	Pulo do Lovo	6116200	37.82	-7.63	60883	72215
Kemijoki	Isohaara	6854700	65.78	24.55	50686	55628

Figure 1. European study domain, tested basins as defined by the model's 0.5 degree resolution, gauging stations and general information on the stations.

Formatted: Caption

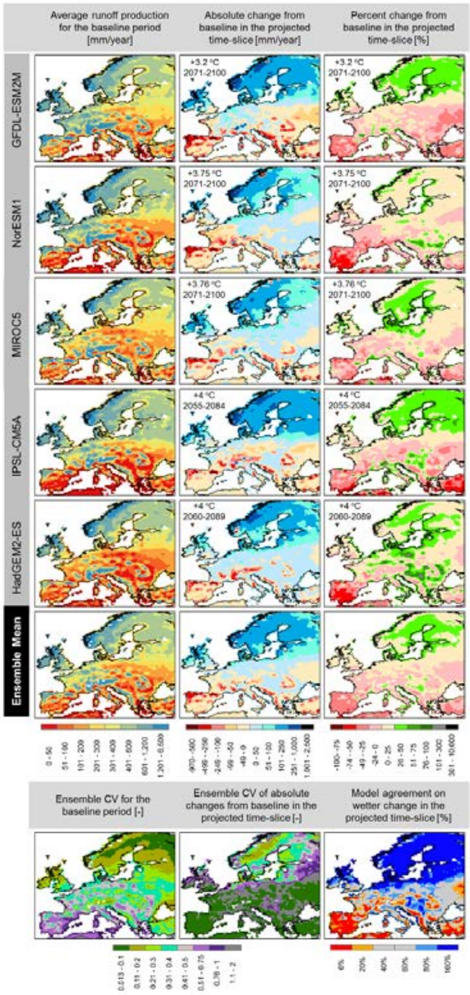
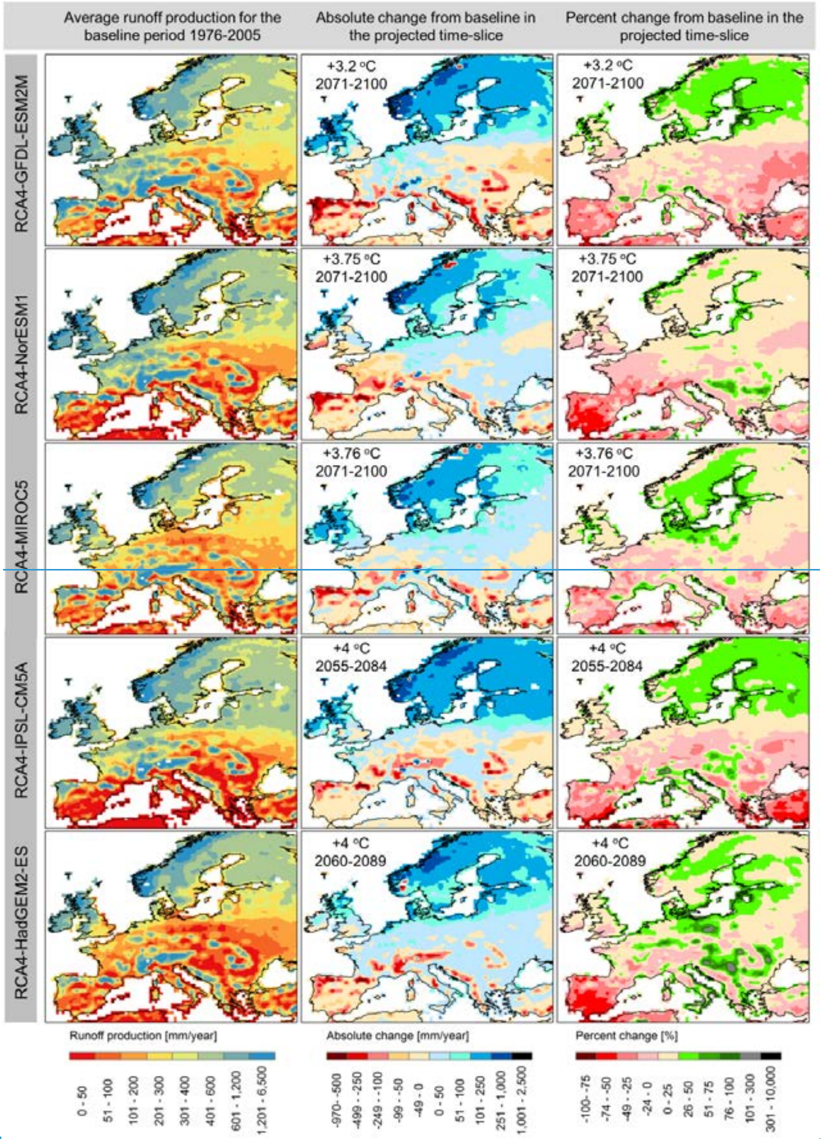


Figure 2. Average runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. Runoff production averaged over the baseline period (1976-2005) (left column), absolute change in runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL

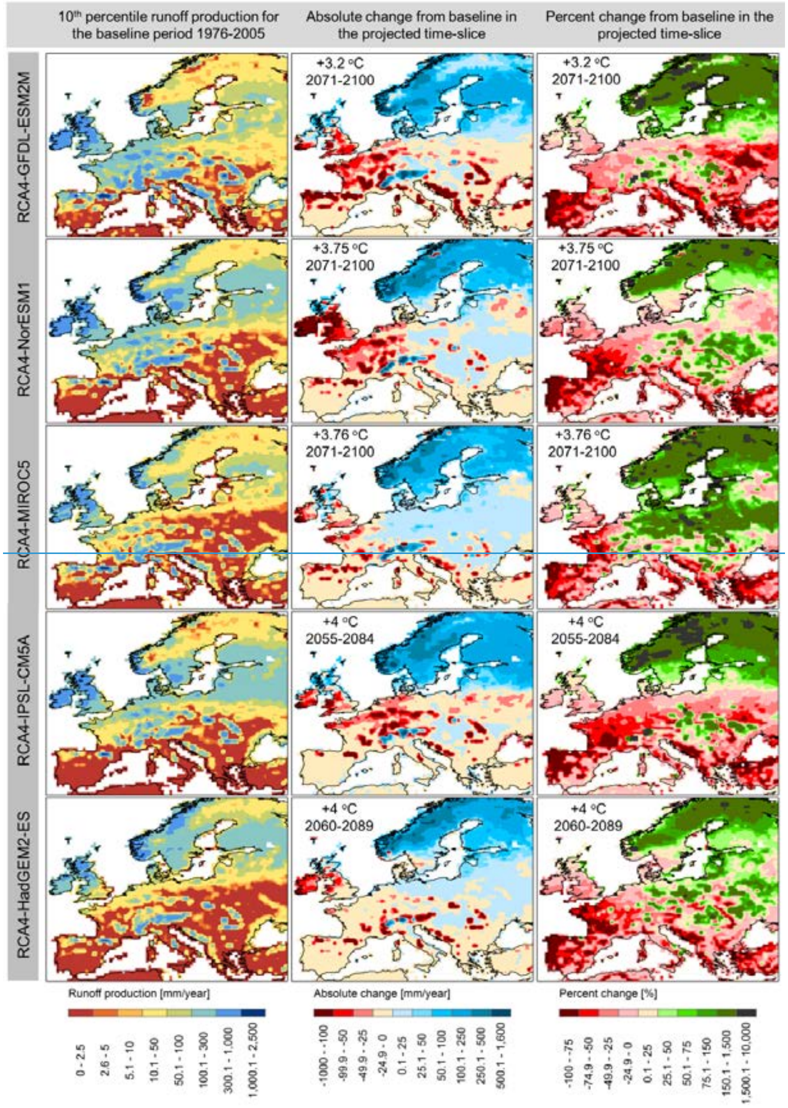
Formatted: Greek, Do not check spelling or grammar projected time-



slice.

~~Figure 2. Average runoff production from Euro-CORDEX data for all dynamical downscaled GCMs. Runoff production averaged over the baseline period (1976-2005) (left column), absolute change in runoff in the projected time-slice (middle column) and percent change in the projected time-slice (right column).~~

Formatted: English (United States)


10th percentile runoff production for the baseline period 1976-2005
Absolute change from baseline in the projected time-slice
Percent change from baseline in the projected time-slice
RCA4-GFDL-ESM2M
+3.2 °C
2071-2100
RCA4-NorESM1
+3.75 °C
2071-2100
RCA4-MIROC5
+3.76 °C
2071-2100
RCA4-IPSL-CM5A
+4 °C
2055-2084
RCA4-HadGEM2-ES
+4 °C
2060-2089
Runoff production [mm/year]
0 - 2.5
2.6 - 5
5.1 - 10
10.1 - 50
50.1 - 100
100.1 - 300
300.1 - 1,000
1,000.1 - 2,500
Absolute change [mm/year]
-1000 - -100
-99.9 - -50
-49.9 - -25
-24.9 - 0
0.1 - 25
25.1 - 50
50.1 - 100
100.1 - 250
250.1 - 500
500.1 - 1,600
Percent change [%]
-100 - -75
-74.9 - -50
-49.9 - -25
-24.9 - 0
0.1 - 25
25.1 - 50
50.1 - 75
75.1 - 150
150.1 - 1,500
1,500.1 - 10,000

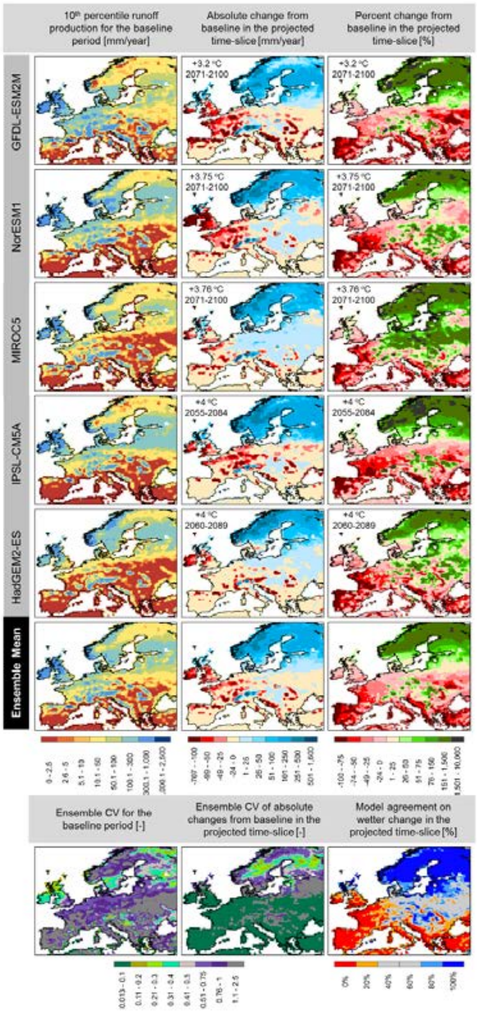


Figure 3. ~~10th percentile of runoff production from Euro-CORDEX data for all dynamical downscaled GCMs. 10th percentile runoff production derived on an annual basis and averaged over the baseline period (1976-2005) (left column), absolute change in 10th percentile runoff~~10th percentile of runoff production from raw Euro-CORDEX data for all dynamical downscaled GCMs and their ensemble mean. 10th percentile runoff production derived on an annual basis and averaged over the baseline period (1976–2005), absolute change in 10th percentile runoff in the +4 SWL projected time-slice (middle column) and percent change in the +4 SWL projected time-slice (right column). Bottom row: coefficient of

Formatted: Normal, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Not Superscript/ Subscript

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Black

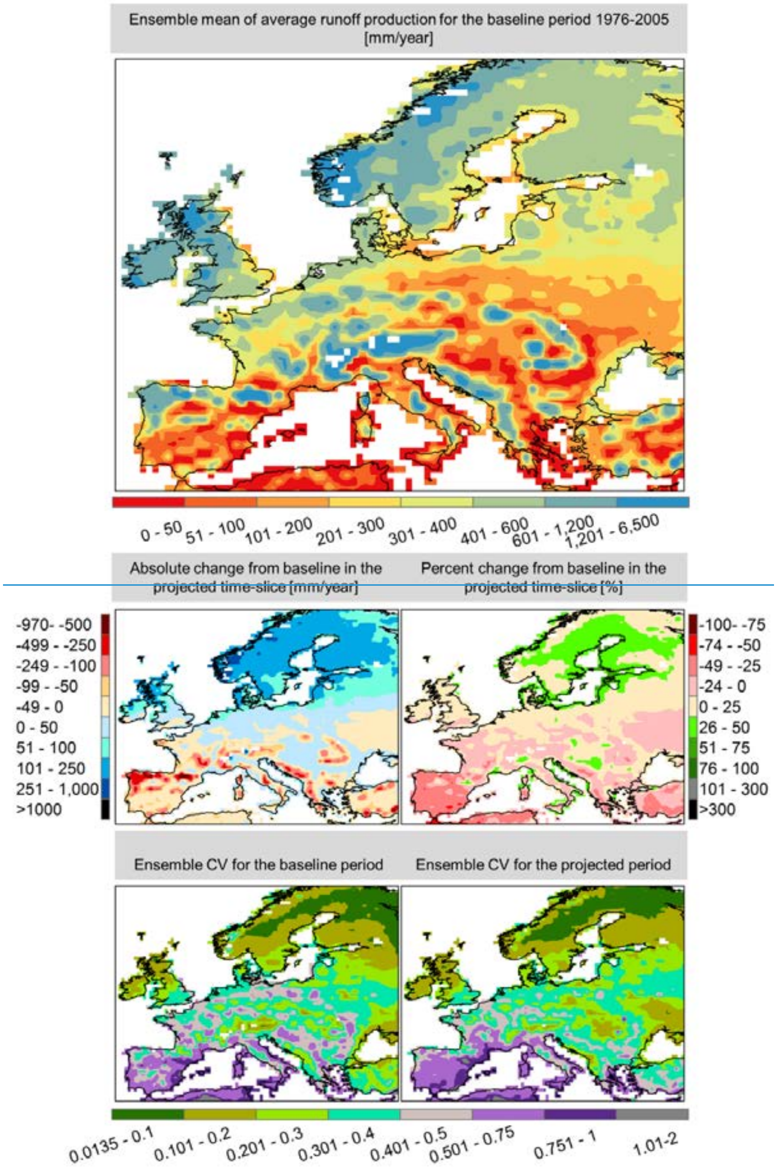
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Black

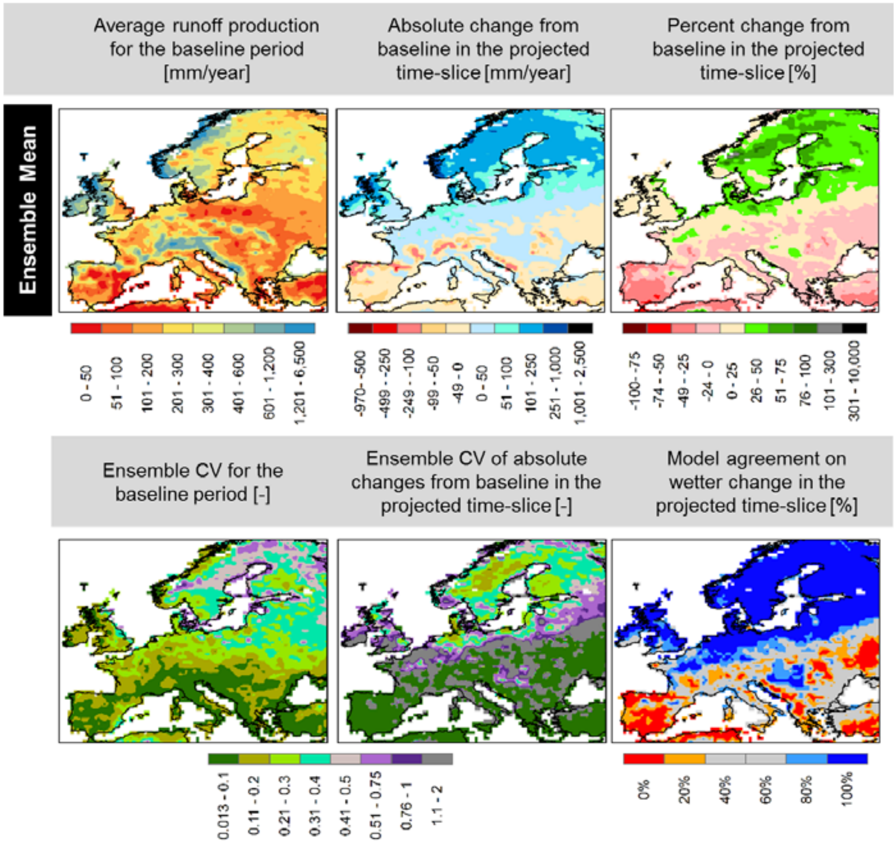
Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Black

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Black

Formatted: Font: (Default) Times New Roman, 12 pt, Font color: Black variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

in the projected time-slice (middle column) and percent change in the projected time-slice (right column).

1 
Ensemble mean of average runoff production for the baseline period 1976-2005 [mm/year]
0 - 50
51 - 100
101 - 200
201 - 300
301 - 400
401 - 600
601 - 1,200
1,201 - 6,500
Absolute change from baseline in the projected time-slice [mm/year]
Percent change from baseline in the projected time-slice [%]
-970- -500
-499 - -250
-249 - -100
-99 - -50
-49 - 0
0 - 50
51 - 100
101 - 250
251 - 1,000
>1000
-100- -75
-74 - -50
-49 - -25
-24 - 0
0 - 25
26 - 50
51 - 75
76 - 100
101 - 300
>300
Ensemble CV for the baseline period
Ensemble CV for the projected period
0.0135 - 0.1
0.101 - 0.2
0.201 - 0.3
0.301 - 0.4
0.401 - 0.5
0.501 - 0.75
0.751 - 1
1.01-2



~~F~~Figure 4. Ensemble mean of average runoff production ~~based on~~from -Euro-CORDEX data bias adjusted against the WFDEI dataset~~sets~~. Top row: Runoff production averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice ~~(middle row)~~. Bottom~~,~~ row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.~~coefficient of variation of the ensemble members for the baseline and projected period (bottom row).~~

Formatted: English (United States)

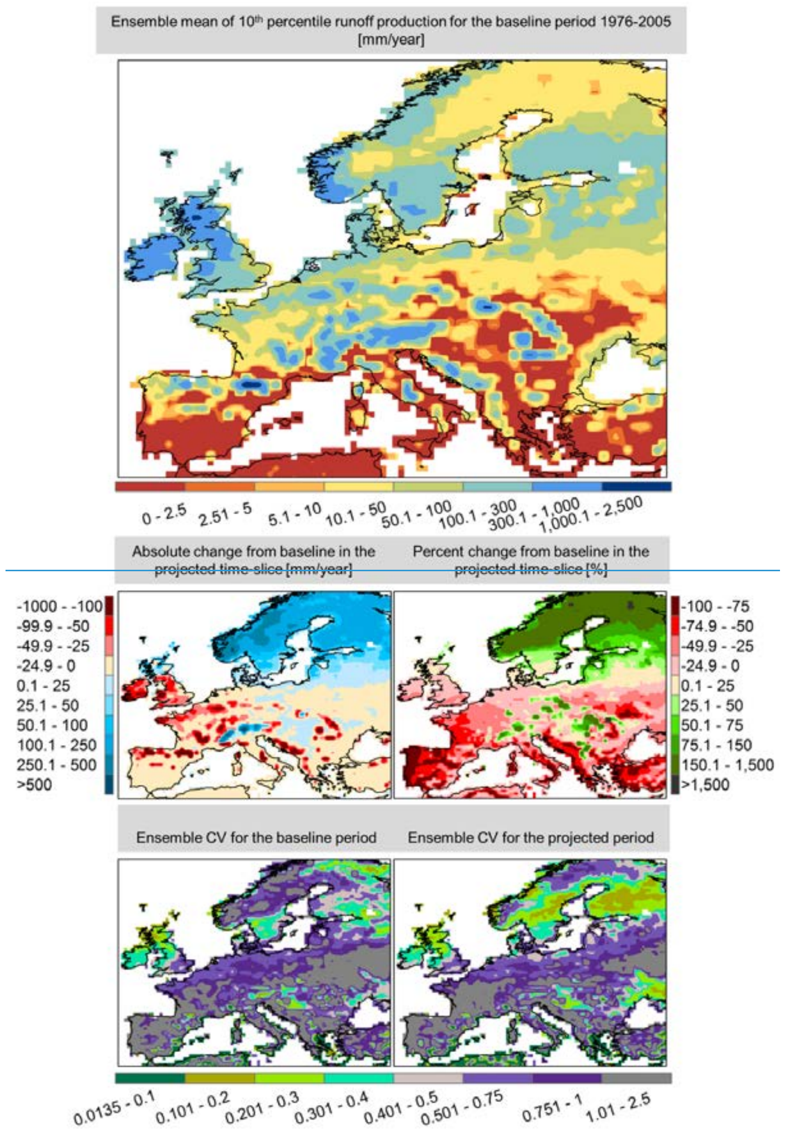


Figure 5. Ensemble mean of 10th percentile runoff production based on Euro-CORDEX datasets. 10th percentile runoff production derived on an annual basis averaged over the baseline period (1976-2005) (top row), absolute and percent change in ensemble mean of 10th percentile runoff in the projected time-slice (middle row), coefficient of variation of the ensemble members for the baseline and projected period (bottom row).

Formatted: Normal

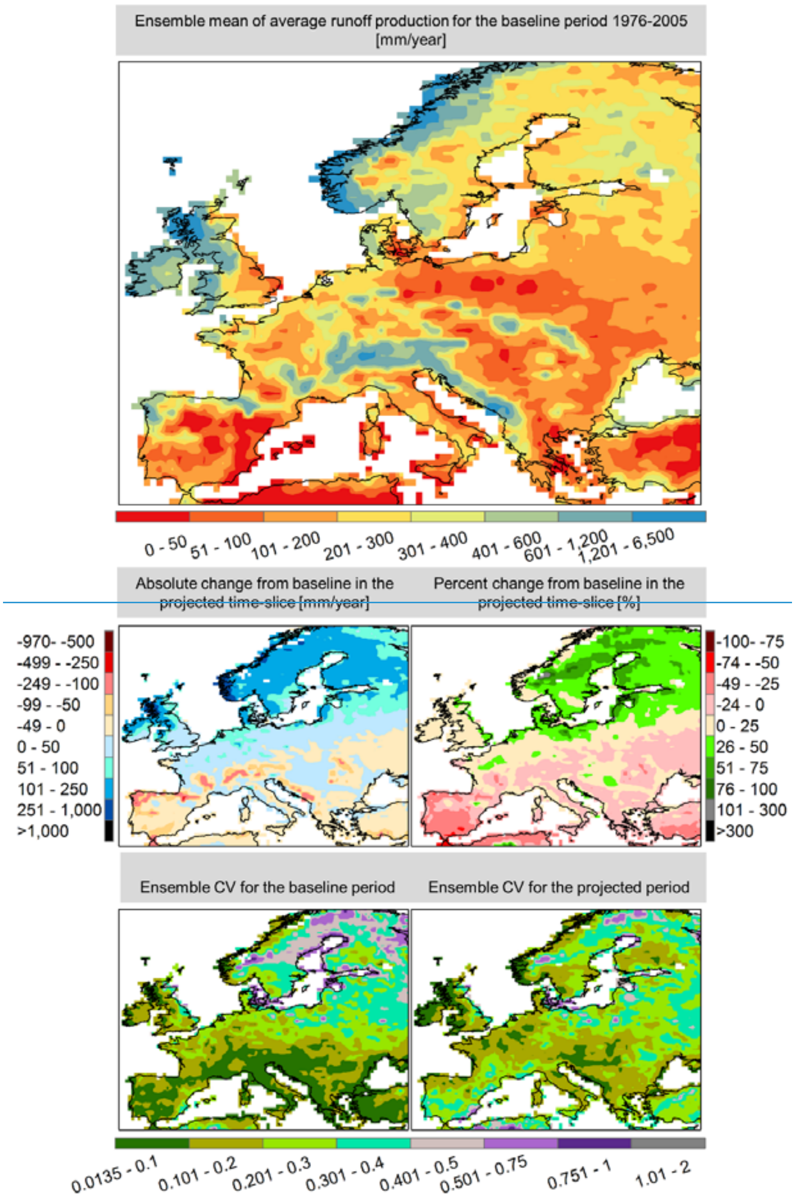


Figure 6. As in Figure 3, but for bias adjusted Euro-CORDEX data (Precipitation and Temperature) against WFDEI data.

Formatted: Normal

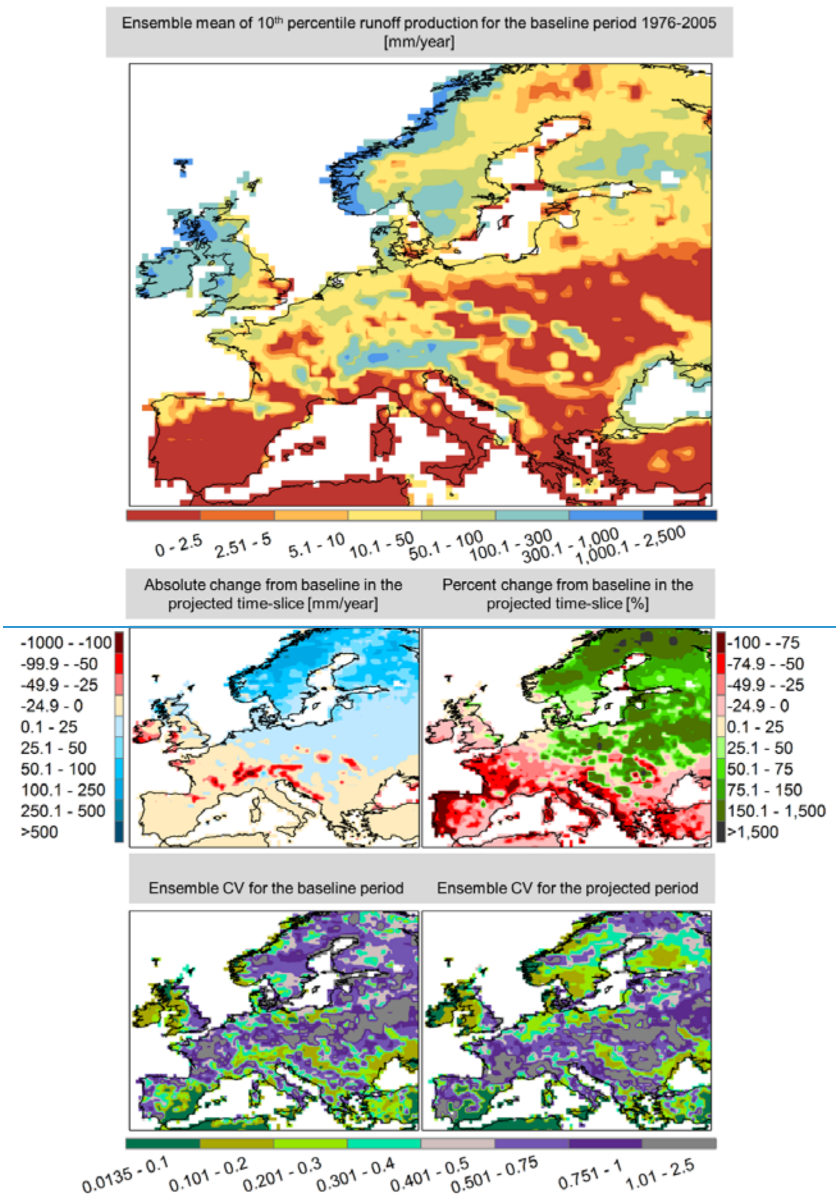
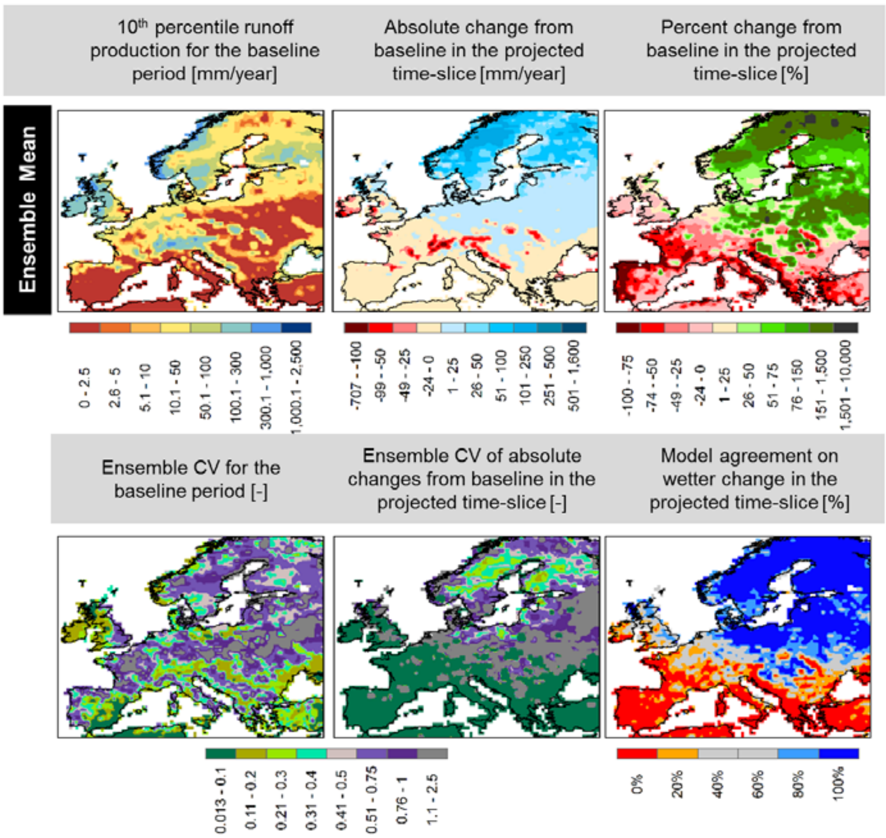


Figure 7. As in Figure 4, but for bias adjusted Euro-CORDEX data (Precipitation and Temperature) against WFDEI data

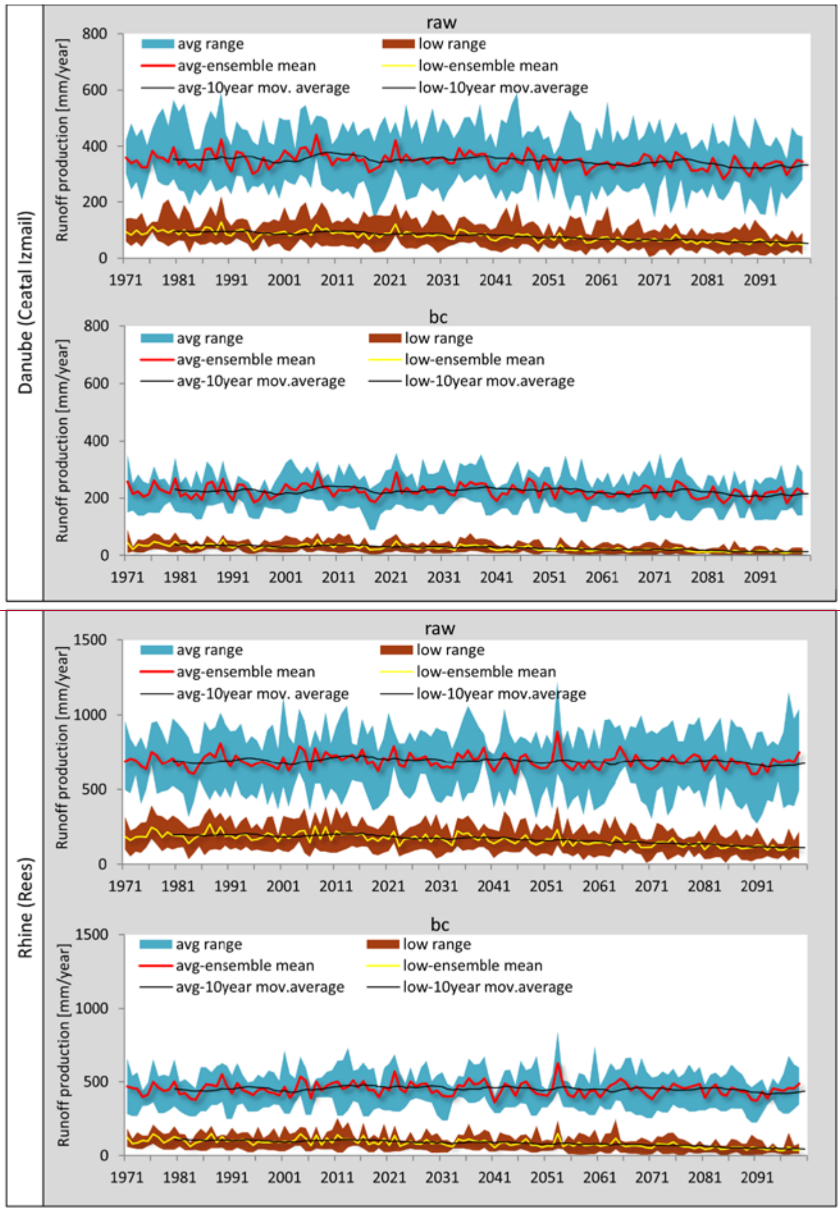


~~Figure 5~~Figure 5. Ensemble mean of 10th percentile runoff production from Euro-CORDEX data bias adjusted against the WFDEI dataset. Top row: 10th percentile runoff production derived on an annual basis averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice. Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

Formatted: Caption

Formatted: Superscript

Formatted: Superscript


Danube (Ceatal Izmail)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
Runoff production [mm/year]
bc
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov.average
low-10year mov.average
Runoff production [mm/year]
Rhine (Rees)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
Runoff production [mm/year]
bc
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov.average
low-10year mov.average
Runoff production [mm/year]
1971 1981 1991 2001 2011 2021 2031 2041 2051 2061 2071 2081 2091

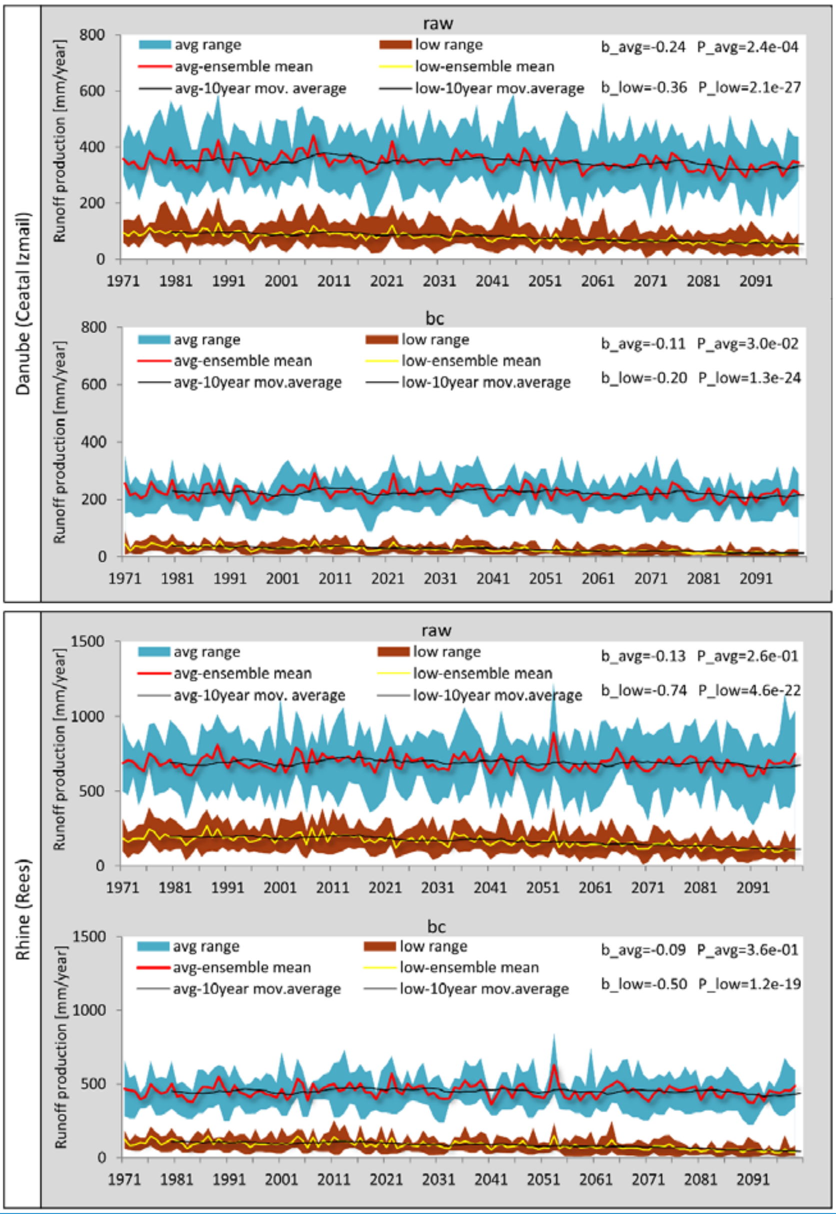
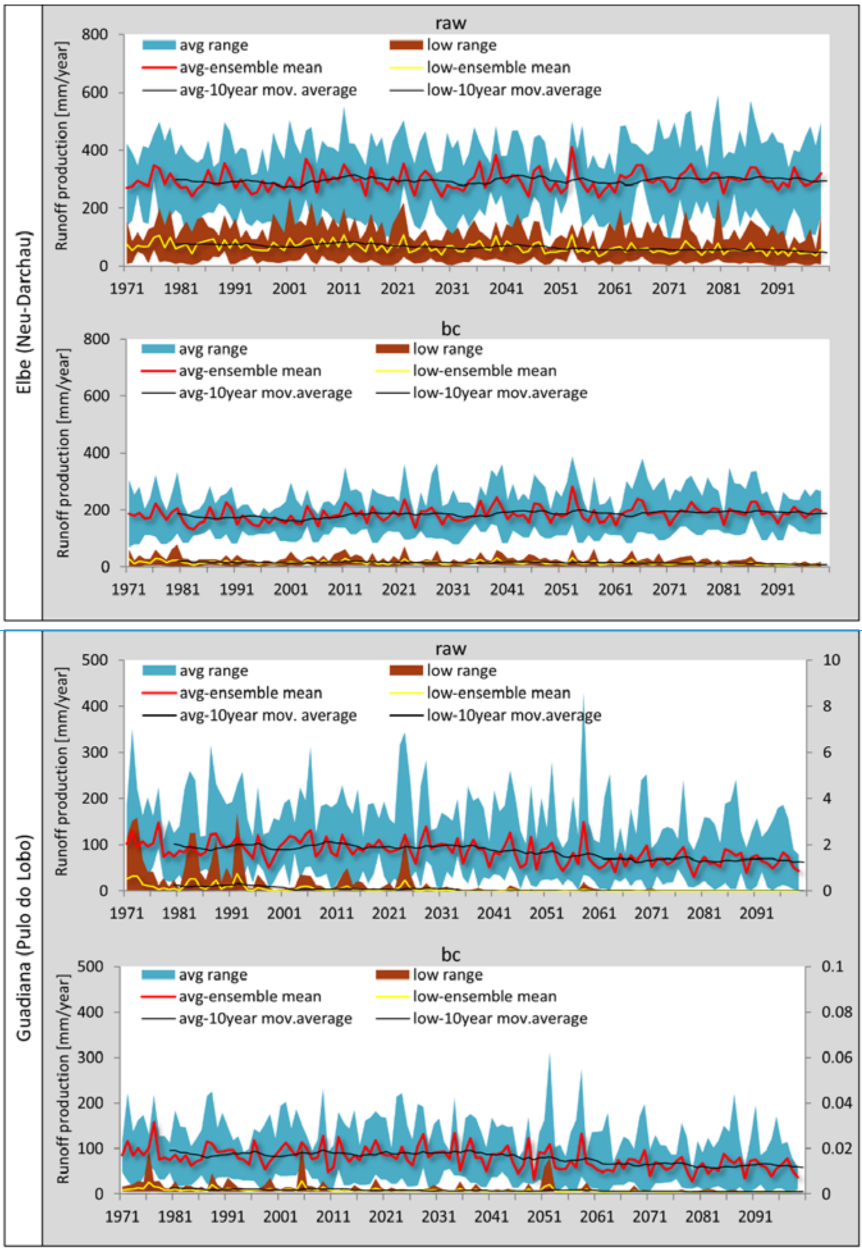
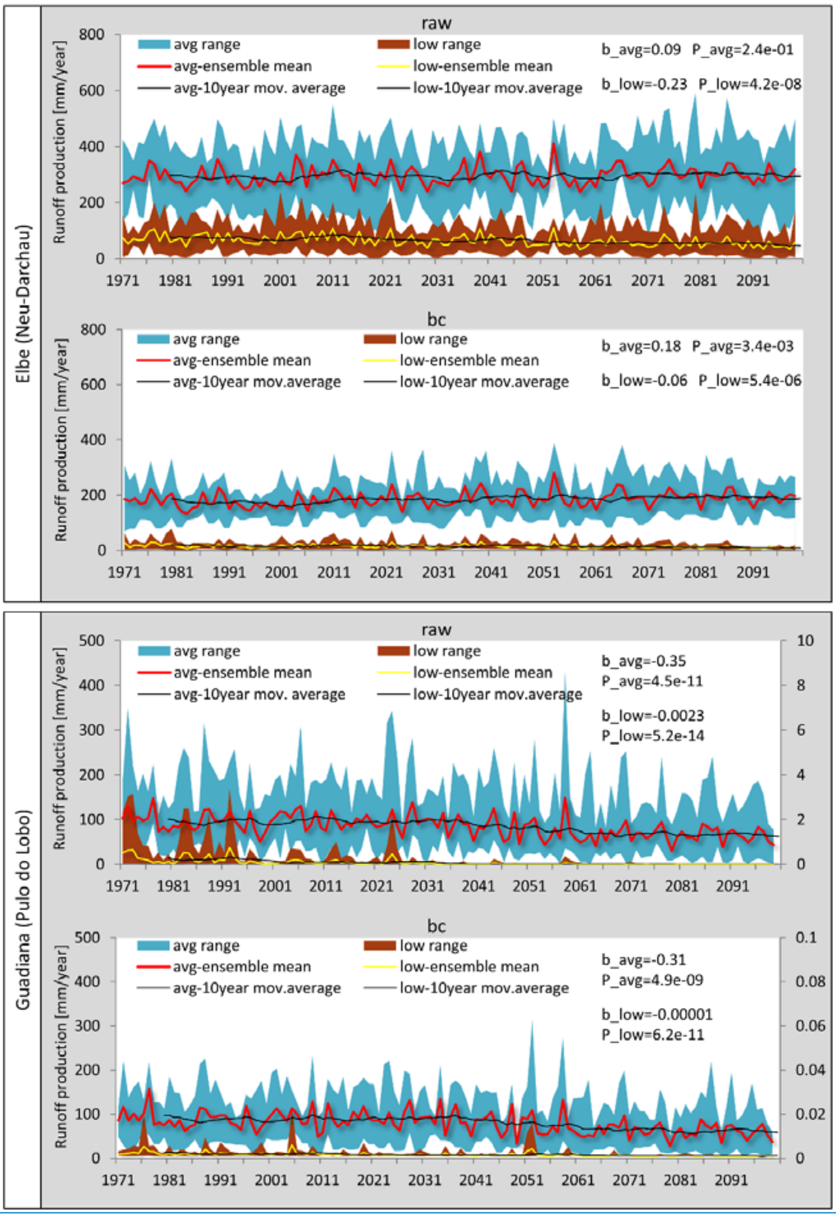


Figure 6. ~~Figure 86.~~ Annual time-series of basin averaged runoff production (average and 10th percentile of annual runoff) for raw and bias adjusted Euro-CORDEX data. For both average

and 10th percentile time-series, the ensemble range, mean and 10-year moving average is shown.

1 
Elbe (Neu-Darchau)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
Runoff production [mm/year]
bc
avg-10year mov.average
Guadiana (Pulo do Lobo)


Elbe (Neu-Darchau)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
b_avg=0.09 P_avg=2.4e-01
b_low=-0.23 P_low=4.2e-08
Runoff production [mm/year]
bc
b_avg=0.18 P_avg=3.4e-03
b_low=-0.06 P_low=5.4e-06
Guadiana (Pulo do Lobo)
b_avg=-0.35
P_avg=4.5e-11
b_low=-0.0023
P_low=5.2e-14
b_avg=-0.31
P_avg=4.9e-09
b_low=-0.00001
P_low=6.2e-11
1971 1981 1991 2001 2011 2021 2031 2041 2051 2061 2071 2081 2091

Figure

6

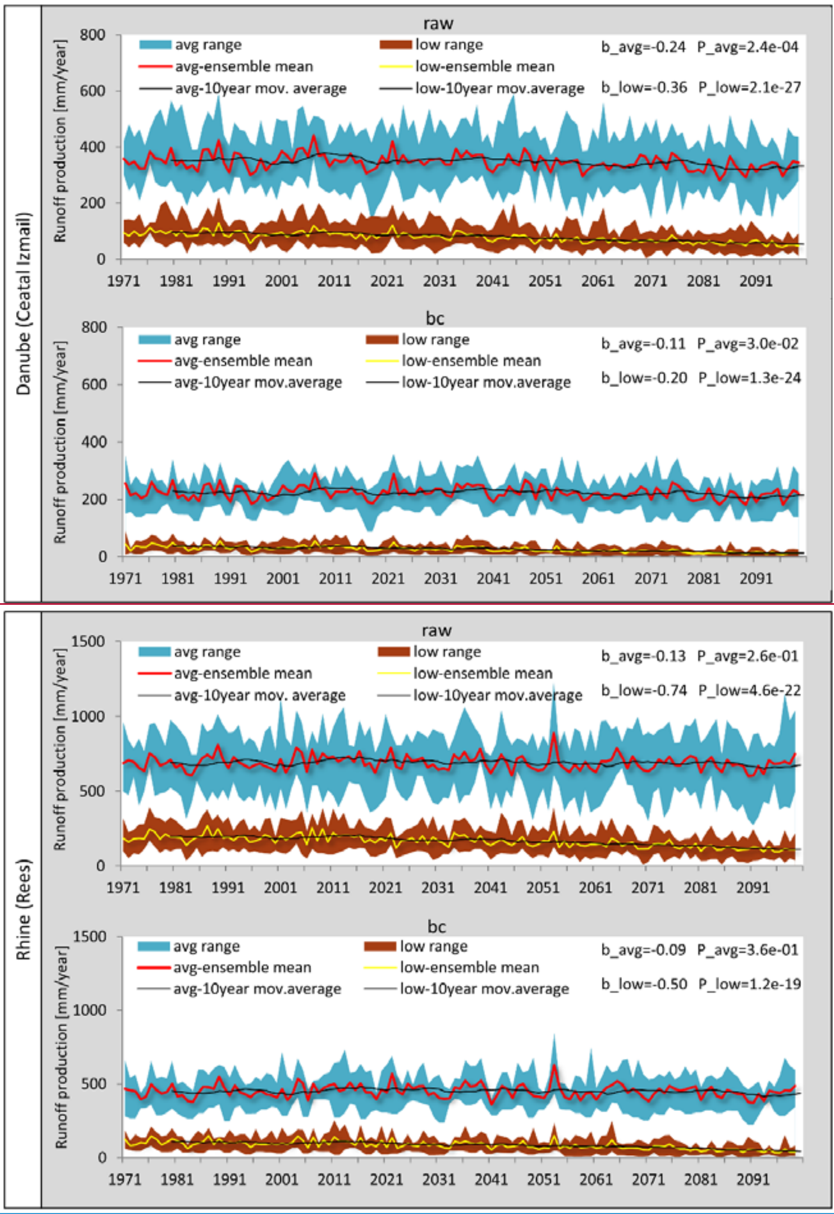
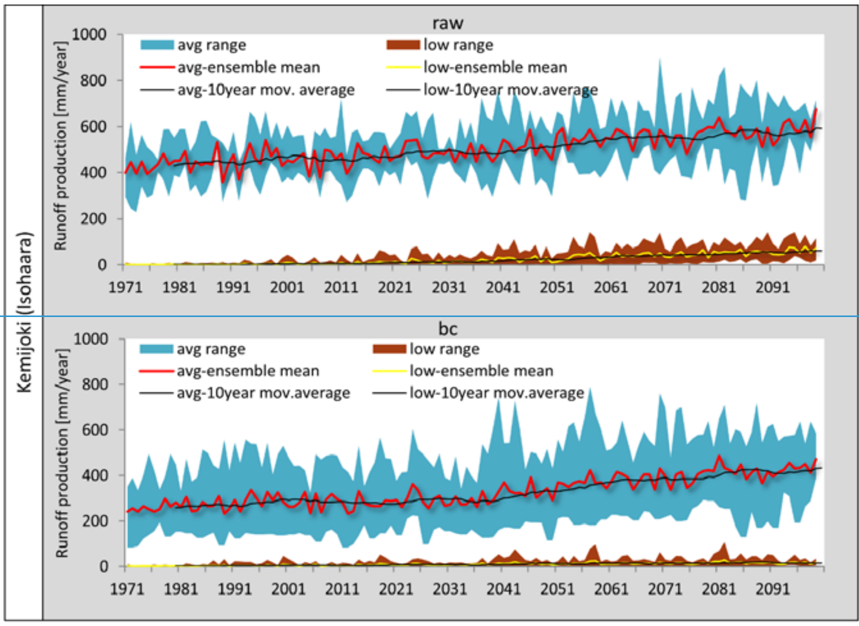


Figure Figure 76 (continued)


Kemijoki (Isohaara)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
Runoff production [mm/year]
bc
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov.average
low-10year mov.average
Runoff production [mm/year]

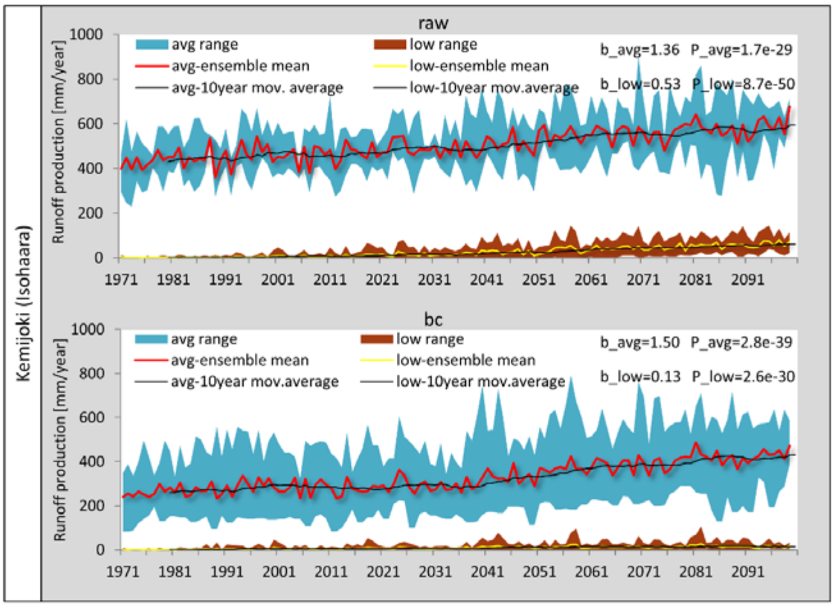

Kemijoki (Isohaara)
raw
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov. average
low-10year mov.average
b_avg=1.36 P_avg=1.7e-29
b_low=0.53 P_low=8.7e-50
Runoff production [mm/year]
bc
avg range
low range
avg-ensemble mean
low-ensemble mean
avg-10year mov.average
low-10year mov.average
b_avg=1.50 P_avg=2.8e-39
b_low=0.13 P_low=2.6e-30
Runoff production [mm/year]

Figure 6

(continued)

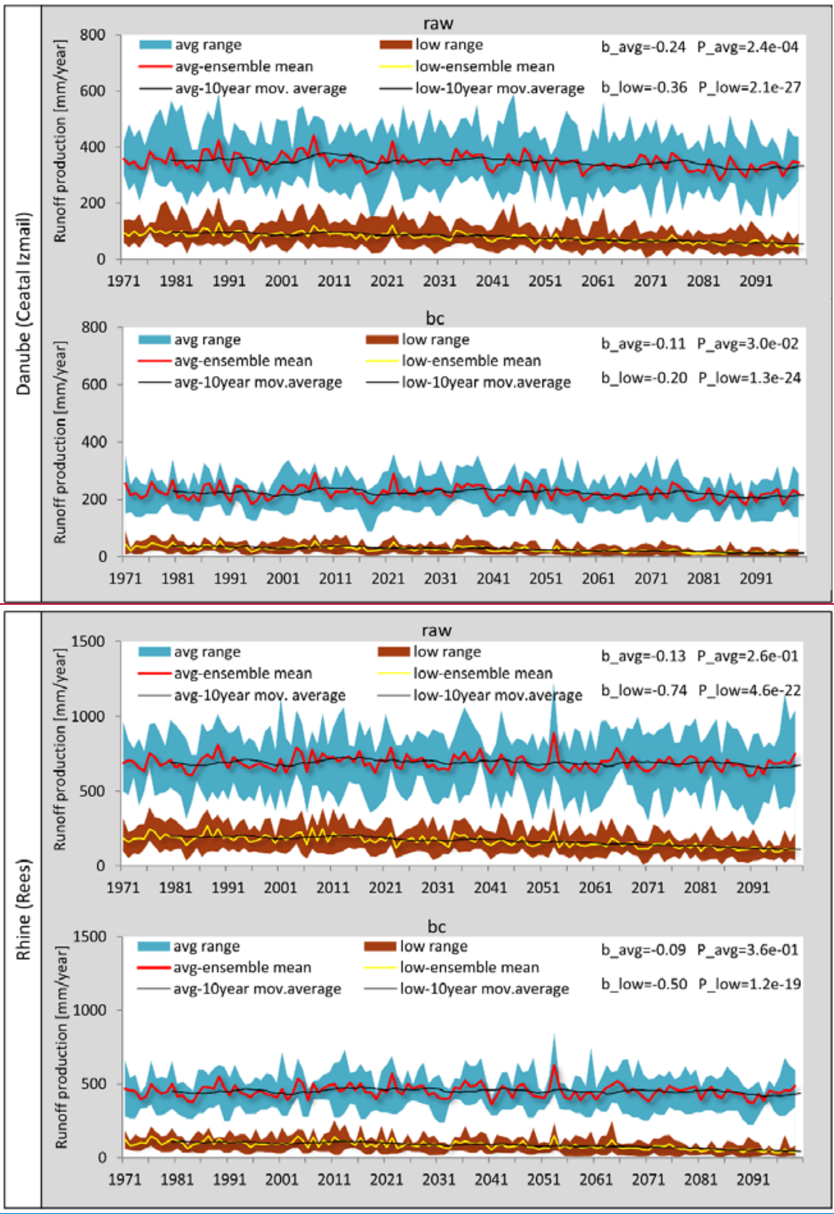
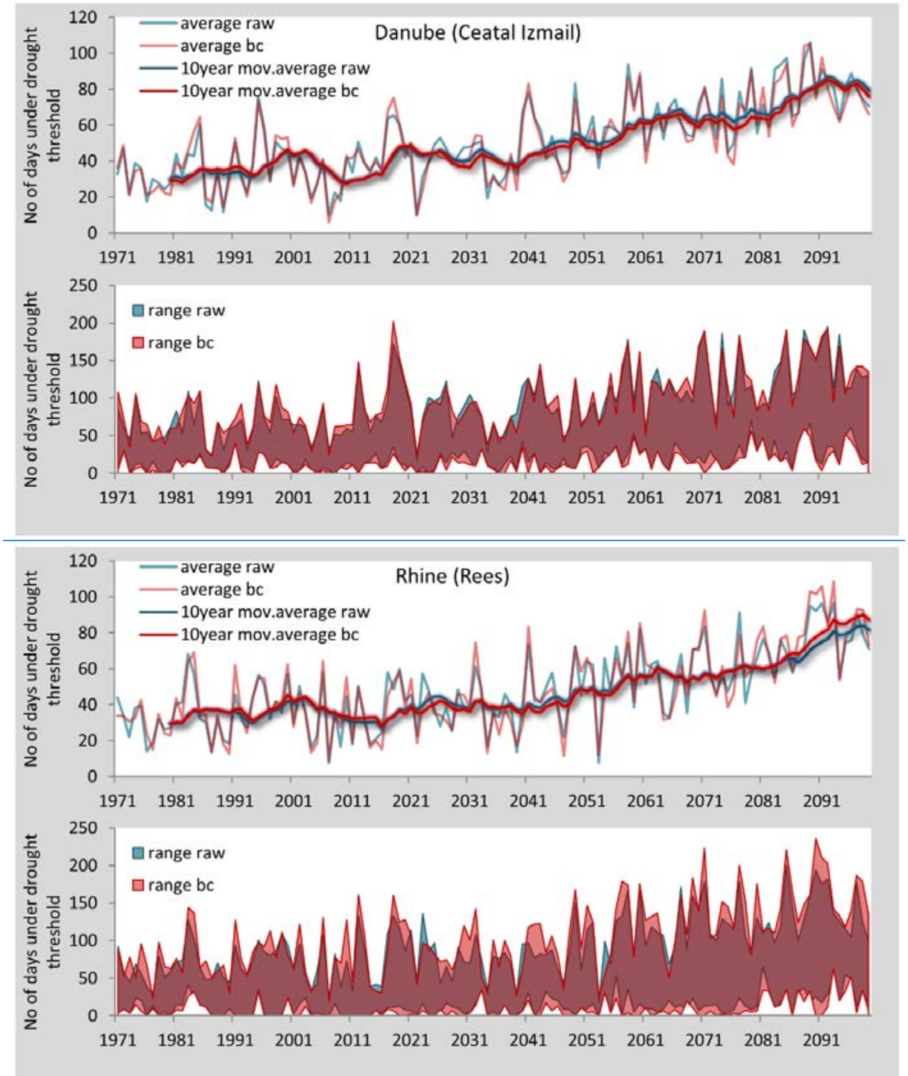


Figure Figure 76 (continued)

Formatted: Caption



Danube (Ceatal Izmail)
average raw
average bc
10year mov.average raw
10year mov.average bc
No of days under drought threshold
range raw
range bc
Rhine (Rees)
average raw
average bc
10year mov.average raw
10year mov.average bc
No of days under drought threshold
range raw
range bc

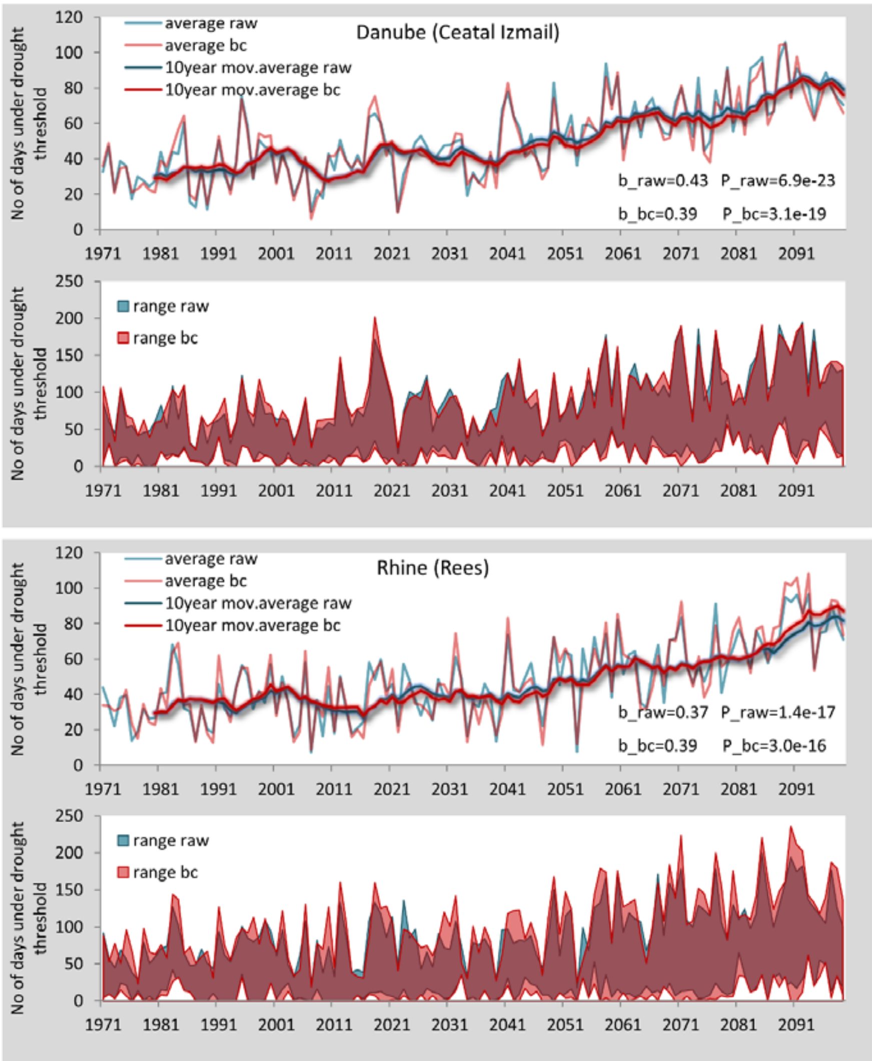
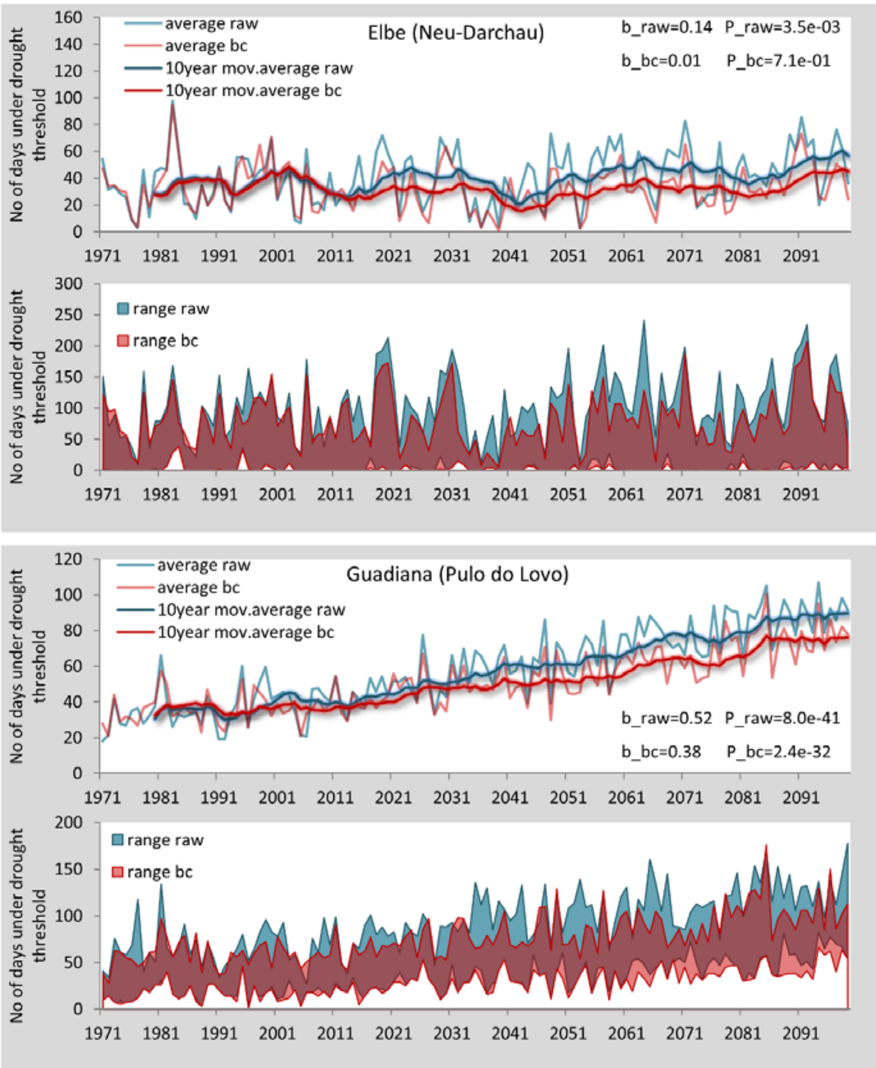


Figure 7. ~~Figure 7~~9. Number of days under drought threshold per year for raw and bias adjusted Euro-CORDEX data. Ensemble mean and 10-year moving average of the ensemble mean (top), ensemble range (bottom).


Elbe (Neu-Darchau)
average raw
average bc
10year mov.average raw
10year mov.average bc
b_raw=0.14 P_raw=3.5e-03
b_bc=0.01 P_bc=7.1e-01
No of days under drought threshold
range raw
range bc
Guadiana (Pulo do Lovo)
b_raw=0.52 P_raw=8.0e-41
b_bc=0.38 P_bc=2.4e-32
1971 1981 1991 2001 2011 2021 2031 2041 2051 2061 2071 2081 2091

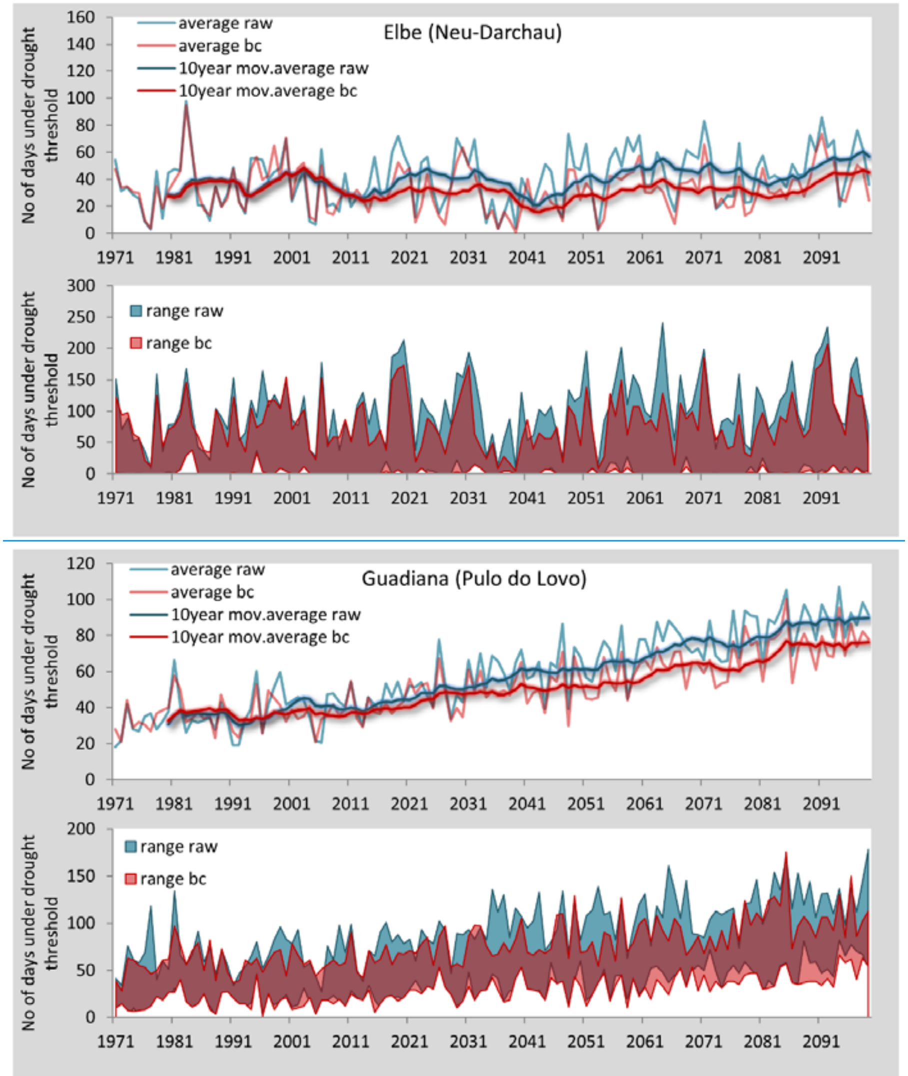
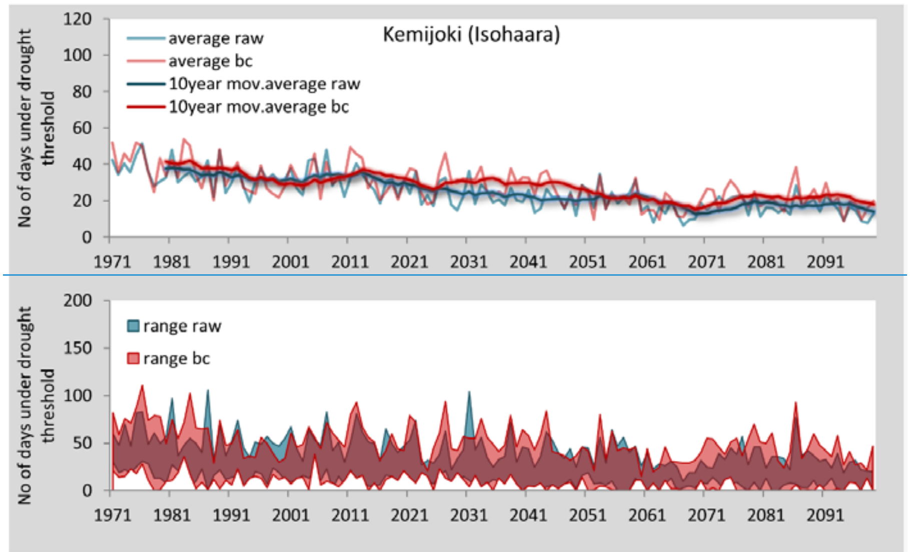


Figure 7 (continued)Figure 7Figure 87 (continued)



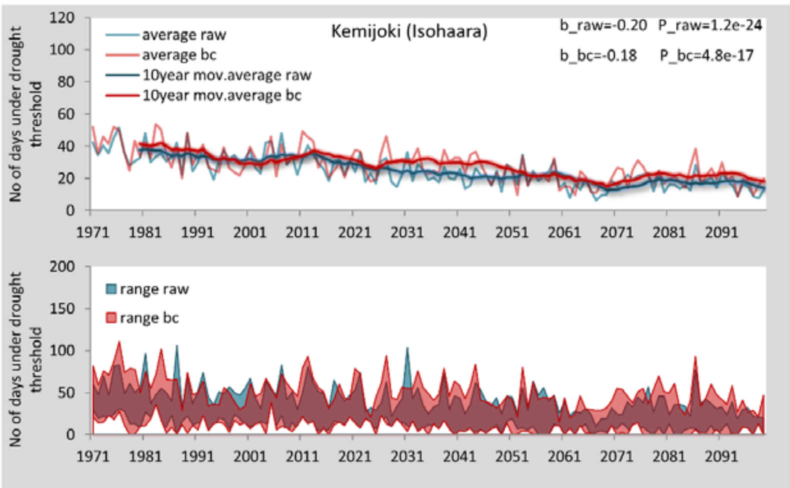


Figure 7 (continued)

Figure 7Figure 87 (continued)

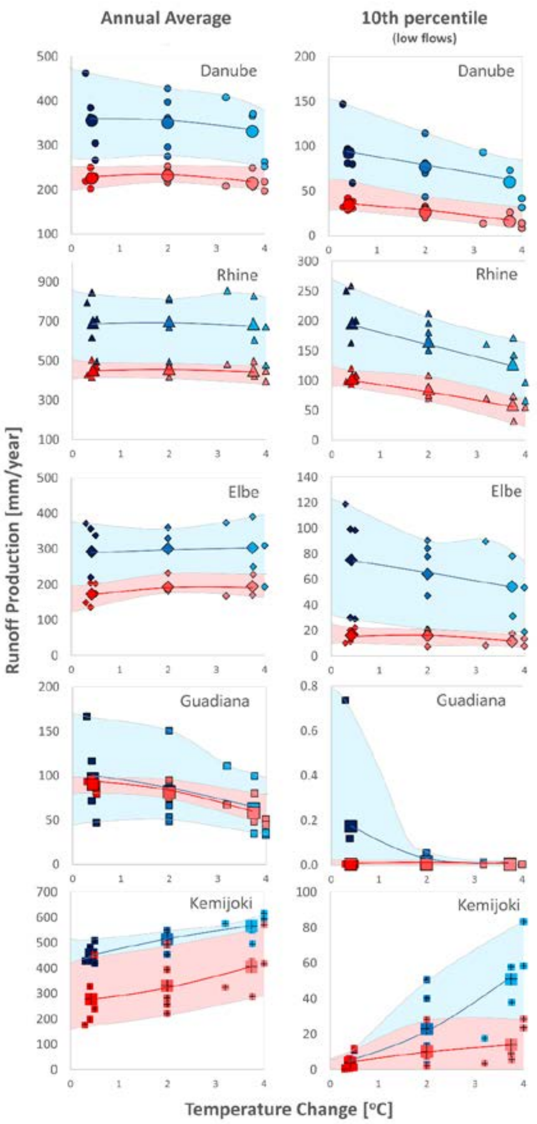


Figure 8. Figure 10 8. Variation of runoff production with respect to temperature change (+2 and +4 SWLs) for raw (light blue) and bias adjusted (light red) Euro-CORDEX data, for both annual average (left column) and 10th percentile (right column) runoff production. Small markers represent the value of each individual model and bigger markers correspond to ensemble mean value.

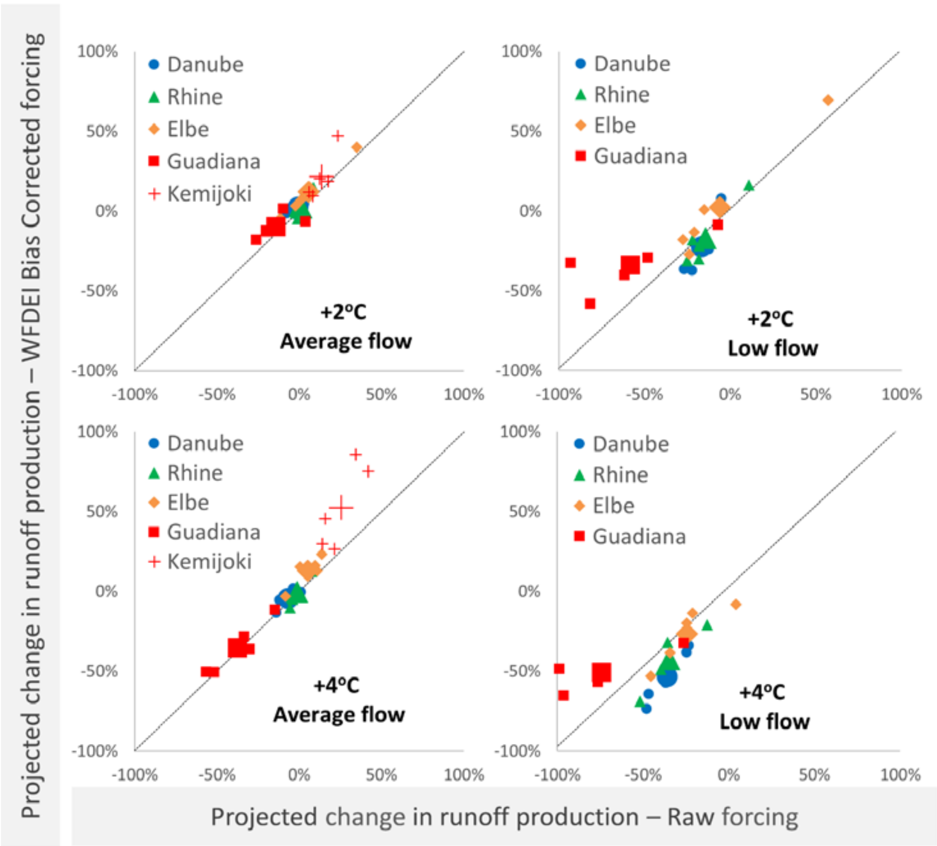


Figure 9. Figure 911. Correlation between projected change in basin averaged runoff production derived from WFDEI-bias adjusted and raw Euro-CORDEX data, for both annual average (left) and 10th percentile (right) runoff production. Correlation is examined at +2°C SWL (top) and at +4°C SWL (bottom). Small markers represent the value of each individual model and bigger markers correspond to ensemble mean value.

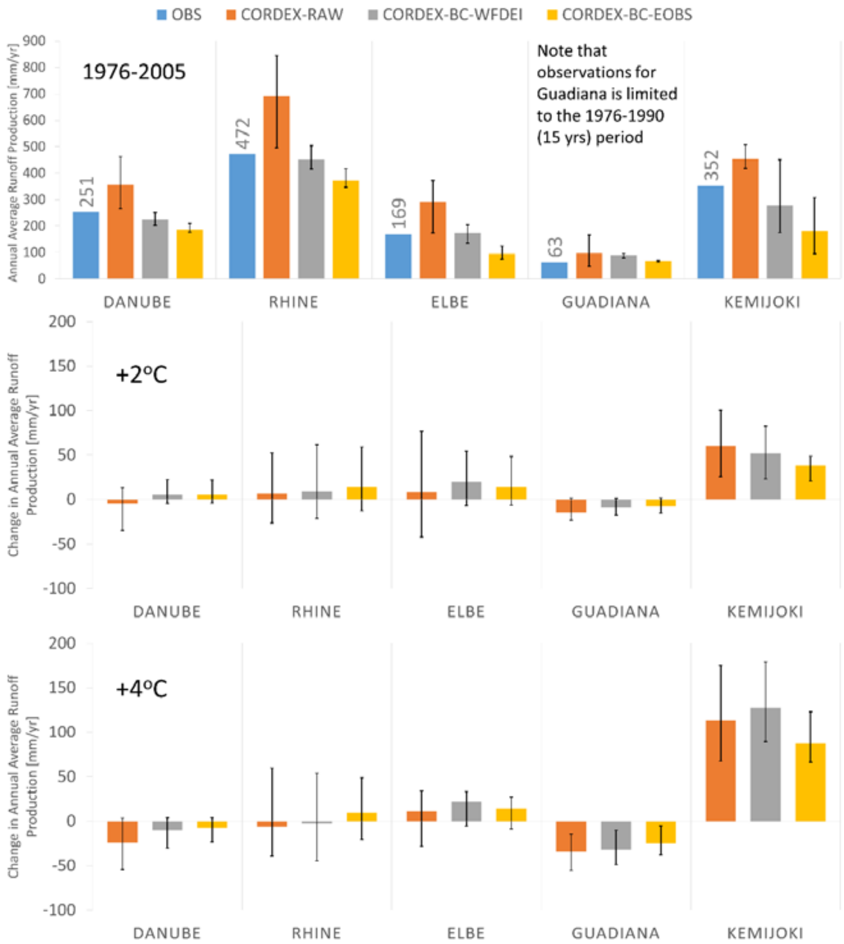


Figure 10. Comparison between the simulations of raw Euro-CORDEX data and bias adjusted against two different datasets (WFDEI and E-OBS) for five study basins. Bars show the ensemble means and error bars the minimum and maximum ensemble member values. (Top row) Annual average runoff production for the period 1976 to 2005.OBS values are derived from GRDC discharge measurements converted to basin averages at the annual time-scale. (Middle row) Percent change in annual average runoff production at the +2 SWL and (bottom row) at the +4 SWL.

~~Figure 5120. Figure 10. Comparison between the simulations of raw Euro-CORDEX data and bias adjusted against two different datasets (WFDEI and E-OBS) for five study basins. Bars~~

Formatted: Keep with next

Formatted: Caption

show the ensemble means and error bars the minimum and maximum ensemble member values. (Top row) Annual average runoff production for the period 1976 to 2005.OBS values are derived from GRDC discharge measurements converted to basin averages at the annual time scale. (Middle row) Percent change in annual average runoff production at the +2 SWL and (bottom row) at the +4 SWL.

SWL.

Supplementary

~~Table S 1. Results of linear regression applied to basin aggregated annual average runoff production for raw and bias adjusted Euro- CORDEX data.~~

Formatted: Left: 2.5 cm, Right: 2.5 cm, Top: 2 cm, Bottom: 3.5 cm

	Basin's Annual Average Runoff Production [mm/year]													
-	Raw							Bias Corrected						
Danube		Coeff.	St. Error	tStat	P-value	r	0.32	-	Coeff.	St. Error	tStat	P-value	r	
Danube	Interc.	829.12	127.91	6.48	1.82E-09	R^2	0.10	Interc.	451.47	104.08	4.34	2.91E-05	R^2	
Danube	X	-0.24	0.06	-3.77	2.45E-04	Adj. R^2	0.09	X	-0.11	0.05	-2.19	3.02E-02	Adj. R^2	
Rhine	-	Coeff.	St. Error	tStat	P-value	r	0.10	-	Coeff.	St. Error	tStat	P-value	r	
Rhine	Interc.	950.24	228.55	4.16	5.87E-05	R^2	0.01	Interc.	640.82	204.57	3.13	2.15E-03	R^2	
Rhine	X	-0.13	0.11	-1.14	2.58E-01	Adj. R^2	0.00	X	-0.09	0.10	-0.93	3.56E-01	Adj. R^2	
Elbe	-	Coeff.	St. Error	tStat	P-value	r	0.10	-	Coeff.	St. Error	tStat	P-value	r	
Elbe	Interc.	112.23	155.05	0.72	4.70E-01	R^2	0.01	Interc.	-171.71	119.48	-1.44	1.53E-01	R^2	
Elbe	X	0.09	0.08	1.18	2.39E-01	Adj. R^2	0.00	X	0.18	0.06	2.99	3.38E-03	Adj. R^2	
Guadiana	-	Coeff.	St. Error	tStat	P-value	r	0.54	-	Coeff.	St. Error	tStat	P-value	r	
Guadiana	Interc.	794.88	98.58	8.06	4.76E-13	R^2	0.29	Interc.	713.59	100.97	7.07	9.31E-11	R^2	
Guadiana	X	-0.35	0.05	-7.21	4.46E-11	Adj. R^2	0.28	X	-0.31	0.05	-6.28	4.87E-09	Adj. R^2	
Kemijoki	-	Coeff.	St. Error	tStat	P-value	r	0.80	-	Coeff.	St. Error	tStat	P-value	r	
Kemijoki	Interc.	-2257.94	186.45	-12.11	6.46E-23	R^2	0.63	Interc.	-2717.09	159.07	-17.08	1.06E-34	R^2	
Kemijoki	X	1.36	0.09	14.83	1.72E-29	Adj. R^2	0.63	X	1.50	0.08	19.16	2.81E-39	Adj. R^2	0.74

Formatted: Position: Horizontal: Left, Relative to: Margin, Vertical: 0.26 cm, Relative to: Paragraph, Horizontal: 0.32 cm, Wrap Around
Formatted: Caption, Left
Formatted: Caption
Formatted: Caption
Formatted: Caption, Left
Formatted: Caption, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption, Left, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted ...
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted ...
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted ...
Formatted: Caption
Formatted ...
Formatted: Caption
Formatted: Caption, Don't keep with next

1

-	Basin's Annual 10th percentile Runoff Production [mm/year]												
-	Raw							Bias Corrected					
Danube		Coeff.	St. Error	tStat	P-value	r	0.78	-	Coeff.	St. Error	tStat	P-value	r
	Interc.	817.99	53.05	15.42	6.94E-31	R^2	0.61	Interc.	442.02	32.50	13.60	1.49E-26	R^2
	X	-0.36	0.03	-13.96	2.09E-27	Adj. R^2	0.60	X	-0.20	0.02	-12.80	1.29E-24	Adj. R^2
Rhine	-	Coeff.	St. Error	tStat	P-value	r	0.72	-	Coeff.	St. Error	tStat	P-value	r
	Interc.	1665.80	127.58	13.06	3.13E-25	R^2	0.52	Interc.	1102.30	94.45	11.67	7.82E-22	R^2
	X	-0.74	0.06	-11.76	4.59E-22	Adj. R^2	0.52	X	-0.50	0.05	-10.78	1.21E-19	Adj. R^2
Elbe	-	Coeff.	St. Error	tStat	P-value	r	0.46	-	Coeff.	St. Error	tStat	P-value	r
	Interc.	530.57	79.89	6.64	8.18E-10	R^2	0.21	Interc.	139.24	26.24	5.31	4.84E-07	R^2
	X	-0.23	0.04	-5.84	4.19E-08	Adj. R^2	0.21	X	-0.06	0.01	-4.75	5.40E-06	Adj. R^2
Guadiana	-	Coeff.	St. Error	tStat	P-value	r	0.60	-	Coeff.	St. Error	tStat	P-value	r
	Interc.	4.70	0.55	8.61	2.35E-14	R^2	0.36	Interc.	0.02	0.00	7.63	4.97E-12	R^2
	X	0.00	0.00	-8.47	5.23E-14	Adj. R^2	0.36	X	0.00	0.00	-7.15	6.16E-11	Adj. R^2
Kemijoki	-	Coeff.	St. Error	tStat	P-value	r	0.91	-	Coeff.	St. Error	tStat	P-value	r
	Interc.	-1048.22	43.96	-23.85	9.80E-49	R^2	0.82	Interc.	-247.59	16.93	-14.62	5.35E-29	R^2
	X	0.53	0.02	24.41	8.67E-50	Adj. R^2	0.82	X	0.13	0.01	15.18	2.62E-30	Adj. R^2

Formatted: Caption
Formatted: Caption, Left
Formatted: Caption
Formatted: Caption, Left
Formatted: Caption, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption, Left, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted: Caption, Left, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted: Caption, Left, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption
Formatted: Caption
Formatted: Caption
Formatted: Caption, Left, Indent: Left: 0 cm, Right: 0 cm
Formatted: Caption
Formatted ...
Formatted: Caption
Formatted: Caption, Don't keep with next

Table S 2. Results of linear regression applied to basin aggregated annual 10th percentile runoff production for raw and bias adjusted Euro-CORDEX data.

Table S 3. Results of linear regression applied to the number of drought under the 10th percentile daily varying drought threshold for raw and bias adjusted Euro-CORDEX data.

-	**No of days under the 10th percentile daily varying drought threshold**											
-	**Raw**							**Bias Corrected**				
Danube		Coeff.	St. Error	tStat	P-value	r	0.73	-	Coeff.	St. Error	tStat	P-value	r	[illegible]
	Interc.	-813.73	71.53	-11.38	4.14E-21	R^2	0.54	Interc.	-747.33	75.15	-9.94	1.39E-17	R^2	[illegible]
	X	0.43	0.04	12.10	6.93E-23	Adj. R^2	0.53	X	0.39	0.04	10.61	3.13E-19	Adj. R^2	[illegible]
Rhine	-	Coeff.	St. Error	tStat	P-value	r	0.66	-	Coeff.	St. Error	tStat	P-value	r	[illegible]
	Interc.	-701.49	75.40	-9.30	5.11E-16	R^2	0.44	Interc.	-747.42	84.71	-8.82	7.40E-15	R^2	[illegible]
	X	0.37	0.04	9.94	1.42E-17	Adj. R^2	0.43	X	0.39	0.04	9.40	3.03E-16	Adj. R^2	[illegible]
Elbe	-	Coeff.	St. Error	tStat	P-value	r	0.26	-	Coeff.	St. Error	tStat	P-value	r	[illegible]
	Interc.	-242.24	94.92	-2.55	1.19E-02	R^2	0.07	Interc.	2.24	80.10	0.03	9.78E-01	R^2	[illegible]
	X	0.14	0.05	2.98	3.45E-03	Adj. R^2	0.06	X	0.01	0.04	0.38	7.06E-01	Adj. R^2	[illegible]
Guadiana	-	Coeff.	St. Error	tStat	P-value	r	0.87	-	Coeff.	St. Error	tStat	P-value	r	[illegible]
	Interc.	-990.84	52.77	-18.78	1.86E-38	R^2	0.76	Interc.	-714.14	47.70	-14.97	7.94E-30	R^2	[illegible]
	X	0.52	0.03	19.89	7.95E-41	Adj. R^2	0.76	X	0.38	0.02	16.05	2.37E-32	Adj. R^2	[illegible]
Kemijoki	-	Coeff.	St. Error	tStat	P-value	r	0.75	-	Coeff.	St. Error	tStat	P-value	r	[illegible]
	Interc.	428.42	31.51	13.60	1.52E-26	R^2	0.56	Interc.	395.55	37.80	10.46	7.34E-19	R^2	0.43
	X	-0.20	0.02	-12.81	1.23E-24	Adj. R^2	0.56	X	-0.18	0.02	-9.73	4.75E-17	Adj. R^2	0.42

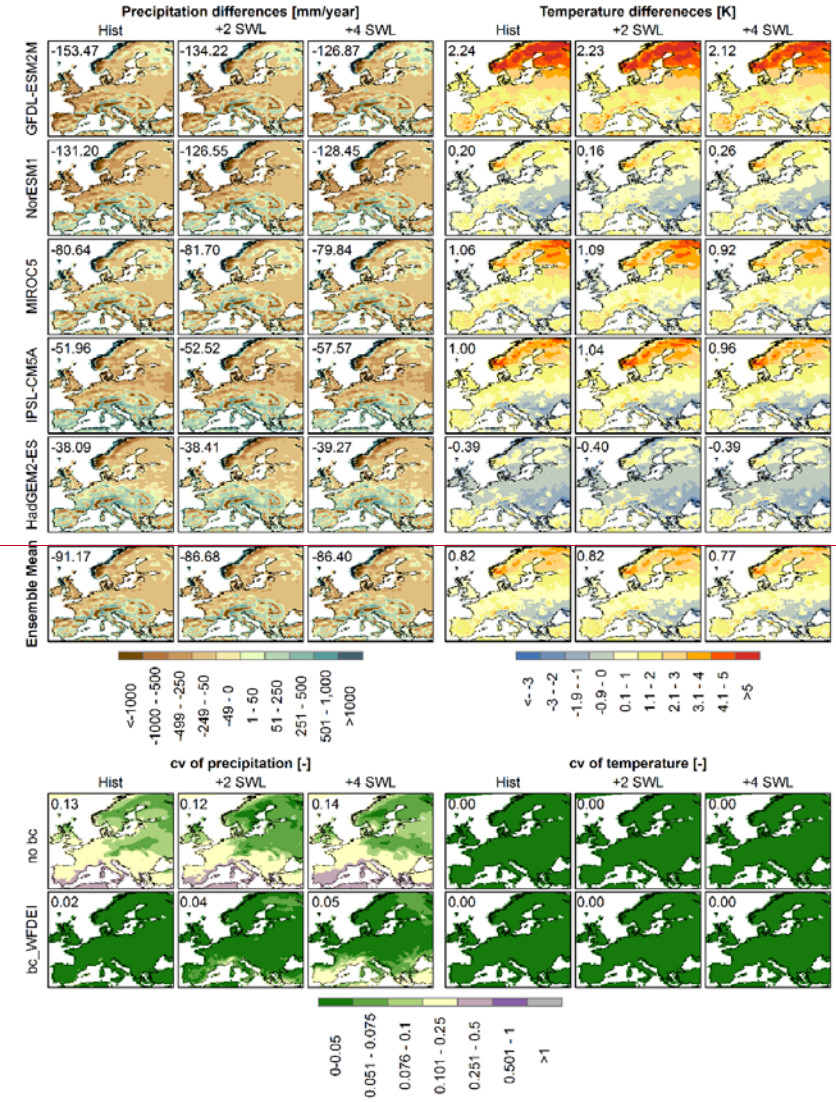
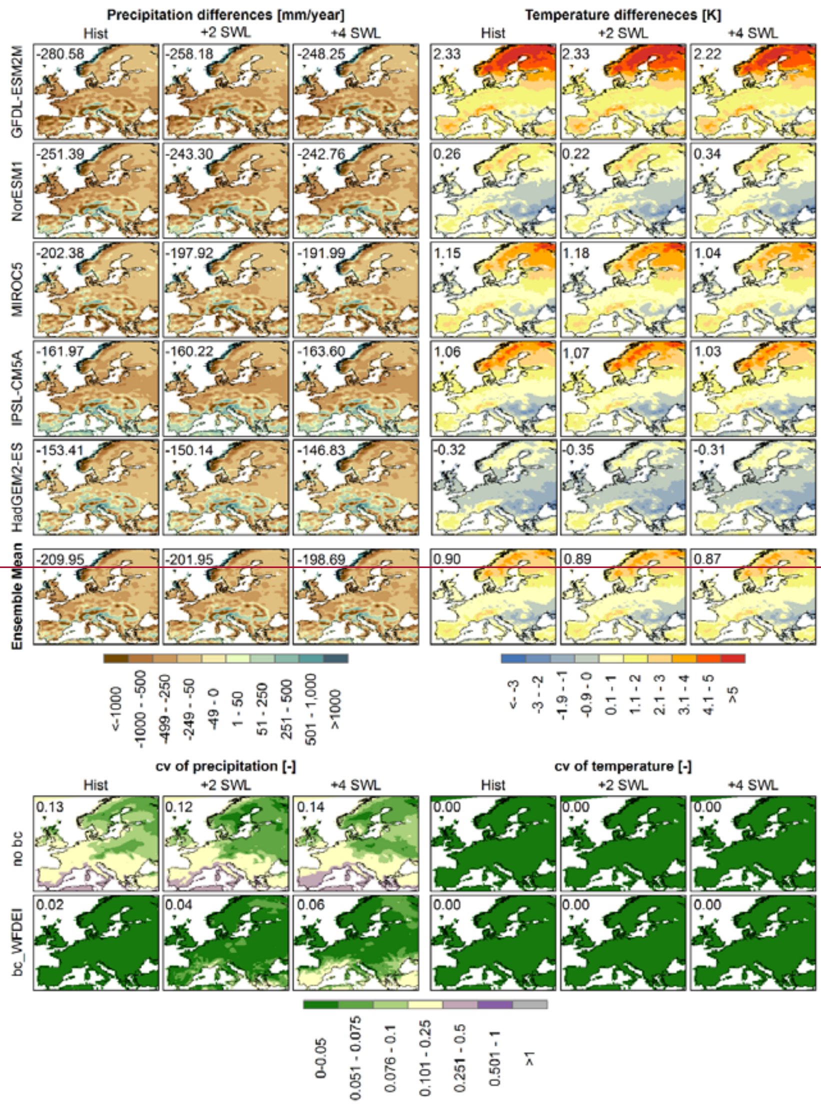
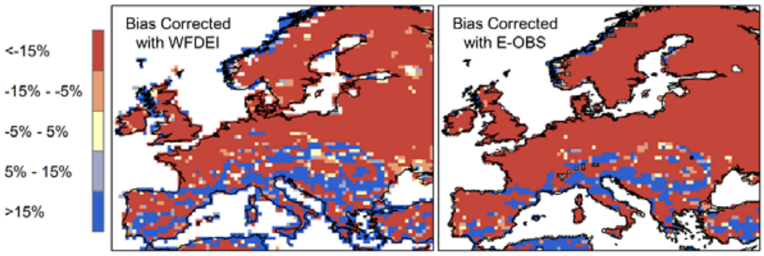


Figure S 1. Absolute differences between Euro-CORDEX data bias adjusted against the WFDEI dataset and raw Euro-CORDEX data, for the variables of precipitation (right block) and temperature (left block). Differences are calculated from the historical (1976-2005), +2 SWL and +4 SWL time-slice averages, for all dynamical downscaled GCMs and their ensemble mean. Bottom block: Coefficient of variation between the ensemble members, for raw and bias corrected against the WFDEI dataset precipitation and temperature forcing variables, for the historical, +2 SWL and +4 SWL time slices. The average value for the pan-European area is shown in each sub-figure.



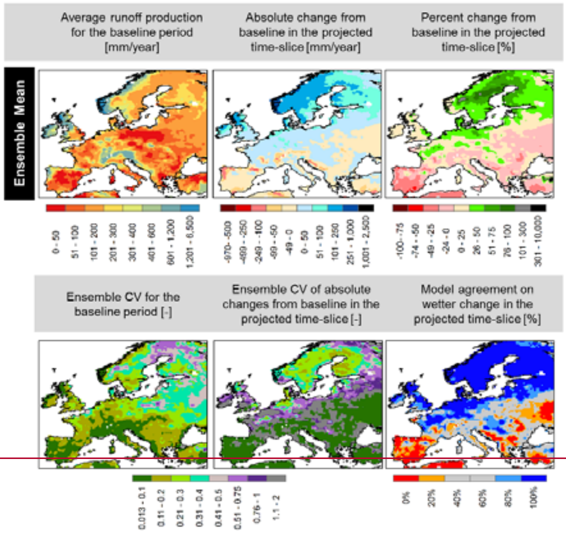
~~Figure S 2. Absolute differences between Euro CORDEX data bias adjusted against the E-OBS dataset and raw Euro CORDEX data, for the variables of precipitation (right block) and temperature (left block). Differences are calculated from the historical (1976-2005), +2 SWL and +4 SWL time slice averages, for all dynamical downscaled GCMs and their ensemble~~

mean. Bottom block: Coefficient of variation between the ensemble members, for raw and bias corrected against the E-OBS dataset precipitation and temperature forcing variables, for the historical, +2 SWL and +4 SWL time slices. The average value for the pan-European area is shown in each sub-figure.



	Bias Corrected with WFDEI			Bias Corrected with E-OBS		
	Drier output	Wetter output	Insignificant change	Drier output	Wetter output	Insignificant change
Percent of pan-European land area	70.40%	26.01%	3.59%	83.62%	14.67%	1.70%
Average percent change	-44.15%	148.77%	-0.53%	-56.10%	215.33%	-0.87%
Average absolute change [mm/year]	-231.44	159.37	-2.96	-285.70	131.97	1.49

Figure S 3. The effect of bias correction on the ensemble mean of average runoff production for the baseline period (1976-2005). Figures: Relative difference between the ensemble means of bias corrected (left:with WFDEI, right:with E-OBS) and raw forcing data. Differences between -5% and 5% are classified as insignificant, differences < 5% as drier output and differences >5% as wetter output after bias correction. Table: percent of land area that falls into each category of change and average of the changes.



~~Figure S 4. Ensemble mean of average runoff production from Euro-CORDEX data bias adjusted against the E-OBS dataset. Top row: Runoff production averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice. Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice~~

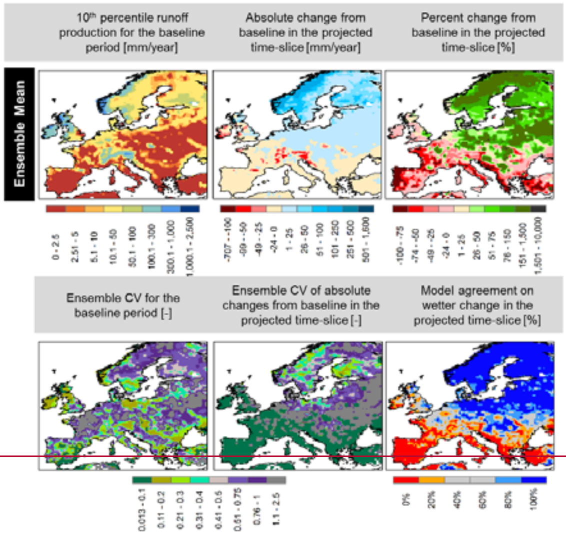


Figure S 5 Ensemble mean of 10th percentile runoff production from Euro-CORDEX data bias adjusted against the WFDEI dataset. Top row: 10th percentile runoff production derived on an annual basis averaged over the baseline period (1976-2005) (top row), absolute (middle row) and percent change (bottom row) in ensemble mean runoff in the +4 SWL projected time-slice. Bottom row: coefficient of variation of the ensemble members for the baseline period (left column), coefficient of variation of the projected absolute changes in the +4 SWL projected time-slice (middle column) and model agreement towards a wetter change in the +4 SWL projected time-slice.

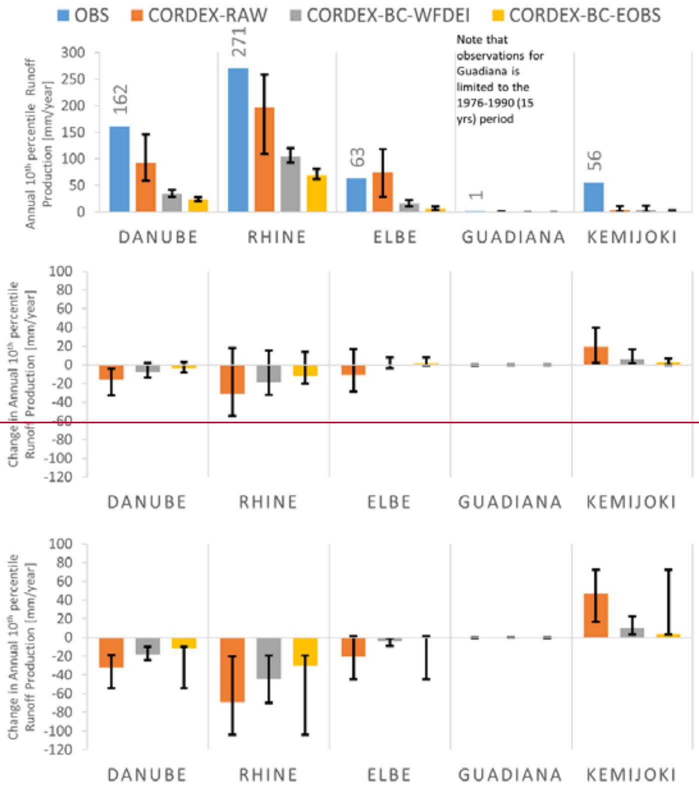


Figure S 6. Comparison between the simulations of raw Euro-CORDEX data and bias adjusted against two different datasets (WFDEI and E-OBS) for five study basins. Bars show the ensemble means and error bars the minimum and maximum ensemble member values. (Top row) Annual 10th percentile runoff production for the period 1976 to 2005.OBS values are derived from GRDC discharge measurements converted to basin averages at the annual time-scale. (Middle row) Percent change in annual 10th percentile runoff production at the +2 SWL and (bottom row) at the +4 SWL.

Formatted: Superscript

Formatted: Superscript