Responses to Referee #2

We thank the constructive insights offered by the anonymous referee, which significatively contributes to the improvement of the manuscript. We respond by highlighting the reviewer's comment in shaded text and our responses in black.

Main Comments:

The manuscript titled "River flow in the near future: a global perspective in the context of a high-emission climate change scenario" investigates the potential effects of global warming on river flows worldwide from 2015-2050 with hydrological simulations. The study aims to provide insights into the potential changes in river flows and their broader socio-environmental consequences. Overall, the manuscript is well structured and clear. However, the following points raise my concerns.

1) There seems to be a lack of explicit validation of simulated river flows by observations from the same historical period. In general, it is important to ensure that the modeling framework accurately reproduces observed river flows in the historical period before trusting projections for the future.

Thanks for making this important point. In response, we will incorporate a comprehensive validation of simulated river flows in the revised manuscript. This validation will involve assessing four distinct metrics: relative bias (RB), overlapping coefficient (OC), correlation coefficient (r), and non-parametric Kling-Gupta efficiency (npKGE) for the 18 hydrological simulations and the ensemble mean. These metrics, which evaluate diverse aspects of the simulations, are computed by comparing our simulations with monthly observations of a selected set of 346 monitored rivers, covering approximately 42% of the global land and contributing to about 45% of the global river discharge.

The ensemble mean presents a relative bias of 7.6%, an overlap coefficient of 0.63, a correlation coefficient of 0.72, and an efficiency of 0.58. These scores improve to -1.8%, 0.62, 0.76, and 0.71 (respectively) when the assessment is restricted to the 20 largest monitored rivers (see new Figure 1). Another important finding of the validation is the consistency of the scores across models. Independent of the resolution and the type of simulation, the scores remain in a narrow range. The validation results compellingly prove the robust performance of the model in the context of this study.

Further details about the validation methodology will be included in the section “2.2 Assessment methodology”, while the scores will be analyzed in the new section “3.1 Validation of the hydrological simulations”.

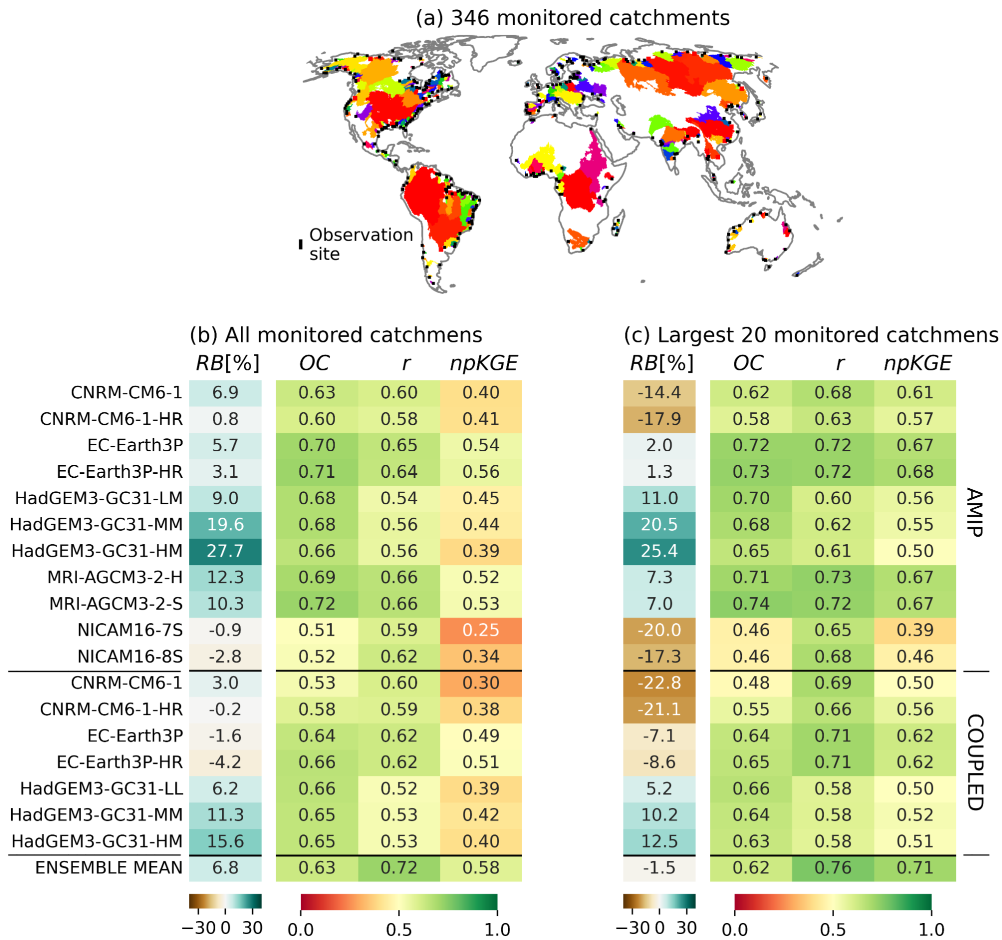

Figure 1. Validation of the global hydrological simulations. (a) Monitored rivers with black dots indicating the observation sites for river flow and colors highlighting the catchment area that contributes to each monitored point. (b) Average relative bias (RB), overlapping coefficient (OC), correlation coefficient (r), and non-parametric Kling-Gupta Efficiency (npKGE) for each simulation and the ensemble mean. (c) As (b) but for the 20 largest catchments. The averaged OC, r, and npKGE are calculated using a weighted average, where the weight assigned to each river is proportional to its contribution to the total observed flow under evaluation.

2. section 3.2.1 is overly descriptive, providing detailed information about individual rivers, their importance to their respective regions, historical context, and model projections. While such detail is valuable to some extent, it may be more than necessary for the main message of the article. I recommend a more concise way to convey the broader implications of this study. For example, group rivers with similar trends and mention the specifics only when they significantly deviate from the general trend. The implications for human settlements, global water systems, and climate systems can be summarized in one final paragraph. In addition, explaining the "why" behind the trend or time of emergence can provide more insightful analysis and make the findings more compelling. Also, the discussion section reads repetitively, which seems to summarize the above finding again.

We value the reviewer feedback regarding the detailed nature of section 3.2.1. While recognizing the value of such details, we acknowledge the need for conciseness in conveying the primary message of the article. In response, we will adopt the suggested approach of grouping rivers with similar trends. We identify four regions with notable trends and clear consensus among model simulation: Central Africa, Arctic, South Asia, and Patagonia (see new Figure 8). This approach offers a concise way to communicate the wider implications of our study. This regional assessment will be complemented with a brief description of specific rivers at their mouths projecting significant changes in their annual cycle, which could be useful for planning purposes given the population living upstream (see new Figure 9).

In response to the reviewer's concern about repetitive elements in the discussion section, we are committed to enhancing clarity and conciseness in our revised manuscript. The updated section will delve into various topics, including choices related to the signal-to-noise ratio technique and their potential impact on results (discussed in our response to reviewer #1 major comment #4), the internal variability of model projections, and the potential sectoral implications of projected changes at the regional and at the local scale. We believe these refinements will effectively address the raised concerns and elevate the overall clarity of our discussion.

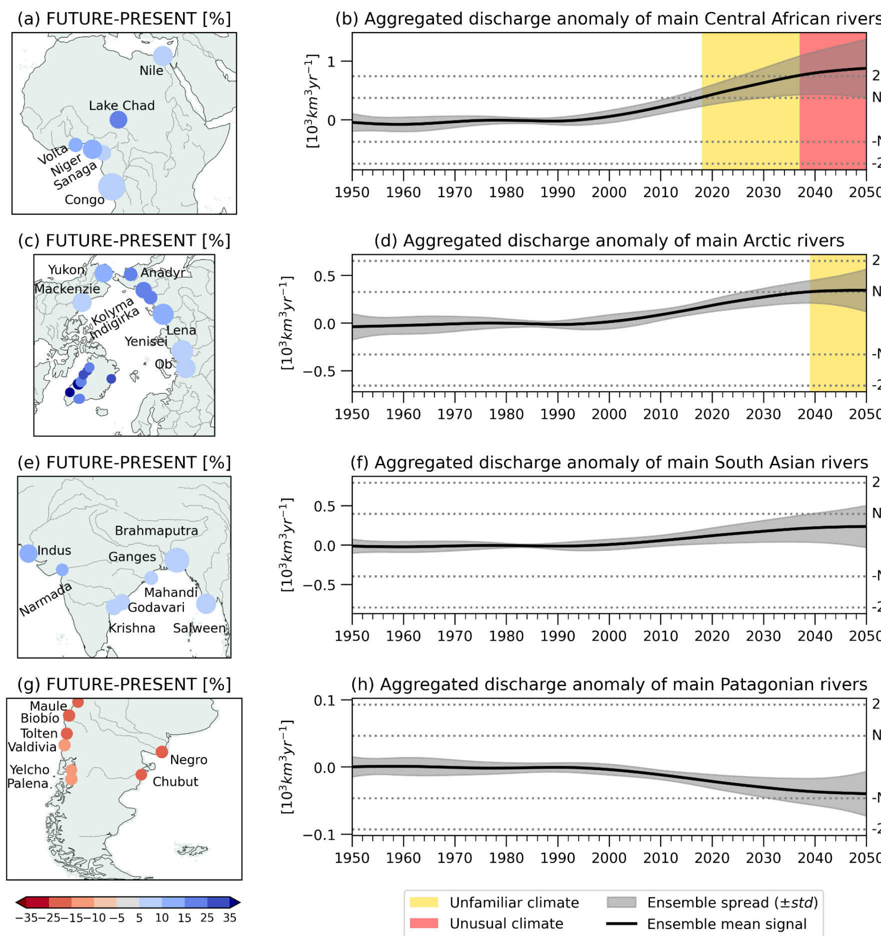

Figure 8. Percentage difference in mean flow at the river mouth between FUTURE and PRESENT for rivers presenting consensus among models regarding their trends in (a) Central Africa, (c) Arctic, (e) South Asia, and (g) Patagonia. The right panels depict the aggregated discharge anomaly signal for the ensemble mean (and the spread across models) for the rivers shown in left panels.

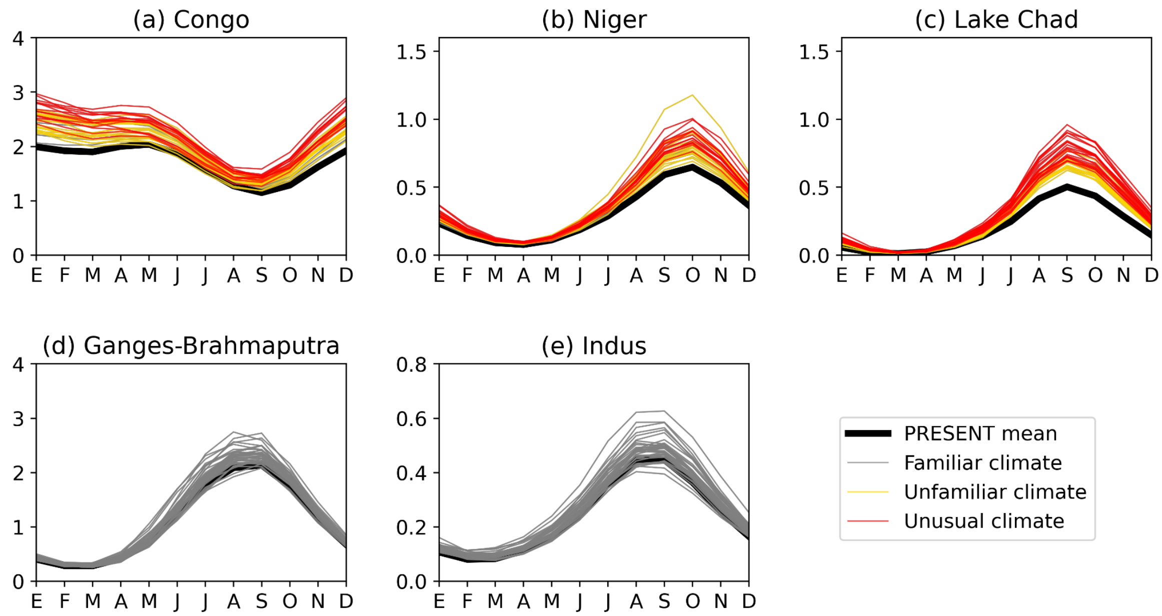

Figure 9. Annual cycles of (a) Congo, (b) Niger, (c) Lake Chad tributaries, (d) Ganges-Brahmaputra, (e) Indus at their mouths. Units are in $[10^3 km^3 yr^{-1}]$.

3. For the information in Appendix A, it's a bit of a leap to conclude that "it may be assumed that our set of simulations is adequate for the proposed objectives and that more realizations would not present substantial alter the presented results" without information about the internal variability of other models. While this may be true for the HadGEM3-GC31 model, it may be premature to state this as a broader conclusion without data from other models to support it.

We agree with the reviewer, the internal variability is only shown for HadGEM-GC31 and that can not be extended to all models. We got access to more realizations of the HighResMIP-CMIP6 from the Laurence Livermore National Laboratory (LLNL) Earth System Grid Federation (ESGF) Node (esgf-node.llnl.gov), which includes total runoff among the variables. Thus, the internal variability analysis was extended to these new GCM simulations, resulting in an assessment of 58 individual realizations in total.

The new results for the CNRM-CM6-1 and EC-Earth3P GCMs family confirm our previous findings for the HadGEM3-GC31 family. While the internal variability tends to rise over time, it is smaller than the inter-model variability and comparable to the variability given by the GCMs' resolution (please see the results in our response to reviewer #1 major comment #2). These new results provide robustness to the assumption that our set of simulations is adequate for the proposed objectives and that more realizations would not substantially alter the presented results.

The new figure showing the internal variability of CNRM-CM6-1, EC-Earth3P, and HadGEM-GC31 GCMs' families along with their interpretation will be incorporated to the Appendix A: “Internal variability in projections of runoff anomaly”.

Minor comments:

1) Figure 4, can you explain why most of the solid lines show a downward trend after 2045?

The "boundary" effects observed in the first and last decades of the several GCMs' signals are caused by the filtering process. The choice of the low-pass filter is a critical step in our analysis. We tested several filters and selected it considering: (a) an effective attenuation of high-frequency noise without introducing phase distortion, and (b) the minimization of the boundary effects. Figure R1.1a shows how the various filter's respond to the river discharge anomaly time-series of a given GCM. From the various options, we selected the Lowess filter in our approach given that it does not introduce phase distortion, unlike Butterworth, Chebysev, or Elliptic (see their phase shifts around 2020), and minimizes boundary effects, unlike the FFT low-pass filter (see the strong blending effect in the first and last decades). Thus, the selected filter ensures realistic climate change signals and noise terms.

A detailed explanation of the choices related to the filtering process and the overall signal-to-noise technique is provided in our response to reviewer #1 major comment #4. The revised manuscript will summarize the main aspects of these topics.

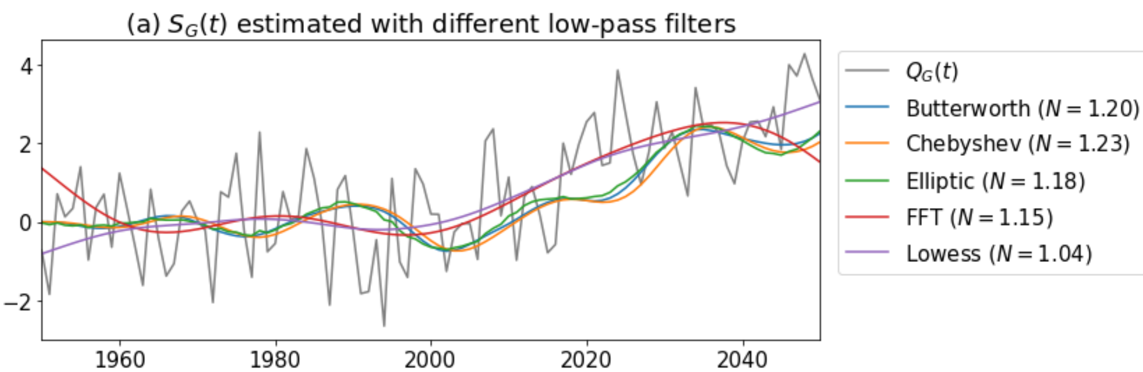

Figure R1.1a. Low-pass filter selection test for a river discharge anomaly time-series (Q_G) using the CNRM-CM6-1-HR model as a case study. The signal S_G is estimated with various filters.

2) Figure 5, I'm curious if the order of calculating the ensemble mean and calculating the signal-to-noise ratio would have an impact, let's say calculating the signal-to-noise ratio of the ensemble mean of the simulation instead.

The order of calculation has an impact on the results. Both alternatives present similar patterns (see Figures 5 original and suggested). However, calculating the signal-to-noise ratio (SNR) of the ensemble mean tends to magnify the SNR values. This effect arises from the ensemble mean time-series, which inherently exhibits minimal interannual variability due to the cancellation of individual model variabilities. The reduced variability results in a small noise value, thereby yielding high SNR values. For this reason, we consider our original calculation to be more realistic.

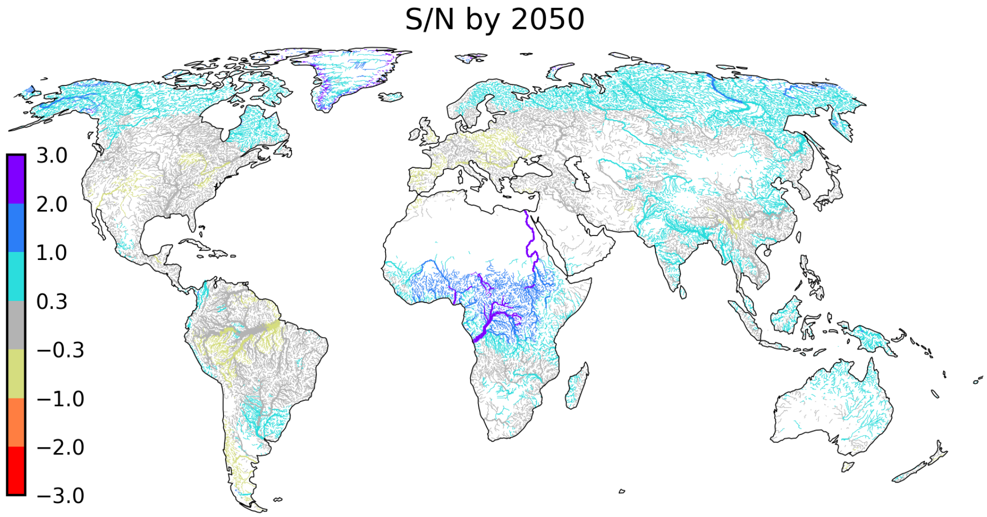

Figure 5 (original). Global map of signal-to-noise ratio of river flow by 2050. <u>The ratio is the ensemble mean of the signal-to-noise ratio of each simulation</u>. Rivers with little climatological flow are masked out.

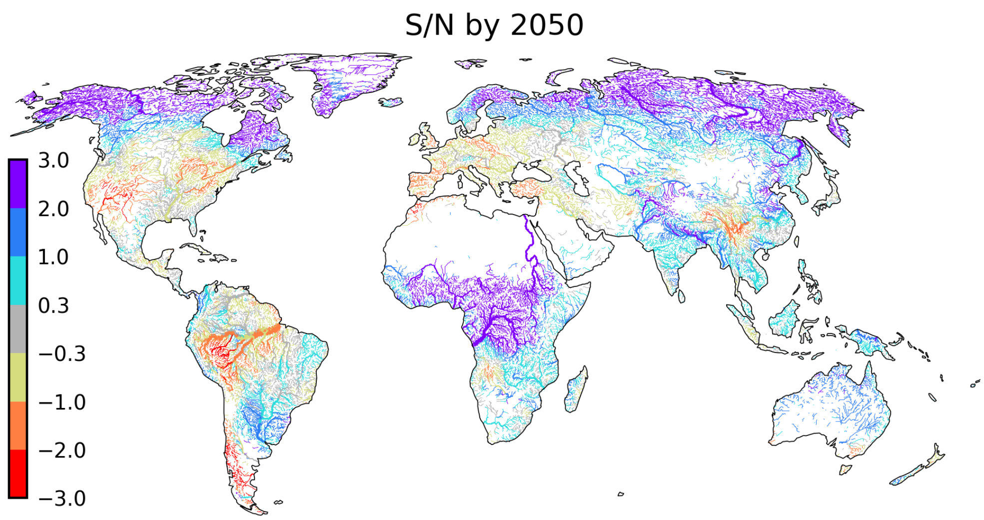

Figure 5 (suggested). Global map of signal-to-noise ratio of river flow by 2050. <u>The ratio is the signal-to-noise ratio of the ensemble mean</u>. Rivers with little climatological flow are masked out.

Editorial comment:

1) Figure 1, there seem to be no squares in the three main plots.

Thanks for noting it. We originally plotted the COUPLED GCMs with squares, but then decided to show them with closed circles, to allow a fair visual comparison of the warming level with the AMIP GCMs. We will fix the legend in the revised manuscript.