Dear reviewers,

Here we provide a list of changes and a short point-by-point response to the reviews. Thank you again for the valuable feedback. We believe that this greatly benefitted the quality of the publication.

Kind Regards,

Jerom Aerts

List of changes:

1. Title:
    Better describes content of the study and uses the term varying spatial resolution instead of scaling.
2. Abstract:
    The abstract is updated based on the extra analyses we performed.
3. Introduction:
    - Restructured.
    - Includes competing basin discretization approaches.
    - Includes a larger body of cited literature that refers to the land surface community, parameter identifiability problem, parameter transferability, and the representative elementary watershed.
    - Clearly stated that we are not using the MPR method.
    - We now use "the effects of varying spatial resolution" instead of scaling as this is only a part of scaling.
4. Methodology:
    - Better description as to why basins are excluded from the analyses.
    - Better description of the calibration methodology.
    - Description of the sampling uncertainty of the KGE score method of Clark et al. 2021.
5. Results:
    - Restructured.
    - Included calibration period CDF of the KGE score and the decomposed components.
    - Added the evaluation period KGE score components to the CDF figure.

- Added objective function uncertainty section. Includes the analyses similar to Clark et al. 2021. Added a table summarizing these results based on the evaluation CDF.
6. Discussion:
    - Restructured. Added headers for readability.
    - Refer to other large domain studies.
    - Discuss the sampling uncertainty results and what this means for the statistical KS-test results.
    - Discuss the effect of using coarse meteorological forcing products.
    - Discuss how vector-based discretization has major benefits when it comes to computational cost and topographic discretization.
7. Outlook:
    - Expended on what needs to be added to this study in order to do a complete scaling assessment as opposed to only a spatial scaling assessment.
8. Conclusions:
    - State that the sampling uncertainty is large and therefore the conclusion based on the KS-test are inconclusive.
    - Added "at the outlet" to the conclusion regarding finer spatial resolution does not always leads to better streamflow estimates.

Point-by-point response to the reviews:

Review #1

- We included the calibration period CDFs of the KGE score and its individual components.
- We included the individual components of the KGE score with the evaluation period CDFs.
- We included the CDFs of the benchmark.
- We added structure to the discussion.
- We included important references for the CONUS domain (Mizukami et al. 2017; Rakovec et al. 2019) in the discussion.
- We resolved all the suggested minor changes.

Review #2

References

- We have extended the cited literature in the introduction by including the land surface community, parameter identifiability, parameter transferability, and the representative elementary watershed.

Lakes and reservoirs

- We reported the amount of lakes and reservoirs used by the model instances. Due to the small amount this did not alter the main conclusions of the study.

Basin selection

- Clearer description as to why basins were excluded from the analyses.

Calibration methodology

- Clearer description of the calibration methodology.

KGE sampling uncertainty

- We performed the suggested methodology of Clark et al. 2021 and discuss the original results in light of the sampling uncertainty.

Discretization:

- We discuss the benefits of vector-based discretization in the discussion.

Scaling:

- We refrain from using the term scaling and use varying spatial resolution instead. Added to the discussion and outlook what is required to analyse spatial scaling.