Dear reviewer,

Thank you for considering this publication after major revision and providing additional feedback. Here, we will provide a point-by-point reply to your latest comments.

Point-by-point reply:

General comments:
I am glad to see a revised version whose scope is reduced and all my comments are addressed. The paper is well organized but I found many typos. Addition of the conceptual models and their performance CDFs is very informative and more practical as not everyone uses physically-based models. The impressive performance of GR4J, given the few number of parameters, is something I did not expect. I only have minor comments now that are stated below, the authors can incorporate them quickly.

Thank you for your feedback. We are pleased to hear that the revised manuscript, with its reduced scope and incorporation of your comments, is now well-organized and informative. We have addressed the typos and improved the grammar throughout the manuscript based on your valuable feedback without altering the core content of the originally revised version.

Specific comments:
L239-240: Are the discharges coming from calibrated models? It is not clear. I am assuming they are.

Yes, the discharges are indeed derived from calibrated models. We have clarified this in the revised manuscript to remove any ambiguity regarding their origin.

L290-296: Here, the conclusion drawn about differences being more evident in north and south as compared the center are subjective (weak). It would be simpler to just divide them via arbitrary (exact angles need to be determined) lines that divide the area into three portions and then comparing either the means or medians of KGE-NP differences. If this point is the focus of the study, then please be thorough. I don’t find it relevant.

We agree that the initial statements were subjective and not well-supported. We have removed the statements regarding visual trends in the results, as they were not essential to the manuscript’s focus. Instead, we retained a brief analysis to complement the large-sample catchment assessments, acknowledging that while the results may not be conclusive, they provide additional context.

L298-305: Yes, but showing relative differences is not a good idea. River width is not shaped by low flows. During low flows, a slight fluctuation in depth can easily translate to above 50% flow increase/decrease. Sedimentation may result in an increase of a few centimeters, depending on the situation. The flow volume that passes through a cross-section is important. Here, another paragraph can be added that translates the relative differences to absolute flow volumes. I think, that would showa different story.

Thank you for raising this valid point. We address this issue in the limitations of the manuscript. The goal of the original plot showing uncertainty percentages is to effectively *communicate that these percentages can be substantial, highlighting the relative uncertainty across different flow categories (low, average, high). We agree that the flow categories are not equally sensitive to the physical phenomena that are the cause of discharge observation uncertainty. Using relative values (percentages) allows us to illustrate how uncertainty varies across different flow regimes, which is crucial given the wide range of flow magnitudes (e.g., from 1 m^3/s to 100 m^3/s) among the catchments. Absolute values result in boxplots with an extensive range, skewing the interpretation due to the large variability in flow volumes.*

Using the three flow categories, we show how uncertainty affects low flows differently than high flows. This approach helps model users understand and interpret the results in the context most relevant to their applications. Given the variation in flow conditions, it is up to the model user to apply this information according to their needs. Therefore, we believe that relative values provide a clearer comparison and better convey the practical implications of uncertainty in our study.

L437-455: The abstract and conclusions deal with the problem stated in the title now. The only problem that remains is getting uncertainty bounds on observed discharges, elsewhere in the world but the results do shed light on what the situation might look like.

Thank you for your feedback. We are pleased that the abstract and conclusions now align with the problem stated in the title. We acknowledge the challenge of obtaining uncertainty bounds on observed discharges globally, which remains a significant issue. This broader challenge should be addressed by the scientific community as we expect discharge observation uncertainty to be substantial in regions where monitoring equipment might be of lesser quality and less stringent maintenance protocols.

While our study provides valuable insights into potential uncertainty scenarios for the specific use case in Great Britain, we recognize that extending these findings to a global context presents additional complexities. The results shed light on what the situation might look like, but further research is needed to address uncertainty bounds on observed discharges worldwide.

We appreciate your understanding and will highlight this in the outlook of the revised manuscript to ensure clarity regarding the scope and applicability of our findings.

Main Changes:
Manuscript: Grammar and spelling of the manuscript. A track-changes document is attached that provides all grammar and spelling corrections.

L239-240: Clarification of the type of model run, calibrated.

L286-293: Removal of text stating subjective spatial trends in the result.

L429-434: Reflection on the limitations of using relative values for expressing discharge observation uncertainty rather than absolute values.

L464-468: Addition of an outlook on the larger challenge of global discharge observation uncertainty estimates and its effect on water resources management.