Reviewer: The manuscript revision is an improvement from the initial manuscript. The revision remains clear and the dataset is well-distributed (i.e. publicly available and accessible). However, the lack of recognition and comparison against other works, as mentioned in the previous reviews (reviewers #2 and #3) remains. Also, the writing in the manuscript should still be improved. Therefore, I would again recommend major revisions for this manuscript. Below is a more expansive description of my main arguments as well as a list of line-by-line comments.

Additional comparison and recognition

Based on the previous round of reviews, the authors have mentioned other similar works in the introduction/discussion and included a thorough harmonized comparison with results from the mHM model.

Nevertheless, the context of the efforts toward hyperresolution modeling remains very limited. Especially in light of the previous round of reviews, the authors use overly strong language when presenting their dataset, aiming to indicate this dataset is a first of its kind. However, several other studies have done similar work. Although this dataset contains novel aspects, there are also downsides (i.e. limited spatial coverage and resolution). Therefore, this dataset should be placed within the context of and compared to other work in the area of hyperresolution modeling.

Note that comparison does not only entail a thorough harmonized comparison as done with the mHM model (especially as the authors could not acquire the outputs from some other studies). Rather, some context on how other studies handle their hydrological reanalyses (i.e. spatial coverage, spatial resolution, temporal coverage, socioeconomic conditions; see Line 647-650), and whether their reported performance is in line with HERA. Note that even though other studies (e.g. GLOFAS-ERA5) use different inputs, their outputs can still be compared.

Reply from authors: We appreciate the reviewer's feedback on the revised manuscript, most of which will improve its quality. We have indeed included references to other high-resolution hydrological reanalysis studies and compared our dataset, HERA, with a run of the mHM model. However, we would like to clarify the statement about HERA being a first-of-its-kind dataset. It refers to the fact that we could not find a publicly available hydrological dataset with characteristics that make a comparison meaningful. This means with comparable (i) resolutions (GLOFAS-ERA5 is 0.1°, meaning that 1 pixel is 30 times larger than HERA), (ii) spatial coverage (Europe), (iii) temporal coverage (HERA starts is 1951, 28 years earlier than GLOFAS-ERA5). We would appreciate any information about publicly available datasets that have characteristics comparable to HERA.

Regarding GLOFAS-ERA5 more in particular, we acknowledge it in the manuscript, and it indeed has a lot in common with HERA (ERA5 as a primary meteorological forcing, LISFLOOD as hydrological model, KGE' used as a skill metric in calibration). We however believe that the two reanalysis, and their validation work, are too different to be compared. GLOFAS-ERA5 is a global reanalysis, mainly developed for large rivers. Most catchments used in its validation have an area above 10 000 km^2 while the large majority of catchments used in the validation of HERA have an upstream area below 10 000 km^2. Performances of GLOFAS-ERA5 are very well described in Harringan et al. (2020), and that article was actually an inspiration for working on a much more detailed reanalysis product with focus on Europe. Although the present article does not exactly reproduces the same figures, same metrics are used, making a quick comparison simple. As stated before, we are doubtful that this comparison makes sense since the catchment sets used to validate both datasets are too different. We believe that the mHM run we used for comparison in the supplement is more relevant in terms of (i) resolutions (ii) coverage, (iii) length and (iv) purpose of the dataset.

We addressed most of the specific comments raised by the reviewer, resulting in the following main changes:

- Addition of a section in the supplement with a table providing a summary of reported performances of HERA, GLOFAS-ERA5, GRFR and EUmHM.
- Modification of the discussion to incorporate reviewer's suggestions
- Addition of new literature.
- Correction of typos and clarification of vocabulary (all manuscript)

Line number in authors' reply refer to the update Track Changes document.

Specific comments (note that line numbers refer to the markup document)

Line 15 '(…) anthropogenic activities have altered (...), soil properties, channel morphologies (…)': these are not changed in this study, which may confuse.

R: We understand the reviewer's concern that the phrase "anthropogenic activities have altered (...) soil properties, channel morphologies" may be misleading since these aspects are not specifically addressed in this study. However, we believe that the abstract's introductory sentence aims to provide a general context for the research, rather than a detailed description of the study's focus. The phrase serves to highlight the importance of hydrological reanalyses in understanding the impacts of human activities on the environment, which is relevant to the broader field of study.

Line 17-19 'The availability of consistent estimates of river flow at global and continental level is a necessity to assess and attribute changes in the hydrological cycle.': one cannot attribute changes based on consistent estimates of river flow, rather several simulations with different forcings/parameters are needed to attribute changes. Although this is possible with the model setup in this study, this study presents a dataset. See also Line 38-40.

R: We removed "and attribute" from the sentence.

Line 37-38 'HERA is the first publicly available long-term, high-resolution hydrological reanalysis for Europe': since there are other highER resolution hydrological reanalysis available for Europe, remove sentence.

R: We appreciate the reviewer's suggestion to remove the sentence stating that HERA is the first publicly available long-term, high-resolution hydrological reanalysis for Europe. However, our statement is based on our thorough search for publicly available datasets with similar characteristics to HERA. Despite our efforts, we were unable to find other publicly available long-term, high-resolution hydrological reanalyses for Europe.

We are aware of the EFAS v5 long run, which shares similarities with HERA, but its shorter temporal coverage (30 years) led us to consider it as a different type of dataset. We would be happy to revise the sentence if the reviewer can provide us with examples of publicly available datasets that meet the same criteria as HERA (long-term, high-resolution, and publicly available).

Line 38-40 'Despite its limitations, it enables (…)': double 'it', rephrase to 'Despite its limitations, HERA enables (…).

R: Sentence changed.

Line 45-47 'These evolving conditions have significantly changed flows in European streams and rivers (…)': requires citation.

R: The following references have been added:

Barker, L. J., Hannaford, J., Parry, S., Smith, K. A., Tanguy, M., and Prudhomme, C.: Historic hydrological droughts 1891–2015: systematic characterisation for a diverse set of catchments across the UK, Hydrology and Earth System Sciences, 23, 4583–4602, https://doi.org/10.5194/hess-23-4583-2019, 2019.

Gudmundsson, L., Boulange, J., Do, H. X., Gosling, S. N., Grillakis, M. G., Koutroulis, A. G., Leonard, M., Liu, J., Müller Schmied, H., Papadimitriou, L., Pokhrel, Y., Seneviratne, S. I., Satoh, Y., Thiery, W., Westra, S., Zhang, X., and Zhao, F.: Globally observed trends in mean and extreme river flow attributed to climate change, Science, 371, 1159–1162, https://doi.org/10.1126/science.aba3996, 2021.

Vicente-Serrano, S. M., Peña-Gallardo, M., Hannaford, J., Murphy, C., Lorenzo-Lacruz, J., Dominguez-Castro, F., López-Moreno, J. I., Beguería, S., Noguera, I., Harrigan, S., and Vidal, J.-P.: Climate, Irrigation, and Land Cover Change Explain Streamflow Trends in Countries Bordering the Northeast Atlantic, Geophysical Research Letters, 46, 10821–10833, https://doi.org/10.1029/2019GL084084, 2019.

Wang, H., Liu, J., Klaar, M., Chen, A., Gudmundsson, L., and Holden, J.: Anthropogenic climate change has influenced global river flow seasonality, Science, 383, 1009–1014, https://doi.org/10.1126/science.adi9501, 2024.

Line 45-47 'These evolving conditions (…), leading to challenges for hydrological sciences, related, for example, to long term variability, climate change, extremes or human alterations of the water cycle': How do changes in the water cycle challenge hydrological sciences?

R: For example:

- Flood protections design: if flood magnitudes are on the rise, societies should reassess the design of flood protection infrastructures and dams spillways. This pushes hydrological scientists into adapting there methods, by using non-stationary statistics for example.
- Water availability: if rivers tend to dry, as it is the case in southern Europe, less water will be available for human usages, and could therefore create further conflicts around water use (socio-hydrology).
- Understanding the complex response of diverse catchments to climate change is also a challenge for hydrological sciences.

Line 49-51 'Observations (…) are lacking at a high enough spatial density (…)': lacking for what?

R: We have update this sentence, now indicating that this hampers analysing pan-European long term trends: "*Observations, despite continuous improvements (Blöschl et al., 2019a; Ekolu et al., 2022), can hamper the analysis of Pan-European long-term trends due to sparse spatial distribution in some regions and temporal discontinuities*".

Line 66-68 'Remote sensing technologies now provide high resolution input for hydrological models (…)': Inputs at the spatial resolution of this study has been provided already for quite some time.

R: We agree with the reviewer. Technologies are rapidly evolving and provide increasingly more accurate and higher resolution information. We updated the sentence accordingly.

Line 76-78 '(…) numerous homogeneous environmental variables (…)': what does this mean?

R: The formulation was not optimal. We have now rephrased this: "Reanalysis products typically provide a large number of variables (e.g., precipitation, wind speed, temperature) that are physically consistent with homogeneous spatiotemporal resolution."

Line 89 'Section 2.2': for me, all sections are unnumbered, so these references (throughout the document) do not help.

R: Yes, we apologize for that. The uploaded revised manuscript has the sections correctly numbered.

Line 104 'Therefore, tis': 'Therefore, this'

R: This has been corrected.

Line 110: 'improvements in processing speed, spatial and temporal resolutions, calibration': 'improvements in processing speed, spatial and temporal resolutions and calibration'

R: This has been corrected.

Line 115-113 'These developments make this dataset the first publicly available long-term Pan-European hydrological reanalysis taking into account the evolving socioeconomic conditions that have altered the hydrological cycle since 1951': since there are other hydrological reanalysis available that take into account the evolving socioeconomic conditions (with explicit results for, for example, human water withdrawals) for Europe, remove sentence.

R: We would appreciate clarification and supporting evidence regarding the existence of publicly available Pan-European hydrological reanalyses that account for evolving socioeconomic conditions, as this would inform our assessment of the novelty and uniqueness of the dataset presented in this work.

Line 170-171 '(…) a 71-year pre-run (longest possible period)': would the spinup not be able to loop the 71 years to get a longer period?

R: The reviewer has a point, in principle, yes. We therefore removed "longest possible period". We believe a 71-year pre-run is sufficiently long.

Line 178 'River initialization in OS LISFLOOD can lead to unrealistic discharge in some catchments': why is this the case after a 71-year spinup?

R: The reason causing the initial months to be unreliable is the fact that water volumes at time 0 in the channels are not known and the model sets a conventional initial volume (LISFLOOD uses half-bankful). This state is expected to "quickly" adjust (it is much shorter memory compared to soil and groundwater). The amount of time required to remove the impact of the initial, fictitious value depends on mainly on river geometry and climate. After additional verifications, we decided to remove the full year of 1950. We added the following sentence to the manuscript (line 169): "*As water volumes at the first time step in the channels are not known, the model sets a conventional initial volume (OS LISFLOOD uses half-bankful), leading to unrealistic initial discharge in some catchments*."

Line 519 ‘Factors that can explain the poor performances (..) include the combination of arid climates and the strong influence of lakes and reservoirs’: this does not explain performance. Does this mean that the model is worse at simulating arid climates and lakes/reservoirs?

R: Yes, that is exactly what it means.

Line 723-725 ‘With its refined spatial and temporal resolution, HERA represents hydrological processes in Europe with more detail than previous publicly available hydrological reanalysis products’: since there are other hydrological reanalysis available that represent hydrological processes in Europe with an even more refined spatial resolution, remove sentence.

R: We understand the reviewer's suggestion. However, our statement is based on our understanding of the publicly available datasets at the time of writing. If there are indeed other publicly available hydrological reanalysis products for Europe with even more refined spatial resolutions, we would appreciate it if the reviewer could provide us with specific examples or references to these datasets. This would enable us to accurately assess the resolution of HERA in relation to other available products and revise the sentence accordingly.

Line 728-730 ‘Parameters in 93.5% of the HERA (…)’: ‘Parameters in 93.5% of the HERA domain (…)’

R: This has been corrected.

Line 730-732 ‘This is a very high calibration coverage for a GHM, which are not systematically calibrated (…)’: many GHMs are systematically calibrated.

R: Models are not systematically calibrated in the following studies:

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Hoch, J. M., Sutanudjaja, E. H., Wanders, N., van Beek, R. L. P. H., and Bierkens, M. F. P.: Hyper-resolution PCR-GLOBWB: opportunities and challenges from refining model spatial resolution to 1 km over the European continent, Hydrology and Earth System Sciences, 27, 1383–1401, https://doi.org/10.5194/hess-27-1383-2023, 2023.

Schellekens, J., Dutra, E., Weiland, F. S., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., Weedon, G. P., and Delft, H.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, 2017.

Nevertheless, our original statement indeed may have caused confusion and the point is that our reanalysis is based on many stations in the calibration, so we have rephrased the sentence: “This is a very high calibration coverage for a GHM (Beck et al., 2017), that can be explained by the relatively high coverage in river gauging stations in Europe.”

Line 734-738 ‘It is difficult to compare HERA with other recent hydrological reanalyses (…) for several reasons: (i) spatial coverage (global vs continental), (ii) spatial resolution (0.25º, 0.05º, 1’), (iii) temporal coverage (iv) dynamic vs static socioeconomic conditions’: results can still be compared. Since this study presents a dataset the reported performance of the datasets can be compared. See also line 758-760.

R: We appreciate the reviewer's suggestion to compare the reported performances of HERA with other recent hydrological reanalyses, such as ERA5-GLOFAS. However, a direct comparison of the performance metrics may not be entirely meaningful due to the differences in the validation approaches, catchment sizes, and spatial coverage between the two studies. While it is technically possible to compare the reported performance metrics, such as KGE', Pearson r, bias ratio, and variability ratio (see table below).

Here is a simple comparison of the reported performances between HERA, GLOFAS-ERA5 and the mHM run used for the detailed comparison:

Dataset	HERA	ERA5-GLOFAS	EUmHM	GRFR
Reference	Tilloy et al. (2024)	Harringan et al. (2020)	Samaniego et al. (2019)	
Spatial coverage	Europe	Global	Europe	
Temporal coverage	1951-2020	1979-Present	1960-2010	Global
validation catchments (N)	2848	1801	357	14698
Median validation catchment area (km^2)	583 (27% of catchment area below 250 km2)	30 046	1 700	Not provided (29% of catchment area below 250 kn2)
KGE' (median)	0.55 (58% > 0.5)	0.33	0.6	Not provided (27% > 0.5)
Pearson r (median)	0.73	0.61	0.8	Not provided
Bias ratio (% of catchments with bias ratio between 0.8-1.2)	50	28	50	44
Variability ratio (% of catchments with variability ratio below 1)	83	61	65	Not provided

In our opinion, the difference in stations used in the validation in different studies (catchment area, spatial coverage) makes this quick comparison rather meaningless. We argue that a meaningful comparison would be to extract data for the same catchments and compare performances similarly to what has been done with the mHM run. Furthermore, this comparison would be on a small amount of large European catchments, which is less meaningful than the comparison done with the mHM run.

We add the table to the Supplementary material with a short explanatory text and mention it in the main manuscript (line 655 of new TC document).

Lines 770-782: this is all just speculation. Please use the results for both the calibrated and uncalibrated station performance (i.e. uncertainty due to model parameter values) and the mHM comparison (i.e. uncertainty due to model structure), as these are quantified.

R: We modified this part of the discussion in line with your recommendation, the following sentence has been removed as it is indeed quite speculative: *"The improvement of overall modelling performance through time could therefore be related to improving climate inputs, as observations in ERA5-land become sparser and more inhomogeneous as we go back further in time (Hersbach et al., 2020; Muñoz-Sabater et al., 2021)"*

The following sentences were added to bring more insight about the impact of calibration in skills:

Line 684: "*In summary, the main strength of HERA lies in its relatively low bias in comparison to the other hydrological datasets considered here (Table S6, Figure S6), while its performances are hampered by its underestimation of variability*."

Line 704: "Calibration generally improves streamflow simulations (Hirpa et al., 2018) and also HERA shows a better performance for stations used in the calibration process (Figure 7.d). The negative biases and variability ratios can be related to the different meteorological forcing (EMO-1) used in the calibration, although an underestimation of the variability was also found in the EFAS v5.0 run (that is forced by EMO-1). The method, parameters and skill metrics used for calibration further affects model performance. Despite its qualities, the skill metric used for the calibration presented in Section 2.1.2 (KGE') is known to result in an underestimation of variability (Brunner et al., 2021b) and to put more weight on high values (Garcia et al., 2017). This could partly explain the reduced performance in reproducing extreme low flows observed in Figure 8 and Figure 9."

Line 842-844 'To our knowledge, no other publicly available hydrological reanalysis currently provides discharge data at similar scales and spatiotemporal coverage for Europe.': This study even compares with mHM, which provides discharge data at higher spatial scales and better spatiotemporal coverage (i.e. all cells, not just >100km2) for Europe. Other studies also exist and were provided during the previous review round. Remove sentence.

R: The mHM run used in the comparison has a lower spatial resolution (5kmx5km) than HERA (1.8kmx1.8km), and provides daily data, whereas HERA provides 6-hourly data. So the underlying spatiotemporal resolution of the hydrological modelling is higher compared to any existing product, and therefore is able to capture more of the variability in space and time relevant for hydrological response in catchments. However, for several reasons we decided to make available only data for river pixels with upstream area larger than 100 km2 (corresponding to aggregation of around 30 grid cells in HERA, or 4 in mHM and 1 in GLOFAS ERA5).). Additionally to reducing storage space, we consider uncertainties being too high at grid level and for very small catchments. We observe a performance decline for smaller catchments, as shown in Figure 7b of the manuscript. The objective of our hydrological run is not to provide an accurate simulation at pixel level, which we argue is not yet possible due to data limitations and conceptualization of catchment processes, but rather to reproduce catchment response at uniform spatiotemporal scale across Europe with reasonable accuracy. We argue that an upstream area of 100 km2 is a fair compromise in this respect.

Line 852-854 'The increased spatial resolution improves the performance due to a better representation of hydrological processes and inputs required to simulate them, including the river network': this conclusion cannot be drawn from this study. The performance increase (compared to what?) could also be due to the changes in inputs this study made.

R: R: This statement is backed with references that also deal with high-resolution GHMs. We are not saying that performance has increased compared to a benchmark, but that increasing the resolution tends to increase performance.

Line 859-862 'The modelling framework developed here further forms a basis for creating alternative (counterfactual) time series of river discharges where climatic or socioeconomic conditions can be kept static, enabling the attribution of changes in hydrological regimes across Europe': indeed, the modeling framework, not HERA itself.

R: We are happy the reviewer agrees with this statement.