We thank the reviewer-1 for the thorough review and comments. We have now addressed all the points raised by the reviewer. Below are our detailed responses to the reviewer's comments, along with the respective changes that we will make in the manuscript. We hereby address them individually. In this document we indicate the Reviewer's comments in *italic red*, while text that was changed in the paper in blue. The line numbers mentioned here is in the track change file of manuscript.

The manuscript by Tiwari et al. presents the use of a novel metric, SPAEF, for estimating the spatial variability of hydrological variables during hydrological model calibration. Various calibration strategies, including stepwise calibration and traditional global optimization methods, were employed. To enhance clarity and conciseness, the abstract and introduction should be streamlined to ensure that readers can easily grasp the take-home messages from the study. Some introductory content may be relocated to other sections. Furthermore, it is essential to clearly and succinctly illustrate the novelty of this study.

In response to the reviewer's comments on enhancing clarity, conciseness, and emphasizing the novelty of the study, we have undertaken the following revisions to the manuscript:

- Refined the abstract to summarize the key findings and contributions briefly.
- Streamlined the introduction by eliminating unnecessary details and relocating introductory content for better alignment to ensure that the key insights conveyed by the study is effectively communicated to the readers.
- Articulated the unique aspects of the study, particularly the innovative use of SPAEF and the impact of different calibration strategies.

Lines 3-6: These two sentences appear somewhat unrelated to the main focus of this study. Consider removing them from the abstract.

Removed the lines 3-6 in updated manuscript.

Line 8: What is SNODAS? It might not be well known (which may not be true). Please provide a complete description of this acronym.

Provided complete description of the acronym for SNODAS (Snow Data Assimilation System)

Lines 9-11: This information may not constitute the primary takeaway from the paper and may not be necessary.

We agree with the reviewer, as this information does not contribute to the primary takeaway ,we removed it from abstract.

Line 21: Consider adding a sentence at the end of the abstract to explicitly emphasize the novelty of this study.

Added a line emphasizing the novelty of the study "The novelty of this study is the implementation of SPAEF with respect to spatially distributed SWE for calibrating a distributed hydrological model. Lines 22-23

Line 35: Consider specifying "during the snow-melt season."

Specified "snow-melt season, typically from March to May" Line 38

Line 52: Please provide specific details regarding the fine resolution.

Added specific details of resolution "GlobSnow resolution (25 km * 25 km) , SNODAS (SNOw Data Assimilation System) resolution (1 km* 1 km)". Lines 52-55

Lines 54-58: Consider relocating these sentences to the discussion or conclusion section to avoid confusion among readers regarding the use of TSMM SWE.

We agree with the reviewer, moved the lines to conclusion for better clarity in updated manuscript.

Line 94: Consider using "variable" instead of "parameter" to clearly distinguish between model parameters and hydrological variables.

Replaced parameters with variables for better clarity between model parameters and hydrological variables.

Lines 115-141: This paragraph could be relocated to the methodology section. Additionally, in the description of the optimization algorithms, please clarify the number of repeated experiments conducted to obtain parameter values.

Moved the paragraph to methodology and added "A total of 1000 iterations conducted for both DDS and PADDS to optimize parameter values." Lines 253-254

Line 115: The final paragraph of the introduction should emphasize the novelty of this study, potentially by comparing it to current methodologies or existing approaches for optimizing model parameters.

"The primary objective in this study is to introduce spatial calibration with SWE data using newly developed metric SPAEF for the calibration of the HYDROTEL hydrological model. We applied SPAEF in combination with other traditionally used objective functions. We conducted seven distinct calibration experiments, each employing a unique combination of objective functions. This allowed us to assess the trade-offs and robustness of these various calibration scenarios by evaluating their performance in terms of both streamflow and spatial SWE patterns. Notably, while SPAEF has been previously applied in studies involving evapotranspiration (Demirel et al., 2018) and soil moisture (Eini et al., 2023), this study uses SPAEF with SWE for the first time." Lines 150-154

Line 164: Add "the variation of" before SWE for clarity.

Corrected in manuscript.

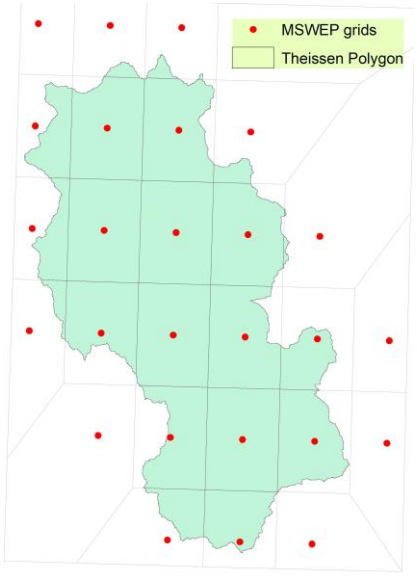
Line 169: Please clarify the time step of the model employed in this study.

HYDROTEL is simulated on a daily basis (opted for this study). Line 183

Line 177: Please explain the selection of the Thiessen polygon method. Does the elevation change the precipitation (lapse rate)?

The Thiessen polygon method is selected to interpolate the meteorological data spatially, the values recorded at the station nearest to a specific cell is allocated to that respective cell. When Thiessen polygon is applied on a gridded data, the polygons generated outlines the area where a specific grid cell is the closest .Therefore, the Thiessen polygons align with the precipitation grid and here we can see that the application of a Thiessen polygon is equivalent to the grid.

"with vertical precipitation gradient of 1mm/100m and vertical temperature gradient of -1°C/100m." Lines 191-192



Line 205: Specify the nature of the data used for March. Is it the mean value of March or daily values?

"average SWE of March" Line 221

Line 223: As mentioned earlier, please specify the time step of the streamflow data.

"The streamflow data is available in a daily basis" Line 215

Line 226: The sentence regarding NSE and similar sentences describing RMSE, KGE, etc., may not be necessary. These details are more suitable for a project report than a science paper.

Removed sentences that were not necessary for the study.

Lines 266-269: Further elaborate on the rationale for selecting March for SPAEF analysis to avoid appearing arbitrary. Consider testing SWE throughout the year or in other months using a similar approach for model calibration and provide this analysis in the Supplementary material.

Thank you for this comment. In response to your comments, I would like to clarify the rationale behind our selection of March as the month for SPAEF analysis in our study. March was chosen as it is the month with the maximum SWE. Our aim was to utilize

the maximum SWE information available during this period. However, we acknowledge that March, despite having the maximum SWE, also coincides with the snow melting period, which could potentially affect the calibration of our analysis. We conducted additional analyses using data from January and February. The results showed that SPAEF performs well with data from both these months. We believe that further research is necessary, with different watersheds and periods used to compute SPAEF, to understand the performance of SPAEF more accurately. The detailed results of these additional calibration can be found in the supplementary material, providing a comprehensive view of the model's performance.

	Calibrated with respect to SPAEF_March & NSE	Calibrated with respect to SPAEF_February & NSE	Calibrated with respect to SPAEF_January & NSE
NSE	0.737	0.739	0.733
KGE	0.764	0.771	0.840
RMSE Spatial	39.38	51.90	50.23
SPAEF wrt SNODAS Jan	0.01	0.077	0.101
SPAEF wrt SNODAS Feb	0.157	0.201	0.181
SPAEF wrt SNODAS March	0.232	0.197	0.167

We added supplementary material with results when calibrated with other months (January and February) and the updated rational in the manuscript.

"For this study, March was selected for SPAEF calibration as it is the month with the highest SWE. Our objective was to leverage the maximum SWE information available during this period. However, we recognize that March, despite having the highest SWE, also overlaps with the snow melting period, which could potentially influence the calibration of our analysis. We performed additional analyses using data from January and February, and the results demonstrated that SPAEF performs well with data from both these months. We believe that further research is necessary, with different watersheds and periods used to compute SPAEF, to more accurately understand SPAEF's performance during the onset of snow accumulation and the snowmelt period. The detailed results of these additional calibrations can be found in the supplementary material, providing a comprehensive view of the model's performance." Lines 583-590

Line 275: Consider placing the KGE metric for model validation in a separate section, distinct from the objective functions used for optimization.

A separate subsection is introduced " Other metric used in this study: One other metric KGE (Kling–Gupta efficiency) is computed for all the calibration experiments. It has been used to assess overall model performance for the various calibration scenarios that were investigated in this study. Lines 322-324

Lines 457-459: The support for this result appears insufficient. Please provide additional information or clarification.

The explanation is added “The sensitivity of RMSE to outliers is a common concern while using it in calibration. Outliers can significantly impact RMSE calculations, and their likelihood of occurrence aligns with the normal distribution that underlies RMSE (Chai and Draxler, 2014). When model biases are pronounced, it may be necessary to address these systematic errors before calculating RMSE. However, the bias insensitivity of SPAEF offers a valuable solution to this challenge (Koch et al., 2018). SPAEF mitigates the impact of uncertainties in observations, providing a more robust and stable approach to model calibration and evaluation. Lines 520-525

Line 530: When drawing the conclusion, be specific and careful about specifying the type of hydrological model and the situations in which this conclusion holds true.

Corrected the conclusion with “distributed hydrological model” Line 607

Figures: The figure axes and the labels should be more obvious. Currently, they are too small.

Updated the figure fonts.

Overall, I think the paper exhibits novelty, especially in introducing a new objective function for SWE in model parameter calibration. However, the writing style of the scientific paper can be further improved. I would recommend accepting this paper after moderate revisions.

Chai, T. and Draxler, R. R.: Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature, Geoscientific Model Development, 7, 1247–1250, https://doi.org/10.5194/gmd-7-1247-2014, 2014.

Demirel, M. C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrology and Earth System Sciences, 22, 1299–1315, 2018.

Eini, M. R., Massari, C., and Piniewski, M.: Satellite-based soil moisture enhances the reliability of agro-hydrological modeling in large transboundary river basins, Science of The Total Environment, 873, 162396, https://doi.org/10.1016/j.scitotenv.2023.162396, 2023.

Koch, J., Demirel, M. C., and Stisen, S.: The SPAtial EFficiency metric (SPAEF): Multiple-component evaluation of spatial patterns for optimization of hydrological models, Geoscientific Model Development, 11, 1873–1886, 2018.