Response to Reviewer RC1

Title: Achieving water budget closure through physical hydrological processes modelling: insights from a large-sample study
Authors: Xudong Zheng, Dengfeng Liu*, Shengzhi Huang*, Hao Wang, Xianmeng Meng
Manuscript ID: hess-2024-230

Reply on RC1：

Many thanks for taking the time and effort to review our paper. All comments from Reviewer RC1 are addressed below with point-by-point responses.

For better readability, replies will start with "**R/**", following the original comments that start with "**C/**" and are shown in **bold**. The revisions to be added into the revised manuscript is highlighted in red. The important parts are highlighted in blue. The quoted content is displayed in *italics*.

Point-to-point response:

C/ This paper emphasizes the issue of decreasing data confidence at the watershed scale in the era of big data, caused by the non-closure of water budget from multiple data sources. In their analysis, the total water budget residuals were quantitatively decomposed into two components, inconsistency and omission residuals, to account for different drivers of water budget non-closure phenomenon. This is an interesting addition, as previous studies have typically given little or only qualitative consideration to the water imbalance caused by omissions in the original water balance equation.

Attempting to close the water balance is valuable, both hydrological inference under climate change and hydrological modeling require data that satisfy the basic assumption of water balance. The PHPM-MDCF proposed in this work employ hydrological model to constrain multisource datasets, which is reasonable because hydrological models are well-known for their water balance capabilities. The correction also seems to be effective, which comes from the validation with results from large sample basins.

R/ Thanks for your positive feedback and recognition of our work. Your comments are valuable for revising and improving our paper. Below, we provide detailed responses to each of your concerns. The corresponding revisions will be incorporated into the subsequent revised manuscript.

C/ However, there are still some concerns that need to be explained in the response or addressed in the manuscript. The authors have observed the typical seasonal pattern of non-closure phenomena but lack corresponding explanations. In addition, although the authors decomposed the closure residuals into two parts, it seems that only the inconsistency residuals were corrected. What is the rationale behind this approach? Why were the omission residuals not corrected?

R/ Your points are very insightful. Adding explanations about the seasonal characteristics of the non-closure phenomenon will indeed strengthen our argument. In addition, as you mentioned, our framework primarily addresses the Resi (inconsistency residuals), without considering the correction of Reso (omission residuals). This is because we consider it as unaccounted-for water in the original water budget equation, which should be explained by other water components. We provide a more detailed clarification in our response to the second major concern and include further revisions to the manuscript. By addressing these two questions, we anticipate that our manuscript will be significantly improved.

C/ In summary, this paper is innovative and aligns with the interests of potential readers of the HESS. After careful consideration and revision, this work has the potential to make a significant contribution to this field. As they described, the underlying Bayesian philosophy is an approach for aligning our understanding of natural processes with real-world observations.

R/ Thank you again for acknowledging our work and perspectives.

Major concerns

C/ Sect. 4.1, the patterns of the Res are of interest to me. The authors identified typical spatial distributions and compared them with previous studies in Sect. 4.4, explaining these patterns through hydro-meteorological conditions and watershed area. From a physical perspective, this explanation is consistent with common sense and is sufficient for me. However, the temporal patterns of the Res are also of interest (Fig. 5). The authors should provide further explanation in this regard or compare them with previous studies, as this could offer valuable insights into the causes of the non-closure of water balance.

R/ Thanks for your suggestion. As we mentioned earlier, we agree with your suggestion to include an explanation of the temporal distribution of Res. This will be addressed in two ways: (1) Comparing the observed seasonal patterns in Res (residuals) with previous studies, and (2) providing an analysis from a physical causation perspective.

Indeed, the temporal distribution of Res shown in Fig. 5 is quite striking. Specifically, as we mentioned in our manuscript, there is a positive bias in Res during the warm season and a negative bias during the cold season. By comparing with previous literature, we found similar temporal distributions and potential influencing factor—namely, the potential underestimation of warm-season evaporation and cold-season precipitation (Kauffeldt et al., 2013; Newman et al., 2015; Lv et al., 2017; Abolafia-Rosenzweig et al., 2020; Robinson and Clark, 2020).

From a physical perspective, the underestimation is related to phenomena such as snowfall, freezing rain, and non-convective precipitation that occur during the cold season, as well as the calculation of evaporation during the warm season.

A further analysis was conducted to examine this by comparing the ratios of evaporation and precipitation for cold and warm seasons separately, along the corresponding Res. Scatter plot shows that basins dominated by cold-season precipitation are more likely to exhibit larger negative Res during cold-season, while basins with higher warm-season evaporation tend to have larger positive Res during warm season.

In both cases, the Res are more sensitive to underestimation of precipitation and evaporation, which is consistent with findings from previous research.

Although it is impossible to obtain true values to evaluate the measurements, these results still highlight potential uncertainties in cold-season precipitation and warm-season evaporation measurements, which could severely impact the assumption of water balance.

The analysis process and corresponding figure to be added into the revised manuscript are given below:

"The pronounced seasonal pattern of non-closure residuals depicted in Fig. 5 is quite interesting. To gain more insight into the observed pattern, we compare it with the temporal factors reported in the literature. The first and foremost reported factor associated with the observed negative biases in Res during the cold season is the underestimation of precipitation (Newman et al., 2015). This systematic bias is related to phenomena such as snowfall, freezing rain, and non-convective precipitation that occur during the cold season, where measurements and simulations are prone to significant errors, including the well-know undercatch phenomenon (Kauffeldt et al., 2013; Robinson and Clark, 2020). Another key factor influencing water budget non-closure is connected to the temperature and evaporation dynamics. Abolafia-Rosenzweig et al. (2020) evaluated the water budget residuals over 24 global basins and found that the likelihood of positive biases in the water balance increases with rising temperatures, which likely induced by the potential uncertainties in evaporation estimates. The research by Lv et al. (2017) also support this perspective, indicating that the underestimation of evaporation is a primary contributor to the water budget non-closure. In summary, according to the literature, cold-season precipitation and warm-season evaporation seem to be the primary drivers of the temporal distribution of Res. To examine this, although it is impossible to obtain the true values, we can provide some evidence by comparing evaporation and precipitation, along with the corresponding residuals, between cold and warm seasons.

Figure 12 depicts this relationship by comparing the ratios of evaporation and precipitation for the cold and warm seasons separately, with the corresponding water budget residuals. For the cold season, the scatter points can be split into two distinct regions along the vertical line where the ratio is 1. The scatter points in the left region indicate basins where cold-season precipitation is lower than in the warm season, leading to relatively smaller absolute residuals (clustered around zero residuals). In contrast, scatter points for basins with dominant cold-season precipitation are dispersed below the zero residual line, with lager negative residuals becoming more prevalent as the proportion of cold-season precipitation increases. In other words, regions where cold-precipitation constitutes a larger proportion of the water budget residuals are more sensitive to the underestimates of precipitation, resulting in larger negative residuals. Furthermore, we observed similar trends in the warm season, where a higher proportion of warm-season evaporation is associated with larger positive residuals. These results confirm the perspective of previous research, highlighting the potential uncertainties in measurements of cold-season precipitation and warm-season evaporation."

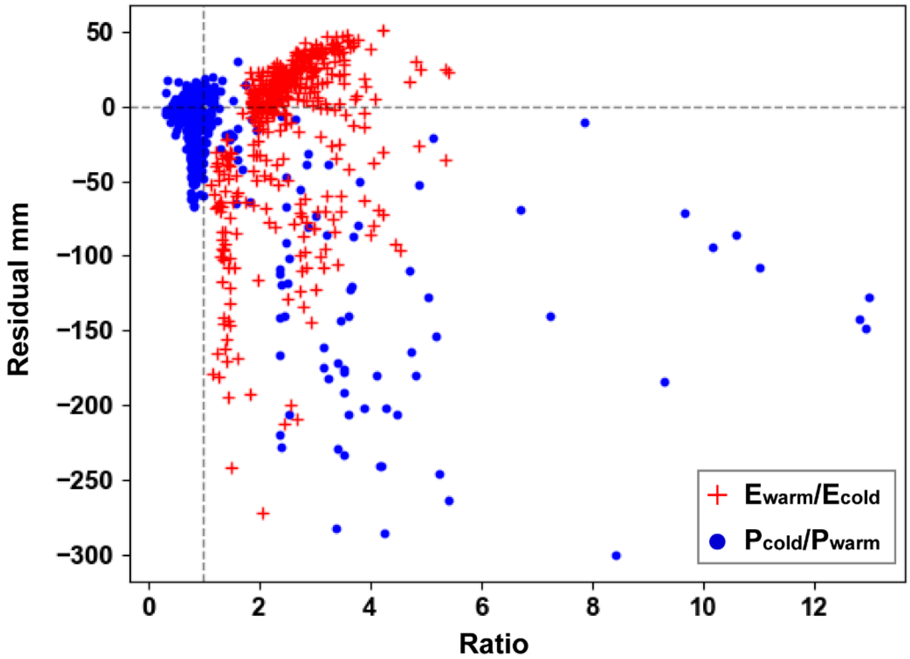

Figure 12. Relationship between the ratios of evaporation and precipitation for the cold and warm seasons separately and the corresponding water budget residuals. Note that blue represents residuals for the cold season, and red represents those for warm season. The seasonal division are consistent with Fig.5. The unit of residuals is "mm".

C/ From Fig. 6, it appears that the Res and Resi have been effectively corrected, but the Reso have not changed significantly. Is this merely a specific case for this basin or a general situation? If it is a general situation, dose this imply that PHPM-MDCF only corrects for Resi and does not account for Reso? I believe that further explanation of this treatment could improve the transparency of the methods used in the paper.

R/ Yes, as you mentioned, our framework only corrects for Resi and does not account for the correction of Reso. This is a general situation for all basins. Essentially, such treatment is guided by the underlying logic of the correction process, as revealed by the residuals decomposition in Eq. 3. Reso is separated from the total water budget residuals to account for water components not considered in the original equation, such as inter-basin exchange.

From a causal perspective, this portion of residuals is less associated with physical inconsistency, as confirmed by the spatiotemporal distribution difference (Fig. 4-5) discussed in Sect. 4.1. Therefore, the framework focused on constraining residuals using physically consistent hydrological model cannot correct this part of residuals. This also explains why Resi decreases significantly after correction in Fig. 6, while Reso remains unchanged.

In addition, the discussion in Sect. 4.2 also highlighted this issue:

"However, despite recalibrating the model with corrected datasets, Res_o driven by the omission in water budget equation exhibited no substantial changes before and after correction (e.g., the monthly mean absolute values maintain around 6.5 mm, see Fig 6f). This phenomenon occurs because we only corrected the inconsistency residuals with reference to the simulation system, while the omission accounting for addition water terms should not be corrected in the existing datasets."

In our opinion, using measurements to describe the theoretical water balance requires two key conditions: (1) physically consistent measurements, and (2) comprehensive description of the water budget equation. However, this is challenging to achieve in practice, whether due to inadequate understanding or limitations in measurement techniques, resulting in residuals corresponding to Resi and Reso.

The framework proposed in this work can, to some extent, enhance physical consistency between measurements through the model, resulting in reduced Resi. However, achieving a more comprehensive description (i.e., reducing Reso) may involve more issues, such as scale effects, more detailed data (both surface and subsurface), and a deeper understanding of the watershed. Addressing these questions is beyond the scope of this study. We look forward to more detailed future research addressing these issues, as mentioned in our discussion:

"*Further investigation would be required to better understand the omission residuals from a physical perspective. For example, a distributed hydrological model with representation of subsurface later flow process will allow us to identify the magnitude of inter-basin interactions; a more detailed description of water budget equation in data-rich environments can help us examine the sources of omission errors. This is undoubtedly important, but not the focus here.*"

Thank you again for your reminder. We recognize the difficulties caused by insufficient explanations and will further emphasize this issue in the revised manuscript.

C/ Although the author has clearly articulated the main scientific problem of the paper, there are still areas that could be further improved, which I have listed in the specific issues.

R/ Tank you for your thorough and detailed review. We have addressed each point and provided responses below.

Specific issues

C/ Line 22-25: According to the results, it seems that humid/wet basins are also prone to larger closure residuals, which needs to be emphasized here.

R/ According to your suggestion, we will revise the phrasing to:

"This emphasizes the importance of carefully evaluating the water balance assumption when employing multisource datasets for hydrological inference in small and humid basins."

C/ Line 36-46: I believe this section should place greater emphasis on the issues of scale mismatch and difficulty in obtaining reference data.

R/ Thanks for your comment. We will revise the manuscript to strengthen the issue of scale mismatches and the challenges associated with obtaining site data. The following statement will be added:

"The issue of scale mismatches and the availability of site data in certain regions also pose challenges for data evaluation."

C/ Line 58-60: It is recommended to cite the review by Beven (2002).

R/ Thanks for your suggestion, we will include this reference to the following sentence:

"Such inconsistency poses an obstacle to robust hydrological inferences (Beven, 2002)."

C/ Line 83-84: It is recommended to add references to support the argument.

R/ We found supporting evidence in the literature Luo et al. (2023) and will include this reference to substantiate our argument.

Luo et al. (2023): *"therefore, the results confirm that increasing the water budget closure accuracy of budget-component data sets reduces the accuracy of individual budget-component products."*

"In the context of applying such closure constraint, it becomes evident that the precision of certain individual components may notably deteriorate, particularly when uncertainties are challenging to quantify (Luo et al., 2023)."

C/ Line 119: "Res" does not appear to be in italics.

R/ Thank you for your reminder. This formatting issue will be corrected throughout the revised manuscript.

C/ Line 126-127: It is recommended to change it to: "(a) How can the total water budget residuals be quantitatively decomposed into inconsistency and omission residuals based on Eq. (3)?"

R/ Thanks for your careful review. We will revise the sentence according to your suggestions:

"(a) How can the total water budget residuals be quantitatively decomposed into inconsistency and omission residuals based on Eq. (3)?"

C/ Table1: The "period" should be "Original Period".

R/ Thanks, we will correct this.

C/ Figure5: The figure caption seems to contain an error. There are no other subfigures.

R/ The caption of this figure does indeed contain errors due to update to the figure, and we will revise it accordingly (i.e., remove redundant subplots sequence numbers):

"**Figure 5.** Temporal distribution of monthly water budget residuals (Res), inconsistency residuals (Res_i), and omission residuals (Res_o) across 475 CAMELS basins with reliable simulations. Boxplot-like diagrams describe variability across catchments, and outliers represent the 10th and 90th percentiles. The unit of residuals is "mm"."

C/ Line 332-334: The argument here doesn't seem to correspond with the figure. Could it be that the figure has been updated?

R/ Thank you for pointing out this error. We will revise the statement while updating the figure caption (remove redundant subplots sequence numbers):

"On the contrary, Res_o tends to be mainly positive except from September to November; its extent of variability is also significantly smaller than that of the other two residuals. In regard to magnitude, Res_i is greater than Res_o, whether considering positive or negative bias."

C/ Line 418: add "which are" before "implemented".

R/ Thanks, we will correct it to:

"This is achieved by the representation of physical hydrological processes underlying the correction strategy, which are implemented in the hydrological model, such as runoff generation and routing."

C/ Line424-426: Change the sentence to "The fact that simultaneous corrections of other variables during extreme runoff noise corrections did not significantly differ from OS-based corrections further enhances our confidence in PHPM-MDCF."

R/ Thanks for your comment. We will revise the statement according to your suggestion.

"The fact that simultaneous corrections of other variables during extreme runoff noises correction did not significantly differ from OS-based corrections further enhances our confidence in PHPM-MDCF"

C/ Line 417: It is necessary to further emphasize the issue of the non-closure phenomenon in humid regions.

R/ Thank you for your suggestion, we will revise the entire manuscript to emphasize the issue of non-closure phenomenon in humid regions. Below are several examples of the revisions will be made:

Abstract: "This emphasizes the importance of carefully evaluating the water balance assumption when employing multisource datasets for hydrological inference in small and humid basins."

Sect. 4.4: "These results highlight the risks of using multisource datasets for hydrological inference in humid and small-scale basins—specifically, potential physical inconsistencies—and underscore the need to carefully test the water balance closure assumption."

Conclusion: "This highlights the need for careful consideration of the water balance assumption when applying multisource datasets for hydrological inference in small and humid basins."

C/ Figure 12: There seems to be a mistake with the R2 values.

R/ Thank you for pointing this mistake. We will correct this mistake in the revised manuscript. The updated figure with corrected R2 values is shown below:

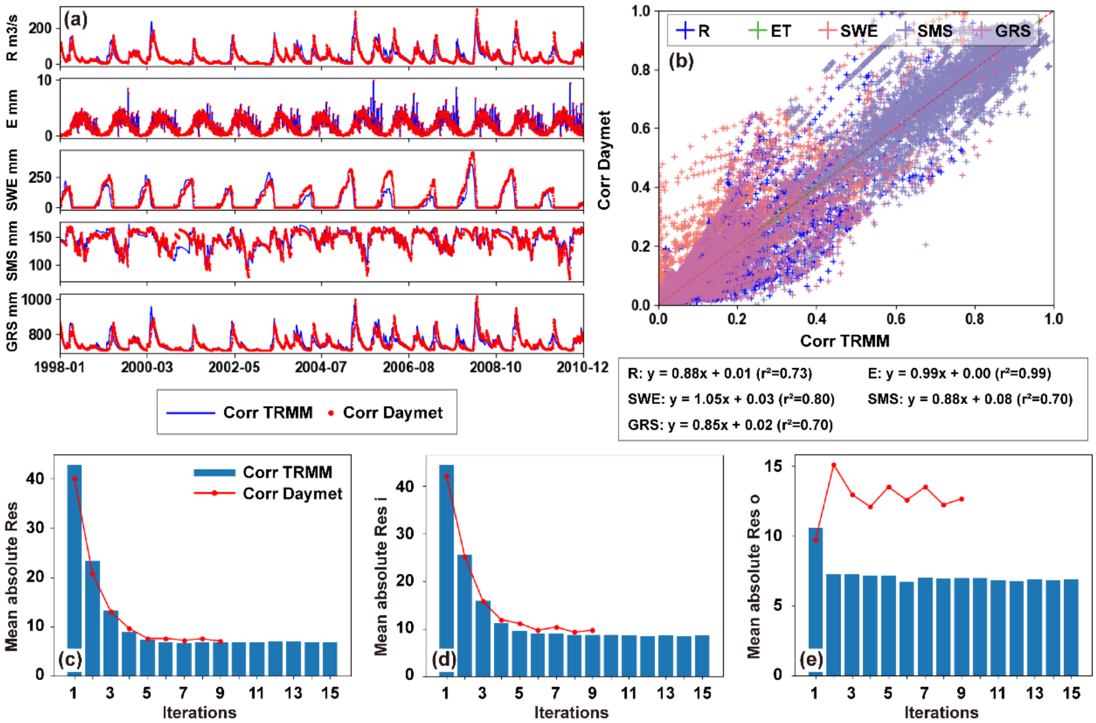

Figure 13. Comparison of correction results based on different forcing datasets (TRMM and Daymet) at basin 1013500. (a-b) Corrected time series of five water budget variables. (c-e) Variation of long-term mean absolute values of three residuals with correction iterations at the monthly scale. The unit of residuals is "mm".

C/ Line 639: Humid regions is a better expression.

R/ Thanks for your suggestion. We will make revisions throughout the entire manuscript. Here is an example:

"This highlights the need for careful consideration of the water balance assumption when applying multisource datasets for hydrological inference in small and humid basins."

Reference

Abolafia-Rosenzweig, R., Pan, M., Zeng, J., and Livneh, B.: Remotely sensed ensembles of the terrestrial water budget over major global river basins: An assessment of three closure techniques, Remote Sensing of Environment, 252, 10.1016/j.rse.2020.112191, 2020.

Kauffeldt, A., Halldin, S., Rodhe, A., Xu, C. Y., and Westerberg, I. K.: Disinformative data in large-scale hydrological modelling, Hydrology and Earth System Sciences, 17, 2845-2857, 2013.

Luo, Z., H. Li, S. Zhang, L. Wang, S. Wang, and L. Wang (2023), A Novel Two-Step Method for Enforcing Water Budget Closure and an Intercomparison of Budget Closure Correction Methods Based on Satellite Hydrological Products, Water Resources Research, 59.

Lv, M., Ma, Z., Yuan, X., Lv, M., Li, M., and Zheng, Z.: Water budget closure based on GRACE measurements and reconstructed evapotranspiration using GLDAS and water use data for two large densely-populated mid-latitude basins, Journal of Hydrology, 547, 10.1016/j.jhydrol.2017.02.027, 2017.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., Viger, R. J., Blodgett, D., Brekke, L., Arnold, J. R., Hopson, T., and Duan, Q.: Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance, Hydrology and Earth System Sciences, 19, 209-223, 10.5194/hess-19-209-2015, 2015.

Robinson, E. and Clark, D.: Using Gravity Recovery and Climate Experiment data to derive corrections to precipitation data sets and improve modelled snow mass at high latitudes, Hydrology and Earth System Sciences, 24, 1763-1779, 10.5194/hess-24-1763-2020, 2020.