Response to Reviewer CC1

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 CC1:

Thank you very much for your interest in our paper and for taking time and effort to review it. All comments from Reviewer CC1 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/ In the context of fast development of measurement techniques, it is our mission to develop methods to leverage the advantages of the measured variables and thus promote the hydrological simulation. This study is a valuable try, which proposed a multisource datasets correction framework, the PHPM-MDCF, to achieve water budget closure with calibration of various variables. This experiment was carried out in 475 COUNS basins, showing great potential to reduce the inconsistency residuals.

R/ We appreciate your recognition of the importance of our work. We hope that this paper can contribute to the data foundations in various fields, such as earth system science and hydrology, within the context of big data. Your comments are very valuable in enhancing the quality of our manuscript. Below, we will provide point-by-point responses to these comments and make the corresponding revisions in the manuscript.

Major concerns:

C/ (1) There are PTRMM in both Eq. (5) and (6), then how do we reduce the inconsistency residuals brought by P in the water budget?

R/ This is a crucial point, but cannot be solved within the current framework. As we have assumed that "*(2) the uncertainties associated with the model forcing and structure can be considered negligible during the modelling process*" in the methods section. In fact, the PHPM-MDCF employs the distance between simulations and measurements to allocate residuals corrections among variables. As a forcing or boundary condition, precipitation cannot be corrected within this framework, or in other words, it cannot be simulated.

Nevertheless, we consider that the uncertainty, or residuals, in precipitation has a minimal impact on the correction of other variables measurements. Some evidences are provided in Sect. 5.2.1, where a comparison of correction results under different precipitation forcing (i.e., TRMM and Daymet) reveals that the correction shows minimal sensitivity to the precipitation forcing.

"In summary, the above results suggest that the correction is minimally sensitive to the choice of forcing, demonstrating the robustness of the correction results."

Theoretically, such behavior stems from the adaptability of hydrological mode to the input data, specifically the calibration compensation capability we described in the introduction (Wang et al., 2023). This enables the model to generate reasonable representation of hydrological process even with imprecise forcing.

However, can the current results offer any guidance or insights for precipitation correction? The answer is affirmative. It is the comparison of corrections with different precipitation presented in Sect. 5.2.1 that highlights the impacts associated with varying precipitation inputs. Starting from this point, we can discern some potential clues.

It is evident that different precipitation products do not impact the correction of inconsistency residuals (Fig. 14c-d) but do results in varying omission residuals (Fig. 14e). On the one hand, discrepancies in precipitation products are compensated by model calibration, result in similar representation of hydrological process and thus similar inconsistency residuals corrections.

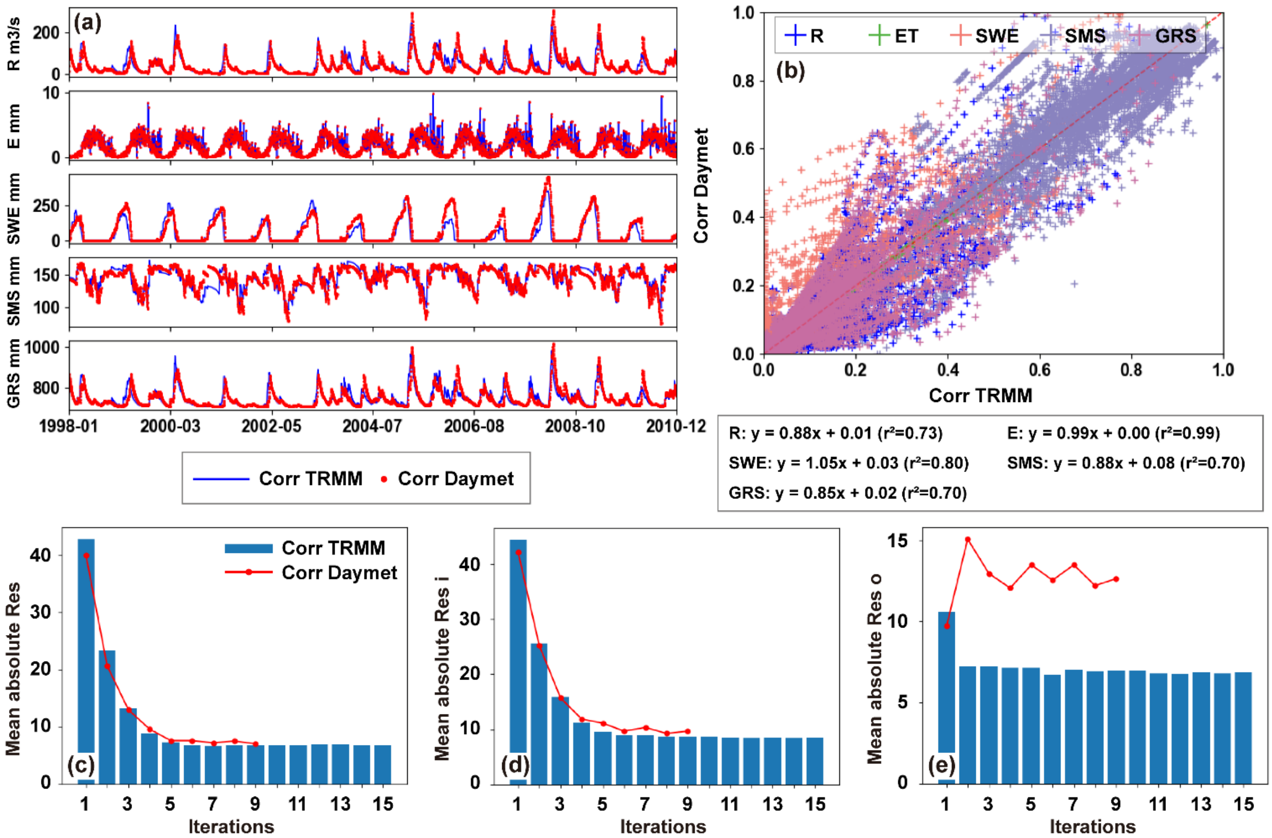

Figure 14. Comparison of correction results based on different forcing datasets (TRMM and Daymet) at basin 1013500. (a-b) Corrected time series of five water budget variables. (c-e) Variation of long-term mean absolute values of three residuals with correction iterations at the monthly scale. The unit of residuals is "mm".

On the other hand, the precipitation products exhibit a systematic bias. In particular, Daymet reports significantly lower precipitation in this basin compared to TRMM (see Fig. S13). Such bias will manifest in the water budget equation, leading to different total input water volumes. Consequently, with the inconsistency residuals of other variables unchanged, maintaining the water balance would require an increase in Res_o (Fig. 14e). Note that the Res_o presented in Fig. 14e represents the mean of absolute values.

Fig. S13 will be added to our manuscript along with the corresponding explanation (Sect. 5.2.1).

"The comparison of the two precipitation products is presented in Fig. S13, where Daymet precipitation is significantly lower."

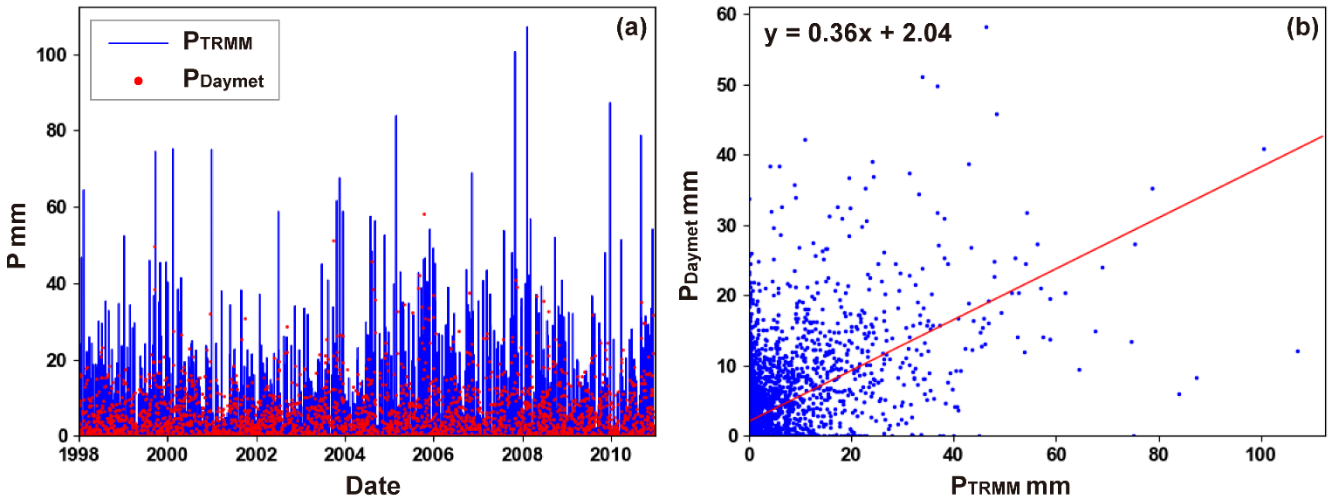

Figure S1. Comparison of TRMM and Daymet precipitation products.

Therefore, it can be inferred that, with other variables unchanged, TRMM precipitation demonstrates superior water budget closure compared to Daymet precipitation, which contains larger inconsistency residuals. This difference in inconsistency residuals is directly reflected in the variations in omission residuals after correction (Fig. 14e). In other words, this portion of the omission residuals (i.e., the difference between the two omission residuals after correction) can be directly corrected in the precipitation.

Note that not all omission residuals can be corrected in the precipitation data, as it still contains residuals from some unknown omitted water content. In other words, such correction must be relative and based on comparisons using multiple precipitation products, as the true values and perfect water balance equation are unattainable. Only through comparisons can the discrepancies in Res_o arising from precipitation inconsistencies be identified.

To focus the attention of this paper on the PHPM-MDCF framework, we have not conducted actual experiments here. Instead, we will introduce this idea in the discussion section.

"It is noted that the PHPM-MDCF has limitations in addressing inconsistency residuals in forcing. The reasons are twofold. On the one hand, this is due to our neglect of uncertainties in the forcing, which, as indicated by the above analysis, appears to have limited impact on the correction for other variables. On the other hand, this is because the PHPM-MDCF allocates residuals based on the distance between simulations and measurement, while the forcing cannot be simulated within the hydrological model. In this case, is there a potential to correct the inconsistency residuals in the forcing? Clues to this possibility are hidden in the above analysis. Systematic biases in precipitation products are directly reflected in the water budget equation, leading to different total input water volumes. Consequently, with the inconsistency residuals of other variables unchanged, maintaining the water balance would require an increase in omission residuals (Fig. 14e). Therefore, it can be inferred that, with other variables unchanged, TRMM demonstrates superior water budget closure compared to Daymet, which contains larger inconsistency residuals. In other words, the differences in omission residuals reflect the discrepancies in precipitation inconsistency residuals. This portion of the omission residuals difference can be directly corrected in the precipitation. However, it is worth noting that not all omission residuals can be corrected in the precipitation, as it still contains residuals from some unknown omitted water content. Such correction must be relative and based on comparisons using multiple precipitation products, as the true values and perfect water balance equation are unattainable. We will explore the approach in future work and extend the PHPM-MDCF based on this idea."

C/ (2) This paper focuses on the terrestrial water balance (Eq. (1)). However, whether this framework is applicable to broader water balances, such as atmospheric water balance or local water balance, or if any adjustments are needed?

R/ The ideas in your comment are very interesting. Although this paper primarily focuses on terrestrial water cycle systems, exploring broader water balance applications is highly valuable for extending the scope of this research.

Through a review of the literature, we found several water balance equations designed for other systems. For example:

- The steady-state hydrological budget equation of the proglacial zone (Cooper et al., 2011):

  $$W_{PZ} = W_P + W_R - W_E - W_{SSS} - W_{SR} \pm \Delta W_S, \tag{R1}$$

  where W_{PZ} is the net proglacial water flux, W_P is the precipitation water flux, W_R is the channel recharge water flux, W_E is the evaporation water flux, W_{SSS} is the sub-surface seepage water flux, W_{SR} is the surface runoff water flux, and ΔW_S is the change in water storage.

- The atmospheric water vapor budget with a focus on the oceans (Penning et al., 2021):

  $$\frac{\Delta W}{\Delta t} = E - P - \nabla \cdot (vq), \tag{R2}$$

  where W being the total column water vapor and $\nabla \cdot (vq)$ the moisture flux divergence.

- The coupled atmospheric–terrestrial water balance equation (Lorenz et al., 2014):

  $$\frac{dW}{dt} + \nabla \cdot \mathrm{Q} = ET_a - \mathrm{P}, \tag{R2}$$

  where W denotes the total column water content in the atmosphere and $\nabla \cdot \mathrm{Q}$ is the net balance of moisture flux (i.e., moisture flux divergence).

Regardless of the water balance system under consideration, the key to applying the PHPM-MDCF is whether the utilized model can represent the components of the water budget equation. The core principle of the PHPM-MDCF is to characterize the physical relationships among water budget components through the model, thereby imposing closure constraints on the measurements. As we noted in the last paragraph of Sect. 4.3.2:

"The physical relationships among various water budget variables, as representation by the model, are also imposed onto the measurements through the correction process. This constitutes the core principle of PHPM-MDCF."

In other words, the application of the PHPM-MDCF to more complex systems to conduct correction can be achieved by replacing the hydrological model (HBV) in the framework with other more suitable models that can output more variables, such as physically distributed models (VIC model; Liang et al., 1994), coupled models (WRF-TOPMODEL; Rogelis and Werner, 2018), or even deep learning models (MCP; Wang and Gupta., 2024).

We will add a statement at the end of Sect. 5.2.2 to emphasize this issue.

"By employing models that generate additional output variables, we can more comprehensively represent the water budget equation and extend the application of the PHPM-MDCF to more complex water budget systems."

C/ (3) Uncertainty plays a crucial role, and this study qualitatively address the uncertainty associated with the model structure. A pertinent question is whether this uncertainty can be quantified. While we know that validating this uncertainty through multiple models may be both challenging and unnecessary within the scope of the current work, it would be valuable if the authors could suggest potential avenues for future research and development.

R/ Thank you for your insightful comment. We would like to address this question from the perspective of Bayesian philosophy. In practical Bayesianism, all models are inherently flawed, yet each model can be assigned a level of confidence that indicates the degree to which we trust it (Hoang, 2020). Only by considering more than one theory and model can we more effectively approach the truth. This is also the core idea of the Beven's Alternative Blueprint (Beven, 2002). As they stated:

"Why should there be any expectation of a single 'real' description when the direct observation of the responses of the most important part of hydrological systems is quite beyond our current capabilities and will be until there is a dramatic improvement in the available geophysical techniques?"

"The fact that there may be no unique answer does not mean that the approach is not science or scientific. Indeed, such an approach has then the additional advantage that we will work more naturally with the many potential worlds of future (and therefore unknown) boundary condition scenarios and the uncertain predictions that should ensue (e.g. Cameron et al., 2001)."

We strongly align with this scientific perspective. The "uncertainty" should be regarded as varying descriptions of the assumed "truth" and the associated confidence levels. Relying on a single theory alone is insufficient.

Due to limitations in data and resources, this study employs only one model for measurements correction, and we acknowledge that this introduces uncertainty. In future work, an effective method for quantifying uncertainty is to use an "ensemble" approach. Specifically, employing multiple models (theories) to describe the same hydrological process enables the range of ensemble corrections to be used for quantifying uncertainty. This is very similar to the ensemble forecasting (Nicolle et al., 2014). Additionally, confidence can be assigned to each correction based on the simulation accuracy of each model, resulting in a unique weighted correction outcome.

The description of the possible approach will be added into the Sect. 5.2.2 as follows:

"Additionally, multiple models can be employed for 'ensemble correction', which aids in quantifying uncertainty and providing more robust correction results."

Minor comments:

C/ (1) Please check carefully of the text, to avoid grammatic errors, e.g. km2 in Line 183.

R/ Thank you for your thorough review. We have conducted a comprehensive check and will make the revision. Below is an example of the revision made.

"However, the assumption is fragile when applied to small basin, leading to significant uncertainty in estimating TWSC for basins with areas less than 63,000 km^2 (Lehmann et al., 2022)."

C/ (2) Line 75-76: The semantics are repetitive; it is recommended to delete "to ensure data consistency".

R/ Thanks for your comment. We will delete the redundant expression. The revised sentence is as follows.

"Other approaches, such as post-Processing Filtering technique (PF) and bias correction method (Munier et al., 2014; Weligamage et al., 2023), can also be helpful in closing water budget."

C/ (3) Line 80: "residuals" is more precise than "bias".

R/ Yes, thank you for pointing that out. The revised sentence is as follows.

"However, the closure constraints imposed by the above methods (hereafter referred to as traditional methods) have been questioned, with Abolafia-Rosenzweig et al. (2020) arguing about the potential incorrect assignment of residuals."

C/ (4) Line 131-134: It seems that these sentences should be changed to the past tense.

R/ Thank you for your comment. We will revise the sentence as follows.

"Furthermore, we developed a multisource datasets correction framework based on decomposition of water budget residuals and multi-objective calibration within hydrological modeling. The presented framework, providing the capability to enhance the water budget closure and hydrological connections among multisource datasets, was applied to a large-sample basins dataset across CONUS."

C/ (5) Line 165: "One of the main aims" might be more appropriate.

R/ Thank you for your suggestion. We will revise the manuscript according to your suggestion.

"One of the main aims of this study is to investigate the decomposition of water budget residuals and correction to datasets, rather than comparing the differences and rankings of closure residuals across different dataset combinations."

C/ (6) Line 167: This sentence should be in the past tense.

R/ Thanks for your comment. We will make the revisions as follows.

"In line with this objective, referring to the work of Petch et al. (2023), we strategically selected single product for each water component to construct water budget equation, thereby laying the foundation for further research."

C/ (7) Fig. 3: It is recommended to add further explanations in the caption of Fig. 3.

R/ Thank you for your valuable suggestion. We will add further explanations to the caption of Fig.3 as follows.

"Figure 3. Illustration of the correction process advancing convergence between the simulation and measurement systems. The measurement system is corrected to approach the simulation system, while the simulation system is refined via parameter calibration to better approximate the measurement system. As a result, the distance between the two systems is reduced, leading to better physical consistency in the measurement system."

C/ (8) Line 458: I suggest emphasizing the spatial distribution of water balance closure.

R/ Thank you for your suggestion. Based on your advice, we will revise the sentence to:

"Therefore, we speculate that the spatial distribution of water budget closure is predominantly influenced by the characteristics of the basin."

C/ (9) Line 619: A "." is missing before the "The major".

R/ Thank you for pointing out this oversight. We will add a period between the sentences.

Reference

Beven, K.: Towards an Alternative Blueprint for a Physically Based Digitally Simulated Hydrologic Response Modeling System, Hydrological Processes - HYDROL PROCESS, 16, 189-206, 10.1002/hyp.343, 2002.

Cooper, R., Hodgkins, R., Wadham, J., and Tranter, M.: The hydrology of the proglacial zone of a high-Arctic glacier (Finsterwalderbreen, Svalbard): Sub-surface water fluxes and complete water budget, Journal of Hydrology, 406, 88-96, 10.1016/j.jhydrol.2011.06.008, 2011.

Hoang, L.: La formule du savoir: Une philosophie unifiée du savoir fondée sur le théorème de Bayes, 10.1051/978-2-7598-2261-4, 2020.

Liang, X., Lettenmaier, D. P., Wood, E., and Burges, S.: A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, J. Geophys. Res., 99, 14415-14428, 10.1029/94JD00483, 1994.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M., Sneeuw, N., and Riegger, J.: Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances, Journal of Hydrometeorology, 15, 10.1175/JHM-D-13-0157.1, 2014.

Nicolle, P., Pushpalatha, R., Perrin, C., Francois, D., Thiéry, D., Mathevet, T., Le Lay, M., Besson, F., Soubeyroux, J.-M., Viel, C., Rousset, F., Andréassian, V., Maugis, P., Augeard, B., and Morice, E.: Benchmarking hydrological models for low-flow simulation and forecasting on French catchments, Hydrology and Earth System Sciences, 18, 2829-2857, 10.5194/hess-18-2829-2014, 2014.

Penning de Vries, M., Fennig, K., Schröder, M., Trent, T., Bakan, S., Roberts, J., and Robertson, F.: Intercomparison of freshwater fluxes over ocean and investigations into water budget closure, Hydrology and Earth System Sciences, 25, 121-146, 10.5194/hess-25-121-2021, 2021.

Rogelis, C. and Werner, M.: Streamflow forecasts from WRF precipitation for flood early warning in mountain tropical areas, Hydrology and Earth System Sciences, 22, 853-870, 10.5194/hess-22-853-2018, 2018.

Wang, J., Zhuo, L., Han, D., Liu, Y., and Rico-Ramirez, M.: Hydrological Model Adaptability to Rainfall Inputs of Varied Quality, Water Resources Research, 59, 10.1029/2022WR032484, 2023.

Wang, Y. H. and Gupta, H.: A Mass-Conserving-Perceptron for Machine-Learning-Based Modeling of Geoscientific Systems, Water Resources Research, 60, 10.1029/2023WR036461, 2024.