

Hydrol. Earth Syst. Sci. Discuss., author comment AC2  
<https://doi.org/10.5194/hess-2021-428-AC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC2

Saúl Arciniega-Esparza et al.

---

Author comment on "Remote sensing-aided rainfall-runoff modeling in the tropics of Costa Rica" by Saúl Arciniega-Esparza et al., Hydrol. Earth Syst. Sci. Discuss.,  
<https://doi.org/10.5194/hess-2021-428-AC2>, 2021

---

Thanks for your suggestions and your comments. We attached our responses to each point starting the line with "R:".

- There is a lack of connection between the supplementary section and the main text. For instance, when the authors introduced the CHIRPS product (~L240), they could link it to Fig. S1 to have a clear picture of the improvement. Another example are tables S1 and S2, which are not mentioned anywhere in the text but would be a useful reference in the discussion section where the authors discuss these hydrological signatures for all models.

*R: Thank you for pointing out that we did miss linking the supplementary material to the text. In the revised paper, we will refer to the additional supplementary materials to improve the description of the results and the discussion.*

- The abstract needs to be improved by including some of the nice statistics and results from the paper that quantify the improvements. Around L23, the authors talk about the hydrological signatures and that using both daily and monthly streamflow is better than just using the daily flows. However, it is not clear by how much.

*R: Thanks for your suggestion. In the revised paper, we will include more details about the improvements of the model configurations already in the abstract, and we will show the comparison of model performance using a %-deviation from the model M1 calibrated with daily streamflow as a benchmark. The latter metric, along with others added to the KGE used for calibration, will be reported already in the abstract and throughout the text, together with a new table.*

- The authors do not specify which model configuration is the baseline (which I assume is M1). Furthermore, while they present performance statistics, it is unclear if these differences are statistically significant to merit the additional data. Moreover, when they discuss the time-series analysis and the differences between the models, they do so in

a descriptive manner to quantify it better. For instance, using a distance metric to evaluate series similarity to the observed data. See DOI: 10.1016/j.rse.2011.06.020 for a summary of some useful metrics. My suggestion would be to make plots of the Mahalanobis distance rather than presenting the original time series (or in addition to Fig. 8).

*R: Thank you for raising this important point. In the revised paper, we will clarify the statistical improvement in the model performance with respect to the baseline (M1) calibrated with daily streamflow data, which usually is common practice. The %-deviation from the benchmark for all models will be used together with comparative metrics such as XXXX in a new table. With respect to your final suggestion, we consider that the distance-metric plots used in Mahalanobis (2011) are interesting but rather hard to interpret for hydrological time series; instead, we will modify Figure 8 to better explain the differences between model configurations.*

- I believe that the first objective should be merged into the other objectives. Running the model (independently of the computer language used) is a trivial objective as it is met from the start of the project.

*R: Thank you for this suggestion. In the revised version, we will merge the first two objectives into a single objective.*

- The authors need to explain how they did the catchment extraction in GRASS by providing additional detail into the used parameters. Also, they need to explain the IDW method in the methods section, define the acronym, and add a reference.

*R: Thank you for your point. Only to clarify, SAGA GIS was used to delineate the catchment boundaries, and GRASS GIS was used to address the space-time series of precipitation and temperature. In the revised version, we will describe the parameters used in SAGA, as well as the description of IDW interpolation.*

- The authors need to improve Fig. 3; interpreting it is confusing. Perhaps it would be best to have it with 4 rows rather than arrows, even if there is a degree of repetition.

*R: We agree and will add a row for model configuration in Figure 3 to clarify the different model configurations in the revised paper.*

- Can the authors modify the presentation of the 86 parameters in L331? It is hard to understand; I would suggest presenting the numbers in parenthesis as the main parameter numbers and then elaborating on how many were linked to soil types, land

cover, etc.

*R: Thanks for your suggestion. In the revised version, we will clarify the presentation of the model parameters.*

- Can the authors add box plots of the other statistics as supplementary? It is hard to visualize them as isolated numbers. Again, can the authors perform tests of significance on the statistics to determine a significant difference between them?

*R: We agree. In the revised version, we will change Table 1 for a boxplot in the supplementary materials section.*

- Can the authors mention what the criteria for defining a KGE of 0.5 as acceptable were?

*R: Thank you for raising this important point. In the revised paper, we will implement a clearer description and discussion on this issue based on recent literature. We will also introduce a clear description of the other performance metrics (KGE, Pearson Correlation Coefficient, MAE, NSE) used for comparison purposes in the supplementary material.*

- Around L605, the authors mention that the corrected temperature improved model performance. The authors need to quantify this performance increase.

*R: Thank you for this suggestion. We will state the performance obtained with the original temperature time series and the comparison using the corrected temperature.*

- The authors mention that the streamflow overestimation can be related to a precipitation bias in CHIRPSc. However, from Fig. S1, this does not seem to be the case.

*R: Thank you for raising this important point. We found that precipitation overestimation persisted in drier environments despite the bias correction. The overestimate was associated with the lack of ground precipitation records to correct the CHIRPS product in headwater catchments such as Rancho Rey. Our Fig. S1 in the supplementary section shows that, in many cases, differences between the water balance fluxes (P, ET, Q) were reduced. In the revised manuscript, we will clarify this point to highlight the issues associated with precipitation bias correction.*

- When discussing model improvement, please quantify it. The authors mention in L635 that M3 and M4 showed better and more realistic results but failed to quantify the improvement. Moreover, from Fig. 10, it seems that even though KGE was higher for M3 and M4, M1 was able to reproduce the actual spatial distributions of PET and AET better, overlapping more with the observed ranges.

*R: That is indeed an interesting point, and we thank you for the comment. Only to clarify, M1 showed high performance for PET but a lower performance for ET in comparison with M3 and M4 (shown in Figure 5). In the revised paper, we will include a statistical comparison between model performances to clarify the improvements between model configurations.*

- In the discussion section, the authors mention that adding PET and AET to the calibration improved model representativeness and link earlier studies. The authors

need to also link this assertion to their study, which is one of their objectives.

*R: Thank you for this suggestion. We will include statements and refer to the performance improvements in the text of the reviewed manuscript.*

- Due to missing tests, I do not see how the authors can conclude that M3 and M4 are better configurations since the statistical significance of the differences has not been evaluated. And in fact, for a lot of the variables, it seemed that M1 performed adequately well compared to M3 and M4. The authors can further support the increased accuracy of M3 and M4 by their link to the FDC information.

*R: Thank you for your suggestion. In the reviewed paper, we will include a statistical comparison using the %-deviation from the model calibrated with respect to the baseline (M1), as well as adding the corresponding metrics of FDCs in Figure 9.*

- Finally, I suggest adding ": A case study in Costa Rica." to the title since it was the only region analyzed in the manuscript.

*R: We agree with this statement that Costa Rica is a limited geographical space and will change the title of the revised paper to "Remote sensing-aided large-scale rainfall-runoff modelling in the tropics of Costa Rica".*

Around L54, the authors mention the opportunities from including additional variables. Please, specify which variables or give a few examples.

*R: We agree. In the revised paper, we will include some examples that were found useful for model calibration in the recent literature and how this could be applied to Costa Rica.*

### **Technical corrections:**

Around L56, the authors mention that more realistic hydrological partitioning comes at the expense of increased computational cost. Can the authors quantify the time penalties involved?

*R: Thanks for your comment. In the revised manuscript, we will clarify that including additional variables to provide a more realistic hydrological partitioning is more related to an increase of complexity for model calibration and validation instead of an increase in the computational cost.*

Around L77, do the authors mean simple bucket models? Any model can be a black-box model.

*R: We agree and will modify this sentence accordingly in the revised paper.*

Around L87, the authors mention that the coarse spatial resolution of the climatological data is an important source of error. Can the authors mention the related uncertainty in the data? (i.e., how much of the model error is associated with the coarse spatial resolution).

*R: That's an interesting point. In our research, we did not compare different products of precipitation and temperature with different spatial resolutions; however, in the revised paper, we will mention the errors found in previous research for Costa Rica and elsewhere.*

L135, the authors mentioned that they merged land covers. Can the authors include how much each merged class contributed to the overall classification?

*R: In the revised paper, we will include a short description of the percent of major classes merged in the land cover classification.*

L159, please remind the reader what both sides are.

*R: Thanks for your suggestion, this will be corrected in the revised paper.*

L175, please quantify the statement; how well did MODIS compare with the ground data? State  $r^2$  or another statistic.

*R: Thanks for your comment. In the revised paper, we will include metrics of MODIS performance from previous research.*

L209, can the authors mention how they chose the soil layer thickness?

*R: Thanks for your comment. Due to the lack of accurate soil thickness maps in Costa Rica, we considered a maximum soil thickness of 3 m in forest land cover and a minimum of 2 m in bare land cover, following the recommendation in the model configuration shown by Arheimer et al. (2020). In the revised paper, we will clearly state these boundary conditions.*

*Arheimer, B., Pimentel, R., Isberg, K., Crochemore, L., Andersson, J. C. M., Hasan, A., and Pineda, L.: Global catchment modelling using World-Wide HYPE (WWH), open data and stepwise parameter estimation. Hydrology and Earth System Sciences Discussions, 1–34. <https://doi.org/10.5194/hess-24-535-2020>, 2020.*

L245, the correction factor appears as B in equation one; it appears as BF and BF2 in equations 2 and 3.

*R: Thanks for your detailed review. In the revised paper, we will correct this error in equation 1.*

L266, from the text, it is somewhat ambiguous if  $y$  refers to each year or the whole period.

*R: Thanks for your comment. We agree and will clarify the equations and text in the revised paper.*

L288, the authors should mention that the parameters for correction are part of a monte Carlo simulation and are set to the ranges in Table 3.

*R: Thanks for your suggestion. Will be included in the revised paper.*

L290, sine function.

*R: We will change this error in the revised paper, thank you.*

L320, can the authors justify why only two years were used as a warm-up?

*R: Thanks for your suggestion. The warm-up period was established as two years before the calibration period in order to avoid issues with initial conditions of water content in soil layers, rivers, and reservoirs. We tested from 1 to 3 years and results did not change from two to three years. In the revised paper, we will include some references to the recommended warm-up periods, and we will clarify our selection of two years.*

L361, please remind the reader which time series.

*R: Thanks. In the revised paper, we will include an explanation of which series are compared.*

L375-379, this information should appear in the introduction.

*R: Thanks for your suggestion. In the revised paper, we will move this sentence to the introduction section.*

L405, can the authors normalize the MAE by the mean precipitation? Doing so would help the reader to understand the relative magnitude of the MAE.

*R: Thanks for this great suggestion. In the revised paper, we will include the normalized MAE in addition to the current value.*

L415, please, specify how they affect the performance.

*R: Thanks for your suggestion. We describe the impact of precipitation issues in different sections, such as L536, where we highlight that higher model performance was obtained in catchments with high cross-correlation on the Pacific slope. In the revised paper, we will link the description section with a sentence in L415 to clarify such an affirmation.*

L532, from Fig. 9, it seems that all the models underestimated the real flows to some extent. Is this due to CHIRPSc?

*R: Thanks for your comment. As you mention, precipitation from CHIRPS is an important factor of error in our results. In the discussion section, we explained how catchments from the Pacific slope showed higher performance in comparison with catchments from the Caribbean slope, related to the performance of CHIRPS to detect rainy and dry years on both slopes. In the revised version, we will clarify the limitation of low flow simulation and the relationship to precipitation performance, and we will also refer to Fig. S2 from supplementary.*

L575, can the authors increase the border thickness of the catchments of Fig. 10? It isn't easy to see them.

*R: Thanks for your suggestion. In the revision, we will increase the line thickness to emphasize the borders.*

L619, is the deviation a positive or negative bias?

*R: Thanks for your suggestion. In the revision, we will include such information.*

L644, what do the authors mean by increased parameter sensitivity?

*R: Thanks for your comment. We agree that this is a confusing message and will change this sentence to clarify the implications of MODIS PET and ET on model parameters.*

L650, Can the authors comment why none of the models at the best performing catchment could reproduce the decrease in water content between 2014-2015?

*R: Thanks for your comment. We assume that do you refer to the water content of Rancho Ray catchment in Figure 8. In this case, Rancho Ray showed the poorest performance of all catchments evaluated, which we will better highlight. The lower capacity to reproduce soil water content by most model configurations is related to the precipitation overestimation that stores water in the soil buckets.*