

**Authors' responses to interactive comment on Paiva et al. "Assimilating in situ and radar altimetry data into a large-scale hydrologic-hydrodynamic model for streamflow forecast in the Amazon" by Anonymous Referee #2
Hydrol. Earth Syst. Sci. Discuss., 10, C1374–C1376, 2013.**

The authors are please to respond to the comments and suggestions by Reviewer in the following text, in which Reviewers' comments are shown in **bold** typeface, and the authors' replies in *italic*.

Authors' responses to comments by Anonymous Referee #2.

Reviewer's general comment:

This study describes the application of a data assimilation scheme in a hydrologic hydrodynamic model that ingests in-situ and remote sensing observations of discharge and water levels in the Amazon River. The topic is very interesting and appropriate for HESS, while the methodology is generally sound. However, there were certain issues that either need to be clarified or be redone.

***Authors' response:** The authors are grateful for the Reviewer's comments. We have made our best efforts to address all corrections suggested*

Reviewer's specific comment: My major concern is with the calibration and validation periods. From the text it seems the model was calibrated using data from the entire period 1998-2009, while the validation period was 2004-2005. Why were the validation years included in the calibration period? How would the results change if these years were left out? I would recommend that the analysis be redone, at least for one of the experiment setups. Also, in section 3.3 it is not entirely clear which stations were used for either calibration or validation. During assimilation, no rationale is given to which stations were selected as observations to be assimilated, and which were kept for validation. Were experiments with different subsets performed? Either some justification for the selection of stations needs to be given, or more experiments need to be performed.

***Authors' response:** We do not intend to validate the calibration of model parameters. We used a model setup for the Amazon that was previously developed and evaluated in Paiva et al [2013], but that still contains errors, as expected. Considering that this model represents a reference, we intended to evaluate if how the data assimilation procedure can improve model estimates. Notice that at this point, data was used to correct model states and parameters were untouched. Also, we choose the period 2004-2005 because it has good data availability (both in situ and altimetry) that could be used for both assimilation and validation of the DA scheme. Also, it contains an important extreme hydrological event – the 2005 drought.*

To make it clearer, we included the following sentence at the end of section 3.6:

"The period 2004-2005 was chosen for its highest availability of in situ discharge and stage data together with altimetry data, and also because it contains an important extreme hydrological event – the 2005 drought."

The DA scheme was evaluated at the sites used for assimilation but also using independent data for validation. This issue is explained in section "3.5. Data

assimilation experiments” for each of the experiments. In the case of experiments 2 and 3, all altimetry data were used for assimilation and independent datasets (in situ data) were used for verification. In the case of experiments 1, only part of the in situ gauges was selected for assimilation or verification. Experiments with different subsets were not performed, considering the high computational costs required. In Exp. 1a and 1c, we choose these stations randomly. In Exp. 1b, we choose 12 stations covering the main amazon tributaries. We now added a few words in section 3.5 to make it clearer.

Reviewer’s specific comment: It would be helpful to add some quantitative information in the abstract.

Authors’ response: Done.

Reviewer’s specific comment: The paper needs to be proofread, there were a number of grammatical errors, e.g. “and discharge with minor degree” should probably be “and discharge to a minor degree” in the abstract.

Authors’ response: Done

Reviewer’s specific comment: p. 2882 “spatial resolution”: I think the authors mean “coverage”.

Authors’ response: Corrected

Reviewer’s specific comment: p. 2887 “the forecast errors”: these are not usually termed forecast errors, rather they are “model innovations” (actual minus predicted measurement).

Authors’ response: Corrected

Reviewer’s specific comment: p. 2888 “In [the EnKF] method... model states are perturbed”: the perturbation of model states is not a requirement, there are a number of different methods to generate an ensemble. In fact, the authors perturb the forcings to generate the ensemble of model states. I suggest replacing this with “ensembles of model states (and in some cases observations) are generated”.

Authors’ response: We corrected this sentence to read as follows:

“In this method, ensembles of model states and/or observations are generated using a priori-known errors and by means of the model operator M , the algorithm generates an ensemble of model trajectories from which the time evolution of model errors and error covariance matrix can be sampled.”

Reviewer’s specific comment: p. 2888 “of P are not required”: do the authors mean the analysis model error covariance P_a ?

Authors’ response: We mean forecast model error covariance P_f . We corrected it to make it clear.

Reviewer’s specific comment: p. 2890, l. 4-5: why are observations expressed as perturbed values? Since the square root version of the EnKF is used,

observations don't need to be perturbed? Were they perturbed prior to assimilation?

Authors' response: *We agree that there is no need to explicitly corrupt observation values when using the square root EnKF. However, it is necessary to know about the observation errors to compute the covariance matrix \mathbf{R} . We corrected this section, changing the term "corrupted values of Q and z " to "observations of Q and z ."*

Reviewer's specific comment: p. 2893 "results were improved in the second experiment": so far, there has been no mention of experiments. Which one is the "second experiment"?

Authors' response: *This sentence refers to the second experiment performed by Getirana and Peters-Lidard [2012]. We rewrote these sentences to make it clear:*

"A second experiment was performed by Getirana and Peters-Lidard [2012] using observed discharges at gauge stations to force the GFR scheme at downstream reaches."

"We assimilated data only from the 287 ASs located downstream of a gauging station where the accuracy of discharge estimates from Getirana and Peters-Lidard [2012] was generally better"

Reviewer's specific comment: p. 2894, l. 11: I don't think the word "sensibility" can be used in this context, probably meant "sensitivity"?

Authors' response: *Corrected.*

Reviewer's specific comment: p. 2895, l. 10-14: how were 12 ensemble members used from years up to 2004? Were years from the future used as well? Was the actual forecast year used?

Authors' response: *We used all data available between 1998 and 2009, assuming that if this forecast system/method were used today, that would be the information available. Then we tested the method for a period in the past, where a large amount of discharge data is available for checking the forecasts.*

We added the following sentence at this section:

"Aiming at exploring the usefulness of such system to provide streamflow forecasts in future applications, we chose to test it for a past period (2004-2005) where a large amount of discharge data is available for verification."

Reviewer's specific comment: p. 2896, l. 20-25: are there any rain gauges available to test the result for the optimal precipitation error? It would be valuable to have an idea of whether the 50% value is close to the actual precipitation errors.

Authors' response: *We now compare the 50% value with a global map of uncertainties in satellite-based precipitation measurements developed by Tian and Peters-Lidard [2010]. We rewrote lines 5-15 from pg 2897 to read as follows:*

"Based on the sensitivity tests, we used the following new parameter values for the further experiments: $N = 200$ (unchanged), $E = 50\%$, $\tau_x = 1.5^\circ$ and $\tau_t = 10$ days (unchanged). However, it is noteworthy that these parameter values related to precipitation errors, although providing better results for data assimilation, may not

realistically represent errors in the TRMM Merge dataset. For example, Tian and Peters-Lidard [2010] developed a global map of errors in daily satellite-based precipitation estimates that show features not represented here: errors that are spatially and seasonally variable and that decrease with precipitation rate (from ~100% at 1mm/d to ~20% at 100 mm/d in South America). On the other hand, the DA scheme had better performance with $E = 50\%$, which is only slightly larger than the values found by Tian and Peters-Lidard [2010] for the Amazon region, that ranges from ~ 20% to ~70%. A possible explanation is the consideration of that model uncertainty coming from precipitation errors and neglecting other sources such as parameter and model structural errors [Liu and Gupta, 2007], making E larger. Therefore, and since the first guess values were not fully justified in the previous studies [Andreadis and Lettenmaier, 2006; Clark et al., 2008], we preferred to use the parameter values where the DA scheme performs better.”

Reviewer’s specific comment: p. 2901, l. 1-6: I realize the dataset is referenced out, but it would be worth mentioning if the altimetry-derived discharge contains any in-situ information or is just derived from a model and the actual ENVISAT data.

Authors’ response: *These details are all given in Section 3.3.*

Reviewer’s specific comment: p. 2901, l. 18-20: were these stations used in the assimilation?

Authors’ response: *Yes. We included the following sentence at this part of the manuscript to make it clear:*

“Notice that these gauges were used by the DA scheme.”