First, we would like to thank the reviewers for having carefully read the paper and provided valuable comments which helped us to improve the quality of the manuscript. We have taken into consideration all the comments raised by the reviewers and changed the manuscript accordingly. The details of our changes are highlighted in the main text. The point-by-point answers to Reviewers #1 and #2 are provided in the following and highlighted in red.

RC1:'Comment on hess-2023-214', Anonymous Referee #1, 09 Oct 2023

General Comments

The authors tackle an important topic, namely how to best combine gauge and satellite precipitation estimates for applications. The hydrological validation they pursue is a reasonable way of testing how the various input datasets perform, and this is clearly the strong point of the manuscript. As such, the existing examples and conclusions related to hydrology seem solid.

The other big point of the manuscript, as the title makes clear, is the combination scheme that employs CA interpolation, but here the manuscript falls short. I would expect a step-by-step demonstration that all the extra mathematical complication produces precipitation fields that are physically meaningful and more consistent with the input fields than some simpler scheme. I would consider it mandatory to address this issue.

The manuscript has been revised. A new figure 4 is added to show the physical consistency of the combination scheme that employs CA.

The English is not ready for publication; I have commented on a few word choices that I found confusing or that might escape a general technical editor, but not otherwise. Other issues needing attention are listed below.

If accepted for publication, we will make use of the English editing service of the journal.

Specific Comments

1. ***CA and regularly gridded data:*** *It is not clear to me how CA handles the regularly gridded data. Does it assign each satellite gridbox value to the finer gridbox closest to that satellite gridbox's center? Conventionally, gridded data give you an average value across the entire box, so if you just assign that value to the box's center, doing an interpolation and then averaging that field back to the original resolution will not give you the right average, in general.*

We Thanks the reviewer for her/his comment. The technique implemented does not pose any constrain on the final interpolated field (i.e., averaging the output field back to the original resolution will not likely give the initial average value exactly). Indeed, what we did is to redistribute spatially the rain information from multiple sources, namely, rain gauges and satellite retrievals. For the former, the measured precipitation is assigned to the closest grid point in our destination grid and then interpolated over the neighborhood grid points using Cressman inverse distance interpolator with a ray of influence of 5 km. From the satellite standpoint, the rainfall values in each Field Of View (FOV) of the satellite are first assigned to the respective FOV center. Secondly, each FOV center is assigned to the closest grind point in the destination grid as done for the rain gauges. Thirdly, each satellite rainfall value thus obtained, is interpolated with Cressman inverse distance with a ray of influence equal to the native FOV size. So doing, the areal information from satellite is spatially redistributed with the inverse distance criteria. Although this can sound a little bit arbitrary because we do not really know how precipitation is distributed within each satellite FOV, the methodology that we implemented has the advantage to treat both rain-gauges and satellite equally in terms of the processing applied. In addition, the approach proposed follows a source-hierarchy, that means that rain gauges are interpolated first and then satellite comes in to play for gap filling only. Eventually the Cellular Automa has the role to make the final reconstructed rainfall filed spatially consistent. Note that the goal of the manuscript is not to propose a new multi-source merging strategy, but instead to demonstrate how the use of multiple sources can improve the hydrological output.

Specifically, locally convex-up areas will be underestimated (particularly sharp peaks), concave-up areas will be overestimated, and the boundary of the precipitating region will spread somewhat into the non-raining area. I consider discussion of these issues to be mandatory.

Certainly, locally convex-up areas will be underestimated (particularly sharp peaks), concave-up areas will be overestimated, and the boundary of the precipitating region will spread somewhat into the non-raining area. However, our results indicates that the distribution of input precipitation field with smoothed peaks and minima but with a more homogeneous spread, improves the estimate of the flow discharge and peak timing in the hydrological model output.

2. *What is the model grid spacing:* Section 4.3 seems closest to stating this, but I missed seeing a declarative statement giving the specific value of the grid spacing. If you're really down at "hundreds of meters", it needs to be made clear that the effective resolution of the precipitation data is back at 5, 10, or more km. Finer than that is just more and more precisely defining smooth variations between the available value locations.

A detailed comment has been inserted on lines L290-293. The hydrological spatial resolution is approximately 900 m, the same as used for the rainfall field. The approach was explained in the previous comment.

3. *Example interpolated precipitation fields:* Given the emphasis on the innovative combination of data sources, I would expect to see a sequence of maps illustrating the process and improvement versus simpler "traditional" schemes. This should happen first, before pointing to the accumulated precipitation in Fig. 5 and the aggregate hydrological results in Fig. 6 (for example). This aspect of the manuscript is where the issues in item 1 need to be addressed.

Thanks for this tip, this is a great idea for future work. Note that the priority of this study is to demonstrate how the use of combined data improves the hydrological model's output. Probably, this can be achieved even using "traditional" data merging schemes. This is the reason why comparisons with other assimilation techniques are not implemented. Figures 5 and 6 are useful to demonstrate that the different data sources give different rainfall fields to the point of improving or worsening the results of the hydrological simulation.

4. *Boundary bias:* This phrase comes up several times, but it was never quite defined as to what boundary was being discussed. It is stated that data outside the basin is used, which presumably should solve a problem at the basin boundary. Perhaps the problem is along the southwest side of the basin, where no additional gauges are shown in Fig. 1. The statement in L.417-418 should appear a lot sooner in the text and be more explicit about how this works.

In general, boundary bias denotes a systematic error or distortion occurring in data or models near the periphery of a study area or dataset. This bias often arises due to differing conditions or factors at the edges of the study area compared to its interior. The presence of boundary bias can impact the accuracy and reliability of analyses and interpretations, underscoring the importance of acknowledging and addressing this bias when working with spatial data or models to ensure more precise and meaningful outcomes.

In the specific context of this study, boundary bias manifests as a discontinuity associated with the absence of observed data beyond the confines of the basin. To mitigate this potential issue, whenever feasible, all available data across the entire domain is utilized. Nevertheless, a discontinuity persists since no observed data beyond the Italian border are accessible. This aspect will be further discussed later in the paper.

To further elucidate this concept, the following sentence was incorporated into lines L229-234: "Furthermore, a strategy used by the work to avoid boundary effects is to extend the spatial domain well beyond the studied basin: this strategy is useful for a better reconstruction of the precipitation field (Figure 1). Many data used, although redundant, lead to a better reconstruction of the rain field. A smaller amount of this data would probably be enough, but the work uses everything that the national rain gauge network has available. Future studies could lead to identifying, given their distribution, enough rain gauges outside the basin deemed useful to overcome the boundary effect."

5. *Fig. 1: I question whether all those gauges outside the basin are really useful and therefore worth depicting. I would suggest that as you move outside the boundary you can stop after you pass about 3 gauges (which of course varies with coverage). Was it really not possible to obtain gauge data to the southwest of the basin? This introduces the boundary effects the manuscript discusses (right?).*

   The aim of this work is to demonstrate how to validate an operational chain for civil protection monitoring and forecast purposes. The rain gauge network that was used is the official national one. Using all the data at our disposal, beyond a defined number (3 or 4) outside the basin for example, helps the model to better distribute the rainfall field, it is certainly not a limit, but rather an added value. A possible future work could be to estimate the optimal number of rain gauges, based on their distribution.

   So, the idea is to use all the rain gauge data available for a possible operational activity, which does not include those outside the national border. These limitations help our validation: despite the many critical issues the model seems to respond very well.

6. *Fig. 7 caption:* The statements after the first sentence are interpretation and belong in the text.
   The sentence has been deleted.

Technical Corrections

1. *L.180, 720:* "O" is actually that author's last name; "Sungmin" is her first name. Done. Thanks for your suggestions.

2. *Nine occurrences:* "IMERG" is mis-stated as "IMERGE". Done. Thanks for your suggestions.

3. *L.52:* Not sure "captative" is the right word. Done. Thanks for your suggestions.

4. *L.88-89:* Awkward phrasing, including that "peculiarity" is probably something like "availability". Done. Thanks for your suggestions.

5. *L.168 and 3 other locations:* I'm not sure what "rain bandwidth" means.

   Shi et al. (2020): "any region has an effective influence (ERI) radius for any rainfall event, which reflects the influence of a rain bandwidth, and such an EIR is not larger than a certain distance". The reference to L269 is wrong.

6. *L.186:* Should this be "...that are usually not instrumented."? Done. Thanks for your suggestions.

7. *L.193:* Fig. 2 refers to a "workflow", not "rationale", which seems better. Done. Thanks for your suggestions.

8. *L.193:* Fig. 2 says three tasks, not four. Done. Thanks for your suggestions.

9. *L.202*: Unclear phrasing; are you saying something like "the model is not calibrated specifically for this study's cases"? The reviewer is right; this sentence is misleading. It was meant that the model was calibrated in past studies on the Po basin of which the Tanaro River is a tributary (Coppola et al 2014) and not for this study. To avoid confusion, it has been deleted from the text.

10. *L.428:* I think "intuitive" could better be "subjective". Done. Now Line 446

11. *Fig. 1:* a) What are the blue lines? b) The hydrometers are nearly invisible; maybe they should be plotted in white? Done. The description of the blue lines has been added in the caption. The figure has been replaced.

12. *Fig. 2 caption:* The phrasing is awkward, perhaps something like "... workflow, consisting of three main tasks:" Done. Thanks for your suggestions.

13. *Fig. 3*: The mostly-dark colors make it really hard to distinguish the basin and gauge coverages. I'd say the Google Earth background and the basin's blue need to be much lighter. Done. A new figure is available.

RC2: 'Comment on hess-2023-214', Anonymous Referee #2, 11 Oct 2023

The paper of Lombardi et al. deals with an interesting topic concerning the reliability and usability of satellite rainfall products for hydrological applications. To this aim, they consider a case study in northern Italy (Tanaro watershed) and test different precipitation fields achieved using only rain gauges, satellites, or several merging options. Precipitation fields are used as inputs for a hydrological model. According to the behaviour of the resulting hydrograph, the precipitation fields are judged in terms of reliability, supporting the discussion with several (very many, indeed) skill scores.

Despite the interest in the topic, I believe several changes in the paper are needed before it can be considered ready for publication in HESS.

First, I have concerns about the methodology followed. As a first step, since the main point is the reliability of the satellite rainfall products in providing quantitative precipitation estimates, I would have set up a validation exercise with observed (rain gauge) precipitation using, e.g., a leave-one-out or cross-validation approach.

Thanks for the comment, it was intentionally decided not to use this approach, relying exclusively on an indirect method of validation through a hydrological evaluation. The work does not aim to do a local validation but a distributed one, thanks to the analysis of data relating to the drainage network.

Then, an indirect validation through simulated hydrographs should be justified more strongly, mainly if performed with a not-calibrated hydrological model (L202). This approach could be tricky and misleading because of the inherent limitations of the non-calibrated hydrological model so that errors can counterbalance and smooth each other. I suggest a preliminary

calibration of the hydrological model with a reference precipitation field (e.g., only rain gauges) and only afterwards assess the changes caused to the modelled hydrograph with other methods.

The reviewer is right; this sentence is misleading. It was meant that the model was calibrated in past studies on the Po basin of which the Tanaro River is a tributary (Coppola et al 2014) and not for this study. To avoid confusion, it has been deleted from the text.

Another concern is the description of the CA-based technique for interpolating and merging precipitation, which is unclear to me. An example should be given. Among other things, it is unclear why the authors also consider the time evolution of the precipitation field.

Regarding the technique adopted in detail, the review can refer to Coppola et al, 2010. Effectively, a detailed description of the CA algorithm is missing in the manuscript and the following section has been added.

The temporal evolution of the rainfall field is fundamental for the hydrological simulation, in particular the CHyM model is forced with the hourly rainfall fields.

Last but not least, the paper could be structured much better.

The reviewer is right; the document has been rearranged.

The different options for achieving the precipitation fields are not clearly presented (e.g., some acronyms are provided in Fig. 2, then they are explained much later -L400- and not in the Methods, but in the Results section).

A description table has been added (table 2)

It is unclear why the authors decided to rely on 17 scores. This choice is somewhat confusing, in my opinion.

The 17 scores are necessary to obtain an objective evaluation of the analysis, given that each score highlights different characteristics. Obviously, if they are all positive, the results are confirmed as reliable.

Furthermore, many of the methods are presented in the results section or even in the conclusions (please refer to specific comments below). The discussion should refer to similar analyses performed by other authors to contextualise the results better. The conclusions section should be more than a summary of the paper.

In summary, I saw the possible added value that this paper can bring to the scientific community. Still, a thorough review is needed regarding the methodological approach and the structuring of the article. Please find below some other minor to moderate comments. I hope my review helps improve the quality of the paper.

1. *LL46-49: please revise. Dealing with predicted rainfall, it's impossible to remove uncertainty (correctly, in fact, the second sentence of the paragraph refers only to observed rainfall data). I suggest focusing on why accurate spatial distribution of rainfall observation is important.*

   To avoid confusion, the sentence has been corrected as follows: "As far as the operational activity is concerned, the hydrological models are usually forced both with observed and forecasted rainfall data, and the uncertainty of hydrological forecasts is strongly related to the uncertainty of the input rain field. Therefore, providing hydrological models with observed precipitation data that is as realistic as possible becomes essential in mitigating uncertainty, during the spin-up phase of the simulation when the hydrological model is forced with observed rainfall data."

2. *L104: "The work": do the authors refer to Shi et al. (2020)?*

   Yes. The sentence has been corrected.

3. *LL111-112: I can't entirely agree with this statement. Indeed, there are a lot of studies dealing with this topic.*

   Done. The sentence has been deleted.

4. *LL118-119 is a repetition of the second main objective declared at LL115-116. If the authors agree with my comment, please consider if the previous sentence (LL117-118) is well-placed and contextualised.*

   The reviewer is correct; the second sentence LL117-118 has been deleted.

5. *L120: are all these 352 stations really useful? I guess the authors only need those lying into or close to the analysed watershed. From this point of view, it's unclear why the authors consider a much broader spatial domain than the investigated watershed (which, moreover, is not at the centre of the domain itself). I guess many stations, for example, lying in the north and northwest, are useless for this case study.*

   The objective of this study is to validate a prospective operational framework designed for civil protection monitoring and forecasting. The rain gauge network employed in this research is the official national network. Utilizing all available data enhances the model's ability to effectively distribute the rainfall field, representing not a limitation but rather an added value. As rightly noted by the reviewer, the inclusion of 352 rain gauge data may be excessive, especially considering that only 73 of these data points fall within the specific basin under consideration. At the state of the art, a study aimed at understanding how much rain gauge data is sufficient to have the same results has not yet been carried out A possible future work could be to estimate the optimal number of rain gauges, based on their distribution.

It is crucial to acknowledge that the algorithm employed in this study applies smoothing to the processed data. Therefore, for scientific rigor, the entire operational domain is utilized, acknowledging both its advantages and limitations. One such limitation is the absence of rain gauge data beyond national borders, which could potentially impact the results. The primary focus of this work is to ascertain whether, despite these constraints, the model's performance improves when incorporating satellite data using this technique. Further investigations could delve into estimating the optimal number of rain gauges based on their distribution, representing a potential avenue for future research.

6. *L146: 1700 m3/s is a peak flow, average daily flow or what else? Some lines below the authors refer to a peak of 4350 m3/s (if "Autorità di Bacino del Fiume Po" is a reference, please add the year; if not, please explain/translate it in English).*

- A maximum flow discharge that can reach 1700 m^3/s in spring and autumn.
- The most significant of these events occurred in November 1994, when the entire river valley was damaged (Marchi et al., 1996; Luino, 2002) and the sensor at Montecastello, located at the outlet of the river recorded a maximum flow discharge peak of 4350 m^3/s (Autorità di Bacino del Fiume Po).

"Autorità di Bacino del Fiume Po" is the official Italian authority that releases information regarding the data relating to the various rivers, the document is among the references and was last viewed in November 2023. "Autorità di Bacino del Fiume Po" has been translated into English as suggested by the reviewer.

7. *L 165: Eq. (1) is quite ambiguous. It is well known that the reference area is hardly a radius, especially in orographically complex regions such as the Alps.*

This equation refers to Shi et al. Estimating a radius of influence is certainly not so immediate, but it can be a good starting point for making various considerations.

8. *L176: Earth*. Done. it has been corrected.

9. *Figure 2 and related caption: please revise. There are several errors: e.g. Guage uncal, "each case studies" [study], "eight [...] setting" [settings]. Furthermore, the terms uncal, cal, uncal1, uncal5, etc. are explained much later (L400). The explanation of the different inputs for the eight simulations should be highlighted much better (maybe with a devoted Table?).*

As suggested by the reviewer, table 2 summarizing the different inputs has been added.

10. *L325: finds*. Done. Now, LL335

11. *LL360-364: I guess this is methods, not results.*

    Section 5.1 also is not results, but data and methods.

    L406: it's 5.2

Thanks for the advice, this section has been reorganized.

12. *LL418-419: however, while some of the stations considered in the study are located much further north of the watershed (as I claimed before, I believe they are useless to this study), after the French border, there are no stations surrounding it. This drawback should be discussed.*

The reviewer is right, as said before, this falls within the limits related to the selected domain, which we voluntarily wanted to maintain. In operational conditions we come across these types of problems, so the use of satellite data helps us overcome the problem, as demonstrated by the improving results.

13. *LL519-528: that's methodology.*

The reviewer is right, this is a repeat. the sentence has been deleted