We thank the reviewers for the time she/he spent in the detailed reading of our work and for recognizing our efforts attempting to improve our manuscript.

Report #1

General Comments

The authors have done a reasonable job of responding to the reviews, even if pursuing approaches that I would have constructed differently.

I will note to the editor that the authors intend to use the journal's editing services to improve the English.

We thank the reviewer, We believe that after this interaction, our work has further improved.

All the changes in the revised version of the manuscript are highlighted in red color. The lines indicated for the corrections refer to the file hess-2023-214-ATC2.pdf

Specific Comments

1. L.185-186: Flow regulation is mentioned at the top of Section 5, but it should also appear here as a major challenge.

We Agree. In section 3.3 we added the following sentence:

LL183-185: *"Furthermore, data on artificial water management are not available, therefore the hydrological model validation presents difficulties in the presence of highly regulated basins since the simulation reproduces the natural flow rate of the river flow without considering the anthropogenic impact."*

We modified the LL 370-372

2. L.278: "GMP" should be "GPM".

Thanks for the correction. We did not find this error; we have checked that all the acronyms and they are correct.

3. L.409-410,463: The calibrated IMERG-F does use *monthly* gauge data, which is what makes it "calibrated". If you are thinking about *submonthly* gauge data, you need to be explicit. This addition of (monthly) gauge data is why CAL performs better than UNCAL.

We thank the reviewer for pointing this out. The rain gauge data we refer to are the local rain gauges, and not the monthly data already used in IMERG-F Calibrated. We modified the sentence.

A new reference and a new detail have been added to the LL178-180

LL454-456: "*Therefore, UNCAL and CAL simulations use only satellite data respectively IMERG-F Uncalibrated and IMERG-F Calibrated, GAUGE simulation uses rain gauge data.*"

To:

"*Therefore, UNCAL and CAL simulations use only satellite data (respectively IMERG-F Uncalibrated and IMERG-F Calibrated), while the GAUGE simulation uses the local rain gauge data*."

Report #2

Although I am very well disposed toward the publication of the article, which I see as potentially valuable for the hydrological community, I am unfortunately dissatisfied with the reviewers' responses, which, on several points, seemed meagre and evasive.

As a preface, I strongly suggest that the replies to the reviewers refer clearly to the changes in the manuscript, indicating precisely the lines and sections changed.

Thanks again for the time you spent in the further revision of our work. In the following we give a precise point-by-point answer to the raised questions. We believe that after this interaction, our work has further improved.

All the changes in the revised version of the manuscript are highlighted in red color. The lines indicated for the corrections refer to the file hess-2023-214-ATC2.pdf

1. Concerning my first concern, if the authors are focused on the hydrological impact, I suggest highlighting it better both in the title (e.g., "on the combined use... to improve hydrological modelling") and the abstract.

   LL1-4: We changed the title in: "On the combined use of rain gauges and GPM satellite rainfall products for hydrological modelling: impact assessment of the cellular automata methodology on the Tanaro river basin in Italy."

   LL26-28: We modified the abstract as suggested adding the following sentence:

   "...the comparison between simulated and reference river flow discharge is crucial for assessing the effectiveness of merged precipitation data in enhancing the model's performance and its ability to realistically simulate hydrological processes."

2. More in general, I believe the authors do not consider the issue of equifinality, i.e., "the principle that in open systems a given end state can be reached by many potential means" (Wikipedia). That is related to my second concern, where I highlighted the possibility that errors can counterbalance and smooth each other. Then, much of the outcome could depend on the initial parameterization. I understand that the model was already calibrated over the whole (much larger) Po River Basin, not specifically on the Tanaro. I think this starting point weakens the paper's outcomes and should be made much more transparent.

   We thank the reviewer for this comment.

   With reference to the hydrological model, CHyM, it is worth noting that it has been widely calibrated using climatological discharge time series of the Po River, as reported in Coppola et al. (2014). To this aim, it is important to note that the conditions of the Po are representative of many alluvial rivers in Europe (Di Baldassarre et al., 2009).

   In any case, the domain used in this work which is the same one used in Coppola et al. 2014, is also the operational domain for hydrological forecasting for flood risk at CETEMPS (Centre of Excellence - University of L'Aquila). As evidenced in Figure 1 the domain does not include the entire Po basin, but only its northern part. Part of the data used in the calibration refers to a hydrometric station on the main branch of the Po River, close to the mouth of the Tanaro river, a tributary of the Po. For this reason, we believe that we can refer to that calibration, published in Coppola et al 2014, since it can be extended to the Tanaro basin without introducing significative biases.

Long simulations were carried out also for the present work to test the model configuration. Given that the purpose of the paper is different, we do not report these results.

LL342-344: This sentence is added: "*The hydrological model is not specifically calibrated over the Tarano basin. However, in this work we refer to the calibration accomplished by (Coppola at al 2014) on the northern part of the Po River, which totally includes the Tanaro basin*."

3. Concerning the third concern (description of the CA-based technique), I can't read about Coppola et al. 2010. I guess the authors refer to Coppola et al. 2007 (I also found Coppola et al. 2014, but it is not referenced).

Thank you for pointing this out. We apologize, the reference was not correct. The correct reference is Coppola et al 2007.
https://www.tandfonline.com/doi/epdf/10.1623/hysj.52.3.579?needAccess=true

4. In their reply, the authors claim that "the following section has been added", but I can't understand where this section is. Anyway, I still believe that some more detail should be added to make this paper more self-consistent.

We apologize for not indicating where the sections were added: the sentence was wrong, because no section had been added in this case.
What was done in Coppola et al 2007 is described in detail in sub-sections 4.1 and 4.2 of our paper. For this reason, the two sub-sections have been reformulated more clearly.

in detail, in section 4.1 you will find the following changes:

LL206-213: *"The precipitation data gridding is necessary to speed-up the numerical processing in the hydrological model, and it defines, on a regular grid, a first guess in terms of precipitation field at the hydrological scale, hereafter termed as Precipitation Background Field (hereafter PBF) (Coppola et al, 2007). The Cressman algorithm (Cressman, 1959) is used to initialize the rain field grid points in the selected domain. Because of its simplicity, the Cressman method can be a useful starting point (Bouttier and Courtier, 1999). According to Li and Shao (2010) the used kernel function determines the accuracy of the fused field and to define a kernel function, it is also necessary to select rain radius. The radius of influence, R, determines the smoothness of the estimated field, containing the spread of the kernel function: a small R corresponds to a rough estimated field and large variance, while a large R corresponds to a smooth surface."*

It was thus modified with the following sentence:

"In hydrological modeling, precipitation data gridding is essential to accelerate numerical processing. It involves creating an initial estimate of the precipitation field at the hydrological scale, known as the Precipitation Background Field (PBF) (Coppola et al., 2007), on a regular grid. The Cressman algorithm (Cressman, 1959) is commonly used to initialize rain field grid points within the designated domain. Due to its simplicity, the Cressman method serves as a practical starting point for this initialization process (Bouttier and Courtier, 1999). The accuracy of the merged field significantly depends on the choice of kernel function, as highlighted by Li and Shao (2010). Selecting an appropriate kernel function involves defining a rain radius. The radius of influence, denoted as R, plays a crucial role in determining the smoothness of the estimated field and controlling the spread of the kernel function: a smaller R results in a more rugged estimated field with higher variance, while a larger R yields a

smoother surface. Thus, the selection of the radius of influence is critical in determining the overall quality and characteristics of the estimated precipitation field in hydrological modeling applications."

LL231-239: *"Obviously, the first difficulty lies in selecting the reasonable value of R. Figure 3 shows the area coverage by rain gauge network, when the algorithm uses a radius of influence equal to 5km. Even if observed data were available for every grid point in the selected domain and no significant errors are found, the rainfall field rebuilt using a direct merging method as the Cressman objective analysis scheme (e.g., Pereira Filho et al., 1998; Goudenhoofdt and Delobbe, 2008) would produce significant bias at the boundary (Li and Shao, 2010; Duque-Gardeazábal et al., 2018); this means that a smaller value of R would lead to a bias, but in a smaller area around the boundary. However, R selection is not a remedy to the boundary bias because the rain bandwidth is likely to be large when observed points are distributed irregularly. The problem of boundary bias is caused by the discontinuity of the background field due to the discretization of the field, while the nonparametric merging method is only able to generate continuous surfaces."*

It was thus modified with the following sentence:

"Selecting a suitable value for R poses the initial challenge in the estimation process. Figure 3 illustrates the area coverage by the rain gauge network when employing a radius of influence, R, equal to 5 km. Even if observed data were accessible for every grid point in the selected domain without significant errors, employing a direct merging method such as the Cressman objective analysis scheme would result in considerable bias at the boundary (Li and Shao, 2010; Duque-Gardeazábal et al., 2018). This indicates that while a smaller value of R may mitigate bias, it would only affect a smaller area around the boundary. However, adjusting R doesn't fully address the issue of boundary bias, as the rain bandwidth tends to be large in cases where observed points are irregularly distributed. The boundary bias issue arises from the discontinuity of the background field due to field discretization, while nonparametric merging methods can only generate continuous surfaces."

The following changes are reported in the last document uploaded in subsection 4.2:

L260-264: *"CA can be described, for example, as identical discrete sites of a lattice, and the state of each grid point evolves according to deterministic rules, conditioned by the values of neighboring cells at discrete time steps.*
CA based algorithm has been developed and implemented in the hydrological model code. According to CA theory, the input grid is considered an aggregate of Cellular Automata, and the status of a grid point corresponds to the value of a rebuilt (i.e. smoothed) precipitation field."

It was thus modified with the following sentence:

"In CA, natural systems are idealized as discrete sites on a lattice, with each grid point evolving based on deterministic rules and influenced by the states of neighboring cells at discrete time steps. This approach provides a structured framework for dynamic systems modelling, reflecting the intricate interplay of elements in nature.
In the hydrological model code, a CA-based algorithm has been developed and implemented. Following CA theory, the input grid is conceptualized as an aggregate of Cellular Automata, where the status of each grid point represents the value of a smoothed precipitation field."
LL279-284:*" The coefficient α assumes a small value (typically from 0.1 to 0.9) to ensure a slight smoothing of the original matrix: all grid points are updated synchronously, and the smoothing is*

performed until the stability is reached, meaning that no significant changes are recorded in the calculated matrix. The grid point associated with the rainfall value available in the considered database is not modified by the algorithm. In terms of time evolution, a regular lattice is updated in discrete time steps according to the previous rule depending on the state of the site and the eight neighboring cells."

It was thus modified with the following sentence:

"The coefficient α assumes a small value, typically ranging from 0.1 to 0.9, ensuring a gentle smoothing of the original matrix. All grid points are updated synchronously, and the smoothing continues until stability is achieved, signifying minimal changes in the calculated matrix. Notably, the grid point associated with the rainfall value available in the considered database remains unaltered by the algorithm. This process enables the hydrological model to refine and stabilize the precipitation data while preserving the integrity of observed rainfall values."

LL295-297:*" The CA method allows to perform the assimilation and spatialization of the rain field, it is useful for the high resolutions necessary for hydrological simulations, and to use different sources of precipitation data."*

It was thus modified with the following sentence:

"The CA method facilitates the assimilation and spatialization of rainfall fields, proving advantageous in achieving the high resolutions required in hydrological simulations and for integrating several precipitation data sources."

5. Concerning the rearrangement of the paper's structure, I don't understand the main changes made from this point of view, and I ask the authors to present them in the next iteration with reviewers.

In the first manuscript uploaded 01 Sep 2023 the structure was:

1 Introduction

2 Study area

3 Observed Data

- 3.1 Rain Gauge data.
- 3.2 Satellite-based rainfall estimates
- 3.3 Observed Flow Discharge data.

4 Methodology

- 4.1 Precipitation data gridding
- 4.2 Precipitation data interpolation and merging: Cellular Automata technique
- 4.3 Hydrological modelling: CETEMPS Hydrological Model
- 4.4 Error Score Metrics

5 Analysis and discussion of the results (L359)

- 5.1 Analyzed case studies. (L366)
- 5.3 Results: hydrological simulation analysis (L406)

6 Conclusions (L495)

As suggested by the reviewer the structure of the paper was rearranged in the previous review.

In the version of the manuscript uploaded on 18 Dec 2023 the reviewer finds the last sections modified, specifically:

5 Analyzed case studies (L372)
6 Results: hydrological simulation analysis (L417)
7 Conclusions (L504)

6. Concerning the choice of using 17 scores, the authors' response is a kind of apodictic. Their very peculiar choice should be justified better. Very seldom I found that this high number of metrics was used in similar hydrological studies.

   Thanks for this remark. The reason why we chose to use 17 statistical scores is that each score highlights different, and somewhat complementary, characteristics of the simulation.

7. Specific comment no.5: in order to shed light on the issue of the number of stations needed for the Tanaro River Basin, performing some more tests by removing at least the rain gauges at the very north border of Italy could be enough. According to the authors' reply, one can conclude that also the absence of the French stations makes the work not "rigorous". Anyway, it is very odd that far-northern gauges are used while not available close western gauges are not.

   We reproduced the operational framework (domain) in which not all potential stations are effectively available. We could have selected only the river basin of interest, but we wanted to recreate the operational setting, with the same conditions and related critical issues, such as the lack of data on the French area, and to highlight that the use of satellite data is essential to overcome these critical issues.

8. Specific comment no.6: not clear yet. If the "maximum flow discharge [...] can reach 1700 m3/s, in spring and autumn", it cannot reach 4350 m3/s in November. Please rephrase.

   The reviewer is right. The 1700 m3/s of river flow discharge refers to standard seasonal conditions, while 4350 m3/s flow discharge refers to the worst condition recorded from 1801 to 2001.
   The sentence has been reformulated at LL145-147.