Universitat de les Illes Balears

Arnau Amengual
Grup de Meteorologia
Departament de Física
Ctra. Valldemossa km 7.5
07122 Palma, Mallorca, Spain
arnau.amengual@uib.es

Palma, 29 January 2024

Dr. Mario Parise
Scientific Editor, NHESS
mario.parise@uniba.it

Dear Dr. Parise,

A revised version of the manuscript NHESS-2023-130 entitled: "*Hydrometeorological controls and social response for the 22 October 2019 catastrophic flash flood in Catalonia, north-eastern Spain*" by A. Amengual, R. Romero, M.C. Llasat, A. Hermoso and M. Llasat-Botija is attached. In the next pages, the authors include a point-by-point response to the comments and concerns risen by the reviewers.

The authors express their gratitude to the scientific editor and the reviewers for their valuable comments, which significantly enhanced the quality of the revised version of this study. Specifically, the abstract and introduction now more clearly emphasize the novelty of the present research. The hydrological model has undergone evaluation through a comparison of numerical results with estimates of maximum peak discharges and times-to-peak. Additionally, the revised manuscript discusses sources of uncertainty, limitations and potential avenues for future research. A more detailed analysis of the social response has been incorporated, and a new discussion section has been added to facilitate comparisons with other case studies, offering insights into elements that may be of broader interest. Conclusions have been revised accordingly.

Sincerely,

Dr. Arnau Amengual

REVIEWER 1

General comments

The paper presents a detailed analysis of the meteorological, hydrological, and social dynamics of a flash flood that occurred in Catalonia (Spain) on October 22 2019, resulting in six fatalities. This very well documented interdisciplinary study combines observation and modelling to understand the meteorological and hydrological context that led to this exceptional event. The study also provides insight into the dynamics of the warning process, the social perception of this warning and of the social response to this event.

The study relies on a comprehensive data set and the modeling study (both atmospheric and hydrological modeling) are well documented. However, some results may be analyzed in greater depth (see specific comments below) and the analysis of the social response is rather qualitative.

Furthermore, the paper lacks a "discussion" section that could allow comparison with other case studies and provide information about the elements of the study that are of interest to a wider audience, beyond the single case study presented in the article. The conclusions and abstract should be revised accordingly.

Therefore I recommend major revision of the paper to allow the authors to address all the issues pointed out in this review.

The authors would like to thank the reviewer for the highly constructive comments made. The revised manuscript addresses all the concerns pointed out by the reviewer, including a new discussion section. The new discussion section compares the results of the present work with previous outcomes and provide information about the elements of the study that are of interest for a wider audience. The conclusions and the abstract have been revised accordingly.

In addition, a more detailed description of the analysis of the social response has been included. The revised version breaks down the social response into two main components, according to Creutin et al. (2009): management activities and human responses. The former includes three different types of actions: information, organization, and protection. The latter encompasses human responses within three groups: Individual, communal and institutional. Consequently, the title of new section 5.2.1 now incorporates the more precise terms "risk management" and "human response". A more quantitative analysis has been conducted based on information gathered during the FLOOD-UP FRANCOLÍ citizen campaign in the Conca de Barberà council. New Table 7 and Figure 12 describe in detail the type of actions and timeline followed during the course of the flash flood.

Specific comments

0/ section 3.1. The atmospheric modeling is based on the new version of a 3D non-hydrostatic model, developed by some of the co-authors, which has been tested and shown to perform as well as more widely used model. This model is use to assess the predictability of the event and the main driving factors of the event. As the authors study the warning process in their paper, it would be interesting to also compare the numerical simulation performed in the paper, with the results of the operational model used in 2019 during the warning process, in particular in terms of positioning of the most intense rainfall. It would allow assessing the data that are crucial to properly anticipate this kind of event. If the TRAM model had been used for the forecasting of this event, would it have improved the forecast?

The TRAM model is forced using ECMWF analyses instead of forecasts, ensuring the application of the most optimal initial and lateral boundary conditions. To achieve this, the TRAM model underwent extensive evaluation against various sources of external large-scale analysis, considering diverse domain sizes, time horizons and various vertical and horizontal resolutions to optimize its performance in terms of the simulated rainfall.

Rather than focusing on the numerical predictability of the event, the model configuration is designed to accurately describe the leading physical mechanisms responsible for the onset and evolution of the convective systems. Consequently, the reference to numerical predictability in Section 3.1 has been removed (line 187 onwards). Furthermore, during the discussion and review process of this study, the study describing in detail the TRAM model was officially accepted in the QJRMS journal. In the updated reference provided, readers can find comprehensive details regarding the features and effective functioning of the model.

In a technical note (reference below), members of AEMET studied the performance of their operational models (deterministic: HRES-IFS and HARMONIE-AREMO; ensemble prediction systems: ENS-IFS ECMWF and AEMET-SREPS) for this case study. The conclusion is quite clear regarding the complexity of the situation: deterministic models faced challenges in locating the signal of extreme precipitation over Tarragona and the southern part of Lleida, due to the difficulty in representing the convergence line and convective train. Only a few members of the ensembles provided some indication of the risk of very heavy rainfall in those areas. The majority of numerical predictions (deterministic and probabilistic) indicated that the most extreme rains would occur in the northeast of the Catalan region (towards the Pyrenees and the province of Girona), performing better during the latter part of the episode.

In the new conclusion section (line 669 onwards), it is highlighted the need of implementing EPSs that produce disturbances at scales as small as those explicitly resolved by convection-permitting NWP models. Producing disturbances at these spatial and temporal ranges has become a paramount task to address the rapid growth of errors associated with convective and mesoscale processes. Furthermore, there is a need for a deeper investigation into ensemble generation strategies that sample uncertainties associated with the formulations of physical processes in NWP models. For instance, apart from considering multiple combinations of different physical parameterizations, exploring alternatives such as the effectiveness of HEPSs based on QPFs generated from stochastic physics parameterizations should be explored in next future. The authors have planned a future study in depth about all these issues and how these strategies can enhance numerical predictability for HPEs.

Cuevas, G., and Pascual. R., 2021: Episodio de tormentas y lluvias torrenciales en Cataluña durante los días 22 y 23 de octubre de 2019. Technical Note 36 of AEMET, 37 pp. Ministerio para la Transición Ecológica y el Reto Demográfico, Agencia Estatal de Meteorología, Delegación Territorial en Cataluña (Barcelona, Spain). https://doi.org/10.31978/666-20-027-X.

Romero, R., 2023: TRAM: A new non-hydrostatic fully compressible numerical model suited for all kinds of regional atmospheric predictions. Quart. J. R. Meteorol. Soc., DOI 10.1002/qj.4639.

1/ Line 185. The authors mention that they optimized their simulation to get a higher performance in terms of Quantitative Precipitation Estimate (QPE). Which performance criteria were used for that?

The optimization of the TRAM model settings for studying the physical factors of the episode was, indeed, primarily qualitative. No quantitative verification index of the simulated precipitation structure against observations was applied; instead, it was assessed by expert judgment that the location of the extreme precipitation core in Tarragona and the south of Lleida, as well as the quantities, were most accurate.

Specifically, the configuration of the HR domain routinely run by the Group of Meteorology at the University of the Balearic Islands on a daily basis was used as a reference (see https://meteo.uib.es/tram). However, for this study, the horizontal resolution was doubled without increasing the number of grid points. The effect of nesting the model in the large-scale fields of GFS and ERA5 was especially investigated, revealing that the latter produced the best results.

2/ line 188: ERA5 reanalysis is used for the boundary conditions of the model. What is the reliability of the reanalysis for the conditions of October 22 2019? Did the authors perform a sensitivity study to uncertainties in the ERA5 reanalysis?

Yes, as mentioned in the previous point, the simulation quality (in terms of precipitation in the affected areas) was investigated based on using GFS-NCEP predictions or analyses, and ERA5-ECMWF reanalyses, for nesting the model.

For this episode, ERA5 data provided better results in the study basin. Interestingly, other examined case studies with TRAM did not reveal significant differences between nesting the model in ERA5 and using ECMWF operational analyses. Therefore, the latter (which are less readily accessible) were not considered.

3/ Section 3.2. The authors present how they derived QPE from radar and ground station data. The method relies on radar rainfall field corrected from beam occlusion and attenuation, using a standard rainfall- reflectivity relationship. If I understood correctly, the radar fields are then corrected for possible biases using the automatic ground stations. It seems that QPE used in this study are different from those presented in Martin-Vide et al. (2023) who used co-kriging with external drift of radar and ground stations data. Have the authors compared the two estimates? Have they performed a sensitivity analysis of the hydrological response with respect to uncertainties in QPE estimates?

Correct. In this study the QPEs are estimated by combing radar-derived precipitation with observations from the automatic rain-gauges. Before estimating precipitation, raw reflectivity volume scans are corrected from beam blockage and attenuation due to heavy rainfall. Next, the standard WSR-88D convective rainfall rate-reflectivity relationship is applied to derive rainfall rate from the radar reflectivity factor. Following this, potential biases in hourly cumulative rainfall and patterns are corrected through a dynamical fitting process involving the automatic pluviometers available across the region of interest. Finally, an independent validation is conducted by comparing the total radar-derived precipitation with observations from the AEMET independent network of daily pluviometers The cumulative precipitation from radar estimates compare well with raingauge measurements (Figure 3).

As highlighted by the reviewer, variations in rainfall estimates derived from radar observations are to be expected when using different methods. Martín-Vide et al. (2023) estimated the rainfall field by combining rain-gauge records and the rainfall field derived by the Catalan Meteorological Service (SMC) from its radar network, consisting of 4 single-polarization C-band radars. These authors combined rain records from the automatic rain-gauges and weather radars by implementing kriging with an external drift. This approach interpolates rain gauge observations using the radar-derived rainfall as a drift, effectively capturing the spatio-temporal variability of the rainfall field.

In this study, the authors use raw reflectivity volume scans by the Barcelona Doppler C-band radar of the Spanish Agency of Meteorology (AEMET) radar network. Firstly, it is important to note that due to differences in technical features and hardware between radar networks, they can perform differently. In this work, the direct use of the raw volume scan data by the Barcelona radar of AEMET enables the application of an integrated set of procedures. These procedures are designed to detect and correct potential sources or errors, such as, partial beam occlusion, signal attenuation, and radar hardware miscalibration.

Even though the radar-derived cumulative precipitation and raingauge measurements compare well in this study, it is important to acknowledge uncertainties in reproducing fine features of the highly variable precipitation patterns. These inaccuracies stem from the assumptions inherent in different correction procedures. While this issue is very interesting, the authors have not compared the two QPE estimates nor have performed a sensitivity analysis of the hydrological response in relation to these uncertainties. The authors agree that it represents a compelling area for future research. Consequently, statements addressing this issue have been incorporated in the revised version of the manuscript (line 423 onwards).

4/ Figure 3: this figure shows that high precipitation rainfall amounts were also recorded in the catchments located in the west of the Francoli catchment. Why did the authors focus on the Francoli catchment only?

Indeed, the HPE on October 22, 2019 HPE also impacted the basins adjacent to the Francolí catchment, leading to floods. However, these floods were less severe and did not produce catastrophic downstream impacts. Additionally, the field campaign conducted by Martín-Vide et al. (2023) focused exclusively on the Francolí watershed. Consequently, there is a lack of detailed information regarding peak discharges and times-to-peak at small drainage scales in the adjacent catchments. Furthermore, the FLOOD-UP FRANCOLÍ citizen campaign was specifically carried out by interviewing the population in the most affected towns inside the Francolí catchment. Therefore, there is also no available information concerning human response, perceptions and specific reactions during the course of this natural hazard.

5/ section 3.3. Hydrological modeling: the hydrological model used in the study is quite simple in terms of hydraulic transfer within the river. Martin-Vide et al. (2023) show that the event led to large geomorphological changes in the riverbed. Is this process taken into account in the hydrological modeling? Does the model make hypotheses about the river bed section (width, depth)?

No, as emphasized by the reviewer, the KLEM model is simple in terms of hydraulic transfer within the river. It relies on a time-invariant velocity value along the channel network. This assumption has found application in several flood modelling studies (Marchi et al., 2010; Nicótina et al., 2008; Zanon et al., 2010, among others). The assessment of simulation results presented in these studies supports the assumption that models of the hydrologic response employing basin-constant channel celerity explain observed travel time distributions, at least for high-flows conditions as observed in Pilgrim (1976). Additional comments have been included in line 225 onwards

Furthermore, hydrological models based on the use of an infiltration equation and invariant flow velocities have been employed by several authors for simulating flash floods (Zhang et al., 2001; Giannoni et al., 2003; Javier et al., 2007; Borga et al., 2007; Sangati et al., 2009). This approach allows for the potential transposition of model results from this study to inform other investigations

Borga, M., Boscolo, P., Zanon, F., Sangati, M., 2007. Hydrometeorological analysis of the August 29, 2003 flash flood in the eastern Italian Alps. Journal of Hydrometeorology 8 (5), 1049–1067.

Giannoni, F., Smith, J.A., Zhang, Z., Roth, G., 2003. Hydrologic modeling of extreme floods using radar rainfall observations. Advances in Water Resources 26, 195–203.

Javier, J.R.N., Smith, J.A., Meierdiercks, K.L., Baeck, M.L., Miller, A.J., 2007. Flash flood forecasting for small urban watersheds in the Baltimore metropolitan region. Weather and Forecasting 22 (6), 1331–1344.

Marchi, L., Borga, M., Preciso, E., Gaume, E., 2010. Characterisation of selected extreme flash floods in Europe and implications for flood risk management. Journal of Hydrology 394 (1–2), 118–133. doi:10.1016/j.jhydrol.2010.07.017.

Nicótina, L., Alessi Celegon, E., Rinaldo, A., Marani, M., 2008. On the impact of rainfall patterns on the hydrologic response. Water Resources Research 44, W12401. doi:10.1029/2007WR006654.

Pilgrim, D.H., 1976. Travel times and nonlinearity of flood runoff from tracer measurements on a small watershed. Water Resour. Res. 12 (3), 487–496.

Sangati, M., Borga, M., 2009. Influence of rainfall spatial resolution on flash flood modelling. Natural Hazards and Earth System Sciences 9, 575–584. <http://www.nat-hazards-earth-syst-sci.net/9/575/2009/>.

Zanon, F., Borga, M., Zoccatelli, D., Marchi, L., Gaume, E., Bonnifait, L., and Delrieu, G. (2010). Hydrological analysis of a flash flood across a climatic and geologic gradient: The September 18, 2007 event in Western Slovenia. Journal of Hydrology, 394(1-2), 182-197.

Zhang, Y., Smith, J.A., Baeck, M.L., 2001. The hydrology and hydrometeorology of extreme floods in the Great Plains of eastern Nebraska. Advances in Water Resources 24, 1037–1050.

6/ Line 238. Could you define the term helicity?

Storm Relative Helicity (SRH) is a meteorological parameter used to assess the potential for rotating updrafts within a storm environment (e.g. Markowski and Richardson, 2010). It quantifies the relative rotation of air near a storm, particularly in the lower atmosphere. SRH is calculated by considering the wind speed and direction at different altitudes within the 1000-700 hPa layer in the case of the TRAM output. A clarifying sentence has been included in the revised version of the paper (line 257 onwards).

Markowski, P. and Richardson, Y. (2010) Mesoscale meteorology in midlatitudes. Wiley-Blackwell, p. 407.

7/ Line 272. The simulation with an erased topography is particularly interesting, as it perfectly highlights the role of topography in enhancing precipitation amounts.

Indeed, these simulations, by excluding the topography in the model, shed light on the physical influence of this factor—often dominant (or at least influential to a considerable extent) in flooding episodes affecting Mediterranean coastal regions

8/ Line 297. "Cumulative precipitation of 29.8 mm to 39.5 mm ...". Does this amount refer to the third part of the rainfall event?

Yes, it does. The dynamics of the successive convective systems impacting the Francolí basin have been further clarified in the revised manuscript. To ease the understanding of the sequence, the text simply alludes to three distinct organized convective systems that sequentially affected the basins.

9/ Line 310 "the relationship between the temporal and spatial scales" and caption of Figure 8d. You should be more explicit on the variable you are plotting in Figure 8d.

The authors agree with the reviewer. The sentence has been rewritten as: "Additionally, the drainage areas impacted by precipitation intensities surpassing 20 mmh-1 and 50 mmh-1 as well as their temporal durations are explored". The caption has been modified accordingly.

10/Line 333. The low runoff ratio over the entire basin is not so surprising as half of the catchment was almost not affected by the rainfall.

Well, the authors partially agree with the reviewer's statement. Indeed, the runoff ratio decreased with an increasing drainage basin due to the limited area affected by exceptionally high rainfall. However, the drainage areas less impacted by the HPE within the Francolí basin still accumulated rainfall amounts exceeding 50 mm, as indicated by radar-derived precipitation. It would have been highly insightful to estimate runoff ratios at the different hydrometric sections (Fig. 1) using the hydrological control simulation. However, the moderate error in the simulated flow volume at the catchment outlet has hindered the execution of this task, as a result of the large uncertainties at these river sections.

11/ Section 4.3.2 The hydrological model is only evaluated by comparing with the outlet discharge, which is not well representative of the area with high specific discharge. The model evaluation would be more convincing if the model results were also compared with the maximum peak discharge and time to peak estimated by Martin-Vide et al. (2023). I would have expected this comparison in the paper. A discussion about the impact of the strong geomorphological changes and large wood transport mentioned by Martin-Vide et al. (2023) and induced by the flood on the hydrological model would also be welcome.

The authors appreciate this comment by the reviewer. We agree that evaluating the hydrological model's performance only at the basin outlet may not adequately represent the drainage areas most severely impacted by the flash flood. Originally, the intent was to assess whether the hydrological model could effectively capture the overall hydrological response based on the measured flood hydrograph at the catchment outlet.

In the revised manuscript, the authors have addressed the reviewer's concerns by comparing flood peak estimates and times-to-peak at the hydrometric sections where estimates are available. The first two paragraphs in Section 4.3.2 have been revised accordingly. The hydrological simulation now provides information for all the hydrometric sections surveyed by Martín-Vide et al. (2023). In addition, a new table 4 has been included to explicitly compare model results with estimates. These modelling outcomes confirms the reliability of the control simulation not only in reproducing the overall basin response, but also in capturing the hydrological response at smaller drainage areas. Figure 11 has been updated to incorporate simulated times for all the aforementioned river sections, facilitating a graphical comparison between model outputs, estimates, and observations. To enhance clarity, the names of the hydrometric sections with available estimates/observations have been included.

Well, KLEM model does not account for geomorphological changes in river channels or for the interaction of woody debris with bridges. These factors primarily influence peak discharges and timing at various river sections in the hydrological model simulations. Transmission losses through the channels are not explicitly considered as well. Infiltration through river beds is indirectly accounted for via the infiltration scheme through the initial abstraction rate, which is considered a calibration parameter. Curve numbers are an input to this work as they are derived from field measurements.

Uncertainties in times-to-peak are minimized using field estimates from Martín-Vide et al. (2023). The hydrological model accurately simulates all times-to-peak within the estimated ranges at the different hydrometric sections (Table 4), providing robust outputs for further analysis of catchment dynamics in Section

5.1. However, the model is unable to consider the interaction of woody debris with various bridges along the river and its consequential impact on peak discharges at surveyed river sections. All concerns raised by the reviewer have been properly addressed in the rewritten section 4.3.2 of the revised manuscript, which is dedicated to hydrological modelling.

12/ Line 444. Was the rainfall hyetograph computed for each sub-catchment where peak discharge was estimated?

Yes, it was. A comment to this respect has been included in the new version of the manuscript, line 485.

13/ Figure 11: why only 3 estimates appear in this figure whereas Table 1 provides 7 estimates of peak discharge from the post-event survey? Could the location of the estimates be plotted in Figure 1?

Figure 11 has been improved to clarify the concerns raised by the reviewer. While estimates of peak discharge are available for all river sections specified in Table 1, their timing could be confidently determined for only 4 of them through interviews with witnesses. The updated Figure 11 now include labels indicating the names of the hydrometric sections where the lag times have been simulated, estimated or derived from observations.

14/ Lines 465-470: the explanation about points of larger areas lying on the Marchi et al. (2010) curve is not very clear.

The explanation of the sharp transition in lag time with increasing drainage size for this flash flood event has been clarified in the revised section 5.1, line 507 onwards

15/line 517: which part of the 44 millions euros was related to the Francoli catchment?

The total compensation disbursed to the municipalities within the Francolí basin amounted to 7.4 million euros (not adjusted for inflation), representing 16.6% of the overall sum. This clarification has been incorporated into the text, line 578.

16: Lines 518-526. It seems that the fatalities in the Francoli catchment are more related to the obstruction of the bridges that led to particular conditions in terms of water height than to the hydrological conditions themselves. Does the model simulation point out to the same locations for critical hydrological conditions?

The authors refer to the answers of specific comments 5 and 11.

17/ Conclusions: the paper lacks a discussion section, before the conclusions section. The latter is quite long and should be shortened.

As mentioned earlier, a new discussion section has been added to compare the results of the current study with previous findings and provide insights into elements that may be of interest to a broader audience. The conclusions and abstract have been accordingly revised.

18/ Line 566 "belatedly"??

The authors apologize for the incorrect use of this word as they are not English native speakers. The word has been changed to delayed.

19/ Data availability. And what about the availability of post-event data, of the modeling results?

The authors have added a statement of the availability of post-event data and modelling results. These can be obtained under request to the authors.

20/ Figure 4: The caption should provide indication on which variable is shown in colors and which variable is shown with lines

21/Figure 12. The caption of this figure must be expanded to be understandable. For instance, what do the colored lines at the top of the figure refer to? Explain what are the rainfall centroids.

The captions of these figures have been improved

REVIEWER 2

General comments

The manuscript is interesting and contributes to our knowledge about the hydrometeorological factors of flash flood events, but the presentation is poor. The manuscript structure needs a lot of work; there is no section explicitly for results and discussion. These make the manuscript hard to follow. The study contribution and research gap that aims to be addressed are also unclear. I have recommended a few edits and comments in the PDF.

The authors would like to express their gratitude for reviewing this study and providing exhaustive and constructive comments to enhance the content of the revised manuscript. The titles of Sections 4 and 5, along with their respective subsections, have been made more descriptive of the presented results. The changes are believed to improve the readability. In addition, a new abstract and discussion and conclusion sections have been included in the revised version of the paper.

Specific comments

Here are additional comments:

1. Overall, the writing is OK but some improvements should be addressed.

2. The illustrations need major improvements.

3. Abstract should be revised to provide information about methods and results. Also, please clarify the unique aspects of this study.

The abstract has been rewritten and the illustrations have been improved to address the concerns raised by the reviewer. The revised abstract now offers insights into the methods employed, present key results and highlights the novel aspects of this study.

4. Introduction: Please explicitly discuss the unique aspects and novelty of this paper.

5. Currently introduction contain some information about flash flood in the case study, but the definition of flash flood is missing. In addition, some examples have been mentioned for small watersheds, but the case study is not as small size as these examples. How is the flash flood dynamic in your case study similar to these cases? Are there other types of floods in these areas?

6. In some paragraphs of introduction section, several references are presented at the end of a paragraph, but these need to be specifically cited throughout the paragraph.

The introduction has been thoroughly revised in the updated manuscript to address all concerns raised by the reviewer

7. The term "social response" is too broad and should be more specific. Do you mean management actions?

Effectively, the term "social response" covers a broad spectrum of social actions, encompassing from the warning procedure to responses at individual, group and organizational levels. The warning procedure itself involves several actions such as, monitoring, forecasting strategies, and the planning of management measures. In this study, the authors opted for the more generic term "social response" because the manuscript not only evaluates management activities but also describes human responses and citizen perceptions during the flash flood.

The revised version breaks down the extensive social response into two main components, according to Creutin et al. (2009): management activities and human responses. The former includes three different types of actions: information, organization, and protection. The latter encompasses human responses within three groups: Individual, communal and institutional. Consequently, the title of new section 5.2.1 now incorporates the more precise terms "risk management" and "human response". In addition, a more quantitative analysis has been conducted based on information gathered during the FLOOD-UP FRANCOLÍ citizen campaign in the Conca de Barberà council. New Table 7 and Figure 12 provide a more detailed account of the types of actions and the timeline followed during the course of the flash flood.

8. Case study section needs to discuss the watershed characteristics such as climate, annual precipitation, land cover distribution, topography and other factors related to flash floods.

The authors value the reviewer's comment and acknowledge the importance of better contextualizing the Francolí basin in terms of climate, annual precipitation, etc. Section 2.1.1 of the revised manuscript contains all this pertinent information.

9. Please add a schematic view of your methodology as a figure at the beginning of Section 3.

10. Sections 4 and 5 should be renamed as results and discussion**.**

As mentioned earlier, sections 4 and 5 have been renamed to provide clearer indications of the results presented within them. Additionally, a new discussion section has been incorporated in the revised version of the paper and the conclusions has been modified accordingly.

12. The control numerical simulation in Section 4.1 should be discussed in detail.

The authors respectfully disagree with the reviewer's comment. This study allocates two pages to discuss the meteorological control simulation. The primary goal of Section 4.1 is to highlight the key physical factors contributing to the development of the HPE. Providing additional details in this section would adversely affect readability and unnecessarily lengthen the study.

12. The models (TRAM, QPEs KLEM etc.) have inconsistent spatial resolutions. How did you handle this inconsistency?

The TRAM meteorological model is devoted to analyse the physical factors at meso- and synoptic-scales that contributed to this catastrophic flash flooding, with a spatial resolution is 3 km. The quantitative precipitation estimates (QPEs) derived from radar observations have a finer spatial resolution of 1 km and a temporal resolution of 10 minutes. These increased spatio-temporal scales enable a more thorough analysis of the key features of the heavy precipitation event that led to the flash flood in the Francolí basin.

The KLEM model operates at a spatial resolution of 25 meters to examine in detail the interaction between the high spatial and temporal variability in rainfall fields and the geomorphological and hydrological factors influencing basin response to heavy rainfall. In our perspective, there is no inconsistency, as different tools and procedures are employed to investigate distinct physical factors and mechanisms at varying spatial and temporal scales. This chain of models and procedures allows for the description of the cascading succession of physical mechanisms and their interrelations that resulted in this event, spanning from the meso- to the micro-scale.

13. My understanding is that the automatic gauges record data at sub-daily timescale but the number of these stations are limited, particularly for streamflow. How did you use daily data for a rapid catastrophic flash flood event? What limitations and uncertainties exist here?

There are 59 automatic rain-gauges located inside or very close to the Francolí basin, recording precipitation at temporal resolutions between 5 and 10 minutes and belonging to different regional or state institutions. In addition, the Catalan Water Agency deployed two automatic stream gauges along the Francolí river in Montblanc and Tarragona, two cities crossed by the river. Montblanc encloses the upper Francolí catchment with a drainage area of 339.9 km^2, while Tarragona measures streamflow near the watershed outlet, covering a basin area of 809.1 km^2. Unfortunately, the flood bore destroyed the stream gauge in Montblanc, resulting in the unavailability of a complete time series of data for the 22 October 2019 episode.

Daily data are only observed from an independent network of pluviometers deployed by the Spanish Agency of Meteorology. These rainfall data have been used in this work solely for conducting a safety validation test of the QPEs. However, it is acknowledged that the automatic stream gauge is limited to just a river section for this event. Consequently, the hydrological model simulates the entire catchment and is calibrated against streamflow observations at the catchment outlet. Thus, evaluating the performance of the hydrological model at the basin outlet may not be entirely representative of the drainage areas that suffered the most catastrophic impacts.

However, Martín-Vide et al. (2023) conducted a comprehensive post-event field campaign, providing peak flood estimates and timing at various river sections in the upper Francolí catchment. To address the concerns raised by the reviewer, outputs from the hydrological simulation have been compared against these estimates. In this regard, the first two paragraphs in Section 4.3.2 have been modified to address the reviewer's concerns, and a new Table 4 has been added to explicitly compare model results with estimates.

These modelling results confirm the quality of the control simulation not only in reproducing the overall basin response, but also in capturing the hydrological response at smaller drainage areas. In addition, Figure 11 has been modified to include the simulated times at all the aforementioned river sections, allowing for a graphical comparison of model outputs with estimates and observations. To enhance the graphical interpretation, the names of the hydrometric sections where estimates and observations are available have been included.

14. Section 4.3.3: Add a table and show the sensitivity scenarios.

New Table 5 and Figure 10 present the results obtained from conducting the sensitivity tests

15. The initial soil moisture is determined based on the antecedent precipitation, as a standard proxy. Why not using global data like ERA5 and CCI that directly present the soil moisture?

This is a very interesting point. Numerous hydrological models are updated with initial soil moisture estimates derived from ERA5 or the soil moisture project from the ESA Climate Change Initiative. Frequently used for real-time hydrological forecasting, these models build on more complex infiltration equations, often resolving the water balance equation.

In contrast, the soil conservation service curve number relies on antecedent precipitation to evaluate initial moisture conditions, and this approach has been adopted in this study. However, it remains as a future task to incorporate more complex infiltration schemes into KLEM, as well as to start the model by assimilating soil moisture fields coming from these analyses.

16. I suggest using CN as a commonly used abbreviation for curve number.

Done.

17. Hydrological model calibration needs details and clarifications. Why CNs were kept invariant? Why did you use an initial abstraction ratio of 0.35 (lambda)? The sensitivity analyses should be extended by evaluating other variables like lambda.

18. What fit metrics (e.g., NSE and PBIAS) were used and how the model performance was judged based on them?

19. Any validation effort on the hydrological model?

The authors appreciate the reviewer's comments and recognize the need for additional details and clarifications regarding the hydrological model calibration. These details and clarifications are now included in the revised Section 4.3.2.

The calibration efforts are focused on reproducing peak discharge, time-to-peak, and runoff volume at the Tarragona hydrometric section where observations are available. In this study, curve numbers represent an input data as they are derived from field measurements. These are set to represent dry antecedent moisture conditions, remaining invariant. However, the initial abstraction ratio is considered a calibration parameter in the infiltration method due to significant soil retention capabilities. The presence of large storativities is associated with exceptionally low initial soil moisture content and the recharge of deep aquifers through infiltration, percolation, and transmission losses along the river beds. This approach allows to correctly simulate the observed water balance.

During the calibration process, the performance of the hydrological model is evaluated against the observed hydrograph using different objective functions, such as the NSE and relative errors in peak discharge and total direct runoff volume. The calibrated value of the initial abstraction ratio is determined to minimize errors in terms of peak discharge and runoff volume, ensuring that KLEM adequately reproduces the overall basin response.

Since calibration is based on the observed flood hydrograph at the catchment outlet, 16 additional river sections have been included in KLEM to explore and validate hydrological response at smaller drainage areas. These 17 hydrometric sections include the 7 river locations surveyed during the post-event field campaign (Table 1). The comparison of numerical simulation results in terms of peak flows and times-to-peak against post-event field estimates serves as a verification test, confirming the quality of the control simulation in reproducing the hydrological response at smaller drainage areas. These information is included in new Table 4 and Figure 11 in the revised version of the manuscript.

20. Some of the error values on Table 4 are high (<65%). How would you interpret these and the efficiency of your hydrological model?

The hydrological sensitivity tests aim to assess the influence of three specific factors on the development of the flash flood event. These are the roles of the: (i) initial soil moisture content; (ii) early rainfall period preceding to the torrential precipitation rates and amounts, and; (iii) variability of the heaviest rainfall period. Once the overall basin response is adequately reproduced by the control hydrological simulation, the sensitivity tests involve varying one specific ingredient at a time and examining its impact on the basin response for the study case. It is essential to maintain the remaining factors invariant during this procedure to ensure consistency with the hydrological control simulation.

The results of the different sensitivity tests highlight the relative importance of each factor in modulating the overall basin response, quantified by the errors in reproducing the control simulation. The highest deviation in simulated runoff volume occurs when considering normal antecedent conditions in sensitivity test 1, even with a smaller total rainfall. In terms of peak discharge, sensitivity test 2 has the most significant impact by neglecting the effect of the early rainfall period on the overall hydrological response. The variability in rainfall during the heaviest precipitation period plays a crucial role in exacerbating peak discharge.

21. Can your results be generalized to other flash flood events in the study area or flash events beyond the study area? Please discuss.

The new discussion section compares the results of the present work with previous findings and provides information about elements of the study that are of interest for a wider audience. The conclusions and the abstract have been revised accordingly.

22. Sources of uncertainty and how they can affect your results should be discussed.

The authors appreciate the reviewer’s comment. In the revised version of the manuscript, the sources of uncertainty related to the quantitative precipitation estimates and the hydrological model simulation have been properly discussed. For further details, the authors refer to rewritten section 4.3.2 and the associated results.

23. Study limitations and potential areas for future research should be discussed.

Limitations of the study and potential areas for future research have been identified and discussed throughout the revised sections. Specifically, it is important to acknowledge limitations associated with the hydrological model performance at different river sections, and uncertainties in reproducing fine features of the highly variable precipitation pattern due to assumptions made in different correction procedures when estimating rainfall from radar observations. In the revised conclusions, potential avenues for future research have also been discussed.

24. Table 1: What does "hydrometric section" mean? Please clarify the duration of total rainfall.

In this study, the term "hydrometric section" was initially used as a synonym for the term "river section". To prevent possible confusion, the former term has been replaced for the latter in Table 1 and throughout the manuscript. The duration of the total rainfall is now explicitly included in the first paragraph of section 4.2

25. Please remove "Color code" column from Table 5.

The authors respectfully disagree with the reviewer's comment. The colour code is aligned with the risk assessment scale used by the Catalan Meteorological Service. This colour scale grading is fundamental for the understanding of section 5.2.1. and Figures 12 and 13. In particular, Figure 13 illustrates the temporal progression of the warning process based on risk assessment and associated colour codes.

26. Figures 1-3 should be improved by considering the size, alignment etc.

Figures 2 and 3 have been improved in response to the reviewer's concerns. Nevertheless, the authors maintain that Figure 1 adheres to the standard configuration found in scientific literature when introducing a study region. Typically, a top-left figure showcases the main features of the region, while a central figure illustrates the main features of the catchment of interest.

27. Section 4.1: Mesoscale processes and role of orography is ambiguous. Please clarify how the simulation works and how Figures 4-5 were produced?

28. Figures 4-5 can be merged.

The authors refer to Romero (2023) for all the technical details of the control and perturbed simulations of this study. Eliminating orography from a mesoscale simulation is a standard procedure in meteorological research that allows studying the role of this factor. For further technical details, see the reference Romero (2011).

The authors kindly disagree about merging figures 4-5 in just one panel. The authors believe that it would negatively affect the structure and readability of the study.

Romero, R., 2023: TRAM: A new nonhydrostatic fully compressible numerical model suited for all kinds of regional atmospheric predictions. Quart. J. R. Meteorol. Soc., DOI 10.1002/qj.4639.

Romero, R., 2011: Application of factor separation to heavy rainfall and cyclogenesis: Mediterranean examples. Chapter 7 in Factor separation in the Atmosphere: Applications and Future Prospects, ed. Pinhas Alpert and Tatiana Sholokhman, Cambridge University Press, 87-119.

29. Figure 8d: What is the main massage of temporal relationship between drainage area and precipitation? Why is the expectation that these two should have a relationship?

Figure 8d explores the basin areas impacted by 10-min rainfall rates exceeding 20 mmh^{-1} and 50 mmh^{-1}, establishing a link between the drainage areas affected by these precipitation rates and their durations. Undoubtedly, these characteristics in rainfall fields are closely connected to runoff generation and subsequent flash flooding. Therefore, it is regarded as an additional metric for describing the spatial and temporal organization in rainfall that led to this flash flooding. This idea has been elaborated further in the revised section 4.2

30. Figure 11 is odd. Why do you have "estimated" uncertainties only on a few data points? This should be for all simulated values. Why do we have only one "observed" value? How can lag time be even observed?

As mentioned earlier, Figure 11 has undergone modifications to enhance clarity. For ease of graphical interpretation, the names of the river sections with available estimates and observations have been incorporated. In addition, the simulated lag times for the 17 river sections serving as control points in the hydrological simulation have been included for comparison. The uncertainties associated with the estimated lag times stem from the ranges of time of peak discharge estimated at the different hydrometric sections, as shown in Table 1, and derived from the post-event field campaign by Martín-Vide et al. (2023). Times of peak discharge were estimated through interviews with eyewitnesses. Their associated uncertainties are shown as vertical bars

Acknowledging that the lag time cannot be directly observed, the authors recognize that the use of the term “observed” might be misleading in this context. In this study, the lag time is computed as the temporal difference between the centre of mass of the rainfall hyetograph (i.e. the rainfall centroid) and the timing of peak discharge. Therefore, the “observed” lag-time is derived from the automatic stream-gauge data available at the catchment outlet and the rainfall fields obtained from the weather radar, justifying its label as observed. An explanatory sentence has been added to the caption of Figure 11 to prevent confusion.

31. Please summarize the key findings of your study (e.g., as bullets) in the Conclusions section.

The key findings of the study are now incorporated into the new discussion section. This section also facilitates a comparison of the results with previous findings and offers insights into elements that may be of interest to a broader audience. The conclusions section has been accordingly revised.

32. Please italicize all parameters and coefficients throughout the text.

33. Please spell out all abbreviations in the figures, tables and headings; these need to stand alone.

Done

I hope the authors find these comments useful in their research. If the authors decide to submit a revision, both sets of my comments, including the above and in the PDF, have to be addressed.

Certainly. The authors would like to express their gratitude once again to the reviewer for his/her valuable comments, which have improved the revised version of the manuscript.