Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2018-42-AC2, 2018

Open Access Earth System Science Data Discussions

Interactive comment on "Integrated high-resolution dataset of high intensity Euro-Mediterranean flash floods" *by* William Amponsah et al.

William Amponsah et al.

wamponsah@knust.edu.gh

Received and published: 20 August 2018

We would like to thank the editor and reviewers for the time spent on our manuscript. Their comments and suggestions were considered in our revised version, which will be largely improved thanks to these recommendations. Our point-to-point responses are reported below. Reviewer's Comments are reproduced and the Authors' Responses are given directly afterward. All reviewer comments are identified using the code RXCY, where X is the reviewer number and Y is the reviewer comment number (for example R1C3 means Reviewer 1 Comment 3). Authors' comments are identified as AR. Revised manuscript refers to the attached marked-up version as supplementary to the

Authors' Comments.

Anonymous Reviewer 1

R1C1: The paper "Integrated high-resolution dataset of high intensity Euro-Mediterranean flash floods" by Amponsah, W. et al. describes a dataset of high intensity flash-floods that occurred in Europe and the Mediterranean region from 1991 to 2015. The paper is quite interesting but I have some important concerns mainly regarding to the dataset itself and the structure of the paper.

AR: We thank the reviewer for his/her thorough review and valuable comments.

R1C2: First of all, I consider that the title should be modified, as in the way it is written it seems that the dataset just contains events from the Mediterranean European region, while the dataset contains also data from the Continental Europe.

AR:Thank you for the suggestion. We have modified the title to 'Integrated high-resolution dataset of high intensity European and Mediterranean flash floods' to adequately represent the spatial coverage of the dataset.

R1C3: Secondly, I guess that a dataset containing 49 events for a 25-years period (i.e. less than 2 floods/yr) might not been enough for publication, so author should consider extending it. Next, I am going to state all my comments in the same order as they appear in the manuscript.

AR:A collection of data for extreme flash flood events is a challenging task (Borga et al., 2014), due (i) to the lack of conventional hydro-meteorological data which characterizes these events (owing to the small spatio-temporal scales at which these events occur), and (ii) to the fact that extreme events are, by definition, rare. Post-flood reconstruction of flood response in ungauged streams has thus a central role in the collection of data on high-intensity flash floods. Collecting rainfall and flood data by means of opportunistic post flood surveys requires the mobilization of a group of researchers (ranging in size from 5 to more than 20 persons) for an extended period of time (ranging from a few days to some weeks). In addition to this, high quality weather radar estimates of extreme events such as the ones triggering flash floods are not easy to be gathered, due to the number of sources of error affecting radar estimation under heavy precipitation and in rough topography environment (Germann et al., 2006; Villarini and Krajevski, 2010). Owing to these reasons, a dataset of almost 50 flash flood events comprising high quality radar rainfall estimates, flood hydrographs, surveyed flood peaks at ungauged sites, and digital terrain models is simply unprecedented in size in Europe and in the Mediterranean in terms of (i) number of events, (ii) variety of provided data, and (iii) the degree of integration.

INTRODUCTION

R1C4: Lines 76-77: ""In the few cases.... major limitation". Is not that strange currently to have catchments of a relatively large size (i.e. 3000 km2) perfectly gauged, so authors should reconsider to modify that sentence.

AR: The focus of the 'major limitation' in this case represents the uncertainties in obtaining peak flows from recorded flood stage data for local extreme floods which may exceed rating curve relations, rather than the density of stations or the accuracy of the stream gauge. It is also worth noting that 3000 km2 is a representative of the upper limit area, with upstream drainage areas of stream gauge stations in our data ranging from as low as 0.27 km2 to 2586 km2. The summary statistics of drainage areas for the stream gauge stations are reported in Table 3.

R1C5: Lines 77-79: "For instance...discharge data". That happens (mainly) when the gauging stations have not been properly designed, but if they were built accordingly to the capacity and competence of the stream it should not happen.

AR: This represents the typical features of the database. Most of the events (at least for some sectors of the impacted areas) recorded flows which were outside the capacity of the main channels. Once again, the database consists of locally rare events that were not easily captured by the hydro-meteorological monitoring networks. However, this is

a rather common conditions when dealing with high-intensity flash floods.

R1C6: Lines 92-94: "The large uncertainty...(Amponsah et al., 2016)". This would be excellent but, if there is no data, how reliable could be a model which cannot be calibrated and validated?

AR: The EuroMedeFF database makes available such data for model calibration and validation. A typical integrated hydrologic flash flood analysis (Amponsah et al., 2016) combines DTM, adjusted-radar rainfall estimates and stream flow data for model calibration. Post-flood peak discharge estimates are used for model validation, taking into accounts the uncertainties affecting the accuracy of the field-based estimates.

R1C7: Lines 111-118. I guess that the structure of the paper must be changed, as I believe the current one is not appropriated. I would move the current section 4 to 3 and the other way around. Moreover, section 2 besides showing the criteria and the summary table, should include the description of the study areas showing their location, so Figure 1 should be here.

AR: We agree with the reviewer's suggestion. We implemented the changes in our revised manuscript.

CRITERIA FOR THE EUROMEDEFF DATABASE DEVELOPMENT

R1C8: Line 123: "digital terrain model... region/catchment". Which is their minimum quality?

AR: We have included the resolution of the DTM (5 – 90 m) in this section of the manuscript.

R1C9: Lines 125-127: "Rainfall data... estimates)". I guess that a 60 min resolution is probably too high for the small catchments (i.e. less than 100 km2).

AR: We completely agree on this point. In fact, the temporal resolution of the radar data is generally higher, including up to 5 min (e.g., ZIN1991, RAMOTMENASHE2006, etc.) and 10 min (e.g., LIERZA2014). In some cases, only data at 60-min resolution were available. However, we believe these situations can be indeed interesting to the scientific community because of the quality of the integrated data as well as the interest of the study case.

R1C10: Lines 129-130: "A unit peak discharge... flood event". A reference is needed here.

AR: By studying flash floods, our focus is on the most extreme events. We needed a simple threshold to define what is an extreme event. We acknowledge that a threshold on unit peak discharge can be simplistic, since it depends on factors such as climate and catchment area (Tarolli et al., 2012). A value of 0.5 can be considered as a lower threshold for flash floods across a variety of climates and studies (Gaume et al., 2009; Marchi et al., 2010; Tarolli et al., 2012; Braud et al., 2014). We have accordingly included these references in this section of the manuscript.

R1C11: Lines 131: "at least one measured flood peak". Could this peak be measured also post event? So, my concern is, how did you decide to measure it. In other words, how did you discriminate which flood events were worth to post-survey between all those for which you did not have measurements?

AR: We stress that this database is not the output of a coordinated international field survey campaign which lasted for 25 years, but it is rather a collection and homogenization (to the extent in which this was possible) of huge work done by several research institutes and national/regional agencies in different countries. Consequently, the choice of surveying a particular event was driven by contingencies, mostly related to the flood response, to the damage caused by the flood, by the lack of data from monitoring networks, or other similar evaluations. In general, the preliminary selection of flash floods was based on rainfall data (amount, intensity) from meteorological agencies and qualitative field recognition of flood response. This led to exclude a number of low-intensity events. Post-flood reconstruction of peak discharge was carried out

for events that passed this preliminary screening. Several of these events were not included in the dataset because they failed to meet the requirements (flood magnitude, spatial extent, storm duration) described in the section 2 of our paper.

R1C12: Line 139: "The upper limit". Which is the lower limit too?

AR: Table 1 includes a range of watershed areas for each event, whereas Tables 2 and 3 report summary statistics of the watershed area on the basis of climatic regions and discharge assessment methods respectively. The selection criteria considered the largest gauged area appropriate to trigger flash floods. Gauged stations also include smaller drainage areas up to 0.27 km2. Post-flood estimates of peak discharge covered areas within this range.

R1C13: Line 145: "The upper limit... up to 48 hours". I guess that value is rather too long; in micro-meso catchments a 48h continuous rainfall will not produce a flash flood, but a "regular" flood (no matter the magnitude). If that is the "official" definition of flash flood, we should re-define it...

AR: We agree with the reviewer that a 48-hour rainfall will likely trigger regular flood instead of flash flood. However, for some relatively large area, the flood-triggering rainstorm lasted up to 48 hours, with varying intense phases triggering flash floods at different locations with different peak time. The duration of most of the events included in the database fall within 6 to 12 hours (which is the official definition of flash floods – e.g., Marchi et al., 2010). Also, if we consider the shape of the flood hydrographs for the long duration rainfall events, the most intense phase of the events (i.e., around the flood peak) fall within this duration.

R1C14: Line 147: "basin-averaged...1 mm/h). Why did you select that value? Please, reference it.

AR: The methodology is similar to Marchi et al. (2010) and Tarolli et al. (2012), where the duration is defined as ‘the time duration of the flood-generating rainfall episodes which are separated by less than 6 h of rainfall hiatus'. We made this threshold explicit to reduce subjectivity. We have accordingly included these references in this section of the manuscript.

R1C15: Line 158: after "Slovenia", you should include a brief description of the study area including a map to show their location.

AR: Done. We have included a brief description of the study area as well as moved Figure 1 to this section.

R1C16: Line 160: "Table 1". How is that table ordered? Could you explain that or re-order, for example, for date or site? Please, include as well the magnitude of each flood event in the table.

AR: We have included a description on how the table was ordered. First on country basis (column 3) followed by date of flood peak for each country (column 4). The magnitude of each flood event is also included.

R1C17: Line 162: "climatic region". This should be explained either here or before in the brief description of the study areas (including a figure indicating that could help too).

AR: Following the reviewer's suggestion under DISCUSSION, we have moved Lines 271-282 of the original manuscript to this section to elaborate on the different climatic region. The authors believe that a link between Table 1 and Figure 1 (in the same section as suggested by the reviewer) should be enough to identify the climatic regions for each study area. This is also to avoid crowded information in Figure 1.

RAINFALL AND DISCHARGE ESTIMATION METHODS

R1C18: Section 3.1: in general, more details about how did you join the radar and the gauged data are needed.

AR: We have included a brief description on how rainfall estimates from radar and

gauges were merged. This really depends on the country (or on the data source), but in all cases the data sources were merged using the same procedure: a mean field bias calculated at the event accumulation scale using rain gauges located in or around the study catchment.

R1C19: Lines 181- 183: "These floods have been kept...(4 km2)". I guess that is a really interesting point. Could you please try to extend the database by following that approach or criteria? It would be really interesting to include in the database some more floods and from some other locations. I am sure that would be possible and would make the study way more interesting and useful.

AR: Extending the data set is not feasible. Collecting more data for past events cannot be done because the high water marks - which represent the most important source of information during post flood surveys – are highly obliterated and cannot be use with confidence when time is passed after the flood. However, the database could be updated to include future events that match the criteria for selection outlined in the manuscript. This will require the collection and elaboration of weather radar data, and the organization and conduction of field surveys right after the flash flood takes place.

R1C20: Lines 187-188. "Discharge data ... through IPEC". How did you join such different methods and reliabilities? Did you calculate/analyze their errors? Did you make any kind of validation of the results? That is just to be sure that all the methods are measuring comparable things and are getting similar results.

AR: Concerning the data homogeneity, we recall again that this dataset comes from the collection of data from multiple sources in multiple countries; data were collected and treated over a 25-year period. This, inevitably involves different measurement instruments, different survey methods and different personnel conducting the surveys and treating the data. Most post-flood reconstructions of peak discharge, however, were performed using the same method, i.e. the survey of cross sections and energy slope, and the application of hydraulic equations under assumption of uniform flow.

Making this dataset as homogeneous and consistent as possible required a large effort, and represents a major asset of the work. Also, the protocol for data collection was standardized through the HYDRATE project, and allowed to have consistency in data collection. The collection of a large variety of data (radar, rain gauges, IPEC, stream gauges, local records) allows to cross validate the data.

R1C21: Line 203: "a wise choice". Not too humble...

AR: We have rephrased to “an appropriate choice”.

THE EUROMEDEFF DATASET

R1C22: Section 4.1: Too repetitive, you do not need to introduce "readme" text file that many times, just at the very beginning.

AR: We have modified the section as suggested.

R1C23: Line 221-222: "with a grid size...problems". I do not think that avoiding data storage problems can be an adequate reason for using grid sizes that coarse. Moreover, in the case of the smaller catchments (less than 1 km2) how representative can be a 90m cell size?

AR: The spatial resolutions of most of the DTMs range between 5 and 75 m with only one each with 80 and 90 m grid sizes. High resolution DTM was not available for all the catchments. Higher resolution DTMs are available for some catchments but cannot be shared because of the policy of the owner institution.

R1C24: Line 224- 225: "DTM is... coordinate system". I guess that it would be much better to use the same coordinate system for all the catchments, for example WGS84. Please, apply that comment to all the places when you state that local country coordinate systems have been used.

AR: Since data were collated and prepared on country basis, we maintained the original data (which came in local country coordinate systems) and we also provided the

corresponding WGS84 system for all data. The main advantage of WGS84 is that it avoids possible conversion problems from local coordinate systems, while providing a homogeneous coordinate system throughout the database.

R1C25: Line 223: "5 m grid". I guess that should be the better resolution for all catchments, no matter the size.

AR: We agree with the reviewer that 5 m grid is a better resolution. However, such high-resolution DTM was only available for few cases in our dataset. This is due to both availability issues (i.e. higher resolution DTMs are not available for the catchment) and policy issues (i.e. higher resolution DTMs are available but cannot be shared since not owned by the research institutes involved in this work). It is worth mentioning again (see answer to R1C23) that the grid sizes of the most DTMs range between 5 and 75 m with only one each with 80 and 90 m grid sizes.

R1C26: Section Radar-rainfall data: adding the raw data could be also interesting.

AR: We agree that adding the original volume scan radar data could be interesting, but is would be out of the purpose of this dataset. We believe what we describe as ‘best available rainfall products’, thus estimates which include the merging of radar and raingauge estimates are the useful data for our target scientific community. Moreover, going back within 25 years to retrieve such data which have been adjusted with ground-based raingauge data might not be feasible in some cases.

R1C27: Line 229: "temporal resolution... 60 min or less)". I would recommend to use a unique a more detailed temporal resolution for all the catchments and events. In the case of not having the same resolution, use the most detailed.

AR: The temporal resolution of the radar rainfall data varies between 10 and 60 minutes depending on the resolutions of the original volume scan and raingauge data obtained from the various meteorological agencies. This represents the most detailed available resolutions for each event.

R1C28: Section 4.3. Flood hydrographs: which is their temporal resolution?

AR: The temporal resolution of available flood hydrographs for each event is stated in the 'Readme' text file. This also varies according to the instrumentation of the stream gauges and to the recording strategy.

DISCUSSION

R1C29: Lines 271-282: please, move all that paragraph to section 2. Introduce and explain there the different climatic regions.

AR: Thank you for the suggestion, the paragraph has been moved.

R1C30: Lines 292-298. How similar are the results obtained for the different methods? Did you carry out any type of uncertainty analysis?

AR: The results from both stream gauge data and indirect methods are compared in Figure 5, with a log-log plot between watershed area and unit peak discharge. The power-law regression of the upper envelope curve is compared to other studies (e.g., Gaume and Borga, 2008; Marchi et al., 2010). We followed the geomorphic impact-based linear error analysis of the slope conveyance discharge determination presented in Amponsah et al. (2016) for the uncertainty assessment of the field-derived peak flood estimates.

Anonymous Reviewer 2

Summary of Review

R2C1: This paper presents and describes a database of selected European flash flood events from 1991 to 2015. This database is a valuable resource for researchers, decision-makers, and other stakeholders across the hydrometeorological community. In particular, I found the unit peak discharge envelope curves quite informative. I have some concerns with a handful of assertions made in the text, some suggestions for improving wording, and a couple of questions that could possibly be answered more fully

in the manuscript. However, these concerns are generally fairly minor, and given the detailed data included in the mooted flash flood dataset, I believe this paper contains worthy contributions to the literature surrounding impactful flash floods.

AR: We thank the reviewer for his/her thorough and positive review of the manuscript and valuable comments on the database.

Specific Comments:

R2C2: Lines 50-51: Saying the dataset is both available to the public and is a "public dataset" seems redundant.

AR: We have modified this sentence in the abstract.

R2C3: Lines 136-138: This wording is slightly awkward. I would recommend "is primarily driven by" instead of "mostly derives from".

AR: Thank you for the suggestion. The sentence has been rephrased.

R2C4: Additionally, the authors should expand upon how the inclusion of an event in the database is driven by the "local observed impact". Are these criteria subjective or objective, and can they be described in the manuscript?

AR: As pointed out in previous replies to reviewer #1 comments, this database is not the output of a coordinated international field survey campaign which lasted for 25 years, but it is rather a collection and homogenization (to the extent in which this was possible) of huge work done by several research institutes and national/regional agencies in different countries. Consequently, the choice of surveying a particular event was driven by contingencies, mostly related to the flood response, to the damage caused by the flood, by the lack of data from monitoring networks, or other similar evaluations. This makes the criteria rather subjective or, at least, with changes in such criteria depending on the country, the time period, and the availability of resources to conduct the surveys. Clearly, these are aspects that cannot be rigorously described in the manuscript.

R2C5: Section 2. iii): The authors restrict causative rainfall events to durations of less than 48 hours and appropriately separate events from one another via objective criteria applied to the basin-averaged hourly rainfall intensity. However, there are many definitions of "flash flood" (for example, the U.S. National Weather Service) that rely upon the time elapsed (often 6 hours, or less) between the causative rainfall and the onset of observable flash flood impacts. This might be difficult to determine for the ungauged events without information from the people impacted, but the database nevertheless includes estimates of the timing of the peak flow for each event. Did the authors calculate the time elapsed between the rainfall and the onset of impacts to determine how the events in the database compare to some of the other flash flood definitions in the scientific literature?

AR: For some relatively large areas, the flood-triggering rainstorm lasted up to 48 hours, with varying intense phases triggering flash floods at different locations with different peak time. The duration of most of the events included in the database fall within 6 to 12 hours (which is the official definition of flash floods – e.g., Marchi et al., 2010). Also, if we consider the shape of the flood hydrographs for the long duration rainfall events, the most intense phase of the events (i.e., around the flood peak) fall within this duration. Similarly, the implementation of a spatially distributed rainfall-runoff model, through an integrated flash flood hydrologic analysis (e.g., Amponsah et al., 2016) for most of the events, permitted to calculate the time elapsed between the onset of rainfall and flood peak.

R2C6: Lines 154-155: This sentence is awkward. I would suggest "This enables us to account for antecedent rainfall in the analyses".

AR: Thank you for the suggestion. The sentence has been rephrased.

R2C7: Lines 160-161: I would suggest the following modification: "to uniquely identify" instead of "to identify unequivocally".

AR: Thank you for the suggestion. The sentence has been rephrased.

R2C8: Section 3.1: The authors summarize their procedures here and refer to several other studies on the topic of processing radar data into quantitative precipitation estimates (QPEs). However, there is some point at which radar data are so degraded (e.g., substantial beam blockage, very high beam height) as to become difficult to use for QPE. Did the authors employ any criteria to ensure that the radar data used were of sufficiently high quality to be used in the QPE process? On a related point: were any flash floods excluded from consideration on the basis that quality radar data were not available? (Though three French events were kept in without radar data, it is not clear if events in other regions were excluded due to a combination of lacking radar data and being in larger catchments).

AR: The flash floods included in the database were selected also because of the availability of high-quality weather radar estimates. This means that (i) criteria to ensure that the radar data used were of sufficiently high quality were actually implemented, but at a local scale (meaning that each country/institution uses different procedures for this), and (ii) a large number of events was not included due to low quality of the weather radar data. The three French events were kept because the authors believe that the absence of radar estimates for such small catchment size (4 km^2) is less detrimental than for floods that hit larger catchments. Also because we had a rain gauge in this 4 km2 basin. In addition, as the radar was quite far from the catchment, located in a zone with complex topography, radar data accuracy was not granted.

R2C9: Lines 221-223: This statement should be revised to something like the following: "Digital Terrain Models (DTMs) with a grid size of 90 m or less are provided. However, to avoid data storage problems," continuing as currently worded.

AR: Thank you for the suggestion. The sentence has been rephrased

R2C10: Lines 228-229: I am curious why the radar data are available in different spatial resolutions in the dataset. Elsewhere in the manuscript it seemed that the authors are starting with the raw polar coordinate radar data, so if the data are being transformed to a rectangular grid why not choose the same grid spacing for each event? I am assuming there is a good reason (i.e. radar characteristics) but I am curious.

AR: The radar estimates are produced from the raw polar coordinate reflectivity data, but were produced by different countries/institutes in different periods and, of course, using data from different instruments. This leads to the inhomogeneity underlined by the reviewer. The temporal resolution is mostly dictated by the radar original resolution, whereas the spatial resolution derives, mostly, from the instrument's resolution. It should be however recalled that the choice of the Cartesian grid spacing can also come from data size issues, in particular when data from these events are extracted from multi-annual national/regional archives in which data size can become an issue (in particular 25 years ago – as the first events date back to).

R2C11: Lines 286-288: I am not sure it is fair to draw this conclusion from this dataset. It's appropriate to say that 90% of the included discharge data is from the Mediterranean region, but this does not necessarily support other research showing that flash floods occur more in this region than elsewhere in Europe, because the authors don't discuss attempts to collect flash flood event data in other parts of Europe. This database does not seem climatologically-representative of all Europe.

AR: We agree with the reviewer on this point. We have accordingly modified this section to reflect on the scientific interests in the collation and analysis of flash flood data in the Mediterranean region compared to other European regions rather than the occurrence.

References

Amponsah, W., Marchi L., Zoccatelli D., Boni G., Cavalli M., Comiti F., Crema S., Lucía A., Marra F., Borga M., 2016. Hydrometeorological characterisation of a flash flood associated with major geomorphic effects: Assessment of peak discharge uncertainties and analysis of the runoff response. Journal of Hydrometeorology, 17, 3063–3077. http://doi.org/10.1175/JHM-D-16-0081.1

Borga M., Stoffel M., Marchi L., Marra F., Jacob M., 2014. Hydrogeomorphic response to extreme rainfall in headwater systems: flash floods and debris flows. Journal of Hydrology, 518, 194–205. http://doi.org/10.1016/j.jhydrol.2014.05.022

Braud, I., and Co-authors, 2014. Multi-scale hydrometeorological observation and modelling for flash flood understanding. Hydrol. Earth Syst. Sci., 18(9): 3733-3761.

Gaume, E.V., and Co-authors, 2009. A collation of data on European flash floods. J. Hydrol., 367, 70–78.

Germann, U., Galli G., Boscacci M., and Bolliger M., 2006. "Radar Precipitation Measurement in a Mountainous Region." Quarterly Journal of the Royal Meteorological Society 132(618): 1669–92. http://doi.wiley.com/10.1256/qj.05.190

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, 118–133, http://doi.org/10.1016/j.jhydrol.2010.07.017

Tarolli, P., Borga, M., Morin, E., Delrieu, G., 2012. Analysis of flash flood regimes in the North-Western and South-Eastern Mediterranean regions. Nat. Hazards Earth Syst. Sci., 12, 1255–1265.

Villarini, G., Krajewski W.F., 2010. "Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall." Surveys in Geophysics 31(1): 107–29.

Please also note the supplement to this comment:
https://www.earth-syst-sci-data-discuss.net/essd-2018-42/essd-2018-42-AC2-supplement.pdf

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2018-42, 2018.