RESPONSE TO REVIEWER

Review of Manuscript No.: hess-2021-352

Title: "Rainfall-runoff relationships at event scale in western Mediterranean ephemeral streams

Authors: Roberto Serrano-Notivoli, Alberto Martínez-Salvador, Rafael García-Lorenzo, David Espín-Sánchez, and Carmelo Conesa-García

We are grateful to the Reviewer for his/her thoughtful and constructive feedback. In this Response to the Reviewer's file, we provide complete documentation of the changes that have been made in response to the reviewer's suggestions and comments. The original comments are shown in **bold text** and the author responses are shown indented in plain text. Quotations from the revised manuscript are shown in *italic text*. Line numbers in the author responses refer to locations in the revised manuscript.

Referee #1

This paper aims at better understanding rainfall-runoff relationships through statistical modelling in two ephemeral streams in Spain (with a focus on rainfall events triggering runoff). The paper is well structured. The objectives are also clearly presented.

Evapotranspiration is probably another driver (see L180) – depending when extreme events occur, response in terms of runoff may differ with the stage of plant growth. Why have you not introduced ET0 data (e.g. https://essd.copernicus.org/articles/11/1917/2019/) in your analyses? e.g. considering P- ET0 as explanatory variable.

> Thank you for your suggestion. We, indeed, considered the inclusion of ET0 which we are sure that could work as one of the main drivers, especially in summertime as mentioned all along the manuscript. However, the suggested dataset does not fit with our approach of rainfall events isolation. The SPETO dataset is at a weekly temporal resolution that considers the division of a month in 4 periods, always starting in day 1 and aggregating the last days (29/30/31) in the 4th week depending on the month. This approach avoids combined weeks among consecutive months. The weekly aggregation, although useful for climatic analysis, is not applicable to our study, where we aggregate rainfall events based on daily precipitation data. Additionally, the dataset ends in 2014, meaning that the last 6 years of our period of analysis are not available.

I have some doubts about the method used for the frequency analysis: obviously, all the episodes have been kept (more than one value sampled each year) and the peak over threshold approach should be carried out to derive return levels. The generalized Pareto distribution is the most suited distribution (instead of GEV adapted for the block maxima method). For example, the empirical return period of the observed maximum and the length of the time series should be in the same order while Figure 8 suggests return periods > 100 years. Consequently, the rainfall events triggering runoff are probably more frequent than those derived from the frequency analysis. The authors have applied the block maxima approach to data resulting from the selection of over-threshold values (threshold = 0). The method and the discussion should both be revised.

> Thank you for your useful comments. Your argument is right, and we have changed the method to calculate return periods through a *peak-over-threshold* approach. As we now state in methods section (L143-150), this is the most suitable approach due to continuous series of rainfall events are available for both watersheds:
>
> *"To contextualize the different thresholds of the RE for different probabilities of generating flow in both watersheds, we estimated the return levels of the RE using the generalized Pareto distribution (GPD) for extreme events using the peak-over-threshold (POT) approach. POT is most suitable when complete time series (as RE) are available due to all values exceeding a certain threshold can serve as basis for model fitting (Coles, 2001). We used four different estimators to fit the POT data to a GPD (Maximum Likelihood Estimation (MLE); Unbiased Probability Weighted Moments (PWMU); Moments (MOM); and Likelihood Moment (LME)) to establish proper and wide confidence levels in the estimate of maximum rainfall per RE. Thresholds for the asymptotic approximation by a GPD in both watersheds were manually selected through the*

graphical representation of Mean Residual Life, the Dispersion index and the scale and shape parameters (see Figure S1 and S2).”

Coles, S.: An Introduction to Statistical Modeling of Extreme Values, Springer Series in Statistics, 208 pp., Springer, London, G. B., 2001.

We added a couple new figures in supplementary material to show the POT threshold selection.

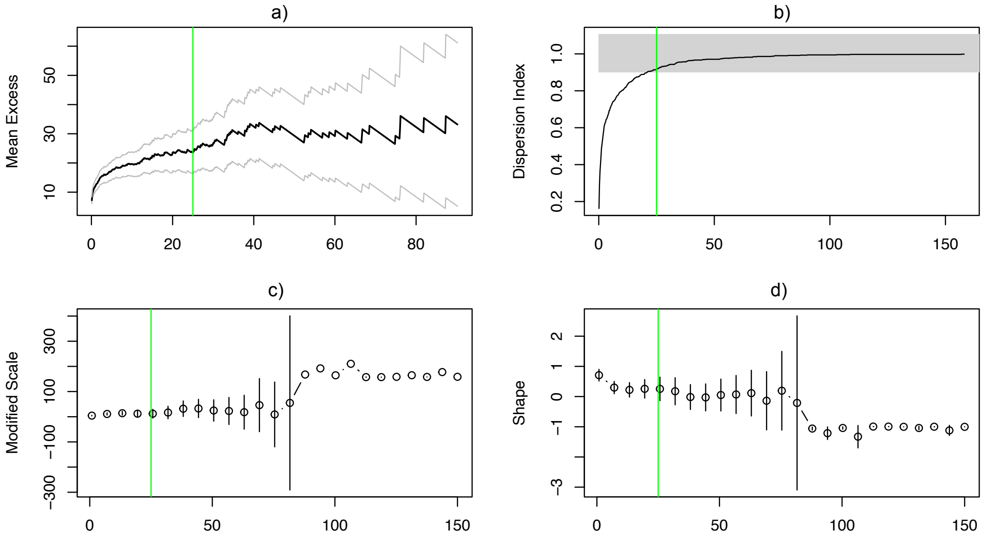

Figure S1: Graphical summary of RE threshold (μ) selection in Algeciras: a) Mean Residual Life; b) Dispersion Index; c) and d) scale and shape parameters estimates from the GPD for a range of values of μ. Green line represents the μ (25 mm) selected, implying a higher variability of its exceeding values in a), c) and d), and posing a limit in b) from which dispersion index estimates are near the theoretical value 1.

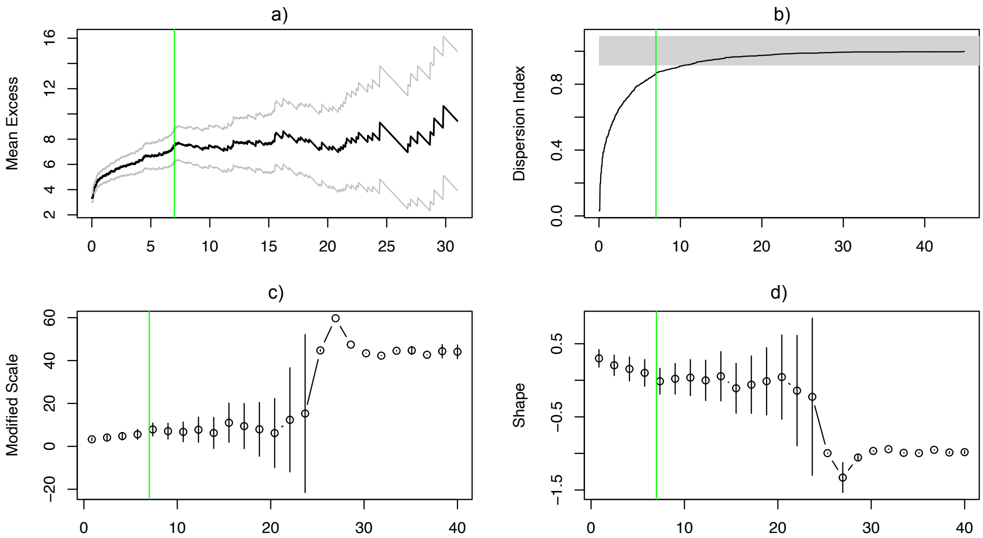

Figure S2: Graphical summary of RE threshold (μ) selection in Algeciras: a) Mean Residual Life; b) Dispersion Index; c) and d) scale and shape parameters estimates from the GPD for a range of values of μ. Green line represents the μ (7 mm) selected, implying a higher variability of its exceeding values in a), c) and d), and posing a limit in b) from which dispersion index estimates are near the theoretical value 1.

Section 3.4 has been rewritten as well as Figure 8 has been changed according to the new method for frequency analysis:

“We calculated the return levels of magnitude of the RE in Algeciras and of cumulated hourly maximums in Mula for different return periods (Figure 8). We used the POT values of RE exceeding a particular threshold (see Figure S1 and S2 for threshold selection) to adjust them to a GPD. Thresholds were 25 mm for Algeciras and 7 mm for Mula that, based on the GAM models,

represent the 95.9% and 96.4% probabilities of flow generation, respectively. Based on the fitted models, the most probable situation in which flow could be generated in Algeciras (99.5% probability) required a magnitude of 158.3 mm, which is approximately a 50-to-100-year return period. However, the return period is dramatically reduced with probabilities, meaning that high-magnitude episodes (e.g., higher than 50 mm) are rare but of key importance to ensure flow generation. The required 3.8 mm of cumulated hourly maximums in Upper Mula to ensure the flow generation at 95% probability are below the selected threshold. However, the great variability of this model increased the probabilities until 98.8% with a maximum of 44.6 mm, which represents a return period higher than 150 years. This large difference reveals the extreme irregularity of flows in Mula and the high uncertainty in prediction based only on the RE."

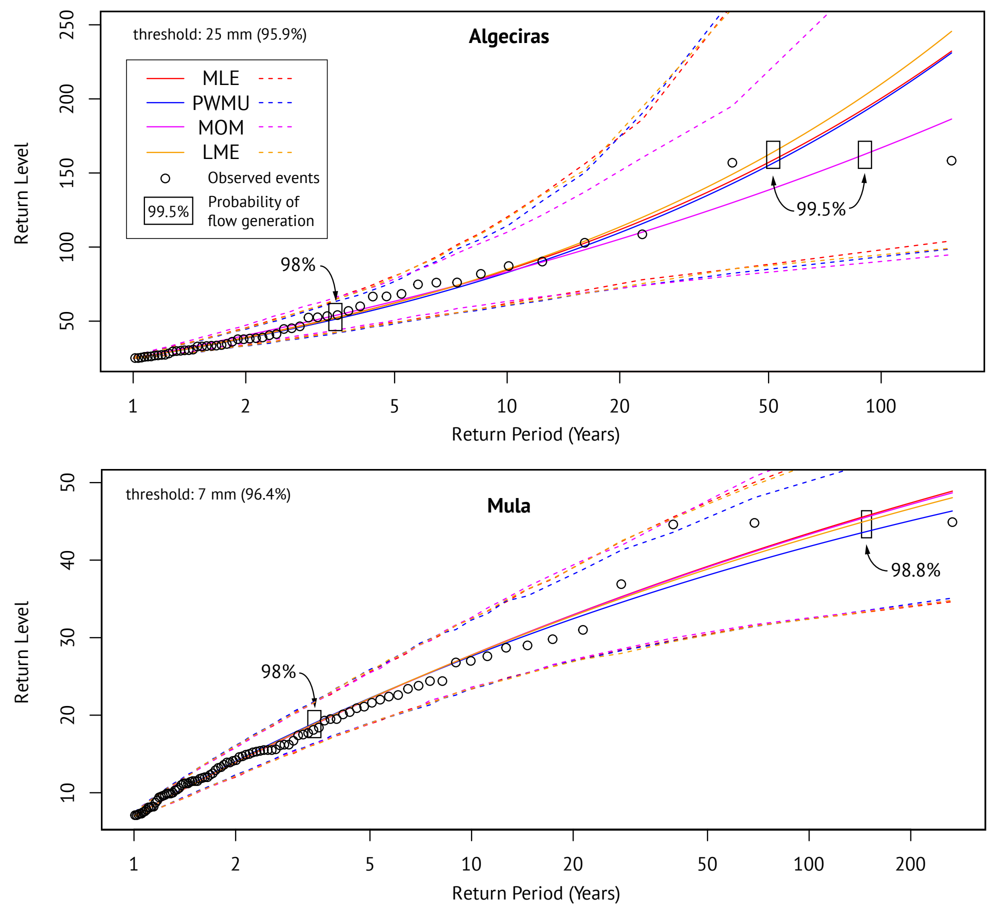

Figure 8: Return levels (RL) of magnitude of the events in Algeciras (top) and cumulated hourly maximums in Mula (bottom). Solid lines show the RL estimated for different return periods with four different methods: Maximum Likelihood Estimation (MLE); Unbiased Probability Weighted Moments (PWMU); Moments (MOM); and Likelihood Moment (LME). Dashed lines show the confidence intervals. Dots are the observed magnitude and maximums of Algeciras and Mula, respectively. RL of 98% and maximum probabilities of flow generation are indicated.

References regarding frequency analysis in the discussion have been also updated (L405-408).

"[…] Additionally, the POT approach implies an assumption of stationarity referred to the fixed character of parameters over time, and climatic series are not stationary. A non-stationary POT approach would be more appropriate, as made in previous works (e.g. Beguería et al., 2010; Agilan et al., 2021), but longer data series are needed to build reliable fittings of distributions."

Agilan, V., Unamanesh, N.V., Mujumdar, P.P.: Influence of threshold selection in modeling peaks over threshold based nonstationary extreme rainfall series. J. Hydrol., 593, 125625, https://doi.org/10.1016/j.jhydrol.2020.125625, 2021.

Beguería, S., Angulo, M., Vicente-Serrano, S.M., López-Moreno, J.I., El-Kenawy, A.: Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006. Int. J. Climatol., 31(14), 2102-2114, https://doi.org/10.1002/joc.2218, 2010.

There are many studies on rainfall-runoff relationships in ephemeral streams. The authors should develop more the peculiarities of their findings for the two catchments regarding these relationships.

Thank you for your suggestion. We included several more references to improve the discussion of the results (L359-363; L373-376)

"[…] For instance, Camarasa (2021) showed that runoff is more dependent on rainfall intensity in the Mediterranean area, and Gutiérrez-Jurado et al. (2019) demonstrated that soil type has the greatest influence on flow generation, as well as Bull et al. (2000) mentioned in a study of a watershed near to our study area. In addition, anthropic interventions such as irrigation, industrial uses, roads, or any water resources change at large scale, can modify rainfall-runoff dynamics, leading to increased consequences of flooding (Conesa-García et al., 2016; Betancourt-Suárez et al., 2021)."

"However, a change in the seasonality of flows is expected under these changing conditions of precipitation, leading to potential alterations that could intensify wet and dry periods (Pumo et al., 2016). In Algeciras and Upper Mula watersheds, climate change scenarios also depict a decrease in water resources caused by the changing seasonality, due to an increased evapotranspiration situation (Martínez-Salvador, et al., 2021)."

Betancourt-Suárez, V., García-Botella, E., Ramón-Morte, A.: Flood mapping proposal in small watersheds: A case study of the rebollos and miranda ephemeral streams (cartagena, Spain). Water, 13(1), 102, https://doi.org/10.3390/w13010102, 2021.

Bull, L.J., Kirkby, M.J., Shannon, J., Hooke, J.M.: The impact of rainstorms on floods in ephemeral channels in southeast Spain. Catena, 38(3), 191-209, https://doi.org/10.1016/S0341-8162(99)00071-5, 2000.

Camarasa, A.: Flash-flooding of ephemeral streams in the context of climate change. Geog. Res. Lett., 47(1), 121-142, https://doi.org/10.18172/cig.4838, 2021.

Conesa-García, C., García-Lorenzo, R., Pérez-Cutillas, P.: Flood hazards at ford stream crossings on ephemeral channels (south-east coast of Spain). Hydrol. Process., 31(3), 731-749, https://doi.org/10.1002/hyp.11082, 2016.

Gutiérrez-Jurado, K.Y., Partington, D., Batelaan, O., Cook, P., Shanafield, M.: What Triggers Streamflow for Intermittent Rivers and Ephemeral Streams in Low-Gradient Catchments in Mediterranean Climates. Water Resour. Res., 55(11), 9926-9946, https://doi.org/10.1029/2019WR02504, 2019.

Pumo, D., Caracciolo, D., Viola, F., Noto, L.V.: Climate change effects on the hydrological regime of small non-perennial river basins. Sci. Total Environ., 512(A), 76-92, https://doi.org/10.1016/j.scitotenv.2015.10.109, 2016.

Details:

L35: There is an inversion between first name and last name in the reference « Thibault et al. 2017 ». = = > Datry et al. is the correct reference.

Modified as suggested.

L40: a reference regarding sediment transport: https://doi.org/10.1016/j.catena.2020.104865

Reference added.

Fig. 1: we do not see the main river network. Please add the location of the two reservoirs, even if we guess that they are the mouths of the two catchments and point out the stations used to compute the precipitation time series.

Modified as suggested.

L102-106: The authors used long time series to perform a stationarity analysis. Are gridded and local data consistent during the concomitant period (correlation, mean, etc.)? This is important to assess the representativeness of the gridded data for the two catchments.

The SPREAD dataset, referenced work as Serrano-Notivoli et al. (2017), spans the period from 1950 to 2012. It was extended until 2020 in the study area using the same data series as used in the rest of the analysis, through the method described in Serrano-Notivoli et al. (2017b) to ensure the reliability of the data. We have added this reference to make clear this point in the methodological section (L110-112).

“[…] we used the SPREAD dataset (Serrano-Notivoli et al., 2017), a daily gridded precipitation dataset covering the whole Spanish territory at a 5x5 km spatial resolution, to analyse long-term trends of annual precipitation of the two watersheds by extending its period coverage until 2020 following Serrano-Notivoli et al. (2017b).”

Serrano-Notivoli, R., de Luis, M. and Beguería, S.: An R package for daily precipitation climate series reconstruction. Environ. Modell. Softw., 89, 190-195, http://dx.doi.org/10.1016/j.envsoft.2016.11.005, 2017b.

L218-219 & S2: Some criteria have been computed, but not commented (please add some comments or delete the values).

Thank you for your comments. We moved the table to supplementary material and referenced in the text the table with the GAM summaries for both watersheds.

Figs 6, 7 and 8: Please use semi-log plots with the y-axis on a logarithmic scale to make the reading easier.

Thank you for your suggestion. Figure 6 has been changed to show all variables in logarithmic scale. As this action increased some Pearson values, corresponding texts in the manuscript have been adapted to the new results. Figure 7 and (new figure) 8 are already in a semi-log scale.

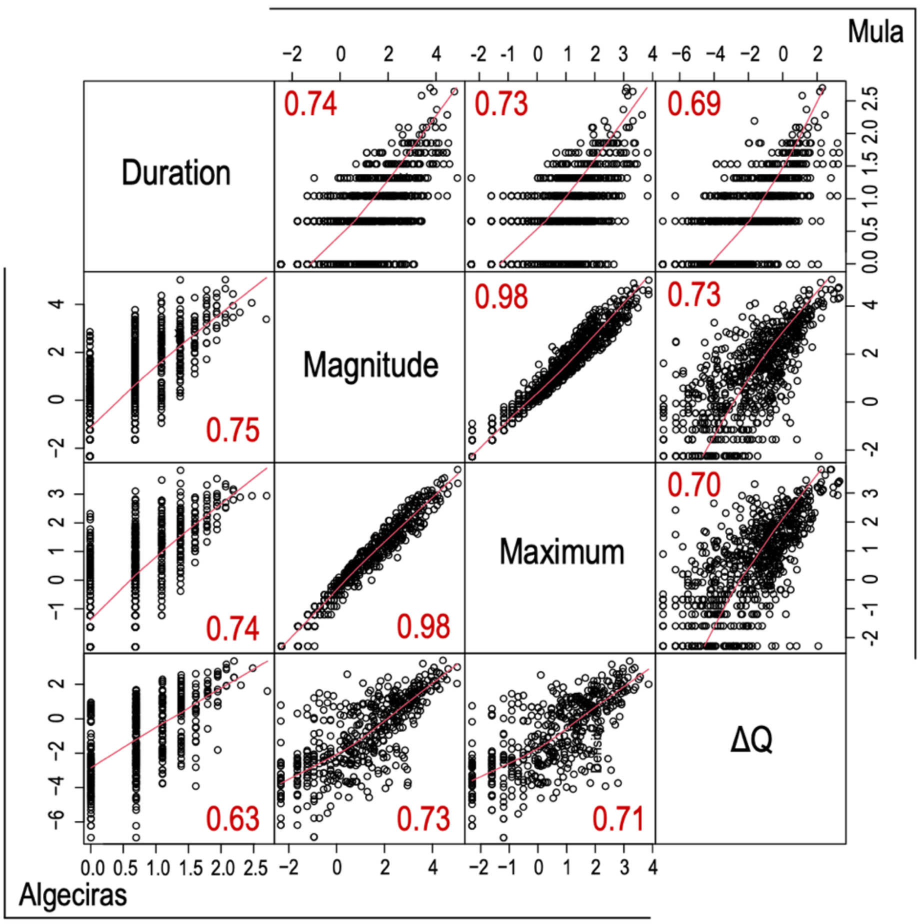

Figure 6: Values of precipitation variables and flow contribution (ΔQ) of all events in Algeciras (bottom left side) and Mula (top right side). Magnitude and maximum variables are in logarithmic scale. Pearson correlations are shown in red (all correlations are significant at α<0.01)