We thank you for your comments that helps us to improve the manuscript. Most of your concern is related to the confuse link of the study to the flash flood context and to the model limitation.

In this new manuscript, we first implemented another year of rainfall-runoff event to consolidate our results. It also allow us to challenge our methodology on the extrem rainfall-runoff of July 2021. Furtermore we included the 2016 and 2018 flash flood events in this new version, including statistics and applying the unit hydrograph on it. Results on moderate rainfall-runoff events could now be introduced against the flash flood context. We hope also, that some rewording will help you to see the connection of our study with this scope.

Concerning the limitation of the model, it is a deliberate choice to consider a very simple model that does not introduce any dependency hypothesis that we could not verify/measure. The objective of applying the model is to obtain a clean way of calculating comparable TTDs over the two catchments. The imperfect fit of the model introduces a bias and informs about the complexity of the hydrological responses, which in itself is already a result. What interests us more specifically is the variability of the TTDs, from one catchment to another, from one season to another. We have now verified by a sensitivity test that the modelling flaw only introduces a bias and does not change the observed variability. Therefore, we consider the model to be sufficient to calculate valid TTDs.

You will find the specific answers of your comments, with the modification applied to the manuscript.

General comments:

The occurrence of extreme events like flash floods are usually be linked to extraordinary catchment system states or precipitation characteristics. They may be a matter of threshold-behaviour. The analysis of the catchment runoff reaction on ordinary rainfall events with a linear model therefore does not necessarily contribute to the understanding of extreme events. Thus, I would strongly recommend to adapt the frame of your manuscript and agree with Prof. Dunkerley's comments.

We agree with you that precipitation properties play a strong role in the generation of flash floods. However, we also know - from the literature (Payrastre et al, 2013, Zanon et al, 2010 for example), that physiographic properties (relief, geology, pedology, land use) can play a key role in the acceleration of runoff processes, and the generation of very fast floods. This is what we study in this article: *What are the characteristics of a catchment that favours rapid and concentrated runoff transfers BEYOND the characteristics of rainfall*? We explore the Ernz Blanche catchment, which has been hit by record-breaking flash floods 3 times (1968, 2016, 2018), to identify what makes it particularly non-resilient to heavy rainfall. We have added the following lines to clarify our objective:

line 105: *„– what is influencing the specific flash flood event patterns, beyond the extrem rainfall properties?“*

In the context of flash floods, we believe that the study can help to identify among the ungauged catchments, those that favour rapid runoff processes.

By choosing to work on moderate events, we assume that the variability of catchment response times is also observed on moderate events. Figure 5 introduced in the new manuscript (and no longer in the supplementary materials) illustrates this assumption. We have added a line in the methodology to make it explicit:

line 127: "*We indeed assume that the hydrological reactivity of the catchment is detectable independently of the magnitude of the precipitation.* "

It is clear that extreme events are very difficult to measure and often do not occur in the short measurement periods available in projects. However, the rainfall-runoff dataset generated from

August 2019 to July 2020 is valuable for understanding the rainfall-runoff reaction of ordinary rainfall and maybe this should be the focus of the paper. In order to be able to better classify the measured events, the following information would be very helpful:

a) information about the flash flood events in 2016 and 2018
b) information about the probability of occurrence of the precipitation intensities. I am not aware whether information on design precipitation with defined return periods is available in Luxembourg. Even if such evaluations are of course subject to uncertainties, they can nevertheless provide guideline values for the classification of the measured precipitation events.
c) Presumably there is a difference in the runoff response between long-lasting precipitation events and very short ones with high intensity. In order to work out these differences, it could be valuable to classify the precipitation events.

a) In line with your comment and those of Professor Dunkerley, we have included the flash flood events of 2016 and 2018 in the manuscript. We introduce their rainfall and runoff statistics in Table 2 and have applied the unit hydrograph model to them. Although the 2016 and 2018 rainfall and runoff measurements were recorded at different location (stream gauge at Larochette and not Medernach) and with different tool (rainfall radar measurement instead of the 4 raingauge network), they provide a benchmark to our rainfall-runoff event database.

b) Our raingauge network has-been installed in 2019. The discharge measurement at Larochette only start from 2014. This short measurement period make it not possible to get robust return period values. At Luxembourg scale there are 14 daily rainfall time series starting in the mid 50‘s. The hydrological network has been installed in the 90s, as well as subdaily rainfall measurements. Using these data, it would have been possible to obtain a return period value for precipitation. However, we preferred to present the events in the form of a principal component analysis. It gives a more encompassing picture of the events, as it includes other properties than those of the precipitation (flow, season of occurrence, soil moisture, …). In order to contextualise this picture, the flash flood events of 2016 and 218 have also been positioned on the graph.

c) We agree with the fact that there is a split between long winter events and their significant hydrological response in terms of volumes and the shorter and "smaller" summer events. Instead of a classification based on the duration of precipitation (which may not be adequate for some events), we preferred to rely on the PCA analysis which combines several rainfall properties and catchment states to reveal 2 rainfall-runoff event groups corresponding to two seasons: the October-April period and the May-September period (see Figure 4).

Specific comments:

Can you please justify in section 3.1. why you used a unit hydrograph model and why you assume a constant runoff coefficient? Even if there is no ideal model, you could also use another methodology, so it would be interesting for readers to know why you chose this one, which has some weaknesses pointed out in the discussion (e.g. line 385, 392).

The idea of working with a unit hydrograph model stems from the fact that we wanted to work with a real black box, making as few assumptions as possible about the hydrological functioning. We have an input signal (the rainfall), We have an output signal (the flow). We apply a transformation function (the gamma function) to go from one to the other. The choice of the transformation function does not imply any hydrological assumption, i.e. that none of the model parameters depend on the soil type, rainfall intensity, etc. This dependence should only be revealed by the correlation analysis in a second step. In other words, we did not introduce any hypothesis into the model that we wanted to verify later.

However, as you mentionned it, there is a strong assumption in the model that the runoff transfer is the same at all times, i.e. that the watershed system is in a steady state. There is actually a transitional phase where the runoff coefficient varies, as you mentioned, from a zero value to a nominal value, but also the transfer times must very likely be longer at the beginning than at the end of the event. The choice is made here to ignore this transitory phase for lack of being able to introduce it without making assumptions on the hydrological processes. In some ways this approach is very similar in assumption to a calculation of median transfer time as being equal to median runoff time minus median rainfall time.

To clarify our purposes we reworded the beginning of the section 3.1:

„We applied a simple unit hydrograph model to reproduce the hydrological responses of each rainfall forcing over each catchment section. The unit hydrograph model assumes (by definition) that each net rainfall unit has the same TTD. We assume that the runoff coefficient (RC) is constant during the event, and we thus consider our catchment in steady state. This strong assumption prevents us from imposing a transient phase (variable RC and TTD) that we cannot measure.“

In line 395 you state that for high intensity events, flood peaks are not well simulated. This again shows the problem that your analysis is not so well embedded in the topic of flash floods.

More complex models with additional parameters will logically give a better reproduction of the hydrological response. However, we could not verify whether the better results will be for the right reasons (i.e. that the new assumptions behind the complex model are true). That‘s why we prefer to use a simple unit hydrograph model, and clearly identified in the manuscript the limitations of the model, being aware of the simplistic view of the catchment that we impose. We now have checked thank to your comments, that this limitation does not impact the assessement of the TTD properties, on which we base our study (see below).

Concerning the difficulties of the model to model the flood peaks of some summer events on the HM section, the description of the June 12, 2020 event (Figure 8 and line 365 - 368, in the modified manuscript) is very important. It shows that the response of a precipitation peak is multiplied into 3 flood peaks. This can only be modelled by integrating a spatialization of the flows, or even a hydraulic model.

It should be remembered that the aim of the model application is to extract response time distributions, not to obtain the best possible model. The model fails to model the three flood peaks, but results in the modelling of a flood peak located in the average of the three flood peaks, so we can assume that the modelling result is sufficient to extract the TTDs.

Concerning the relatively s on HM section and : this has to be lokked

A sensitivity study with changed runoff coefficients, as mentioned by Prof. Dunkerley, would at least be very helpful.

We haven‘t made sensitive analysis before getting your comment and the one mentioned by Prof. Dunkerley. Without making it exhaustively, we tested on the KOE catchment – i.e the one that seems to be more affected by the constant RC assumption – a variable RC along the event. Two Runoff Coefficients – RC^m and RC^p – have been defined: the first one characterizes the re-invigoration of the soil drainage at the beginning of the events and the second one characterizes the hydrological response in the heart of the flood respectively. Arbitrarily, the re-invigoration period is fixed to 20 h. RC^m and RC^p are calculated as indicated below:

- $$\frac{RC^m}{RC^p} = SCW\,20$$
- $$RC^m = \frac{RC \cdot V^{tot}}{V^m + SCW\,20 \cdot V^p}$$

where V^{tot} is the total rainfall amount, V^{m} the rainfall amount occurring during the first 20 hours, and $V^{p} = V^{tot} – V^{m}$. The impact on the FDC scores are presented in the supplementary materials, table S3. The TTD properties resulting in variable RC are compared to the TTD properties with constant RC on figure S4.

According to the FDC score, there is indeed an improvement of the results with a mean decrease of 2%. More specifically the results are significantly better with a variable RC for 11 out of the 40 events. Those events occurs during the November-May period.

Considering the TTD properties, TTD50 and TTDpk decreases in average by 0.5 h and 0.4 h respectively. The decrease is homogeneous on the data set. A largest difference appears during the April-Mai period, resulting in smaller range of transfer lag times uncertainties. Nevertheless, the seasonal variations of the TTD properties can be similarly observable on both unit hydrograph model simulations. The comments about TTDs properties thus based on the simulations with the constant RC hypothesis are still valid.

We recognize that the simulation results could be improved taking into account variable RC, in terms of scores and absolute values. Nevertheless, we assume – according to the presented test, that the general TTD properties variability observed over the seasons (and which is the subject of the paper) is consistent.

Please comment on possible differences between the LAI survey period and the period of your data (e.g. land use changes). Please explain how you compressed the LAI data with a spatial resolution of 1 km2 into a value for the correlation analysis. Did you use the mean value over the catchment area?

We recognized that the use of LAI that we dispose as an indicator of any influence of the vegetation cycle was not appropriate. We then removed it and replace it by the calendar day (DAY) as an indicator of the seasonal state of the catchment.

Please explain which soil moisture values you used in the correlation analysis? Did you use the value from the station situated in the respective catchment? If so, there are two stations in the HM catchment? Did you use the mean of both stations?

This is specify line 190: „***The observed soil humidity measurements were weighted according to the cover rate of each soil texture to account for their spatial variability***.“

As example in the HM section, the soil texture distribution is:

- 40% of sandy soil which covers Luxembourg sandstone (Li2, figure 2),
- 35% of clay sol which covers marls of Strassen (Li3, figure 2)
- 25% of clay sol which covers variagated marls of midlle Keuper (Km3, figure2)

We applied those rates on the three related soil moisture sensors (only the one located in KOE catchment is then not used here).

The lithological abbreviations (Km3, Li2, Li3) are confusing if one does not know the context (and therefore cannot assign the numberings – e.g. what is Li1?). In Fig. 1, some of the lithological units have an abbreviation, some do not. This looks very inconsistent. Would it be possible to do without the abbreviations (Km3, Li2, Li3)? If not, it would be important to at least cite the geological map from which these designations originate. In any case, I would delete the lithological abbreviations from the abstract.

Km3 : third layer from midlle Keuper (Triassic) period

Li2, Li3 : second and third layer from Lias (Jurassic) period

We choose to maintain the abbreviation as they are widely used at national level. As you suggested, we clarified the legend, systematically introducing the geological abbreviation. Furthermore, we sorted the substrates, from the youngest to the oldest, in order to be readable for the uninitiated. As example the marls plateaus correspond to the yougest substrate.

We added a reference on Figure 1 related to the geological properties of the area:

„Figure 1: Ernz Blanche catchment (102 km^2). Discharge and rainfall monitoring network; Left: geological substrates (see Kausch & Maquil (2018) for more details). [...]“

384: What do you mean with: “the model overestimates the rising limb of the flood wave” – is it the duration of the rising limb, its slope or something else?

We mean that the simulated discharge was higher than the observed one during the rising of the flood wave. We reworded the sentence to be understandable:

Line 476 – 477*: „the model overestimates the discharge during the rising limb of the flood wave for the KOE catchment, while it underestimates and delays the flood peak“*

Fig. 10 and 11: please explain the size of the circles

The size, as the color of the circles, are related to the Kendall or Hoeffding ceofficients. This is now specified in the legend.

504: Is „pseudo“ necessary?

We report here the conclusions of another publication, and we have decided to keep the vocabulary proposed by the authors in order not to insert any interpretation of their result.

Technical comments:

We thank you for all the detailed corrections mentioned in this section. We have incorporated all of them, some of which are commented on below where other changes have been applied.

- “et al” should be “et al.” in the whole paper (see https://www.hydrology-and-earth-system-sciences.net/submission.html#references)
- 15/17: delete “Km3, Li2, Li3” . We are keeping the information on geological substrates in the summary, as we think it is of primary importance. The full name is now preferred to be understandable to all.
- 16: add “(HM)” after lower catchment.
- 62/576: „Bronstert“ (instead of „Bronstaert“)
- 106: add “(TTD)” after transfer time distribution
- 110: delete “(TTD)” after transfer time distribution
- Fig 2 left and right are interchanged (Depending on whether the picture arrangement or the text is changed, the text in line 162 may have to be corrected.)
- Fig 3 light brown is not clearly visible, please use a darker colour
- 246, 322, 336, 347, 353: correct “VOL1H” to “VOL1h”
- 270: “one event” instead of “one events”
- Fig. 8: The labelling of the x-axes is wrong: it should be ”median transfer time”, “peak flow lag time”, “runoff response concentration”. You could also omit these words because they are explained in the figure caption and just write the abbreviations.Fig. 9: I cannot find an x-axis for reading the LAI-values.
- Fig. 10 upper line: SWC20 should be green, RC should be black
- 458: „McGlynn et al. 2004“ instead of „McGlynn, McDonnell, Seibert, and Kendall 2004“

- 478: “Scaini et al. 2018” instead of “Scaini, et al. 2018”
- 500: “Hortonian flow” instead of “hortonian flow”
- You often use „note that“, I would avoid this phrase (at least I would not use it so frequently), but this is a matter of taste. We reworded some sentence to decrease the use of this formulation.

References:

Olivier Payrastre, Eric Gaume, Pierre Javelle, Bruno Janet, Patrick Fourmigué, Philippe Lefort, André Martin, Brice Boudevillain, Pascal Brunet, Guy Delrieu, Lorenzo Marchi, Yoann Aubert, Elisabeth Dautrey, Laurence Durand, Michel Lang, Laurent Boissier, Johnny Douvinet, Claude Martin & l'équipe « enquêtes post-événements » d'HyMeX (2019) Hydrological analysis of the catastrophic flash flood of 15th June 2010 in the area of Draguignan (Var, France), La Houille Blanche, 105:3-4, 140-148, DOI: 10.1051/lhb/2019057

Zanon, F., Borga, M., Zoccatelli, D., Marchi, L., Gaume, E., Bonnifait, L., & 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), 182–197. https://doi.org/10.1016/j.jhydrol.2010.08.020