--- #RC1 ---

This paper presents an analysis of some high-quality rainfall and streamflow data collected in two adjacent catchments in Luxembourg (the Ernz Blanche basin). The objective, as reflected in the title of the manuscript, was to understand better the hydrological mechanisms resulting in flash flooding in this catchment. The paper is generally clear and straightforward to read, though I think that the main focus should have been more strongly on large rainfall events than on hydrologic response under more usual events.

We thank you for you review that was really helpful to improve our manuscript. Our main change according to your comments were:

- to include the 2016 and 2018 flash flood events in our study. The unit hydrograph model was applied on those two flash flood events. We can then connect our results on moderate rainfall-runoff events in the flash flood context. We hope also, that some rewording will help you to see the connection of our study with this scope.
- to insert a sensitivity analysis into the supplementary materials, in order to argue the small impact of our constant RC assumption on our results.
- Present more cautiously our explanations for the impact of the hydrophobicity of forest litter and the soil surface. We acknowledge that we do not have data to validate these explanations, but we believe that this open discussion could be beneficial to the scientific community, by opening up hypotheses to be tested.

In addition, we have taken advantage of this rewrite to add an event year to ensure the consistency of the correlation analyses and to work on two contrasting hydrological years (one being rather dry, the other rather wet).

Here below, we answer to each of your specific comments.

Oddly, though the authors mention the occurrence of several historical flash floods, including one in 2016 and another in 2018, they do not describe those events in any detail. They provide no discharge data, no runoff coefficients, and no rainfall event data. In order to find something of these events, I consulted an EGU Abstract (Iffly et al. 2018) by some of the same authors. There, I was able to learn that the 2016 event had much more intense rainfall than anything that the authors investigate in the present ms., recording 20 mm in 10 minutes (=120 mm/h), 50 mm in 1 hour, and up to 70 mm in 6 hours (=almost 12 mm/h). In contrast, in the present paper the most intense event reported had a maximum rainfall rate of ~ 27 mm/h. All but one of the remaining events listed in Table 2 had maximum intensities of < 10 mm/h. These seem unlikely to be responsible for flash floods. I was not able to locate information on the 2018 flash flood event for additional comparison. I think that it would help readers place the results of the current ms. in context, if some information on the historical flash floods could be provided, at least in summary.

As suggested, we added the 2016 and 2018 flash flood event properties (rainfall amount, flood peaks, runoff coefficients) in table 2, although the data is not exhaustive, with the highest impacts and flows of both floods being located downstream of the presented measurements. We also added a PCA analysis of the rainfall-runoff event dataset (Figure 4), on which the 2016 and 2018 flash flood statistics are positioned, which further contextualizes the study database in the context of flash floods.

I think that the focus on 'ordinary' events needed some comment. How can a study of much more ordinary rainfall events shed light on what occurs during the seemingly far more intense rainfalls that seemingly accounted for the historical flash floods?

We assume that there are not only the extreme properties of precipitation but the intrinsic catchment properties and its hydrological state that cause the hydrological response to be rapid and concentrated in a relatively large flood peak. In order words, we suppose that the catchment "hydrological reactivity" is independent of the rainfall magnitude enough (although this will make the high hydrological reactivity to be problematic) to be detected on a moderate rainfall-runoff event database. In order to clarify this assumption, we added two sentences in the introduction:

lines 104: "*Here, we ask – in the context of a Central European study area – what is influencing the specific flash flood event patterns beyond the extrem rainfall properties?*"

line 123-124: "We indeed assume that the hydrological reactivity of the catchment is detectable independently of the magnitude of the precipitation. The same model is also applied on the 2016 and 2018 flash flood events, with the aim of having reference transfer times characteristic of flash floods.*"*

Additionnaly we added one year of rainfall-runoff measurements to our study which results to enrich the database with 17 additional rainfall-runoff events. Among them, the event that occurred on the 13th July 2021 consists in an extrem event in terms of rainfall amount (129 mm) and discharge peak (the highest water level was recorded during that event since the oldest hydrometric station has been installed at Larochette in 2014). Although this event is not a flash flood, it enables to apply the TTD properties using extrem rainfall statistics, making possible to question the study result independency from the rainfall magnitude.

Did these [the flash floods], for instance, occur when the soil had been thoroughly wetted by antecedent rainfalls? Does surface runoff overtop ground surface roughness elements when the rain is sufficiently intense (above some threshold?), allowing a smoother and more direct path downslope? What was the nature of the precipitation? I assume that the flash floods were the result of shorter, more intense, convective events, and therefore were likely to have occurred in summer (this information is missing from the current ms.). I imagine that these were late afternoon events, but this would also be relevant information. Were there very local runoff sources located close to the stream channels, perhaps? Could the movement of convective cells parallel to the long, narrow catchment be significant? Did that occur (perhaps Doppler radar might shed some light on this)? Catchment response to intense convective cells might be quite different from that in stratiform rain, for instance, and different parts of the catchment might show altered hydrologic responses under those different rainfall inputs.

You are right, these are general characteristics of flash floods. But again (and perhaps this was poorly expressed), we are trying to determine what favours rapid and concentrated flooding, beyond the properties of rainstorms. the fact that they are convective events, of high intensity are recognized characteristics that favour flash floods, the fact that the rainstorm is located downstream of the catchment where the hydrographic network is strongly defined also. Here we seek to understand why two catchments react differently to the same rainfall event, whether it is intense or not.

Iffly et al. 2018 refer to lag times to runoff peak of just 90 minutes, whilst in the present study these lags extend to many hours.

In this study, the runoff peak response (TTDpk) varies from 0.5h to 13.7h, and more specifically on HM catchment and during the dry condition (15th April - 15th October), TTDpk's average is 1.9± 0.9h. Those results are actually in agreements with the lag times to runoff peak of 1.5h mentionned in Iffly et al. 2018. Furthermore, in the updated manuscript, we added TTDpkvalues for the 2016 and 2018 flash flood events, which are 0.9 ± 0.1h and 0.1h respectively. Those values still corresponds to the order of magnitude of the HM section's TTDs during the 15th April - 15th October period

The study is weakened by the assumption of a constant runoff coefficient through the duration of rainfall (mentioned in line 200 and elsewhere). This seems particularly inappropriate for long events of several days duration, such as were examined in this ms., and even for events of a few hours duration, when breaks in rainfall (e.g. shown in Figure 6 and Figure 7) allow soil drainage and the re-invigoration of soil infiltrability. It would have been interesting and informative to have seen at least some preliminary sensitivity testing to see how important an effect a changing runoff coefficient might have been to the hydrologic modelling. Perhaps the authors have done such tests and could comment?

We haven‘t made sensitive analysis before getting your comment. 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.

It would also strengthen the argument of the paper if the authors could present some data on hydrophobicity in the forested areas, that they appeal to as a mechanism to account for more runoff there. Was hydrophobicity actually present, or was this not investigated? If present, does it dissipate in longer events, so that perhaps it differentially affects runoff behaviour in short convective events in summer?

Unfortunately we have not carried out any hydrophobicity measurements on site, neither during the period nor elsewhere. We can only answer your question indirectly. There is indeed a notable difference between convective summer events and winter events: for the first one, the maximum intensities arrive at the beginning of the rainfall event, whereas for the latter there is a progressive increase in the intensity of the precipitation over time. We can think that the arrival of strong intensity without an initial humidification on a dry and therefore hydrophobic soil, inhibits infiltration. Runoff is then favoured all the more if the ground is sloping (hydrophobicity prevents water from attaching to the ground AND gravity leads to runoff). However, we have not yet been able to verify this hypothesis. While looking for references on this subject, I read your three interesting papers (Dunkerley, 2012, 2016, 2021), which tend, with experimental justification, to the opposite conclusions (a rain peak at the end of an event favours runoff). Nevertheless, as you mentioned in your article, your experiments are carried out on a flat terrain, and the results ultimately assess the variable infiltration capacity of soils. In a sloping configuration, there will be a first barrier to this infiltration which is the "adhesion" of the water to the soil surface before infiltration. If there is no adhesion (the hydrophobic property of the soil), then the relief might play a key role. An argue in this hypothesis is the fact that the very fast runoff are only observed on the HM section where steep slopes close to the drainage network are present. Finally, as those statement can not be verified with our current data set, we clearly specify at the beginning of our discussion that this is a plausible explanation that must be subsequently checked:

“Since our dataset appears to be (too) limited for validating our hypothesis, we propose here a list of plausible explanations – based on examples from scientific literature – for the drastic decrease in response times observed in summer on the HM section, as opposed to the KOE section.“ (lines 535 – 538)

The authors identify LAI as an important factor in the hydrologic response (TTD) (lines 491-492). Though without comment, the authors appear to use LAI data from 2002-2006, many years prior to their field data collection. This warrants some comment. Further, the LAI seems to be very small, to judge from Figure 9 (left panel), seemingly the only data presented on this variable. The authors only appear to link LAI to the speculation about litter layers and wettability, evidence for which is not provided. Could the authors offer a fuller comment on why LAI might relate to TTD? Do they consider this to be a real, physical effect, or merely a chance statistical correlation (for instance, via some other seasonally-varying parameter)? Their comments and thoughts would be helpful. They could also perhaps consider presenting LAI data for their catchments (as a map) if they have it available. It would appear to be very variable among fields, forests, etc.

We agree that the use of LAI as an indicator of any influence of the vegetation cycle was awkward. We then removed it and replace it by the calendar day (DAY) as an indicator of the seasonal state of the catchment. Having no specific indicator related to forest litter condition, its role discussed in the discussion is presented as a hypothesis to be tested, as said before.

Finally, I wondered whether there is a role for roofs, roads, drains, culverts, etc., in the catchment response. I do not know this area, but Figure 2 suggests that, at least locally, the villages may have impervious areas that are efficiently drained. The main stream channels also warrant at least some description. Have they been modified, perhaps to flow between artificial banks or walls? How significant is the channel travel time from the upper to the lower catchment? In the same way, landuse could helpfully be described, especially whether fields are tilled seasonally.

Relating to the soil sealing in connection with the presence of urban areas (essentially in Larochette), we know from our field knowledge that rapid flows come from both sides of the lateral tributaries, which are not very urbanised apart from a few villages with a few houses on the plateaus.

The transit time from the top to the bottom of the catchment area would require chemical or isotopic tracing measurements which were not carried out for this study.

Concerning land use, an additional figure has been added in the supplemental materials. The steep slopes are mainly covered by forests, the downstream part of the KOE catchment is mainly grassland. There are only crops (mainly corn) on the marly plateaus of the HM section. The seasonal development could have an influence. Nevertheless the shortest times are obtained at the end of September - beginning of October when corn is most developed. This is why we have not detailed their impact.

Overall, this is a solid study, containing some interesting results. However, I am not sure to what extent these actually bear on the factors accounting for flash flooding.

We added some information about the 2016 and 2018 events which must help to link our result to the flash flood issue. Furthermore we really believe that our study help to highlight how fast and concentrated runoff can be processed in specific catchment, beyond the rainfall properties. Consequently we think to give insights to identify catchment prone to flash flood, even if they are ungauged.

_

Minor errors:

- line 13: should be 30 km 2 (space is required between numerals and symbol for unit of measurement)
- line 46 and throughout the paper: 'et al' should be 'et al.' (as a contraction of et alia)
- line 93: should end sentence with a question mark
- line 120: omit the parentheses
- line 145: Captions are reversed (left to right)
- line 162: should say 'Figure 2 left', not right
- line 180: it would be preferable to refer to time-aggregated data as rainfall rates (they are equivalent mean rainfall rates, not true intensities)
- line 253: again, space required following numerical quantity
- Figure 9: there are two dashed lines, only one is listed in the legend
- line 500: Hortonian (capital H after the family name of Robert Horton)

The minors comments has been applied. Thanks for those detailed corrections.

-- #RC2 --

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:

- $\frac{RC^m}{RC^p} = SCW20$
- $RC^m = \frac{RC \cdot V^{tot}}{V^m + SCW20 \cdot V^p}$

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

--- #RC3 ---

We really thank you for your comments and questions on our article; It helped us to see how our discourse deviated from the importance of geology (which was our first assumption before we really analysed the data) to the importance of the catchment layout. So we put more emphasis on presenting the topography and landscape elements in the introduction which allowed us to be a bit more comfortable in the discussion.

Hereafter you will find our specific answers to your specific comments.

This study compares storm runoff processes between two catchments with similar size based on the analysis of transfer time distributions (TTDs). The authors present a seasonality in TTDs, which had a different trend between the two catchments. Quick runoff transfers occurred under dry condition in a catchment. The authors attribute the rapid flows to marly plateaus, hydrophobic forest litter, and the absence of a riparian zone in the catchment.

This paper deals with an important topic. I think their analysis of TTDs using a unit hydrograph model is effective for comparing the storm runoff characteristics between the neighboring catchments. Seasonality in TTDs (Figure 8) is especially interesting. However, data for discussing the causes of the seasonality and inter-catchment differences in TTDs are insufficient. More information about groundwater dynamics and topographic analysis is needed to discuss the causes of rapid flows. My major concerns are listed below, followed a list of specific comments and technical corrections.

General comments

Which is the novelty of this study, analytical methodology or the estimated causes of rapid runoff? If the TTD-based comparison of storm runoff characteristics is a novel approach, the authors should emphasize this content in Introduction and Discussion. If they think the causes of rapid runoff are the main findings, they should increase reliability of estimating the causes. In this manuscript, relationships between runoff mechanisms and TTDs are unclear and the factors causing rapid flows in their study sites are only speculated from the results in TTDs.

The TTD-based comparison is not a novel approach as it is already used either in old studies comparing the shapes of the hydrograph to classify a set of catchments or more recently to study the impact of catchment management (before and after restoration management, as example see Memberu et al, 2018, figure 6). From our point of view, the novelty presented here is a clear demonstration of the distinct impact of dry conditions on runoff transfer processes. It is thus shown that beyond the precipitation characteristics, the configuration of the catchment will be a key factor in the generation of a flash flood following a summer storm. In order to emphasize this idea, we reworded the paragraph summarizing the results:

“We observed a seasonality of the TTDs for both catchments, with dry conditions having an opposite impact on them. The KOE catchment reacts less quickly and more spread out under dry conditions. On HM catchment on the contrary, response times are significantly shorter and concentrated (-59% ± 33%) and (+33% ± 87%). ***This opposite seasonality*** *leads us to consider/hypothesize different control factors of the runoff transfer processes in relation with the topographic and geological layout of the catchment areas.”*

(REF:Memberu, M. W., Haghighi, A. T., Ronkanen, A.-K., Marttila, H., & Kløve, B. (2018). Effects of Drainage and Subsequent Restoration on Peatland Hydrological Processes at Catchment Scale. Water Resources Research, 54(7), 4479–4497. doi.org/10.1029/2017WR022362)

Although the authors focus on bedrock geology, groundwater dynamics in each geology are unclear. The rapid flow due to hydrophobic forest litter was also not observed in their study sites. Moreover,

differences in riparian topography between the catchments were not presented despite mentioning riparian buffering. Due to lack of these data, they only speculate the causes of rapid flow. If they want to discuss the causes based on the data in these two catchments, more detailed presentation of groundwater flows and topographic characteristics in the catchments is needed.
Unfortunately we do not have groundwater flows data, nor hydrophobic forest litter measurements. The main idea of the paper was to highlight the distinct behavior along the Ernz Blanche catchment on dry conditions. We orientating the discussion toward the causes because it seemed to us coherent to enumerate the possible factors - even though we cannot justify otherwise than by the literature. The idea here is to propose several avenues for further investigation.
Concerning the topography, it is true that we have reduced its presentation while it occupies a major part of our conclusion/discussion. We have therefore added some elements in the presentation section of the watersheds, in particular the iso-contours on figure 1 and a map superimposing slopes and Heights Above Nearest Drainage (HAND) highlighting the riparian zone on KOE and the steep slopes on the perimeter of the hydrographic network on HM.

I can't understand why the authors compared only two catchments despite the observations in six nested catchments (Figure 1). How different were TTDs between the six catchments? I think examination of relationships between TTDs and catchment characteristics (including geology, topography, catchment size, and vegetation) using the data of six catchments can provide more valuable implications. Even though groundwater flows were not observed, the causes of rapid flow may be estimated with reliability if the comparison of six catchments is conducted.
We carried out the analyses on only two catchment areas and not on all six embedded catchment areas, because on the one hand the flow series at two stations (Reisdorf, Hessemillen) are uncertain due to the backflow of a confluence and a dam respectively, and on the other hand the comparison of transfer time distributions of catchments of different size seemed less convincing to us. We have nevertheless mentioned them because the shape of the hydrographs clearly indicates a break in behavior between what happens upstream of Heffingen and what happens downstream (Larochette, Médernach, Hessemillen, Reisdorf). This is particularly showed/visible on Figure S1 in the supplementary materials.

Specific comments
Title: If the main theme of this paper is causes of flashy runoff, the causes should be examined more deeply based on other groundwater and topographic data or the comparison of TTDs between more catchments.
We changed the title for: **„Flood patterns in a catchment with mixed bedrock geology and a hilly landscape: identification of flashy runoff contributions during storm events“**
The initial objective was to find the causes of flashy contributions. Although we believe that we have identified and highlighted when and where rapid contributions occur on the White Ernz, we recognize that we do not have enough evidence to clearly identify the causes.

L14-17: Although the geology of the catchments is well described, there is little information about their topographic characteristics. I want the authors to clearly present the difference in topography between the catchments.
As suggested, We gave more details about the topographic characteristics:
The upper catchment (KOE) is dominated by **a low land area (38% of the catchment is located less than 30 m above the river network) consisting in variagated** marly bedrock (Midle Keuper Km3) and moderately steep Luxembourg sandstone outcrops (Lower Liassic Li2). The lower catchment (HM) has its drainage network deeply cut into the Luxembourg sandstone, with half of it being covered by marly plateaus (Lower Liassic Li3, **located between 80 m and 100 m above the river network)** featuring heavy clay soil.

L27-29: These causes are only the speculation and remains hypotheses. As these hypotheses were not verified, this is inappropriate as the conclusion. It may also be possible that the quick runoff under dry condition was caused by direct precipitation on stream channel and/or rapid runoff from riparian zone. As the catchment got wetter, hillslope runoff with long transfer time may contribute to stream water, which can be a possible mechanism of longer TTDs in wet conditions.
It seems to us that we are defending the same hypothesis: dry conditions imply direct and rapid runoff to the river. The real question is why this has more impact on HM than on KOE. We thus clarify our hypothesis by assuming that: I) the impact of dry conditions is stronger on marly plateaus; ii) the dry litter of sloping forests favors runoff rather than water retention and infiltration, and iii) the riparian zone when sufficiently wide allows a buffer effect on this direct runoff. We agree that those hypotheses not actually proven by experimental results, but they are suggested as such: line 30: "stand as **our main hypotheses** in this respect".

L58-60: Whereas the authors wrote "The numerous faults and cracks support quick water transfer through the weathered bedrock and explain fast hydrological responses" in this sentence, they also wrote "Less permeable bedrock will lead to ... smaller catchment mean transit times." in L80-81. Whether the weathered bedrock can contribute to fast responses (smaller transit times) or not?
...both : depending at which time scale we look at. At event scale, we observed fast flows, but at seasonal scale the baseflow release is low. „catchment mean transit times" in Pfister et al. 2017 actually refers to „baseflow mean transit times". I replaced it to clearly make the distinction.

L89: If the main problem of previous research is the lack of observations in extreme events, this should be clearly presented in Background section. The event magnitude should also be emphasized in the Results and Discussions.
You're true. Our study is based on moderate events and the event's magnitudes do not make the our study specific. The specificity of our study is to focus on what impacts the speed and amplitude of runoff processes beyond rainfall properties. The last studies in Luxembourg does not enables to answer this question. We changed the lines 90-95 to clarify this idea.
"To date, all investigations focusing on rainfall-runoff transformation processes in the Luxembourg context have been limited to small experimental watersheds (< 5 km^2) or dedicated to storage and catchment release. While these studies have substantially improved our understanding of physiographic controls on runoff generation, we still have poor knowledge of the processes leading to quick runoff on catchments with a genuine river network as in flash flood events."

L98-99: I could not understand the difference in flash flood type between Central Europe and MA regions. Please describe the difference more clearly in Background.
The lines 74-78 explain the different context between Central Europe and MA region which makes impossible to transfer knowledge of the flash flood processes in MA to the Central Europe. We thought it is clear enough but we give more details here:
"While most flash flood related literature published to date refers to the Mediterranean area (MA), the processes underlying flash floods in Central Europe remain poorly understood. This mainly relates to the fact that :

- *in these catchments (i) the climate forcing is not primarily controlled by topography (as opposed to MA),* → in MA, the Alps, the Pyrenees, the Cévennes foothills consists in elevated mountains which can block and induce convective and stable (in space) storms. In Central Europe, there is no such topographical barrier. Although the 2018 and the 2016 flash flood event were induced by relatively high rainfall amount, this is not the same order of magnitude than in MA (~50 mm compared to ~200 mm)
- *(ii) catchment storage filling states are very different between early summer (storage levels being still high when flash floods occur in Central European catchments) and autumn (storage levels being low when flash floods occur in MA catchments)* → the catchment

conditions are not the same as the season of occurrence is different so the conclusion found in MA could not be transfer to Central Europe.

- *and (iii) the underlying bedrock geology is very different between Central European and MA catchments." → As well, some studies on MA relate the flash flood processes to specific bedrock geology. Here in Central Europe, the bedrock is linked to other mountains formations, so the impact could not be deduced again.*

L116-162: I could not understand which catchments had more permeable bedrock and larger storage capacity. According to Table 1, geology seems similar between the catchments: Both had the main geology of sandstone and second geology of marls. If the authors focus on the geological features, geological difference between the catchments should be explained more clearly. Information about vegetation is also required because the effects of litter are discussed.
At a first glance the catchment seems to get the same geology; half part being marls and the other part Luxembourg sandstones. However there are two big differences:

- First, the marls layers consists in two different geological substrates with different properties. On Koe catchment the marls is a middle Keuper stratum (km3, Trias superior) and on HM the marls date from the Sinemurian period (Li3, mars of Strassen). The middle Keuper marls actually includes conglomerates and thin beds of dolomite, which can in turns include aquifers that are significant enough to be mentioned (Bouezmarni et Debbaut, 2006). Conversely, the Strassen marl is revealed in the landscape by the appearance of numerous springs at its upper limit. These features tend to reveal a relative permeability of km3 in comparison to Li3.
- Second, the arrangement of the sandstones and the marls layers is reversed. On HM, the marls of Strassen consist in the plateau, i.e the top of the relief, while the midlle Keuper marls on KOE consists in the riparian zone. Because of it respective location, the marls layer on HM will be more sensitive to dry condition.

We make several changes in section 2.1 and table 1 to be more accurate on that point. When speaking of km3 marls, we said "variagated marls". On table 1, instead of speaking of the first and second geology in terms of size area, we speak about "lower geology" and "overhanging geology".

(REF: Bouezmani, Debbaut. Carte géologique de Wallonie, Tintigny, Etalle, Notice explicative. Université de Liège, 2006.)

Figure 1: Please add contour lines in the figure. Addition of the map of slope angle or topographic index is also helpful to understand the topographic features of study sites. As soil moisture was observed at the points of raingauges, "Raingauges" should be changed to a phrase such as "Raingauges and soil moisture observations".
The suggestions have been applied to improve the catchment's presentation: the topographic contour lines have been added to the figure and the legend has been modified. Furthermore a map which integrates the slope display and the heights above nearest drainage map. The latter information highlights the buffering area on KOE catchment and the closeness of the slopes to the main river network on HM catchment.

Table 1: What is the difference in river width and riparian area between the catchment sections? Similar area, elevation range, and slope range does not necessarily mean that the two catchments have similar topography.
We added a statistic about the height above nearest drainage that characterizes the difference in surface area of the riparian zone between both catchments.

L139: Does "deeply cut" mean that valley was deeper in HM section than KOE catchment? If so, this topographic characteristic should be quantitatively presented.

The surface area close to the river network in terms of elevation is 3 times smaller on HM catchment in comparison to the KOE one. This figure presented now in table 1 supports the qualification of “deeply cut” apply to the river network in HM catchment. In an illustrative way, the slope display in figure 1 now illustrates this characterization.

L212 “net rainfall amount after infiltration”: How did you determine the amount of loss (i.e., total rainfall – net rainfall)?
This is based on the observation of the average runoff coefficient: the net rainfall volume is equal to the discharge volume observed. Assuming a constant RC along each event the net rainfall at each time t_i is : $R_{net}(t_i) = RC^* R_{total}(t_i)$ (this is already described in equation 7)

Figure 7: Hydrographs in Heffingen catchment were very clearly different from those in Koedange and Medernech catchments. Why was the runoff delayed in Heffingen? I think the comparison between various catchments may provide clearer insights into runoff mechanism than the comparison between only the two catchments.
See figure S1 in supplementary materials. The hydrographs in Heffingen catchments are different from the Medernach catchments but not from the Koedange catchment. With the latter there is only a delay and a spread, which is an expected behavior when looking at a downstream measurement. This has been already described lines 187-189 (first submission): *“The headwaters (as expressed through the Koedange and Heffingen stream gauges) consistently triggered rather attenuated hydrological responses. Further downstream, the stream gauges located downstream of Larochette exhibited a much more responsive behavioural pattern.”*

Figure 8: Please add the results in runoff coefficients of each event. I also recommend the authors to conduct statistical in order to examine whether the difference in the TTD values between the two catchments was significant in each event.
We are not sure to understand your suggestion. The runoff coefficient is part of the observation and not calculated using the model. That’s why it does not appear in the figure. Or do you mean to range event according to the runoff coefficients?

L418: I want the authors to show the location of “large flat terrain” in Figure 1 based on topographic map with the spatial distribution of slope angle or topographic index. It would also be helpful if the area of this flat terrain can be shown in Table 1.
This is now visible through the heights above nearest drainage statistics (table 1 and map (figure 1 right).

L421-422: Why does the limited permeability of underlying bedrock lead to large storage capacity? I think permeable bedrock has larger storage capacity because groundwater is stored within weathered layer or fractures in bedrock.
It is rather a misunderstanding of the connection between the two parts of the sentence. We reworded the sentence: *“This almost 100% marly (km3) catchment has a rather large storage capacity,* ***despite of*** *~~considering~~ the limited permeability of its underlying bedrock.”*

L490-510: Although only the effect of litter layer is discussed, discussion about evapotranspiration is also necessary for the impact of the vegetation because LAI directly affects it.
As focusing on the time transfer distribution variability and not on the runoff coefficient ones, we do not discuss on evapotranspiration.

L513-514: Differences in geological substrates and landscape features between the catchments should be more clearly presented throughout the manuscript.
Thanks to your suggestion, we added several description that - we hope – will help the catchment’s characterization.

L516-517: There is no evidence that main runoff source in the KOE catchment was groundwater and deep soil water.
We reworded this part of the conclusion*: " In the KOE catchment, the water transfer get a seasonal variation disconnected from precipitation characteristics (except for one summer event)."*

L521-523, L529-530: It seems that the authors attribute the difference in runoff characteristics to topography in slope and riparian zone rather than geology. If so, stories focusing on the topographic features may be better.
As said before, changes have been made in line with this suggestion.

L531-532: Runoff coefficients were one of magnitude smaller in summer than in winter (L439). I think this result indicates that runoff during dry summer season had small risk of flooding even if the rapid flows occurred. Both results of runoff coefficients and TTDs should be considered to provide conclusion for flood risk management.
Our conclusion deals with the specific context of **flash** floods. In this context, the short timing and the peak magnitude rather – than the runoff coefficient – are of first importance. As a prove, the 2016 and 2018 **flash** flood events get not so high runoff coefficients (14% and 20%), while their flood peaks are among the three highest recorded.

Technical corrections
L24: "Another catchment" would be better than "The HM section" because I could not understand this is the name of catchment when I firstly read Abstract.
The HM acronym presentation has been added line 18.

L100: Does "mean summer and winter runoff" mean baseflow runoff in summer and winter?
It means "runoff coefficient". This has been specified.

L111-114: I think these sentences are unnecessary. ok.

L154-156: The order of Figures 2 and 3 is reversed. The caption has been corrected.

L177: Although it was written that "the rainfall amount had to exceed 10 mm", there is an event with the rainfall amount of 9.8 mm (Table 2).
The threshold of 10 mm was actually applied on the raingauge observation average that make a slight difference with the rainfall amount on Medernach catchment (obtained with weights on raingauge according to the Thiessen polygons).
We changed the description: *"[…] according to the following criteria: i) the rainfall amount average on the 4 raingauges had to exceed 10 mm, [...]"*

Figures 6 and 7: Please check if the date of (c) is true. Were they really different between the two figures?
The dates of (6c, 7c) are right. The figure 6 shows the result of the KOE catchment simulation, and the figure 7 shows the same but for Medernach. On panel, We choose two different events, because it seems to us more illustrative of the model "weaknesses" on each catchment respectively.

L308-310: Were these values the ranges in both catchments?
Sorry but we do not understand the question.

Figure 10: The color of SWC20 and RC may be wrong.

This has been changed.