<u>Revision notes on the paper by Grandjouan et al. « An original approach combining biogeochemical signatures and a mixing model to discriminate spatial runoff-generating sources in a 2 peri-urban catchment ».</u>

Reviewer 2

We thank the reviewer for the time dedicated into reviewing our manuscript and for the constructive feedback provided. Our detailed responses and the corresponding changes to the manuscript are presented below. Reviewer comments are shown in black, our response in blue, and proposed modifications to the manuscript in italic blue.

1. This an interesting study putting forward the value of rich sampling of geochemical parameters in streamflow to trace back the spatial source of streamflow in a small peri-urban catchment with heterogeneous land use, in various hydrometeorological conditions. Such a data-based approach is valuable and could undoubtedly contribute to methodological advances and policy-level demands linked to water quality and runoff management in (peri-)urban areas.

I think the study has potential for publication in HESS. However, I have some concerns about the way some parts of results and discussion is carried out, which sometimes makes it hard to follow what the authors are getting at, although the overall message seems very clear at other times. I therefore suggest a major revision.

We sincerely thank the reviewer for the positive evaluation of our work. We also greatly appreciate the constructive comments. The reviewer offered relevant suggestions for carefully revising the manuscript and making the narrative more concise and easier to follow.

General comments:

- **2.** In the methodology and first part of the results, I am not sure to grasp why the authors present the full set of parameters they have measured (52) and even describe their patterns in the typology of source (sect. 3.1), since a significant portion of them is discarded in the subsequent signature build up used for source identification. In my view it introduces some confusion and the unnecessarily lengthen the manuscript. I provide some suggestions in the specific comments.

  Please see response to comment n°11.

- **3.** I have mixed feelings about the robustness the end member mixing analysis conducted here. The way Section 4.1 seems to cast serious doubts on the ability to distinguish sources, especially regarding sewer system, urban & road runoff, and surface runoff. I appreciate the transparency of the authors in pointing at the limits of their approach, but as it stands I am left wondering whether the results presented can then be interpreted into a perceptual model, as is done in Sect 4.2. I think it owes in great part to the somewhat concatenated aspect the writing sometimes takes in the Discussion, where (sometimes strong) statements/ideas and comparison with the literature are given in one sentence without really connecting with a conclusion. This concerns both Sect 4.1 and 4.2. Examples include:
  - L480-481: where wastewater source is concluded to be indistinguishable from URB without much nuance or discussion of the implications

    We made the choice to sample water from the sewer system during rainfall events, in order to characterize the biogeochemical signature of the water transferred to streamwater during overflows. However, this led to the sampling of a heterogeneous sample, characterized by a mixture of wastewater and urban and road surface runoff. This sampling strategy had a strong influence on the contributions estimated for the SEW and

URB sources. The mixing model faces a first limitation as it is unable to distinguish wastewater alone from urban and road surface runoff. Indeed, the SEW signature may have been diluted and influenced by the URB signature, which already showed a variable biogeochemical composition. As a consequence, we may have overestimated the SEW contributions during events. Moreover, the results for dry weather conditions are less reliable, as only wastewater is released through leaks in the sewer system. Ideally, we should have built the wastewater signature using samples collected from the sewer system under dry weather conditions, to better distinguish URB contributions from wastewater.

- L495-501 where the robustness of using precipitation as a proxy for runoff signature is questioned, then a reference is used to say that it is sometimes valid, without stating whether the referenced study was done in a similar context, making the overall argument somewhat weak.

  We acknowledge that the hypothesis that surface runoff represents a strong assumption, which was necessary due to the lack of data on direct surface runoff composition outside urban areas. It appeared as the most consistent approach in order to apply the mixing model. We assumed that surface runoff does not have enough time to acquire significant biogeochemical elements from the soil it flows over. We understand that such hypothesis does not take into account the enrichment of water by soil leaching, as these waters can quickly accumulate elements (Langlois & Mehuys, 2003). Yet, Fröhlich et al. (2008) conducted a similar study in the Dill Catchment (Germany), aimed at identifying runoff sources, including wastewater, groundwater and stormwater flow, in which they regrouped surface and subsurface runoff. To do this, they sampled streamwater from the outputs of sub-catchments characterized by specific geological formations, during baseflow and hydrological events. They showed that the geochemical composition of stormflow was similar to the composition of precipitation, characterised by low-mineralization. Their results suggest the predominant contribution of low-mineralized waters for several events, which support the use of the composition of rain to represent the quick surface runoff source, in cases where runoff water could not be sampled. In any case, our study could benefit from a proper sampling of quick surface runoff in order to better estimate their contributions to streamwater. Several studies analyzed direct surface runoff water collected on soil surface during hydrological events (e.g. Le et al., 2022; Omogbehin & Oluwatimilehin, 2022), but these studies are often conducted in tropical areas, where direct surface runoff often occurs out-side of urban areas. Such sampling appears to be difficult in temperate areas, with less intensive rain-falls.

  In any case, the assumption of a composition of surface runoff close to the composition of precipitation may lead to an underestimation of the quick surface runoff contribution when applying the mixing model for hydrological events. Therefore, we should analyze the model outputs for hydrological events while explicitly considering the potential influence of this assumption.

- L530-536: where the authors start from describing a fill-and-spill (which was not only described in the Panola catchment, see Spence & Woo, 2003; McDonnell et al., 2021) to arguing that "quick surface runoff is favored in grasslands" (as compared to what?), based on a lower water demand than forest and crops (an assertion which is not always true, see for example Houspanossian et al., 2023), and then a last sentence about dry soils...connected to the previous statement? What is the conclusion then?

  In this section, our goal was to explain the variable contributions from grasslands and agricultural areas by the highly variable thickness of the saprolite horizon downwards from forest - 1 to 20 m (Goutaland, 2009). We thus explained the absence of runoff for the GRA and AGR sources under low flow conditions in dry weather by the existence of throughs at the saprolite-gneiss interface in which water can be stored and released discontinuously. As mentioned, this process was described as “fill-and-spill” by McDonnell et al. (2021), and observed in the Panola catchment by Tromp-van Meerveld & McDonnell (2006), and in the Pocket lake catchment by Spence & Woo (2003), both being characterised by a similar crystalline bedrock. Both studies showed that the generation of subsurface and surface flow in this context can be delayed, as it requires to meet sufficient rainfall amount to increase water storage at the soil-bedrock boundary. When these conditions were not observed, Spence & Woo (2003) and Tromp-van Meerveld & McDonnell (2006) noticed intermittent flow, which is similar to what we observed at the Mercier and Ratier catchments. Indeed, contribution from agricultural lands are low or absent during summer storm events, and major during major events, when rainfall amounts are sufficient. However, grasslands showed quicker and more frequent responses under storm conditions. This difference may be linked to lower interception by vegetation, shallower root systems, and reduced water demand in grasslands compared to forests or crops (Madani et al., 2017; Robinson & Dupeyrat, 2005). Indeed, Robinson & Dupeyrat (2005) showed that forest display a higher total evapotranspiration on a yearly average, thanks to interception evaporation, although grassland display higher rate of evapotranspiration during dry summer conditions.

I would thus suggest to rework the Discussion thoroughly, to improve fluidity and clarity.

We acknowledge the reviewer for this constructive feedback on the Discussion section, as we understand that several ideas could have been developed further. In addition to the responses to the examples mentioned above, we understand that we could reinforce the whole Discussion in this way.

- **4.** Also, Grandjouan et al. (2023) considered three subsurface sources in their analysis : colluvium aquifer, but also a fractured aquifer and saprolite, with significant, seasonally-variable contribution from all of them. I am guessing that the latter two (fractured and saprolite) being widespread over the catchment (as opposed to the colluvium), the more spatialized analysis proposed in this study implicitly includes a spatialization of fractured and saprolite contribution to streamflow embedded in FOR-1, FOR-2, AGR and GRA sources. Yet, it does not seem to me that this link between the sources identified in Grandjouan et al. (2023) and the present study is clearly made, although this previous is cited several times especially in Sect 4.2. This link should thus be more thoroughly put forward in the Discussion, to better show how this study builds up on Grandjouan et al. (2023) and the associated benefits.

We understand that the link between this previous study and our paper may seem unclear. In their study, Grandjouan et al. (2023) built an initial perceptual hydrological model of the Ratier catchment, describing the general hydrological behaviour of the catchment and the main contributions to streamflow. That model identified three main sources including colluvium groundwater, fractured gneiss groundwater and the saprolite layer. They observed positive correlations between discharge and saprolite contribution, and negative correlations between discharge and gneiss groundwater contribution. However, they also showed unclear boundaries between both contributions, and suggested that the hydrological variability could not be fully explained by the vertical contributions to streamflow (i.e. saprolite, gneiss and colluvium aquifer). They suggested that a spatial decomposition of streamflow, according to land use, could give more insight on the hydrological behaviour of the Ratier catchment.

In our study, we used the extensive dataset obtained with the source sampling to improve the initial representation of the catchment hydrological behaviour by Grandjouan et al. (2023). Our results showed different hydrological behaviour linked to land use, that could not be explained only by the geological characteristics. For example, the forest source is characterised by a shallow or absent saprolite depth with groundwater playing as the dominant hydrological process. Contributions from forest is stable at dry weather, and remain minor during summer storms, due to strong canopy interception and high evapotranspiration in this part of the catchment. For grasslands and agricultural lands, generation of runoff is generally driven by a fill-and-spill mechanism within the saprolite layer, producing intermittent sub-surface contributions. Their contributions therefore strongly depend on the topography of the saprolite-gneiss boundary. Although grasslands and agricultural lands share a similar geological context, they show significant differences in their contributions, partly because of higher crop water demand for the agricultural lands.

Specific comments:

5. L139-140: can the authors justify the choice of 3mm-threshold to distinguish dry and wet periods? It seems quite a low threshold, and 3 mm in 5 days is not intuitively "wet" in my view.

We chose to distinguish dry and wet periods according to the median of daily rainfall recorded for the period 2011-2023 at the Pollionnay station. We ackowledge that the term "wet" could be confusing but the objective was to separate rainfall data between two equal groups, based on cumulative antecedent rain.

6. L148- : Braud et al (2018) is missing from the reference list, so it's hard to know what the original method was.

We thank the reviewer for point out this omission. The reference will be added to the reference list.

Braud, I. (2008). *Analyse des données pluie-débit sur les sous-bassins du Mercier et de la Chaudanne. Période1997-2007. Données de base, critique de données pluies, programmes Fortran et utilisation de R pour cetteanalyse, résultats.* Irstea Lyon-Villeurbanne.

7. L150-151: Can the authors give more details on the configuration of HCA used here, e.g. how the optimal number of classes was found, was is done using absolute or relative concentrations, etc.

In order to build our Hierarchical Clustering Analysis (HCA), we identified an optimal number of three classes using the "elbow" method (Thorndike, 1953) and associated them to different types of events: small winter events, summer storm events and major events. We used absolute centred and scaled concentration as input data to the HCA.

8. L156-157: please provide a reference and/or justification for this API, notably the k=0.8 factor. As it stands, it seems somewhat arbitrary.

We understand that this value seems arbitrary, as methods of optimization exist to estimate more credible factors (Li et al., 2021). However, these methods are complexe and time-consuming. The value of the k multiplying factor was chosen from the work of Sarrazin (2012) on the same catchment we studied. It is a value generally used in the calculation of API (Dingman, 2015; Viessman & Lewis, 2003).

9. L157-158: I am not sure to understand the reason behind choosing two events per class, instead of 3, 4...please develop if there was a specific reason.

We targeted a minimum of 2 events per class to study the variability of the results within each class identified. However, we faced some difficulties to sample the expected events. First, the lack of rain between 2021 and 2023 led to less hydrological events: we measured around 500 mm of rain for the year 2022 at the Pollionnay station compared to 750 mm for the mean precipitation rate for the 2010-2023 period. Then, we faced some difficulty to install and program the automatic samplers at the right times. Weather forecasts were not always optimal and the samplers could not be left on site for long time periods due to vandalism issues. Technical issues on these samplers prevented us to sample the March 2019 and 22 June 2022 at Ratier and Mercier, respectively. Finally, the laboratory had constraints to be able to process the samples, which restricted the number of working days for which we could sample the events. Three or 4 events may have been better to study intra-class variability, but the objective of two events per class seemed reachable.

10. L223: is it hierarchical classification, or clustering?

Authors meant "Hierarchical Clustering Analysis". The error will be corrected in the text.

11. L233-238: since a part of parameters can be eliminated from the target signatures even before measuring them, just by their nature (non-additive, undefined relations with abiotic factor, too reactive), I do not see the point of initially listing them in the previous sections nor describe how their are measured, as they do not serve the current study. I would therefore recommend to move this reasoning earlier in the methods and remove the corresponding parameters' (UV*, Fe, phosphates, etc.) protocol from the main text and even from Table 3. This may streamline the manuscript. The case is different for the parameters removed as per screeing test described in L238-244, as I understand that measurement was necessary before screening out these parameters.

We thank the reviewer for pointing out that aspect of the paper, as we understand that numerous parameters that were initially analysed were not useful for the source signature building. We chose to analyse a wide set of parameters, including those which did not meet the criteria for mixing model (conservativity, additivity), as we considered that they could provide useful information for characterising the source or validating the signature building. However, it is true that only a few of them were ultimately used. We acknowledge that we could retain only the parameter that meet the additivity and conservativity criteria, before applying the range-test. This is the case for 33 biogeochemical parameters, including 6 major ions, silica, 15 trace metal elements, Dissolved Organic Carbon (DOC) concentrations, 2 Ultra Violet-Visible (UV-Vis) indicators and 5 High Pressure Size Exclusion Chromatography (HPSEC) indicators. The parameters that we could discard prior to the study include NO_3^-, NO_2^-, PO_4^{3-}, NH_4^+, Fe, Mn, DTN, E2:E3, E2:E4, E3:E4, E4:E6, SR, SUVA, Mw-254, disp 254 and the microbial indicators. However, we could retain the 2 host-specific microbial DNA targets *HF183* and *rum-2-bac* as they can trace the origin of water in case of fecal contamination. These makers were used to evaluate the biogeochemical signature built and the estimations of the contributions calculated by the mixing model. The other 5 microbial parameters (G16S, BTT, integron class1 and class 2, BTS) were not clearly used in the paper and could probably be removed. The full set of 55 biogeochemical parameters is available at: https://entrepot.recherche.data.gouv.fr/dataverse/chypster/.

12. Figure3: this is a figure with a lot of information. I would suggest to keep the parameters effectively used for later source identification (which are not even put put in the axis/caption at present), which would make it much more readable and easier to interpret the "relevant" patterns and clustering for the later analysis. Perhaps the "full" figure can go to the Supplementary materials.

As suggested, we can remove from this figure the parameters that we discarded on the basis of their nature (see response to comment n°11). However, we would keep the remaining 35 parameters for building this heatmap, including the two microbial parameters *HF183* and *rum-2-bac*. Indeed, we built this figure to give a global visualisation of the biogeochemical composition of the source water, and to better characterise and identify these sources. This heatmap indeed shows that the BOU and VRY samples from the two forest sample points do not follow the same biogeochemical trend. We therefore chose to consider two different sources linked to a forest land use. It was only based on these results that we were able to apply the reductionist tracer approach and select the 15 definitive tracers that were used in the mixing model. The new figure that will be used in the revised paper is presented below.

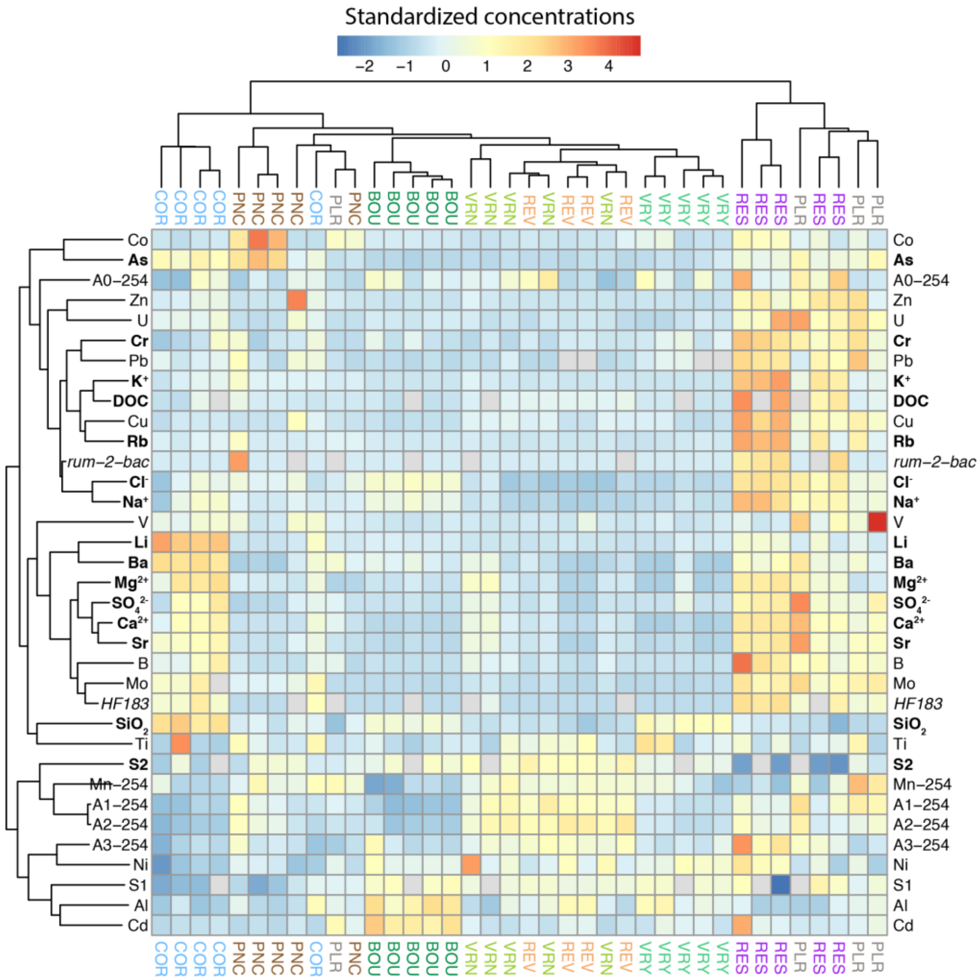

Figure 1 – Heatmap representation of the median concentrations of the biogeochemical parameters in source samples. Standardised concentrations are shown in a range of colours from blue for negative values to red for positive values. Positive values represent high concentrations for a specific parameter and source sample, compared with the other samples. Negative values represent low concentrations for a specific parameter and source sample, compared with the other samples. Biogeochemical parameters and source samples are classified into groups based on Hierarchical Classification Analysis. Quick surface runoff (SUR) was not considered as all biogeochemical parameters were not available for this source.

13. Fig. 3 & 5: how were the concentrations standardized? It seems like an important point of methodology to interpret the results, all the more that the range are quite different between the two figures, hence the method?

The two figures were built to highlight different aspects of the dataset, which explains the use of distinct standardization methods. For the heatmap, we used a centring and scaling method to emphasize relative differences among tracers across sources. This method allows to compare tracers with different units on a common scale, and to visualize contrasting patterns of over- or under-representation. In contrast, the radar plot was intended to represent the relative magnitude of each tracer within a given source, and to compare overall shapes between sources. In this case, we preferred a min-max normalization (scaling between 0 and 1), as it preserves proportional relationships among tracers and facilitates visual interpretation of source-specific signatures.

14. Fig. 6 & 7: It may be interesting to add the corresponding figure with absolute flow amounts (perhaps in the Supplementary Materials) to put these results in perspective of m3, as is already done somewhere in Sect. 3.3.2 (e.g. for SEW contributions in m3 during events), but without graphics.

We thank the reviewer for this suggestion, which could indeed facilitate the interpretation of the results, and present them in a more concrete way. We will add the corresponding figures with absolute flow amount in appendix. These figures are presented below. Figure A1 represents the equivalent contributions in daily volumes (in m^3) that we calculated considering that the discharge measured at the time of sampling was representative of the daily discharge. Figure A2 represents the total volume contribution that we calculated based on the relative contributions for each source and the total flow in m^3.

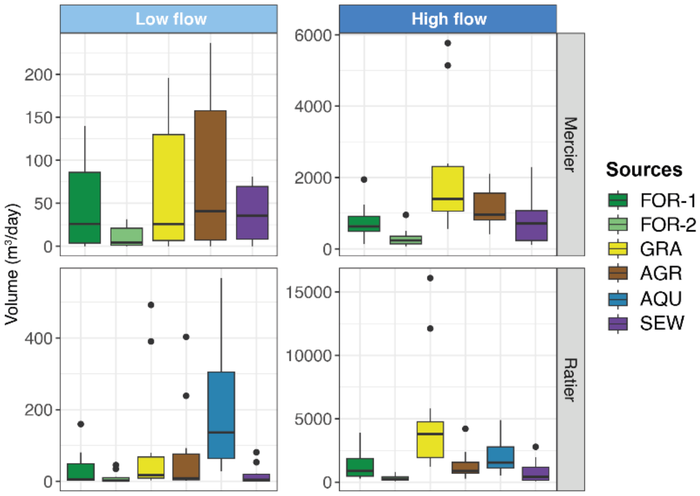

Figure A1 – Daily volume contributions in m^3 estimated for dry weather samples by the application of a biogeochemical decomposition using a Bayesian approach for the Mercier and Ratier catchments. Contributions in terms of volume were calculated based on the relative contributions for each source and the total flow for each sampled day in m^3. Boxplots represent the median contribution, interquartile range (1st and 3rd quartiles), minimum and maximum values. Low flow samples correspond to a mean daily discharge lower than 20 L/s and high flow samples to a mean daily discharge higher than 20 L/s.

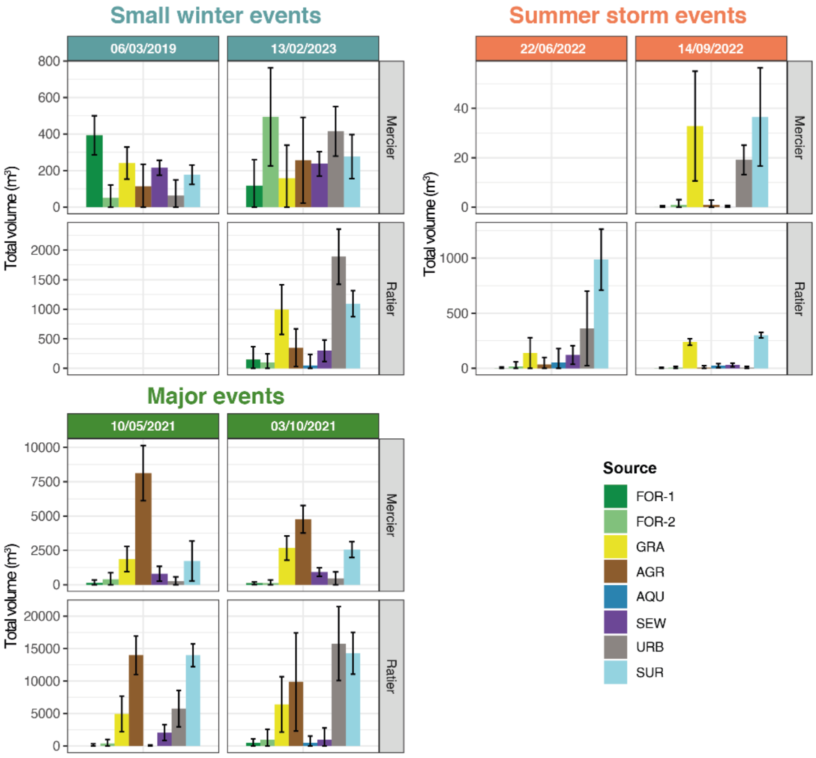

Figure A2 – Total volume contributions to the hydrological events sampled between March 2019 and March 2023 at the outlets of the Mercier and Ratier catchments. Contributions in terms of volume were calculated based on the relative contributions from each source and the total flow in m^3. The contributions correspond to the mean of the results obtained for each samples decomposition by the Bayesian mixing model approach. The error bars correspond to the mean of the standard deviation calculated from the sum of the squares of the deviation. The events of 6 March 2019 at the Ratier station and 22 June 2022 at the Mercier station were not collected.

We could use these figures to put the relative contributions estimated by the mixing model in perspective with their corresponding m^3 volume flow. For example, for dry weather, the median volume contributions of wastewater for the Mercier can be estimated to 30 m^3/day at low flow and 800 m^3/day at high flow. As a comparison, Dubois et al. (2022) estimated the average daily wastewater flow from a French household around 0.311 m3/day, and Aussel et al., (2004) the wastewater discharge per inhabitants in France around 0.2 m3/day. Wastewater contribution to the Mercier stream therefore represent the equivalent of a contribution of 100 households or 150 inhabitants. For hydrological events, we estimated SEW volume flows around 900 and 2 000 m^3 at the Mercier and Ratier stations during the May 2021 event, and around 1 000 m^3 for both stations during the October 2021 event. Such volume of wastewater transferred to the stream are equivalent to the mean daily wastewater discharge for 3 000 to 6 500 French households, or for 5 000 to 10 000 inhabitants (Aussel et al., 2004; Dubois et al., 2022).

15. Fig. 7: Also, is it stated somewhere why 6 March 2019 and 22 June 2022 were not sampled at Ratier and Mercier, respectively ?

The March 2019 and June 2022 events were not sampled at the Ratier and Mercier station, respectively, due to technical issues on the automatic samplers. See response to comment n°9.

16. L389-391: an estimate of the rainfall amount leading to sewer overflow would be informative.

A quantitative rainfall amount needed to cause a sewer overflow is difficult to estimate, as it depends on the volume of wastewater present in the sewer system prior to rainfall, on the intensity of the rain and on the settings of the sewer overflow device itself that can evolve with time. However, field observations and discussions with the sewage network managers suggest that the sewer system is clearly undersized, and that overflows often occur with low rainfall amount. We indeed observed indications of overflow during the small winter events with less than 10 mm of cumulative precipitations.

17. L409-411: This sentence seems more of a discussion point. I would suggest to move it to Sect. 4, possibly developing this interesting point, especially if local or comparable references on rainfall spatial variability and variable sources activation are available.

We acknowledge that we could improve and move the whole paragraph to the Discussion section, as the comparison between the Mercier and Ratier results is indeed an interesting point that could be discussed in relation with comparable references. The spatial rainfall variability can have, indeed, a strong influence on the URB contributions. This phenomena is particularly relevant during convective summer storm events, where precipitations are localised and lead to quick response of urban areas, as showed by Kermadi et al. (2012) for the Yzeron catchment (which includes the Ratier catchment). The influence of rainfall spatial distribution on hydrological response in urban areas is undergoing increasing study, especially through hydrological modelling (Cristiano et al., 2017). Such studies encourage the use of high spatial resolution radar weather radar images for studying rainfall spatial variability in small peri-urban catchments, although this remains uncommon (Emmanuel et al., 2012).

18. L412-414: Again, this seems like a Discussion item, but as it may already be somewhere else, I'd recommend to remove this sentence.

As suggested, we will removie this sentence, as the general idea was already clarified in the Discussion part.

19. Fig. 8: this is a nice figure, combining important information. Yet, it is hard to tease out the relative contribution of sources since the "total" bars sometimes drastically change in height. I wonder if as companion, supplementary figure wouldn't be useful to visualize the temporal changes described in Sect 3.3.3, with a 0-100% y-axis for stacked contributions. If technically possible, superimposing the discharge time series on a econd y-axis (same as now) would be nice.

As suggested by the reviewer, we will show the relative contributions with a 0-100% y-axis. We preferred to enriched the initial figure in order to see both representations at the same time. We therefore divided the initial figure into 3 figures for each class of event:

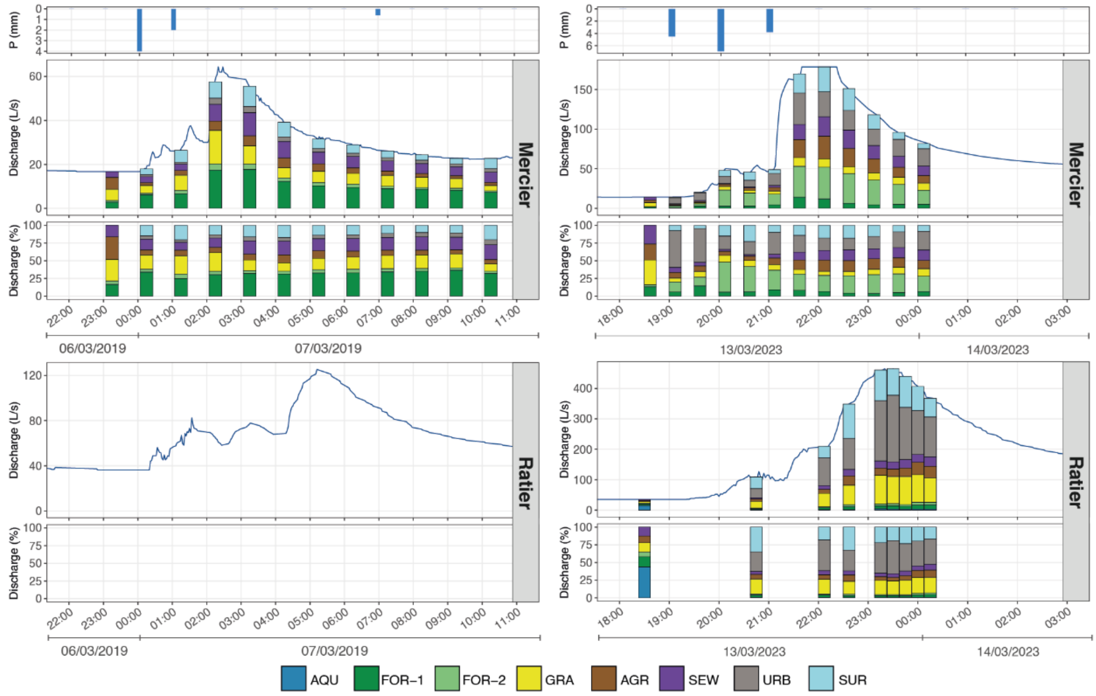

Figure 2 – Precipitation and hydrograph separation results for the sampled events at the Mercier and Ratier stations for the small winter events of March 2021 and March 2023. The results are shown according to the total outlet discharge and according to the relative contributions of discharge. The size of the bars for the upper plots corresponds to the instantaneous discharge associated to each sample.

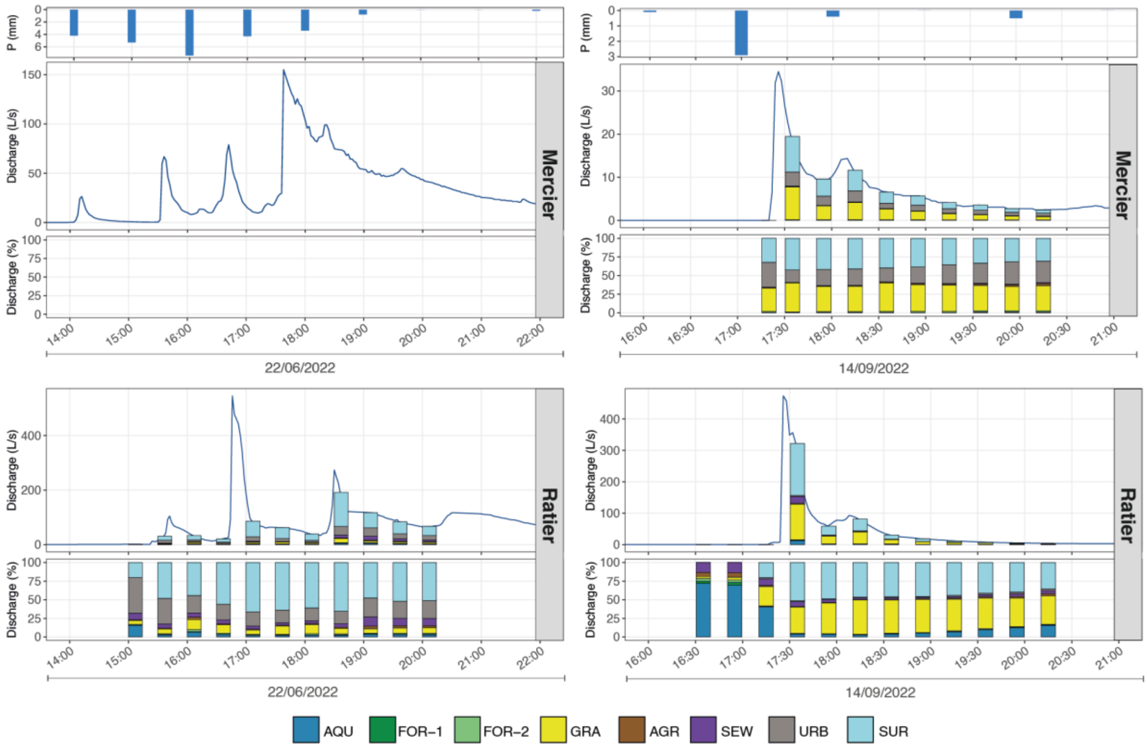

Figure 3 – Precipitation and hydrograph separation results for the sampled events at the Mercier and Ratier stations for the summer storm events of June 2022 and September 2022. The results are shown according to the total outlet discharge and according to the relative contributions of discharge. The size of the bars for the upper plots corresponds to the instantaneous discharge associated to each sample.

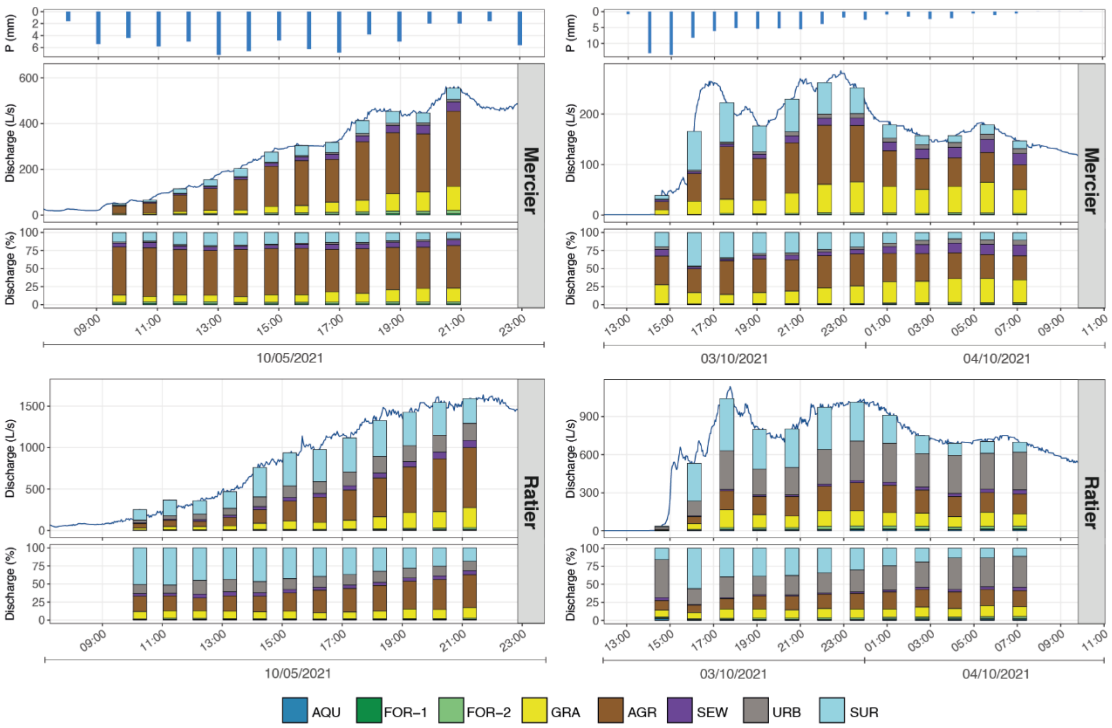

Figure 4 – Precipitation and hydrograph separation results for the sampled events at the Mercier and Ratier stations for the major events of May 2021 and October 2021. The results are shown according to the total outlet discharge and according to the relative contributions of discharge. The size of the bars for the upper plots corresponds to the instantaneous discharge associated to each sample.

20. L450-452: these absolute numbers are very informative, such comparison would have welcomed already in Sect 3.3.2 circa L401-402.

We took into account this suggestion and understand that we could use such comparison earlier in the Results section. We could used two references that estimated the wastewater daily discharge per French household (Dubois et al., 2022) and per inhabitant (Aussel et al., 2004). These estimations could allow us to make concrete comparisons. See response to comment n°14.

21. L456-461 : the statement are phrase in a very general way, this introductory paragraph may benefit from (subtle) reminder it applies to the current study: "In the studied catchments, …" or "In this study, …"

Following the reviewer's suggestion, we could add reminder to the text to clarify this introductory paragraph:

[…] In our study, these signatures seem to have been particularly well defined for forests and grasslands. […] Our results question the representativeness of these signatures and the initial assumptions on which the identification and sampling of these sources were based.

22. L481-485: this is an interesting point, it seems to me that this research group jointly analyzed stable isotopes and other chemical signatures in their end member analysis. At any events, I am not sure to understand what the authors are here concluding, please clarify.

The point of this section was to highlight the use of stable isotopic tracers, while showing that they could have a great potential when used jointly with geochemical tracers. These are the conclusions from Kuhlemann et al. (2021), in which they used stable isotopes jointly with conductivity to estimate the contribution of wastewater in the Erpe peri-urban catchment (Germany), in an Bayesian mixing model (using MixSIAR). As they faced high variability in source composition, due to the similarities in concentrations between the composition of wastewater and other runoff sources, we suggested that more precise tracers could have been used.

23. Fig. 9: what is the yellow storage? Storage amplitude between dry and wet conditions?

The yellow storage initially represented the storage amplitude between low water and high water. We acknowledge that this could be confusing as we only precise low and high water at dry weather. Such distinction is thus only relevant under dry weather conditions. We will therefore modify the figure to represent a single water storage in the case of events, while distinguishing low from high water in the case of dry weather. We will also modify the storage colours for clarity, and the water levels of each stream were revised to match the represented hydro-meteorological conditions and the stream responses. The new figure that could be added in the revised paper is presented below.

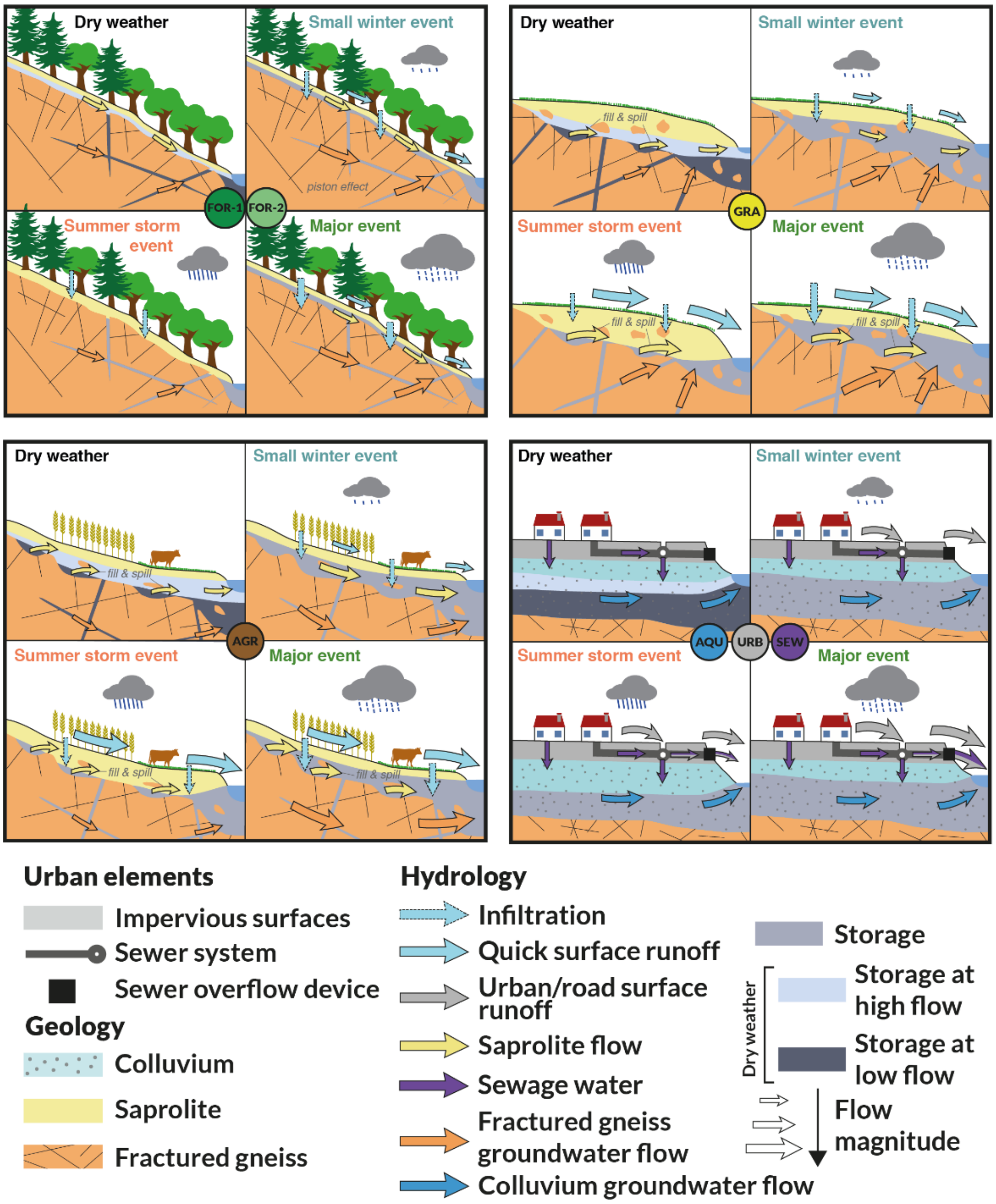

Figure 5 – Improved perceptual model of the Ratier catchment, initially build by Grandjouan et al. (2023). Main contributions, estimated by the mixing model, are illustrated according to the nature of the source and the four hydro-meteorological conditions studied, including dry weather, small winter event, summer storm event, major event. FOR : forest; GRA : grassland; AGR : agricultural; AQU : aquifer; URB : urban and road surface runoff; SEW : wastewater.

24. L537-539: Having the same geological features does not necessarily result in having the same hydrological behavior ; topography, land use (and thus water demand and timing). The next sentence seems in contradiction. I think this part needs some rephrasing.

We could rephrase the text to clarify our idea as follows:

[…]This difference may be linked to lower interception by vegetation, shallower root systems, and reduced water demand in grasslands compared to forests or crops (Madani et al., 2017; Robinson & Dupeyrat, 2005).

25. L555-567: This reads as a summary rather than a conclusion ; this is more the function of the abstract, so I would suggest to significantly reduce or remove this part.

We acknowledge that this paragraph was too summary. We could remove the whole paragraph and add broader perspectives at the end of the conclusion:

The objective of this study was to identify runoff-generating sources in a small peri-urban catchment, and estimate their contribution to streamwater with a mixing model based on a biogeochemical dataset comprised of classical and original tracers. This approach showed the potential of the use of biogeochemical tracers to perform a spatial decomposition of water, based on the physical characteristics of a catchment, in addition to a more traditional vertical decomposition. This study also showed the need for precise methods to identify the runoff-generating sources and their biogeochemical signatures.

[…]

More broadly, the application of mixing models in relation to land use remains relatively unexplored in the literature. This study highlights the potential of such an approach when incorporating biogeochemical parameters and highlights the need for further research in this direction.

This work illustrates the broader potential of mixing models to identified the spatial origin of streamflow and improve our understanding of catchment hydrological behaviour. Such approaches could provide valuable insights for validating spatially distributed hydrological models, which often face difficulties in adequately representing source contributions. More generally, combining mixing models with land-use and hydro-meteorological data may help to better anticipate the impacts of land management, urbanisation, or climate change on runoff-generation processes. Future research should therefore focus on integrating tracer-based source characterisation with modelling frameworks, to improve both process representation and predictive capacity in heterogeneous catchments.

Technical comments

26. L394: from Fig. 7, I read a 32%-contribution of URB to the event on Oct '21 at Ratier, not 22% as in the text.

The error will be corrected in the revised paper.

27. Figure 8: the color code is the same as Fig. 6-7, but not the legend order, it would help to homogeneize.

Fig. 4 to 8 will be modified in order to homogenise the order of sources in all figures and to represent them in a logical vertical order: the groundwater source (AQU) first (on the far left or at the bottom, depending on the figures), followed by sources mainly marled by land-use (FOR-1, FOR-2, GRA and

AGR), the sewage system (SEW) and finally the two surface runoff sources (on the far right or at the top).

References

Aussel, H., Bâcle, C. L., & Dornier, G. (2004). *Le traitement des eaux usées*. INRS.

Cristiano, E., ten Veldhuis, M.-C., & van de Giesen, N. (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. *Hydrology and Earth System Sciences*, *21*(7), 3859-3878. https://doi.org/10.5194/hess-21-3859-2017

Dingman, S. L. (2015). *Physical Hydrology : Third Edition*. Waveland Press.

Dubois, V., Falipou, E., & Boutin, C. (2022). Quantification and qualification of the urban domestic pollution discharged per household and per resident. *Water Science and Technology*, *85*(5), 1484-1499. https://doi.org/10.2166/wst.2022.064

Emmanuel, I., Andrieu, H., Flahaut, B., & Furusho, C. (2012). *Influence of rainfall spatial variability on rainfall runoff modelling : Case study on a peri-urban catchment*.

Fröhlich, H. L., Breuer, L., Frede, H.-G., Huisman, J. A., & Vaché, K. B. (2008). Water source characterization through spatiotemporal patterns of major, minor and trace element stream concentrations in a complex, mesoscale German catchment. *Hydrological Processes*, *22*(12), 2028-2043. https://doi.org/10.1002/hyp.6804

Goutaland, D. (2009). *Programme ANR AVuPUR - Prospection géophysique par panneau éléctrique de trois parcelles d'un sous-bassin versant de l'Yzeron*. CETE de Lyon.

Grandjouan, O., Branger, F., Masson, M., Cournoyer, B., & Coquery, M. (2023). Identification and estimation of hydrological contributions in a mixed land-use catchment based on a simple biogeochemical and hydro-meteorological dataset. *Hydrological Processes*, *37*(12), e15035. https://doi.org/10.1002/hyp.15035

Kermadi, S., Braud, I., Jacqueminet, C., Branger, F., Renard, F., & Michel, K. (2012). Quels liens entre climatologie, occupation des sols et inondations dans le bassin versant de l'Yzeron (ouest Lyonnais) ? Apport de l'analyse conjointe de données hydroclimatiques et d'images satellitaires très haute résolution. *Climatologie*, *9*, 83.

Kuhlemann, L.-M., Tetzlaff, D., & Soulsby, C. (2021). Spatio-temporal variations in stable isotopes in peri-urban catchments : A preliminary assessment of potential and challenges in assessing streamflow sources. *Journal of Hydrology*, *600*, 126685. https://doi.org/10.1016/j.jhydrol.2021.126685

Langlois, J. L., & Mehuys, G. R. (2003). Intra-storm study of solute chemical composition of overland flow water in two agricultural fields. *Journal of Environmental Quality*, *32*(6), 2301-2310. https://doi.org/10.2134/jeq2003.2301

Le, H. T., Pommier, T., Ribolzi, O., Soulileuth, B., Huon, S., Silvera, N., & Rochelle-Newall, E. (2022). Overland flow during a storm event strongly affects stream water chemistry and bacterial community structure. *Aquatic Sciences - Research Across Boundaries*, *84*(1), 7. https://doi.org/10.1007/s00027-021-00839-y

Li, X., Wei, Y., & Li, F. (2021). Optimality of antecedent precipitation index and its application. *Journal of Hydrology*, *595*, 126027. https://doi.org/10.1016/j.jhydrol.2021.126027

Madani, E. M., Jansson, P. E., & Babelon, I. (2017). Differences in water balance between grassland and forest watersheds using long-term data, derived using the CoupModel. *Hydrology Research*, *49*(1), 72-89. https://doi.org/10.2166/nh.2017.154

McDonnell, J. J., Spence, C., Karran, D. J., van Meerveld, H. J. (Ilja), & Harman, C. J. (2021). Fill-and-Spill : A Process Description of Runoff Generation at the Scale of the Beholder. *Water Resources Research*, *57*(5), e2020WR027514. https://doi.org/10.1029/2020WR027514

Omogbehin, M. H., & Oluwatimilehin, I. A. (2022). Changes of water chemistry from rainfall to stream flow in Obagbile Catchment, Southwest Nigeria. *Regional Sustainability*, *3*(2), 170-181. https://doi.org/10.1016/j.regsus.2022.07.006

Robinson, M., & Dupeyrat, A. (2005). Effects of commercial timber harvesting on streamflow regimes in the Plynlimon catchments, mid-Wales. *Hydrological Processes*, *19*(6), 1213-1226. https://doi.org/10.1002/hyp.5561

Sarrazin, B. (2012). *MNT et observations multi-locales du réseau de drainage d'un petit bassin versant rural dans une perspective d'aide à la modélisation spatialisée* (Theses 2012GRENU042, Thèse de doctorat, Université Grenoble Alpes; p. 283). https://tel.archives-ouvertes.fr/tel-01548105

Spence, C., & Woo, M. (2003). Hydrology of subarctic Canadian shield : Soil-filled valleys. *Journal of Hydrology*, *279*(1), 151-166. https://doi.org/10.1016/S0022-1694(03)00175-6

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*(4), 267-276. https://doi.org/10.1007/BF02289263

Tromp-van Meerveld, H. J., & McDonnell, J. (2006). Threshold relations in subsurface stormflow : 2. The fill and spill hypothesis. *Water Resources Research*, *42*(2). https://doi.org/10.1029/2004WR003800

Viessman, W., & Lewis, G. L. (2003). *Introduction to Hydrology*. Prentice Hall. https://books.google.fr/books?id=pSOsQgAACAAJ