Reviewer #1

This paper investigates spatio-temporal variability of end-member chemistry in a mountainous catchment. In a second step, EMMA is performed for four runoff events determining that soil sources contribute in addition to baseflow and precipitation, but groundwater being the dominating component. Additionally, the authors tested whether concentration of geochemicals could be calculated from conservative mixing. This was not the case. The authors also discussed the potential link between chemistry and changing hydrological connectivity. I find the study and the data set quite interesting. The paper is well written and data analysis is clearly described. While I like to overall paper good, there are several limitations.

We appreciate the overall positive assessment of our work and the helpful suggestions to improve the manuscript.

1. While I like the research questions and the introduction, I do not think that the research gaps for questions 1 and 2 are convincingly presented. For question 3 (first part), I believe that literature shows that this is not the case for most catchment where three component EMMA is performed.

We thank the reviewer for this comment. We recognize that the content and structure of the introduction did not logically lead to the research questions. We will restructure the introduction to provide more background for research questions one and two. We will also add other literature to present the research gaps more clearly.

Regarding the third research question, we agree that the answer to the first part of the question is obvious in some situations. This is exactly why we often use three-component EMMA. We included the first part of the question for completeness but we also see that we need to rewrite the question to emphasize the novelty of our work, rather than reminding the reader about the results from other studies. Therefore, we propose to change the third research question as follows:

"In how far does conservative mixing of baseflow and rainfall explain the changes in stream solute concentrations and do discrepancies from this mixing indicate when other sources become connected to the stream?"

2. The connectivity part is a little bit weak. It is only loosely linked to the results and could be made stronger in results and discussion. The study also lacks a clear definition of hydrological connectivity. Is it mass transport here?

We agree that the discussion of the connectivity part can be improved. We plan to rewrite the discussion (see below) and will include more linkages to the connectivity part. We plan to include a paragraph about the assumption that a connected water table indicates hydrologic connectivity (see below) and discuss the implications of this study for connectivity studies. To improve the link with the results, we plan to expand the paragraph in the discussion that discusses the connectivity results (L412 – 422) as well.

We agree that defining hydrological connectivity is important and that this will help the reader understand the study better. In this manuscript, we refer to connectivity as the flow of water between different locations in the catchment. We plan to include the following definition: "Hydrologic connectivity is the linkage of separate regions of a catchment via water flow" (Blume and van Meerveld, 2015).

As connectivity here is linked to GW level rising close to the surface. Several recent papers challenged such a simplified assumption (e.g. Jackson et al., 2014, Klaus and Jackson, 2018, Gabrielli and McDonnell, 2020). I guess this is still somewhat in the debate, but clearly data on bedrock permeability should be presented to check whether the assumed connectivity from GW levels can be realistic. Maybe other proof can be provided that GW level can be used to infer connectivity?

This is a very interesting question. Klaus and Jackson (2018) indeed showed based on the contrast in soil and bedrock permeability and Darcian flow principles that groundwater will infiltrate the bedrock before reaching the stream in many catchments. However, they focused on situations where there is a perched water table that occurs above the soil-bedrock interface during events. In the Studibach, there is an almost permanent water table in the low conductivity gleysols in most locations. We assume that significant lateral flow occurs when this water table rises into the near-surface layers, where the hydraulic conductivity is much larger (c.f. Schneider et al., 2014). However, flow to deeper soil layers is also substantial (Feyen et al., 1999). We will make it clearer in the manuscript that we talk about groundwater flow in the more permeable layer of the soil, i.e., above a saturated soil.

Gabrielli and McDonnell (2020) calculated which regions of several catchments can or cannot contribute to streamflow for four geologic settings. Both the Klaus and Jackson (2018) and Gabrielli and McDonnell (2020) paper show that the difference between conductivity of the conducting and impeding layer (in these papers soil and bedrock) determines the downslope travel distance, and thus connectivity to the stream. They assume Darcian flow. Preferential flow will significantly increase the upper layer conductivity, and thus increase the distance that a water parcel can travel before infiltrating to the bedrock. Chloride and bromide tracer studies have shown that preferential flow is an important transport mechanism in the Alptal (Feyen et al., 1999). Feyen et al. (1999) estimated that the effective saturated hydraulic conductivity ($K_{sat}$) is roughly $10^3$ times larger than the $K_{sat}$ of the soil matrix (0.0062 m/s vs. 7.2 $10^{-7}$ m/s for the soil at 3-25 cm depth and 0.2 $10^{-7}$ m/s for the soil at >40 cm depth). Therefore, water might be able to laterally transfer much quicker than expected. van Meerveld et al. (2018) report a surface $K_{sat}$ of 2.8-5.6 $10^{-7}$ m/s for the wetland sites (n=2), 5.6 $10^{-7}$-1.1 $10^{-5}$ m/s for a steep meadow site (n=2) and >1.1 $10^{-4}$ in the forest (n=1) (data from Sauter, 2017), which shows that the surface infiltration rate depends on land cover.

If we would assume that the low conductivity part of the lower part of the soil profile is unsaturated (which is not the case), we can calculate the downslope travel distance according to Klaus and Jackson (2018) and Gabrielli and McDonnell (2020). Slug tests for the Studibach groundwater wells (Zehnder, 2013) suggest that the $K_{sat}$ of the soil at the soil-bedrock interface ranges from 1.71 $10^{-6}$ to 3.62 $10^{-9}$ m/s (median: 2.33 $10^{-7}$ m/s). If we use these permeability values to calculate the downslope travel distance using Eq. 1 from Gabrielli and McDonnell (2020), assume an average well depth of 1.05 m, a depth of the more conductive soil layer of 0.3 m, and a slope of 30 degrees (which is the average slope in the catchment) or 10 degrees (the slope in the areas close to the stream), we obtain the flow distances in the upper part of the soil (Table 1).

Table 1: Downslope travel distances of a water parcel through the upper layer of the soil before percolating into the lower soil layer if this layer was unsaturated (which is not the case). Each column shows the $K_{sat}$-values that were assumed for the upper (first row) and lower (second row) soil layer, and the downslope travel distance (m) for a slope of 10˚ (the slope in areas close to the stream) and 30˚ (the average slope in the catchment).

| | **Assuming no preferential flow** | | | **Assuming only preferential flow** | | |
|---|---|---|---|---|---|---|
| $K_{sat}$ upper soil layer (m/s) | 2.33E-7 | 1.71E-6 | | 0.0062 | | |
| $K_{sat}$ lower soil layer (m/s) | 3.62E-9 | 2.33E-7 | 3.62E-9 | 1.71E-6 | 2.33E-7 | 3.62E-9 |
| Downslope travel distance 10˚ | 8 m | 1 m | 55 m | 423 m | 3102 m | 200 km |
| Downslope travel distance 30˚ | 14 m | 2 m | 100 m | 768 m | 5634 m | 360 km |

This back of the envelope calculation suggests that the downslope travel distance through the upper soil layer is large enough for a water parcel to reach the stream via preferential flow pathways. This is in part because the flowing stream network density in the Studibach is high (8.5 to 23.9 km/km$^2$ during dry and wet conditions, respectively; van Meerveld et al., 2019) and travel distances to the stream are thus relatively small. If we do not consider preferential flow, the water only reaches the stream if the permeability of the deeper soil is as low as the minimum measured $K_{sat}$. But the vertical gradients are likely much smaller than the unit gradient assumed by Klaus and Jackson (2018) and Gabrielli and McDonnell (2020) because the lower layer is saturated and thus the travel distances would be smaller.

We do not have $K_{sat}$ data for the bedrock but assume that this is even lower than for deeper soil layers, as otherwise there would not be a permanent water table above the bedrock in such a large part of the catchment.

3. While I think that the paper is quite good, the discussion is currently weak. While the authors are discussing the data and their variation in detail (which is appreciated), I miss discussion of the broader impact of the study, as well as a better link to the introduction or the literature in general. Right now, the discussion refers to only a few studies, mainly related to processes in the same catchments. The authors need to present the broader implication of their work, and make their general contribution to the state of the art outside their study site clearer. At the end of the read I was a little unsure on the take home message. I really think the impact of the paper would be much better if that is achieved.

Upon re-reading the manuscript, we also recognize the weaknesses highlighted by the reviewer and the lack of a broader discussion and clear take-home messages. To overcome this, we plan to rewrite the discussion so that it also includes:

- A comparison of the results to literature from other study sites, as we did in the introduction. We think that, for instance, linkages to the studies of Ladouche et al. (2001) and Soulsby et al. (2007) would be useful here.

- A section that shows the broader impact of the study. We think that such a section should include how our results fit with current concepts of hydrologic connectivity (e.g., Blume and van Meerveld, 2015). It should also address the assumptions made with calculations of connectivity as mentioned above (e.g., Jackson et al., 2014).

In addition, we will rewrite parts of the existing discussion and emphasize the take-home message. We can achieve this with a section on the broader impact of our study as the final paragraph, and finish with a take-home statement. This could be something along the following lines: “The combination of hydrometric and hydrochemical data can be useful to identify hydrological connectivity and aid the interpretation catchment-scale runoff generation. However, we have to take the variability of the tracer concentrations in different water sources into account, as they can be large compared to the change in streamwater concentrations. The observed gradual deviations in the concentrations that are expected based on mixing of baseflow and precipitation are likely the result of increases in the contributions from many (small) landscape elements in the catchment and thus reflect the gradual increase in connectivity during events."

4. The majority of the figures need to be reworked (3, 5, 6, 8). They lack the quality that is needed for publication.

We realize that the font size might not have been sufficiently large and that the contrast of the color scheme was not sufficient when the manuscript is printed. Therefore, we will enlarge the font and points, and revise the color scheme so that the figures will be more readable.

All our figures (except Figure 2) are vector images and thus have fulfill the DPI requirements. However, by including them in the .pdf document the quality was reduced. We will make sure to fulfil the required DPI when uploading the original figures.

Minor comments:

L35: typo "McGuire"

Thank you for pointing out this mistake.

L47: The authors present catchment size and location for the Maimai; one could do the same for the Rietholzbach. The introduction generally good; the research gaps for the first two research question should be made more clear.

We will mention the catchment sizes and locations throughout the manuscript. See our comments above regarding the first two research questions.

L92: Why should it only be baseflow? The literature is quite clear that, if tested, this is barely the case. So why asking a question we know to be not true?

See our comments regarding the third research question above.

L150: That is a valid assumption; but how variable is soil water chemistry (yes, the data is partly presented, but it could be stated)? Additional some more information on the choice of geochemicals and their commonly observed behaviour would be nice.

We recognize that more information on the variability of tracer concentrations in different water sources (particularly in soil water) and the choice of tracers and their characteristics can aid the reader in understanding the manuscript. Therefore, we will add a short description of the tracers and their variability in the last paragraph of the methods section. This section currently describes how we interpret the changes in streamwater solute concentrations.

Additionally, we will follow the suggestion of reviewer two to include a figure that shows the spatial-temporal variability of the tracer concentrations in each water source. We are testing different figure types but most likely it will be a figure that has a panel for each tracer and shows a boxplot per source (Figure S1). We prefer to add this figure to the supplementary information but of course will mention it in the text.

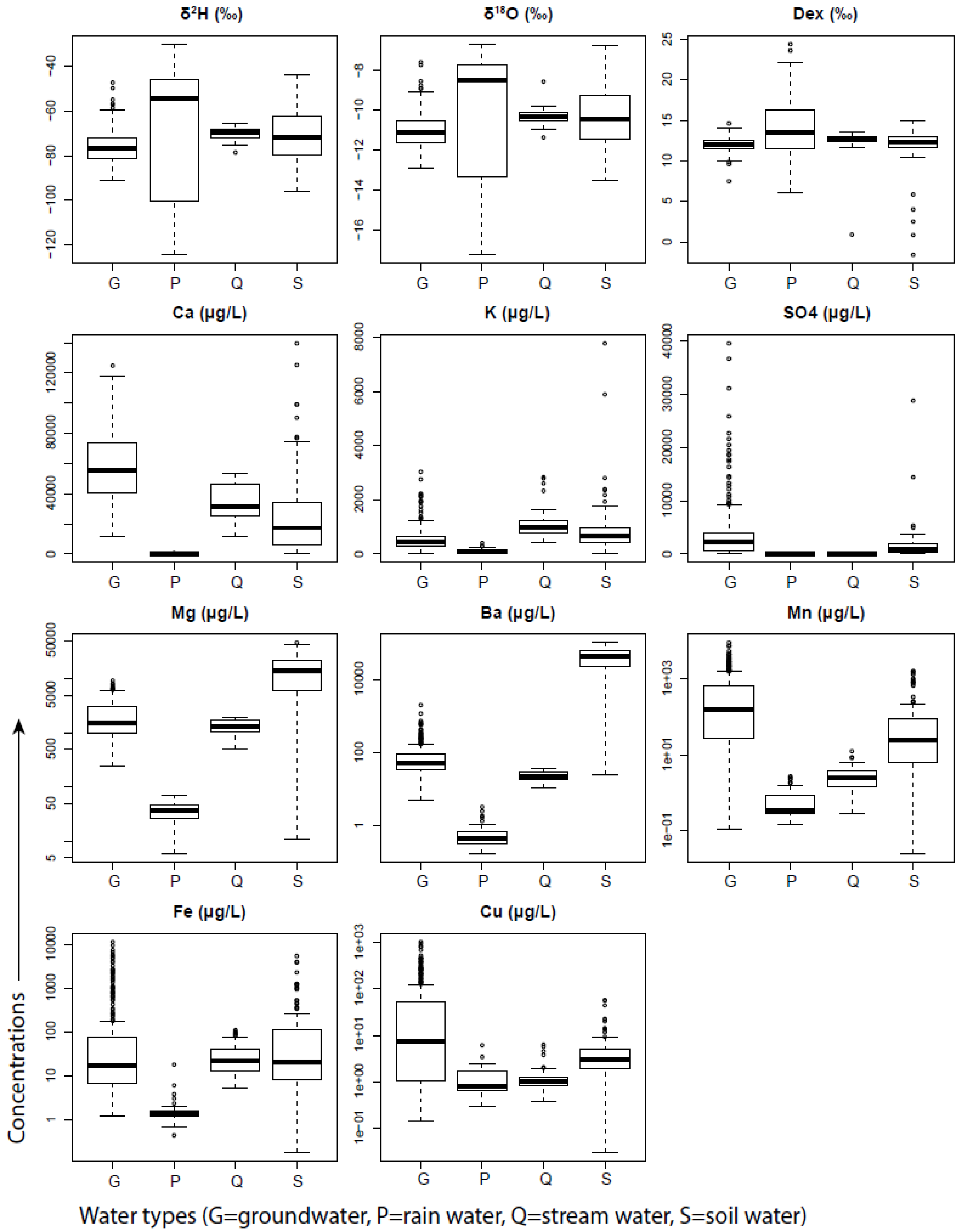


Figure S1: Box plots of the tracer concentrations for the different water types: groundwater (G), rainfall (P), streamflow (Q) and soil water (S). Each boxplot contains all streamflow or rainfall samples taken during the four events presented in this study, or all soil water or groundwater samples taken during the snapshot campaigns used in the study. Isotopic tracers are shown in ‰, values for chemical tracers in µg/L. Please note that y-axes differ for each panel, and that the y-axes of the panels on the bottom two rows are logarithmic.

L188: A clear definition of connectivity is needed, especially when not investigating the mass flux directly.

We agree that adding a definition of connectivity will be helpful (see answer to the comment above).

L198/199: You can only assume connectivity in cases where one have a low permeable of underlying bedrock (cf. Jackson et al., 2014; Klaus et al., 2018; Gabrielli and McDonnell, 2020).

We realize that we did not explicitly mention or comment on this assumption and will address this assumption in the revised manuscript. See also our answer to the comment above.

L219: Define “similar”

The difference in the event water fraction for the two-component hydrograph separation using $\delta^2H$ or $\delta^{18}O$ as a tracer was 0.05. We will include this information in the manuscript to better define the similarity.

L251ff: There is a nice paper by Harris et al. (1995) that looked into changing end-member contributions. The idea is not too different from the one here.

Thank you for highlighting this paper. We were not aware of it. Indeed, it discusses the calculations that match the idea that we present. Thus incorporating it in our discussion is a good suggestion.

L251ff: There is a range of studies that looked (e.g. McCallum et al., 2010), related to hydrograph separation, how GW chemistry is different from baseflow chemistry.

Thank you for recommending this paper. The McCallum paper indeed shows how baseflow chemistry can be different from groundwater contributions during a rainfall event. It is interesting because they showed through their numerical model and field-observations that the spatial variability of groundwater is important for hydrograph separation. We will certainly include the reference to this paper and also investigate other papers that discuss this topic, such as Chanat and Hornberger (2003) and Jones et al. (2016).

L345: Or does that indicate a much less pronounced connectivity compared to the model?

We appreciate this suggestion, and recognize that indeed, a less pronounced connectivity change might also be a valid reason for a smaller change in streamflow composition than expected. We will include this alternative interpretation in the manuscript and add a more general comparison of the model results and observations. Such a statement could be along the following lines: “The limited change in streamwater composition might also be the effect of a discrepancy between the model results and the actual expansion of the hydrologically connected area, for instance, because the change in connectivity is smaller than predicted by the model. Additionally, the transport of a water parcel to the catchment outlet takes time, and the expansion of the hydrologically connected network within the catchment and change in streamflow composition at the catchment outlet might thus not be synchronous.”

L365: Is that surprising? The spatial variability is the maximum extend of the mixing diagram of endmembers. Thus, changes in the stream must be smaller, if the sampling was representative. I am missing the bigger picture here. The discussion is very detailed and evolves around the data being non-conclusive. It would be nice to expand this section and discuss what the key contribution to the field of runoff generation is. How do you go beyond studying this catchment? How does your work related to previous work? What is the key novelty? You may also think of linking your discussion better to the introduction and the used references there.

Indeed, this is not surprising but very few studies have characterized the spatial variability for groundwater and soil water. We mention that this is not surprising in the final part of the discussion (L442) but we recognize that these two statements are rather far apart and will adjust the text accordingly.

L448: but for some? And what do you infer from that?

Indeed, the contribution of soil water is important during some of the events, but not all events, which further complicates the analysis of connectivity based on stream chemistry. We will discuss the implications of soil water contributions to streamflow in the description of the hydrologic functioning (section 5.2).

Figures 3, 5, 6, 8 are not very well done. While the content is fine, the presentation, choice of colours, font size, and point type should be revised.

Thank you for pointing this out. We will improve these figures following the suggestions.

**References;**

Blume T, van Meerveld HJI. 2015. From hillslope to stream: methods to investigate subsurface connectivity. Wiley Interdisciplinary Reviews: Water 2 (3): 177–198 DOI: 10.1002/wat2.1071

Chanat, J. G., and Hornberger, G. M. ( 2003), Modeling catchment-scale mixing in the near-stream zone—Implications for chemical and isotopic hydrograph separation, Geophys. Res. Lett., 30, 1091, doi:10.1029/2002GL016265, 2.

Feyen, H. & Wunderli, H. & Wydler, H. & Papritz, A.. (1999). A tracer experiment to study flow paths of water in a forest soil. Journal of Hydrology. 225. 155-167. 10.1016/S0022-1694(99)00159-6.

Fischer, Benjamin & van Meerveld, Ilja & Seibert, Jan. (2017). Spatial variability in the isotopic composition of rainfall in a small headwater catchment and its effect on hydrograph separation. Journal of Hydrology. 547. 10.1016/j.jhydrol.2017.01.045.

Gabrielli, CP, McDonnell, JJ. Modifying the Jackson index to quantify the relationship between geology, landscape structure, and water transit time in steep wet headwaters. Hydrological Processes. 2020; 1– 12. https://doi.org/10.1002/hyp.13700

Inamdar, S., Dhillon, G., Singh, S., Dutta, S., Levia, D., Scott, D., Mitchell, M., Van Stan, J, and McHale, P. ( 2013), Temporal variation in end-member chemistry and its influence on runoff mixing patterns in a forested, Piedmont catchment, Water Resour. Res., 49, 1828– 1844, doi:10.1002/wrcr.20158.

Jackson, C. R., Bitew, M., & Du, E. (2014). When interflow also percolates: Downslope travel distances and hillslope process zones. Hydrological Processes, 28(7), 3195–3200. https://doi.org/10.1002/hyp.10158

Jones, J. P., Sudicky, E. A., Brookfield, A. E., and Park, Y.-J. ( 2006), An assessment of the tracer-based approach to quantifying groundwater contributions to streamflow, Water Resour. Res., 42, W02407, doi:10.1029/2005WR004130.

Klaus, J., & Jackson, C. R. ( 2018). Interflow is not binary: A continuous shallow perched layer does not imply continuous connectivity. Water Resources Research, 54, 5921– 5932. https://doi.org/10.1029/2018WR022920

Ladouche B, Probst A, Viville D, Idir S, Baqué D, Loubet M, Probst J-L, Bariac T. 2001. Hydrograph separation using isotopic, chemical and hydrological approaches (Strengbach catchment, France). Journal of Hydrology 242 (3–4): 255–274 DOI: 10.1016/S0022-1694(00)00391-7

Sauter, T. 2017. Occurrence and chemical composition of overland flow in a pre-alpine catchment, Alptal (CH). Department of Geography. University of Zurich, Zurich, 79 pp.

Soulsby, C., Tetzlaff, D., Van Den Bedem, D., Malcolm, I. A., Bacon, P. J., & Youngson, A. F. (2007). Inferring groundwater influences on surface water in montane catchments from hydrochemical surveys of springs and streamwaters. Journal of Hydrology, 333(2-4), 199-213. https://doi.org/10.1016/J.JHYDROL.2006.08.016

von Freyberg, J., Studer, B., Rinderer, M., & Kirchner, J. W. (2018). Studying catchment storm response using event- and pre-event-water volumes as fractions of precipitation rather than discharge. Hydrology and Earth System Sciences, 22(11), 5847-5865. https://doi.org/10.5194/hess-22-5847-2018

van Meerveld H.J.I., Fischer B.M.C., Rinderer M., Stähli M., Seibert J. 2018. Runoff generation in a pre-alpine catchment: A discussion between a tracer and a shallow groundwater hydrologist. Cuadernos de Investigación Geográfica 44 (2): 429–452 DOI: 10.18172/cig.3349

van Meerveld, H. J. I., Kirchner, J. W., Vis, M. J. P., Assendelft, R. S., and Seibert, J.: Expansion and contraction of the flowing stream network alter hillslope flowpath lengths and the shape of the travel time distribution, Hydrol. Earth Syst. Sci., 23, 4825–4834, https://doi.org/10.5194/hess-23-4825-2019, 2019.

Zehnder, M. 2013. Hydraulic Conductivity Estimation and Analysis using Slug Tests - Relations with Site-Characteristics in a pre-alpine torrent catchment, Alptal (SZ). Department of Geography. University of Zurich, Zurich