This manuscript reconstructs debris-flow magnitude and frequency in the Horlachtal, Austria, since 1947. It uses extensive geomorphological mapping using historic and recent orthophotos. The authors show that debris-flow activity in this area was dominated by short-term variations rather than consistent increasing or decreasing trends. Furthermore, the analyses point to local thunderstorms triggering debris flows in the Horlachtal.

In my opinion, this work is strongly relevant for the journal of Natural Hazards and Earth System Sciences. The manuscript is based on a solid and extensive analysis. In total 834 debris flows have been mapped, leading to strong statistics. Furthermore, the manuscript is well-written, although figure presentation may be improved. Below I list a number of suggestions for improvement.

**Main:**

*Transport-limited vs supply-limited hillslope systems.* In lines 101-103 the authors state that the debris flows in Horlachtal occur in transport-limited hillslope systems. However, in the discussion the authors argue that highly active periods affect debris-flow activity in the following years by reducing magnitude and frequency as a result of depleted sediment storages (e.g., lines 472-482). This is a textbook example of supply-limited conditions, and therefore the statements in lines 472-782 and 101-103 are in direct contrast with each other.

*Relation between rainfall magnitude and flow magnitude.* On a related note, in transport-limited systems one would expect a correlation between triggering rainfall magnitude and
debris-flow magnitude. In contrast, such a correlation becomes weaker and would typically be absent in supply-limited systems because flow magnitudes also are limited by the volume of sediment available. It would therefore be of interest to compare triggering rainfall to flow magnitudes, or given the data availability perhaps maximum rainfall magnitudes in a given period versus the maximum debris-flow magnitudes in the same period (this should at least be possible for sub-catchment ZT judging from the information in section 5.2.2). It may be needed to normalize by catchment size or another morphometric characteristic of the source catchment as this also affects flow magnitude.

*Catchment morphometry versus flow magnitudes.* In section 4.3 the volume of 404 debris flows is compared to a number of morphometric parameters of their source catchments. A key component of such an analysis is information on how many events are generated from each studied source catchment. If each catchment in the dataset generates multiple flows there is stronger statistics, while if each catchment only produces 1 flow this introduces uncertainty since this one flow may have been relatively small or large. It would be good to elaborate on this in the manuscript.

*Conclusions.* I suggest to shorten the conclusions and also remove the subsections.

**Details:**

Lines 97-98: Please elaborate on how these type 2 debris flows in Zimmermann (1990) or type 1 in Wichmann (2006) and Rieger (1999) are defined, e.g., describe their characteristics.

Lines 162-163: Two times "The approach".

Line 2019: Parameter should be parameters.

Lines: 270-274: “The mapping of the debris flows showed a concentration of these processes in the parallel sub-catchments GT, LT and ZT. As those debris flows show such a different picture when comparing them to the activity in the other sub-catchments, and because of the similarities in the geomorphological and geographical settings, the analyses concerning deposition volumes were carried out exclusively in GT, LT and ZT.” It is unclear how the debris flows of sub-catchment GT, LT, and ZT differ from those of the other catchment. As such, this statement raises a lot of questions. Please clarify.

Line 300. Include space in “many events”.

Lines 297-301. To me the most striking feature in Fig. 8 is the strong increase in flow volume around 1990. Therefore, it would be good to also describe that here.

Line 532. Include space in “adetailed”.

**Figures:**

Overall: In many figures the font sizes are too small and should be enlarged.

Figure 2: It would be more informative to not only plot mean temperatures and precipitation, but rather plot a band indicating the values range. For example, mean +- std or 25-50-75 percentiles.

Figure 5: For readability the font size on the axes should be enlarged.

Figures 6 and 8: Given the unequal intervals of the time slices the left panels of these figures are not informative. I therefore suggest that the authors only present the data on the right panels, and combine debris flow frequency (Fig. 6) and magnitude (Fig. 8) in one figure. Do not denote panels as left and right, but annotate as “a” and “b”. In addition, also for these figures the font size is too small and I suggest enlarging the font. In addition, for the magnitudes it would be beneficial to also include uncertainties with dashed lines.

Figure 12: For comparability it would be better to present the magnitude-frequency curves in panels b and c together in one panel. Also gridlines would help interpretation of the figures. Again, font sizes should be enlarged in this figure.