

Comment on hess-2021-567

Anonymous Referee #1

Referee comment on "Agricultural intensification vs. climate change: what drives long-term changes in sediment load?" by Shengping Wang et al., Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-567-RC1>, 2021

I found the preprint titled "Agricultural intensification vs climate change: What drives long-term changes of sediment load?" an interesting and worthwhile contribution to the exploration and identification of potential drivers of long-term changes in suspended sediment. I applaud the authors on their well written abstract and introduction, use of data analysis techniques, and helpful interpretation of their results. I found much of the writing clear and concise. I have a few major comments, and some minor ones, that I hope the authors will consider.

Major comments:

- Only 1 metric of "climate change" was evaluated in this paper and this metric, monthly mean precipitation, does not capture the change in climate most likely to effect sediment transport. Metrics such as change in the magnitude, frequency, or duration of the highest precipitation events (> 75th or 90th percentile? Max event?) would be more appropriate for assessing the effect of changes in climate on sediment transport. Furthermore, I don't see any discussion related to the type of precipitation (i.e., changes in the proportion of precipitation as snow versus rain). The catchment is located in Austria, so surely snow fall is a consideration. Is it possible there has been a shift from snow to rain which could be driving some of these changes in sediment?
- The differences in sampling methods between the two time periods needs more attention. As presented, I am not convinced we aren't just seeing changes in sediment due to changes in sampling methods (collection frequency and also techniques). Sediment is particularly sensitive to changes in sampling (see couple of references at bottom). Lines 228-230 suggest an approach for dealing with differences in sampling but it is not clear to me how this technique takes care of the potential issue. Even if the difference in sampling methods/techniques cannot be resolved, I think this analysis would still be worthwhile, as long as this issue was thoroughly discussed in the context of the results.
- Statistical techniques can and should be used to help identify and quantify differences between periods. The text indicates a t test was used for assessing differences in

monthly erosivity density (line 189) but how differences between the SRCs and the sediment loads were determined isn't described. I appreciate the development of the theoretical SRCs but even the differences between these are discussed qualitatively. For example, on lines 433-435, how is the distinction between "considerably higher" and "not different" being made? There is considerable overlap in the 5 and 95th percentiles in both panels of Figure 7. Specifically, on panel (a), the line for the 50th percentile for Period II falls outside the gray area for Period I. Using a statistical test like ANCOVA or a regression equation with categorical variables for season and time would help to determine which slopes and intercepts are (statistically) significantly different.

- Given the structure of the dataset, how were the sediment loads presented in this report calculated? What technique was used? Relatedly, a table of these loads might be worth including since the term "load" is featured in the title. Also, Lines 459-460 used the word "load" to reference to Figure 8 but this figure is showing sediment concentration. In several places in the paper it feels like the terms "concentration" and "load" are being used interchangeably. Sometimes this can be ok since the SRCs are always positive and increases in concentration can be inferred to mean increases in load, but they are not the same thing and it would be prudent to be clear about which term is used when presenting the results of these analyses (i.e., these SRC are build using sediment concentrations not loads).
- I'm wondering about the choice to average concentration and streamflow data by month.... I wonder if some of the important variability (related to the magnitude of the events) is being muted? The report states that a majority of the sediment load is transported in just a few high flow events in each period, but these important events are being averaged with all the available data for each month.

Minor comments:

Line 245: Is this saying the solid and dashed lines in 2b represent a shift in the sediment transport regime such that concentrations at a given discharge are relatively lower for periods A', B', and C' compared to A, B, and C? Consider rewording.

Line 226: The text indicates SRCs were also fit by season-and-year however the results and discussion only present and discuss the SRCs by season. Suggest removing this since the by year and season results aren't discussed.

Line 269: Figure 4 is used to support the statement that the coefficient of log a follows a normal distribution, but Figure 4 is a flow duration curve. Also "shown" would probably be a better word than "proven" in this context.

Figure 3: Consider including labels that indicate which months are in the Growing vs. Dormant seasons. Are the months included in the growing and dormant seasons available

somewhere in the text? I can't seem to locate it. My apologies if I just missed.

Figure 1: Caption indicates a "black hatched area in b", but I don't see it. How different are the catchment sizes between the 2 time periods?

Line 330: Can you support this statement more? I think the authors are saying that decreases in streamflow cannot account for the observed increase in sediment transport, because if that was the case, then we'd expected to see increased streamflow, correct? Consider rephrasing

Figure 6: Keep the colors used for Period I and II consistent between Figures 5 & 6. What are the arrows for? The text describes the right points being the "left-upper area" but this is not true for (a).

Lines 436-440: The +/- ranges for the loads given in this paragraph would result in negative sediment loads. Also, +/- ranges are quite large suggesting perhaps there isn't a statistical significant difference between these loads?

Figure 8: What does the dashed vertical line at zero represent?

References on sediment sampling in rivers:

Awal, R., et al., 2019, A General Review on Methods of Sediment Sampling and Mineral Content Analysis, *Journal of Physics: Conference Series*, <https://doi.org/10.1088/1742-6596/1266/1/012005>

Groten, J.T., and Johnson, G.D., 2018, Comparability of river suspended-sediment sampling and laboratory analysis methods: U.S. Geological Survey Scientific Investigations Report 2018-5023, 23 p., <https://doi.org/10.3133/sir20185023>

Harmel, R.D., et al., 2010, Impact of Sampling Techniques on Measured Stormwater Quality Data for Small Streams, *Journal of Environmental Quality* 39:1734-1742, doi:10.2134/jeq2009.0498