Earth Surf. Dynam. Discuss., author comment AC1
https://doi.org/10.5194/esurf-2021-85-AC1, 2022

Reply on RC1

Lena Katharina Schmidt et al.

Author comment on "Suspended sediment and discharge dynamics in a glaciated alpine environment: Identifying crucial areas and time periods on several spatial and temporal scales in the Ötztal, Austria" by Lena Katharina Schmidt et al., Earth Surf. Dynam. Discuss., https://doi.org/10.5194/esurf-2021-85-AC1, 2022

Replies to the general comments:

There is a bit of a mix-up regarding terminology in the article. In the Introduction and Methodology section, you discuss "suspended sediment yields" (SSY, t/km²). However, later in work, you describe sediment mass fluxes expressed as suspended sediment loads (SSL, t/yr). Moreover, you didn't mention how you calculated sQ, SSY, and SSL. This should be added in the first instance to understand what is going on.

Reply: Thank you for pointing this out. Indeed, we decided to use both suspended sediment yields (SSY, for comparability among the sub-catchments) and suspended sediment loads (SSL, to enable the readers to get a feeling for the absolute magnitudes). We are happy to add the respective equations to the methodology and to review whether it will be more easily understandable if we change from SSL to SSY in some instances.

The paper's abstract is poorly written and does not tell the story well. Undeniable statements (like L24ff, L27ff) are mixed with results so that after reading the abstract, it is not clear if anything new has been done. Try to be more specific, highlighting material you use (e.g., water discharge and suspended sediment concentration series from 2006 to 2020). Names of gauging stations are worth mentioning. It would be interesting to the reader to have some descriptive statistics in the abstract (e.g., mean annual SSC, Q, or SSY) and main results. A paper by Mensh and Kording (2017) might be helpful.

Reply: We are happy to rewrite the abstract as you suggested (more specific description of data, gauging stations, etc.). However, we fail to understand what you mean by “undeniable statements”. Both lines you are referring to describe aspects of our results.

Correct the structure of the article. Dissolve the results from methods and discussion. See minor comments for some suggestions.

Reply: We agree, thank you for the detailed suggestions.

You are saying (L144-145) that you have measured turbidity at all stations and then recalculated NTU to SSC. However, Fig2 and the corresponding equation only describe *lower distribution bounds (0-20 SSC or 0-10 NTU). From Supplementary materials(Vent_Q_SSC), I can see that SSC increased up to 1000 g/l. How did you calculate suspended sediment concentration for values above 20 g/l? Using the same equation for extrapolating the linear model to a high-value area usually leads to significant errors and uncertainties. This case should be corrected and critically discussed. This is the weakest part of the research, questioning your conclusions.*

Reply: Thank you for this valuable comment, which highlights that we apparently did not make our description clear enough. Firstly, we only refine the calibration of turbimetry at one station (Sölden) and use data provided by the Hydrographic Service of Tyrol for the other two gauges. Secondly, the output given by the turbidity probe in Sölden is already in (tentative) concentration units, i.e. (m)g/l (not NTU), but had to be recalibrated to the actual concentrations using samples. Thirdly – and perhaps most importantly – SSC in Fig. 2 is given in **g/l** (and refers to the station Sölden, see figure caption) while SSC in the supplementary file you refer to is in **mg/l** (and refers to the station Vent) as stated in the metadata. This should resolve your remark about extrapolation. We will improve the description in lines 144 ff. to be more clear.

The second weakest point of your research is the visual identification approach of the strongest sediment flux events. This approach is described by you too vague. I insist on adding some criteria, and event statistics. Addinf og descriptive statistics of all Q (m³/s), SSC (g/l) and SSL (t/event) events will help us (readers) to understand was is «strongest» mean.

Reply: Thank you for this comment. We agree that the visual approach may seem somewhat subjective, but we have put considerable thought into how to best identify suspended sediment events. Unfortunately (and as you also point out in your comment on hydrograph demarcation later on), it is not straightforward to automatically identify events. Due to the considerable intra-annual variation in SSC, a threshold-based approach would overlook events early in the year, which stand out against concentrations or yields that occur that time of the year, but with absolute concentrations or yields much smaller than during summer. In turn, monthly thresholds are problematic because interannual differences are high. To program our logger for sampling, we came up with an event detection routine (as briefly outlined in line 161 f.) considering the concentrations in the days ahead and the increase in concentrations, which worked well enough to sample events, but which is unsuitable for complete event identification. For instance, we observed that the occurrence of one event masked the detection of subsequent events in the days after, which disqualifies this approach for the problem at hand. Similarly, the beginning and end of each event are not easily put into a formula, as not all events do show clear points of inflection before and after the peak. Thus, we came to the conclusion, that visual identification (i.e. expert opinion) is the most straightforward approach to this problem. We will however improve our description on which criteria guided the delineation. We already describe the identified events with respect to Q and SSL in lines 431f (and your comment brought to our attention that the superscripts (e.g. $6*10^6$ to $13*10^6$ m³) were lost due to formatting, we will correct that as well).

I understand that there is some evidence that the sediment load at the stations is simultaneously changing. However, what about water runoff? Figure 7 shows that the mean annual Parde's coefficient for Soelden and Tumpen varies equally. Why does Vent stand out like this? It would be interesting to compare the water runoff with the snowfree area too. However, the visual technique you use in Figure 8 requires some quantification. Maybe compare the week of the year of the beginning of the increase in sediment load and water runoff (i.e., the inflection point) with the beginning of the snowmelt at different elevations for different years?

Reply: The difference in Pardé coefficients is due to the higher elevation of gauge Vent (as we explain in L 366 to 375) and was not surprising given the existing knowledge on this (e.g. Gattermayr, 2013; Kormann, 2016; Weber and Prasch, 2016; as cited in the manuscript).

The respective figure shows mean annual cycles. As such, we do not think somewhat soft features such as "increase in sediment load" should (nor could) be precisely pinpointed to specific week. Therefore, we propose to leave the comparison on the qualitative level.

Replies to specific comments and technical corrections:

L128-129 — The sentence is unrelated to the rest of the text. What slope is meant? I guess catchment slope like the one from Table 1. Consider removing or improving the phrase. Personally, I find it redundant here.

Reply: Thank you for this detailed comment, we will remove the sentence.

L130 — I don't understand where the footnotes in Table 1 are headed. Is it like sources of the data used for calculations? Then additional column named «Data Sources» with references may be the better way of presenting. Otherwise, consider moving the phrase from the Title to the Table's bottom (or footnote).

Reply: Thank you for this comment, we understand. Indeed, the footnotes are headed towards the data sources. We chose not to add an extra column, since the table already is quite wide and the data sources are the same for several entries, so a data sources column would include a lot of repetition. We will move the data sources to a footnote at the bottom of the table, as commonly handled in other tables in the journal.

L151 — You said (L151) that 2019-2020 data are preliminary. Why? What makes it preliminary? Is it needed to be checked by authorities?

Reply: Indeed. We received data of the stations Vent and Tumpen from the Hydrographic Service of Tyrol, who quality check the data eventually. That has not happened for the data of 2019 and 2020. We will specify this more clearly here.

L206 — The first mentioning of the SSY «…we visually identified SSY peaks…» needs abbreviation decoding. Moreover, I suppose you meant SSC here.

Reply: Thank you, we will add the explanation. However, we did indeed mean SSY, since our analysis is aimed at identifying the events with the highest sediment output.

L144-145 — I'm just curious what turbidity sensors did you use. E.g., model and Manufacturer

Reply: The sensors are Solitax sensors by Hach at all gauges. We will add that to the description in the manuscript.

L156 — Again, please mention the model and manufacturer of the automatic sampler

Reply: Thank you, we will add that. It is a P6 L Vacuum by MAXX.

L175 — write it like an equation

Reply: We will add the respective equation.

L180 — I'm surprised with the Turbidity dimension. Shouldn't it be NTU or FTU?

Reply: We understand the confusion since many turbidity probes give results in units of NTU or FTU. However, as mentioned above, the raw data of the turbidity probe used here are already given in concentration units as a preliminary calibration happens within the sensor – yet still need to be calibrated with concentrations from a sufficient number of samples.

L196 — Please, explain why did you choose a 3 mm threshold. Indeed, you are correct to note that 3 mm is not enough to consider an event as erosional (Renard et al. (1997) suggested a 12.7 mm threshold, for example). Nevertheless, at the same time, it seems that we should separate the snow from the rain more by the air temperature. And that threshold, according to the 2018 study (Jennings et al., 2018) for Tumpen, should be around 1.5 °C, not 0 as you used.

Reply: We agree that the temperature threshold should be changed, thank you for the very helpful reference. However, this will not change the analysis result, since temperature is in fact >1.5°C in all cases.

However, the identification of precipitation events is more intricate. Our precipitation data are point measurements at the gauge in Vent and we know that precipitation within the almost 100 km² catchment above the gauge can be highly variable and is affected by the topography. On the one hand, this is reflected in the precipitation gradient (e.g. L113f.) of about 5% per 100m. Assuming this would be applicable to individual events, a 12.7 mm precipitation event in front of the Vernagtferner glacier at about 2850 m elevation would correspond to about 6.7 mm at the gauge in Vent at roughly 1900 m. This is also reflected in the differences in mean annual precipitation (L303ff: "The mean annual precipitation recorded close to the Vent gauge is 666 mm while areal precipitation of the whole catchment is estimated between 1200 and 1500 mm, and for the 11.4 km² Vernagtferner sub-catchment […] even 1525 to 1900 mm are reported"). Thus, we can generally expect the precipitation measured in Vent to be a lower bound of precipitation falling in the entire (sub-) catchment. Adding to this, considering the possibility of rain on snow events and fluvial erosion, we doubt that the Renard threshold can be meaningful here. Instead, we used the hydrograph shape as additional information (as described in L198f) and used the low threshold of 3mm. We will improve the explanations.

L200-205 — some additional visualization may be helpful. The hydrograph demarcation by water sources is a very discussable topic, and your way to demarcate it is a bit complex.

Reply: Thank you for this helpful comment. We understand that this part needs more explanation. We have prepared a visualization suggestion (see page 1 in Supplement.pdf; a schematic diagram of an event in 2020). However, any visualization can only be exemplaric and will fail to describe all possible types of events. Thus, we believe that it would be more helpful to improve the description in the text (as mentioned above). For example, it is very insightful that you are referring to hydrograph demarcation, when we are aiming at SSY events, so we will make this more clear.

L239-242 — This chunk belongs to the Methods section

Reply: Thank you. We intentionally placed this paragraph here so that readers would not have to jump back to the methods to understand the graph that follows. However, we agree that we need to describe this in the methods. We will reduce the paragraph here and add the description to the methods.

L254-261 — While this part is a discussion.

Reply: Thank you, we will move this part to an appropriate place in the discussion.

L249 — Mean annual discharge per area or specific discharge? I suggest using the same wording in the whole paper. Otherwise, it is confusing. Moreover, mm/a is it mm per annum? It is more common to write mm/yr or mm/year

Reply: Thank you for this detailed remark. We will harmonize the wording and change the units to mm/yr.

L284-289 — This is a discussion

Reply: We agree and will move this to the discussion.

L291 — Are both p-values equal to 0.001? This is surprising considering the various scattering in Fig4.

Reply: Indeed, the p-values are <-2.2e-16 and 0.001857. Thus, both are (well) below the significance level (commonly denoted as α) of 0.01 (not 0.001, this is not a common significance level and was a typing mistake). We will add that in the form of "(significance level α = 0.01)" for clarification.

L297ff — This is a discussion

Reply: We agree and will move this to the discussion. As a result, section 3.3 will be very short, so we will merge sections 3.3 and 3.4.

L341ff — It is necessary to add the corresponding section in Methods. How did you calculate Parde coef?

Reply: Thank you, we agree that this is missing in the methods. However, as we switched to % of annual runoff (see comment on figure 7 below), this is now obsolete.

L411 — This is a discussion

Reply: We are sorry, in the version we submitted L411 is the heading of section 3.8. We assume that is not what you meant.

L434 — sediment load not yield, I guess

Reply: Yes, thank you for the attentive comment. We will adjust that.

L473ff — you have already mentioned your aim in the Introduction

Reply: That is correct, however we intentionally repeated it here to make reading easier. We suggest to adjust the first paragraph to rather summarize the important findings from the results section as suggested by Mensh and Kording (2017).

Fig4. These are exciting results, but I'm not sure that linear regression is the right way to analyze the SSY-Glacier area relationship in your case. Or maybe I understand your graph wrong because of the legend absence. First of all, you should mention that you hypothesize that the glacier area didn't change significantly during 2006-2020. However, from table 1, we know that this is not true (up to 6 % for less than ten years). I guess that the actual distribution of sQ and SSY along the glacier area would be different if you compare yearly SSY with yearly glacier area. The plot like on Figure 4 can make sense only if you compare mean annual values for 2006-2020 with the mean annual glacier area for 2006-2015. That will make your plot look less significant (i.e., only 5 points) but will *make more sense.*

Reply: Thank you, we will add a legend. We agree, of course the glacier cover changed during this time and ideally we would use annual values here. However, annual glacier area measurements are simply not available, all there is are the glacier inventories of 2006 and 2015. Thus the mean annual glacier area for 2006 – 2015 would be the mean of two measurements, which is not very informative in our opinion.

Fig5. It would help if you avoided your qualitative assessment in the figure caption. Better to add R2 on a graph.

Reply: Thank you. The figure caption is a result of a recommendation by Mensh and Kording (2017) (which you also recommended): "[...] the title of the figure should communicate the conclusion of the analysis". However, we are happy to add the R^2 to the figure.

Fig6 — This is a good illustration for the discussion ð□□□

Reply: We are not entirely sure, what this comment is targeting, maybe because the last characters of your comment have been lost. We are discussing the relationships between glacier area and mass balances and Q and SSY in the paragraph starting line 483. If you are suggesting to refer to figure 6 in the discussion, we can add a reference (e.g. in L494).

Fig7 — Why are you using Parde coef and not the same % of annual runoff as for the suspended sediments? Maybe adding standard errors or standard deviation will be more valuable than the min-max range. Again, there is wrong wording: the second graph should be % of annual SSL (sus. sed. load), not SSY.

Reply: We used the Pardé coefficient since it is a standard hydrological index for streamflow seasonality. However, we agree that % of annual runoff will be more consistent and thus easier to follow for the readers and thank you for the helpful comment. We will adjust the figure accordingly, which however does not alter the message (page 2 in Supplement.pdf). Additionally, we used the 25 % and 75 % quartiles instead of min-max here, as a suggestion. However, we disagree on the "wrong wording", since the percentage of annual SSL [t] is the same as percentage of annual SSY [t/km²].

Fig8 — Is this multiyear average % of SSY and Snow free area? Can you add confidence intervals on lines, then? It is correctly to label the dashed lines simply by the station name as they represent not SSY but the ratio of annual SSY.

Reply: Thank you for the helpful comments. Yes, the version in the manuscript before showed multiyear averages. Since we assume that your question is directed towards the desire to visualize interannual variation, we adjusted the figure by adding interquartile ranges (25% and 75% quartiles, see page 3 in Supplement.pdf). Further, we plotted the median of snow free area per week of year instead of the mean, as this reduces the influence of singular classification errors within the original snowcover data (as e.g. above 3500 m between week 20 and 30 in the earlier version of this plot). We will adjust the figure description and the description in the text accordingly.

Please also note the supplement to this comment:
https://esurf.copernicus.org/preprints/esurf-2021-85/esurf-2021-85-AC1-supplement.pdf