



EGUsphere, author comment AC3
<https://doi.org/10.5194/egusphere-2022-616-AC3>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC3

Lena Katharina Schmidt et al.

Author comment on "Reconstructing five decades of sediment export from two glacierized high-alpine catchments in Tyrol, Austria, using nonparametric regression" by Lena Katharina Schmidt et al., EGU Sphere, <https://doi.org/10.5194/egusphere-2022-616-AC3>, 2022

RC3: 'Comment on egusphere-2022-616', Anonymous Referee #3, 21 Nov 2022

Dear anonymous Referee #3,

We would like to thank you for the very thoughtful and detailed comments, questions and suggestions. Below, we provide our response as direct answers to each comment and hope that our suggestions will be to your satisfaction. We also provide figures in the attached pdf for better understanding.

Best,

Lena Katharina Schmidt on behalf of all authors

Review of the manuscript (egusphere-2022-616): Reconstructing five decades of sediment export from two glaciated high-alpine catchments in Tyrol, Austria, using nonparametric regression by Lena Katharina Schmidt, Till Francke, Peter Martin Grosse, Christoph Mayer, Axel Bronstert

Summary: In this manuscript, the authors apply quantile regression forest (QRF) to simulate suspended sediment concentration (SSC) at the outlet of two nested glacierized catchments in Upper Ötztal in the Tyrolean Alps, in Austria. As predictors, they use discharge, precipitation and temperature. The QRF model(s) are used to generate long-term (1967-2020, 1974-2020) time series of mean daily SSC and specific annual suspended sediment yield (sSSY), which are later analyzed for trend analysis and point change detection. To identify causality for such trends and abrupt changes, the authors apply the same statistical analysis to the observations of precipitation, temperature, discharge, and mass balance of the two largest glaciers within the study area.

General comments: I think that the aim of the manuscript of understanding the potential of machine learning techniques to model SSC in alpine catchments, including climate variables as predictors, is valuable. However, in my opinion several aspects require substantial revision.

Major revisions: The methodology is not sufficiently explained. The authors should

clarify better mainly: (1) the quantile regression forests approach and the selection of the antecedent conditions for the predictors, (2) the procedure followed to fill-in the missing data (how did you compute the correction factors?) as well as (3) to disaggregate the data. Likewise, the availability of data and their resolution is quite confusing and requires clarification.

The authors frame some parts of the manuscript in a way that is conceptually questionable and potentially misleading. First, when the authors talk about 'reconstruction of sSSY', they should clarify well that the analysis of sSSY is based solely on simulations of SSC derived with a QRF model. Not only the QRF model cannot reproduce values outside of the range of values of the training dataset, but also the processes of sediment production and transport might have changed over time. Second, given the nature of the model, it is expected that trends and changes in the predictors lead to trends and changes in SSC. Therefore, I suggest that the authors discuss the trend analysis of the predictors before or together with the trend analysis of sSSY .

Answer: Thank you for this valuable comment. To avoid being misread, we suggest to replace the term "reconstructing" with "estimating". Of course, the processes of sediment production and transport might have changed over time. That is why we designed the predictors in a way that they can be seen as proxies for these processes. For example, sediment production in these areas will be a function of temperature (glacier melt and movement, sub- and proglacial sediment transport, potential permafrost thaw), as well as potentially precipitation and antecedent moisture conditions (hillslope erosion, slope destabilization) and sediment transfer is tightly linked to discharge. As mentioned in the answers to reviewer 2, it is not necessarily the case that changes in the predictors lead to trends and changes in SSC, as QRF is not a linear model. Since we analyze trends and change points in the predictors but also in the glacier mass balances – which are independent data and not part of the predictors – we suggest to leave the order as it is, since it would otherwise be necessary to open a new chapter for the glacier mass balances.

I think that it would be interesting to quantify the trends and shifts in SSC, to analyse how much the change in sSSY is related to a change in discharge, in SSC or in both. This would allow understanding if the increase in sediment load is due to an increase in transport capacity, in sediment supply or a combination of the two.

Answer: Thank you for this comment. We agree that it is an interesting question. To put it briefly, we find roughly the same change points and trends in mean annual SSC as in sSSY (see figure 1 in attached pdf): mcp yields change point around 1980/1981 for both locations, while the Pettitt test disagrees in Vernagt (likely due to its limitation with respect to the beginning and end of time series). However, mean annual SSC considers low concentrations in spring (when discharge and flux are also low) with the same importance as days in August (with higher SSC but also much higher discharge and higher fluxes). That is why we focus on annual yields (and changes therein), because we find them to be more adequate and meaningful to aggregate to annual resolution.

In both Validation A and B, models fail to capture the largest SSC values. As discussed by the authors, this is likely related to the inherent limitation of QRF in extrapolating beyond the range of values of the training data. It would be important to quantify the impacts of this limitation on the total suspended sediment yield. I suggest that authors compute the fraction of total suspended sediment yield transported during these 'extreme' days.

Answer: Thank you for this important point. It makes us aware, that we need to improve our explanations and point out more clearly, that both figures you are referring to show out-of-bag estimates, i.e. the model estimate for a day with very high SSC is based only on those trees, that do not "know" this day. Thus, it can be seen as a quite rigorous validation, and means that the performance of the full model for these days (or days with

similar conditions) will be better.

We calculated the difference between daily Q_{sed} based on turbidity and daily Q_{sed} from the out-of-bag model estimates of the full models, to assess the underestimation in annual SSY for the 10 days with the highest Q_{sed} in the turbidity time series. The underestimation on these 10 days represent 0.6 to 2.8 % of the annual SSY at gauge Vernagt and 1.7 to 19.1 % of annual SSY at gauge Vent. However, the 19 % underestimation stem from the most extreme event in the time series in August 2014, where 26 % of the annual SSY was exported within 25 h – likely associated with mass movements (Schmidt et al., 2022). *The full (i.e. non-OOB) model estimate for this day only shows an underestimation of 6 %. We will add that to the discussion.*

It is not clear to me, which is the added value of using P and T? This could be quantified by running the QRF models excluding either precipitation or temperature and evaluating their performances. Likewise, I think that it would be interesting to run the QRF model without discharge. This would contribute to understand the relevance of the predictors and to estimate the potential of using such models in ungauged catchments.

Answer: Thank you for this interesting question. Variable importance can be analyzed for QRF models, by interpreting the so-called "variable importance" for the related RF-models, e.g. by quantifying the decrease in model performance, if a predictor is permuted (see figure 2 in attached pdf). At both gauges, discharge is the most important predictor, but at gauge Vent, temperature and the derived predictors and the day-of-year are also above 10 % IncMSE (average increase in squared residuals if the variable is permuted). Precipitation and the derived predictors (such as precipitation of the antecedent 24 and 48 h) are less important. At gauge Vernagt, short-term precipitation is also less important, but long-term antecedent precipitation (up to 53 days) is the second most important predictor.

However, the interpretation of these analyses is not straightforward because the predictors are partially correlated (as can easily be imagined with temperature and discharge, as glaciers melt) and thereby "share" some importance. That implies that if we perturb one predictor, some of the information would still be present in the correlated predictor. Secondly, predictor importance is also likely to vary throughout the season, calling for a more elaborate analysis. Thus we suggest not to include this in the manuscript.

We would not recommend applying QRF in ungauged catchments, firstly, because discharge is a very important predictor, and secondly, because the model needs to be trained for each site, and needs training data, also and especially of suspended sediment concentrations.

Specific comments:

Ln. 166-171: Which is the resolution of the discharge data? Please, specify.

Answer: The answer is a bit complex, there are different periods of times for the gauges where different temporal resolutions are available. That is why we did not state it here but prepared the table in the Appendix. We will add a reference to it here.

Ln. 174-176, Ln. 181-183: How did you compute the 'conversion factors'? Over which time period?

Answer: We derived linear relationships between e.g. the available Temperature at gauge Vent and Vernagt for all dates when data were available at both measurement stations and used this linear model (as stated in the brackets in the text) for conversion. We will clarify this in the revised manuscript.

Ln. 179: which resolution?

Answer: 60 min resolution since 1974, 10 min in 2000 and 2001 and 5 minutes ever

since. We stated this in the table in the Appendix and will add a reference to it here.

Ln. 234-244: Please, in addition to the reference to Zimmermann et al., 2012 provide clarification for the antecedent predictors.

Answer: Thank you, we will make this more clear.

Ln. 208-210: How did you disaggregate the data? How did you use the 10-min data? In the gap-filling part?

Answer: Thank you, we will provide details on how we disaggregated the data. The disaggregation only refers to the gap-filling model at gauge Vent, where precipitation and temperature data from 2000 and 2001 had to be disaggregated from 60 to 10 min. resolution. Discharge and temperature, that were given as hourly means before, were adopted as is for the 10 min timesteps and precipitation sums were divided by 6. Exactly, the 10-min data were used in the gap-filling model. We will clarify this.

Ln. 211-214: I find this paragraph confusing. Please, clarify.

Answer: Thank you, we will clarify this.

Ln. 259-265: Please, move this chapter to chapter 3.1

Answer: You probably refer to the chapter up until line 285 and including figure 2 and are suggesting to first describe the general approach, then the data and then the model in detail? Thank you, we will do that.

Ln. 272-274: Does it make sense to first use a model to estimate the SSC data, and later use the modelled SSC to estimate a model? I think that it would be more correct to exclude from the QRF the time steps in which SSC is not available.

Answer: Thank you for this question. For our purpose, these steps were indispensable to supply the model with the full range of values, which were originally lost in the limited turbidimeter data. Moreover, the modelled SSC from the gap-filling model are only a small part of the training data of the validation and reconstruction models, and we train the gap-filling model on different data, i.e. suspended sediment concentration samples instead of turbidity. These samples include times when the turbidity probe failed but also when it reached saturation. Keeping in mind the range-sensitivity of QRF, we argue that it is important to add these data.

Ln. 277-278: Did you train the QRF models on all available data? Please, clarify.

Answer: Thank you, we assume you are referring to the models in Validation A? Yes we did and we will clarify.

Ln. 281-282: Is this model different from the daily model of Validation A?

Answer: Yes, it is different (see above), we will make this more clear.

Ln. 291: Please, clarify q-weighted.

Ln. 300: Please, clarify equation 2.

Answer: Both comments refer to the Q-weighted SSC. We will clarify.

Ln. 349-350: I understood that at gauge Vernagt predictors were available at hourly resolution (see Ln. 259-260). Do you mean that the predictant, SSC, is daily? Please, clarify better in chapter 3.1. Data availability and resolution is very confusing.

Answer: Thank you, we will clarify this. It is correct that predictors at gauge Vernagt are available in hourly resolution, but at gauge Vent, long-term data are only available in daily resolution. Since we wanted to ensure comparability between the gauges – and Validation A showed that the loss of model skill is small, we focused on daily resolution, which also helped keeping computational times reasonable.

Ln. 369-370: I agree that NSE and BE are quite good, in the context of suspended

sediment transport. However, I wonder how much the largest values, which are substantially underestimated by the model, contribute to the total suspended sediment yield. Quantifying this would help assessing the model performance.

Answer: *Thank you; we answered this question above (in our third answer to your comments).*

Ln. 385-401: Please, move this chapter to the chapter about data (3.1).

Answer: *Thank you, this has also been suggested by reviewer 1. We will do that.*

Ln. 471: CP is not defined previously.

Answer: *Thank you, we will define CP here.*

Ln. 476-477: please, describe more in details the mass balance record.

Answer: *Thank you, we will do that.*

References

Schmidt, L. K., Francke, T., Rottler, E., Blume, T., Schöber, J., and Bronstert, A.: 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, *Earth Surf. Dyn.*, 10, 653–669, <https://doi.org/10.5194/esurf-10-653-2022>, 2022.

Please also note the supplement to this comment:

<https://egusphere.copernicus.org/preprints/2022/egusphere-2022-616/egusphere-2022-616-AC3-supplement.pdf>