Dear Genevieve Ali and referees,

Many thanks for another valuable round of reviews. We highly appreciated the opportunity to improve our work. We remain eager to address any further comments and suggestions.

First, we would like to state that we have uploaded the model input and output data, raw discharge and sediment concentration data, and the R scripts used for i) adjusting the sediment rating curves and propagating the regression uncertainty, ii) running the WaTEM/SEDEM model, iii) performing the random forest analysis (Figure 5), and iv) summarising model results (Figure 7) to the Zenodo data repository (https://doi.org/10.5281/zenodo.6560226). We hope the data, along with our responses and revisions, will address the concerns from Reviewer #2 regarding the scientific quality of our work.

Below we answer to all your comments. For clarity and improved visualisation, the editor and reviewer comments are shown from here on in black. The authors' replies are in blue font below each of the reviewers' statements. The changes in the revised manuscript are displayed in green. Line numbers refer to the revised tracked manuscript.

Editor's comments:

Dear Authors,

Two referees provided comments on your revised manuscript. One reviewer recommended that some technical corrections be made. The other reviewer was of the opinion that your manuscript was much improved, compared to the previous version, but they raised that some comments from the previous round of review had not been properly or sufficiently addressed. I do agree with their two main major comments:

1) It is not appropriate to state that data are not available if they are available (with respect to crop information, in your case). Unless the statement made to that effect in the revised manuscript was not properly phrased and therefore not purposefully misleading. If data are available in the right format, then it would be important to incorporate them into your manuscript. If some data are available but they do not meet a quality criterion (or a spatial coverage criterion or a spatial resolution criterion) that you are using, then it would be better to state so explicitly.

We completely agree it is inappropriate to state that data are not available if they are so. Our reasoning to state that open source geodata on crop information were not available within our study catchment was based on our interaction with the Swiss Geoservices portal on 09-10.08.2021. After the first review round, we contacted the geoportal to request access to the data mentioned by Reviewer #2 in their initial comments. However, as stated in our previous response letter, we were told the data were not freely available and access would cost 4,260 CHF (even though we mentioned the data would be used for research purposes). We can provide proof of this interaction with the Geoservices portal, if the reviewers and the editor would like so, in order to clarify any misunderstandings. We had no intention to make any sort of misleading statements, and we apologise if that was somehow conveyed.

The parcel-specific crop data for the Baldegg catchment are indeed now freely available, as stated by Reviewer #2, and we were able to access it. As mentioned by Reviewer #2, canton Lucerne has changed their open data policy since our inquiry, and we were unaware of that. We are highly thankful to the reviewer for pointing this out.

However, as also mentioned by Reviewer #2, including such data to our manuscript at this point would be of little added value – mostly because i) we have dealt with the uncertainty in the landuse classification within our Monte Carlo simulation; ii) our aim with this paper is to investigate the effects of linear structures on sediment transfer and to explore connectivity scenarios, not to produce an ultrarealistic soil erosion map for the Baldegg catchment or to perform a rejectionist model test; and iii) although crop data from 2021 are available, we would still need to make a lot of assumptions for a model representing long-term (~20 years) average annual conditions, which would lead to very a similar parameter sampling approach to one we employed so far.

Hence, we have removed the sentence about data availability and explain our rationale as it follows:

"Due to the difficulties involved in accurately representing long-term average agricultural landuse patterns and farming management practices per field parcel, pastures and cropland were considered a single arable land category, using only the information available from the land cover map (Table 2) (Swisstopo, 2018). In this case, minimum and maximum values were relaxed to represent a wide possible combination of crops and support practices. Such combinations were assessed with the *CP*-Tool (Kupferschmied, 2019), which allows for the calculation of *CP* values considering common crop rotation systems in Switzerland. The minimum *CP* values were particularly reduced to include typical values for permanent grasslands in Switzerland (~0.01) (Schmidt et al., 2018b). This simplified approach should be appropriate considering i) our focus on connectivity scenarios and linear landscape structures, and ii) the use of the Monte Carlo simulation with the sampling of a wide parameter space that accounts for the uncertainty in the landuse classification". (L250-261)

We have also included additional information about the temporal scale represented by the model:

"The model is by default executed in an average yearly time step, as typical in RUSLE applications, which predict long-term (~20 years) average annual soil losses (…)" (L188-189).

"Suspended sediment loads are a product of a complex interaction of hillslope and channel remobilisation processes, which are not fully represented by WaTEM/SEDEM. In addition, since the model is RUSLE-based, the soil redistribution rates represent long-term average annual values, which hampers a straightforward comparison with annual sediment transport rates." (L331-333)

Since we have not used the parcel-specific crop data, we believe it would be confusing to the readers to mention such data in the manuscript. However, if the editor and the reviewers think it would be important to mention the availability of the parcel-specific crop data, we would of course be happy to do so.

2) Some of the comments made by one reviewer regarding the results reported in Figure 8, Table 4 and Figure 5 need to be addressed/discussed in the manuscript, because other readers may raise similar questions (personally, I know that, I, too did not understand how a better model fit could lead to a larger percentage of out-of-bound observations).

We apologise for the blunter in Figure 5 and for not providing a more thorough response regarding Table 4/ Figure 8. We have corrected the mistake in Figure 5, provided a detailed explanation to the reviewer's question, and addressed the topic of the rating curves and the percentage of out-of-bound observations in the text. Please see our response to Reviewer #2 below.

Reviewer #1

We are highly thankful for another round of thoughtful comments. We have incorporated your suggestions to the manuscript and replied to all your questions below.

L 45: Not to mention transport of phosphorus, as you point out below.

Thank you for pointing this out, we have included a mention phosphorus delivery to surface waters (L 42).

L 49: Is the transport to downstream systems not considered part of sediment connectivity? I guess it would depend on the defined spatial and temporal scale.

Indeed, we believe this depends on the spatial and temporal scale of the analysis, and perhaps on how one interprets the definition from Heckman et al. (2018). But in general, we understand that sediment connectivity research is more focused on internal sediment transport processes than on catchment sediment yields, or sediment export to downstream systems.

L 149: A bit more information on the inclusion of a hysteresis parameter could be included here - Is this referring to hysteresis in the sense of Sheriff et al. 2016? Sherriff, S. C., Rowan, J. S., Fenton, O., Jordan, P., Melland, A. R., Mellander, P. E., & Huallachain, D. O. (2016). Storm event suspended sediment-discharge hysteresis and controls in agricultural watersheds: implications for watershed scale sediment management. Environmental Science & Technology, 50(4), 1769-1778.

Thank you, we agree we could have included some more information regarding hysteresis and seasonality in sediment export patters and rating curves. Following your comment, we made the following adjustments "The rating curve partially accounts for hysteresis and seasonality (Table 1), which can have a significant impact on sediment export patterns and reflect the catchment landuse, hydrological connectivity, and internal sediment source dynamics (Sherriff et al., 2016). To derive the parameters in equation 1 we used a parsimonious multivariate regression which does not require separate calibration for different seasons (Cohn et al., 1992; Vigiak and Bende-Michl, 2013)." (L149-153)

L 230: It might be useful to state why LS isn't varied – I'm guessing there isn't too much uncertainty associated with the DEM or the LS equation?

Apologies for the missing information and thanks for raising this question. We do believe the LS equation is a source of uncertainty, although a lesser one compared to the CP and K factors (see Batista et al., 2021). However, since we used a high-resolution DEM, the variance associated to the LS factor is much reduced, compared to the typical set up in which soil redistribution models are applied. We included an explanation to this matter in the text: "Finally, the LS_{2d} factor was calculated with a slope (rad) and an upslope contributing area (m^2) grid, which were obtained by processing a 2 m x 2 m resolution DEM from SwissALTI3D (Swisstopo, 2014a). In this case, the error in the LS_{2d} factor was not incorporated into Monte Carlo simulation due to the use of the high-resolution DEM, which should considerably reduce the uncertainty associated to the parameter estimation." (L261-265)

L 417 – 473: this is a bit confusing. Is there a typo here?

Thank you for noticing this. We corrected to: "In a wider context, our study has demonstrated how structural sediment connectivity patterns can be investigated with a conceptual model such as WaTEM/SEDEM, provided that model spatial resolution is sufficiently fine to represent relevant features and processes." (L503-505)

Reviewer #2

General comments

I would like to thank the authors for revising the manuscript and for answering most my questions. The quality of the manuscript was improved largely by the revisions. However, some of my comments in the first review round were in my opinion not sufficiently addressed. I think the topic and scope of this manuscript are highly scientifically significant and I would like to see this paper published at some point. However, especially the answer to the second insufficiently addressed comment (regarding Figure 8) makes me question the scientific quality of the manuscript, and I think the manuscript still needs major revisions. However, I may be on my own with this judgement and I am looking forward to see the opinion of the other reviewers and of the editor.

Thank you for another thorough review of our manuscript, we highly appreciate it. We are sorry our replies from the previous review round were did not sufficiently address your concerns. As stated above, we have uploaded all our data and code to the Zenodo data repository, making it fully transparent and reproducible. We hope this, in combination with our responses below, will convince you of the scientific quality of our work. We highly welcome the constructive criticism and remain eager to address any points you find necessary. Below we provide more detailed responses to your comments.

Insufficiently addressed comments from the first review round

1) L251-252: I still strongly disagree with your statement that open source geodata of crop statistics are not available for the Baldegg catchment. This is simply wrong. As mentioned in the previous review, these open source geodata are actually available in high resolution. They can be downloaded within few minutes using the link below and can be used for free (creative commons by license): https://geodienste.ch/services/lwb_nutzungsflaechen (Additional information is provided here: https://daten.geo.lu.ch/produkt/lwnfmgdm_ref_v1) Since in the dataset some of the agricultural areas (probably 10-20%) are missing a crop classification, you would still have to take some assumptions, but you would be able to represent the cropping reality much better.

The canton of Lucerne recently changed his law with regards to open data. Therefore it is correct that most probably this dataset was not completely open source by the time when you handed in the manuscript (e.g. for commercial use there was indeed a fee which had to be paid). However, I am pretty sure that the dataset was already available freely for governmental institutions and universities by that time. Here, I must say that I have the feeling that, after being told about this dataset in the first revision, the authors did not really put a lot of effort into obtaining this dataset. I understand that using this dataset would mean a lot of additional effort for the authors, and that probably the added benefit to the manuscript would be small. However, it should be at least stated correctly in the manuscript that this dataset is available, but was not used due to reason XY.

We apologise if we gave the impression that we did not put effort into accessing the crop data. Following your comments in the previous review round, we contacted the Swiss Geoservices portal on 09.08.2021, requesting access to the "Landw. Bewirtschaftung: Nutzungsflächen" data for the Baldegg catchment, for scientific purposes and for the University of Basel. We received a reply on 10.08.2021, stating that the data would cost 4,260 CHF, unless the research was being done on behalf of or in cooperation with Canton Lucerne. This was our reasoning for stating in the revised manuscript that open source geodata on crop statistics were not available for the study area, as this was indeed the case at the time. We can provide proof of these interactions with the Geoservices portal, if the reviewer and the editor would like so, in order to clarify any misunderstandings.

Indeed, the crop data is now freely available and we were able to access it – many thanks for pointing this out. As you mentioned, however, including this dataset to our analysis at this point would have little added value to the manuscript, as we already dealt with the uncertainty in the land use classification with the Monte Carlo simulation, in which we sampled parameter values from both arable land (considering different crops) and grasslands. In addition, our aim here is not to provide an ultrarealistic soil erosion map for the Baldegg catchment, but whether to explore connectivity scenarios and investigate the influence of linear landscape structures on sediment transfer. Furthermore, although crop data from 2021 are available, we would still need to make a lot of assumptions considering crop rotation (including temporary pastures) for a model representing average annual conditions, which would lead to very a similar parameter sampling approach to one we employed so far.

Hence, we have removed the sentence about data availability and explain our rationale as it follows:

"Due to the difficulties involved in accurately representing long-term average agricultural landuse patterns and farming management practices per field parcel, pastures and cropland were considered a single arable land category, using only the information available from the land cover map (Table 2) (Swisstopo, 2018). In this case, minimum and maximum values were relaxed to represent a wide possible combination of crops and support practices. Such combinations were assessed with the *CP*-Tool (Kupferschmied, 2019), which allows for the calculation of *CP* values considering common crop rotation systems in Switzerland. The minimum *CP* values were particularly reduced to include typical values for permanent grasslands in Switzerland (~0.01) (Schmidt et al., 2018b). This simplified approach should be appropriate considering i) our focus on connectivity scenarios and linear landscape structures, and ii) the use of the Monte Carlo simulation with the sampling of a wide parameter space that accounts for the uncertainty in the landuse classification." (L250-261)

We have also included additional information about the temporal scale represented by the model:

"The model is by default executed in an average yearly time step, as typical in RUSLE applications, which predict long-term (~20 years) average annual soil losses (…)" (L186-187).

"Suspended sediment loads are a product of a complex interaction of hillslope and channel remobilisation processes, which are not fully represented by WaTEM/SEDEM. In addition, since the model is RUSLE-based, the soil redistribution rates represent long-term average annual values, which hampers a straightforward comparison with annual sediment transport rates." (L327-331).

Since we have not used the parcel-specific crop data, we believe it would be confusing to the reader to mention such data. However, if the editor and the reviewers think it would be important to mention the availability of this data, we would of course be happy to do so.

2) L443-445: In the previous review round, I already asked the following question: In the "Ron" stream, the 95% prediction interval seems much narrower than in the other rivers (Figure 8). Therefore, the observed values are mostly outside of this interval and the out of bound percentage is much higher than for the other streams. Can you explain this? Your answer was: The interval is narrower because the model fit was better, and the residuals were lower. This led to a lower proportion of the observed values being outside the prediction interval.

However, in Table 4, you report that the percentage of observations outside of the prediction interval was 61%, which is much higher than for the other catchments. This strongly contradicts your answer. In my opinion, there is something wrong with your fitted rating curve and you should fix this. (It is possible that I simply did not understand your answer correctly. In this case, I would be very happy if you could explain it to me in a version "for dummies". Additionally, I would suggest that add a discussion to the manuscript regarding the influence of this high percentage of observations outside the prediction interval on your results.)

We agree we could have given a more detailed answer to this question, and we are sorry you had to ask again. In all honesty, perhaps we did not give much attention to this question because the sediment load estimates were a somewhat secondary topic in the manuscript. Your questions made us realise that the way the manuscript was drafted put too much focus on the sediment rating curve and the uncertainty estimation of the sediment loads, which in the end created confusion and deviated the reader from our main findings. That is, we only included the sediment load estimates (with a quantification of the regression uncertainty) to provide a generable idea of the plausibility of the model simulations, as we explain in the manuscript. Therefore, below we answer to your questions in detail and explain how we edited the manuscript to improve the clarity of our paper.

First, we would like to point out that we have re-calculated all the sediment rating curves and sediment loads in order to be completely sure there was nothing wrong with the estimates. Our curve fitting and error propagation calculations were correct, but we decided to remove the covariate number 5 from the sediment rating curves, as upon reflection following one of your minor comments, we realised that we ourselves were having trouble understanding the reasoning behind the parameter z in the covariate equation. We had used a value of 15 days to consider short-term exhaustion, as we understood was suggested in Vigiak and Bende-Michl (2013), but in the end we were not happy making this assumption as we were not completely confident regarding the correctness of our interpretation. As you see, the exclusion of covariate number 5 led to small changes in the uncertainty bands from Figure 7 and in some of the evaluation metrics in Table 4. Please note that the differences in the estimated average annual sediment loads, which we compared to model outputs, were minimal. Although some of the evaluation metrics showed a decrease for some streams, we preferred to stick to the covariates we were completely satisfied with and that would not add unnecessary confusion to the manuscript.

Now, regarding your question: the approach we used to quantify the uncertainty in the sediment rating curves is based on the residual errors from the regression models displayed in equation 1/table 1. Once the models were adjusted, we used the *sim* function from the R package *arm* to draw posterior simulations of the model coefficients (the intercept and the coefficients). If the residuals from the initial model fit were large (poor fit), then the posterior distribution of model coefficients was broad. Similarly, if the initial model fit had low residuals, the spread of the posteriors was narrow.

For example, see below the quantile values for the posteriors of the model coefficients for the rating curves for the Mulibach and the Ron. The sediment rating curve for the Mulibach had a poorer fit (Residual Standard Error = 1.2, R^2 = 0.52) compared to the Ron (Residual Standard Error = 0.90, R^2 = 0.60) and hence the posterior distributions of the model coefficients were much wider (in particular for the intercept and x_1).

Stream	Quantile	Coefficients				
		(Intercept)	x_1	x_2	x_3	x_4
Mulibach	0%	8.26	2.01	0.14	-0.73	-0.59
Mulibach	25%	9.12	2.49	0.21	-0.51	-0.48
Mulibach	50%	9.42	2.64	0.23	-0.44	-0.42
Mulibach	75%	9.77	2.85	0.25	-0.37	-0.34
Mulibach	100%	10.64	3.52	0.36	-0.20	-0.09
Ron	0%	3.52	0.91	0.08	-0.38	-0.42
Ron	25%	3.71	1.06	0.14	-0.22	-0.26
Ron	50%	3.75	1.11	0.16	-0.16	-0.21
Ron	75%	3.78	1.14	0.19	-0.13	-0.15
Ron	100%	3.91	1.24	0.27	0.04	-0.03

If we take the range from the posteriors of model coefficients to provide a range of model realisations, then the regression models with greater residuals and poorer fits will have larger uncertainty bands and the other way around. If the uncertainty bands from the model realisations are very wide, they might encompass a larger number of observations, even if the regression model had poorer fit and the predictions from the median of the realisations showed a higher RMSE. Hence, one could understand this a trade-off between accuracy and precision, as seen for instance in the example from the rating curves for the Ron and the Mülibach, and their respective proportion of observations not encompassed by the uncertainty bands, as well as their RMSE, ME, and NSE values (Table 4).

We have now addressed this topic in the text, starting in the methods:

"In order to estimate continuous daily sediment concentration values, later used to produce average yearly sediment loads for each tributary, we used a rating curve approach (Equation 1), combining the roughly triweekly sediment concentration measurements with continuous discharge measurements. The rating curve partially accounts for hysteresis and seasonality (Table 1), which can have a significant impact on sediment export patterns and reflect the catchment landuse, hydrological connectivity, and internal sediment source dynamics (Sherriff et al., 2016). To derive the coefficients in equation 1 we used a parsimonious multivariate regression which does not require separate calibration for different seasons (Cohn et al., 1992; Vigiak and Bende-Michl, 2013)." (L146-154)

"To analyse the uncertainty in the regressions we simulated posterior distributions of the model coefficients (β_0, β_k) with an informal Bayesian function of the R package 'arm' (Gelman and Hill, 2007), as in Batista et al. (2021). This function produces realisations of model coefficients based on the residual standard error of the regression, which means that models with poorer fits will yield broader posterior distributions of regression coefficients. The posterior distributions were used to simulate 1000 sediment concentration values for each day *i*. These were transformed into daily distributions of sediment loads (Mg), considering the mean daily discharge measurements from the gauging stations. Sediment loads were ultimately aggregated into average annual values (Mg yr^{-1}) with uncertainty bands, which should allow for a general comparison with the different sediment connectivity scenarios simulated by WaTEM/SEDEM." (L165-175)

Moreover, we addressed the observations out-of-bound from the regression uncertainty bands in the results. In this case, we moved Figure 8 to the supplementary material, in order to avoid an excessive focus on the sediment rating curves. We think this will improve the readability of the paper and we appreciate how your comments prompted these changes.

"It is important to note that the median daily sediment concentrations calculated from the 1000 realisations of the rating curves (Equation 1) underestimated the high sediment concentration measurements, for all tributaries. This resulted in the positive mean error of the median estimates (Table 4). Moreover, the Nash-Sutcliffe model efficiency coefficient for the median calculations was unsatisfactory considering the usual thresholds for model performance (e.g. Moriasi et al., 2015). However, the 95 % prediction interval of the rating curves encompassed a large proportion of the sediment concentration observations for the tributaries with poorer fits and wider uncertainty bands (i.e., the Höhibach, Mülibach, and Spittlisbach) (Table 4, Supplementary Material Figure 1). The sediment rating curves for the tributaries which displayed a better fit (i.e., the Ron and Stägbach) encompassed a much lesser proportion of the observed sediment concentration values (Table 4, Supplementary Material Figure 1). That is, the regressions with the lowest residual standard errors had narrower uncertainty bands, which albeit produced more accurate median predictions, led to a greater proportion of observations out-of-bound from the 95 % prediction interval. In any case, the largest errors were associated to underestimates of extreme events, and therefore, it is likely that actual sediment loads from the tributaries are contained within the long right side of the skewed distributions resulting from the error propagation of the rating curves (Figure 7), which would increase the overlap with the shortcut scenario." (L425-442)

Major comments

L357: In Figure 5, the MSE increase for CParable is missing. Additionally, the MSE increase reported for road connectivity does not correspond to the numbers in the text.

Apologies for this blunter in Figure 5. The label for the CP_{arable} was missing and the Road Connectivity label was misplaced, thus the divergence from the numbers in the text. This was probably caused by re-ordering the factors in the facet plots. Thanks for picking this up, we completely missed it!

In all honesty the first author could not find the last piece of code he used for this figure, which led to again some new numbers in the random forest analysis importance ranking, due to the randomness of the tree building and because we increased the number of trees (L351-356). This did not change the order of the importance ranking or the interpretation of the analysis. As we mentioned above, the code for reproducing the analysis and the figure is now available, with a set seed to insure the same results in case of another run of the random forest.

Minor comments

L133-134: What are these 10 degrees relating to? Average slope? Please clarify in the manuscript.

Apologies for the missing information. We updated the information in the manuscript to: "Steeper slopes (average values above 10°) (…)". (L123)

L249-250: Do you mean the dataset swissTLM3D (as referenced here) or the Swiss Map Vector 25 BETA (which you mention in your answer to the review)? In the first case, it does not make sense to me to mention the "1:25'000" scale, since – to my knowledge – there is no scale related to the swissTLM3D dataset. Additionally, to improve reproducibility of your work, you should in my opinion specify here that you first converted the roads to polygons by using buffers considering their widths.

Apologies again with the confusion regarding the Swisstopo references. Indeed, the data for the roads was retrieved from the swissTLM3D (Swisstopo 2020), and then converted to polygons using a buffer. This polygon merged with the Swiss Map Vector 25 BETA (Swisstopo, 2018) and rasterised to produce a land cover grid which included roads, tree lines and hedges. We made these corrections and included the information you requested in the text:

"For the *C* and *P* factors, here combined in a single *CP* parameter, uniform distributions were created for each landuse class in the catchment, based on commonly used values from the literature and a land cover map (1:25000) (Swiss Map Vector 25 BETA) (Swisstopo, 2018), which we rasterised to the model resolution (2 m x 2 m)." (L246-248)

"Hedges and tree lines within field blocks were already classified in the large-scale topographic landscape model of Switzerland (swissTLM3D) (Swisstopo, 2020) and required no additional processing apart from a merge with the land cover map (Swiss Map Vector 25 BETA) (Swisstopo, 2018). Furthermore, three road connectivity assumptions were assessed for each model iteration. For such, we first converted the roads from polylines (as available in the swissTLM3D) to polygons, using a buffer distance based on the road widths. Next, these polygons were rasterised and incorporated into the land cover grid used for modelling." (L289-296)

Finally, we included the correct citation for the Swiss Map Vector 25 BETA to the reference list:

Swisstopo. Swiss Map Vector 25 Beta, Das digitale Landschaftsmodell der Schweiz. 2018.

L357: Explain the abbreviation RFA here. I know that you explain it in the text, but each Figure should be readable by itself.

Thanks for noticing this, we spelled out Random Forest Analysis (RFA) in the figure caption:

"Figure 5. Mean squared error (MSE) increase associated to model input factors for the Random Forest Analysis (RFA). Larger relative errors indicate the input factors were more important for estimating model outputs." (L359-361)

L365: From what you write in the manuscript, I did not make the link to Table 3. This was only clear to me after your explanations to my questions. I therefore suggest to refer to Table 3 here.

Thank you, we now refer to Table 3 (L370).

L367: The combined use of hyphen and minus makes the legend of Figure 6 rather confusing. Please use another symbology for this (e.g. use "to" instead of the minus).

Thank you for the suggestion, we changed the symbology in the legend from Figure 6 using "to" instead of minus.

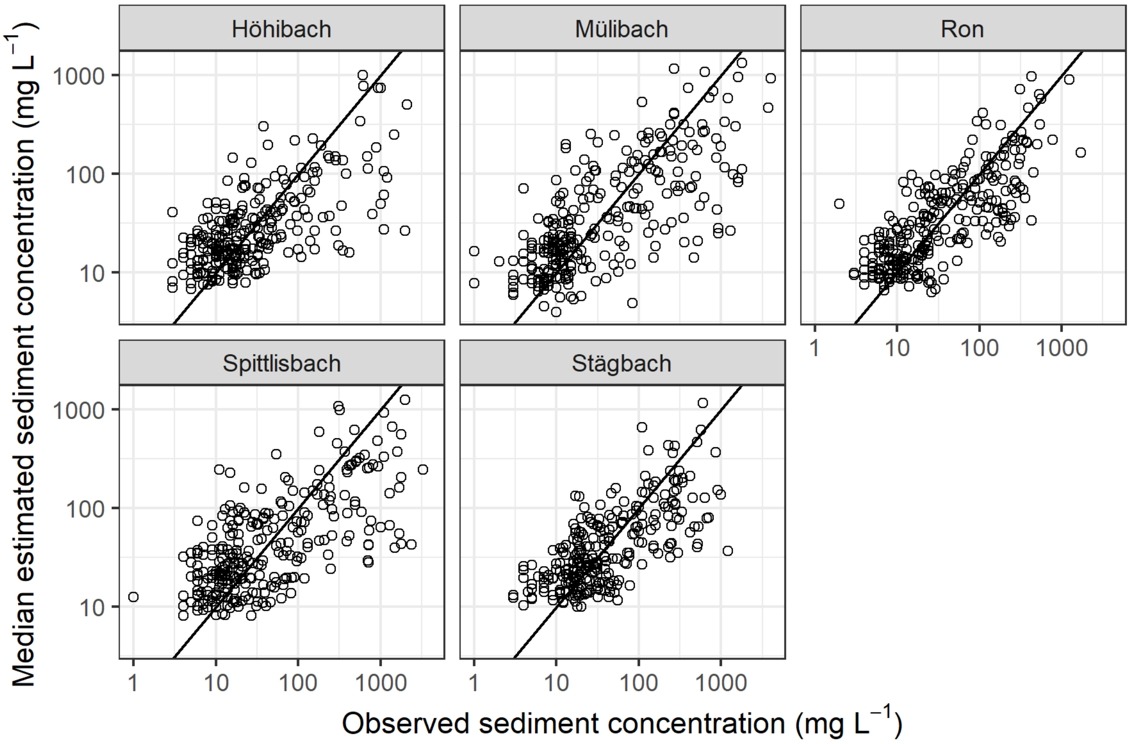
L420: Figure 7 is now much more convincing to me than it was in the previous version. Thanks for the adaptations.

Thank you for suggesting these adaptations, we agree this is more convincing.

L428-429: An additional question to your model came up when trying to understand why the model does represent the Ron catchment well. In L428-429 you write: “It is important to note that the median sediment concentrations calculated by the rating curve (Equation 1) underestimated the actual observations, for all tributaries.” If your model aims on predicting mean annual sediment loads, should the median of your model not have a mean error of approximately zero?

Again, we apologise, as we could have been more precise regarding these statements. The regression model is used to estimate daily sediment concentrations, which are then used to calculate the sediment loads, in combination with the daily water discharge measurements (please see our response to your previous question about the rating curve above).

An unbiased model should indeed have mean error of approximately zero, but, as we now highlighted in in the text, the regression models particularly underpredicted the high sediment concentrations (L423-424, L435-436). Hence, there is positive mean error (observations higher than predictions), even for the median of the realisations of the rating curves. This can be also visualised in the figure below (solid line is the perfect fit), which we prepared for the reviewer:



L476-478: Thanks for clarifying which characteristics you analysed. You should not only mention that you determined the corresponding numbers for each catchment (stream density, road density, fraction covered by agricultural land/forest/infrastructure), but also provide them somewhere in the manuscript or in the supplementary information. Furthermore, it is unclear to me what you mean exactly with land cover. Do you mean “(agricultural land, forests, and infrastructure (e.g. settlements, developed areas, and roads)”, as you mention it in L113-114? In the next sentence, you are hypothesizing about the influence cropping specifities, which also relate to land cover. This makes this paragraph rather confusing. You should state more clearly, what you mean with "land cover" here.

Apologies for not including such information beforehand. We supplied the characteristics of the sub-catchments in the supplementary material:

Supplementary Material Table 1. Land cover and physiographic characteristics of the analysed sub-catchments draining into the Lake Baldegg.

Attribute		Sub-catchment				
		Höhibach	Mülibach	Ron	Spitllisbach	Stägibach
Land cover (%)	Arable land	72	56	71	52	76
	Forest	13	40	12	25	8
	Infrastructure	4	3	12	11	10
	Orchards	11	1	4	12	6
	Other*	0	0	1	0	0
Road density (km km^{-2})		5.9	6.0	6.9	8.4	6.7
Stream density (km km^{-2})		2.0	3.0	1.9	3.0	1.2
Mean elevation (m a. s. l.)		617	708	568	634	545
Mean slope (%)		12	16	9	12	9

* Water surfaces and rock outcrops.

We refer to the table in the text and explain the following: "Of note, the contrasting results for the Höhibach sediment loads (Figure 7), which are much closer to the sink and patch-connector simulations, do not seem to be explained by any physiographical characteristic of the sub-catchment (Supplementary Material Table 1). Hence, we speculate that this different pattern could be caused by a lower inlet drainage density or specific farming practices within the Höhibach contributing area." (L480-485)

Technical corrections

L147-149: Not sure if this sentence is grammatically correct.

We rephrased to: "On average 274 grab samples were taken from each tributary, which corresponds to one sample every 22 days, in addition to the samples collected during high-flow events (10 – 13 per year) (Figure 3)". (L136-138)

L154: y-axis: Sediment Concentration -> Sediment concentration

We changed the y-axis title in Figure 3 to "Sediment concentration".

L157: Equation -> equation

In this case we kept the capital letter, as in the rest of the manuscript.

L162: xk -> xk,i

Updated to $x_{k,i}$ (L155).

L165: The variable z (row 5) is not explained. Please explain.

As explained above, we removed covariate number 5 from the rating curve.

L171: R package version? L321-322: R package version?

Sorry we did not add these package versions before. They are now available in the references:

Gelman, A. and Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge Univeristy Press, New York., R package version 1.12.2, 2007.

Liaw, A., Wiener, M.: Classification and regression by randomForest. R News, 2, 18–22, R package version 4.7.1, 2002.

References

Batista, P. V. G., Laceby, J. P., Davies, J., Carvalho, T. S., Tassinari, D., Silva, M. L. N., Curi, N. and Quinton, J. N.: A framework for testing large-scale distributed soil erosion and sediment delivery models: Dealing with uncertainty in models and the observational data, Environ. Model. Softw., 137, doi:10.1016/j.envsoft.2021.104961, 2021.

Heckmann, T., Cavalli, M., Cerdan, O., Foerster, S., Javaux, M., Lode, E., Smetanová, A., Vericat, D. and Brardinoni, F.: Indices of sediment connectivity: opportunities, challenges and limitations, Earth-Science Rev., 187(December 2017), 77–108, doi:10.1016/j.earscirev.2018.08.004, 2018.

Vigiak, O. and Bende-Michl, U.: Estimating bootstrap and Bayesian prediction intervals for constituent load rating curves, Water Resour. Res., 49(12), 8565–8578, doi:10.1002/2013WR013559, 2013.