We thank the referee for his constructive review. Our replies are embedded in bold italics in his original comments.

1 General Comments

This paper evaluates coarse sediment transfer in the Strimm creek, Eastern Italian Alps, using a large number of PIT-tagged particles. The results of the measurements are used to estimate, via the virtual velocity approach, the total and annual volume of bedload transfer. The 3-year measurements include snow-melt-dominated periods, events of (heavy) precipitation in summer, and mixed periods; hence, they are analysed with respect to the hydrogeomorphic regime between the surveys (in total 8 and 10 in the two sections, respectively). Another focus of the paper lies in comparing the bedload transfer in the Upper vs. Lower section of the Strimm creek (termed US and LS, respectively) that are separated by a bedrock step and differ with respect to their morphology and lateral sediment connectivity.

The paper presents an original study that I consider very relevant for the scope of ESurf. It is generally very well written and contains an appropriate number of tables and figures of good quality. The conclusions are well supported by the results. I suggest that the manuscript be accepted for publication in ESurf after moderate revisions. My biggest concern is that the step from weight-classed Q thresholds and virtual velocities to "overall" volumetric transport rates is poorly described.

2 Specific Comments

The *introduction* contains a more technical part related to bedload transport estimation and measurement, and a part that refers to the application to landscape evolution, specifically with respect to the issue of glacial inheritance.

• p419 l26 pls define "relative sediment transport rates"

With the term "relative" we intended season-specific sediment transport rates, but we agree with the reviewer that using such an adjective is not necessary here. The wording was modified to "sediment transport rates".

Study area section

• p422 l10: pls explain on what basis (single-year or short period measurements ? annual precipitation regressed on elevation ?)

The estimation draws on several multi-year studies – still unpublished – on precipitation variation with elevation in the Venosta valley, by means of XPOL radar and rain gauges placed both within the Gadria-Strimm basin and in nearby catchments. As worldwide in mountain areas, the biggest issue is the assessment of winter precipitation. We think more details in the ms are not that useful as mean annual precipitation is not relevant for data analysis.

• p422 l22,26: you give the slope range for the "valley floor" - the definition and delineation of the valley floor influences the slope values found within this area. Was it delineated using a geomorphological map ? Is there always a conspicuous breakline between the valley floor and the hillslopes ? Moreover: In l22, you write "the large hummocky moraine" without having introduced this landform and its position in the landscape. I'd suggest that you either do so, or write "a large...".

We realize that the original text was misleading. Slope values refer to the Strimm channel flowing through the hanging valley floor. The relevant text has been modified as follows:

"Perennial fluvial transport regime characterizes the channel reaches draining the lower half of the hanging valley floor (channel slope range: 1–8 %)". The large hummocky moraine, since there is only one, has been labelled in Figure 1.

• p423 l7,9: do the slope values refer to the channel gradient (i.e. to the longitudinal profile) or to the valley bottom (i.e. to an area) ? Pls clarify

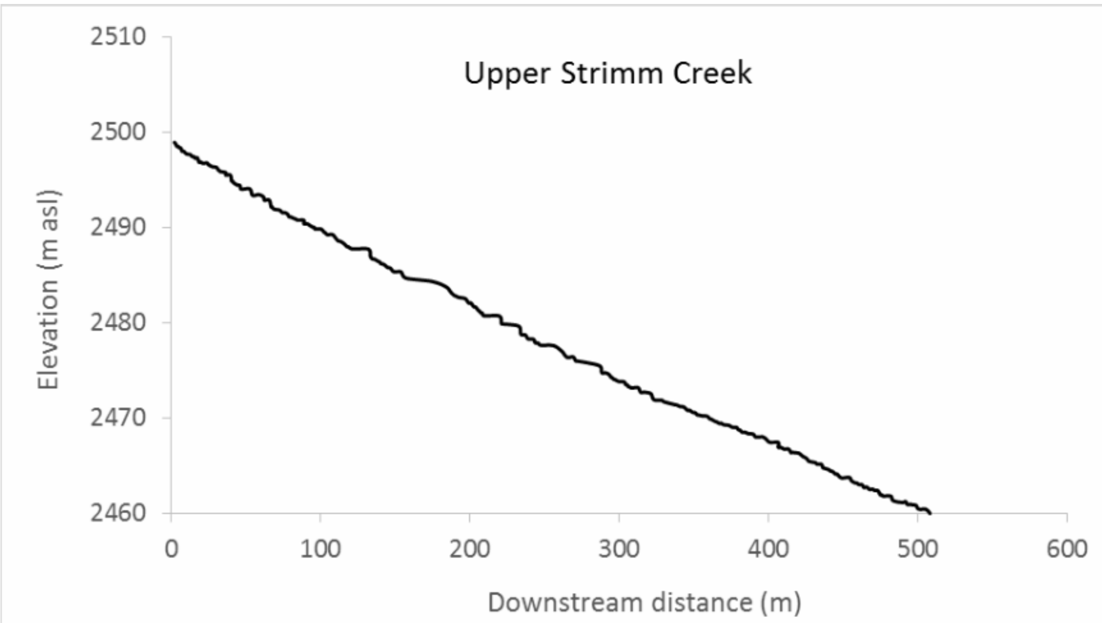
Same as above, the slope range refers to in-channel slope values. The new text now reads as follows:

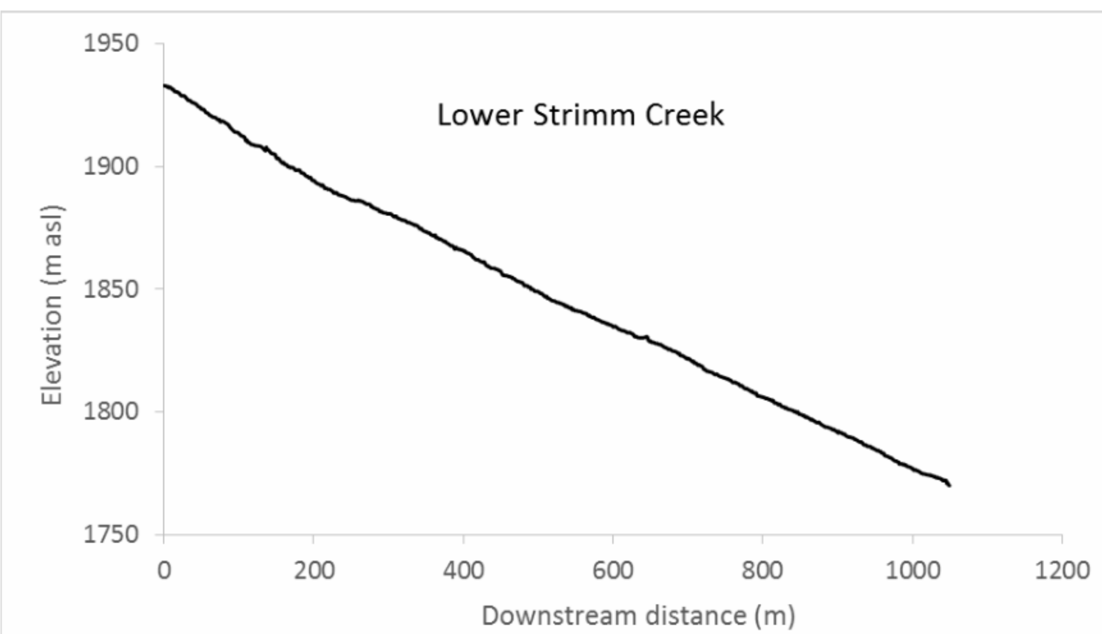
p422 l26: "In the middle part of the basin (slope range of channel main stem: 10–25 %) [...]"
p423 l7-9: " This is the steepest part of Lower Strimm, with channel slope ranging between 12 and 42%".

• Can you offer any evidence that the two sections are reasonably homogeneous and representative?

In terms of representation, the two channel "sections" are representative in that they constitute about 50% of Strimm Creek channel length that could be monitored. Specifically, in US it was not possible to move further upstream as the channel becomes ephemeral. In LS, we avoided channel stretches containing colluvial channel junctions (see location of red linework in Figure 1) and stretches affected by the presence of check dams (i.e., located downstream of R1; Figure 1). In addition, downstream of R1 the channel becomes too steep to wade.

In terms of homogeneity, we show that US and LS consistent surficial grain-size distributions (Figure 5). To address the reviewer's concern we report the longitudinal profiles of US and LS obtained by field-based topographic surveys. These profiles show no discontinuity and/or break in slope within US and LS. In addition, the dominant channel morphology stays the same within each of the two study areas. We do not think that it is necessary to add the two field-based long profiles in the manuscript, but we would be happy to do so, should the reviewer (or the editor) have a different opinion.





Methods section

• p423 l23 consider re-wording "flows in WS1 are on average 60% higher"

Comment accepted, the text has been changed accordingly.

• p424 l5ff: you indicate the range and the SD (please introduce this abbreviation; I assumed it is "standard deviation") of grain sizes. I think that you should give the mean also, because the SD directly relates to it. Moreover, the SD has the same units as the range (and the mean). The value of SD (2.1 mm) sounds very low to me for a total of 34 sediment samples with a range across three orders of magnitude. Is it correct ?

Comment addressed. Since the mean is not customarily used when dealing with grain size distributions, we have decided to provide additional quantitative information by reporting the D_{16}, D_{50} and D_{84}. These percentiles are immediately comparable to published values.

The manuscript was modified as follows:

"In the upper study area, sediments range from 4 to 512 mm, with D_{16} = 26.7 mm, D_{50} = 62.3 mm, and D_{84} = 118.6 mm. The lower study area is comparatively coarser, with b-axis varying from 5.6 to 1024 mm, and a much more heterogeneous distribution (D_{16} = 24.8 mm; D_{50} = 76.1 mm; D_{84} = 281.3 mm)"

• p425 l23: pls indicate how the relative position was measured (tape measure? laser distancer?)

The relative position was measured with a laser range finder; lines in the manuscript describing this methodology have been modified as follows:

"The relative position of each clast was measured by means of a TriPulse 360 B laser rangefinder (Laser Technologies Inc) from a series of georeferenced control points along the topographic net created during the 2010-2012 surveys, so that, after each survey, it was possible to derive the absolute position of the clasts."

• p426 l2-4: consider re-wording; what are "statistical relations within site" ? Can statistical relations exhibit scatter?

The text has been reworded as follows:

“Even though we have measured the three axes of each tracer stone, we will show the results of travel distance in relation to tracer weight only. This is because exploratory analysis of travel distance distributions in relation to grain-size classes displayed high variability among classes (likely due to their highly variable shapes), and assessing the effect of particle shape on sediment transport is beyond the scope of this work.”

• p426 l7: how do the (arbitrary ?) classes compare to the empirical distribution of grainsizes (weights) of the natural sediment ?

In Figure 3 we report the grain size distributions of the natural sediment (Figure 3a) with those of the released tracers (Figure 3b). The size distributions of the tracers mimics the natural ones, in that tracers released in US were comparatively finer than in LS. The following limitations apply: (i) given the dimensions of the PIT tags (23 mm long) it was not possible to drill holes in clasts with a B-axis < 23 mm (cf.,Table 1); (ii) in LS the tracer range does not cover the coarsest natural fraction (B-axis > 230mm).
On p 424 line 26 of the original manuscript we state: "The grain size distributions of the tracer particles released in the two study sites mimic the coarser nature of LS in comparison to US (Fig. 3b), with the minimum tracer size dictated by the size of the PIT tags (i.e., we cannot trace clasts with B-axis < 23mm)".
Because we have not conducted any bulk grain-size sampling, we cannot perform a direct comparison between the weight distribution of the natural sediment and the tracers. To partly address the reviewer's concern we add numbers and percentages of the tracers in the different weight classes (see amended Table 2). In addition, the sentence on p424 line26 is now followed by a statement where we acknowledge that the coarser tail of natural grain sizes is not traced: "While in US we are able to trace the entire natural range of grain-size variability (i.e., > 23mm), in LS for logistical reasons we could not tag clasts comprised between 230 and 1024 mm.

“In our analysis, the tagged particles were grouped into eight weight categories, (W1 to W8, Table 2; Table 2 also provides information about the clast count / weight class both at LS and US)”

Caption in Table 2 has been also modified as follows:
“Table 2 – Tracer clast distributions across weight categories (g) in US and LS.”

W1	W2	W3	W4	W5	W6	W7	W8		
<200	201-400	401-600	601-800	801-1000	1001-1500	1501-3000	> 3000		
72	91	43	13	7	6	0	0	**clast count**	**US**
31.03%	39.22%	18.53%	5.60%	3.02%	2.59%	0.00%	0.00%	**%**	
21	48	44	35	28	36	16	4	**clast count**	**LS**
9.05%	20.69%	18.97%	15.09%	12.07%	15.52%	6.90%	1.72%	**%**	

• p426 l15-21: I'd suggest that you add the Q_t values for each weight class as a new row in table 2. You could then discuss the consistency of the so derived thresholds in addition to fig. 6ff

Thanks for the suggestion. However, we think that adding threshold values for Q at this point of the manuscript could be confusing for the reader: in the methods section no reference is made to numerical values for variables derived in the field (Q_t values are introduced only in the section 4.4 of the manuscript). Additionally, such values are class-independent (at LS, Q_t is 0.38 for all weight classes; the same happens at US, with Q_t = 0.32). These values now appear at p432 l15-18.

• p426 l28ff: If G = d*w*v(1-p), the dimension is L*L*L/T = L^3/T, so the unit must be m^3/Time, not m^3, and G must be termed a bedload transport rate, not a bedload transport volume (also to avoid duplication of volume/volumetric). Moreover, it is not clear to me if G refers to a weight class and should hence read G_i, or if it refers to the total quantity of sediment. This is important because it is needed to reproduce the estimation of total sediment flux rates. If the latter is true (i.e. G refers to total sediment flux), I'd suggest that you change the notation of v, because it would then refer to the mean virtual velocity computed across all grain sizes; should this mean be weighted with the proportion of the respective weight classes ? Can you give a reference for the assumption that p=0.3 ?

Thanks to the reviewer. Indeed, the notation used in the manuscript assumed an annual computation (T=1 yr). From each G_i, class-specific transport values were derived, and finally the total volume of mobilized sediment was derived. Hence, L28ff will be modified as follows:

"Following Liébault and Laronne (2008) among others, the volumetric bedload transport rate for each weight class (G_i [m^3s^{-1}]) was assessed as:

$$G_i = d_s w_s \overline{v_i}(1-p) \quad \textit{(i=1,…, 8)} \qquad (2)$$

where w_s [m] is the mean width of the active channel bed, d_s [m] the depth of the active layer of sediments, $\bar{v}_i$ the mean virtual velocity of the i-th-class transported clasts, and p the fractional porosity of channel sediment (here assumed to be equal to 0.3, a typical value following Bunte and Abt, 2001). The volumetric transport rate of each class was then multiplied by the integral transport time (i.e., total time for which the threshold discharge value was reached or trespassed). The total inter-survey volumetric transport rate was then derived as

$$G = \sum_{i=1}^{8} G_i \qquad \textbf{(3)}$$

• p427 l11: Pls report a last a short summary of depth of burial to make your assumption more reproducible. The cited literature suggests it is reasonable, though.

Digging tests (n = 17) performed in US returned burial depths comprised between 1 and 14 cm (median depth = 9 cm), therefore within the D_{50} of the surficial grain size distribution (Houbrecht et al., 2012). In LS, where no digging was conducted, we did not observed any major morphological unit rearrangement, even after the most important bedload events, so that the values for the active layer used in the paper appear sensible and in line with those reported by other researchers working on steep mountain streams (excluding large flood events).

The new text in the manuscript reads as follows:
"This approximation, which is based on digging tests conducted on a subset of buried tracer stones in US (n = 17; burial depth range: 1-13 cm; median burial depth = 9 cm), is in agreement …"

Results section

• p429 l1 and anywhere you use "intra-survey" throughout the paper: If a survey is defined as the act of measuring the relative position of the tracer particles, then Qmax during the time between two surveys is relevant. This period should be termed "inter-survey" rather than "intra-survey" in my opinion. In the caption of Fig 7, the term "inter-survey" is used correctly.

Comment accepted, indeed there was a mistake. Intra-survey has been replaced by inter-survey in the whole manuscript.

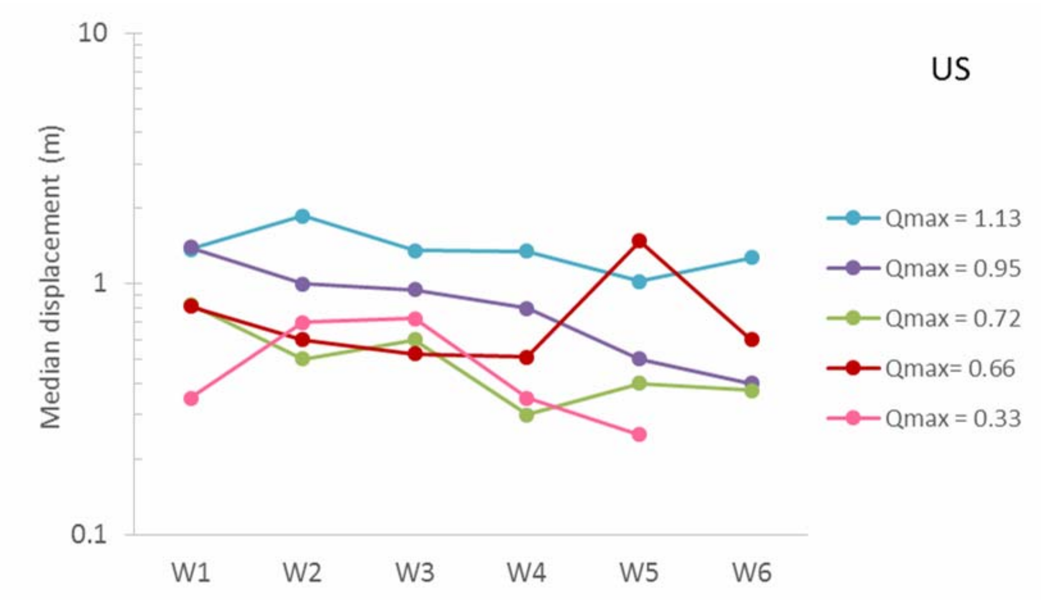
• p430 l4 and Fig. 6: the naming of the small figures is inconsistent as the letters j and k are left out. I'd suggest that you use the letters a-h for the US diagrams in the left columns and the letters i-q for the LS diagrams in the right column. Moreover, 10 surveys can be counted for LS in Fig 4 (however the first two are very close to one another). If there are 10 surveys, then mention that and why one of these was left out. If there are nine, pls correct Fig 4.

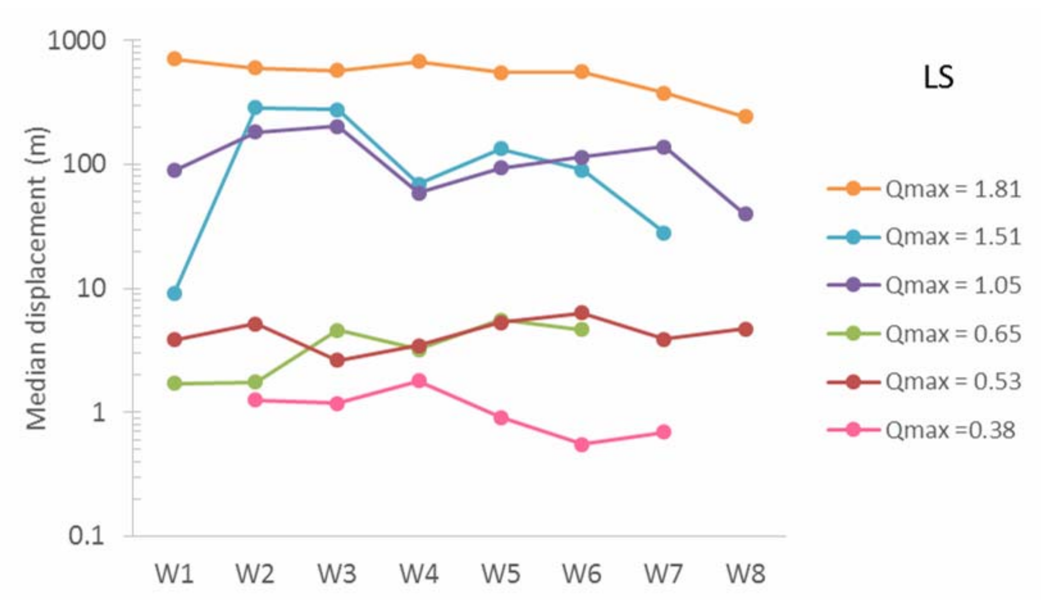
The lettering in the Figure 6 panels have been changed following the reviewer's suggestion. Fig. 4 had a blur that left the impression of two very close surveys, but in fact only one survey was conducted on October 2011. The figure was corrected accordingly.

• p430 l6: which of the plots are meant by "selected plots" ? Please be more specific by explaining your point by example of a specific diagram

We agree with the reviewer, "selected" was not the proper term to use. In any case, the two figures (7 and 7b) have been replaced with the two figures below. New Fig 7a and 8a report median travel distance across weight classes for selected inter-survey periods. In consideration of this new representation, comments to the figures have been changed accordingly.

The new text reads as follows (p430, l6ff):
"For each five intra-survey values of Qmax, weight classes were plotted against the median travel distance during that period (Fig. 7a). These plots suggest that median transport lengths are only weakly influenced by peak discharge (all values within one order of magnitude). Boxplots of tracer travel distance (d) across weight classes (W) (Fig. 7b) show that median values (and maximum values even more) decrease progressively (with the exception of class W5) for heavier particles."





The text at p.432, line 3-11 has been changed as:

"Median travel distances at LS are more sensitive to Qmax (Fig. 8a) compared to what observed at US (Fig. 7a), as they range over three orders of magnitude. In contrast, less pronounced than at LS is the decrease in travel distance (max, median, quartiles) with increasing particle weight (Fig. 7b). Indeed, displacement data at LS indicate that all the clasts within the analyzed size range (representing the "ordinary" bedload flux) move along similar distances for a given flow condition."

• p430 l9: I suppose that the box plot contains data from all tracer particles and all surveys; hence, there must be data with zero movement; I understand that these cannot be shown on a logarithmic y axis. Please make clear (both in the text and in the caption of Fig 7b,8b) if you only evaluated particles that moved, or if you set the tracer travel distance to a minimum (20 cm according to your error estimation) even if the respective particle has not moved at all between two surveys. I think that both possibilities imply potential problems with the visual interpretation of the boxplots that you should discuss briefly.

The boxplots contain data from mobile tracer particles only (hence, observed displacement greater than 20 cm in US and 50 cm in LS). To avoid any possible misunderstanding, the manuscript will be changed as follows:

"Boxplots of tracer travel distance (d) across weight classes (W) (Fig. 7b), in which "no motion" counts are not included, show that median values decrease progressively for heavier particles […]"

Following the same logic, on p432 l6ff, the manuscript will be changed accordingly:

"Boxplots of clast displacement (d) across weight classes (W) during the study period (Fig. 8b; "no motion" counts are not included) show comparable ranges of variability, both in terms of inner (25–75 %) and outer (< 25 and > 75 %) quartiles."

Captions of figures 7b and 8b will be changed as follows:

Figure 7 –Boxplot showing the distribution of tracer travel distances across weight classes in Upper Strimm (US). No motion counts are not included.

Figure 8 – Boxplot showing the distribution of tracer travel distances across weight classes in Lower Strimm (LS). No motion counts are not included.

We think that the exclusion of immobile clasts from our boxplots does not invalidate the analysis. The main objective of this paper is to evaluate and contrast bedload fluxes; for this reason we have placed more emphasis on those tracers that have moved during a given inter-survey period. Quantitative data on immobile tracers are reported in Figure 6 and in supplementary Tables S1 and S2.

• p430 l11: Do you have any idea why W5 forms an exception here ? Is it a random effect ?

Only seven clasts belong to class W5, so that, indeed, random effects tend to cover the real movement patterns. The manuscript will be changed as follows:

"[…]. class W5 being the only exception with the highest median (1.0 m). This value is probably related to the limited number of tracers in this class (only 7).

• p430 l12: Why "interestingly" ? The upper whisker of the boxplots in Figs B should exactly reproduce the highest data points of Figs A, so that the correspondence of Figs A and B is not surprising. And the intepretation that there is a decline in maximum travel distances with increasing weight is not counter-intuitive I think. Probably just delete "interestingly"?

Comment accepted. Interestingly was deleted from the manuscript.

• p430 l18: "no motion recorded in W1", similarly to US (p429 l5). Is it considered a random effect, or could it be that the smallest particles may be caught between larger particles and therefore may be immobile ?

We hypothesize that the main reason for no mobility in W1 class at LS might be due to entrainment and some limited shielding effects induced by the presence of bigger clasts in the riverbed. These effects are much more evident at LS than at US, where coarse calibers are rare.

• p431 l2f: "only 9 clasts (in W1 and W4) moved further than 10 m" - the maximum travel distance in W3 was larger, too (as mentioned in the same line)

Comment accepted, the new sentence reads as follows:

"Maximum travel distance was highest in W3 (21.7 m), lowest in W8 (1.7 m), and 9 clasts belonging to classes W1 and W4 moved further than 10 m."

• I was asking myself if it was really necessary to report in the text with so much detail on the movement of selected W classes between each survey. I think that the sections 4.2 and 4.3 could be shortened to highlight only the most important findings together with the relevant characteristics of Q in the respective intersurvey period.

Shortening the text could be very difficult because it would mean that some information gets lost; nonetheless, we agree with the reviewer that the form our data are presented could benefit from a better organization. Hence, we added a tables (see amended Table 3) where data have been reorganized so that they could work as a summary for the in-text data.

"Table 3 – Recovery rates per survey in US and LS. Rates of the last two surveys are considerably lower than what previously recorded, due to several clasts moving beyond the outlet section of the study area."

Survey (US)	Qmax [m^3s^{-1}]	Mobilized clasts [%]	Mobilized classes	Max travel distance [m]	Intra-survey dominant regime	Recovery rate
2011.09.28	0.36	3.8%	W2, W3, W5	1.2	Rainfall	100%
2012.06.14	0.66	27.7%	W1 to W6	4.5	Snowmelt	100%
2012.07.03	0.55	4.6%	W1, W2	3.4	Snowmelt	100%
2012.09.11	0.33	17.3%	W1 to W5	3.9	Rainfall	98.5%
2012.10.04	0.32	4.3%	W1 to W3	1.6	Rainfall	97.7%
2013.06.27	1.13	50.8%	W1 to W6	35.0	Snowmelt	93.4%
2013.10.01	0.72	38.1%	W1 to W6	12.0	Mixed	92.3%
2014.06.13	0.95	40.6%	W1 to W6	23.9	Snowmelt	90.3%

Survey (LS)	Qmax [m^3s^{-1}]	% mobile	Mobilized classes	Max travel distance [m]	Intra-survey dominant regime	Recovery rate
2011.09.27	0.58	14.1%	W2, W3, and W5 to W8	33.9	Rainfall	100%
2012.05.18	0.53	49.4%	W1 to W8	17.2	Mixed	100%
2012.06.28	1.05	86.5%	W1 to W8	936.8	Snowmelt	100%
2012.08.21	0.65	37.6%	W1 to W8	21.7	Mixed	97.1%
2012.10.25	0.51	36.6%	W1 to W8	21.1	Rainfall	96.5%
2013.05.21	0.38	26.8%	W2 to W7	10.8	Snowmelt	96.5%

2013.07.02	1.81	96.3%	W1 to W8	926.9	Snowmelt	96.5%
2013.10.02	1.01	60.0%	W1 to W8	959.3	Mixed	56.0%
2014.07.16	1.51	64.6%	W1 to W8	958.3	Snowmelt	54.7%

• p432 l 14-19: This paragraph suggests that one single Qt was sought for which all W classes were mobilised. However, the formula refers to each W class. Please consider my comment referring to naming Qt for each W class in Table 2. I think that the reader can also better assess the consistency of the different Qt's (Qt should at least not decrease for increasing particle weight), and the significance of the difference between the two "overall" thresholds for US and LS named here. In l20ff, you refer to the different W classes again.

The formula was used for each weight class to determine independently a Qt value. It turns out that the threshold values are almost identical (0.32 m^3s^{-1} at US, 0.38 m^3s^{-1} at LS). We agree that this is not evident from the way we put it in the original text . Accordingly, the text on p426, l18 and Eq 1 (p426, l19) was changed into:

"The threshold discharge value associated with the motion of each weight class was defined as:

$$Q_{ti} = \min\left\{\max\left(Q_i\right); \min\left(\overline{Q_i}\right)\right\},,$$

Furthermore, lines 14-19 on p432 were changed as follows:

The flow competence method (Eq. 1) allowed identifying entrainment thresholds for the two study sites. At US, we found that discharges > 0.32 m^3s^{-1} were able to mobilize clasts from any weight class (W1 to W6). This figure is slightly lower than what estimated at LS, where entrainment across all classes (W1 to W8) occurred for flows > 0.38 m^3s^{-1}. Given the uncertainty associated with water discharge at LS, we consider the two values not significantly different.

• p432 l25f: while the Fig shows virtual velocities in cm/min, you use cm/s here. Please stick to the same units in the text and in the table. In l25, Qmax < 0.36 m3/s should read "Qmax <= 0.36 m3/s".

Units were changed in the text so that virtual velocities are now expressed in cm/min

On p432, l25: Qmax was changed to Qmax $\leq$ 0.36 m^3s^{-1}.

• p432 l28f: I would write "may include" instead of "includes", because you simply do not know to what extent this is the case, even if you constrain the inter-survey time to the time where $Q > Q_t$, your data shows that particles may rest even if $Q > Q_t$.

Comment accepted. "Includes" was replaced with "may include".

• One of my biggest concerns is that you do not clearly state, in my opinion, where/whether you use Qt for single W classes or for the whole bedload. This needs to be clarified anywhere you refer to bedload rates and transport velocities. As you later refer to the "bedload transport volume", I think that you refer to the total sediment. Then, you have to describe and explain how you come from the different virtual velocities of the single W classes to the overall transport rate.

For comments about the determination of Qt, please refer to the answer given to the reviewer's observations on p432 lines 14-19. The following changes were made to the manuscript to clarify when we refer to the total transport volumes:

- p434, l7: "we estimated bedload transport volumes in US (Table 3) and LS (Table 4)" was changed into "we estimated the total bedload transport volumes in US (Table 3) and LS (Table 4)"
- p434, l9: "Cumulatively, in the monitored portions of US and LS bedload volumes amounts to respectively" was changed into "Cumulatively, in the monitored portions of US and LS total bedload volumes amounts to respectively"

• p433 l17: the progressive increase is only visible for the last three boxes (and strictly, not for the median) in Fig 9B, the first snowmelt-related box in Fig 9B "destroys" this "progressive increase"

Thank you for this comment. We are aware of the high virtual velocities associated with Qmax = 0.38. In the original text on page 433 lines 19-21 we had written: "The highest median virtual velocity is associated with an early secondary snowmelt peak (Qmax = 0.38 m^3s^{-1}), characterized by no rainfall inputs (cf. Fig. 4), occurring before the main snowmelt runoff in spring 2013."

We understand, however, that the sentence on line 17 by itself was misleading. For this reason, we have changed the ordering of the sentences dealing with snowmelt-related virtual velocities. The new paragraph now reads as follow:

"The highest median virtual velocity is associated with an early, short lived secondary snowmelt peak (Qmax = 0.38 m^3 s^{-1}), characterized by no rainfall inputs, occurring before the main snowmelt runoff in spring 2013 (cf. Fig. 4). For longer snowmelt cycles, similarly to what observed in US, virtual velocities tend to increase progressively with peak flow rate".

• p433 l28f: Please specifiy for which event you see this "direct correlation". I spotted a slight positive (please add direction of the assumed correlation) one only for the red triangles in Fig 10A. And it would be interesting to discuss this counterintuitive behaviour (heavier clasts moving faster than small ones). I also think that the smaller velocity of the smallest particles could be due to shielding by larger particles, right ?

Indeed, we meant that the correlation is visible only for rainfall-induced events at US, but we see that this linkage was not clear in the original text. The manuscript will be changed as follows:

"Results show that rainfall-induced peak flows have almost identical effects in the two study sites by imparting virtual velocities clustered between 0.1 and 3 cm s^{-1}. In these two instances, we observe a direct weak correlation between tracer weight and virtual velocity (red triangles in Figure 10). In these instances, we reckon that heavier clasts move faster than the others because once they are in motion, they are subjected to higher water velocities (due to their higher protrusion into the flow) and are more capable than finer particles to pass over locally slow, rough bed patches, thanks to their higher momentum. In this way, they find fewer opportunities to stop once they are entrained."

• p434 l9: Please be consistent with "transport volume" vs. "transport rates"; here, you convert a volume (m3/3 years) to a mean annual mass flux (t/yr), which might be fully correct, but is a bit confusing.

Comment accepted. The manuscript will be reworded as follows:

“Cumulatively, in the monitored portions of US and LS bedload volumes amounts to respectively 3.3 m^3 and 590 m^3 (Figure 11b), which correspond to a transport rate of approximately 2.9 and 511 tonnes per year.”

• p434 l13f: This seems to anticipate the discussion; the rates are not only controlled by the event dynamics, but also by sediment availability and morphology, right? This is discussed in the next chapter. Concerning the term "higher efficiency" in l13: Spontaneously, I would say that "long distance transport in short time" is efficient; here, you use "high efficiency" for snow melt periods for which you have already stated that they span a long time, and virtual velocities are (well, perhaps spuriously) low. Are snow-melt periods perhaps "effective" by contributing most of the annual budget ?

High efficiency was indeed used to stress the impact of snowmelt-driven events on the total amount of transported sediment. In order to avoid any misinterpretation, the word "efficiency" has been replaced with "dominance". The manuscript now reads:

"In agreement with what noted above in terms of virtual velocity clustering for different hydro-meteorological forcing (cf. boxplots in Figure 9a and 9b), this two-order of magnitude difference is mainly controlled by the dominance of snowmelt-driven bedload transport in LS (Figure 11a), and to a lesser extent by rainfall and mixed events. Specifically, snowmelt-related transport accounted for 75% in US and 93% in LS of the total bedload flux."

Discussion chapter

• p434 l27f: "decoupled conditions" refers to within-channel connectivity, right ? Later, you address lateral sediment connectivity; please refer to sediment availability (which is partly a consequence of lateral connectivity) also.

Decoupled conditions refer to the presence of a buffer between the hillsides and the Strimm channel. The term geomorphic coupling was first introduced by Brunsden and Thornes (1979), and refers to the degree of proximity between hillslopes and channels. When hillslopes are delivering material to the channel, they are "coupled", in that their sediment flux contributes to the sediment yield measured at the basin outlet.
In the original text, lines 23-26 (page 434) are meant to set the stage for the subsequent sentences. It follows that "decoupled conditions" refers to the degree of hillslope-channel coupling (or lateral connectivity, Fryirs et al., 2007). In order to address the reviewer's concern the new text on page 434, line 27 now reads as follows: " (…) decoupled conditions (i.e., buffered from lateral hillslope sediment inputs Brunsden and Thornes, 1989; Fryirs et al., 2007) ".

• p435 l27 compare annual rates of sediment, 200 $m^3 year^{-1}$ and 1 $m^3 year^{-1}$, not 1 m^3

Comment accepted. The text has been changed accordingly.

• p436 l24f: I do not understand the conclusion that Strimm Creek is a bedload-dominated system because, remarkably, the single debris flow accounts for 20-25 years of “ordinary” bedload transport. Please explain

The fact that a similar large event was not recorded in the provincial debris-flow database, which includes historical records dating back as far as three centuries ago, suggests that events of similar magnitude have a recurrence interval of the order of 100 years. Following this logic, we think that bedload transport in Strimm Creek dominates over debris-flow transport.

Conclusions

• p437 l19ff: I think that the conclusion reported here has not been developed in the results and/or discussion chapter. In order to appear in the conclusions, it should be in the discussion chapter.

Comment addressed. Implications for long-term landscape evolution have been added to the discussion on page 436, line 26:

“The implications for long-term landscape evolution in Strimm Creek are intriguing as we are looking at two portions of the same basin that are functioning at markedly different speeds, and therefore characterized by a different geomorphic sensitivity to change (Brunsden and Thornes, 1979). Upper Strimm Creek is a typically “slow response” sub-system, in which, due to generalized hillslope-channel disconnection and to inefficient bedload transport, any perturbation through the hillslopes and the drainage network is transmitted relatively slow. By contrast, Lower Strimm Creek represents a highly sensitive, fast responding system thanks to the high degree of hillslope-to-channel sediment delivery, together with much higher bedload transport. In consideration of the limited bedload transport recorded at US, which implies a limited supply of clasts to the rocky valley step separating US from LS, we postulate that the valley step will undergo extremely slow bedrock incision, due to the limited tool effect acting on the sub-vertical rocky channel (e.g., Finnegan et al., 2007; Turowski et al, 2007). This hypothesis is supported by the little bedrock incision achieved by Strimm Creek in the last 12,000 years since deglaciation.”

Figures and captions

• Fig 1: Add legend symbol for lakes; is it correct that the ephemeral section of the US creek is discontinuous – if so, what is the reason ?

Comment addressed. Reference to the lakes has been made in the figure caption. In Fig 1 now appears also the outline of the ummocky moraine ridge which causes water loss to subsurface in US i.e., the channel disappears completely for some tens of meters.

• Fig 7: Weight classes W7 and W8 are not part of 7a - not represented in US area? If so, please mention in caption and text. The X axis is not to scale, i.e. you use a seemingly metric axis for a categorial variable (in turn, the lines connecting the dots are misleading to some degree). Either use a metric axis and the mean of each category, or remove the lines to make clear(er) that a categorial variable is displayed on the x axis.

Only classes from W1 to W6 are represented in the US area. To address the reviewer's concern the text (p426, l6ff) will be modified as follows:

"In our analysis, the tagged particles were grouped into eight weight categories: W1 to W6 in US, and W1 to W8 in LS (Table 2)."

And the caption for the figure 7 will be modified as follows:

Figure 7 – (a) Tracer mean displacement across weight classes and stratified by inter-survey peak discharge (Qmax) in Upper Strimm (US), where only classes W1 to W6 apply.

• Fig 8A: The data for Qmax=0.41 m3/s displays a non-consistent relationship, i.e. the maximum displacement is very low for the smallest class and appears to slowly decline for the heavier ones. This could/should be addressed in the text. Do you think that is due to random effects, or is it a systematic observation because the smallest particles are shielded by larger ones, at least for comparatively low Q ? - similar in Fig 7A

See reply to previous comment (i.e., page 433, line 28 and following).

3 *Technical Corrections*

• p422 l4 write "Ötztal" or "Oetztal" instead of "Otztal"

• p422 l28 duplication of "colluvial", I'd suggest you write "convey colluvial material from hillslope tributaries"

• p423 l2 insert comma between "Here" and "open-slope"

• p425 l16 consider re-wording "foregoing" to "aforementioned"

All the above technical corrections were implemented in the manuscript