Reply on the comments of the two referees on the manuscript "Bedload transport controls intra-event bedrock erosion" by A. R. Beer and J. M. Turowski (esurf-2014-44)

Reply by A. R. Beer and J. M. Turowski

We cordially want to thank both referees for their effort on reviewing our manuscript and their helpful comments! In the following we will comment on all their remarks and give explanations, corrections, discussions and changes we will apply. For better readability we cite the words of the referees in normal text style and add our replies in *indented, in italics and another font* respectively.

Reply to comments by referee 1 (anonymus):

… , it is sometimes hard to follow between the many model classes, model names, studied periods, exponent values..., and the main conclusions are not always easy to identify.

> *We will clarify our wording and add another figure with an overview of the event analysed, as well as a table with associated numbers. We will further rename the "long period" in "event period" and add the model equations to Table 1 (see below) to better guide the reader and simplify our explanations.*

In section 3, the variables that are directly measured on site, or extrapolated from various assumptions, should be more clearly listed (a table would be appreciated).

> As mentioned above, w*e will provide an additional figure and an additional table with these variables.*

I also regret that the actual value of the erosion rate (that is, the parameter K) is not discussed throughout the paper. Though the relative influence of the 4 different parameters is certainly of importance, an accurate prediction of E would also be strong criterion to discriminate between models (in particular, the time-integrated prediction of models with no tools effect could differ more significantly from measurements).

> *The aim of this manuscript is to analyse the transient performance and adequacy of bedrock erosion models with the specific focus on the consideration of the four proposed sediment effects. Hence, we did not attend to give absolute erosion values (whose stepwise measurements are given in Beer et al., 2015), but review model behaviour on scaled values of discharge, bedload transport and erosion. For many models there are no real standard values for K as they always need to be calibrated for specific sites; hence, there is no use of doing this here for our purposes in this article. Yet, we intend to analyse them in comparison to the published values (as e.g. given in Sklar and Dietrich, 2006), but want to do this based on more erosion events to attain more robust mean values, which goes beyond this paper.*

p55 l26: "the suspension effect term regulates the fraction of particles in suspension": what does "regulates" mean here? Is S_e equal to this fraction, or how does it depend on it?

> *Pardon, this is confusing indeed. We mean "… gives the fraction of particles in suspension" and will correct the sentence this way.*

For sake of self-completeness of the article, it would be easier for the reader if a few more words were added about the studied stream (typical width, slope, pebble size...).

> *We will give these values in the additional table mentioned above.*

p57 l18: the sentence "can be determined to 1kg" is not clear to me.

This part will be changed to "can be determined to an accuracy of 1 kg ± 30 %".

p57 l24: it does not seem relevant to give 3 significant figures for the peak discharge if the expected uncertainty is 15%.

We`ll change it to "1.1 m^3/s".

p58 l18: since H_y is said to be the "hydraulic parameter", it sounds weird to see it listed amongst the "sediment effects", even if it can include motion threshold.

We will change the introduction of H_y (p.55 l24) from "Hy is a placeholder for an effective hydraulic parameter (e.g., discharge, stream power, bed shear stress) incorporating the grain motion threshold." to "H_y is a placeholder for an effective hydraulic parameter (e.g., discharge, stream power, bed shear stress), where effective means it can incorporate a threshold for grain motion. Thence, Hy represents the sediment motion effect in these cases.".

p59 l1: "as representative of other members..." would be lighter.

Thank you, we will change that.

p59 l16-17: the units of the bedload transport rate and discharge are incorrect. Shouldn't the unit stream power be expressed directly in W.m^{-2}?

Yes, it is W/m^2; we will change that.

l.18: Shields take a capital S. The Shields number should be quickly defined.

We will shortly define the critical Shields stress.

p 59 l20 to 22: I do not understand that l20 the authors "focus on continuous bedload" but shortly after that they "focus on transient behaviour".

We will change the term "focus" in l20 to "restrict". We want to say that we restrict the time of bedload transport to the period with measured transport rates in excess of 1kg/minute and will focus our analysis on the transient behaviour of the erosion model predictions based on these rates. Hence, there is no conflict in our explanations.

In figure 1, is the grey region "bedload transport" an observation or the outcome of a model?

This is the measurement of bedload transport based on the geophone sensor as explained in chapter 2 (observation site an data).

p60 l16: "wavy" pattern does not sound very precise.

At the point of introduction this notion is only an observation from the graph and is not meant to be precise. The description is only used to highlight the different between the two groups of models.

p60 l18-22: this paragraph about the threshold of motion is not very clear.

We will change the paragraph from: "Threshold of motion: the timing of bedload transport is of distinct importance, especially if the tools effect is ignored in addition to the threshold, since models neglecting both effects (USP and LD) predict erosion even when none was detected (cf. Sklar and Dietrich, 2006). This is particularly obvious for the end of the observation period (Fig. 1a and c)." to "Threshold of motion: the period of bedload transport is of distinct importance, especially if the tools effect is ignored. Models neglecting both, motion and tools effects (USP and LD,) predict erosion even when none was detected (cf. Sklar and Dietrich, 2006). This is particularly obvious for the end of the observation period (Fig. 1a and c).".

p60 l26: how is the value of S_e obtained?

We will add a note here: "(for details on calculation see APPENDIX A)".

p61 l 4-5: the sentence "there was negligible time of full bed cover..." does not sound clear to me.

We will change the sentence "the cover term Fe averaged at 0.91 +/- 0.13, and there was negligible time of full bed cover calculated during which erosion was completely prohibited" to "the fracture of exposure Fe averaged at 0.91 +/- 0.13. During negligible time full bed cover was calculated where erosion was completely prohibited". Here we already take note of the comment of the second reviewer that the notion "Fe" stands for the "fracture of exposure" and change our wording.

p63 l28: "the ratio of active over critical shear stress": I think that this is the parameter that is commonly defined as the transport stage. It would be clearer to define it explicitly by an equation.

We will add the notion "(the transport stage; cf. Sklar and Dietrich, 2004)".

p64 l10-15: it seems normal that the cover effect has no influence in the measurement over a short event and a localized spot. Since the erosion rate is normalized to 1, only the influence of the time variations of F_e (and not the value of F_e itself) can be assessed. Since F_e appears to vary over a rather narrow range during the event, the extracted value for the exponent c is certainly dubious. Therefore it seems too bold to conclude that there is "an absence of the cover effect". This could be discussed more extensively.

We see no reason why the cover effect basically should not appear during a short event and at a single point, since commonly there are huge short-term fluctuations of sediment transport observable (cf. Beer et al., 2015 fig.4 for the event on hand). The conclusion of an absence of the cover effect is based on (i) the fact that we calculated high values of Fe (> 0.9) during most of the time and (ii) on visible observations of no sediment residuals on site over years. As a consequence, the changes in the exponent c for model optimization purposes did not have a dominant influence on model performance. Hence, we do not think that we should add more discussion on this topic.

p64 l20: according to table 2, in the TO model optimized for "the erosion period" the erosion rate scales as Qs^{1.9} rather than Qs^1. This nonlinearity is of potential interest, since it would increase the influence of large hydraulic events on bedrock incision; this could be discussed, even if the results presented here are for the "long period".

Thank you for your observation! We attribute this discrepancy to the other periods to be an artefact of scaling in the erosion period. See the already existing notions on this in the section on "model sensitivity on time period". We will gladly consider your suggestion and add discussion on this nonlinearity in this section, too.

In the list of all model classes, it would be clearer to state for each of them the common form of equation (1).

We will provide a further column with this information for every model class in table 1.

p68 l7 not all four terms are between brackets

We will give the tools effect Qs in brackets too for easier readability and comparability to Equation 1.

p68 l18 "...were based on the following reasons..."

Yes, we will change "was" to "were".

Reply to comments by referee 2 (Phairot Chatanantavet):

The method section is short and may need more detail and clarity.

> *We will review it for more clarity and add more information on the models, usage of data and analysis by referring to the new figure and table already mentioned above.*

… the conclusion section needs to be polished, I think, and maybe added the points I make below.

> *We will revise the conclusion section based on our comments below.*

The idea of the suspension term (Se) in your equation 1 was proposed by Sklar and Dietrich 2004 without any empirical data to support the term (first arisen in their saltation length formula), and this term incorrectly affect the saltation length for bigger grains. They just came up with it to support their idea that all grains (cobble or sand) should be suspended when shear velocity exceeds particle fall velocity, which is not necessarily true (from my flume observation as detailed in Chatanantavet et al. 2013 JGR-ES). And I have calculated a few times before also. For example, in Johnson and Whipple 2007 data. If you use the hydraulic data with high transport stage in J-W 2007, and use the formula in Sklar-D 2004 to calculate the step length Ls, you came up with Ls about 2 meters in the 2.5 m long flume which is unrealistic. Later, they seemed to realize this and completely dropped the suspension term in Lamb, Sklar, and Dietrich 2008 (e.g. the Lamb et al's model). I am glad to see that your results in P61L1-3 supports my claim. I think it would be great if you look into details of [the suspension term (Se)] in your discussion. This terms does not seem right nor needed, as you can see your results in table 2 and figure 2. The SA model needs to drop this term once and for all.

> *Thank you very much for your suggestion and explanations! We will add a short discussion regarding the outcomes and discussion of Hodge et al., 2011 (JGR) and Chatanantavet et al., 2013 (JGR).*

When you state/imply that the stream power type models work as well as the SA model by Sklar-D 2004 (e.g., P67L7), beware that 1) the exponents in the former have been said to be site specific or process specific, but in the SA model the exponents are fixed from physics based analysis, and 2) importantly in your study you ignore erosivity k, which is a black box in the stream power models but in Sklar-D model, it is quantifiable and derived analytically from sensible parameters (rock properties etc). Basically what you are testing here is only the exponents of each sediment effect term in your specific site of tools dominated channel. So, beware not to let this aspect escape in your summary of the results.

> *You are right with the differences in 1) usage of the model exponents and 2) determination of the erosivity factor K for the stream power models in contrast to the saltation abrasion model. We will gladly discuss this. However, we consider all models as implementations of the generic erosion model form given in Eq. 1 and therefore analyse their performance despite potential mechanistical parametrizations. Also, both facts do not have an influence on the overall patterns observed: 1) our sensitivity analysis (changing the exponents) did not improve performance of the models relative to each other and 2) the erosivity is a fixed/constant number that has no impact on our evaluation that uses scaled values of erosion, exactly to get rid of this influence.*

I think since the title is short, it should be noun instead of a sentence, i.e. "Bedload transport controls in intra-event bedrock erosion" may be better.

Thank you; we will potentially adopt this change, since it will better set the focus on the relative meaning of the different sediment effects.

P 55, L26; throughout the paper, Fe stands for "fraction of exposure" in previous works, as opposed to Fc "fraction of cover". I understand that you mean "cover effect" but beware and try not to confuse the readers. For example, in Figure 1c, you wrote "cover term"; people may misunderstand it as a coverage fraction, but in fact it is Fe, fraction of exposure where 1 means fully exposed bed. Pls check throughout the paper. You may say in p55 L26 that Fe is defined as fraction of bedrock exposure, describing the cover effect.

Yes, this is confusing, thank you for pointing this out! We will add an explanation to P55L26: "F_e is the fraction of exposure, which is based on the cover effect". In Figure 1c we want to show the transient course of the four terms in equation 1 in terms of how they contribute to the erosion. Hence, we will give the "exposure" here and will change the notations of Fe from "cover effect" to "fraction of exposure" in the text.

Section 2; I think you may want to quickly state some basic information here about the Erlenbach Creek without demanding the readers to look at the previous papers, e.g., the slope, the channel width, D50, D90 (but I saw later that you listed D50 in section 3). A range of flow depth during the study period, if possible, should be listed somewhere.

As mentioned in the reply to the first referee, we will add information on this in an additional table.

P58L10; "before" is a better choice of word than "until"

We will change that.

P58L10; unknown; hence, only the ….. is a correct grammatical use of "hence".

Thanks, we`ll change that as well.

P59 L 17; discharge but a unit of velocity. Did you mean m3/s? Check somewhere else in the paper too.

Thanks, we`ll change that as well.

P59L18; Shields (with capital S)

We will change that.

P59L19; Figure 1 therein …. Use the word "therein" to distinguish from your own figure 1.

Thanks for the suggestion.

P59L21; were scaled to unity …. is a better choice of word.

Thanks for the suggestion.

P59L18; the critical Shields stress here is unusually high especially on smooth bedrock surface. Lamb's formula was derived from alluvial channels for the most part. Pls discuss or recheck this why it was so high. Did you have to estimate the associated flow depth? Was it reliable data?

Correct, the critical Shields value is too high. The reason is that we used the calculated shear stress when transport started on site of the erosion scales in the chute channel that is unnaturally smooth compared to its high slope (thence, here Lamb`s formula is not applicable). However, actual detachment of the grains that are transported on site happened in the natural channel bed section upstream of the upper stream gage (see Beer et al., 2015 Figure 2c). Since the bed slope there is lower, critical Shields stress was lower too and grains removed upstream didn`t stop downstream on site with higher slopes.

Therefore, we estimate the critical Shields stress on site of the measurements applying an approach by Rickenmann et al., 2006 (Proceedings of the River Flow Conference, 2006: eq. 18) that was developed for bed slopes < 20 %. Using this, we get a Shields value of 0.006, which is in range with other values for steep and smooth beds (0.007;

Chatanantavet et al., 2013 table 2, JGR for bed slopes < 9%), but lower than common values for alluvial beds (lowest values here are 0.03; cf., Buffington and Montgomery, 1997, Figure 3b, WRR). This will change the effective values of shear stress and stream power used in the modelling, but not the transient patterns of the model evaluation that is based on scaled values. Further the difference on USP and EUSP will disappear, since during the whole event considered the critical discharge was exceeded. However, there will be no fundamental change in the overall results, discussion and conclusions.
For flow depth we use a regression based on measurements of a rotating laser distance sensor (cf. Beer et al., 2015).

P59L24; Pls add a sentence justifying why you need to show three separate simulation time periods.

We will change the sentence "Model sensitivity on bedload transport was assessed using three separate simulation time periods." to "Only part of the models incorporate bedload transport and we only have reliable erosion data sometime after its onset. Hence, we evaluated model performance sensitivity on actual bedload transport using three separate simulation time periods".

P61L4; again, see above. When you say "the cover term Fe averaged at 0.91". It sounds like the alluvial cover fraction is at 0.91; in fact, it is the bedrock exposure is at 0.91. Rephrase this.

Yes correct, this is misleading! As mentioned above, we will change this notion to "fraction of exposure" throughout the paper.

P61L1-3; this is great. It proves again that the suspension term proposed by Sklar-Dietrich 2004 was incorrect. It had no empirical data to support it. See main comments above.

Actually, our conclusion was only drawn from analysis of a single erosion event at a site with specific conditions. It needs to be further proved for other conditions if this notion is general.

P61L19; It sounds odd to optimize the exponents for the SA model, which is from physics based analysis. The exponents should be fixed. Maybe mention that as well.

Thanks for the suggestion. We will add here: "This can be related to the fact that the exponents of the SA model are based on physical process analysis and therefore should remain fixed.".

P63L25; explicitly state "exponent" 3 instead of 1

Here we actually mean the threshold number at the beginning of the suspension term (see inside the squared brackets in equation A1), not the exponent. We will change the sentence "Applying a suspension threshold of 3 instead of 1 in the suspension effect term Se (see Appendix A) did not change the predictions of the SA model." to "Applying a suspension threshold of 3 instead of 1 in the suspension effect term Se (see the first number in the squared brackets in Eq. A1) did not change the predictions of the SA model." for easier readability.

P66L26; the threshold of suspension may be only suitable for sand size (e.g. 2 mm) with high flood flow, but Lamb et al model took care of that by proposing the inadequacy of Sklar-D model (downplayed step length and thus the suspension term) but promoting "turbulence intensity" as a major role in bedrock abrasion by suspended load. Maybe you can discuss this a bit.

*Transport in suspension is not restricted to a specific grain size but related to the acting stresses or the shear velocity. Hence, also cobbles could be transported in suspension, if the streamflow is fast enough. For the event on hand we calculated Rouse numbers P between 3 and 13 (particle fall velocity / (Karman constant * shear velocity); Rouse, 1937), which indicate that the grains of mean size here (0.02m) were very likely transported as bedload, referring to the common assumption of bedload transport with*

P > 2.5. Scheingross et al., 2014 plead for potential higher rates by suspended erosion as by bedload erosion that could be predicted by the total erosion model (Lamb et al., 2008) during high flood events. However, these higher rates during suspension are for a fixed grain size with varying slope. With higher slope also larger grain sizes are likely to be transported as bedload (if existing) which in these cases may outpace suspended load abrasion again. The whole topic on this is still not clear and we will add more discussion like this to our text.

P67; cover effect is more pronounced in a channel scale. So, when you test this term in your process model with concrete slab in predominantly smooth bedrock reach, it does not show much besides some patch of gravel collide sometimes resulting in slightly non fully exposed bedrock in your fig 1c. I am not sure though if this is equivalent to alluvial cover traditionally studied in many previous works, which are more in a channel scale. In your abstract, it maybe even makes more sense to state that your study was under an artificial bedrock reach and there was no permanent static cover. Your cover was more like dynamic alluvial cover. (I see later that you said something like this in P64L10-14. Pls elaborate more.)

Also a temporal dynamic bed cover (cf. explanation in Turowski et al., 2008, Geomorphology) would be visible in Figure 1c and inhibit erosion on site, despite it would have not appeared on the channel scale. So, both dynamic and static cover would locally prevent erosion in a natural setting. But yes, our study site is clearly detachment-limited due to the smooth and steep channel (cf. discussion on the threshold of motion above), which will prohibit dynamic cover. We will point his out more carefully.

P68L17; bedload dependent model

Thanks, we`ll correct this.

P68L18; parameterization is misspelled

Thanks, we`ll correct this.

P68L26; “in lieu of” is a better choice of word than “actually”

We will change the sentence “The single exponent a to scale stream power (actually discharge) is mainly set to 0.5 …” to “The single exponent a to scale stream power (which is proportional to discharge) is mainly set to 0.5 …”, to for better readability.

P69L10; initiation is misspelled

Thanks, we`ll correct this.

Table2; it is odd to see that you optimized the exponents for the SA model which is physics based as stated previously. You thus obtained unrealistic exponents of 20 and 13. The first one of 20 was because the suspension term is problematic as mentioned above. Also, the cover exponent of -1.4 for LD and AB is odd. Maybe this is because the Fe term was intent for channel-scale or static cover, but you are applying it for dynamic cover here in a process scale.

We applied the sensitivity analysis on the model exponents for all models due to the reason already given above (we consider all models as based on a generic formula; cf. Sklar and Dietrich, 2004, WRR). Actually there would be no difference if dynamic or static bed cover would emerge, erosion would be damped. The formulation of the fraction of exposure term (which is related to the cover term) is not based on a mechanistic argument nor data, but is an assumption (linear connection between 0 and 1). There are other ideas (exponential model based on probabilities; Turowski et al., 2007, JGR) and studies on the consequences of channel bedrock coverage (Hodge et al., 2011, JGR), but this goes beyond our focus here. We will add some words like this to enhance the discussion.

Figure 2 and 3, y axis, “deviation from” not “deviation to”

Thanks, we`ll correct this.

tau*/tau*c is commonly defined as “transport stage”.

We first thought to skip to mention this additional term, but it might be better to cite it, since it is well-known. Hence, we will add a notion on that.