We thank the reviewer for the thoughtful and constructive comments.

General comments

Given the focus on rainfall variability in the introduction text, I expected a paper that would advance our knowledge on the impact of rainfall variability on long term incision rates. Essentially what I read was a paper that concludes that lithological strength variability is very important in correctly predicting erosion rates and that accounting for rainfall variability also helps some (results in table 5, especially).

The reviewer her/his main concern is on the role of lithology versus rainfall variability in controlling erosion rates. The reviewer concludes that lithology dominantly controls erosion and that rainfall helps some in explaining spatial patterns of erosion. In fact, that is indeed one way of looking at the problem: The Area-Based Stream Power Model (A-SPM) does a good job in predicting spatial pattern of erosion rates after correcting for lithological heterogeneity. However, the goal of this paper is not only to come up with just *a* model that describes the spatial variation in erosion rates. What we aim to do, is to explain the spatial pattern of erosion rates and to identify the factors controlling it. Therefore, we do not propose to use process-based erosion models as tools to 'better' predict incision, rather we use them as tools to get additional insights in the existence of a non-linear relationship between CRN-derived erosion rates and river steepness (k_{sn}). That said, we agree with the reviewer that we were not entirely clear in passing on that message to the reader, which is why we will rewrite those parts of the manuscript where we justify the use of the different incision models. In the abstract, for example, we will add the following lines to clarify this:

" … By combining a state-of-the-art hydrological dataset with a stochastic-threshold stream power model, we show that the observed non-linear relationship between river steepness and erosion rates can be attributed to the existence of an orographic runoff gradient in combination with incision thresholds. The role of temporal runoff variability, represented by river discharge magnitude frequency distributions, could not be identified. Our findings have two main implications for the overall interpretation of CRN-derived erosion rates and the use of river incision models in general: (i) applying process-based river incision models for predicting erosion rates at the landscape scale is only relevant when accounting for the confounding role of environmental factors such as lithology and (ii) orographic runoff gradients and incision thresholds explain the observed non-linearity between river steepness and CRN-derived erosion rates. We present a methodology to study spatial variability in incision rates which can be used as a framework to study the coupling between river incision, lithological heterogeneity and climate at larger continental to global scales. "

We will also add the following paragraph to clarify the objectives of the paper:

"The main objectives of this paper are twofold: we want (i) to assess the impact of lithological heterogeneity on river incision and (ii) to unravel the role of allogenic (spatial and/or temporal runoff variability) versus autogenic (incision thresholds) controls on river incision. We develop and evaluate our approach in the southern Ecuadorian Andes where detailed lithological and hydrometeorological information is available as well as CRN-derived erosion rates (Vanacker et al., 2007, 2015).

…

In the following sections, we first describe the study area, characterize the lithological configuration by developing a new lithological erodibility index and compile a database to represent runoff variability. Second, we present the methods and assumptions used for calibrating and simulating river incision. In a third section the modelling results are presented: we start by evaluating the impact of lithological heterogeneity on river incision rates using an area-based stream power model (A-SPM). We then evaluate to what extent the variability in erosion rates can be explained by spatial and/or temporal runoff variability and the existence of incision thresholds using process-based incision models (R-SPM and ST-SPM). Note that the goal of using process-based models is not to improve the explanatory power of the A-SPM but rather to get insights in the potential drivers of incision variability which are otherwise lumped. In a final section we discuss our findings, highlight implications of our work and discuss further perspectives. "

I think the introduction needs to be revised somewhat to better reflect the results presented in the paper. The abstract does a better job of communicating the essence of the paper. Generally, the manuscript is very heavy on the methodology and too light on the discussion of the results and why these results matter.

We agree with the reviewer that we can organize our introduction somewhat better. As suggested in the line specific comments, we will also add some additional sentences throughout the manuscript to guide the reader better through the paper and to maintain a good flow in general. Our updated paper will be reorganised using the following section headers:

1. Introduction
    - 1.1. Background
    - 1.2. River incision models
        - 1.2.1. Area-based river incision model
        - 1.2.2. Process-based river incision models
2. Study area
    - 2.1. Geology
        - 2.1.1. Tectonics and geomorphic setting
        - 2.1.1. Lithological strength
    - 2.2. CRN-derived erosion rates

2.3. River morphology

2.4. Runoff variability

2.4.1. Spatial runoff patterns

2.4.2. Frequency magnitude distribution of orographic discharges

3. Methods

3.1. CRN-derived erosion rates to calibrate river incision

3.2. River incision models

3.3. Optimization of model parameters

4. Comparing model results with CRN-derived erosion rates

4.1. Area-based river incision model

4.2. Process-based river incision models

4.2.1. Runoff-based SPM (R-SPM)

4.2.2. Stochastic-Threshold SPM (ST-SPM)

5. Discussion

5.1. Are CRN-derived erosion rates representative for long term river incision processes?

5.1.1. Equilibrium between river incision and hillslope denudation

5.1.2. Integration timescales of ECRN and ksn

5.2. Environmental control on long term river incision rates

5.2.1. Geology

5.2.2. Rainfall

6. Conclusions and Implications for landscape evolution

7. References

We will also expand the result section and remove some of the methodology sections where possible. We will keep the section on the lithology since we think this part is necessary for the paper.

I also think the authors sometimes overreach on the significance of some results.

In the context of the clarified focus of the paper, discussed before, we will frame the results more clearly and mention the limitations explicitly.

It seemed like a long slog through the methodology section with many figures that did not seem terribly relevant OR were uninterpretable (Figures 3,5,7,8,9,11,13). Not all of these need to be relegated to Supplementary Material, but it would be helpful if some of them were and the important figures referenced more prominently in the text.

We will reduce the number of figures to 9, by merging some and moving others to the SI. We will remove the figure on PGA as suggested by the reviewer. In the updated version of the manuscript, we will also provide more details in the subscripts of the figures to make them understandable and readable as stand-alone objects.

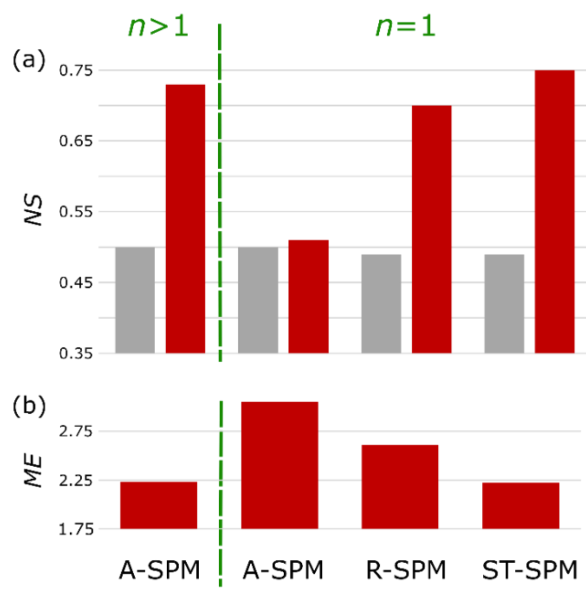
I often felt like I had to hunt down the authors motivation for a methodology or intuit the reasons why results were significant. The authors need to be clearer throughout the manuscript on both of these points.

Noted, we will revise.

With some substantial improvements to this manuscript, particularly in cutting down the methodology section and refining and expanding the results section, I think it can be published as a valuable contribution to the geomorphology community.

We appreciate that the reviewer sees our work as a valuable contribution to the community. As mentioned before, we will try to cut down the methodology section where possible. Two essential parts of this paper – the high-resolution hydrological product and the lithological erodibility index will be kept in the main part of the manuscript, although we will cut down the text and move non-essential methodological aspects to the supplementary materials.

We will streamline the results section by consistently documenting all model parameters in one table (Table 4 rather than table 4 and 5) and will consistently refer to the scenarios as documented in the table. For the sake of clarity, we will present the model fits of all the scenarios in the supplementary information. Moreover, we realized that the discussion section could benefit from an additional graph reporting the overall model performance of the different models and will include this new figure (see below) in the revised document.



Additional figure: Comparison of model performance of four selected river incision models. (a) Nash Sutcliffe model efficiency (*NS*) and **(b)** Model Error (*ME*) associated to four different model scenarios is shown when (i) not considering lithological heterogeneity (grey bars) and (ii) considering lithological heterogeneity (red bars). The Area-Based Stream Power Model (A-SPM) performs well when lithological heterogeneity is considered and all parameters are left free for calibration, resulting in an *n* value of 1.62 (for a full overview of model parameters, see Table 4). When *n* is fixed to the theoretically derived value of 1 (see text), model performance strongly deteriorates when using the A-SPM (see main text). Process-based models, explicitly incorporating runoff variability (R-SPM) and incision thresholds (ST-SPM) perform well when *n* is fixed to 1 and when considering lithological heterogeneity. The highest model performance (highest *NS*, smallest *ME*) is obtained for the ST-SPM where both runoff variability and river incision thresholds are considered.

I have many specific comments on science issues and several technical corrections that are included in an annotated PDF that I will attach.

We will address all the specific comments in the revised version of the manuscript.

Line specific comments

Line 16: ‘enable to assess’: typo

Where is the typo? Sentence rephrased.

Line 18 ‘variability of rock strength and its resistance to incision’: wc

Isolating the role of rainfall variability remains difficult in natural environments, in part because environmental controls on river incision such as lithological heterogeneity are poorly constrained

Line 22 ‘Using 10Be catchment-wide erosion rates, meteorological and hydrological data, as well as data on bedrock erodibility, we provide quantitative constraints on the importance of rainfall variability and lithological variations’: main point of the paper

That is right…

Line 29 wc (word choice, reconsider)
Research on how rainfall variability and tectonic forcing interact to make a landscape evolve over time has long been limited by the lack of techniques that measure erosion rates over sufficiently long timespans

Line 64 several small grammar mistakes: subject verb agreement
Noted, we will revise.

Line 88: spelling
We will write a new ‘objectives’ paragraph (see above)

Line 124 this needs to broken up into completely separate equations or at least labeled 4a, 4b, 4c. Psi is not defined in words and needs to be, as the threshold parameter.
Noted, we will revise.

Line 127 a little confusing here as I was looking for the second component in equation 4. starting new paragraph should solve the problem
Noted, we will revise.

Line 142. this is a big assumption. What needs to happen to make this true? some additional info from discussion can be moved up here.
Good suggestion, we will add a paragraph with assumptions and revisit them in the discussion.

Line 152 model set 1: trad stream power
That's right. We will name this one A-SPM consistently

Line156 second set of model runs with R-SPM. Above, should describe sets of model runs in the same order as incision models
In the theoretical section we first describe the ST-SPM because the R-SPM is a simplification of the ST-SPM. In the result section however, we believe it makes more sense to first present the R-SPM because it assumes constant *k* values and no thresholds. By presenting R-SPM first and then moving on to ST-SPM we gradually add layers of complexity which we find easier to navigate the reader through the result section.

Line 179: I see values of k (little k, no subscript) in table 2, but I don't see where it comes into the incision models. This needs to explained and clarifed. There are many parameters that are some version of big or little k with subscripts, superscripts, exponents, etc.
Noted, we will revise.

Line 184: I think the paper would flow better if the organization was like this:
1. Introduction
2. all study area section
3. Explanation of River Incision models
4.1 Model runs using river incision models
4.2 Optimization of model parameters.
Good suggestion, see structure of updated manuscript as a response to general remark before.

Line 191 It doesn't seem like all of this information is necessary to arrive at the critical information in the final sentence of the paragraph.
We removed part of this section.

Line 206 possible to include a MAP map of the study area? That would be useful given the focus on characterizing the impacts of rainfall variability. If not feasible, at least give a value for MAP on the western slopes as well.
MAP is represented in Figure 6 of the updated manuscript, we will clarify.

Line 212 At some point in the paper (and maybe it's coming later), I would like to see a summary of the catchment erosion rates from these 30 sub-basins.
The erosion rates are given in Table 2. We will increase the size of the labels in Figure 1 to enhance clarity.

Line 227 Are these important for supporting the results/interpretations of this paper? Seems out of place to mention them here.
We suggest keeping the reference to these plots as they are key to understand the discussion on transient incision pulses (in the discussion section of the paper). We will however, move these lines to the new section 'River morphology' where they will be better in place.

Line 239 it would be helpful to readers to explain at the beginning of section 4.4 that the reason you do all of this is to get this regional kw value.
Noted, we will revise.

Line 244 why is 45 stations with 10 years of data not enough? are they all clustered together?
Indeed, most of them are in the centre of the basin and do not cover the catchments where CRN derived erosion is measured (shown on Figure 1). We will clarify.

Line 248 this is bout 28 km resolution. pretty coarse.
While daily temporal resolution is really fine resolution to drive models that evolve over thousands of years.
That is right, therefore we develop a HR product by downscaling the 0.25° WaterGAP3 data. See also next reply.

Line 252 I don't easily grasp the relevance of this section, especially the second half of the paragraph, starting on line 248. What needs to happen in this paragraph is a more succinct explanation of the data sets used to get a pdf of daily runoff and more importantly, why using these data sets is an improvement on the data from the monitoring networks on the ground.
We will shorten and rephrase this paragraph as:
"To estimate runoff variability for all 30 sub catchments, we use hydrological data derived in the framework of the Earth2Observe Water Resource Reanalysis project (WRR2; Schellekens et al., 2017) available from 1979 to 2014. More specifically, we use the hydrological data calculated with the global water model WaterGAP3 (Water – Global Assessment and Prognosis: Alcamo et al., 2003; Döll et al., 2003) at a spatial resolution of 0.25° and a daily temporal resolution (earth2observe.eu). In the following paragraphs, we explain how we derive (i) a high-resolution runoff map by spatially downscaling this coarse data and (ii) catchment-specific magnitude frequency distributions of discharge (pdf_Q) characterising the temporal variability of runoff."*

Line 255 nice intro and motivation for methodology here. But, before you get into the detailed explanation of methods, refer readers to figure 6 so they get a visualization of where you're going and why you do this.
Thanks.
We will point the readers to figures 5 and 6:
"The procedure consisted of the following steps and is presented in Figures 5 and 6:"

Line 281 This sentence unnecessarily confusing. Use more words to explain. this section needs an introductory sentence to orient readers.
We will resolve by adding the following text:

"Runoff variability is typically casted in terms of spatial runoff variability (section 2.4.1). However, also the temporal pattern of runoff might influence river incision and is typically represented by discharge magnitude frequency distributions. Constraining the shape of these distributions is important, because the number of large storm events determine the frequency by which thresholds for river incision to occur are exceeded (see section 1.2.2 and references therein)."

Line 285 here little k is finally defined. this needs to happen earlier where it is first mentioned.
Noted, we will revise.

Line 294 how important are daily variations in discharge over 9 million years of uplift and erosion?
Good point, we will mention this earlier in the assumption section coming with the river incision models and revisit the issue in the discussion section of the paper.

Line 297 this is all fine and good, thorough work, but the summary/motivation of why you do this needs to be at the beginning of the paragraph. otherwise makes for very heavy reading.
Agreed, see reply above.

Line 300 is this section necessary? Does it really contribute to the main goal of the paper, which I understand to be evaluating the role of rainfall variability on incision rates. this sections feels like overkill. I recommend moving to supplementary materials.
As explained earlier, we do believe this section is critical given the importance of lithological heterogeneity in controlling river incision rates. Therefore, we will keep it in the methodology section.

Or at least the seismicity section can go to supp mat.

Agreed, we kicked it out.

the lithological strength section should actually stay as it's very important later in the paper.
Indeed… Should now also be clear to the readers when they arrive at this point, given the enhanced focus on lithology in the abstract/intro

Line 317 where are these data of measured uniaxial compression strength? OK, I have found it now, but this section is confusing. it needs to be more clear and use more words to explain
We will explicitly mention that the uniaxial compressive strength data can be found in Table S4 to enhance clarity.

Line 319 this part definitely seems irrelevant to the main focus of the paper. There are so many other things already going on, this just feels like a distraction. Unless seismic activity is really playing a huge role, in which case maybe the focus of the paper should be on that.

Agreed, we kicked it out.

Line 327 reference figure 12, not table 4.
We will redo the figure numbers in the revised version of the paper and point the reader to both the table and the figure.

Line 331 more explanation and description of figure 12 would be helpful here before launching into another lengthy description of another methodology.
Agreed, we describe all scenarios now in more detail.

Line 336 what about Bayes factors of 1.06 vs 1400 tells us that the data fit a model with variable erodibility better? Needs more explanation.
See comment before.

Line 346 coming back to an earlier comment from near the beginning of the manuscript, it seems the focus of this paper is equally on the effects of spatial variation in both erodibility and runoff. this is not clear/emphasized in the abstract and introduction.
Agreed, we will resolve this by rewriting the abstract, and objectives of the paper. See comments before.

Line 351 spell out model names for section title.

Noted, we will resolve. To clarify, we will also break up this section in two subsections:

4.2. Process-based river incision models

4.2.1. Runoff-based SPM (R-SPM)

4.2.2. Stochastic-Threshold SPM (ST-SPM)

Line 353 nice explanation of what just happened and what will happen next. manuscript needs more of this in places.

Thanks, and agreed, we will resolve.

Line 360 also good emphasis.

Thanks

Line 362 this under prediction/over prediction trend is not obvious to me. Looks to me like the observational data is scattered somewhat evenly about the modeled data line. If the Nash Sutcliff number tells us that there is under/over prediction, then explain how that happens. Otherwise, I think such a claim is not supported.

We will rewrite this paragraph

Line 369 these two scenarios evaluate role of little k.

True, we will clarify.

Line 372 little k apparently not so important.

Correct, we will stress this as: *"In scenario 5, k is fixed to the average value for all catchments (k = 1.01) whereas in scenario 6, k is set to the catchment specific values as listed in Table 2. Both scenarios (5 and 6) perform well with an NS value equalling 0.71 indicating that temporal runoff variability (k) is not influencing model performance"*

Line 375 here finally runoff variability is evaluated.

No runoff is evaluated before. This will be clearer in the revised result section.

Line 378 this is misleading, as scenario 7 performs equally well!

We will explicitly mention that both scenarios perform equally well.

Line 379 what's the significance of these lower threshold values?

We will explain and frame it with some data from literature.

Line 380 I agree that the model vs. measured erosion rates look better in 14b compared to 14a. But it's an over-reach to say that ST-SPM "correctly" predicts low erosion rates. There's still a good bit of scatter and error in model vs. measured ero rates. Just modulate the word choice a little here.
Noted, we will revise.

Line 386 some of this context would be helpful earlier in the paper, e.g. near line 142 or section 4.2
Noted, we will revise by pointing the readers to this section

Line 401 how does this relate to the data presented? where do you suspect over/under estimation of ero rates from Be10?
This will be discussed in the next paragraph.

Line 406 these ids are hard to see and find. note that they are in teh northern section of field area and also refer readers to figure 2, can see these areas are steeper.
All good suggestions, we will do so.

Line 409 confusing. is the variability in agreement due to differences in drainage area in each catchment or to over/under estimation of CRN erosion rates? This needs to be more precisely worded, as the following sentences make clear.
We will rephrase this sentence.

Line 427 good point that will have been on the mind of many readers through out the paper. maybe acknowledge this timescale mismatch earlier.
Good suggestion, we will do so.

Line 437 a bit much detail at this point.
Agreed, we removed part of this paragraph.

Line 444 I must have completely missed this point. Refer readers back to the relevant model runs/figures.
Noted, we will revise.

Line 447 allowed us
Noted, we will revise.

Line 449 given what you just said, now useful are more advanced methods likely to be? where would they be useful?
We will discuss this by adding a sentence.

Line 456 this seems to be the major, clear finding of this work.
We will revise the text to clarify the main findings.

Line 467 if this is your conclusion, recommend removing earlier discussion of seismicity.
Agreed

Line 475 this is also a good and significant point that could be highlighted more prominently.
We will do so and mention it in the abstract of the paper.

Line 484 enables us
Noted, we will revise.

Line 485 am i missing something? this scenario does NOT include variation in runoff.
Hmm, we do not exactly know what the reviewer means here. But we will rephrase the sentence and the overall paragraph in general.

Line 486 above error/confusion makes it hard to evaluate this very important claim.
Agreed, we will rephrase this sentence and the paragraph in general.

Line 494 OK, good point here.

Line 495 yes, but the numbers are only slightly higher for scenario 6 vs scenario 4. How significant is this seemingly small difference?

We will rewrite this paragraph, as well as the paragraph in the results section dealing with these scenarios. In the results section we will add a sentence explicitly mentioning the similarity between these scenarios:

"Note however that differences in model performance between R-SPM scenario 2 and ST-SPM scenarios 5-8 are existent but not very pronounced. To evaluate the significance of these differences, our analysis should be repeated on larger datasets capturing a wider variability in erosion rates and hydrology."

Line 503 i would say that mainly spatial variation in rock erodibility controls river incision patterns. I have to say, I'm not convinced of that rainfall variability matters a huge amount from the data presented here.

Adding rainfall variability (actually, spatial runoff variability) does not improve explanatory model power in comparison the empirical A-SPM. However, we do not propose to use process-based erosion models as a tool to 'better' predict incision, rather we use them as tools to get additional insights in the existence of a non-linear relationship between CRN-derived erosion rates and river steepness (k_{sn}). Applying the ST-SPM enables us to quantify the role of spatially variable runoff and incision thresholds. This will be better documented and discussed throughout the updated version of the manuscript.

Line 518 medium? anyway, medium is not a great word choice to describe basin size.

Noted, we will revise.

Line 527 But the simplest version of the model, A-SPM does almost as good of a job! R^2=0.73, NS=0.73! You must explain this and justify why variable R actually matters!

See comments before. Paragraph will be rewritten.

Line 535 I think this conclusion is fair that this study shows potential, but more research is still needed for definitive answers about R variability.

We will keep this message in the revised version of the paper.