Dear Referee:

We thank you for the detailed and constructive review on our manuscript. We have considered all the comments and we will edit the manuscript in a revision where the description and explanation are unclear, misleading, or incomplete. First, we address the general comments, and then provide responses to specific comments below.

In our study, short-term erosion rates were estimated from suspended sediment yields, which likely underestimate basin-averaged erosion rates due to the lack of bedload and dissolved load data. Unfortunately, data of the latter two types of sediment transportation are too scarce, and the geographic distributions of these data are too uneven to meta-analyse them between climate zones at the global scale. Therefore, we did not include these data for the analyses. In addition, the environmental controls of bedload and dissolved load are complex and may show different characteristics from the suspended load. For example, the percentage of bedload to the total load tends to be higher in mountain regions and drylands (Dedkov and Mozzherin, 1996; Singer and Dunne, 2004), but the percentage of dissolved load seems to be higher in tropical regions and lower in drylands (Milliman and Farnsworth, 2013). Therefore, a systematic correction of short-term erosion rates is not possible. In order to enhance the reliability of short-term erosion rates, we suggest providing the globally-averaged percentage of bedload and dissolved load in the Methods. Previous studies estimated that the bedload typically accounts for < 10% of the total load (Milliman and Meade, 1983), and the average dissolved load accounts for around 20% globally (Milliman and Farnsworth, 2013). Thus, the short-term erosion rate maybe on average ~30% higher than what we presented in the study. Given that the short-term erosion rates, estimated only by suspended sediment yield, are higher than long-term rates in each climate zone (Fig. 2; except for the Cold K–G zone), the potential underestimation of short-term rates should not alter the conclusion of anthropogenic influences on short-term erosion in this study.

In terms of the quality checking short-term erosion rates, the data compiled from the USGS have been confirmed by the data source and labelled as 'Approved for publication: Processing and review completed'. The data collected from publications cannot be checked systematically and uncertainty ranges will be highly variable due to several reasons (Milliman and Farnsworth, 2013): the variety of measuring techniques over different periods of time, inadequate monitoring period (several rivers with historic records < 5 years, especially the shorter ones), watershed modification (resulted from dam construction or climate change), erroneous transcription of the data. Therefore, the quality of data is hard to assess and overcome, for which we will address the issue and limitation in the text. More generally, meta-analyses such as this one are inherently challenged by variability in the quality and substantiation of the underlying data (we will state this explicitly in the Methods). We have done our best to include data from the most reliable sources available (i.e. USGS, OCTOPUS).

In the reply below, the Referee's comments are in *black italics*, and our responses in blue.

Sincerely,

Shiuan-An Chen, Katerina Michaelides, Michael Singer and David Richards

---

**Intro:**

*What makes this study unique over the previous studies that compared these two different methods? (Besides, maybe, a larger dataset now available?). Why do we need yet another comparison? Is the comparison actually leading anywhere, as both methods have different biases… and the uncertainties associated might be too large to say anything beyond something that is better than a factor of 2 comparison? That alone could result in differences that are beyond the uncertainties.*

While some previous studies have either compared short- and long-term erosion rates between drainage basins or analysed erosion rates in particular regions (e.g. Kirchner et al., 2001; Portenga and Bierman, 2011; Wittmann et al., 2011), we are not aware of any studies comparing compiled short- and long-term erosion rates at the global scale. One of the key findings of our study is that a non-linear relationship exists between **long-term** erosion rates and climate, reflecting the balance between precipitation and vegetation cover (erosion rate peak in the semi-arid regions) which corroborates early theoretical work by Langbein and Schumm (1958) and Walling and Kleo (1979). However, we find that this relationship does not hold for short-term erosion rates as proposed by these former studies. This result was not mentioned by the reviewer but is a key finding and take-home message from the paper which directly addresses the reviewer's question as to what makes this study unique.

**Methods:**

*What does "compiled from published literature" mean for suspended sediments? Was there some initial quality check performed? For the USGS data, 2 criteria were used to confine the data (monitoring time and a basin area threshold). But, were there similar criteria for the other station data? Often, data is published were sediment rating curves are really poor, or monitoring times are really short. Especially in remote terrains, suspended sediment data is very sparse due to inaccessibility (in glacially impacted terrains) or due to infrequent rainfall and low discharge in general (dry regions). Hence, a rigorous data quality control and resulting means to use only the best data is needed first. Otherwise, any comparison can only be qualitative in nature and a quantitative comparison that even includes statistical analysis, as attempted by the authors, is useless. A useful endeavor for making short-term erosion rates better comparable with cosmogenic nuclide denudation rates would be to associate an uncertainty to the former. Perhaps this could be done by MonteCarlo Simulation or so, but without having an uncertainty associated, the comparison remains qualitative. What does a factor of e.g. "1.4 higher" mean? Is this beyond uncertainty? As you may have guessed by now, in my view anything that this < factor of 2 between the two methods is actually a quite acceptable agreement. The problem is that not much more to be drawn if one of the methods does not have an uncertainty….*

Erosion rate, suspended sediment yield, and sediment flux data were collected from published papers, textbooks, and the USGS website for estimating short-term erosion rates. The USGS data have been quality checked. However, there are difficulties in estimating the quality and uncertainty of data from other sources because the data were collected with various methods at different time periods. The two criteria used for USGS data were not used for collecting data from other sources. While the majority of suspended sediment yields from publications were recorded over 5 years, several records were shorter than this period of time. The reason for setting a basin-area threshold for USGS data is to enable comparison to similar basin sizes from which the long-term erosion rates were obtained (i.e. from the OCTOPUS database). The basin size of short-term erosion rates from other sources are generally larger (e.g. rivers included by Milliman and Farnsworth, 2013 include those discharging to the ocean, not the tributaries). We acknowledge the challenges in summarising suspended sediment data over a longer period of time, especially in cases where the rating curve between concentration and discharge shifts (we have published on this issue, Singer and Dunne, 2001). Nevertheless, we have relied on published sediment flux values from highly cited and well-regarded studies (Milliman and Farnsworth, 2013), which contain descriptions of data quality control. Therefore, we feel the conclusions from our meta-analysis of published short-term rates are robust. We will edit the explanation in the Methods.

**Results:**

*Another issue is that once datasets are compared to each other (short- vs. long-term rates), one should use the individual data from each basin/river only, meaning the data should be compared 1:1, i.e. only compare stations where there is actually short-term AND long-term data measured within an acceptable range of distance, or better even measured at the very same station). Only when trends with e.g. climate are analyzed for each short- or long-term dataset individually, the entire dataset might be used.*

We agree that it would be ideal to compare short- and long-term erosion rates at the same location. However, there are minimal data available ($n < 100$) that satisfy this criterion (Table 1 below). We also tried to extract long-term erosion rates within 200 km of all short-term rate points. Similarly, the number of paired data is generally less than 100 for each Aridity Index categories (Table 1 below). The comparison becomes meaningless if the distance threshold is further extended. Therefore, we chose to combine all data within each climate zone, assuming that with a large number of data, the variability within drainage basins will be reduced and some climatic signal on erosion rates may be revealed.

**Table 1: The number of paired data of short- and long-term erosion rates at the same location and within 200 km.**

| Number of paired data | Aridity Index categories | | | | | Sum |
|---|---|---|---|---|---|---|
| | Hyper-arid | Arid | Semi-arid | Dry sub-humid | Humid | |
| Same location | 1 | 5 | 6 | 15 | 52 | 79 |
| < 200 km | 2 | 12 | 76 | 53 | 439 | 582 |

Section 3.4: *This area-grouping makes sense and should have been done prior to the entire analysis. Otherwise, there is always the question of whether any trend observed may be due to the different number of observations within each bin….*

Below (Fig. 1) we present the erosion rates of basins smaller than 2,500 km$^2$. The patterns displayed in this figure are the same as the patterns using ALL data presented in the manuscript. One difference is that the short-term erosion rates are more variable between climate zones compared to the short-term rates using ALL data (main manuscript), which is likely a result from the limited short-term erosion data from smaller basins. Therefore, breaking down the data into even more area-group categories would substantially limit the number of data points within each climate zone.

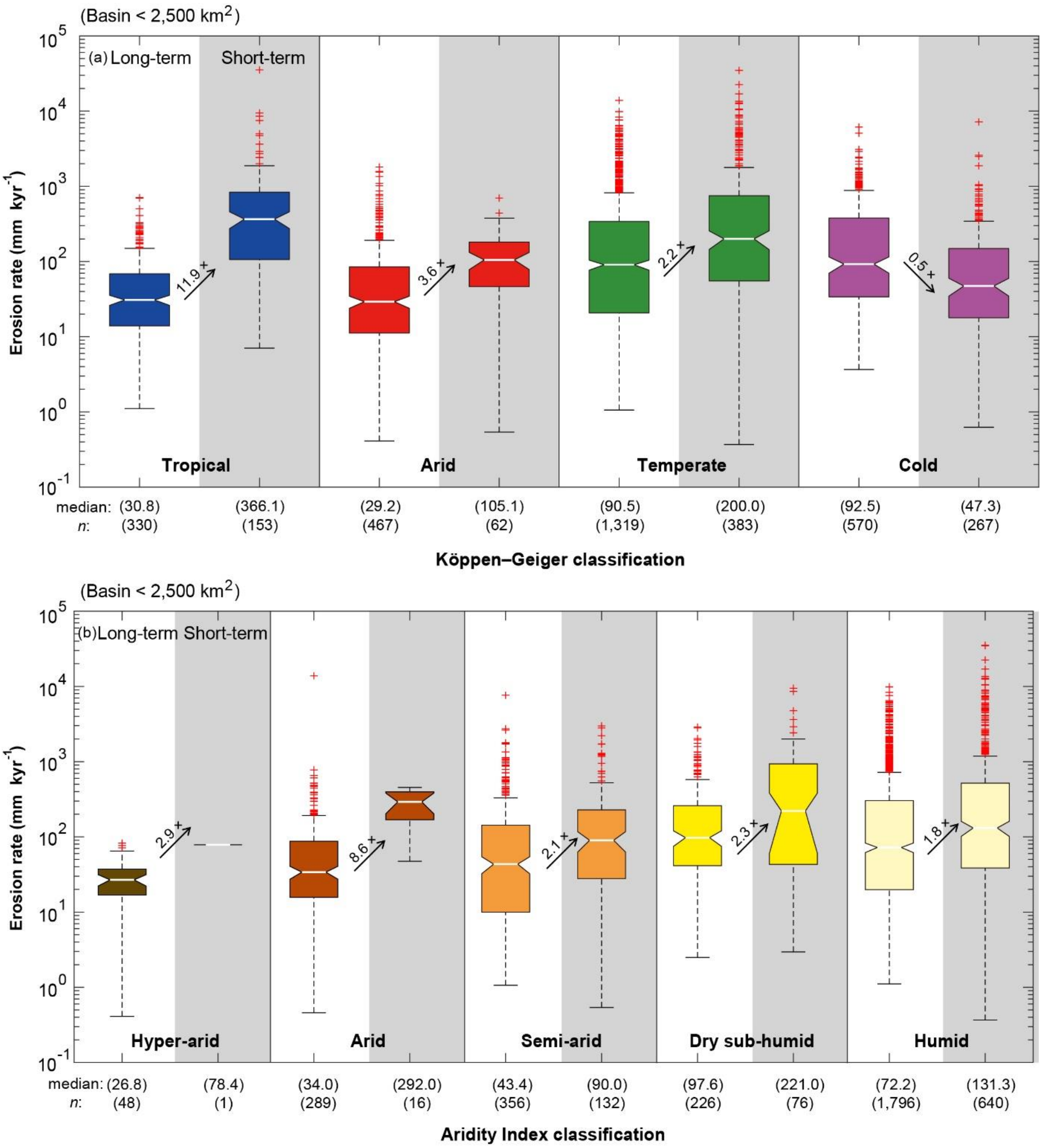


**Figure 1: Long- and short-term erosion rates of basins smaller than 2,500 km$^2$ for climate zones of Köppen–Geiger climate classification (a) and Aridity Index classification (b).**

Fig. 3: *This trend found between the US-derived long-term erosion rates and MAP - is this trend also present in the entire dataset? If not, why is it present only in this dataset and how can then a global general interpretation be drawn if the global dataset does not show the same trend? (In line 376, the usage of 3,074 datapoints is mentioned in this regard. I´m confused, as in Fig 3, only the US data is used…Is the red line in Fig. 9 now using the entire dataset, or only a US-subset?)*

We only analysed the relationship between MAP and erosion rates in the USA because of the widespread systematic availability of MAP data across the USA – and mainly for comparability with other studies which tend to use MAP instead of Aridity Index. Nevertheless, the relationship between global long-term erosion rates and Aridity Index categories (Fig. 2b), show a similar result to the relationship with MAP across the US, with a peak in erosion rate within the Dry sub-humid zone. The only difference between these two results is that the Aridity Index classification cannot resolve erosion rates in the humid regions since all data are lumped into one bin. We will edit the legend of Fig. 9 to emphasise that the line is drawn from the USA data.

Fig. 4: *Glacially impacted denudation rates higher than non-glacially impacted rates: That is nothing new. See reviews by Dixon et al. (2018) and Delunel et al. (2020, ESR) for the European Alps and the study by Ganti et al. 2016 (Sci Adv) that shows that cosmogenic denudation rates are likely affected by a time scale averaging bias. It´s a pity that these studies were not cited.*

Thank you for the suggestion. We will add the literature in the Introduction and Discussion.

Fig. 5: *I don´t think that an increase of 1.4 has any significance without analyzing uncertainties.*

The magnitude of a difference between distributional medians does not indicate whether or not the difference is significant. Significance in this case is tested via the Kruskal–Wallis test, which showed that these two distributions are indeed statistically different (see manuscript). Regarding the issue of uncertainties in the underlying data, given that: 1) the short-term erosion rates in Fig. 5 are all generated with a consistent data collection procedure; and 2) the numbers of data points in agricultural and non-agricultural regions are extensive ($n$ = 826 and 732, respectively).

Fig. 6: *What kind of figure is this? What does the bar legend indicate? Number of observations???*

Figure 6 is a density scatter plot. The bar legend indicates the number of data points in each pixel. We will clarify this in the figure caption.

**Discussion:**

Section 4.1: *A key point for the relation between long-term erosion and MAP is the LOWESS smoothing method. However, there is no reference nor any other further information given how this smoothing works (averaging window?). Given that the resulting shape of the pattern is so much different than that found by* *others, I would encourage these authors to provide more information on it. See also my comments to Fig. 3 that are relevant here.*

The LOWESS smoothing method uses locally weighted linear regression to smooth data, similar to the moving averaged method, which also smooth data locally by neighbouring data points, but the LOWESS uses linear polynomial regression rather than the average value (Cleveland, 1979). We drew the LOWESS regression by the built-in function, smooth, in Matlab, to highlight the pattern of erosion rates (Fig. 3a). We set the polynomial as “linear”, the span as “30% of data points”, and the robust as “off”.

Section 4.2: *What are the actual apparent ages (integration time scales) of the long-term data? Given that denudation rates are typically high (>0.5 mm/yr or so) in glacially- impacted regions, the resulting integration time scale are low (<1200 yrs), and do therefore not integrate over the last 25-15 ka. Same problem for Section 3.2.*

According to Brown et al. (1995), the timescale of erosion rate of ~200 mm $kyr^{-1}$ (in past glaciated regions of the Temperate zone; Fig. 4) is around $10^4$ years, which may or may not have experienced the last ice age directly. However, the enhanced sediment production by former glacial processes can be transported by following fluvial processes during warmer periods, as shown by Ganti et al. (2016) (the erosion rates increase over shorter timescales from millions of years to decades). Since a significant difference in long-term erosion rates between past glaciated and non-glaciated regions is shown in our data, we believe that the erosion rates in mid- and high-latitude regions are, at least, indirectly enhanced by former glacial processes. However, we fully acknowledge that this is a subject of considerable debate in the literature, so we will add a fuller explanation in the Discussion.

Section 4.3: *Same here as for Fig. 5.*

Section 4.4: *Sorry, I don´t get where this leads to. I find the section too general to be useful. Why make such a fuss about an absent relation between erosion and drainage area? Usually people use such an absent relation to show that there data is NOT influenced by sampling location… This section jumps from one topic to another without any clear red thread….The second para is ok for what the first-order observation is… (the fact that the larger the basin, the better the agreement between short- and long-term erosion rates). Last para: An $R^2$ value of 0.24 or 0.29 does not describe a significant relationship.*

We feel it is necessary to explain the absent relationship between erosion rate and drainage area because the existing theory (i.e. stream power incision law) predicts a positive influence of drainage area on erosion rate, whilst some other studies present an inverse relationship (e.g. Milliman and Syvitski, 1992; Milliman and Farnsworth, 2013). Therefore, we proposed potential controls on the unclear relationship to highlight the complexity of drainage basin environments. We will add some clarification to the beginning of this section to explain why we are presenting this information. $R^2$ values do not “describe significance” of a relationship, but rather they quantify the proportion of the variance explained by the relationship. The significance of a regression is measured by a *P*-value, and even relationships with very small $R^2$ values may be significant. The *P*-values of the relationships between long-term erosion rates and the slope gradient and total relief are both lower than 0.01 ($P = 4.23\times10^{-36}$ and $1.5\times10^{-32}$, respectively). We will add these values in the manuscript.

**Reference:**

Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, Journal of the American Statistical Association, 74, 829–836, https://doi.org/10.1080/01621459.1979.10481038, 1979.