

Review for “Global analysis of short- versus long-term drainage basin erosion rates” by Chen et al.

General comments:

This paper by Chen et al. is the latest attempt to gain a better understanding on the controls of climate and tectonics on erosion from the comparison of modern river loads with longer-term cosmogenic nuclide-derived denudation rates. The authors have my respect for attempting this comparison, as it is time-consuming to do. Nevertheless, I find the approach too general to be useful. For example, no attempt has been made to quantify uncertainties for short-term erosion rates, but knowledge of which should be crucial for when comparing two independent methods. Short-term rates are erosion rates from suspended sediments, while rates from cosmogenic nuclides are denudation rates, thus integrating over erosion and weathering. Strictly speaking, the two cannot even be compared. One way out would be to include dissolved river fluxes in short-term rates. While for rates from cosmogenic nuclides the authors rely on a previous compilation which has been carefully quality-checked, a quality check for the compilation of short-term rates is missing (see below).

As such, this study does not provide in my view significant scientific advances over previous studies that have carried out this comparison, and it is missing several substantial characteristics of a solid scientific manuscript (like uncertainty assessment). I am therefore against the publication of this manuscript at it is now.

Specific/extended comments for each section

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.

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 where 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....

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.

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....

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?)

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.

Fig. 5: I don't think that an increase of 1.4 has any significance without analyzing uncertainties.

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

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.

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.

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.