

Dear editors,

Thank you for the valuable and constructive feedback on our revised manuscript. We have submitted a second revision, with substantial changes that reflect your comments and suggestions. In addition to clarifying the text in multiple places, we have conducted two additional lab analyses (pH in CaCl₂ and ⁹Be concentration) and modeled the potential for meteoric ¹⁰Be loss due to pH dependent desorption. We believe the manuscript is much stronger and fully addresses the comments raised by the reviewers.

Comments from the Editor 2nd round (May 2026)

Lines refer to the “authors track changes” version.

I decided to give the authors some more feedback, instead of sending the MS again to the reviewers. This will save them some time, but I would expect to treat my comments as if from a reviewer. The manuscript itself is in my opinion still in parts unstructured and needs clarification. Also, discussion of the results of at least one major issue is pending. I would strongly suggest that the authors follow these instructions, as the MS in its current form not very accessible.

Thank you for your helpful feedback on the manuscript.

Major comments

I think the introduction needs some re-structuring. It's too detailed, and with the sub-heading removed now, reads more like a “methods” section. The balance between the soil-based inventory approach and the other approaches is not even, hence I would move the equation and the detailed description to the methods part. The paragraphs dealing with the assumptions (half-life, 84-90), retentivity & dust (92-96; 97-104) I would also move to the methods, or create a subheading “testing assumptions” and clearly explain your approach. Reading the 3.1. methods / laboratory methods, it's just weird to have no equation at this point. Are Tables S1 and S2 in the data repository? Please provide a link to the ESS-DIVE archive, it's not available.

Based on these comments we have substantially re-structured the Introduction and have moved some of the introductory text to the Methods section.

All supplementary tables are in the Supplemental Information document.

We have provided a link to the ESS-DIVE repository. For reasons we do not understand, the doi alone does not properly connect to the repository, but the link that includes the doi does.

Regarding soil pH, I ask to cross-check the pH values derived here by using ultra-pure water with a method using CaCl₂ (and using a reference material for pH, such as IRMM433-7).

We have conducted pH measurements in CaCl₂ and in the revised text we indicate that the standards we used to calibrate the pH probe are traceable to U.S. National Institute of Standards and Technology-traceable standards. As expected, the CaCl₂ pH values are lower than those measured in water. We report the methods and a summary of the results in the main text, and report all of the data in a supplementary table. Note that we use the pH values measured in water to model desorption losses, as the published empirical equation we used that relates pH and partition coefficients is based on pH values measured in water.

I don't understand why the measured fluxes correlate so well with precipitation and precipitation-related variables, but then the rainfall-relation from Graly et al. (2011) gives much higher fluxes. How can that be? The watershed itself is not that large, any rainfall-dependence should not change drastically. Was there perhaps something lost without accounting for (see pH comment)? I would like to see a more in-depth discussion on this issue, perhaps geared towards additive versus dilution effects?

Since the Graly et al. (2011) model is empirical it is difficult to obtain a deep, mechanistic understanding of the differences between measured and predicted values. Although the fluxes we measured are well-correlated with precipitation, it doesn't necessarily mean that the Graly et al. model should also predict those values well, because the empirical coefficients that scale flux as a function of precipitation very likely differ between the sites Graly et al compiled data from and the East River. To discuss this possibility, we have added the following text:

"The empirical formulation of Graly et al. (2011) predicts fluxes as a function of both latitude and precipitation rate. Since our sites have essentially the same latitude, the variability in the predicted fluxes solely represents differences in precipitation. Hence the cause of the discrepancy in measured and predicted fluxes may simply reflect a different proportionality between precipitation and flux in the Graly et al. (2011) dataset, which included a world-wide compilation of sites, and the East River dataset."

Overall confusion results for me for several aspects centered around the term "predictions". In that regard, please explain how the "watershed-scale fluxes" are calculated, what the "forced-to-zero" predictions actually are. I had asked the first question already in my comments in the first round. Please address.

We have explained how the watershed-averaged fluxes are calculated in Section 3.4 'Comparing fluxes against estimates from other methods', a new section we have added to the manuscript. The specific text is:

"We applied the regression equations described in section 3.2 to generate a prediction of flux for each pixel in the watershed. We then predicted watershed-average fluxes by calculating the mean value of all pixels."

Minor comments

31: remove "garden variety" (whatever that means)

There is a fair amount of precedence for the use of the term "garden variety" in the literature. Garden variety means "common" and hence differentiates meteoric ^{10}Be from the from the less common in situ-produced form of ^{10}Be .

*See for example, page 136 of: Nishiizumi, K., Lal, D., Klein, J., Middleton, R. and Arnold, J.R., 1986. Production of ^{10}Be and ^{26}Al by cosmic rays in terrestrial quartz in situ and implications for erosion rates. *Nature*, 319(6049), pp.134-136, which states:*

" ^{10}Be studies can be made in any mineral as long as it can be established that it does not absorb rainwater-scavenged atmospheric ^{10}Be , which we designate as the 'garden-variety' ^{10}Be to distinguish it from that produced in situ."

Other examples of papers that use the term "garden variety" include:

*Nishiizumi, K., Winterer, E.L., Kohl, C.P., Klein, J., Middleton, R., Lal, D. and Arnold, J.R., 1989. Cosmic ray production rates of ^{10}Be and ^{26}Al in quartz from glacially polished rocks. *Journal of Geophysical Research: Solid Earth*, 94(B12), pp.17907-17915.*

*Monaghan, M.C. and Elmore, D., 1994. "Garden variety" ^{10}Be in soils on hill slopes. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 92(1-4), pp.357-361.*

*Small, E.E., Anderson, R.S., Repka, J.L. and Finkel, R., 1997. Erosion rates of alpine bedrock summit surfaces deduced from in situ ^{10}Be and ^{26}Al . *Earth and Planetary Science Letters*, 150(3-4), pp.413-425.*

*Fifield, L.K., Wasson, R.J., Pillans, B. and Stone, J.O.H., 2010. The longevity of hillslope soil in SE and NW Australia. *Catena*, 81(1), pp.32-42.*

*Kaste, J.M. and Baskaran, M., 2011. Meteoric ^7Be and ^{10}Be as process tracers in the environment. In *Handbook of Environmental Isotope Geochemistry: Vol I* (pp. 61-85). Berlin, Heidelberg: Springer Berlin Heidelberg.*

Willenbring, J. (2013). Meteoric ^{10}Be . In: Rink, W., Thompson, J. (eds) *Encyclopedia of Scientific Dating Methods*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6326-5_179-1

Hence we have opted to retain the wording “garden variety”.

35: “circulated through the atmosphere” – not sure what this statement is good for? Please remove or improve clarity, eg. by referring to the residence time of Be in the stratosphere vs troposphere

We have edited the text to state:

“ $^{10}\text{Be}_{\text{met}}$ resides in the stratosphere and troposphere for 2 years and 55 days, respectively (Yamagata et al., 2019). Hence atmospheric circulation transports $^{10}\text{Be}_{\text{met}}$ prior to its deposition on Earth’s surface by precipitation or dry fallout (Field 2006; Heikillä et al., 2008).”

36: “binds with minerals” is equally vague. It is coprecipitated, not only “adsorbed”.

We have edited the text to indicate:

“Once delivered to Earth’s surface, $^{10}\text{Be}_{\text{met}}$ adsorbs to and is co-precipitated on mineral surfaces, hence the highest concentrations are associated with secondary minerals and fine-grained particles with high surface areas (e.g., Graly et al., 2010).”

165: pH measurements: I would ask you to again check your pH measurements. Using ultra-pure water can result in readings to be very unstable. Using CaCl_2 leads to more reliable results (e.g. Minasny et al. (2011)).

We have done these measurements and report the findings in the Results and in Table S2. However, following Deng et al (2021 QSR) we use the formulation of Campforts et al (2016) to determine partition coefficients as a function of pH. Since the Campforts et al function uses pH measured in water, we use the pH measurements in water to estimate the partition coefficient.

215-220: I would move the list of uncertainty treatment to a footnote of a table.

We have moved that text to a footnote of Table 2.

244- units for erosion rate now differ from those used above. Streamline, please

We have clarified the variables and streamlined the text. It now states:

“We assume dz/dt values from eq. 4 are equivalent to ε values in eq. 3 and convert from units of length per time to mass per area per time using bulk density values inferred for the $N_{surface}$ sample (Table S3).”

247: “We also modelled F_{corr-e} using ε values ranging from....” is statement showing up well below the eq. 4, but in the caption of Fig. 4, I get the impression that you used eq. 4 for the modelling. So, what have you exactly done and can you please explain more properly how erosion rates were modelled and how a single model erosion rate (Fig. 4) can lead to regression lines shown in Fig. 4?

We have decided to remove Figure 4 and report the dz/dt values modeled with equation 4 (or the single value measured by Quirk et al., 2024) in Table 1. The lines in Figure 4 were based on the range of dz/dt values from Quirk et al and are from the initial submission. Since we now use the diffusion equation to model dz/dt for each site, the lines and the figure are not needed.

253: Now you present other methods. Subheading?

We have added several new subheadings (3.1 to 3.4) to the Methods to help separate and better explain the different analyses

275: I don't quite understand what you mean with “generate spatially continuous predictions of F_{met} ”.... Are you simply referring to correlations in Fig. 5?

Yes, we use the empirical regression relationships to predict flux for every pixel in the watershed, which results in spatially continuous predictions. To clarify the methodology, text in a new section 3.4 now indicates:

“The elevation, MAP, snow depth, and SWE data for each sample location were extracted from gridded datasets that span the entire watershed, including the Washington Gulch and Slate River tributaries. We applied the regression equations described in section 3.3 to generate a prediction of flux for each pixel in the watershed. We then predicted watershed-average fluxes by calculating the mean value of all pixels.”

At the end of this paragraph, you mention details on the GCMs used for comparison. This is out of place here. Move to e.g. results or even discussion section.

We have moved the description to a new paragraph in section 3.4, which is called: “Comparing fluxes against estimates from other methods”

307: Please refer to as “AMS-measured $^{10}\text{Be}/^{9}\text{Be}$ ratios”.

We have made this change.

411&412: Graly “et al.”

We have corrected that omission.

449: This is only unsurprising if MAP, snow depth and SWE equivalent haven't changed through time. How sure is this actually?

Rates of precipitation have surely changed through time. However, it is less likely that changes in precipitation rates would cause the spatially patterns of orographic precipitation to change. Even if average precipitation rates in the watershed were to halve or double, precipitation rates would still be higher at high elevation and lower at low elevation due to physical relationships among elevation, temperature, saturation vapor pressure, and the rate of water vapor condensation. We have added text to better inform readers that precipitation within the watershed is orographically-controlled:

“Observations of enhanced precipitation rates on the windward side of mountains demonstrates that orography plays a strong role in controlling precipitation patterns within the watershed (Heflin et al., 2024). Hence like other mountain landscapes with orographically-influenced rain and snowfall (e.g., Roe, 2005), mean annual precipitation within the East River watershed increases with elevation (Figure 2).”

452: I don't see a factor of 2-3 in table 4. Or I am I misunderstanding? Is this the reason for giving “forced” values so that they are comparable among the different variables?

Given that we have now modeled desorption losses, values in the table have changed. However the values for the 100% columns of (now) Table 5 range from 1.10 to 2.34 ($\times 10^6 \text{ atom cm}^{-2} \text{ yr}^{-2}$). We now indicate the values differ by about a factor of two:

“However, different slope and intercept values for regression relationships result in different F_{met} and F_{cor} predictions for the same location and watershed-averaged fluxes that differ by a factor of about two.”

So, why are these then all different if they represent the “same” forcing?

We think something may have been “lost in translation”. A common regression method involves making (forcing) the regression line pass through the origin, such that the y-intercept value of the regression is equal to zero. In American English, this method is commonly called “forcing a regression through the origin” or a “forced regression”. In this case, the term “forcing” does not refer to the factors driving the relationship between variables, it simply indicates that it is not a standard regression where the y-intercept can have a non-zero value. To clarify, we have edited the text to remove the term “forced regression” from the manuscript and replaced it with “regression through the origin”, which is longer, but perhaps more universally understood.

454: So that means that the integration time scales of the data is much shorter than the F_{met} from soil inventories....

Yes, this sentence (line 426) nicely summarizes that: “The precipitation data average over only a few decades and snow depth and SWE datasets cover only two winters, which are only small fractions of the post-glacial period over which 10Bemet has accumulated in soils.”

However, the integration times do not need to match in order to develop an empirical relationship between mean annual precipitation (or snowfall) and flux, provided the spatial patterns do not change very much over time.

488: Zheng’s flux values are for the modern atmospheric conditions only. Perhaps also use the modern value only from Heikkilä & FvB instead of using the average?

We have clarified this text; we do not use an average value from Heikkilä and von Blanckenburg (2015). Since our measured fluxes are Holocene-averaged values, not modern values, we don’t think comparison against modern values is the best approach. As explained in the response to the next comment, we have clarified the text in this section.

501: Now you refer to the “pre-industrial” value. I’m confused. If at all, I would compare the modern values between the GCMs (mainly because there are no pre-industrial values from Zheng’s model).

We have clarified the text to indicate:

“We compared our watershed-averaged flux predictions against three GCM-based predictions. The ECHAM5-HAM flux predictions of Heikkilä and von Blanckenburg (2015) are from a three-decade simulation that used pre-industrial aerosol and greenhouse gas concentrations and a modern solar modulation factor (ϕ) of 501.76 MeV. The solar modulation during the Holocene was lower ($\phi = 280.94$ MeV), hence Heikkilä and von Blanckenburg (2015) multiplied the fluxes by a factor of 1.23 to re-scale them to average Holocene values that represent the last 9,400 years. The GEOS-Chem 14.1.1 and ECHAM6.3-HAM2.3 model predictions of Zheng et al. (2024) are modern predictions that average from 2005 to 2013 and use a ϕ value of 488 MeV. Since our flux estimates from the East River average over the Holocene, we multiply both of the Zheng et al. (2024) flux predictions by 1.23, which approximately re-scales them to Holocene-averaged values and makes them more comparable to our predicted watershed-average values.”

Since the goal of our study is to compare Holocene-averaged fluxes against predictions from GCMs (and not compare one GCM prediction against another GCM prediction), we do not think comparing modern flux predictions from two GCMs is the best approach for this paper.

505: there is something missing in that sentence.

We did not detect anything missing, but the sentence has been edited to reflect the results of our new analyses:

“Hence, the GCM-based predictions generally fall within the range of our empirically estimated watershed-averaged flux values, although the degree of agreement depends on the predictor variable and whether or not fluxes are corrected for erosion and desorption.”

573: Your measured fluxes are lower, Graly rainfall-prediction gives higher fluxes....Can you give the Graly values in some comparison?

We have changed the wording – thank you for pointing that out. We have added the following text to help explain the discrepancy:

“The empirical formulation of Graly et al. (2011) predicts fluxes as a function of both latitude and precipitation rate. Since our sites have essentially the same latitude, the variability in the predicted fluxes solely represents differences in precipitation. Hence the cause of the discrepancy in measured and predicted fluxes may simply reflect a different proportionality between precipitation and flux in the Graly et al. (2011) dataset, which included a world-wide compilation of sites, and the East River dataset.”

Eq. 3: I thought the erosion correction approach was from Clow et al., 2020. Please cite.

We have added the citation.

Table 4: May be I am ignorant but what does a forcing to zero mean ? It gives quite different fluxes, so I would wonder about the physical meaning of these values.

Please see our detailed response above.

Also, I don't get these values in general: Have you simply calculated the mean values for each variable from the 4 different sites which you then termed "predicted"?

Yes, the values are the average value of the rasters shown in Figure 5. In addition to the updated text described above, we have also added a footnote the table (now Table 5) to further clarify the calculation.

Fig. 6: The 4 different figures all look pretty similar. Choose a different color -scheme?

We have stretched the color bar to accentuate differences. However, because each panel is a prediction of the same quantity and the regression relationships are not all that different, the maps should look similar.

Fig. 7: The citation of Graly et al., 2011 is not in the citation list.

We have corrected this omission.

Fig. 8: Given that I am confused regarding your "predictions", I would certainly spend some words on explain what the "watershed-scale" fluxes are that you present here, and how they are calculated.

We have added a new sub-heading to the Methods and the text there clarifies the calculation.

Comments from the Editor (first round)

I carefully reviewed the replies to the comments made by the reviewers. I do see the need to go into more detail, and need to see this manuscript again after these comments. I would also send this out again to the reviewers, as I think that in parts, the reviewers (mainly reviewer 1) comments were not adequately addressed.

The comment “Regression between ^{10}Be flux and environmental variables in Fig. 5. All these regression lines are strongly affected by the extremely high ^{10}Be flux data from Copper Creek. If this data point is removed (as demonstrated above), the regression becomes much less obvious”). “ was largely ignored, but I do agree with the reviewer that this needs to be addressed.

How solid is this regression?

We have conducted the regression analysis with and without the Copper Creek site and report the results in the main text and all of the regression coefficients and alternative regression figures in the supplement. The text now states:

“The Copper Creek site has substantially higher elevation and precipitation values than the other sites and hence strongly influences regression relationships. When the Copper Creek site is omitted from the analysis, R^2 values range from 0.25 to 0.93 (Table S7).”

Omitting the Copper Creek site lowers R^2 values, but they are still quite high (>0.8) for six of the fourteen relationships we explore. We are not sure what constitutes a “solid” regression, but there are still high correlation coefficients (and p values <0.05) when the Copper Creek site is removed. Hence our interpretations are not dependent on the results from this single site.

Do the symbols contain the uncertainties (hidden?) or what uncertainty has been propagated?

We have added error bars to the regression figures. We already report our uncertainty propagation methods for determining uncertainty in inventories and fluxes in several parts of the text:

“We assumed the concentration measured in the lowest depth interval was equal to the inherited $^{10}\text{Be}_{\text{met}}$ concentration and subtracted that value from concentrations measured

in overlying samples (e.g., Ruessner et al., 2010) and added the 1σ AMS analytical uncertainties in quadrature.”

“We propagated uncertainty for individual depth increments by adding uncertainty associated with the inheritance-corrected $^{10}\text{Be}_{\text{met}}$ concentrations and an assumed 5% relative uncertainty in both ρ and V_i in quadrature.” Note that the 5% relative uncertainty is what the Deng et al. study mentioned below did.

“Uncertainty in F_{met} for each calibration site was calculated by adding uncertainty in I_{met} values and moraine age in quadrature.”

In Fig. 5b, it looks to me as if the 4 samples cluster, and there is the one at high fluxes (the middle 3 values are certainly the same within uncertainty). It's clear that the other variables are just a derivative of precipitation, so that these, too, showing the same correlation does not support the argumentation. I would not elevation here, it might be totally misleading to people from the *in situ* community who might think that there is a sort of “scaling” with elevation like there is with *in situ*-produced nuclides. Please remove it accordingly.

The assessment of where the points plot is correct and it is also correct that all of the variables are related to one another, as we state in the text:

“ F_{met} is highly correlated with elevation, MAP, snow depth, and SWE, which is unsurprising because these variables are related to one another via orographic controls on precipitation rates.”

However, this comment does not consider a key aspect of our argument, which is that orography (e.g., elevation) is what controls precipitation, and hence the flux of meteoric ^{10}Be in the watershed. Removing the elevation data would substantially weaken our analysis and we have opted to keep it in the manuscript.

I can see however that snow depth and SWE might be relevant, as these could present a more time-average of MAP. Within this possibility, could you discuss why the MAP-derived flux is lower than the snow-based ones? (see comment by reviewer 1). Why would MAP data underpredict precipitation? Could you discuss this?

As we state in the manuscript, snow depth and SWE data are only available for two winters, so we don't think we can argue that they represent long term precipitation rates. However, they likely capture long term patterns in precipitation, which is sufficient for generating regression relationships that can be used to predict flux for areas of the watershed that lack a measurement.

Because the flux predictions are based on empirical regression relationships, there is little physical insight to be gleaned. The snowfall and MAP data don't exactly track one another one to one, so the differences in the slopes and intercepts of the regression

equations result in different watershed-averaged predictions. If we take the inheritance-corrected fluxes as an example, the maximum snow-based prediction is 1.43×10^6 atoms $\text{cm}^{-2} \text{yr}^{-1}$, whereas the lowest MAP-based prediction is 1.10×10^6 atoms $\text{cm}^{-2} \text{yr}^{-1}$. Given the empirical basis for the predictions, the differences are expected. Whether the differences are large or small will likely depend on the perspective of the reader.

Regarding ‘why would MAP underpredict precipitation’; there is not a weather station at each study site, so we do not know what the true rates of modern-day precipitation are. Because we lack data for each site we use a gridded MAP product, which is calibrated using data from weather stations, but the modeled data undoubtedly over- and under-predicts some of the calibration data, and so also undoubtedly over- and under-predict precipitation at sites with no measurements.

Regarding pH, I also agree with reviewer 1 that at a pH lower than 6, “only minimal chemical loss of $^{10}\text{Be}_{\text{met}}$ ” cannot be assumed. The reason I do not agree with the reply (which is based on Graly et al. ’s soil compilation) is that in Graly compiled data, only 1 profile is included that is developed on glacial till, but there, pH hasn’t been measured. Hence, the quoted pH cut-off is likely valid for other materials, where humic acid (shifting the pH dependency) and grain size (supplying adsorption space) might all matter. Clow et al. ’s moraines by the way have quoted pH values of 5.5 to 8, so on average (much) higher. I would therefore expect to see a calculation on potential loss due to low retentivity, and its effects on the depositional flux.

We have conducted an analysis of desorption loss in our revised manuscript. Please see section 3.2 of the manuscript, which includes substantial additions.

Next, as mentioned by the reviewer, the clast size $>2\text{mm}$ were not quantified, and I am not satisfied by the reply. What is then the estimated proportion of the large clasts (and what is the consequence for the ^{10}Be concentration?). Please clarify. I cannot see any discussion on this in the MS, although it reads as if there is data on this.

This comment seems to be focused more on the initial manuscript we submitted, rather than the revised manuscript. Based on the comments of the reviewer on our initial submission, we conducted an analysis of $> 2 \text{ mm}$ clasts that closely follows the methodology of Deng et al (2021) and included those results in our revised manuscript.

For example, the manuscript indicates:

“The inventory was determined by multiplying the inheritance-corrected $^{10}\text{Be}_{\text{met}}$ concentration (N_{corr} , atom g^{-1}), by soil bulk density (ρ , g cm^{-3}), the thickness of the sampling interval (z , cm), and the volumetric fraction of fine ($< 2 \text{ mm}$) soil V_f , and

summing the values for all (n) individual sample depth increments (i) (e.g., Deng et al., 2021b):

$$I_{met} = \sum_{i=1}^n (N_{corr} \cdot \rho \cdot z \cdot V_f)_i \quad (1)."$$

and

“The volumetric fraction of < 2 mm soil particles was calculated for each depth interval following Deng et al. (2021b) as: $V_f = (1 - V_{c1}) \times (1 - V_{c2})$, where V_{c1} and V_{c2} are volumes of two sizes of coarse clasts. V_{c1} was determined by sieving each sample to separate < 2 mm and > 2 mm clasts and the volume fraction of > 2 mm clasts was calculated assuming a density of 2.65 g cm^{-3} . V_{c2} was estimated by using field photos of each soil pit to digitize the area of large cobble- to boulder-sized clasts that were too large to be sampled by our 44.7 mm diameter steel ring. The volume fraction of large clasts was calculated by multiplying the area fraction by two-thirds (e.g., Deng et al., 2021b).”

The data are reported in supplementary tables. We previously reported the fine fraction only as that is all that is needed for the calculation and the coarse fraction can be easily calculated as 1 minus the fine fraction. However, based on this comment, we now report the fine and coarse fraction in Table S3. Since the reviewer indicated we should do the analysis a different way (which we did), rather than ask us to compare the results from two different approaches, we don't believe any discussion is warranted.

Related to this is the assumption on bulk density value: How good is this assumption? The argument here is that the same bulk density as a till from a previous study in the valley (is this now the Copper Creek till?) was used. How variable is the bulk density in tills? Is there any effect of post-depositional compaction in tills, and if so, could that increase/ change with age?

This comment seems to relate to our initial submission and not to the revised manuscript we submitted. Bulk density was measured for each soil horizon in each profile (not just from the in situ-produced profile at the Gothic moraine site), we have updated the text to make that clear:

“Bulk density was determined at each site by driving a 90.8 cm^3 steel cylinder into the soil at depths corresponding to different horizons and weighing the dry soil (Table S1).”

and

“Since we measured bulk density by soil horizon, and not sample depth interval, we assigned bulk density values to each sample interval based on soil horizons identified in the field (Table S3).”

Table S1 indicates the variability of density values.

The tills have not compacted since they were deposited, though even if they had, it would not change the way inventories are calculated. Rather than post-depositional

compaction that increases density, the density of the tills has declined since deposition, due to chemical weathering and volume changes associated with bio- and cryoturbation.

What rationale is being used for assuming that the measured values at depth correspond to the inheritance? (Concentrations could further decline, no? You don't know, you haven't sampled).

This is all very speculative. Of course the concentrations might further decline, but they also might not decline. We have followed the methods of prior studies (e.g., Reusser et al., 2010; Deng et al., 2021) in using the lowermost sample to assess inheritance. Those studies could of course also have collected deeper samples and reached different conclusions regarding inheritance.

In a scenario of temporally-uniform erosion, the measured inheritance should represent the erosion rate measured by in situ ^{10}Be . Converting this meteoric ^{10}Be concentration into an erosion rate is possible, but might be prone to bias (grain etc.). But, perhaps worth a try?

Without some theory to support this statement, it is unclear if this is the case. In an in situ-produced ^{10}Be profile (from a fluvial terrace, for example), the inherited concentration equates to a paleo-erosion rate, which would equal the present-day erosion rate if erosion were temporally uniform. However, if a glacier excavated regolith from a valley bottom that had been exposed to meteoric ^{10}Be deposition for 1,000 years the resulting moraine would have a very different initial (inheritance) concentration than a moraine formed in a valley where meteoric ^{10}Be had accumulated in regolith for 100,000 years prior to glacial erosion. Similarly, a moraine formed from material eroded from a few meters depth versus many tens of meters depth would have different initial (inherited) concentrations of meteoric ^{10}Be . The later example is a more likely scenario in the East River, hence it is unclear to us how the concentration of inherited ^{10}Be would yield information on the erosion rate of the moraine.

Some minor comments:

Fig. 3: Please label each profile at the top with the name of the till.

We have made this addition.

I would support the citation of Deng et al., 2021, Deposition and retention of meteoric ^{10}Be in Holocene Taiwan river terraces. Quaternary Science Reviews 265.

We agree. We've cited the paper extensively in the manuscript but mistakenly did not include it in the reference list.

I don't quite understand the term "watershed-scale" ^{10}Be fluxes- how is this calculated? Is this based on the distribution of elevation within the watershed and then the average flux is computed? (So, sort of an average of each map shown in Fig. 6?).

Please clarify.

Yes, that is how it is calculated. We have revised the text to clarify:

"The elevation, MAP, snow depth, and SWE data for each sample location were extracted from gridded datasets that span the entire watershed, including the Washington Gulch and Slate River tributaries. We applied the regression equations described in section 3.2 to generate a prediction of flux for each pixel in the watershed. We then predicted watershed-average fluxes by calculating the mean value of all pixels."

As soils can have relatively high ^9Be concentrations, I wonder if the $^{10}\text{Be}/^9\text{Be}$ (AMS) should be corrected for the natural ^9Be concentration?

It depends on how one interprets 'relatively high'. We dissolved an aliquot of each sample in HF and measured ^9Be – please see new paragraph at the end of section 3.1. The maximum amount of ^9Be we measured was equivalent to 0.30% of the amount of ^9Be added as carrier. The average was 0.18%. The values are smaller than the uncertainty in the ^9Be measurement of the carrier we used for this study (1.2%, based on multiple measurements of multiple dilutions). Hence we don't think they are something that we need to worry about for this study and have added this text:

"Concentrations of ^9Be were ~1 ppm indicating the average mass of ^9Be in the samples was ~0.2% of the mass added as carrier (Table S5). Given that 0.2% is less than the uncertainty of the Be concentration of the carrier we used (1.2%), no corrections were made to the AMS-measured $^{10}\text{Be}/^9\text{Be}$ ratios."

Had the samples been smaller and had we added less carrier, then it would have been something that needed to be accounted for.