Dear Prof. Vermeesch:

Thank you for your acknowledgement of our effort to revise our first submission, and for the effort in searching for qualified reviewers for our manuscript. We appreciate the comments and suggestions from Dr. Balco and you, and we have made changes to our manuscript accordingly.

We concede to Dr. Balco's and your argument that the inheritance and effective exposure age must be constrained as positive during inversion. Therefore, we have rewritten the related section (3.1.2) along with appropriate simulation examples to emphasize the importance of doing a constrained least squares inversion when either exposure age or inheritance are small and negative results may arise. We have also revised the codes that we uploaded to GitHub to return constrained results.

We also appreciate your suggestion of changing the inheritance and effective exposure age into exponential forms. We have investigated this transition, but we are unable to transform it into a linear inversion problem. We therefore decided to keep the original form in order to preserve the simplicity of the approach.

We also made some minor revision following Dr. Balco's suggestions, as described below:

Reply to comments from Dr. Balco:

Why needed? One of the issues that has come up in the discussion of this paper is that from the applications perspective there is not a strong need for a simplified inversion. Depth-profile data are not collected in great quantity and there is not a science application that is currently seriously hindered by the computational time needed to do a full forward model inversion. For myself I am not worried about this issue and I don't think it's an obstacle to publication. For one thing, it is potentially useful for making sure that a more complex inversion scheme is working correctly. Also, I can envision a fast inversion method being useful for database applications in which one seeks to compare a lot of depth-profile results using different production rate scaling methods, or something of that nature. Of course it's actually not that fast because you still need to estimate all the production rates due to muons, which requires a site-specific muon production calculation, which in turn requires a bunch of numerical integrations no matter what. Regardless, however, even though a simplified inversion method isn't of dire immediate need for any current application, it is certainly something that is potentially useful.

Thank you for your comment. We appreciate for your acknowledgement of the potential of our approach. We envision our approach as a compliment to existing forward models, and it is a convenient tool to explore the trade-offs between exposure age and erosion rate, as we state in the revised introduction: "As an inverse approach, linear regression directly returns a best estimate, providing a clear way to explore how model inputs, such as erosion, affect the result."

The issue of negative inheritance.

1. As written, several parts of section 3.1.2 are not correct, for example "imposing the physically reasonably prerequisite...may lead to underestimation of the exposure age." It's not a physically reasonable but optional prereequisite, it's a requirement. It's also inaccurate to say that it leads to *underestimation of the exposure age – what it actually does is lead to an incorrect uncertainty distribution, and the lower value of the age appears to be the result of improperly using the mean to represent an asymmetrical distribution. In the constrained regression problem, the uncertainty distribution is not expected to be symmetrical about the true value. Thus, this section needs revision so that it correctly outlines how (i) the actual regression problem is a constrained linear regression that is expected to lead to complicated uncertainty distributions when the constraints are operative, and (ii) applying an unconstrained regression is a simplification that is only correct when Cinh >> 0 and Te >> 0.*

2. It's also probably worth a brief discussion of the case where a surface is quite young, so that the uncertainty distribution for t runs into zero. The unconstrained regression is also inappropriate in this case. This could possibly occur even when t is fairly large if the inheritance is also large, such that TeP < Cinh.

Thank you for the detailed analysis and explanation of the negative inheritance problem. We agree with the reviewer, and we have rewritten section 3.1.2 accordingly. We have also corrected the related statements in section 4.2.3.

3. The Beida River / Fig 8 analysis, which improperly uses an unconstrained regression when constraints are operative, should be redone with the correct, constrained regression. Of course for the Lees Ferry analysis, the constraints are inoperative, so the unconstrained regression is fine.

We have updated the data, figure, and analysis for the Beida River, based on the constrained model.

Other items.

The most important one is that the description of the effective attenuation lengths for muon production (line 62) is oversimplified and therefore somewhat misleading. There is no single attenuation length for either fast muon production or negative muon capture production, because the nature of the production process is such that as depth increases, the energy of the remaining muons increases, so the instantaneous attenuation length for the production process also increases. Thus, describing m1 and m2 as the attenuation lengths for these processes is not correct. The values of 1500 and 4320 g cm2 that are in Table 1 in this paper, which were given by Heisinger as approximate values that could be used in simplified erosion rate integrations, are not correct at any site or depth, except possibly by accident. Using these values in an application with significant muon production would most likely yield a result that was quite wrong. However, it is true (see Balco, 2019, section 8) that it is usually possible to represent total production by muons in a finite depth range at a specific site as the sum of two exponential functions. Although the authors' response to the reviews stated that "We updated muon production rates used in the pseudo profiles and in the Beida River case,' this is not evident from the text, and as far as I can tell, it appears likely that what they actually did was modify the surface production rates and not the attenuation constants. More seriously, it is not possible to tell whether they used site-specific values for the Beida River and Lees Ferry example, or if they used the incorrect values in Table 1. At the very least, the authors should modify the text near line 62 to indicate that the values of P and Λ pertaining to muon production are not, in fact, true production rates or attenuation lengths for any *particular process, but instead are site-specific constants that come from fitting to a more complicated production model.*

Thank you for pointing this out. We realize that the statement we made in the manuscript is misleading, therefore, we modified the text originally in L62 as "Note that the two exponential terms for muogenic production used here are an approximation of a complex muogenic production path, because the muon energy spectrum hardens continuously with depth (Helsinger 2002a, 2002b; Marrero et al., 2016; Balco 2017)." (L64-66 in the revision), we also added a footnote to Table 1 and S2 to clarify this. For both the Beida River and the Lees Ferry examples, we updated the inversion results using site specific muon production rate and attenuation length based on Model 1B of Balco, 2017. This leads to less than 1% of difference between the old and revised age estimations (Section 3.2).

Starting in section 3 of the paper the authors contrast the results of their regression scheme with a full forward-model-fitting scheme that they describe as 'Bayesian.' This is confusing because the important point is that this method employs a full forward model to predict the observables – the contrast between this method and the regression method would be the same whether the approach to choosing the best values of the model parameters approach was Bayesian, frequentist, or something else. Likewise, it would be possible to perform a Bayesian linear regression. Thus, it would be more helpful to the reader to describe this alternative method as 'forward model fitting' or 'forward model optimization' rather than 'Bayesian' by itself.

Thank you for pointing this out. We agree with the reviewer that using Bayesian in our previous revision is confusing for the readers. We have replaced 'Bayesian' with 'forward model fitting' and 'forward approach' in the revision.

Near line 40, the authors use the term 'derivative' to describe refined approaches that are derived from the initial regression approach. The use of this word is confusing here, because in the context of a paper like this one that is mathematical in nature, it implies to the reader that these approaches will involve derivatives in the mathematical sense. For example, in line 44, the use of 'second derivative approach' indicates to the reader that the approach will involve second derivatives of some function or field, which is not the case.

Thus, 'derivative' should not be used here. Possible improvements would be to use 'specific approaches' or 'special cases' to contrast with the 'general approach' in line 40.

Thank you for the suggestion, we have changed the 'derivative' approaches into 'specific approaches'.

Finally, a thoroughly insignificant point is that I disagree with reviewer Alan Hidy about the abbreviation 'TCN.' I see no reason that the method in the paper could not be applied to depth profiles in lunar or Martian regolith! The authors should use their best judgement here.

We have replaced 'TCN' with 'CN'.