Universität Potsdam • Institut für Geowissenschaften
Karl-Liebknecht-Str. 24-25 • 14476 Potsdam-Golm GERMANY

Editors – Earth Surface Dynamics

University of Potsdam
Institute of Geosciences

Dr Mitch D'Arcy
Alexander von Humboldt Postdoctoral Fellow
mdarcy@uni-potsdam.de

Response to reviewer comments of ESurf-2019-21: "Inferring the timing of abandonment of aggraded alluvial surfaces dated with cosmogenic nuclides"

Dear Prof. Mudd,

Thank you very much for the two reviews of our submission to ESurf. These are thoughtful and constructive reviews that have helped us to improve the manuscript, and we are grateful to both reviewers for considering our work so carefully. We are pleased that they both liked the essence of the manuscript.

We have incorporated many, though not all, of the recommended changes. Please find attached our detailed responses and edits. Page/line numbers refer to the revised version of the manuscript with tracked changes showing "all markup". Line numbers refer to the final revised version of the manuscript with tracked changes.

We hope that our manuscript is now suitable for publication in ESurf. Please do not hesitate to contact me for any further information.

Yours sincerely,

Mitch D'Arcy

Dr Mitch D'Arcy
(on behalf of all authors)

Reviewer 1: Dr Luca Malatesta

Dear Editor, I have read the latest a manuscript by D'Arcy and colleagues with pleasure. They offer a new probabilistic approach to identify the likeliest age of abandonment of an alluvial surface based on series of exposure-dated samples at its surface. They build a power law that predicts the likeliest amount of time elapsed between the youngest surface age and the effective fluvial incision based on the distribution of surface ages assuming their uniform distribution during the period of activity of said surface. It is a useful contribution that can be applied widely and is definitely worthy of publication in ESurf! In my opinion, the manuscript is ready for publication pending minor clarifying modifications. The article is very well written and easy to read. I would however encourage the authors to consider modifying their probabilistic approach and follow an explicit derivation of their probability power law without requiring the use of "artificial data" for empirical fitting.

We thank Dr Malatesta for his helpful review. We are pleased that he likes our work, and we respond to his comments below.

Below I briefly describe an alternative approach for the probability law and I provide line by line comments on the text.

Probability The approach using synthetic data has the advantage of mimicking a field situation with n dated boulders out of a larger number. However, it seems to me that using an explicit approach would be much more advantageous. There is no need to graphically fit the powerlaw and deal with the associated error margins, the term "artificial data" can be avoided altogether, and the theoretical framework would be reinforced. Further it would become a more flexible platform, for example to introduce non-uniform distribution of surface ages. I have asked Quentin Berger, probabilist at Paris-Sorbonne, for some help as to how the explicit derivation can be made. I include a document that summarises his explanation hereby. The derivation would replace section 4.1 and provide a definitive and clean solution for this approach. I think it would improve the impact of the manuscript. That being said, it is not a necessary modification and the manuscript stands on its legs as is. It is for the authors to decide whether they want to follow an explicit approach or not.

We agree that it's worthwhile to consider an analytical solution, and we're grateful that Dr Malatesta and Dr Berger have taken the time to derive these suggested equations. Nonetheless, there are several reasons why these equations cannot replace the approach we take using artificial data.

First, let us quickly summarize how we understand the derivation in the document. Essentially, the probability is evaluated that a particular sample is older than the time of abandonment (set to zero in the document) plus a specified duration τ (eq. 4). This period τ is identified with the time difference between the youngest sample and the time of surface abandonment (which is also called τ in our manuscript). Because we assume uniform distribution of sampled boulders with equal likelihood of sampling, the probability is given by 1- τ/T, where T is the length of surface activity. Next, it is required that all samples fulfil this criterion, and consequently, the probability is raised to the power of n, where n is total number of samples (eq. 5). Finally, assuming $\tau << T$, the equation is expanded to first order, using a Taylor series, noting that the result is equal to a Taylor expansion of an exponential function to first order.

There are a number of points that can be made in response.

Firstly, the suggested equations are incomplete. A comparison of eq. 5 in the derivation and eq. 2-3 in our manuscript shows that our equations yield more detail, for example a prefactor to the exponential term (Eq. 2) and the dependence of the exponent on the specified percentile (Eq. 3). Further details of our results, for example the relationship between T and the spread of sample ages (Eq. 6) are not dealt with. Therefore, the derivation may provide a first step, but further steps still need to be worked out.

Secondly, we do not think that the derivation actually reproduces what we are simulating with the artificial-data approach. The youngest sample age in our approach is not older than τ, but determines τ (i.e., we require a sample with age τ). Thus, the determination of the probability is not correct (Eq. 4 in the derivation). We have since developed some ideas of how to correct the equations, but this is far from giving a usable or publishable result. Whether the exponential approximation that arises from the derivation is coincidental or whether there is a relationship to our approach is not yet clear to us.

Thirdly, even if a complete analytical solution is possible (which it might not be), it seems likely that a numerical solution or artificial data are necessary to provide other elements of a workable approach. An example is the estimation of T, where only the span $a_{max} - a_{min}$ can ever be measured empirically (i.e., Fig. 6 and section 4.2). There are actually several advantages to choosing an artificial-data approach, which we mention at the end of section 2. One additional reason that is not mentioned in the manuscript is that the artificial data may be easier to understand for researchers who do not have a rigorous mathematical background. Our equations are correct, the error margins of our parameterisation are very small (see Fig. 5), and importantly, our equations do not require approximations such as $\tau/T << 1$ such as in the suggested derivation, which Fig. 4 shows is not realistic.

We are very open-minded about the possibility of developing a full analytical solution in the future, but this is a complicated problem to solve and the suggested equations only provide a starting point. For these reasons, we believe it is advantageous to continue with our approach using artificial data. We decided to not include analytical derivations in the present paper.

Line by line

p. 2 L. 31. “These approaches risk circular or inaccurate interpretations.” Can you elaborate or give a few examples of these risks?

Yes, we have clarified the text. We changed “*approaches*” to “*assumptions*”, because this sentence is referring directly to the previous sentence where we open this point with two specific examples, including citations. We added an additional citation to Macklin et al. (2002). We now explain in the text that these examples (1) assume that abandonment coincides with palaeoclimate events, and then conclude that climate controls aggradation/incision cycles (risking circularity); and (2) assume that the youngest sampled age approximates abandonment, which our analyses show will often not be the case (risking inaccuracy).

p. 3 L. 13-15. I suggest to indicate that these ages are arbitrarily selected to produce the scenarios. The reader (or at least I) might think that they are lucky draws from random rounds and that you are already talking about experiment results. It’s a small detail but it would help focusing on the examples you are building.

Done, this is a good suggestion.

p. 4 l. 4. “In this study, we use artificial data [...]” At this point it can be unclear whether you use artificial data on virtual surfaces or if you populate a real “geomorphic surface” with artificial data. I suggest to maybe include the purpose of the approach here already: e.g. “we use artificial data to simulate the characteristics of surveyed surfaces” (which you bring up only later at the end of the paragraph on l. 9-10.) This entire paragraph is actually paramount as it frames the use of “artificial data” for the first time. I suggest to carefully edit it such that the combination/coexistence of artificial data and field sites is clear. At this point in the text, Many readers will be asking themselves “ok i understand the problem and motivation but how is that useful for my field site?”.

These are good points. We have edited the paragraph to make our approach and its utility clearer, and elaborate on how our synthetic data approach can inform real field studies.

p.4 l. 27. Missing coma after “T”

Corrected.

p. 4 l. 27-28. I suggest to indicate the uniform distribution of the ages here already. The readers might be wondering about it.

We agree that it’s important to point out the uniform distribution of selectable ages, but we think a better place to discuss the assumptions of our approach is section 3.3 ‘Experimental assumptions’ (now at p.6, l.12), once the reader is familiar with our overall approach involving artificial data. We now address this particular assumption explicitly in sections 3.2, 3.3, 5.1, and 5.3, which we think is in good context.

p. 5 l. 13. tau = a_min - t_aban is an important relation, I’d suggest to give it a full equation line.

Good idea, we have done this. We put the equation at the start of section 3.1 and re-numbered the other equations accordingly.

p. 7. The lines of equations lack punctuation.

We’re not sure what punctuation is missing from the equations. If our article is accepted then we are of course happy for it to be formatted following ESurf style guidelines.

p. 7 l. 6. “then tau = 12 kyr for P = 0.95.” I’d suggest to paraphrase the end of the sentence in plain english for clarity.

Done.

p. 7 l. 11-15. the parameter k has a negative value. It should be mentioned here (important for what happens when n grows to infinite). Potentially even better âA˘T and ˇ I believe in accordance with the convention for such parameters âA˘T give k a positive ˇ value with an explicit negative sign in the equation.

We now explicitly point out that *k* has a negative value (p.7, l.34).

p. 8 l. 21: section 4.3 is very good and will be very useful.

Thanks!

p. 8. l. 24: using the parameters values listed above I assume? It might be worth specifying it.

Yes, the parameters (and equations) we derive from our artificial data are universal. Rather than specifying this here, we have added a line, "*Equations 2 through 6 are thus calibrated using our artificial data*…) to the end of section 4.2 above (p. 9, l. 11-11), to make this clear.

p. 9 l. 11-17. this paragraph reads a little like conclusion material. I am not sure it is necessary.

We disagree. This is the opening paragraph of the Discussion and we think it should briefly outline the key implications of our work, namely that abandonment ages will often be more informative than average surface ages, and that our probabilistic approach provides a new way of constraining abandonment. Given that readers thinking about their own field sites will likely jump to this subsection (5.1, "Implications for surface dating"), we think that a very brief discussion of the key points is helpful.

p. 9 l. 23. "significantly" probably needs defining since you provide a quantity of "one order of magnitude" thereafter.

We changed "*significantly*" to "*substantially*", as this sentence is only supposed to be a qualitative statement.

p. 9, l. 25. There is no figure 5d.

This was a typo, we have corrected it to Fig. 5b. Thanks for spotting it!

p. 12 l. 10 Without much context, I don't see why that would be a "conundrum".

We changed "*This conundrum could be partly resolved*…" to "*More realistic values can be obtained*…" (p.13, l.19).

p. 13 l. 2-4. I am not sure that this characterisation is fair to previous work, many authors showed the importance of timing abandonment and not mean ages. The standout "finding" of the present manuscript is to propose a simple and efficient method to get there using incomplete datasets. It's an important step.

We think this is referring to p. 14, l. 2-4, rather than p. 13 (which does not refer to previous work). The majority of studies that date surfaces such as alluvial fans do simplistically represent the surface with an average age (whether a mean, mode, or the peak of a frequency distribution) and rarely attempt to infer the subsequent age of abandonment (although we do explicitly acknowledge several examples in the Introduction). We're certainly not claiming to be the first to consider abandonment, but we do feel it is fair to conclude in our paper that "*the timing of surface abandonment may provide more informative and more precise interpretations than taking an average of measured surface ages*", because one of the novel contributions of our work is quantify the precision with which abandonment can be inferred.

For example, Fig. 4 demonstrates that for desirable probabilities, the timing of abandonment can indeed be pinned down more precisely than the period of surface formation, T. Similarly, the example application to the younger Q4 surface on the Baja California fans (Fig. 8, left) shows that surface abandonment likely overlaps with the Younger Dryas, offering a more precise and informative interpretation than the average surface age (for reasons we elaborate on in section 5.2.1). Both of these results demonstrate the value of inferring abandonment, and the potential precision with which this can be accomplished, in a new way that hasn't been demonstrated before.

Reviewer 2: Anonymous

In this manuscript, D'Arcy et al. propose a new probabilistic approach to constrain the timing of alluvial surface abandonment using cosmogenic radionuclide dating of surface boulders. Using randomly sampled surfaces ages from a hypothetical alluvial surface, a distribution of surfaces ages are obtained where both the number of samples and age of the depositional surface are varied. The discrepancy between the ages sampled and the true timing of surface abandonment are then determined. The relationships drawn from this analysis are then also applied to an independently dated alluvial fan system in Mexico to infer the timing of surface abandonment. The manuscript is motivated by better constraining the timing of surface abandonment; the authors suggest that this may be a more useful constraint than an average surface age which is unlikely to relate to any particular forcing or event of interest. In contrast, the timing of abandonment will likely reflect changes in climate, base level change, tectonic forcing or major drainage reorganization. I enjoyed reading this manuscript – it addresses a well thought-out set of questions, is very well written and I believe is a valuable contribution to the field. I would recommend the manuscript for publication pending a few very minor clarifications.

We thank Reviewer 2 for his/her thoughtful review, and we respond to his/her comments below. Reviewer 2 has raised some very interesting questions that we hope will inspire future studies!

I have a general query about boulders and age distributions and their representation. In the conceptual model, it is assumed that boulders are evenly distributed across the surface and that there is a uniform probability distribution of selectable ages. This is mentioned in the experimental assumptions too (section 3.3). I am curious as to how much these two assumptions are likely to modify the modelling results, and whether these assumptions are actually more likely to be the norm in reality. Is this by any chance something that has been examined or tested? The fact that many alluvial surfaces are not characterized by large numbers of large boulders does indeed suggest that their delivery downstream of their source areas may be temporally clustered and correspond to very large events – this may not be relevant given that it is only the youngest ages which matter here. In general, the authors do a thorough job of highlighting the assumptions and limitations of their approaches.

We're pleased that Reviewer 2 thinks we have thoroughly highlighted the assumptions and limitations of our approach, because we want to be upfront about these.

We have not yet performed explicit tests with different distribution shapes of selectable surface ages. One of the reasons for this is simply because it would multiply the analyses in our current manuscript by x number of different distribution shapes, and each would require a substantial amount of consideration (to work through the predictive relationships and equations), and too many figures and analyses for one paper. However, we do agree that it would make a very interesting question for a future study, which could also bring in a compilation of age clusters from well-sampled alluvial fans in order to empirically look at what shapes these distributions tend to have in the real world.

For our work here, the main implications of changing the shape of selectable age distributions would be to (i) change the number of samples needed to get an accurate estimation of T; and (ii) change the value of τ for a given probability/number of ages/T. We can speculate here with two cases:

1. **Selectable ages are normally-distributed with a peak in the middle of T.**

A slightly greater number of ages would be needed to estimate T accurately, because you're more likely to end up sampling ages that cluster around the middle of the depositional timespan, as opposed to distributed randomly throughout T like in our scenarios. Therefore, everything in Fig. 6 would probably be shifted down to slightly lower ratios of $(a_{max} - a_{min})/T$.

Next, a_{min} would presumably fall further from t_{aban} in many cases, which would make τ slightly larger for a given value of P. The magnitude of this effect might in turn depend on n and T (i.e., a bit like Fig. 2). So for small values of n changing the distribution shape might have a bigger effect, but as n increases, perhaps the results would become more similar to ours.

2. **Selectable ages are biased towards the youngest end of *T* with a long tail decaying towards the older end of *T*.**
   Again, slightly more ages would be needed to estimate T accurately, simply because the sampled ages will always be biased by clustering (wherever the cluster sits within T). However in this scenario, you're more likely to densely sample near to the timing of abandonment, which should result in a smaller value of τ for a given value of n, T, and P, i.e., more accurate estimates of abandonment timing.

So, we speculate that different distributions of selectable ages would result in different effects, probably changing the size of τ by some small amount. A dedicated study would be needed to pin these effects down. We think that a good way to go would be to start with a compilation of measured ages from natural fans, to see if ages appear to be randomly distributed throughout a timespan, or with a particular distribution shape of the tails.

Section 3.1 – The first time I read this section I was a little confused – it felt like the second paragraph was more observations made from the data rather than a description of methods (line 20-25 in particular). Perhaps some re-phrasing or reordering of material may help with the flow of this section.

We agree that the text needed some clarification here, which was also raised by Reviewer 1. We have edited both section 3.1 and the preceding section 2 (see response to Reviewer 1). We chose to keep the reference to the example case in Fig 1 (referred to as p.4, l.20-25 above) because it illustrates the key advantage of using artificial data and bridges the Justification and the Methods section. However we have added a sentence after this point that explicitly points out why an artificial data approach makes sense, and edited the section to make the text clearer.

Fig. 1B – Could you add a y-axis on the kernel density plot?

In principle we could, yes, but it would be somewhat meaningless because the values would only reflect the size of the x-axis bins used to make the frequency distributions. This is only a cartoon illustration so we feel that would complicate the figure unnecessarily, and we left out the y-axis. Of course, later in the paper the y-axis becomes meaningful when we develop the probabilistic approach, so we do then add y-axis labels.

P4. L4-11. Again, I had to read this paragraph a couple of times over to work out what was a 'true' timing and a 'real' surface – some confusion on what you have modelled and what is a 'real' example. You also do not mention/introduce that you apply the modelling to a case study in either section 1 or 2. Instead, it does feel like it pops slightly out of the blue during the paper discussion. Perhaps integrate this into the end of section 1 where you outline what you are going to present with respect to the artificial data and generation of probabilistic equations.

These are good points. We have edited the paragraph (now p.4, from l.8), also in response to Reviewer 1, to make it clearer. We have rephrased "*real surfaces*" as "*natural surfaces*" for clarity. We also agree about flagging up the case study earlier on; as suggested, we now mention this at the end of section 1.

P9. L13 – I don't think it is unreasonable to say that an average age does not/should not correlate with an external forcing.

This is a good point, this sentence can be phrased in a better way. We have changed (now at p.10, l.16):

> "*…our findings indicate that averages of sampled surface ages are likely to be imprecise representations of the mid-point of surface formation, **and** may not **correlate** with **any** external forcing event…*"

to:

> "*…our findings indicate that averages of sampled surface ages are likely to be imprecise representations of the mid-point of surface formation, **which** may not **coincide** with **a particular** external forcing event…*"

We agree that our results do not explicitly demonstrate that average surface ages will not correlate with external events. They do demonstrate that average ages will often be imprecise representations of the actual average surface age (i.e., Fig. 3a), so we have kept the first part of the sentence. For the second part, we now simply point out that the average surface age might not coincide with external forcing events, for the reasons we discuss in section 2 (Justification). The mid-point of surface formation is, by definition, in the middle of a period of stability when a surface continues being deposited, unlike when the switch from activity to abandonment occurs.

P.9 L 21. This is probably more for my own curiosity. For the variables you have modelled, you state that 6 to 7 ages are sufficient to characterize the timing of abandonment when T = 30 kyr. You also touch on this in section 5.3 but was wondering if you could just clarify/expand. In your artificial case, the period of surface activity is defined. What if you turn up at a new field site without any indication of how old/period of time each surface has been active for? How many samples are needed/adequate to estimate the timing of surface abandonment to a high degree of probability? Perhaps some idea of the periodicity of forcing mechanism needs to be known (if climatically driven) – but then the argument becomes somewhat circular! Or should we just grab as many samples as we can and state the uncertainty/probability?

These are great questions, and we have been thinking about this issue of how many samples to collect too! It's true that Fig. 3 is referring to the case where T = 30 kyr, but Fig. 4 on the other hand is looking at a wide range of values of T, and we still see that the curves really start to level off after about 6-7 samples. The value of τ is larger when T is larger, but collecting a few extra samples (say, 10) rarely counteracts the effect of increasing T. In other words, you don't know what T is when you're sampling in the field, but whatever number of samples you collect it's unlikely to make much of a difference anyway after ~6 or 7 ages (unless you collect hundreds, which is not feasible). Even if *T* happens to be very large, collecting 10 ages rather than 7 still probably won't be enough to offset the effect of *T*.

We think it's a really interesting idea that climatic periodicity might be driving the formation of fan surfaces, and therefore might provide a guide for the size of *T*. Reviewer 2 is right to point out the risk of circularity, which is why we don't speculate about these questions in this paper, but we are certainly planning to explore this topic in future papers containing data from real alluvial fan systems. We hope our work here inspires other groups to date more fans, because to tackle this question we really need more field examples with well-dated fan surfaces.

Regarding how many samples to collect, our view (based mostly on Fig. 4 and Fig. 6) is that going up to 6 ages will always provide benefits whatever the age/duration of the surface. Collecting more than 6 ages will give smaller and smaller returns, so while it might be useful when especially precise constraints on abandonment are required (e.g., comparing with millennial-scale climate events), it would probably be better to spend those additional resources dating a different surface. The case study from the Baja California fans illustrates this well – the Q4 surface is only dated with 5 ages, but that's still enough to show fairly convincingly that abandonment overlaps with the Younger Dryas (which only lasted ~1 kyr, so is as short as most palaeoclimate events come). Collecting another 5 ages from that surface would narrow the probability distribution a little bit, but it would probably be a better use of time and money to use those 5 samples to date something else.

One caveat here is that an old outlier was discarded from the Q4 dataset (attributed to nuclide inheritance). So if the goal in the field is to measure ~6 'good' ages, then gathering 1 or 2 extra samples might still be useful in case there are some outliers that need to be thrown out. Given unlimited resources, our strategy would be to process 6 or 7 samples per fan surface, but collect another 1 or 2 samples to keep in reserve. Even if 1 or 2 of the ages turned out to be outliers then the dataset would still probably be fine for inferring abandonment timing (for most purposes). If 3+ outliers turned up, or the project required very precise estimates of abandonment age, then you could go back and process the backups.

This is a bit tangential, but if you're going out in the field to sample fans, it's worth taking some free Landsat-8 imagery with you to help choose your sampling sites. We published a paper in *Remote Sensing of Environment* (2018) titled "Alluvial fan surface ages recorded by Landsat-8 imagery in Owens Valley, California", where we talk about these opportunities. Landsat-8 imagery is a really powerful resource when sampling fans and can make a big difference to ensuring you collect samples from the right patches of the surfaces, and ultimately get robust datasets.

P9. L25 – There is no Fig. 5D.

This was a typo, we have corrected it to Fig. 5b. Thanks for spotting it!

P11. L9 – Should this be Figure 8D?

Corrected.

P12. L12 – If displacement can only occur after surface abandonment, do you have any constraint on a minimum age of displacement onset? Could this estimated time-averaged slip still only be a minimum rate? If so, perhaps state somewhere.

We assume that a faulted surface would start to accumulate a displacement as soon as it's abandoned (i.e., as soon as it stops being actively resurfaced). That perspective in turn assumes that a fault is continuously slipping, or at least that the time interval between slip events is insignificantly small compared to the age of surface abandonment. It might be that in some cases there is an additional lag time, which would make the time-averaged slip rate a minimum estimate as Reviewer 2 suggests, even when calculated using the abandonment age instead of the average surface age. However we really don't know whether there is likely to be a lag time in a significant number of cases, so we prefer not to make a general suggestion that time-averaged rates will be minimum estimates. Applied studies will probably need to evaluate this possibility on a case-by-case basis.

P12. L 28-30 This is a really good point – I also feel that this shouldn't just be in a limitations section! Deriving an average surface age would certainly be biased by burial of older material but by focusing on the timing of abandonment this bias is removed. Perhaps bring this up earlier in the manuscript as an additional strength of this method.

Thank you! We prefer to avoid repetition in different parts of the manuscript, but we agree that this is a valuable point so we have emphasised it more clearly in the text at p.14, l.4-12.

Additional Edits

We decided to add a list of mathematical notation (now section 7). This doesn't add much length, but we think it will make it easier for readers to get to grips with our equations.

We have made some very small text edits throughout to improve the wording and clarity of a few sentences. These edits are all shown with tracked changes.

We renamed sub-section 4.3 from '*Application to **real** surface ages*' to '*Application to **measured** surface ages*', related to the comments by Reviewer 2 above. We also added a small amount of clarification to this section, just to make sure the text is very clear about how readers could go about applying our approach to their own data (which is ultimately our goal!). For example, here we explain that "*discrete values of τ can be converted into a probability distribution by calculating the density of P within fixed increments of τ*". We decided that it would be helpful to briefly expand on this point with two additional sentences that clarify what we mean and also explain how the Matlab script can be used to implement our approach. The section is still concise, but we think it will now be easier for readers to follow Fig. 7 and reproduce our approach with their own data.

Related to Fig. 7, we decided to take out the equations because it isn't necessary to reproduce them in the figure and was a waste of space. Figure 7 is now a single-column figure.

We added a citation to Terrizzano et al. (2017) at p.12, l.17.

Inferring the timing of abandonment of aggraded alluvial surfaces dated with cosmogenic nuclides

Mitch K. D'Arcy[1,2], Taylor F. Schildgen[2,1], Jens M. Turowski[2], Pedro DiNezio[3]

[1]Institute of Geosciences, University of Potsdam, Karl-Liebknecht-Strasse 24/25, 14476 Potsdam, Germany
[2]GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany
[3]Institute for Geophysics, Jackson School of Geosciences, University of Texas at Austin, J.J. Pickle Research Campus, Building 196 10100 Burnet Road (R2200), Austin, TX 78758, USA

Correspondence to: Mitch K. D'Arcy (mdarcy@uni-potsdam.de)

Abstract. Information about past climate, tectonics, and landscape evolution is often obtained by dating geomorphic surfaces comprising deposited or aggraded material, e.g., fluvial fill terraces, alluvial fans, volcanic flows, or glacial till. Although surface ages can provide valuable information about these landforms, they can only constrain the period of active deposition of surface material, which may span a significant period of time in the case of alluvial landforms. In contrast, surface abandonment often occurs abruptly and coincides with important events like drainage reorganisation, climate change, or landscape uplift. However, abandonment cannot be directly dated because it represents a cessation in the deposition of dateable material. In this study, we present a new approach to inferring when a surface was likely abandoned using exposure ages derived from in situ-produced cosmogenic nuclides. We use artificial data to measure the discrepancy between the youngest age randomly obtained~~sampled~~ from a surface and the true timing of surface abandonment. Our analyses simulate surface dating scenarios with variable durations of surface formation and variable numbers of~~ sample~~ exposure ages from sampled boulders. From our artificial data, we derive a set of probabilistic equations and a Matlab tool that can be applied to a set of real sampled surface ages to estimate the probable period of time within which abandonment is likely to have occurred. Our new approach to constraining surface abandonment has applications for geomorphological studies that relate surface ages to tectonic deformation, past climate, or the rates of surface processes.

1 Introduction

Geomorphological studies that link the formation of landforms to past changes in climate or tectonic deformation depend on the accurate dating of surfaces comprising aggraded or deposited material. Surfaces commonly targeted for dating include alluvial fans, fluvial fill terraces, glacial till, pediments, and volcanic flows, among others. For example, the ages of fluvial fill terraces and alluvial-fan surfaces have been used~~are widely dated in order~~ to (i) decipher how erosion and sedimentation have responded to past hydroclimate changes (Owen et al., 2014; Schildgen et al., 2016; Tofelde et al., 2017); (ii) derive time-integrated slip rates for active faults (e.g., Frankel et al., 2007, 2011; Gosse, 2011; Hughes et al., 2018); and (iii) quantify the rates of surface processes such as weathering, landform erosion, or channel avulsion and incision (Schildgen et al., 2012; Regmi et al., 2014; Bufe et al., 2017; D'Arcy et al., 2018).

A common assumption is that a geomorphic surface can be represented by a single formation age. Surfaces are usually point-sampled and dated in multiple locations, e.g., by cosmogenic nuclide exposure dating of surface boulders. Typically, sampling is~~a~~ limited to a small number of (often fewer than 10) large, stable surface boulders ~~are sampled for exposure dating~~, which exhibit no evidence of weathering, rotation, or disturbance. From the set of exposure ages obtained, an average surface age can be~~is~~ calculated with an uncertainty that reflects both analytical uncertainty and the spread of sampled ages. However, many geomorphic surfaces are active for an extended period of time, during which material is continually deposited until the surface is abandoned (e.g., Savi et al., 2016; Denn et al., 2017; Foster et al., 2017). Alluvial-fan surfaces provide one example. Rather than being formed instantaneously, fan surfaces are typically active for thousands or tens of thousands of years before being abandoned when the channel avulses or incises (e.g., Dühnforth et al., 2007). This prolonged period of activity results in a meaningful spread in ages collected from a single surface (e.g.,~~,~~ Owen et al., 2011). For any geomorphic surface with a non-negligible period of formation, a set of surface ages will capture a portion of the full timespan over which that surface was active. An average of those ages will sit somewhere within the true timespan of surface deposition, whereas~~but will overlook information such as~~ the maximum age~~, which~~ might approximate the onset of surface activity, and~~or~~ the minimum age~~, which~~ might approximate the timing of surface abandonment.

In some cases, the timing of surface abandonment may be a more useful constraint than an average surface age. In contrast to surface deposition, abandonment occurs at a particular moment in time (e.g., coinciding with a switch to incision) and so can, in principle, be defined with greater precision. For surfaces with an extended period of formation, the timing of abandonment is more likely to coincide with events of interest such as reorganisation of a drainage network (Bufe et al., 2017); changes in climate, sediment supply, or base level (Steffen et al., 2009; Tofelde et al., 2017; Mouslopoulou et al., 2017; Brooke et al., 2018); or tectonic deformation such as faulting, uplift, or subsidence (e.g., Frankel et al., 2007, 2011; Ganev et al., 2010). Abandonment ages would also benefit any study that uses surface exposure dating to measure the rates of ~~a~~ post-depositional processes, such as in situ weathering (e.g., White et al., 1996, 2005; D'Arcy et al., 2015, 2018), the topographic decay of landforms (e.g., Hanks et al., 1984; Andrews & Bucknam, 1987; Spelz et al., 2008), or channel avulsion and incision (e.g., Schildgen et al., 2012; Finnegan et al., 2014; Malatesta et al., 2017). Yet the abandonment of a surface represents a cessation in the deposition of dateable material, and therefore cannot be directly dated. Instead, the timing of abandonment must be inferred. Some studies make assumptions about when geomorphic surfaces were abandoned based on independent information such as palaeoclimate records (e.g., Macklin et al., 2002; Cesta and Ward, 2016); others assume that the youngest sampled surface ages fall close to the timing of surface abandonment (e.g., Sarıkaya et al., 2015; Foster et al., 2017; Ratnayaka et al. 2018; Clow et al., 2019). These ~~approaches~~ assumptions risk interpretations that are circular (in the former case) or potentially inaccurate (in the latter case)~~interpretations~~, highlighting the need for a robust method to quantitatively infer the timing of surface abandonment from a set of sampled surface ages.

Here, we introduce a new probabilistic approach to constraining when a depositional surface was abandoned, based on what is known about its activity. We use artificial data to randomly point-sample the ages of virtual surfaces, in scenarios that are representative of studies dating natural geomorphic landforms such as alluvial fans. We quantify how close the youngest obtained age is likely to fall to the true timing of abandonment, depending on the overall period of surface activity and the number of samples collected. From these artificial data, we derive a set of probabilistic equations and a Matlab tool that can be applied to real geomorphic surfaces to estimate when they were abandoned. Finally,~~, and~~ we demonstrate the application of these equations and the Matlab tool to natural surfaces with a case study of dated alluvial-fan surfaces in Baja California, Mexico.~~.~~

2 Justification

Here, we present a hypothetical example of a dated alluvial-fan surface to illustrate why the timing of abandonment may, in some cases, be more useful than an average of sampled surface ages.

Consider an alluvial-fan surface that was active for a 30 kyr timespan, starting at 80 ka and ending at 50 ka, when the surface was abandoned due to fan incision (Fig. 1). In this example, deposition occurred on the fan surface ~~was deposited~~ throughout a period of climatic stability and abandoned when the climate changed, and we make the assumption that there is an equal likelihood of obtaining any age within the entire period of deposition. A distribution of surface ages can be obtained by point-sampling the fan surface; an approach analogous to ~~studies using cosmogenic nuclides to measure~~studies that measure the exposure ages of boulders atop landforms. We present two arbitrarily-selected possible outcomes in Fig. 1, where 6 surface ages are obtained. In scenario 1, the ages are distributed relatively evenly through time, producing a mean age of 65.8 ka, which~~that~~ closely approximates the true average surface age of 65 ka, with~~and~~ a standard deviation of 10.5 kyr. In scenario 2, the ages obtained are unevenly distributed through time, producing a slightly older mean surface age (71.4 ka) and a smaller standard deviation (5.2 kyr). These scenarios are plotted against time in Fig. 1b as data points and kernel density plots, and they resemble equivalent natural datasets (e.g., Owen et al., 2014).

Sample set 2 is more tightly clustered than sample set 1, despite being less representative of the average surface age, illustrating that greater clustering of ages is not necessarily an indicator of accuracy. Furthermore, neither average age corresponds to any meaningful event. The fan surface was equally active for the entire period between 80 and 50 ka, the average ages sit within a period of climatic and depositional stability, and the peaks in the kernel density plots are artefacts created by randomly sampling a linear series.

In contrast, the abandonment of the fan surface does occur at a precise moment in time when deposition ends at 50 ka. In this example, abandonment coincides with an abrupt change in climate that triggered an incision event (cf., Simpson and Castelltort, 2012), so it is arguably a more informative target for dating than an average age that imprecisely approximates the mid-point in the duration of surface deposition. However, the abandonment of the surface represents a cessation in the deposition of dateable material, so its timing instead must be inferred from what is known about the surface activity. Given that (i) the sampled ages constrain the timespan over which the surface was formed, and (ii) abandonment occurred sometime after the youngest age, it could be assumed that the youngest sampled age best approximates abandonment. In scenario 1, the youngest age falls within ~1 kyr of surface abandonment, which would enable a correct interpretation of correlation between fan incision and the climate change event. In scenario 2, however, there is a ~14 kyr discrepancy between the youngest sampled age and the timing of surface abandonment, which would probably fail to demonstrate the correlation between climate change and fan incision. Therefore, the question becomes: how close is the youngest age obtained from a surface to the actual timing of surface abandonment?

This question cannot currently be answered for a natural dataset, yet the ability to reliably estimate when a surface was abandoned has important implications for many geomorphological ~~studies~~ applications (see section 1). In this study, we use artificial data to simulate natural surfaces that undergo active deposition~~are deposited~~ over variable periods of time and dated with ~~by sampling~~ a limited number of surface ages. These artificial data are analogous to studies that date natural geomorphic surfaces, for example, with cosmogenic nuclide exposure ages obtained from sampled boulders ~~collected~~on ~~from~~ alluvial fans or fluvial fill terraces. However, unlike field-based datasets, artificial data ~~uniquely~~ enable us to constrain the ~~likely~~ time difference between the youngest age obtained from a geomorphic surface and the true timing of surface abandonment, which is generally unknowable for natural surfaces.~~.~~ There are several additional advantages to taking an artificial-data approach. First, we can repeat the random sampling of surface ages (e.g., as depicted in Fig. 1) many~~a large number of~~ times to probabilistically determine where the youngest sampled age tends to fall with respect to abandonment. Second, we can prescribe the surface parameters, meaning the exact timing of abandonment and the full period of surface activity are known. Third, we can select surface properties that are representative of natural~~real~~ geomorphic surfaces and numbers of samples commonly obtained in geomorphic studies. Fourth, we can perform a thorough quantification of the uncertainties in our analyses. For the above reasons, the artificial-data approach allows us to derive a set of equations and develop a Matlab tool that can then be applied to natural datasets (a set of surface ages) to determine the probability of surface abandonment occurring within a specified window of time.

3 Methods

3.1 Artificial-data approach

We used artificial data to constrain the temporal discrepancy, which we denote τ, between the youngest age sampled on a surface (a_{min}) and the actual timing of surface abandonment (t_{aban}):

$$\tau = a_{min} - t_{aban} \qquad (1)$$

Our experiments are designed to be representative of natural alluvial-fan surfaces, but the results are more widely applicable to any abandoned depositional surface that has been ~~subsequently~~ dated.

In the absence of additional information (e.g., the existence of an ~~additional~~ independent constraint, such as a younger alluvial-fan surface with an intermediate age), the abandonment of a surface could have occurred at any time between the youngest sampled age, a_{min}, and the present, or within a particular time window after a_{min}. In the example case (Fig. 1), the data in sample set 1 would require a time window, τ, of 1.1 kyr~~, (and 14.4 kyr for sample set 2),~~ placed immediately after the ~~the~~ youngest ~~sampled~~ age~~s,~~ to overlap with the correct timing of surface abandonment, t_{aban}. However ~~(f~~for sample set 2, a 14.4 kyr window is required~~)~~. For natural cases, the abandonment timing is unknown; we know the temporal discrepancy between a_{min} and t_{aban} in these artificial-data examples because we impose t_{aban}. ~~; for real-world cases, this information is unknown. As such, artificial data provide the unique advantage that~~An advantage of the artificial-data approach is that knowledge of t_{aban} ~~is known for a given surface and can thus be compared against a sampled set of surface ages~~allows us ~~in order~~ to quantify ~~$\tau$the time difference between t_{aban} and a_{min}~~ in every tested scenario, which in turn enables us to determine probability distributions of τ. ~~In principle~~Conceptually, the size or τ, i.e., the proximity of a_{min} to t_{aban}~~, will~~ depends on the number of surface ages obtained, *n*. The greater the number of samples, the closer the youngest sampled age is likely to come to the abandonment age (Fig. 2a). The size of τ also depends on the total duration of surface activity, which we denote as *T*. If *n* ages are randomly-sampled from a longer time span, *T*, then a_{min} is likely to fall farther from t_{aban} (Fig. 2b).

Our artificial-data experiments simulate surfaces with durations~~a length of the period~~ of activity, *T*, between 10 and 50 kyr, sampled with numbers of surface ages, *n*, between 2 and 10. These values are representative of natural alluvial-fan surfaces and typical dating studies involving a small number of ages. For each combination of *T* and *n*, we randomly sampled a set of surface ages 10,000 times, allowing us to ~~reliably~~ constrain the probability that a_{min} falls within a certain temporal window (τ)~~distance~~ of t_{aban} in each scenario.

3.2 Implementation

We first implemented our experiments using discrete sampling within a spreadsheet. For each surface, we created a list of selectable surface ages spanning the total period of surface activity, *T*, and placed at equal intervals of 0.1 kyr. For the example case (Fig. 1), this would mean a list of selectable ages of 80.0 ka, 79.9 ka, 79.8 ka, etc., to a minimum value of 50.0 ka. We chose periods of surface activity, *T*, equal to 10, 20, 30, 40, and 50 kyr. From each list, we randomly selected *n* unique values, and repeated this exercise 10,000 times for each integer value of *n* between 2 and 10. For example, if *n* = 6 and *T* = 20 kyr, then we extracted 10,000 different datasets, each comprising 6 randomly-selected~~ampled surface~~ ages, from the 20 kyr-long list of selectable ages available at 0.1 kyr intervals. This process is analogous to random sampling of 6 boulders for cosmogenic nuclide exposure dating~~ages, e.g., from surface boulders,~~ on an alluvial fan surface that formed over a 20 kyr period and deposited a 'selectable' boulder once every 100 years. We extracted 10,000 sets of ~~surface~~ ages for each of the 45 different combinations of *T* (5 unique values) and *n* (9 unique values). For each dataset, we calculated the mean value of the selected~~ampled~~ ages, $\bar{a}$, and the time difference between the youngest age and the abandonment time, τ. We then determined~~extract~~ cumulative frequency distributions of τ in each scenario with a given combination of *T* and *n*.

To test whether 10,000 iterations are sufficient to produce reliable statistics and whether the discretization of ages has an important effect, we repeated all our artificial-data experiments using a non-discrete approach in a Matlab script. We defined *T* as a time range, from within which any point in time could be randomly sampled, i.e., ~~an excess number~~tens of thousands of 'selectable' surface ages were available rather than ~~a list of~~ hundreds of discrete values. Performing 100,000 iterations with the Matlab script produced ~~identical~~ results that are indistinguishable from ~~to~~ the discrete spreadsheet-based approach with 10,000 iterations. All data analyses are provided by D’Arcy et al. (2019) in an online data repository. Finally, we explore the assumptions and limitations of our analyses in section 5.3.

3.3 Experimental assumptions

In designing our artificial-data experiments, we make several assumptions. First, surface ages are randomly selected from the total period of surface activity. Therefore, when constructing our experiments, we assume that when ages are obtained~~collected~~ from real geomorphic surfaces, they are randomly point-sampling the full timespan of surface formation, and that this timespan represents a uniform probability distribution of selectable ages. A~~This~~ uniform probability of ages may not ~~always~~ be realistic~~the case~~ in certain natural cases, for example, if boulders on an alluvial-fan surface are spatially clustered by age and all samples are taken from one part of the surface. Second, the entire period of surface activity is assumed to be available for sampling, i.e., no subset of the surface history is missing as a result of processes like burial or erosion. Third, all selectable ages within the period of surface activity have an equal likelihood of being sampled; this implies that the surface formed with a constant deposition rate and there are no pulses of activity that increase the probability of sampling a particular age. Finally, we do not explicitly factor in processes like nuclide inheritance, erosion, or incomplete exposure, which can affect exposure ages derived from cosmogenic nuclides. We consider the implications of all these assumptions for ~~real~~ natural datasets in section 5.3.

4 Results

4.1 Random sampling of surface ages

To illustrate the results of our experiments, we first present one artificial-data~~example~~ scenario in Fig. 3, in which the surface is formed between 80 ka and 50 ka (i.e., T = 30 kyr) and is randomly sampled with n = 2, 4, 6, or 8 ages (with 10,000 repeat experiments for each value of n). Figure 3a shows how a frequency distribution of the mean value of all sampled ages, $\bar{a}$, changes with n. The distribution is centred on the true average surface age of 65 ka and narrows as a greater number of ages are sampled. If only 2 ages are sampled, then $\bar{a}$ can occupy almost any age within the full period of surface activity.~~,~~ A~~and a~~s n increases, $\bar{a}$ tends to fall closer to 65 ka. The distribution of $\bar{a}$ approaches a normal distribution as n increases. This observation is compatible with the central limit theorem and the law of large numbers:~~, and~~ $\bar{a}$ converges on the true average surface age as the number of samples increases, despite the dataset randomly sampling a linear series.

A frequency distribution can also be plotted for the youngest age, a_{min}, randomly selected~~ampled~~ from the surface (Fig. 3b). If only 2 ages are obtained~~collected~~, then the youngest can fall almost anywhere between 50 and 80 ka, although the distribution is asymmetric and younger values of a_{min} occur slightly more frequently than older values. As n increases, the distribution of

a_{min} shifts towards 50 ka such that when n = 8, a_{min} falls within 5-10 kyr of t_{aban} (i.e., τ is equal to 5-10 kyr) in the majority of sampling experiments. ~~As t_{aban} is known in our experiments (50 ka), τ can be calculated for each set of ages sampled.~~ Cumulative frequency distributions of τ reveal how close the youngest sampled age comes to the known timing of surface abandonment (Fig. 3c). For example, if only 2 ages are obtained, then in 60% of experiments, $\tau \leq 12$ kyr, i.e., the youngest age falls somewhere within 12 kyr of abandonment. If 6 ages are obtained, then in 90% of experiments, $\tau \leq 10$ kyr. Any percentile of τ can be measured from Fig. 3c, allowing τ to be plotted against n (Fig. 3d). As a greater number of ages are obtained~~sampled~~, the value of τ associated with a given percentile decreases, i.e., the youngest sampled age comes closer to the timing of surface abandonment~~ as the number of samples increases~~. However, the decrease in τ is non-linear and diminishes with increasing n. For example, as n increases from 2 to 4 ages, the 95th percentile of τ falls from ~23 kyr to ~16 kyr, but collecting another 2 ages (n = 6) only reduces τ to ~12 kyr. The 95th percentile of τ falls below 10 kyr when n exceeds 7 ages. In other words, if 7 ages are randomly-sampled from a surface, abandonment will have occurred within 10 kyr after the youngest age in 95% of cases.

We equate the percentiles of τ in Fig. 3c with the probability, P, of abandonment occurring within a time window defined by τ. Thus, if P = 0.9, the window of time τ (placed immediately after a_{min}) is large enough that in 90% of our experiments, the true timing of surface abandonment would fall within it. This is equal to the 90th percentile of τ, which would be 7.5 kyr for the scenario T = 30 kyr and n = 8, for example (Fig. 3d). Note that in this scenario, τ does not imply that the surface was abandoned exactly 7.5 kyr after a_{min}, but rather that there is a 90% likelihood that abandonment occurred anywhere within a 7.5 kyr window after a_{min}. The probable window of abandonment, τ, increases with P because a larger window of time is required to capture the true timing of abandonment in a greater proportion of cases.

At the same time, τ is inversely and non-linearly related to the sample size of ages obtained, n (Fig. 4). The dependencies between τ and n, T, and P are illustrated in Fig. 4 for all tested scenarios that are representative of natural alluvial fan surfaces (n = 2 to 10; T = 10 to 50 kyr), with probabilities between 0.50 and 0.95. For example, if 6 ages are obtained from a surface that formed over a 30 kyr duration, then τ = 12 kyr for P = 0.95 (Fig. 4a). In other words, in 95% of cases the youngest of 6 ages obtained from such a surface will fall within 12 kyr of the true timing of abandonment. Only in 5% of cases does the discrepancy between a_{min} and abandonment exceed 12 kyr. If P decreases to 0.5 (Fig. 4f) then τ decreases to 3 kyr for this particular scenario (n = 6 and T = 30 kyr).

The results of our artificial-data experiments (Fig. 4) can be described by one equation that allows τ to be calculated for any scenario:

$$\tau = \tau_0 + PTe^{kn}$$ (2~~1~~)

Here, the parameter k is a decay constant with a negative value that increases exponentially with P (Fig. 5a):

$$k = a + be^{cP}$$ (3~~2~~)

Constants a, b, and c can be derived empirically using our artificial data. Note that we calibrate all our equations with time in kyr.

a = -0.425 ± 0.029 b = 0.011 ± 0.011 c = 2.830 ± 0.885

The parameter τ_0 increases linearly with T, but with a slope that increases exponentially with P (Fig. 5b), and can therefore be described by a pair of relationships:

$$\tau_0 = mT \quad (4\cancel{3})$$

$$m = m_0 + ge^{hP} \quad (5\cancel{4})$$

Parameters m_0, g, and h are constants with values again determined empirically from our artificial-~~-~~data experiments:

m_0 = 0.019 ± 0.008 g = 0.005 ± 0.002 h = 3.784 ± 0.406

Given that τ_0 signifies the value of τ as n trends towards infinity, it represents the ~~most precise possible constraint on~~limit of precision that can be obtained on the abandonment period, τ~~.~~, ~~-~~when inferring the timing of surface abandonment with this probabilistic method. For the scenarios shown in Fig. 5b, which represent reasonable values of T for natural alluvial fans and desirable values of P, τ_0 varies from a few centuries to ~10 kyr. ~~These τ_0 values illustrate the limits to precision when inferring the timing of surface abandonment in this probabilistic way.~~

4.2 Total period of surface formation

Equation~~s~~ ~~1~~2 can be solved for τ (using the parameterization of Eqs. 3 through 5) with knowledge of only the number of ages sampled, n, and the total period of surface formation, T, as well as a chosen probability, P. We are able to parameterise ~~equations 1~~2 ~~through 5~~τ using artificial data because we know the value of T in our experiments. However, when sampling natural~~real~~ depositional surfaces, T is unknown and instead only the span of sampled ages, $a_{max} - a_{min}$, can be measured. This span might approximate T, but some fraction of time will ~~remain unsampled~~not be captured. Fortunately, ~~To resolve~~our artificial-data experiments also allow us to determine ~~this conundrum, our artificial data experiments also determine~~ what~~the~~ fraction of T ~~that~~ is captured by $a_{max} - a_{min}$ in scenarios of varying n (Fig. 6).

The artificial data indicate that, for example, 6 randomly-distributed ages will span ~70% of the total timespan of surface activity, T, in the average case. In the 1% of 'worst' (most clustered ages) cases, $a_{max} - a_{min}$ will only represent ~30% of T, and in the 1% of 'best' (least clustered ages) cases it will represent more than 95% of T. In half of all experiments for n = 6 (from P25 to P75), $a_{max} - a_{min}$ falls within 60-85% of T. There is a diminishing improvement with an increasing number of sampled ages, such that by n = 10, the average span of ages has only increased to ~80% of T and 50% of all experiments fall between 75-90% of T. An order of magnitude more ages (hundreds) would be needed for the mean $a_{max} - a_{min}$ to come within 95% of the full period of surface activity.

A regression can be fitted to the distributions in Fig. 6, taking the form:

$$\frac{a_{max} - a_{min}}{T} = q + re^{sn} \quad (6\cancel{5})$$

Parameters q, r, and s are empirical coefficients derived graphically from our artificial data. For the mean case (the solid black line in Fig. 6), they take the values:

q_{av} = 0.838 ± 0.007 r_{av} = -1.035 ± 0.030 s_{av} = -0.366 ± 0.016

Equation ~~5~~6 is also fitted to ±1 standard deviation (σ) above and below the mean values in Fig. 6 (dashed black lines). For 1σ above the mean, parameters *q*, *r*, and *s* take the values:

$q_{+1\sigma}$ = 0.928 ± 0.005 $r_{+1\sigma}$ = -0.983 ± 0.055 $s_{+1\sigma}$ = -0.512 ± 0.027

For 1σ below the mean, parameters *q*, *r*, and *s* take the values:

$q_{-1\sigma}$ = 0.764 ± 0.007 $r_{-1\sigma}$ = -1.196 ± 0.015 $s_{-1\sigma}$ = -0.296 ± 0.008

Equation 6~~5~~ can therefore be used to estimate the size of *T* in the average case plus ±1σ bounds, given the measured span of ages collected from a surface. Equations 2 through 6 are thus calibrated using our artificial data, and can be used to probabilistically calculate the window of time during which any dated surface was likely abandoned.

~~Equations 12 65 are thus calibrated using our artificial data, and can now be used to probabilistically calculate a window of time during which any dated surface was likely abandoned.~~

4.3 Application to ~~real~~ measured surface ages

Given that Eqs~~.~~ 2~~1~~ through 6 are~~is~~ probabilistic (i.e., *P* is a variable), ~~it~~ our artificial-data approach can be used to infer a probability distribution of abandonment ages from a set of measured surface ages. We illustrate the steps involved in applying ~~Eq. 1~~Eqs 2 through 6 and the Matlab script to real data in Fig. 7.

To solve for τ at a discrete probability value (*P*), *T* is first calculated with Eq. 6 and τ is then calculated using Eq. ~~equations~~ 2, ~~1~~(with parameters defined in Eqs 3 through 5) ~~65~~ ~~with discrete values of *P*~~ (Fig. 7a)~~, resulting in discrete windows of time in which abandonment is likely to have occurred with different probabilities~~. These discrete values of τ can be converted into a probability distribution by calculating the density of *P* within fixed increments of τ. For example, in Fig. 7a 50% of the probability of abandonment falls within a relatively small window of time (the light blue bar for *P* = 0.5), whereas a longer window of time is required to contain an additional 45% of the probability of abandonment (the light pink bar for *P* = 0.95). Thus, the density of *P* is greater within the window τ for *P* = 0.5, and this density diminishes as τ and *P* increase. The Matlab script (provided as supplementary information) enables determination of the ~~full~~continuous probability distribution of τ. After generating artificial data based on *n* ages and a duration of deposition *T* (from Eq. 6), the script calculates ~~These values of τ can be converted into a probability distribution (Fig. 7b) by calculating~~ the density of *P* within fixed increments of τ. If the sampled surface ages are known with exact precision, then the resulting distribution of τ values provides a probability distribution of times that ~~would~~ directly postdate the youngest age and yield a probability distribution of surface-abandonment ages (Fig. 7a).

However, real surface ages have associated uncertainties that must also be incorporated into the estimated abandonment ages (Fig. 7b~~c~~). The Matlab tool is designed to incorporate this uncertainty, and is explained in the following steps. F~~F~~irst, we use ±3σ uncertainty on a_{min} to characterise the probability distribution of potential a_{min} values. In the example schematic (Fig. 7~~c~~b) we assumed a normal distribution, as is typical for exposure ages of individual boulders, but alternative distributions could be used. This distribution of a_{min} values is then discretised, and separate probability distributions of τ are calculated for each potential value of a_{min}, i.e., repeating ~~steps illustrated in Fig. 7a-b~~Fig. 7a. The resulting, temporally shifted probability distributions of τ are weighted according to the probability distribution of a_{min} and summed, ~~resulting~~ producing in an overall probability distribution of likely abandonment ages that accounts for uncertainty on the youngest age (Fig. 7c).

If the ~~1σ~~ uncertainty on a_{min} is small compared to τ calculated using Eq. 2~~1~~, then incorporating age uncertainty will have little impact on the resulting probability distribution of abandonment ages. If the ~~1σ~~ uncertainty on a_{min} is large, it will have a greater influence on the final probability distribution of abandonment ages. ~~In the supporting information, we provide a Matlab script that can be used to input a set of measured surface ages and output a probability distribution of abandonment timings following the steps outlined in Fig. 7.~~

5 Discussion

5.1 Implications for surface dating

Our artificial data provide new information about what measured ages represent when collected from aggraded surfaces that formed over non-negligible timespans. Crucially, our findings indicate that averages of sampled surface ages are likely to be imprecise representations of the mid-point of surface formation, ~~and~~ which may not ~~correlate~~ coincide with ~~any~~ a particular external forcing event (Fig. 1). In contrast, surface abandonment typically occurs at a discrete moment in time and is more likely to coincide with external forcing events such as changes in climate or tectonics. By using artificial data, we have derived a set of probabilistic equations for inferring when a surface was likely to have been abandoned, based on a distribution of randomly-sampled surface ages. These equations can complement and enhance interpretations based on any dataset comprising surface ages. The spreadsheet and the Matlab tool allow for quantification of the full probability distribution of τ, and the Matlab tool additionally allows for the incorporation of uncertainty on the youngest age, a_{min}~~.~~

While a distribution of ages is required for dating surfaces that have formed over extended periods of time, our analyses reveal that an increasing number of ages yields diminishing returns ~~from sampling an increasing number of ages; these diminishing returns apply to constraints on~~for constraining the timing of abandonment (Figs 3d and 4) and the total duration of surface activity (Fig. 6). An appropriate number of surface ages will depend on the desired precision, but our results indicate that there is little to be gained by obtaining more than 6 to 7 ages per surface (Figs. 3, 4, and 6), assuming no outliers, for the purposes of most geomorphological studies. T~~Indeed, t~~o obtain ~~significantly~~ substantially more information about a surface, an order of magnitude more ages would be required. As explained in section 4.1, τ_0 represents the maximum precision with which the abandonment age can, in principle, be inferred. For many natural surfaces, τ_0 can range from a few centuries to ~10 kyr (Fig. 5b~~d~~), depending on the period of surface activity and the desired probability. Our methodology thus provides a new way to~~of~~ quantify~~ing~~ ~~the limits to the precision with which~~how precisely a distribution of surface ages can be interpreted. Thi~~s~~s~~e~~ limit to precision~~s~~ should be considered~~are~~ in addition to the age uncertainty associated with cosmogenic nuclide exposure dating;~~,~~ ~~and~~ both are ~~an~~ important considerations when inverting landforms and sedimentary deposits for palaeo-environmental information (Foreman and Straub, 2017).

When sampling in the field, it may ~~sometimes~~ be advantageous to target different parts of an aggraded surface ~~in order~~ to capture as much of its period of activity as possible. This strategy applies to surfaces upon which the locus of deposition has systematically migrated during deposition. For example, if channel migration on an alluvial-fan surface resulted in ~~particular fragments~~a portion of its overall history being recorded in particular parts of the surface (e.g., Savi et al., 2016; Schürch et al., 2016; D'Arcy et al., 2017a,b), then greater spatial coverage would capture a greater range of ages. However, if each deposition event followed a random trajectory on the surface, resulting in all potentially 'selectable' ages being spatially mixed, then it would be unnecessary to distribute sampling locations across the surface.

5.2 Case study: Alluvial fans in the Laguna Salada Basin, Mexico

Here, we use a case study of alluvial-fan surfaces in the Laguna Salada Basin, Mexico, to demonstrate how our findings can be applied to real surfaces to gain new information about when they were abandoned.

The Laguna Salada Basin is a half-graben in northern Baja California, Mexico. This basin contains well-preserved alluvial fans eroded from the neighbouring Sierra El Mayor and Sierra Cucapa, with at least 8 generations of distinct fan surfaces formed by a sequence of aggradation and incision cycles. The ages of two of these fan surfaces—mapped as Q4 and Q7—were estimated by Spelz et al. (2008) using ^{10}Be exposure ages of stable surface boulders with no evidence of erosion or disturbance (Fig. 8). We used the CREp calculator (Martin et al., 2017) to update the exposure age estimates of Spelz et al. (2008) using the time-corrected Lal/Stone scaling scheme (Lal, 1991; Stone, 2000), the ERA40 atmosphere model (Uppala et al., 2005), the atmospheric ^{10}Be-based VDM geomagnetic database of Muscheler et al. (2005) and Valet et al. (2005), and the current global reference SLHL ^{10}Be production rate of 4.13 ±0.20 at g^{-1} yr^{-1} in the ICE-D database (Martin et al., 2017). We assume a sample density of 2.7 g cm^{-3} and no boulder erosion. The oldest age measured on the Q4 surface was excluded as an outlier by Spelz et al. (2008), and we maintain this interpretation. The remaining exposure ages span from 14.4 ±1.1 ka to 32.1 ±2.9 ka for Q4 (n = 5), and 188.6 ±22.7 ka to 246.9 ±13.7 ka for Q7 (n = 6) (Fig. 8b, yellow bars). On both fan surfaces, the dispersion of ages greatly exceeds the age uncertainty, suggesting~~confirming~~ that each surface was deposited over an extended period of time.

For both distributions of fan surface ages, we used equations ~~1~~2 through~~o~~ 6~~5~~ to calculate probable abandonment windows, τ, for different values of P. For example, on the Q4 fan surface with an a_{min} of 14.4 ±1.1 ka, τ = 3.3 kyr when P = 0.5, suggesting a 50% probability that the surface was abandoned within 3.3 kyr after a_{min}, i.e., between 14.4 ka and 11.1 ka. The size of τ increases with P, as explained in section 4.1, such that τ = 12.0 kyr for the Q4 surface when P = 0.95, i.e., the abandonment window ranges from 14.4 ka to 2.4 ka. The full ~~If Eq. 21 is applied to the data without accounting for age uncertainty, the resulting~~ probability distribution of τ is highly asymmetric (Fig. 8c, red dashed lines). Of course, the uncertainty on a_{min} must also be accounted for~~.,~~. To do so, we used ~~so following the approach outlined in Fig. 7~~the Matlab script with the required inputs—~~—~~T, n, the desired number of iterations, a_{min}, and the 1σ uncertainty on a_{min}—~~—~~to~~,~~ ~~we~~ derive a continuous probability distribution of τ for ~~the~~ each ~~Q4 and Q7~~ fan surface that incorporates age uncertainty ~~using the Matlab tool~~ (Fig. 8c, solid black lines). The~~se~~ probability distributions of τ~~-~~ incorporating age uncertainty, derived for both the Q4 and Q7 surfaces, illustrate how the likelihood of surface abandonment is distributed over time for two representative natural datasets. On the Q4 surface, the measured age uncertainty is small compared to τ, so the resulting τ distribution has an asymmetric shape that is primarily determined by the form of Eq.~~s~~ 2~~1 to 6~~5 and our artificial~~-~~data calibration (Figs. 3 and 4). The majority of the Q4 τ distribution occupies a short timespan that is smaller than the spread of sampled surface ages; this result supports our reasoning that the timing of surface abandonment can, in some cases, be constrained more precisely than a representative age of surface formation (see section 2). The age uncertainty on a_{min} is significantly larger on the older Q7 surface and therefore dominates the probability distribution of τ, giving it a wider and more symmetrical shape despite the greater number of measured ages, n. This result underscores the importance of accounting for age uncertainty when using our equations to infer the likely timing of surface abandonment~~, which our supplementary Matlab tool incorporates~~.

5.2.1 Climatic implications

Our estimates of when the Laguna Salada fans were abandoned have important climatic implications. Spelz et al. (2008) speculated that the aggradation and incision of the fan surfaces was partly controlled by past climate changes, and there is growing evidence that alluvial systems can be highly sensitive hydroclimate recorders (D'Arcy et al., 2017a,b; Terrizzano et al., 2017; Tofelde et al., 2017; Ratnayaka et al., 2018; Wickert and Schildgen, 2019). We explore this idea by comparing the surface age data with two palaeoclimate proxy records (Fig. 8d): the GRIP ice core $\delta^{18}O$ record from Greenland (Johnsen et al., 1997) and the LR04 global benthic $\delta^{18}O$ stack (Lisiecki and Raymo, 2005). These records primarily reflect the growth and decay of continental ice sheets, which are generalised into Marine Isotope Stages (MIS).

The obtained ~~sampled~~ Q7 ages clearly coincide with the broadly interglacial conditions of MIS 7, so we interpret that the surface was deposited throughout this stage. Our statistical analyses indicate that the Q7 surface was abandoned—in this case due to fan incision—during the subsequent MIS 6 and coinciding with a climatic transition to ~~a~~ more glacial ~~climate~~conditions. Indeed, 71% of the area beneath the Q7 τ distribution falls within MIS 6 (191-130 ka), which we interpret as a 71% likelihood that the surface was abandoned and incised during this stage. For the Q4 fan surface, the sampled ages alone indicate that abandonment coincided with the end of the Last Glacial Maximum (MIS 2) and the global shift to interglacial conditions in the Holocene. Spelz et al. (2008) interpreted this observation (fan incision during a shift to interglacial climate) to contradict the Q7 data (fan incision during a shift to more glacial conditions~~limate~~). However, supplementing the measured ages with our probabilistic analyses reveals that Q4 abandonment is likely to have occurred during the Younger Dryas, a short-lived climate episode from~~between~~ 12.9 to ~~-~~11.7 ka during which the northern-hemisphere climate returned to a cooler state (Carlson, 2013). ~~In Fig. 8c,~~We find that ~~36% of the τ distribution falls within the Younger Dryas and~~ the peak of the τ distribution—i.e., the single most probable abandonment age—falls at 12.7 ka (Fig. 8c). This interpretation reconciles the Q7 and Q4 surfaces on the Laguna Salada fans, which would have both been incised as a result of climatic shifts towards more glacial conditions. This case study also demonstrates how our probabilistic approach, ~~uniquely~~ enabled by our use of artificial data, can be used to quantify the likelihood of individual abandonment scenarios and strengthen palaeoclimatic interpretations derived from alluvial deposits.

5.2.2. Tectonic and weathering implications

The results in Fig. 8 also have tectonic implications. The Laguna Salada fans are dissected by fault scarps related to the Laguna Salada fault and the Cañada David detachment; the largest Q7 scarp has an offset of 9.9 m (Spelz et al., 2008). Typically, studies divide the fault offset by the mean surface age (which for Q7 is 215.9 ka) to estimate a time-averaged slip rate, which would be 0.046 mm yr^{-1} in this example. However, as a scarp can only accumulate displacement once the surface has been abandoned, i.e., when it is no longer being resurfaced, the estimated age of abandonment may be a more appropriate timescale for determining a displacement rate. Accumulating a 9.9 m offset since 177 ka (the most likely abandonment age, Fig. 8c) would produce a time-averaged slip rate of 0.056 mm yr^{-1}; an increase of 22%. Following this logic, the probability distribution of τ could be translated into a probability distribution of time-averaged slip rates. For the Q4 fan surface, calculating a slip rate with a most likely abandonment age (e.g., 12.7 ka) instead of the mean surface age (23.3 ka) would result in an even larger increase in the calculated displacement rate of 83%. Underestimating fault slip rates by this magnitude could have important~~major~~ implications for tectonic and fault hazard analyses.

Spelz et al. (2008) also measured the diffusional decay of fault scarp geometry over time, and used the calculated mean fan surface ages to derive time-integrated scarp mass diffusivities between ~0.01-0.10 m^2 kyr^{-1}. Intriguingly, the authors interpreted these diffusivities to be anomalously slow. ~~This conundrum~~ More realistic values can be obtained ~~-could be partly resolved~~ by, again, using the estimated surface abandonment ages to calculate scarp mass diffusivity, rather than average surface ages. This approach would result in faster diffusion rates, as Spelz et al. (2008) expected, while ~~simultaneously~~ recognising that a fault scarp can only form and erode once a fan surface has been abandoned.

The alluvial fans of the Laguna Salada Basin provide a representative example of natural, aggraded geomorphic surfaces that~~, which~~ are formed over a non-negligible period of activity and are dated with a small set of exposure ages that randomly ~~-~~sample the duration of surface activity. This case study demonstrates tha~~t~~t ~~Eqs 21 through to 65, together with an incorporation of exposure age uncertainty provided by the Matlab tool,~~our artificial-data approach can provide valuable constraints on the timing of surface abandonment based on a set of exposure ages, which ~~. These constraints complement the sampled surface ages and~~ can improve~~enhance~~ interpretations involving palaeoclimate, tectonics, and landform evolution.

5.3 Limitations to the probabilistic approach

Our artificial~~-~~data~~,~~ approach and ~~therefore~~ the resulting parameterisation of Eqs. 2~~1~~ through~~o~~ 6~~5~~ ~~,~~assume that a distribution of surface ages are obtained by randomly sampling the full duration of surface activity. In some cases, this assumption might be realistic. For example, ~~, e.g.,~~ the Q7 surface on the Laguna Salada fans (Fig. 8) was sampled in different places and produced ages spanning all of MIS 7, suggesting the full duration of surface activity might be well-represented. If so, ~~Eqs 1 to 5~~our approach could be symmetrically applied to the oldest sampled age to estimate the onset of deposition. In contrast, the Q4 surface was sampled entirely at the fan apex, where enhanced vertical aggradation makes it likely that the earliest deposits from this depositional episode have been buried. In practice, this sampling approach would improve estimates of when abandonment occurred. By clustering the surface ages towards the end of the depositional period, T would effectively shorten, given that our approach derives T empirically from the ages that are actually obtained from a surface, and τ would be constrained more precisely as a result. ~~Because our approach derives T empirically from the ages that are actually sampled on a surface (Eq. 6~~~~5), the burial of early deposits does not matter for estimating abandonment. Indeed, any average surface age~~The total duration of deposition would be biased toward younger ages by the burial of older deposits, but this bias is unimportant when focusing on the timing of abandonment, which is a strength of our approach. However, vertical burial would mean that T (solved with Eq. 6) would no longer represent the total duration of deposition, and it would therefore be inappropriate to use our ~~equations~~ approach to estimate the onset of deposition.

Like burial, subsequent erosion of part of a surface might hide a fragment of the period of deposition from sampling. The implications of erosion depend on how spatially-homogenous the surface is, i.e., whether erosion has randomly eliminated ‘selectable’ ages from throughout the duration of activity, or instead eradicated complete fragments of the timespan of activity. Again, erosion would only impede our method of inferring the abandonment age if the youngest part of the duration of activity were destroyed. Given that burial and erosion are site-specific, they cannot be universally incorporated into our equations and must be considered on an individual-case basis.

~~Equations 2~~~~1 to 6~~Our approach~~5~~ assumes that all sampled surface ages are true ages. In reality, incorrect ages are sometimes encountered when dating surfaces. For example, cosmogenic nuclide exposure ages may be biased towards older ages as a result of nuclide inheritance, as is interpreted to be the case with the oldest exposure age on the Laguna Salada Q4 fan surface (Fig. 8a). Including old outliers in our analyses would lead to an over-estimation of the size of both T and τ, and therefore unnecessarily imprecise estimates of the abandonment window, but would not change the position of a_{min}. A more serious error would arise from incorrect young ages, e.g., resulting from erosion or shielding of boulders targeted for cosmogenic nuclide exposure dating. The inclusion of spurious young ages could expand the apparent period of surface activity T past the true timing of abandonment, leading to estimates of τ that are both too large and, more importantly, too young. Therefore, equations 2~~1~~-6~~5~~ and the Matlab tool should be applied to ‘clean’ datasets that do not contain spurious ages, and particularly not spuriously young ages when attempting to calculate abandonment times.

Finally, our approach derives the true period of surface activity, T, from the measured age range $a_{max} - a_{min}$, based on the results of our artificial ~~-~~data experiments (see section 4.2 and Fig. 6). This step is necessary because the true duration of T is ultimately unknowable for natural surfaces, so we parameterise Eq. 6~~5~~ using the mean ratio of $(a_{max} - a_{min})/T$ among our artificial ~~-~~data experiments. Of course, any given set of real surface ages might happen to capture a greater or smaller fraction of T than the mean case. For this reason, we also provide parameterisations of Eq. 6~~5~~ for ±1 standard deviation (σ) above and below the mean ratio of $(a_{max} - a_{min})/T$, thus allowing $\pm 1\sigma$ uncertainty on T to be tested. In practice, the uncertainty associated with T has little effect on the probability distributions of τ produced by Eq. 2~~1~~, and so is likely to be insignificant for most geomorphological applications. To illustrate the sensitivity of τ to the uncertainty on T, we re-calculate the probability distributions of τ for the Q4 and Q7 Laguna Salada alluvial fan surfaces with the Matlab tool (Fig. 9) using the $\pm 1\sigma$ bounds on T (Fig. 6).

The uncertainty on T has a negligible effect on the probability distributions of τ, for both the young and precisely-dated Q4 surface where the τ distribution is most sensitive to the form of Eqs 2~~1~~ through~~o~~ 5, and the older, less-precisely dated Q7 surface, where the τ distribution is most sensitive to the measured age uncertainty. This sensitivity analysis demonstrates how the conversion of $a_{max} - a_{min}$ to T has little bearing on the estimated timings of surface abandonment. Nonetheless, our artificial-data calibration allows the $\pm 1\sigma$ uncertainty on T to be calculated, if desired.

6 Conclusions

Our study uses artificial data to simulate depositional geomorphic surfaces that form over a non-negligible timespan, and are subsequently dated with exposure ages on a set of randomly-sampled ~~surface ages~~boulders. We investigate scenarios that are representative of natural alluvial fans, which are commonly targeted for surface dating, however our results may be more broadly applicable to other depositional landforms that form over protracted periods of time. Our findings suggest that, for a variety of different purposes, inferring the timing of surface abandonment may provide more informative and more precise interpretations than taking an average of measured surface ages. We use our artificial data to derive a set of probabilistic equations that can be applied to a distribution of real sampled surface ages to estimate a period of time within which abandonment is likely to have occurred with a given probability. These equations account for site-specific variables including the number of ages and the duration of activity for a particular surface, and ~~they~~ our artificial-data approach can be used to generate a probability distribution of likely abandonment ages. We ~~furthermore~~ provide a Matlab script that generates a probability distribution of abandonment ages for a given surface, and furthermore allows for the uncertainty associated with measured ages to be incorporated ~~in the probability distribution of abandonment ages~~. The ability to constrain the timing of surface abandonment has useful applications for geomorphological studies that relate surface ages to tectonic deformation (e.g., deriving fault slip rates), climate (e.g., reconstructing past hydroclimate changes), or the rates of surface processes (e.g., weathering and landform evolution), a subset of which we demonstrate using a case study of alluvial fan surfaces in the Laguna Salada Basin, Mexico. The statistical framework we introduce in this paper offers a new method of probabilistically estimating when a surface was abandoned, which can complement and enhance interpretations of any distribution of sampled ages obtained from surfaces that experienced a non-negligible period of deposition.

7 Notation list

a_{min} — Youngest sampled age (kyr)

a_{max} — Oldest sampled age (kyr)

a, b, c — Empirical constants relating k to P

k — Exponential decay constant relating τ to n

m_0, g, h — Empirical constants relating τ_0 to P and T

n Number of sampled ages

P Probability (between 0 and 1)

q, r, s Empirical constants for estimating T from a_{max} and a_{min}

σ Standard deviation

t_{aban} Age of surface abandonment (kyr)

T Timespan of active surface formation (kyr)

τ Abandonment window; the difference between a_{min} and t_{aban} (kyr)

τ_0 Parameter relating τ to P and T (kyr)

~~7~~8 Code availability

The Matlab script used to analyse the data is provided in the supplement.

9~~8~~ Data availability

All data in this article are presented in the main paper and are freely available online via the Figshare repository at: https://figshare.com/s/8001bf97556ae078d1e8.

10~~9~~ Author contributions

MKD, JMT, and TFS conceived of the idea for this article. MKD performed the analyses and wrote the main text. TFS created the Probabilistic Sampling Matlab script and contributed to interpretations. JMT contributed to the statistical analyses and PDN contributed to climatic interpretations. All authors commented on and contributed to the final article.

11~~0~~ Competing interests

They authors declare that they have no conflict of interest.

~~7~~12 Acknowledgements

M. D'Arcy was supported by an Alexander von Humboldt postdoctoral fellowship, a Research Focus: Earth Sciences grant awarded by the University of Potsdam, and by the Emmy-Noether- Programm of the Deutsche Forschungsgemeinschaft grant number SCHI 1241/1-1 awarded to T. Schildgen. We thank Luca Malatesta and an anonymous reviewer for constructive reviews that improved the quality of the manuscript.

References

Andrews, D.J., and Bucknam, R.C. Fitting degradation of shoreline scarps by a nonlinear diffusion model. Journal of Geophysical Research: Solid Earth, 92, 12857-12867, https://doi.org/10.1029/JB092iB12p12857, 1987.

Brooke, S.A.S., Whittaker, A.C., Armitage, J.J., D'Arcy, M., and Watkins, S.E. Quantifying sediment transport dynamics on alluvial fans from spatial and temporal changes in grain size, Death Valley, California. Journal of Geophysical Research: Earth Surface, 123, 2039-2067, https://doi.org/10.1029/2018JF004622, 2018.

Bufe, A., Burbank, D.W., Liu, L., Bookhagen, B., Qin, J., Chen, J., Li, T., Thompson Jobe, J.A., and Yang, H. Variations of lateral bedrock erosion rates control planation of uplifting folds in the foreland of the Tian Shan, NW China. Journal of Geophysical Research: Earth Surface, 122, 2431-2467, https://doi.org/10.1002/2016JF004099, 2017.

Carlson, A.E. The Younger Dryas climate event. Encyclopedia of Quaternary Science, 3, 126-134, https://doi.org/10.1016/B978-0-444-53643-3.00029-7, 2013.

Cesta, J.M., and Ward, D.J. Timing and nature of alluvial fan development along the Chajnantor Plateau, northern Chile. Geomorphology, 273, 412-427, https://doi.org/10.1016/j.geomorph.2016.09.003, 2016.

Clow, T., Behr, W.M., and Helper, M.A. Pleistocene to recent geomorphic and incision history of the northern Rio Grande gorge, New Mexico: Constraints from field mapping and cosmogenic ^{3}He surface exposure dating. Geosphere, 15, 1-19, https://doi.org/10.1130/GES02017.1, 2019.

D'Arcy, M., Roda Boluda, D.C., Whittaker, A.C., and Carpineti, A. Dating alluvial fan surfaces in Owens Valley, California, using weathering fractures in boulders. Earth Surface Processes and Landforms, 40, 487-501, https://doi.org/10.1002/esp.3649, 2015.

D'Arcy, M., Whittaker, A.C., and Roda-Boluda, D.C. Measuring alluvial fan sensitivity to past climate changes using a self-similarity approach to grain-size fining, Death Valley, California. Sedimentology, 64, 388-424, https://doi.org/10.1111/sed.12308, 2017a.

D'Arcy, M., Roda-Boluda, D.C., and Whittaker, A.C. Glacial-interglacial climate changes recorded by debris flow fan deposits, Owens Valley, California. Quaternary Science Reviews, 169, 288-311, https://doi.org/10.1016/j.quascirev.2017.06.002, 2017b.

D'Arcy, M., Mason, P.J., Roda-Boluda, D.C., Whittaker, A.C., Lewis, J.M.T., and Najorka, J. Alluvial fan surface ages recorded by Landsat-8 imagery in Owens Valley, California. Remote Sensing of Environment, 216, 401-414, https://doi.org/10.1016/j.rse.2018.07.013, 2018.

D'Arcy, M., Schildgen, T.F., Turowski, J.M., and DiNezio, P. Artificial data – Probabilistic age sampling, Figshare data repository, https://doi.org/10.6084/m9.figshare.8014583, 2019.

Denn, A.R., Bierman, P.R., Zimmerman, S.R.H., Caffee, M.W., Corbett, L.B., and Kirby, E. Cosmogenic nuclides indicate that boulder fields are dynamics, ancient, multigenerational features. GSA Today, 28, 4-10, https://doi.org/10.1130/GSATG340A.1, 2017.

Dühnforth, M., Densmore, A.L., Ivy-Ochs, S., Allen, P.A., Kubik, P.W. Timing and patterns of debris flow deposition on Shepherd and Symmes creek fans, Owens Valley, California, deduced from cosmogenic ^{10}Be. Journal of Geophysical Research, 112, F03S15, https://doi.org/10.1029/2006JF000562, 2007.

Finnegan, N.J., Schumer, R., and Finnegan, S. A signature of transience in bedrock river incision rates over timescales of 10^4–10^7 years. Nature, 505, 391-394, https://doi.org/10.1038/nature12913, 2014.

Frankel, K.L., Brantley, K.S., Dolan, J.F., Finkel, R.C., Klinger, R.E., Knott, J.R., Machette, M.N., Owen, L.A., Phillips, F.M., Slate, J.L., and Wernicke, B.P. Cosmogenic ^{10}Be and ^{36}Cl geochronology of offset alluvial fans along the northern Death Valley fault zone: Implications for transient strain in the eastern California shear zone. Journal of Geophysical Research, 112, B06407, https://doi.org/10.1029/2006JB004350, 2007.

Frankel, K.L., Dolan, J.F., Owen, L.A., Ganev, P., and Finkel, R.C. Spatial and temporal constancy of seismic strain release along an evolving segment of the Pacific–North America plate boundary. Earth and Planetary Science Letters, 304, 565-576, https://doi.org/10.1016/j.epsl.2011.02.034, 2011.

Foreman, B.Z., Straub, K.M. (2017) Autogenic geomorphic processes determine the resolution and fidelity of terrestrial paleoclimate records. Science Advances, 3, 1-11, https://doi.org/10.1126/sciadv.1700683, 2017.

Foster, M.A., Anderson, R.S., Gray, H.J., and Mahan, S.A. Dating of river terraces along Lefthand Creek, western High Plains, Colorado, reveals punctuated incision. Geomorphology, 295, 176-190, https://doi.org/10.1016/j.geomorph.2017.04.044, 2017.

Ganev, P.N., Dolan, J.F., Frankel, K.L., and Finkel, R.C. Rates of extension along the Fish Lake Valley fault and transtensional deformation in the Eastern California shear zone–Walker Lane belt. Lithosphere, 2, 33-49, https://doi.org/10.1130/L51.1, 2010.

Gosse, J.C. (2011) Terrestrial Cosmogenic Nuclide Techniques for Assessing Exposure History of Surfaces and Sediments in Active Tectonic Regions. In: Busby, C., Azor, A. (Eds) Tectonics of Sedimentary Basins: Recent Advances. Blackwell Publishing Ltd, Chichester, UK, pp. 63-79.

Hanks, T.C., Bucknam, R.C., Lajoie, K.R., and Wallace, R.E. Modification of wave-cut and faulting-controlled landforms. Journal of Geophysical Research: Solid Earth, 89, 5771-5790, https://doi.org/10.1029/JB089iB07p05771, 1984.

Hughes, A., Rood, D.H., Whittaker, A.C., Bell, R.E., Rockwell, T.K., Levy, Y., Wilcken, K.M., Corbett, L.B., Bierman, P.R., DeVecchio, D.E., Marshall, S.T., Gurrola, L.D., Nicholson, C. Geomorphic evidence for the geometry and slip rate of a young, low-angle thrust fault: Implications for hazard assessment and fault interaction in complex tectonic environments. Earth and Planetary Science Letters, 504, 198-210, https://doi.org/10.1016/j.epsl.2018.10.003, 2018.

Johnsen, S.J., Clausen, H.B., Dansgaard, W., Gundestrup, N.S., Hammer, C.U., Andersen, U., Andersen, K.K., Hvidberg, C.S., Dahl-Jensen, D., Steffensen, J.P., and Shoji, H. The δ^{18}O record along the Greenland Ice Core Project deep ice core and the problem of possible Eemian climatic instability. Journal of Geophysical Research: Oceans, 102, 26397-26410, https://doi.org/10.1029/97JC00167, 1997.

Lal, D. Cosmic ray labeling of erosion surfaces: In situ nuclide production rates and erosion models. Earth and Planetary Science Letters, 104, 424-439, https://doi.org/10.1016/0012-821X(91)90220-C, 1991.

Lisiecki, L.E., and Raymo, M.E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. Paleoceanography, 20, PA1003, https://doi.org/10.1029/2004PA001071, 2005.

Macklin, M.G., Fuller, I.C., Lewin, J., Maas, G.S., Passmore, D.G., Rose, J., Woodward, J.C., Black, S., Hamlin, R.H.B., and Rowan, J.S. Correlation of fluvial sequences in the Mediterranean basin over the last 200 ka and their relationship to climate change. Quaternary Science Reviews, 21, 1633-1641, https://doi.org/10.1016/S0277-3791(01)00147-0, 2002.

Malatesta, L.C., Prancevic, J.P., and Avouac, J.-P. Autogenic entrenchment patterns and terraces due to coupling with lateral erosion in incising alluvial channels. Journal of Geophysical Research: Earth Surface, 122, 335-355, https://doi.org/10.1002/2015JF003797, 2017.

Martin, L.C.P., Blard, P.-H., Balco, G., Lavé, J., Delunel, R., Lifton, N., and Laurent,V. The CREp program and the ICE-D production rate calibration database: A fully parameterizable and updated online tool to compute cosmic-ray exposure ages. Quaternary Geochronology, 38, 25-49, https://doi.org/10.1016/j.quageo.2016.11.006, 2017.

Mouslopoulou, V., Begg, J., Fülling, A., Moraetis, D., Partsinevelos, P., and Oncken, O. Distinct phases of eustatic and tectonic forcing for late Quaternary landscape evolution in southwest Crete, Greece. Earth Surface Dynamics, 5, 511-527, https://doi.org/10.5194/esurf-5-511-2017, 2017.

Muscheler, R., Beer, J., Kubik, P.W., and Synal, H.-A. Geomagnetic field intensity during the last 60,000 years based on ^{10}Be and ^{36}Cl from the Summit ice cores and ^{14}C. Quaternary Science Reviews, 24, 1849-1860, https://doi.org/10.1016/j.quascirev.2005.01.012, 2005.

Owen, L.A., Frankel, K.L., Knott, J.R., Reynhout, S., Finkel, R.C., Dolan, J.F., and Lee, J. Beryllium-10 terrestrial cosmogenic nuclide surface exposure dating of Quaternary landforms in Death Valley. Geomorphology, 125, 541-557, https://doi.org/10.1016/j.geomorph.2010.10.024, 2011.

Owen, L.A., Clemmens, S.J., Finkel, R.J., and Gray, H. Late Quaternary alluvial fans at the eastern end of the San Bernardino Mountains, Southern California. Quaternary Science Reviews, 87, 114-134, https://doi.org/10.1016/j.quascirev.2014.01.003, 2014.

Ratnayaka, K., Hetzel, R., Hornung, J., Hampel, A., Hinderer, M., and Frechen, M. Postglacial alluvial fan dynamics in the Cordillera Oriental, Peru, and palaeoclimatic implications. Quaternary Research, 91, 431-449, https://doi.org/10.1017/qua.2018.106, 2018.

Regmi, N.R., McDonald, E.V., and Bacon, S.N. Mapping Quaternary alluvial fans in the southwestern United States based on multiparameter surface roughness of lidar topographic data. Journal of Geophysical Research: Earth Surface, 119, 12-27, https://doi.org/10.1002/2012JF002711, 2014.

Sarıkaya, M.A., Yıldırım, C., and Çiner, A. Late Quaternary alluvial fans of Emli Valley in the Ecemiş Fault Zone, south central Turkey: Insights from cosmogenic nuclides. Geomorphology, 228, 512-525, https://doi.org/10.1016/j.geomorph.2014.10.008, 2015.

Savi, S., Schildgen, T.F., Tofelde, S., Wittmann, H., Scherler, D., Mey, J., Alonso, R.N., and Strecker, M.R. Climatic controls on debris-flow activity and sediment aggradation: The Del Medio fan, NW Argentina. Journal of Geophysical Research: Earth Surface, 121, 2424-2445, https://doi.org/10.1002/2016JF003912, 2016.

Schildgen, T.F., Cosentino, D., Bookhagen, B., Niedermann, S., Yıldırım, C., Echtler, H., Wittmann, H., and Strecker, M.R. Multi-phased uplift of the southern margin of the Central Anatolian plateau, Turkey: A record of tectonic and upper mantle processes. Earth and Planetary Science Letters, 317-318, 85-95, https://doi.org/10.1016/j.epsl.2011.12.003, 2012.

Schildgen, T.F., Robinson, R.A.J., Savi, S., Phillips, W.M., Spencer, J.Q.G., Bookhagen, B., Scherler, D., Tofelde, S., Alonso, R.N., Kubik, P.W., Binnie, S.A., and Strecker, M.R. Landscape response to late Pleistocene climate change in NW Argentina: Sediment flux modulated by basin geometry and connectivity. Journal of Geophysical Research: Earth Surface, 121, 392-414, https://doi.org/10.1002/2015JF003607, 2016.

Schürch, P., Densmore, A.L., Ivy-Ochs, S., Rosser, N.J., Kober, F., Schlunegger, F., McArdell, B., and Alfimov, V. Quantitative reconstruction of late Holocene surface evolution on an alpine debris-flow fan. Geomorphology, 275, 46-57, http://dx.doi.org/10.1016/j.geomorph.2016.09.020, 2016.

Simpson, G., and Castelltort, S. Model shows that rivers transmit high-frequency climate cycles to the sedimentary record. Geology, 40, 1131-1134, https://doi.org/10.1130/G33451.1, 2012.

Spelz, R.M., Fletcher, J.M., Owen, L.A., and Caffee, M.W. Quaternary alluvial-fan development, climate and morphologic dating of fault scarps in Laguna Salada, Baja California, Mexico. Geomorphology, 102, 578-594, https://doi.org/10.1016/j.geomorph.2008.06.001, 2008.

Steffen, D., Schlunegger, F., and Preusser, F. Drainage basin response to climate change in the Pisco valley, Peru. Geology, 37, 491-494, https://doi.org/10.1130/G25475A.1, 2009.

Stone, J.O. Air pressure and cosmogenic isotope production. Journal of Geophysical Research, 105, 23753-23759, https://doi.org/10.1029/2000JB900181, 2000.

Terrizzano, C.M., García Morabito, E., Christl, M., Likerman, J., Tobal, J., Yamin, M., and Zech, R. Climatic and tectonic forcing on alluvial fans in the southern Central Andes. Quaternary Science Reviews, 172, 131-141, https://doi.org/10.1016/j.quascirev.2017.08.002, 2017.

Tofelde, S., Schildgen, T.F., Savi, S., Pingel, H., Wickert, A.D., Bookhagen, B., Wittmann, H., Alonso, R.N., Cottle, J., and Strecker, M.R. 100 kyr fluvial cut-and-fill terrace cycles since the Middle Pleistocene in the southern Central Andes, NW Argentina. Earth and Planetary Science Letters, 473, 141-153, https://doi.org/10.1016/j.epsl.2017.06.001, 2017.

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., Bechtold, V.D.C., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., and Li, X. The ERA-40 re-analysis. Quarterly Journal of the Royal Meteorological Society, 131, 2961-3012, https://doi.org/10.1256/qj.04.176, 2005.

Valet, J.-P., Meynadier, L., and Guyodo, Y. Geomagnetic dipole strength and reversal rate over the past two million years. Nature, 435, 802-805, https://doi.org/10.1038/nature03674, 2005.

White, A.F., Blum, A.E., Schulz, M.S., Bullen, T.D., Harden, J.W., and Peterson, M.L. Chemical weathering rates of a soil chronosequence on granitic alluvium: I. Quantification of mineralogical and surface area changes and calculation of primary silicate reaction rates. Geochimica et Cosmochimica Acta, 60, 2533-2550, https://doi.org/10.1016/0016-7037(96)00106-8, 1996.

White, A.F., Schulz, M.S., Vivit, D.V., Blum, A.E., Stonestrom, D.A., and Harden, J.W. Chemical weathering rates of a soil chronosequence on granitic alluvium: III. Hydrochemical evolution and contemporary solute fluxes and rates. Geochimica et Cosmochimica Acta, 69, 1975-1996, https://doi.org/10.1016/j.gca.2004.10.003, 2005.

Wickert, A.D., and Schildgen, T.F. Long-profile evolution of transport-limited gravel-bed rivers. Earth Surface Dynamics, 7, 17-43, https://doi.org/10.5194/esurf-7-17-2019, 2019.

Figures

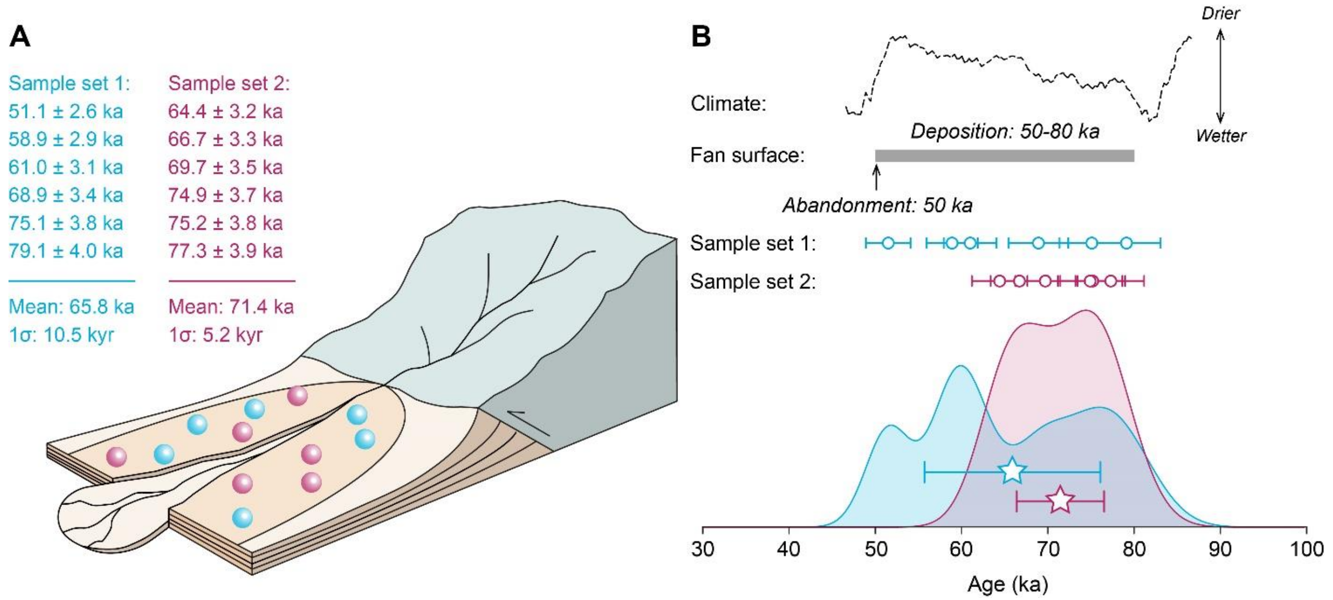

Figure 1: (A) Conceptual alluvial fan surface that was formed over a 30 kyr period, from 80 ka to 50 ka, after which it was abandoned, e.g., due to incision. Two different dating scenarios (sample sets 1 and 2) are shown in which 6 surface ages are randomly selected. (B) The true period of surface activity (grey bar), compared with the sampled ages presented as data points (circles), kernel density plots, and mean surface ages ±1 standard deviation (stars). A hypothetical climate scenario is depicted as a dotted line.

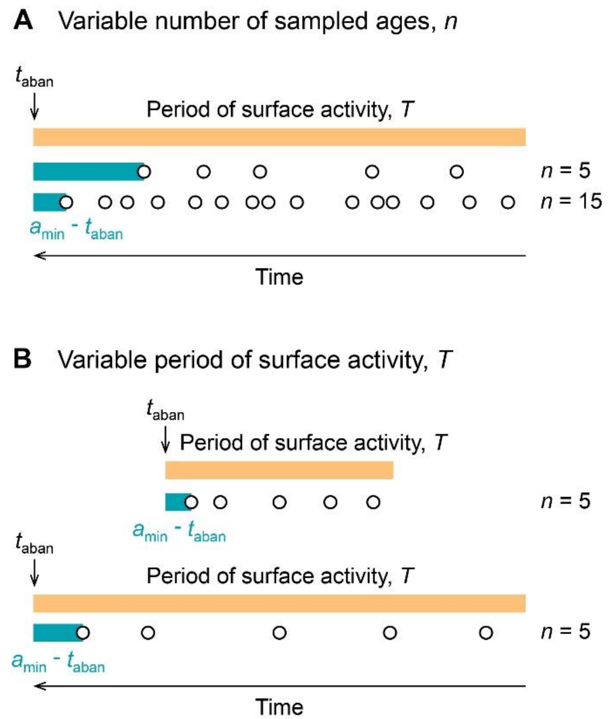

Figure 2: Schematic surface with a period of activity, T (orange bar), abandoned at t_{aban}, and randomly sampled with n ages (circles). (A) If n increases, the youngest sampled age, a_{min}, is likely to fall closer to t_{aban}. (B) If T increases, the youngest sampled age, a_{min}, is likely to fall farther from t_{aban}, even if the same number of ages are sampled.

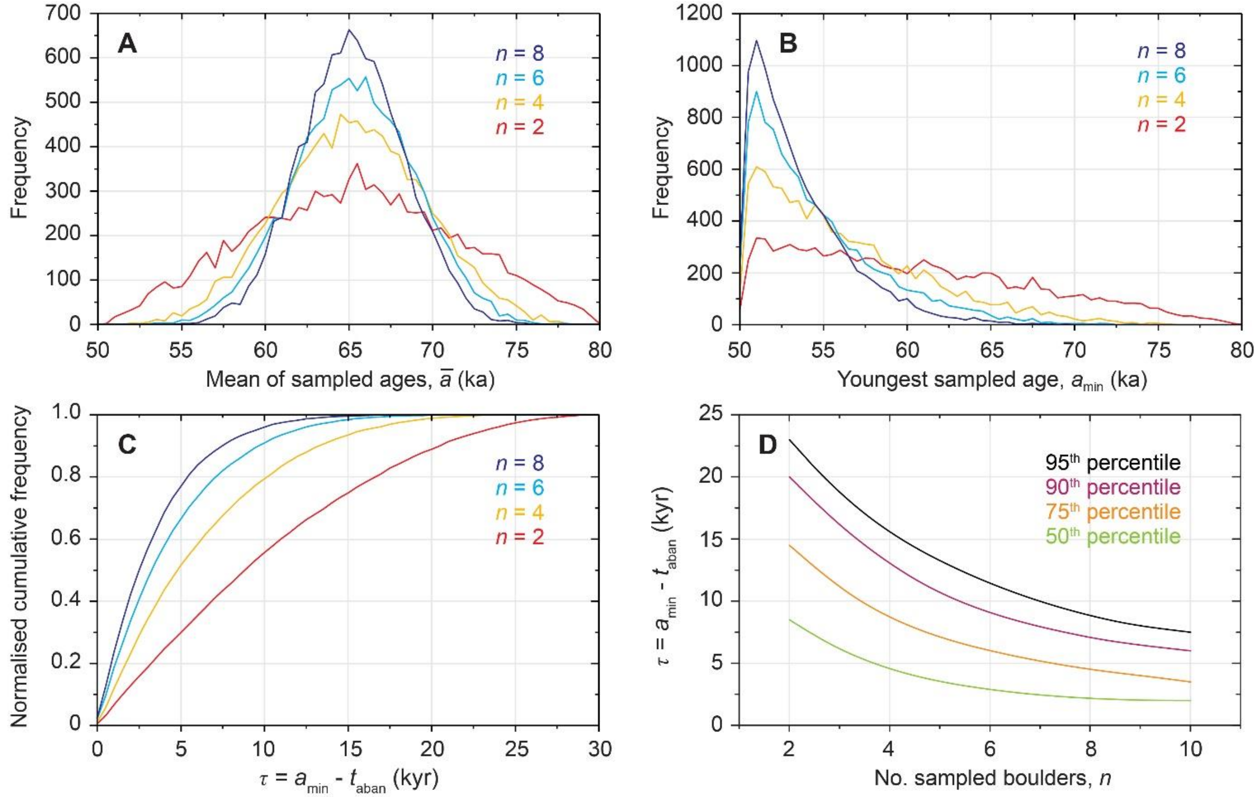

Figure 3: Example results of the artificial-data experiments for a surface active from 80 to 50 ka (T = 30 kyr). A number of ages, n, were randomly sampled from the surface 10,000 times. (A) Frequency distributions of resulting mean sampled age, $\bar{a}$. (B) Frequency distributions of the youngest sampled age, a_{min}. (C) Cumulative frequency distributions of τ normalised to a sum of 1. (D) Selected percentiles of τ plotted against n.

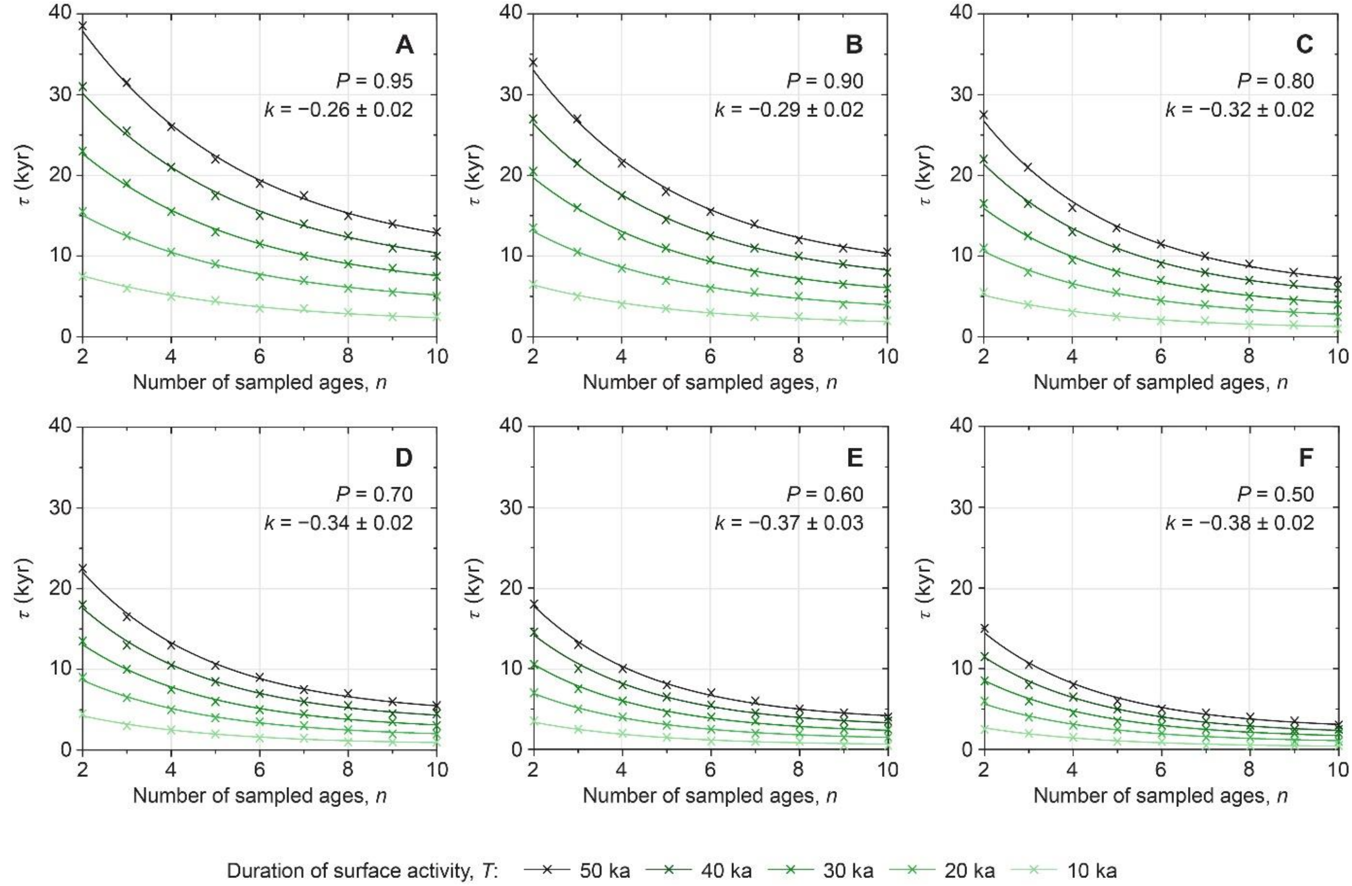

Figure 4: The probable abandonment window, τ, as a function of the number of boulder ages, n. Data are shown for different probabilities, P (panels), and durations of surface activity, T (colours). Parameter k is a decay constant that depends on P (see text for details).

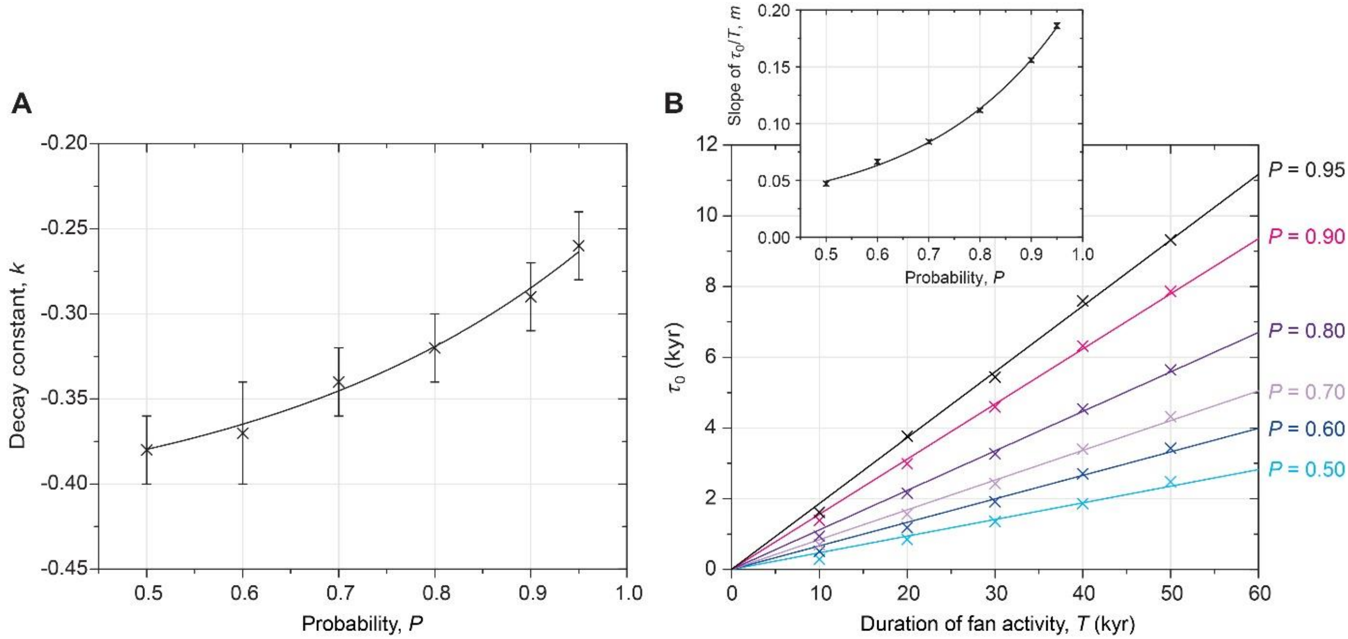

Figure 5: (A) Variation in the decay constant, *k*, as a function of the probability, *P*. Error bars show the standard error on *k* when Eq. 2~~1~~ is fitted to the data in Fig. 4. The regression corresponds to Eq. 3~~2~~. (B) Variation in τ_0 as a function of *T* for different values of *P* (indicated by colours). Linear regressions are fitted corresponding to Eq. 4~~3~~. Inset: Variation in *m* as a function of *P*. An exponential regression is fitted corresponding to Eq. 5~~4~~.

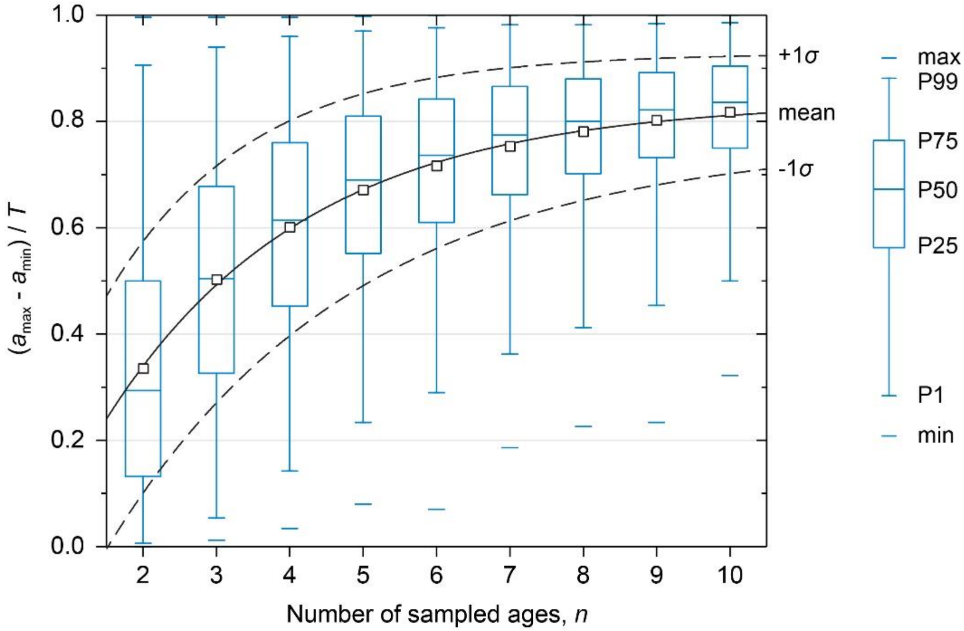

Figure 6: Box plots showing the fraction of the total period of fan activity, *T*, captured by the span of sampled boulder ages in the artificial-data experiments, $a_{max} - a_{min}$ plotted against the number of sampled ages, *n*. Each box represents 10,000 experiments. As a greater number of ages are sampled, the span of the set of ages is more likely to capture a greater fraction of *T*, although with diminishing returns for increasing *n*. Black lines show exponential regressions corresponding to Eq. 6~~5~~ and fitted to the mean values (solid) and ±1 standard deviation (dashed).

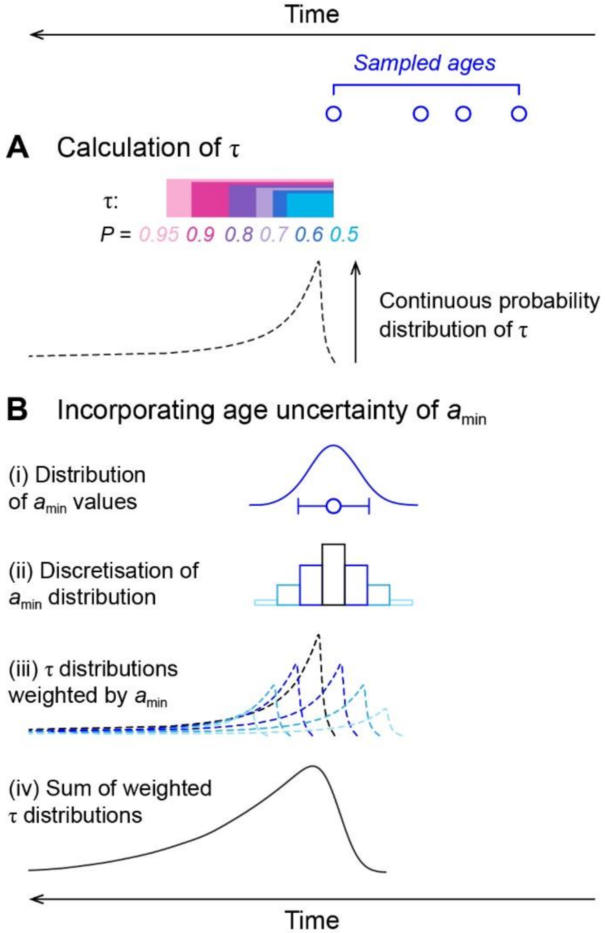

Figure 7: Schematic demonstrating how to infer the timing of surface abandonment from a set of sampled ages. (A) Probable abandonment windows, τ, are calculated using Eq. 2~~1~~ for discrete values of P (coloured bars).~~ (B)~~ A continuous probability distribution of τ is ~~calculated as~~equal to the density of P within each discrete ~~window~~ increment of τ ~~in~~ (dashed line); this can be generated using the Matlab tool provided~~A)~~. (B~~C~~) In reality, a_{min} is not perfectly known, and has an associated age uncertainty that must be accounted for. (i) The ±3σ uncertainty on a_{min} provides a distribution of probable values of a_{min}. (ii) The distribution of a_{min} values is discretised. In the Matlab tool, we have set this discretization to be 1/10 the 1σ uncertainty on the youngest age, a_{min}, to provide a highly~~-~~-resolved result (note that the cartoon illustration here shows much wider discretisation bins for ease of visualisation), but this discretisation value can be modified. The discrete window of τ used to calculate the density of P~~ in (B)~~ is set to the same width. (iii) Probability distributions of τ are calculated for each potential value of a_{min} (as per panel A~~B~~), and weighted according to the probability distribution of a_{min} values. (iv) The weighted, temporally shifted τ distributions are then summed to produce a final probability distribution of surface abandonment timing that incorporates uncertainty in the youngest age. ~~(D) Equations used to infer the timing of surface abandonment, calibrated with our artificial data.~~

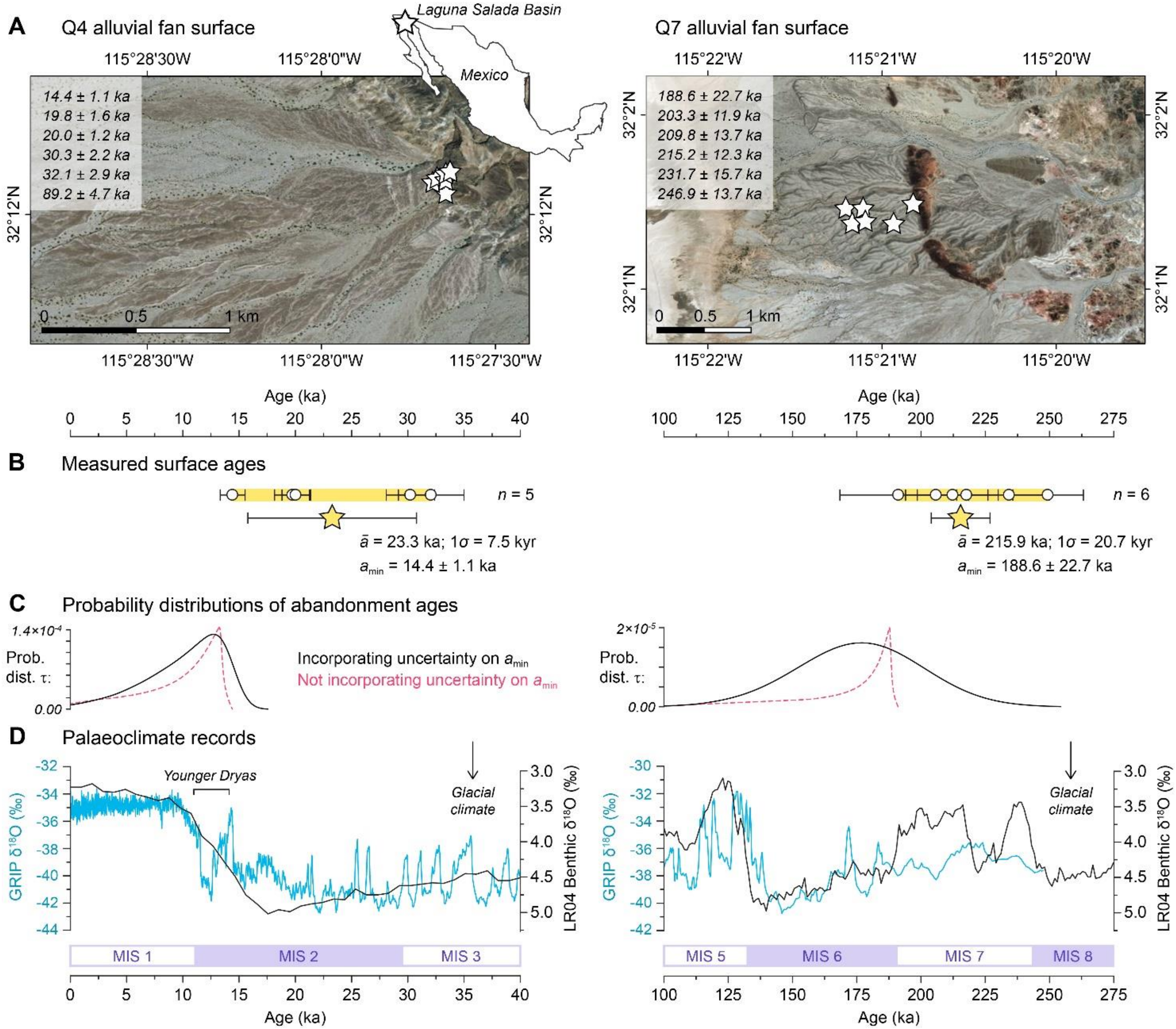

Figure 8: Two alluvial-fan surfaces in the Laguna Salada Basin, northern Baja California, Mexico. Left: Q4 surface; right: Q7 surface, after Spelz et al. (2008). (A) Locations of surface boulders sampled for ^{10}Be cosmogenic nuclide exposure dating. (B) Boulder exposure ages recalculated after Spelz et al. (2008) (white circles) and mean surface ages ±1σ (yellow stars). (C) Probability distributions of τ calculated using ~~Eq. 21~~the Matlab tool and incorporating uncertainty on a_{min} ~~following Fig. 7~~ (black). For illustrative purposes, probability distributions of τ are shown if uncertainty on a_{min} is not incorporated (red dashed). (D) Selected palaeoclimate proxies: the GRIP ice core δ^{18}O record from Greenland (blue; Johnsen et al., 1997) and the LR04 global benthic δ^{18}O stack (black, Lisiecki and Raymo, 2005). Marine Isotope Stages (MIS) are indicated by boxes.

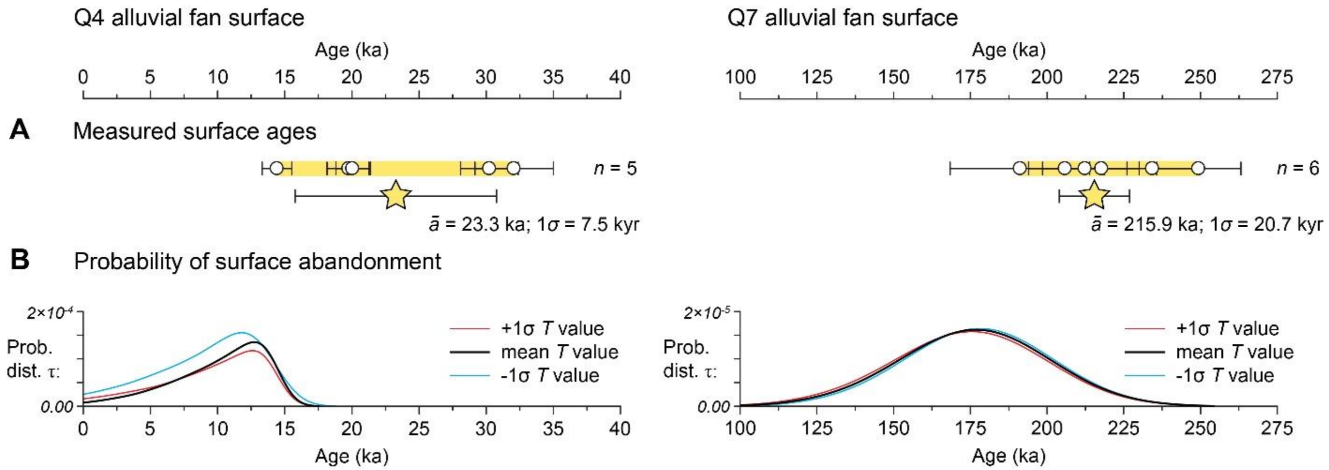

Figure 9: (A) Measured surface ages for the Laguna Salada alluvial fan surfaces, following Fig. 8b. (B) Probability distributions of τ calculated using Eq. 2~~1~~ and incorporating age uncertainty, where T is derived from the measured spread of surface ages using the mean case in Fig. 6 (black curves) and $\pm 1\sigma$ uncertainties on T (red and blue curves).