We thank Chelsea Mackaman-Lofland for the thoughtful, supportive, and constructive comments on our manuscript. In this document, reviewer comments are replicated in black, and our replies are in blue. Following the request from *GChron*, at this stage we indicate only the changes we intend to make in the revised manuscript, which will be submitted soon.

General comments

The authors present a new, open-source Python package designed to forward model large numbers of low-temperature thermochronometric ages from time-temperature paths generated by geodynamic numerical models. The manuscript, and Python package, comprise timely and potentially substantial contributions to the fields of geochronology and geodynamics – as the authors mention in the paper, advances in geodynamic numerical modeling have enabled increasingly high-resolution simulations of geological and thermal events across various tectonic settings and thermochronology data provide quantitative constraints against which such models can be validated. The GDTchron Python package presented here expands on thermokinematic modeling approaches (which predict time-temperature histories and low-temperature thermochronology ages based on deformation kinematics and topographic evolution but do not directly incorporate dynamics; Braun, 2003; Braun et al., 2012; Almendral et al., 2015) and previous strategies for integrating time-temperature information from geodynamic models with thermochronology data (which has involved exporting t-T paths from geodynamic models and importing them into thermal history modeling software like HeFTy or QTQt; Ketcham, 2005; Gallagher, 2012) in that it allows for extraction of time-temperature histories, and thermochronology age predictions, for a large number of particles from multiple steps in a geodynamic model.

The paper is well written, the accompanying Python notebooks are well documented/commented and readily accessible via Github, and the figures are generally illustrative.

We warmly thank the reviewer for the supportive comments!

However, this research has several areas that I consider in need of major improvements before manuscript publication and widespread implementation of the Python package by users in the geodynamics and thermochronology communities:

Specific comments
1. The evolution of topography can profoundly impact thermal histories and predicted thermochronology cooling ages, particularly for the low-temperature systems that this contribution focuses on (e.g., Braun, 2005; Gilmore et al., 2018). The authors briefly mention this and caution users to implement geodynamic models that account for erosion and sedimentation – and even make changes to the parameters used to approximate topographic evolution to "provide more realistic surface processes" in their rift-inversion orogenesis example model. However, particularly considering potential readers/GDTchron users who may have specific expertise in either geodynamic modeling *or* thermochronology, the authors need to include more detailed information on i) how surface processes and other parameters that affect thermal evolution, like heat production, are incorporated into geodynamic numerical models, ii) the extent to which those choices

may impact t-T histories and predicted cooling ages, and iii) related recommendations or considerations when integrating geodynamic models and thermochronology datasets.

In the revised manuscript, we will be happy to include more background on the implementation of surface processes and heat production in geodynamic models, their impact on cooling ages, and considerations about comparing models with thermochronology data. We provide some initial answers below:

i) Until recently, most geodynamic models approximated surface processes using some version of hillslope diffusion, in which material at higher elevations is redistributed to lower elevations as a basic approximation of erosion/sedimentation. In more recent models, the geodynamic model is coupled with an external surface processes model (such as Fastscape), which may include the stream power law and marine transport/deposition in addition to hillslope diffusion. The inclusion of these latter two processes (especially the stream power law) in a dedicated surface processes model is what we describe as "more realistic."

ii) As our original manuscript and the reviewer allude to, these choices can have a significant impact on cooling ages, since exhumation requires both rock uplift and erosion to bring hotter material at depth to the surface while cooling it. If erosion is too inefficient, rocks with young ages will never be exposed at the surface because they will accumulate too much daughter product as they cool, and if erosion is excessively efficient, ages will be too young because rocks will never accumulate daughter product before being eroded away.

iii) As a result, a major challenge in comparing models with observed ages lies in whether the chosen surface processes model produces patterns of ages that are realistic and whether the patterns or absolute ages are comparable with a given natural system. For example, if given surface process parameters produce reasonable patterns of ages but changing the underlying geodynamic evolution alters those patterns significantly, that could be used to discriminate between competing geodynamic scenarios in a given region. However, unless multiple surface processes implementations have been tested and/or there are observational constraints on the surface history, one could not rule out a surface processes control on the change in pattern.

A more in-depth comparison of the strategies for incorporating surface processes in the Vasey et al. (2024) rift-inversion model, versus the new rift-inversion model showcased in this study, would provide a valuable, and cautionary, case study. Both of these models incorporate the same geodynamic evolution, but produce notably different low-T cooling ages – a change solely related to the mechanism of topographic evolution. These results raise questions regarding the sensitivity of predicted cooling ages, and measured thermochronology data, to geodynamic mechanisms versus surface processes that need to be addressed in this manuscript to help educate potential GDTchron users and establish best practices.

2. Though the authors highlight advances to understanding of thermochronometer systems and diffusion/annealing kinetics in their motivation for this work, they have not incorporated the most widely used and up-to-date kinetics for predicting (U-Th)/He or fission track cooling ages or track length distributions (e.g., Enkelmann and Garver, 2016; Flowers et al., 2009; 2023a; 2023b; Guenthner, 2021; Guenthner et al., 2013, 2014; Ketcham, 2019; Ketcham et al., 2007; Tamer & Ketcham, 2020) in their Python package. This is flagged as a potential future direction, but until these updates are made, users will either be limited in comparisons of geodynamic model predictions to thermochronology data or existing thermal history models, or will need to follow the same workflows of previous studies that exported t-T paths predicted by the geodynamic models for interpretation using other software like QTQt or HeFTy (as performed in Ternois et al., 2021). I encourage the authors to incorporate apatite fission track annealing kinetics after Ketcham et al. (2007), and options to use radiation damage and annealing models for the apatite and zircon (U-Th)/He systems (e.g., Flowers et al., 2009; Guenthner et al., 2013; or more recent updates), in this publication. Baseline parameters for U and Th concentration, grain size, Dpar, and other parameters could draw from standard values– though users would benefit from options to include specific values at specific sample locations to compare with thermochronology data. Routines for the implementation of these kinetic models, and grain parameters, in the PecubeGUI (Bernard et al., 2025) may provide a useful framework for such updates.

2. For many geodynamic models that are approximating general processes rather than recreating specific systems, the patterns of ages will ultimately be more important than the absolute ages, and the uncertainties inherent in other parameters within the geodynamic models outweigh the relatively small changes in individual ages that will result from the choice of kinetics. As a result, we disagree that results will not be meaningful unless updated parameters are used. Additionally, the improvements in computation time when using simpler kinetics for large numbers of particles may be more important than the inclusion of radiation damage or other more complex kinetics in some models.

3. For the apatite radiation damage accumulation and annealing model (RDAAM) and its zircon counterpart (ZRDAAM) specifically, although these kinetic models do a better job of approximating He behavior in many populations of apatite/zircon (i.e., when there are multiple protracted phases of burial and exhumation and correlations between effective U and age), there may still be situations where the diffusion kinetics employed in GDTchron are appropriate (i.e., when there is monotonic cooling and correlations between grain size and age). Thus, it is not necessarily the case that these models have been entirely superseded by the more recent radiation damage accumulation and annealing models.

We note that GDTchron does allow users to set U, Th, grain size, and mean etch pit diameter (Dpar) for the existing models, although currently only for the entire model domain. We see allowing conditional changes in these parameters throughout the model domain as a logical next step for development.

3. While related references are briefly mentioned and cited (e.g., Ketcham, 2005; 2024; 2025; Gallagher, 2012; Braun, 2003; Braun et al., 2013; etc.), I'd like to encourage the authors to include a more in-depth discussion of how the GDTchron package fits into the framework of related work, including 3-D and 4-D thermokinematic and 2-D (time-depth and time-temperature) thermal history modeling. This may be particularly relevant to discussions and recommendations related to recommendations/best practices for incorporating crustal thermal parameters, surface processes, and disentangling the effects of each system on predicted cooling ages.

We will gladly provide more context about how GDTchron fits within the ecosystem of modeling software for thermochronologic data. In our view, its primary function is to enable quantitative testing of the effects on thermochronometric age of parameters that can be captured in geodynamic but not thermal history or thermokinematic models. These could include far-field driving forces (e.g., extension/compression rate/magnitude, slab pull), rheology (e.g., diffusion vs. dislocation creep, flow laws for different crust/mantle materials), and impact of fluids, melt, and/or phase changes, to name a few examples. This is fundamentally different from thermokinematic models like Pecube in which both motion and topography are generally prescribed. The downside of geodynamic models is that since these exhumation kinematics and surface evolution are not prescribed, it can be significantly more

difficult (though not impossible) to generate models that approximate the surface observations of real systems in detail.

4. Finally, the tendency of the implemented interpolation scheme to assign new syndeformational deposits the same thermal history as adjacent, pre-existing particle paths may be a real limitation in areas that involve significant accumulation of such deposits during polyphase tectonic and geodynamic histories (like the Pyrenees studies cited as examples in this manuscript, Capaldi et al., 2022; Curry et al., 2021). Are there any changes to the interpolation program, as implemented, that would help set up for future changes in ASPECT or other geodynamic models that might include better tracking and ID of particles related to syndeformational sedimentation?

We agree that this is a real limitation, and we will be happy to expand on it in a revised manuscript. We note, however, that this limitation is derived primarily from the geodynamic models themselves rather than from GDTchron. In its current version, GDTchron can track a particle throughout a model run as long as it maintains the same particle ID. In ASPECT, any new particles created (such as those representing deposited sediment) typically receive a new ID, decoupling them from their source. Future implementations of ASPECT-Fastscape coupling would need to have a way to track the provenance of a given sediment particle, either by maintaining the same ID as the source material (in which case no change to GDTchron is necessary) or through another field that connects the particle to its source material. Updating GDTchron to capture this field would be straightforward but would have to be done in response to a corresponding change in ASPECT.

Technical corrections

Line 54: Specify that the simplified model involves block uplift here?

We will make this change in the revised manuscript.

Line 55: Because routines incorporating erosion and sedimentation are so important in predicting cooling ages for low-T thermochronometers, I strongly recommend providing more information about what these routines were, & what geological or other evidence supports their use, here in the intro in addition to later in the paper.

We will specify the use of hillslope diffusion, the stream power law, and marine transport/deposition in the revised manuscript.

Line 109 / Section 3.1: What kind of computational resources are required to run GDTchron? Can you provide more information here regarding the types of scenarios that might run on a personal computer/workstation, versus HPC? (I recognize that related information is documented within the Github & other package resources, but would be valuable to include for readers in this section of the paper too)

We will gladly expand on this point in the revised manuscript. The choice of resource will depend on the scale of the model, but in general, we find that determining ages for all particles in the crust (upper ~50 km) at all timesteps in a publishable geodynamic model will likely require use of a high-end desktop workstation (>30 cores) or an HPC cluster node. On the other hand, running a more restricted set of particles or using a shorter or more simplified model may be doable on a more standard personal computer. We see optimization of the code to enable more of the workflow to be possible on a personal computer as an important next phase of development.

Line 141: As a reviewer with expertise in thermal history and structural/thermokinematics modeling but no hands-on experience with geodynamic software like that listed above, I am curious what common inputs and outputs to these software are, particularly as they relate to the temperature field, partical motion, and the development/decay of topography. Even though your target audience may already be familiar with these programs, I recommend including a more comprehensive background/context section that helps frame geodynamic model capabilities, and limitations, for a broader audience. E.g., how is topography typically handled in these systems, and what are recommendations/considerations for evaluating incorporation of topographic change and sedimentation in geodynamic models? In addition, there may be readers with expertise in thermochronology, but not ASPECT etc. modeling, who would like to use GDTchron to extract t-T paths and predict cooling ages based on published geodynamic models to help inform their sampling strategies. Providing a little more context for the geodynamic modeling component would help clarify the intended workflow and capabilities of the GDTchron package.

We appreciate this perspective and will gladly expand on the sorts of parameters that go into these geodynamic models in the revised manuscript. We outline much of the topography and surface processes components in the responses to the "specific comments" above. In general, we note that these models often contain hundreds of input parameters and are highly flexible, which is both their strength and weakness. Such parameters include a wide range of initial and boundary conditions, material properties, and process parameters such as the size and resolution of the model domain, a material model describing how the model deforms, velocity boundary conditions, and temperature models that may incorporate processes like radiogenic or shear heating. The strength of this approach is the ability to have a physical model that is quasi-realistic in which you can isolate any one of these given parameters to see its impact on the final result. The weakness is that many of the choices for these parameters are often semi-arbitrary or poorly constrained, making it potentially difficult to deconvolve their effects on the model outcome.

Line 169: Choose a different word than "flow" when discussing how material exits the top of the box? "Flow" has implications for rheology/deformation style that do not necessarily apply here.

This is a good point, and we will change this to "exit" to avoid potential confusion.

Line 170: By "temperature structure of the box remains constant," do you mean that there is no advection or change in isotherms due to particle uplift & exhumation? It appears so, from the model diagrams? If not, please consider revising this simple example to allow for isotherm advection, which will produce a more geologically accurate case study?

This is correct; this highly simplified model has a static temperature structure. In line with the reviewer's comment above about community familiarity with geodynamic models, we made the choice in this model to substantially reduce the number and complexity of parameters both to make it more easily comprehensible for those who do not do geodynamic modeling and to isolate the effects of rock uplift and mineral/isotopic system on age. We recognize that this is not a geologically accurate model in many respects and will emphasize this more in the revised manuscript.

Line 170: Specify "particle / rock uplift" or "exhumation" here, as there is no surface/topographic uplift in this scenario that maintains flat topography (e.g., has complete erosional efficiency).

We will change this to "rock uplift" in the revised manuscript.

Lines 176-177: Language related to Fig. 5 here is a little confusing – could be interpreted to mean that the timesteps were generated with GDTchron, but they were defined in ASPECT, right? Then the outputs of that model were processed in GDTchron?

This is correct – we will revise this sentence to make it clear that GDTchron is postprocessing the outputs of the ASPECT model.

Lines 179-180, sentence starting with "Ages young with depth…": Phrasing makes it seem like 0 Ma is reached spatially below the PRZ/PAZs, but the unit is time, not space. Change to "Ages young with depth through the partial retention/annealing zones of each thermochronometric system. At the base of these zones, cooling ages are 0 Ma because diffusion and annealing outpace He and fission track production" or similar?

This is a very good suggestion, and we will revise this sentence along these lines.

Line 190 / Fig. 5: I strongly recommend devising a way to show the movement/change in particle positions in this figure - as vectors, a box with no fill showing before/after positions, positions of partial annealing/retention zones before and after uplift, etc. For Y-axis for Figs. 4 and 5, I strongly recommend flipping the axis such that 0 km is at the surface, and values increase with depth (or negative elevation). This recommendation uses the same perspective as structural cross sections and other geodynamic models, and I think would make both thermal structure and predicted cooling age plots more intuitive for readers and potential GDTchron package users.

We will invert the Y-axes in these figures as suggested so that the values reflect depth. We will also add horizontal lines to each plot showing the location of the model surface at the final timestep, which will illustrate when/how much rock uplift is taking place.

Lines 194-196, sentence starting with "One could explore the effects…": Predicted cooling ages for all of these scenarios also depend on the mechanism for incorporating or approximating topographic evolution – which may have just as much impact on the distribution of predicted cooling ages as the aforementioned factors, especially for low-T thermochron data in regions of

relatively slow rock uplift/exhumation (Gilmore et al., 2018; Ketcham, 2025; & references therein).

We agree and will make this point clear both here and in some of the earlier parts of the text where the reviewer suggests emphasizing the importance of the choice of surface processes models.

Lines 200-204: Completely agreed – though I think it would be helpful to articulate how these factors are incorporated into geodynamic models to help educate readers and ensure potential GDTchron users make effective choices when coupling geodynamic models and interpretations of thermochronology data.

We will add to the subsequent paragraph some details of how these features are incorporated into the more complex rift-inversion model. In general, there is considerable variation in how these factors are used across geodynamic models, so we will not be able to describe all possible permutations.

Lines 214-215, sentence starting with "Although thermokinematic models…": Consider including a brief discussion of the similarities & differences of each approach, e.g., Pecube versus GDTchron? What is each good at (e.g., Pecube incorporates a sophisticated framework for estimating topographic evolution, Pecube-D tracks thermal histories of particles involving large amounts of horizontal shortening/extension and complex faulting, statement on how these relate to the conditions of most geodynamic models?)

As described in the response to "Specific Comment" 3, we will be happy to elaborate on the differences between geodynamic and thermokinematic models. As suggested, Pecube does an excellent job of handling topographic evolution and complex prescribed fault systems, whereas geodynamic models can test the effects of a broader range of parameters but in a more generalized way.

Line 220 / Fig. 6: Please clarify that ages are relative to the start time/duration of the model. Consider flipping Y axes for both the model grid (so that 0 km is at Earth's surface and depth/negative elevation increase downwards) and the Surface Ages plots (so that 0 Ma is at the bottom, and older ages are on top - this is the framework commonly used for Pecube-D and other thermokinematic models). What do you mean by "Maximum Age (Ma)" as Y-axis label on bottom charts? Seems like these are just predicted cooling ages?

We will invert the Y-axes of the requested figures in the revised manuscript. We will relabel the Y-axis on the bottom row to be "Age (Ma)" to avoid confusion – the "Maximum Age (Ma)" was included because we use the greatest age at each X-value in the model results as a proxy for the surface age.

Please consider adding some vertical exaggeration to the middle plots, and perhaps intermediate ticks/values to the Plastic Strain and AHe Age gradient keys – predicted cooling ages only use a third of the space, and it is challenging to read these gradient values without zooming in 200+% into the figure.

We will add additional ticks to the colormaps for strain and AHe age. Although we understand that vertical exaggeration would make the middle plots easier to read, it would also distort attempts to compare ages with the locations and geometries of major faults/shear zones (indicated by plastic strain). Our goal with the bottom plots is to make the gradients that may be challenging to see at the surface in the middle plots more readable.

Line 229, sentence starting with "This 2D model…": Which model? That in Vasey et al. (2024) using hillslope diffusion to approximate surface processes, or the revised model in this study that uses Fastscape? Here, it would be valuable to use this example, and the changes made to "provide more realistic surface processes", as a framework for discussing what changes were made to the mechanisms of approximating topographic evolution and their influence on predicted cooling ages.

The parameters described in this sentence apply to both the Vasey et al. (2024) model and the Fastscape-coupled model presented here. We will revise to make this explicit and provide more detail on the implementation of surface processes.

Line 231, sentence starting with "At 16 Myr…": Time reference is unclear here - do you mean after 16 Myr into the model run, or at 16 Ma? At the end of the sentence, it would be helpful to add "… with a total model run time of 36 Myr" or similar.

In general, we are explicit about using Myr to refer to model runtime and Ma to refer to ages. As suggested by Reviewer 2, we will change all figures and text to use Ma to avoid confusion.

Lines 236-237, "resulting in slightly younger AHe (~10 Ma) and AFT (~14 Ma) ages…": Here, can also emphasize that these predicted ages are partially reset with respect to inherited model age?

This is a good point, and we will make this point explicitly in the revised manuscript.

Line 240: Update "…lithospheric thickening and uplift" to "lithospheric thickening, uplift, and erosion"

We will make this change in the revised manuscript.

Lines 259-261, "Implementing these alternative kinetic models…": Why not incorporate the most widely used kinetic models in this first release, especially if kinetic models are styled after those used in Ketcham (2005)? Doing so would significantly expand the potential for comparisons between predictions based on geodynamic models and thermochronology data for specific study areas.

We agree that including these kinetic models would be useful, though we do not think this initial release is the best venue to do so. See our response to "Specific Comment" 2 for details.

Lines 263-265, "Importantly, the ages produced when applying outputs from geodynamic models are dependent on parameters that would be highly variable among samples within real systems (e.g., U and Th concentrations, grain size, Dpar)…": Here, viable solutions might include incorporating values based on best practices (e.g., for grain sizes for He analyses) or values based on standards, then allowing for users to adjust GDTchron sample parameters to best match thermochronology data of interest? This suggestion is inspired by similar options in thermokinematic modeling programs including PecubeGUI (Bernard et al., 2025).

Currently, GDTchron does allow the user to modify U and Th concentrations, grain size, and Dpar from default values for a particular model run, and we will make that explicit in the revised text.

References not included in the manuscript:
Almendral, A., Robles, W., Parra, M., Mora, A., Ketcham, R. A., & Raghib, M. (2015). FetKin: Coupling kinematic restorations and temperature to predict thrusting, exhumation histories, and thermochronometric ages. AAPG Bulletin, 99(8), 1557-1573.

Bernard, M., van der Beek, P., Colleps, C., & Amalberti, J. (2022, May). PecubeGUI: a new graphical user interface for Pecube, introduction and sample-specific predictions of apatite (U-Th)/He and 4He/3He data in the Rhone valley, Switzerland. In EGU General Assembly Conference Abstracts (pp. EGU22-2277)., 524–525, 1–28, https://doi.org/10.1016/j.tecto.2011.12.035, 2012.

Gilmore, M. E., McQuarrie, N., Eizenhöfer, P. R., & Ehlers, T. A. (2018). Testing the effects of topography, geometry, and kinematics on modeled thermochronometer cooling ages in the eastern Bhutan Himalaya. Solid Earth, 9(3), 599-627.

Ketcham, R. A. (2025). Incorporating topographic deflection effects into thermal history modelling. Geochronology, 7(3), 449-458.