



This paper is published under the terms of the CC-BY-NC license.

© 2022 The Authors

Thermal history modeling techniques and interpretation strategies: Applications using HeFTy

Kendra E. Murray¹, Andrea L. Stevens Goddard², Alyssa L. Abbey³, and Mark Wildman⁴

¹Department of Geosciences, Idaho State University, 921 South 8th Avenue, Pocatello, Idaho 83209, USA

²Department of Earth and Atmospheric Sciences, Indiana University, 1001 E 10th St., Bloomington, Indiana 47408, USA

³Department of Geological Sciences, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, California 90840, USA

⁴School of Geographical and Earth Sciences, University of Glasgow, 8NN, University Avenue, Glasgow G12 8QQ, UK

ABSTRACT

Advances in low-temperature thermochronology, and the wide range of geologic problems that it is used to investigate, have prompted the routine use of thermal history (time-temperature, tT) models to quantitatively explore and evaluate rock cooling ages. As a result, studies that investigate topics ranging from Proterozoic tectonics to Pleistocene erosion now commonly require a substantial numerical modeling effort that combines the empirical understanding of chronometer thermochemical behavior (kinetics) with independent knowledge or hypotheses about a study area's geologic history (geologic constraints). Although relatively user-friendly programs, such as HeFTy and QTQt, are available to facilitate thermal history modeling, there is a critical need to provide the geoscience community with more accessible entry points for using these tools. This contribution addresses this need by offering an explicit discussion of modeling strategies in the program HeFTy. Using both synthetic data and real examples, we illustrate the opportunities and limitations of thermal history modeling. We highlight the importance of testing the sensitivity of model results to model design choices and describe a strategy for classifying model results that we call the Path Family Approach. More broadly, we demonstrate how HeFTy can be used to build an intuitive understanding of the thermochronologic data types and model design strategies that are capable of discriminating among geologic hypotheses.

INTRODUCTION

Thermochronology is used by researchers across the geosciences to measure the thermal history of rocks and thereby address questions that range from faulting and rock deformation, to climate-driven erosion, to lithosphere-scale changes in the geothermal gradient. Along with methodological advances in producing and interpreting thermochronologic data sets (e.g., Ault et al., 2019; Gautheron and Zeitler, 2020, and references therein), thermal history (i.e., time-temperature, tT) modeling has emerged as an important and, in

many cases, required step in thermochronologic data interpretation (Fox and Shuster, 2020). The need for quantitative analysis of rock cooling ages has prompted the development and dissemination of modeling software programs (e.g., Ketcham et al., 2000; Ketcham, 2005; Braun, 2003; Zeitler, 2004; Hager and Stockli, 2009; Gallagher, 1995, 2012; Fox et al., 2014), including the commonly used HeFTy (Ketcham, 2005) and QTQt (Gallagher, 2012) programs (Fig. 1). As an increasingly broad range of geoscientists generate thermochronologic data sets, they join the ranks of thermal history modelers but commonly lack experience with the non-uniqueness of rock cooling ages and the tools used to quantitatively interpret thermochronologic data.

Thermal history models are critical tools for interpreting cooling ages because individual ages, calculated from the measured abundance of parent radionuclides (e.g., U and Th) and their daughter products (e.g., ⁴He and fission tracks), are not very meaningful unless considered in a geologic context. Unlike geochronologic ages, single thermochronologic ages are non-unique because the retention of the daughter products, and thus the age, is controlled by a sample's net temperature history. As such, many different thermal histories can result in the same measured age. This was demonstrated early in the development of U-Th/He thermochronology by Wolf et al. (1998), who

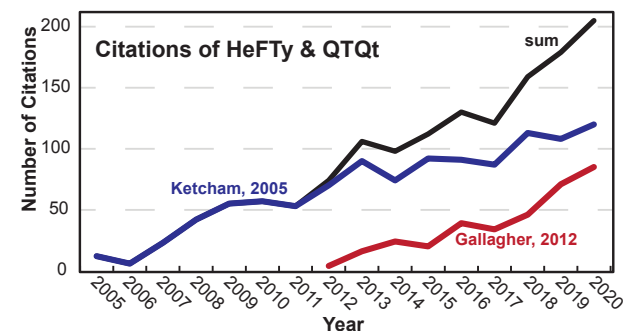


Figure 1. Illustration of the steady increase in citations of the two most commonly used thermal history modeling programs, HeFTy (Ketcham, 2005) and QTQt (Gallagher, 2012), from 2005–2020.

Kendra Murray <https://orcid.org/0000-0003-4008-1645>

presented five very different tT histories that all yield a 40 Ma apatite U-Th/He (AHe) age (Fig. 2A). Wolf et al. (1998) emphasized the need to use multiple samples or thermochronometric systems to reconstruct rock cooling histories, because additional data narrow the range of possible tT histories. The range of possible tT histories can also be narrowed by what is independently known about a sample's history. Thermal history models bring all of this information together; they are the numerical tools we use to interpret thermochronologic data in the context of the temperature sensitivity of thermochronometers (i.e., kinetics derived from experimental observations) and independent knowledge or hypotheses about a study area's geologic history (Fox and Shuster, 2020). Advances in our understanding of additional factors that control individual cooling ages, such as radiation damage (e.g., Flowers et al., 2009) and crystal composition (Carlson et al., 1999) as well as fission-track etching and annealing (Ketcham and Tamer, 2021; Aslanian et al., 2022), have only added to the need for numerical tools to guide the assessment and interpretation of thermochronologic data sets.

Given the integral role tT models play in modern thermochronologic interpretations (Fox and Shuster, 2020; McDannell and Flowers, 2020), it is important for the expanding user community to understand the opportunities and limitations of tT modeling and have access to information about modeling strategies and best practices. As has been discussed in recent literature, any model offers only an imperfect representation of the data, and inverse tT models are commonly "ill-posed" such that modelers must make many choices while attempting to retrieve meaningful geologic information from thermochronologic data (Fox and Carter, 2020). Indeed, many debates that hinge on thermochronologic data are in fact disagreements about modeling choices and not geologic processes or histories (e.g., Fox and Carter, 2020; Flowers et al., 2015). Thus, although there is general agreement about the

importance of conducting tT modeling with careful consideration of thermochronometer kinetics, inversion algorithms, statistical methods, model design, data input, and geologic context, exactly how this should be done is extensively debated (e.g., Vermeesch and Tian, 2014, 2018, 2020; Gallagher and Ketcham, 2018, 2020; Green and Duddy, 2021). Although valuable recommendations have been made about the reporting of modeling results (Flowers et al., 2015, 2016, 2022; Gallagher, 2016), most thermochronologic studies do not provide explicit descriptions of their modeling practices and interpretation strategies. The absence of these details in the published literature means that many new and experienced modelers must reconstruct or reinvent modeling and interpretation strategies for themselves. It also relegates meaningful discussion of modeling practices to relatively inaccessible spaces, such as the manuscript review process or the occasional short course.

Here, we aim to fill this resource gap by presenting a series of concrete examples, accompanied by guided exercises in the Supplemental Materials¹, that demonstrate the fundamentals of thermochronometer behavior and present a range of thermal history modeling activities in the HeFTy program (Ketcham, 2005). We use synthetic data and generalized examples from our own work in order to focus this contribution on modeling methods and to not call out specific study areas. Although there is no one "right" way to model data, there are best practices that are common among thermal history modeling approaches and programs. We briefly summarize the features that characterize robust modeling practices in HeFTy that are presented in this contribution as well as in previous publications (Ketcham, 2005; Flowers et al.,

¹Supplemental Material. HeFTy modeling tutorial, Tables S1–S4, Figure S1. Please visit <https://doi.org/10.1130/GEOS.S.19991567> to access the supplemental material, and contact editing@geosociety.org with any questions.

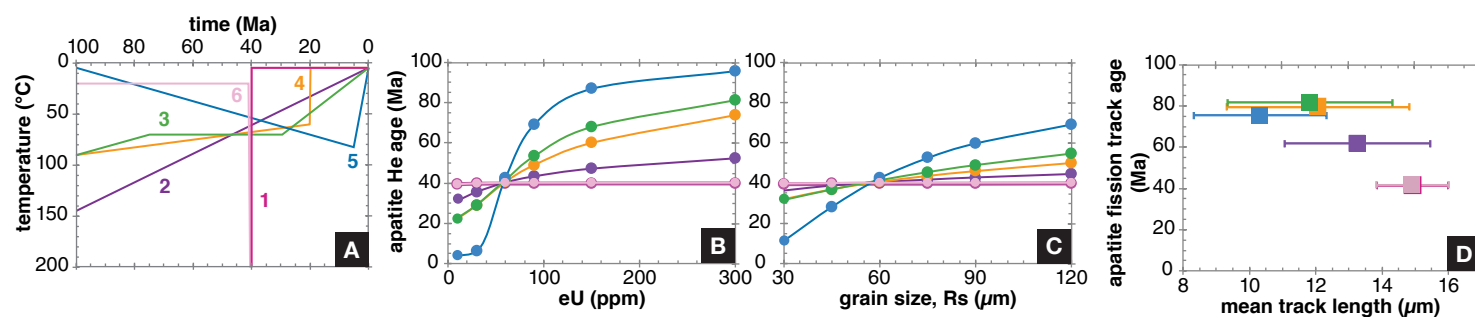


Figure 2. (A) Six tT paths (adapted Wolf et al., 1998) represent thermal histories associated with common geologic settings, including volcanic cooling (Path 1), craton erosion (Path 2), craton rifting (Path 3), thrust belt erosion (Path 4), basin burial and exhumation (Path 5), and transient localized heating next to a near-surface igneous intrusion (Path 6). See Table 1. For each path, an apatite grain with an equivalent spherical radius (R_s) of 60 μm and parent nuclide composition ([eU]) of 60 ppm will record a He cooling age of ca. 40 Ma using the radiation damage accumulation and annealing model (RDAAM) kinetics (Flowers et al., 2009). Equivalent thermal histories could be designed to illustrate the non-unique behavior of other thermochronometers. (B) Predicted apatite He age-[eU] trends for the six paths. This [eU] range over-represents the variability in a typical sample. Apatite grains with [eU] as high as 150 ppm and 300 ppm are uncommon, and most grains in a typical sample have [eU] between 10 ppm and 90 ppm. (C) Predicted apatite He age versus grain size trends for the six paths. (D) Predicted AFT age versus mean track length relationship for $D_{\text{par}} = 2.05 \mu\text{m}$.

2015; Fox and Carter, 2020; Fox and Shuster, 2020). We also refer the interested reader to our forthcoming companion paper (Abbey et al., 2022) that aims to fill the same gap using the QTQt program (Gallagher, 2012).

■ FORWARD AND INVERSE MODELING OF SYNTHETIC DATA: EXPLORING DATA UNIQUENESS AND MODEL RESOLUTION

We demonstrate fundamental aspects of using thermochronologic data to reconstruct a rock’s “true” time-temperature history in HeFTy with both forward and inverse models of synthetic data. Following the lead of Wolf et al. (1998), we use six different thermal history paths (Fig. 2A) that each yield a 40 Ma He age for an apatite grain of specific size and composition.

We also expand upon Wolf et al. (1998) by modeling the AHe data trends that arise from one well-documented source of He age variability that is built into the current kinetic models (Flowers et al., 2009; Gautheron et al., 2009; Willett et al., 2017) and commonly leveraged in tT history models: radiation damage accumulation and annealing. Radiation damage produced during the decay of U, Th, and Sm substantially impacts the mobility of He and thereby the measured He age. The extent of this impact depends on the abundance of damage in a grain, which is controlled by the grain’s U-Th-Sm composition (commonly combined into one parameter called effective U, or [eU] = [U] + 0.234[Th] + 0.0047[Sm]) and the amount of time a grain has been accumulating damage. Radiation damage abundance is also a function of the grain’s thermal history because it heals (i.e., anneals) at high temperatures. These effects manifest in AHe data sets as positive-slope relationships between AHe age and [eU] (Fig. 2B) because for some—but not all—types of thermal histories, higher [eU] apatite grains accumulate more radiation damage, and as a result they retain more He and have older measured ages than lower [eU] grains that experienced the same thermal history. Programs like HeFTy implement various radiation damage accumulation and annealing models and are

thereby practical tools not only for interpreting data but also for training one’s intuition about how these convoluted effects should manifest in real data sets.

Our approach of using synthetic data is deceptively simple, for it simultaneously offers an accessible entry-point for using HeFTy and illustrates the central features of thermal history modeling that form the foundation of the rest of this paper. As a complement to this text, in the Supplemental Material we provide a set of activities useful to readers interested in replicating and exploring these exercises for themselves.

Forward Modeling

The forward modeling capabilities of HeFTy enable the user to predict cooling age(s) for the apatite fission track (AFT) and zircon fission track (ZFT) systems and zircon (ZHe) and apatite (AHe) (U-Th)/He systems and/or AFT lengths given a single specific tT path (Fig. 2; Table 1). These predictions can be compared with measured data to test hypothesized cooling histories and eliminate geologic scenarios that predict thermochronometric ages that are inconsistent with observations. Forward model predictions are also useful for learning and research planning. The user can, for example, test whether hypothesized rock tT histories can actually be distinguished from each other using thermochronology, evaluate how specific parameters (such as length distributions in the AFT system) or multi-chronometer approaches may more clearly constrain tT histories, or explore the consequences of choosing to use one set of published kinetics over another during thermal history modeling. Here, we illustrate the utility and limitations of forward modeling using six specific tT paths that are each consistent with a different geologic setting or process (Table 1, Fig. 2A). This exercise is useful for developing intuition about what types of thermal histories produce distinctive thermochronologic data trends and why.

We forward model six hypothetical tT paths that include rapid cooling, slow cooling, episodic cooling, or non-monotonic cooling over a 100 m.y. time

TABLE 1. SUMMARY OF THE SIX tT PATHS

		Path 1		Path 2		Path 3		Path 4		Path 5		Path 6	
Geologic process		Volcanic cooling		Erosion of a stable craton		Erosion of a stable craton followed by rifting		Erosion of thrust belt topography		Basin burial and exhumation		Local igneous intrusion and cooling	
Thermal history		Time	Temp	Time	Temp	Time	Temp	Time	Temp	Time	Temp	Time	Temp
		40	200	100	145	100	90	100	90	100	5	100	20
		39.9	5	0	5	20	60	75	70	5	82.5	41	20
		0	5			19.9	5	29.5	70	0	5	40.5	200
						0	5	0	5			40	5
												0	5
Predicted	AHe age (Ma)*	39.8		40.4		40.5		41.4		42.5		40.1	
	AFT age (Ma)**	41.3		61.8		79.5		81.6		75.3		41.6	
	AFT mTL (µm)**	14.92 ± 1.07		13.25 ± 2.19		12.06 ± 2.75		11.83 ± 2.49		10.32 ± 2.00		14.91 ± 1.09	

Note: *Rs = 60 µm, [eU] = 60 ppm. **D_{par} of 2.05 µm, etched in 5.5M HNO₃ at 20° C for 20 s, corrected for c-axis projection. AFT—apatite fission track.

period (Fig. 2A; Table 1) and are designed to demonstrate the temperature sensitivity of the AHe system (after Wolf et al., 1998). The AHe system has a bulk closure temperature of 80–40 °C (Flowers et al., 2009). By choosing a model start date of 100 Ma, we are assuming that the apatite grains formed 100 m.y. ago. Forward models for each of the six tT paths all predict the same AHe cooling age of ca. 40 Ma using the kinetics of Flowers et al. (2009) for a single 60 ppm [eU] apatite grain (i.e., a standard Durango-like composition) and a 60 μm radius (R_s , Fig. 2; Table 1).

Forward models of these six tT paths in HeFTy can also predict cooling ages for various other chronometric systems, additional observations such as track lengths (Fig. 2D), or He age trends that arise from well-documented sources of age variability such as grain size (Fig. 2C; Reiners and Farley, 2001) or U-Th composition (Fig. 2B; Flowers et al., 2009; Gautheron et al., 2009). In the examples below, we use forward models to predict He ages for six apatite grains with variable parent-nuclide concentration ([eU], 10–300 ppm) and compare the predicted AHe age variability produced from each of our six tT paths (Figs. 2–3). We also predict and examine AFT ages and lengths for each history.

Grain [eU] Composition Effects

As previously described, variations in grain [eU] can result in significant intra-sample He age variability for some types of thermal histories; for the AHe system, this variability takes the form of a positive-slope relationship between He age and [eU] (Flowers et al., 2009; Gautheron et al., 2009). To evaluate if variable grain [eU] can produce distinctive data trends that permit differentiation among our six hypothetical tT paths, we use forward models to predict the AHe ages of grains with a size of 60 μm and [eU] of 10 ppm, 30 ppm, 60 ppm, 90 ppm, 150 ppm, and 300 ppm (Fig. 2B). By design, grains with an [eU] of 60 ppm yield an AHe cooling age of ca. 40 Ma for each tT history. Paths 1 and 6 yield ca. 40 Ma ages for all grains regardless of composition (Fig. 2B). The AHe ages predicted for these paths have variability that is $\leq 15\%$ of the standard deviation of the mean age, which is well within the range that we define as “reproducible” for He thermochronometers (Flowers et al., 2015; see also Flowers et al., 2022). Path 2 yields ages of 31–52 Ma (18% variability). In contrast, the predicted AHe ages for Paths 3, 4, and 5 have 42%, 46%, and 78% variability; Path 5 yields He ages from 4.1 Ma to 96 Ma (Fig. 2B).

The predicted age-[eU] trends can be used to qualitatively distinguish among some, but not all, of the six paths considered here (Fig. 2B). For example, although the age-[eU] trend is similar for Paths 3 and 4, this data trend would clearly be inconsistent with the tT histories of Paths 1, 2, 5, and 6. Similarly, Paths 1, 2, and 6 have negligible age-[eU] trends (i.e., they produce “reproducible” ages), but such data sets are clearly inconsistent with the predicted positive age-[eU] trends for the tT histories of Paths 3, 4, and 5 (Fig. 2B).

More broadly, these forward model results concretely demonstrate the substantial impact radiation damage accumulation and annealing can have on AHe ages for particular thermal histories (Flowers et al., 2009; Gautheron

et al., 2009). Age variability and age-[eU] trends arising from radiation damage effects are most extreme in thermal histories during which (1) grains reside in the partial-retention zone (PRZ) for tens of millions of years (Paths 2–4), enabling differential accumulation of radiation damage and retained radiogenic He, or (2) grains are colder than the PRZ for tens of millions of years, differentially accumulating radiation damage, and are then reheated and partially reset with variable He retention as a function of each grain’s accumulated radiation damage. For Path 5, the lowest [eU] grain (10 ppm) has a predicted age of 4.1 Ma; this crystal was entirely degassed of He during peak reheating at ca. 5 Ma. In contrast, the highest [eU] grain has a predicted age of 96 Ma and thus retained nearly all of its He, despite experiencing the same reheating event. Higher reheating temperatures (or longer time at the peak temperature) would be required to reset the He content of the high-[eU] apatite grain; even more significant reheating would be required to anneal the radiation damage accumulated in these grains.

Predicted AFT Ages and Lengths

Although the six paths are tuned specifically to exploit the partial-retention behavior of the AHe system, we can use HeFTy to predict the ages of other low-temperature thermochronometers using these same paths and thus evaluate whether these other systems could help discriminate between these tT histories. For example, here we report the predicted ages and track lengths from the AFT system (Fig. 2D; Table 1), which has a bulk closure temperature of 140–90 °C (Green et al., 1989; Ketcham et al., 2007). We assume that the apatite grains have an average D_{par} of 2.05 μm and etched in 5.5M HNO_3 at 20 °C for 20 s. The data are corrected for c-axis projection.

The AFT data predicted using the six paths vary more than the AHe age versus grain-size trends and in a way that is similar to the age-[eU] trends (Fig. 2; Table 1). Like the predicted AHe ages for Paths 1 and 6, the predicted AFT ages and track lengths for these paths are indistinguishable (41.3 Ma and 41.6 Ma and 15.63 μm and 15.62 μm , respectively). Path 2 yields significantly older AFT age and shortened mean track lengths (61.8 Ma and 14.47 μm) than Path 1, so AFT analysis would likely be useful in distinguishing Path 2 from Paths 1 or 6 for a rock that had a limited range of apatite [eU] compositions (and thus limited apatite age variability as a function of either [eU] or grain size). In contrast, the AFT mean ages and lengths predicted for Paths 3 and 4 are similar (79.5 Ma and 81.6 Ma and 13.87 μm and 13.82 μm , respectively); thus, distinguishing between these paths is difficult even with the addition of AFT data.

Forward Modeling Take-Aways

A forward modeling approach has clear utility for exploring, learning about, and evaluating intra- and inter-thermochronometer cooling age variability. It is

a particularly valuable tool for assessing whether observed cooling age variability in a real data set is quantitatively consistent with these known sources of data “dispersion.” For example, in real He data sets, it is common for both grain size and [eU] to vary in the AHe and ZHe systems (e.g., Flowers and Kelley, 2011). In the AFT system, track annealing is dependent on chemical composition (Green et al., 1986; Carlson et al., 1999). Especially for rock thermal histories with reheating events or long residence times at conditions under which a system is partially retaining or annealing daughter products, the compound effects of natural grain variability can result in “dispersed” cooling ages that do not, on first appearance, have clear age trends. Such convoluted variability makes it difficult to qualitatively assess the source(s) of age variability on simple bivariate plots; however, constructing forward models is an effective way of evaluating these trends. Forward models can also serve as important guides for constructing inverse thermal history models, understanding inverse model results, and designing a geologic study using thermochronologic data sets.

What is the range of possible tT histories that could produce a specific He age-[eU] trend or a specific multi-chronometer data set? How much better constrained is a model result with the addition of the AFT data, or with an apatite [eU] range of 10–300 ppm versus 20–60 ppm? These questions are best explored using an inverse approach.

Inverse Modeling

Unlike forward modeling, inverse modeling takes data inputs and finds the range of tT histories that are consistent with them. An inverse thermal history model result is the collection of tT paths that all result in statistically significant fits to the input data, given the model's design. To demonstrate how data trends and model design choices impact HeFTy model results, we present a series of inversion models for synthetic AHe ages and age-[eU] trends predicted by the forward models of Paths 1–6 in the previous section (Fig. 3).

HeFTy Inverse Modeling Basics

In HeFTy, inverse model results are produced by proposing thousands to millions of tT paths, predicting cooling ages and other data (e.g., track-length distributions) for each path, and then comparing predicted data to each piece of input data using a goodness of fit (GOF) statistical test (Ketcham, 2005). For each attempted tT path, the results of each statistical test are considered together (Ketcham, 2005; Ketcham et al., 2009), and HeFTy accepts paths that pass specific statistical criteria as either “acceptable” or “good” fits. An acceptable-fit path has $\text{minimum}(\text{GOF}) > 0.05$. A good-fit path has both a $\text{mean}(\text{GOF}) = 0.5$ and $\text{minimum}(\text{GOF}) = 1/(N+1)$, where N is the number of statistical tests (i.e., data inputs; Ketcham et al., 2009). Inverse models that include multiple chronometers, track-length distributions, or a range of grain sizes or [eU] will generally be better constrained than those that lack these

data and/or features, as we demonstrate next, although more data inputs also offer more opportunities to fail the GOF test. More information and published discussions of these fit criteria can be found in Ketcham (2005); Ketcham et al. (2009); and Flowers et al. (2015).

Additionally, HeFTy users dictate an inverse model's exploration of time-temperature space by creating boxes through which tT paths must pass and prescribing the behavior of the tT paths between these boxes. When used deliberately, these boxes are one of the strengths of HeFTy because they narrow the range of tT histories generated by a non-learning Monte Carlo algorithm in a simple, explicit manner. HeFTy does not use an optimization algorithm to search for candidate thermal histories (Ketcham, 2005), unlike the Bayesian Markov Chain Monte Carlo approach employed by QTQt (Gallagher, 2012). The simple Monte Carlo approach makes HeFTy particularly well-suited to investigating complex or deep-time problems, when many degrees of freedom are useful for identifying the range of possible thermal histories. Unlike QTQt, HeFTy has no limit on the number of constraint boxes. Constraint boxes in HeFTy typically represent independent geologic information; for example, they can require a sample to be at surface temperatures at the time of an independently dated unconformity. Boxes can also be used to explore (or prescribe) a particular tT path behavior during testing of a hypothesis, for example, sample reheating at a specific time. For other important discussions of the role of constraint boxes in HeFTy, see Ketcham (2005), Vermeesch and Tian (2014, 2018, 2020), Gallagher and Ketcham (2018, 2020).

Inverse Model Design for a Simple Exploration of “Perfect” Data

Here, we illustrate how data trends and constraint boxes control HeFTy inversion results. Each inverse model has the same set up, which is deliberately designed so that the tT paths explore a very wide range of monotonic and non-monotonic cooling histories within the 100–0 Ma and 200–0 °C tT space. This requires two constraint boxes (Fig. 3A). Box 1 spans 100–95 Ma and 0–180 °C, which creates tT paths that begin above, within, and below the AHe PRZ. Box 2 spans 95–0.1 Ma and 0–200 °C; this large box is necessary because we are interested in considering non-monotonic cooling histories (i.e., Paths 5 and 6), and HeFTy will only generate tT paths that include reheating if the maximum temperature between boxes increases with time (i.e., in this case Box 1 has a maximum temperature of 180 °C and Box 2 has a maximum temperature of 200 °C) or the maximum temperature of two side-by-side boxes is the same. We prescribe a present-day surface temperature of $5^\circ \pm 5^\circ \text{C}$, which includes the present day temperature used to generate these synthetic data (5 °C). We choose “episodic” and “variable” path behaviors between the tT constraints to allow, but not require, sudden changes in path behavior and both heating or cooling of paths between constraints, which ensures that the model attempts the widest possible range of paths.

To input our synthetic data, we assigned a 10% error to each predicted He age. We set each inverse model to run until it found 100 good-fit tT paths

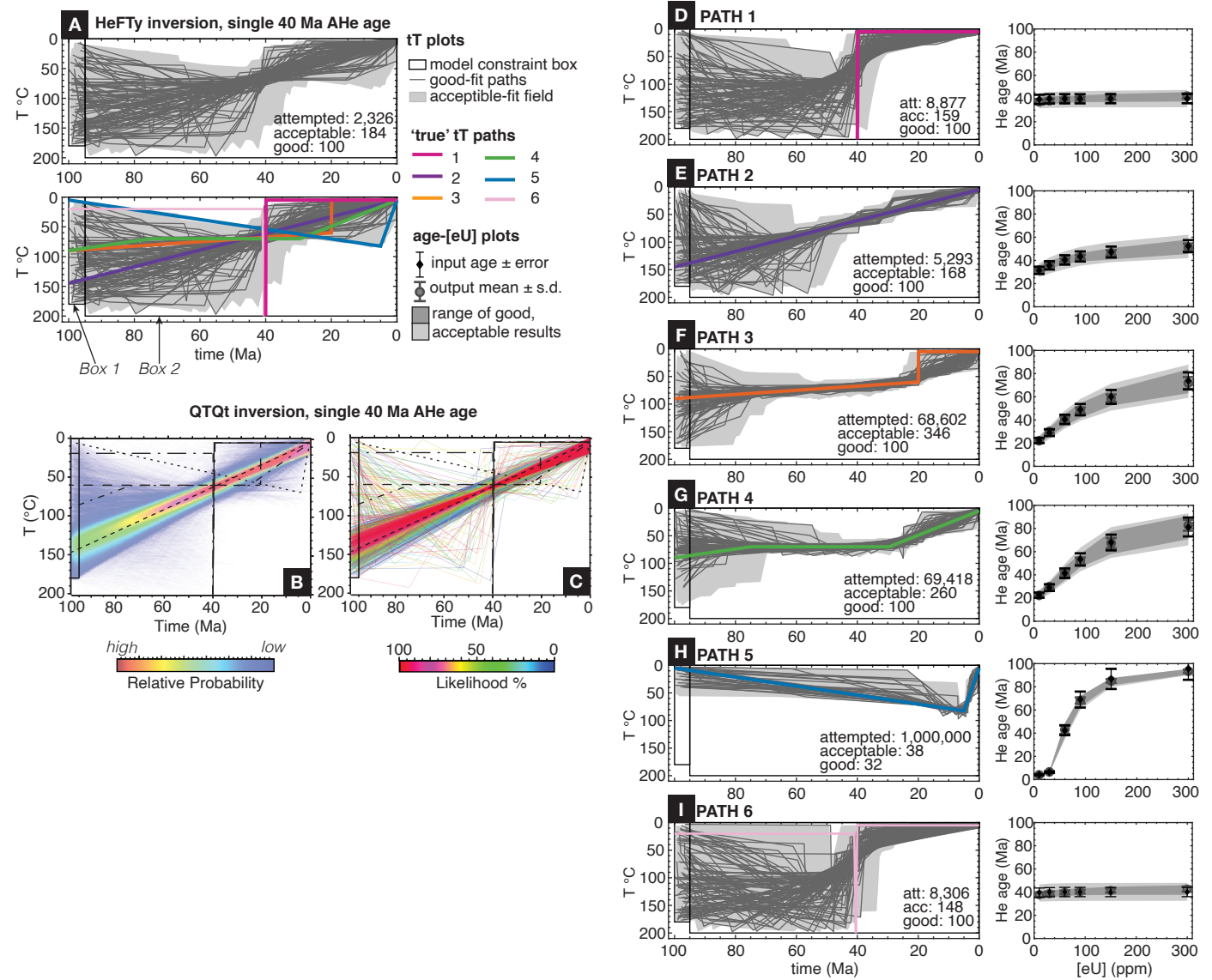


Figure 3. (A) Inversion result for a single 40 ± 4 Ma AHe age ($R_s = 60 \mu\text{m}$, $[\text{eU}] = 60 \text{ ppm}$); the “true” paths from Figure 2 (overlaid on the lower panel) are among many histories that yield a single 40 Ma age. (B) Expected thermal history and (C) maximum likelihood plots for the same 40 Ma AHe age modeled in QTQt demonstrate key differences in the modeling philosophy and model visualization of these two commonly used programs. (D–I) How well can an inverse model resolve six “true” thermal histories with synthetic apatite He age-[eU] trends (Fig. 2B), given identical inverse model designs in HeFTy? Left column shows inversion results for Paths 1–6; the path used to generate the synthetic data (i.e., the “true” path) is overlaid in color. The right column shows the age-[eU] relationship of the synthetic grains for each path and the model fit to those data.

so that model runs can be compared. The number of total attempted paths required to reach this end-condition ranges from <2500 to >1,000,000. We visualize the model results both in tT space and in age-[eU] trends (Fig. 3). We note that because He data inputs for these inverse models are synthetic, the model results reflect an ideal scenario in which one has “perfect” He data that span a remarkably wide range of [eU] and vary only as a function of [eU]. For readers interested in exploring this inverse modeling for themselves, we have provided a set of activities in the Supplemental Materials (see footnote 1) that replicates and expands upon this demonstration.

Inverse Model of a Single 40 Ma Apatite He Age

We first present the result of an inverse model of a single 40 Ma AHe age ($R_s = 60 \mu\text{m}$, [eU] = 60 ppm). This illustrates how the nonuniqueness of a single cooling age, together with our generic model design, looks in HeFTy (Fig. 3A).

As expected (Wolf et al., 1998), a very wide range of histories yields a single 40 Ma age, including the paths we forward modeled in the previous section (Figs. 2A and 3A). Because the inverse model lacks additional thermochronologic data (inputs) or geologic information (narrower constraint boxes or restricted path behavior), it does not constrain the thermal history prior to ca. 40 Ma. After ca. 40 Ma, the model result overall requires the grain to have been colder than 90 °C after ca. 35 Ma, and the good-fit paths support cooling to <70 °C (i.e., cooling to at or below partial-retention temperatures) by 35 Ma.

Good-fit tT paths that exactly follow Paths 5 and 6 are missing from this HeFTy inversion result (Fig. 3A); this is due to our model design. One simple consequence of many different paths yielding a single 40 Ma age is that 100 good-fit paths are found after only 2,326 attempts. Thus, all of the specific paths that yield a 40 Ma age will not be attempted during this short model run. Paths 5 and 6 are different from Paths 1–4, because to find a good fit, HeFTy needs to randomly generate a path nearly identical to these distinctive “true” paths. Very similar paths that are shifted only a few degrees or millions of years would not yield a 40 Ma age. The generic design of our constraint boxes is not optimized for exploring these specific histories, although it does not exclude them. As we discuss below, a Path 5-like tT history only accounts for ~0.01% of the attempted paths with this very generic model set-up. Importantly, this is not a statement about the likelihood that a sample would experience this tT history in nature.

A comparison of this HeFTy inversion (Fig. 3A) to a QTQt inversion of the same single 40 Ma AHe age (Figs. 3B–3C) illustrates several fundamental differences in the way these programs search tT space and present inversion results to the user. In contrast to the non-learning random Monte Carlo search of tT space that HeFTy employs, QTQt’s Bayesian transdimensional Markov Chain Monte Carlo algorithm prefers and thus converges on simple, Path 2-like solutions. As a result of this focused sampling of Path 2-like histories, the color map of relative probability presented in an expected thermal history plot—the QTQt visualization commonly used for geologic interpretations (Fig. 3B)—may suggest to someone unfamiliar with Bayesian modeling that a

Path 2-like history is preferred or most likely to fit this single 40 Ma age, which we know is not the case (Wolf et al., 1998). Indeed, individual tT paths similar to Paths 1–6 have >80% likelihoods, as is visualized in the maximum likelihood plot (Fig. 3C), which is similar to a HeFTy visualization (e.g., Fig. 3A). As with the HeFTy result, the small number of Path 5-like histories in the QTQt result reflects the way the program searches tT space—that is to say, the design of this particular inversion and the underlying philosophy of the modeling program. See also Fox and Carter (2020) and the commentaries related to Flowers et al. (2015) (Flowers et al., 2016; Gallagher, 2016) and Vermeesch and Tian (2014) (Vermeesch and Tian, 2018, 2020; Gallagher and Ketcham, 2018, 2020).

Inverse Models of Synthetic Age-[eU] Trends

Paths 1 and 6 involve very rapid cooling (>100 °C/m.y.) at ca. 40 Ma from >140 °C to 5 °C, which yields invariant 40 Ma AHe ages regardless of grain [eU] or the previous history of these grains. When this “flat” age-[eU] trend is an inverse model input, the thermal history prior to ca. 40 Ma is unconstrained by the model result (Figs. 3D and 3I). The thermal history after 40 Ma is better constrained, and it is also better constrained than the model of a single 40 Ma grain (Fig. 3A). The model result for Path 6 (Fig. 3I) does not include the exact “true” tT history because Path 6 is a very extreme reheating event at temperatures hotter than the closure temperature of the AHe system; consequently, HeFTy is unlikely to randomly generate such a “just right” path, given the constraint box configuration used here, even when set to explore “episodic and variable” histories.

Paths 2, 3, and 4 have slow monotonic cooling through the AHe PRZ, which produces positive-slope AHe age-[eU] trends (Figs. 3E–3G). This is because the longer a sample resides at partial-retention conditions, the more variable the AHe ages are. In each inverse model result, the early part of the tT history is poorly constrained, but it is still better constrained here than in inverse models of the Path 1 and Path 6 data that have no age-[eU] trend. The age-[eU] trends that arise from Paths 3 and 4 are very similar (ages of 22–74 Ma and 22–81 Ma, respectively; Fig. 2B), and thus the results of inverse modeling these trends, given the identical model design, are indistinguishable (Figs. 3F–3G). In other words, even with age variability that arises from radiation damage accumulation and annealing effects, it is not possible for our models to distinguish a history with rapid cooling from PRZ to surface conditions at 20 Ma (Path 3) from a history in which slow, steady cooling (2.2 °C/m.y.) from PRZ to surface conditions starts at 25 Ma (Path 4) (Fig. 2A); these are two thermal histories from which very different geologic interpretations would be made.

In Path 5, gradual heating from the surface to the PRZ prior to rapid cooling (15.5 °C/m.y.) during the last 5 m.y. produces an extreme and distinctive apatite He age-[eU] trend (Fig. 2B). The inverse model set-up for this exercise is not designed to efficiently find this tT history, and a narrow range of tT histories yields the observed age-[eU] trend, so therefore it takes tens of thousands of attempted paths to find a single good- or acceptable-fit path (Fig. 3H). Critically, this apparent difficulty in finding acceptable or good paths in HeFTy is

a useful indicator that the observed age-[eU] trend requires a distinctive tT history and may provide a narrow, and thus powerful, constraint on the tT history of the sample.

If available, independent geologic information or assumptions in the form of smaller constraint boxes can significantly narrow the range of attempted tT paths and improve the efficiency (i.e., optimize) of the search of tT space in HeFTy (Fig. 4; Table S1, see footnote 1). For example, if it were known from the geologic record that the Path 5 sample was at surface conditions

at 100–95 Ma, we could narrow the temperature range of Box 1–0–20 °C (Fig. 4A). All attempted tT paths would start at near-surface conditions and, because of the position of Box 2 relative to Box 1, most paths would explore a non-monotonic history. An inversion with this model set-up finds 100 good-fit paths after 906,473 attempted paths (Fig. 4A). In contrast, changing the maximum temperature of Box 2 from 200 °C (Fig. 3H) to 100 °C (Fig. 4B)—which could be justified given the observed 96 Ma AHe age—narrows the range of attempted paths but does not improve the efficiency of finding good-fit paths.

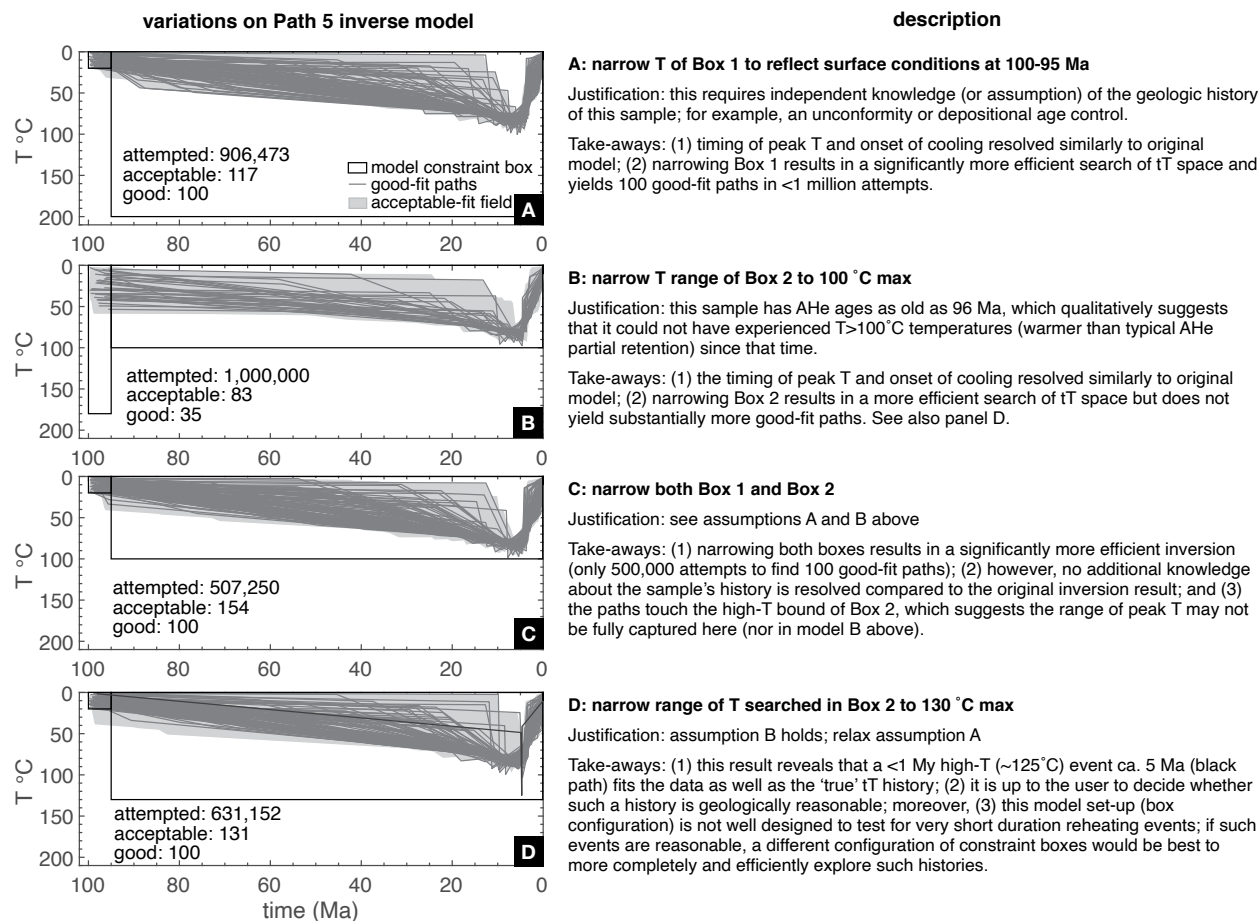


Figure 4. (A–D) Examples of testing the sensitivity of an inversion result (Fig. 3H) to the configuration of constraint boxes. In each case, the model result (i.e., the peak burial temperature and timing of the onset of final cooling at ca. 5 Ma) is similar. This analysis suggests that so long as it is not important to consider (D) very short duration (<1 m.y.) high-temperature thermal events, the model result is not sensitive to the design of the constraint boxes; moreover, narrowing the constraint boxes can optimize the time it takes to find good fits to the data (Ketchum, 2005) without excluding possible alternative families of tT paths. However, it is always important to clearly articulate the assumption(s) and/or justification(s) inherent in a particular constraint box design.

If both narrow constraints, along with their corresponding assumptions, are combined (Fig. 4C), the inversion is substantially more efficient at finding good-fit paths. However, the “true” Path 5 is equally well resolved in each case (Figs. 3H and 4A–4C).

In Path 5 models with a narrower Box 2 (Figs. 4B–4C), several of the good- and acceptable-fit tT paths touch the high-temperature (100 °C) bound on this box, which suggests that the box is limiting the full range of tT paths that fit the data. Indeed, if the maximum temperature bound is 130 °C (Fig. 4D), the model randomly finds a good-fit tT path with a geologically rapid (<1 m.y.), high-temperature (125 °C) heating event at ca. 5 Ma. This path's fit to the age-[eU] trend predicted from Path 5 reflects the fundamental trade-off between rock cooling rate and a thermochronometer's effective closure temperature (Dodson, 1973). The details of this good-fit path, as with Path 6 (Fig. 3I), had to be just right to fit the data. Whether it is important for a thermal history model to explore such histories depends on the sample's geologic context. In fact, HeFTy is useful for modeling very short-duration thermal events (e.g., Min and Reiners, 2007), but effective consideration of geologic histories that include such events requires a deliberate search of tT space using constraint boxes.

As with all model design choices, optimizing the boxes and the number of attempted or good-fit paths in HeFTy ultimately depends on a model's purpose and geologic context.

Inverse Modeling Take-Aways

HeFTy inverse models serve the user best when they offer a comprehensive range of possible tT paths that fit a given set of thermochronologic data and geologic constraints. It is then up to the model user to consider which histories are reasonable, given the geologic setting and the scientific questions of interest. However, as we have begun to demonstrate in this simple exercise with synthetic data, inverse model results can be extraordinarily sensitive to details of the model set-up in ways that may not be clear to the user—even the experienced user—without systematic exploration of alternative models. This requires an iterative modeling approach that documents, justifies, and explains decisions that affect both model design and thermal history interpretation.

We refer readers to the Supplemental Material (see footnote 1) for complementary and additional guided exploration of these simple synthetic examples in HeFTy. Indeed, we suggest that everyone should perform a modeling exercise with synthetic data, whether in HeFTy or another program of choice, before modeling real data for the first time.

STRATEGIES FOR AND EXAMPLES OF THERMAL HISTORY MODELING USING HeFTy

Next, we present a range of model design and interpretation strategies in HeFTy, using the synthetic data generated from Path 5 and examples from

real (measured) thermochronologic data sets from our own work. The details of these examples are tailored to HeFTy's approach to searching tT space, to the fact that HeFTy inversion results are suites of individual tT histories that each fit the data, and to the important role of constraint boxes in controlling HeFTy model results. However, many of the approaches are applicable to other thermal history modeling programs and support other recent discussions of thermal history modeling practices (e.g., Fox and Carter, 2020; Flowers et al., 2022). We also emphasize that there are many appropriate approaches to handling thermal history modeling results; so, one's preferred method will depend on the data and modeling tools available as well as the geologic problem(s) of interest.

Examples of Sensitivity Testing

Whenever a thermal history model resolves a tT history that appears useful for geologic interpretation, we argue that the next step is to clearly identify “the why.” That is to say, the reason(s) that a model produces a particular result, be it the power of a geologic constraint, a data trend, a spatial relationship among samples, the choice of kinetic model, etc., should be articulated. Thermal history modeling programs like HeFTy are tools for not only generating preferred model results that are geologically interpretable but also for systematically testing the sensitivity of the results to model design and data inputs. Here, we illustrate this sensitivity testing approach in HeFTy using the synthetic AHe data predicted from Path 5 (Fig. 3H) and then a real data set modeled in deep time.

When a model result is only sensitive to parts of the model design that are supported by geologic observations or assumptions that are clearly articulated and justified, we consider that result robust. In other words, an arbitrary choice should not significantly change the result, and modeling choices that do impact the result should be identified and explained. Although sensitivity testing is most efficient when guided by specific ideas about the factors that likely control the model result, this part of a modeling project—even for experienced modelers—is commonly an iterative learning process because of the complex, non-unique nature of thermochronometer behavior and how this manifests on geologic timescales.

Sensitivity Testing an Inverse Model Result

Path 5 offers an excellent synthetic example of the power of sensitivity-testing the result of an inverse model in HeFTy. As previously discussed, Path 5's simple burial history produces a distinctive age-[eU] trend ([eU] = 10–300 ppm, AHe = 4–96 Ma; Fig. 2B) that when modeled resolves the “true” tT history—particularly the requirement of reheating, the peak reheating temperature, and the timing of cooling to surface conditions—remarkably well (Fig. 3H). We use HeFTy to investigate how robust this model result is using suites of additional models in which we vary the model design and data input.

To assess the robustness of the Path 5 inverse model result (Fig. 3H), as if we did not know the “true” tT history, we first test its sensitivity to the inclusion of individual AHe ages. Systematically removing the lowest and highest [eU] grains reveals that the model’s ability to resolve the “true” path is strongly controlled by young, low-[eU] grains (Figs. 5A–5C). Indeed, if the two low-[eU] grains are removed (Fig. 5C), the model is unable to resolve any details of the

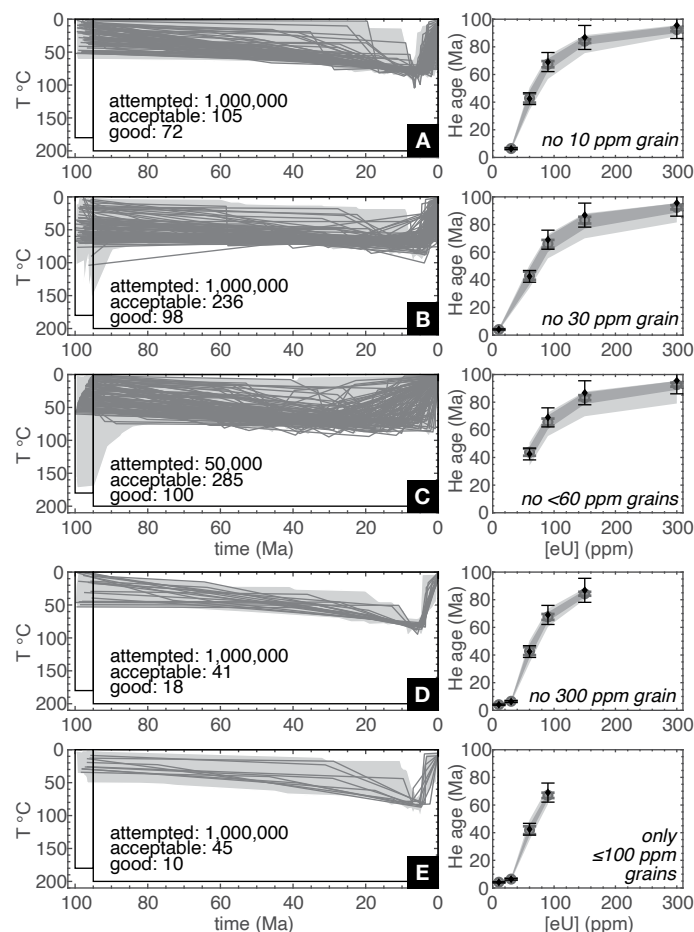


Figure 5. Systematically removing individual data inputs from a model permits the user to explore which parts of a data set most impact the model result. Here, we systematically remove apatite He grain ages from the age-[eU] trend generated from Path 5, remodel the remaining data, and compare the result to Figure 3H. (A) Without the lowest [eU] grain, the timing of cooling onset is poorly resolved. (B) Without the 30 ppm grain, both the heating and cooling history are poorly resolved. (C) When both low-[eU] grains are removed, the remaining data are not sensitive to the burial history. (D–E) In contrast, removing the high-[eU] grains does not substantially change the model result.

thermal history, except a requirement that the sample must have been colder than ~90 °C since ca. 80 Ma. Less intuitively, it appears to be the 6 Ma, 30 ppm [eU] grain (Fig. 5B), and not the youngest, lowest [eU] apatite grain (Fig. 5A), that is most important for documenting the “true” timing of peak heating at ca. 5 Ma. In contrast, when the oldest, highest [eU] grains are removed from the model (Figs. 5D–5E), reheating is still required to fit the remaining data, and the model result is largely unchanged. If we were interpreting a real data set with an unknown tT history, these additional model results would guide our assessment of the robustness of the original result. For example, we could (1) articulate why the apatite radiation damage model (Flowers et al., 2009) is most sensitive to the lowest [eU] apatite data and consider if alternative kinetics for the AHe system (Gautheron et al., 2009; Willett et al., 2017) would yield a different fit to the data, or (2) interrogate the quality of the lowest [eU] grain analyses (Peyton et al., 2012; Murray et al., 2014), or (3) offer a clear explanation for why some rocks from a field area preserve this reheating signal, while others do not, because samples with no low-[eU] apatite grains would not document this reheating event even if they experienced it.

A second useful test is to remove Box 2, which permits us to assess the requirement of reheating to produce the observed age-[eU] trend. Previously (see section above on Inverse Models of Synthetic Age-[eU] Trends), we examined the impact of changing the maximum temperature of Boxes 1 and 2 on the Path 5 model result (Fig. 4). Now, by removing Box 2 entirely, we ask HeFTy to assume only monotonic cooling from Box 1 to the modern surface temperature condition. This model design yields no good or acceptable fits to the Path 5 age-[eU] trend. This result may appear obvious here because we know that the “true” answer requires reheating. However, in the case of modeling an unknown sample, demonstrating this reheating requirement would offer valuable insight. In this way, sensitivity testing can also be used to reject alternative hypotheses or to demonstrate that two hypotheses are, in fact, distinguishable from one another using a particular data set and/or model design.

Removing Box 2 from the Path 5 model design, and finding no fits to the perfect synthetic data, also highlights a critical feature of HeFTy model results: when a HeFTy inversion does not return good or acceptable paths, even after hundreds of thousands of attempts, it does not necessarily mean that the data are of poor quality, too over-dispersed to be interpreted, or inconsistent with the kinetic model being used (Ketcham, 2005). Thus, it is critical to consider each piece of the model design and the assumptions inherent in those design choices. Systematically testing alternative model set-ups as a part of routine tT modeling practice is a useful way of revealing such assumptions.

Sensitivity Testing a Deep-Time Model

Testing the sensitivity of inverse model results is particularly important when investigating deep-time (i.e., >500 m.y.) thermal histories, because deep-time models are most successful when they leverage multiple chronometers and a large number of geologic constraints (McDannell and Flowers, 2020).

These model components are then integrated over hundreds of millions or billions of years, so it is remarkably easy to accumulate unintentional assumptions. HeFTy permits the user to impose an unlimited number of geologic constraints, so it is a useful tool for deep-time applications, when the resolving power of models in Precambrian time commonly leverages the fact that much of a sample's Phanerozoic tT history is already well known.

We illustrate a sensitivity-testing approach for evaluating thermal history models with an example of a deep-time model with three samples from Paleoproterozoic crystalline basement in the Colorado Rocky Mountains, which incorporates He and Ar data (Murray et al., 2022) in the context of multiple geologic constraints from local and regional field relationships (Table S2; see footnote 1). The model starts at 1.75 Ga and 800 °C, when the rocks formed (Premo and Fanning, 2000), and ends at modern surface conditions. Hornblende and biotite $^{40}\text{Ar}/^{39}\text{Ar}$ ages from nearby samples are 1450 Ma and 1350 Ma, respectively (Shaw et al., 1999). Zircon He ages from three samples form a negative-slope relationship between age and [eU] (Fig. 6A), which is qualitatively consistent with radiation damage accumulation and annealing in zircon (Guenther et al., 2013) and suggests that these data document a deep-time history.

To quantitatively test this assessment of the data and resolve parts of the Precambrian history of these rocks, we combine three samples into one thermal history model (Fig. 6A). By jointly modeling these data, we make the assumption that these samples share the same thermal history from 1.7 Ga to present. The model design uses high-temperature constraints derived from separate HeFTy models of the Ar ages (Figs. 7A and S1, see footnote 1). It also incorporates low-temperature constraints from local and regional field relationships, which place the sampled basement rocks in near-surface at three times prior to Cenozoic time (ca. 700 Ma, 500 Ma, and 300 Ma) and document their Paleozoic and Mesozoic burial history. Little is known about the Proterozoic history of these rocks, so we use exploration boxes to investigate a wide range of monotonic and non-monotonic histories from 1300 Ma to 700 Ma and 700–500 Ma. The resulting model (Fig. 6A, hereafter referred to as the preferred model result) requires a distinctive Neoproterozoic history, in which the rocks are heated to ~235–285 °C after ca. 700 Ma and cooled to near-surface conditions by ca. 500 Ma. But why is this?

First, we assess which parts of the ZHe age-[eU] trend the result is most sensitive to, because like in the Path 5 synthetic AHe data set, the age-[eU] trend is central to resolving the distinctive Neoproterozoic history (e.g., Orme et al., 2016). As is typical when modeling He data trends in HeFTy (Flowers and Kelley, 2011; Flowers et al., 2015), the 21 zircon He ages were grouped by [eU] and five average synthetic zircons were used as the HeFTy data inputs to capture the first-order age-[eU] relationship and smooth out the second-order age variability. This grouping and averaging is done at the discretion of the modeler, so it is important to ask how these choices impact the model result.

One way to address this question is to systematically remove the synthetic grains that represent the low, middle, and high [eU] parts of the age-[eU] trend (Figs. 6B–6D, respectively) and remodel the remaining data. This sensitivity

test demonstrates, for example, that the model result is not different when the middle-[eU] part of the ZHe trend (Fig. 6C) is removed. Thus, we would not expect modest changes in the choice of synthetic grain that represents this part of the trend to impact the results. Additionally, the models without the highest and lowest [eU] grains reveal that, counterintuitively, it is the young, ca. 70 Ma highest [eU] grains (Fig. 6D) and not the oldest Paleozoic–Proterozoic grains (Fig. 6B) that most strongly require the distinctive Neoproterozoic history. This gives us insight into how the zircon radiation damage accumulation and annealing model (ZrDAAM; Guenther et al., 2013) is fitting the age-[eU] trend.

The high-[eU] zircon grains are those that accumulate the most damage; if sufficient damage accumulates, the grains cannot retain He, and the ZrDAAM predicts a very young or 0 Ma age (Guenther et al., 2013; Guenther, 2021). In this data set, the high-[eU] grains do retain He and have measured He ages of ca. 70 Ma. Thus, these high-[eU] grains provide an upper bound on how much damage they have accumulated since 1.7 Ga and thereby the time that has elapsed since both the He content and radiation damage were completely reset, which is no more than 600–700 m.y. Forward models (Fig. 6E) support this assessment and further demonstrate that known heating during Cretaceous time cannot reset the highest [eU] grains without eliminating the age-[eU] pattern entirely. These forward models also explicitly test, and reject, an alternative hypothesized thermal history in which these rocks exhumed to surface conditions by ca. 717 Ma and then stay in the near surface until Phanerozoic time (i.e., Flowers et al., 2020). Thus, the one reason the preferred model resolves a heating event is that there is no more than 600–700 m.y. of radiation damage accumulated in the zircons from this sample, and therefore any damage accumulated before ca. 700 Ma had to be subsequently annealed during a significant thermal event. The ability of the current ZrDAAM to fit high-[eU] ages is an area of active debate (Guenther, 2021, and references therein), so the impact of the high-[eU] grains would be particularly important to note as a part of a geologic interpretation of this model result.

Next, we examine how the Ar data contribute to resolving the preferred deep-time history. Mesoproterozoic hornblende and biotite Ar ages require that these rocks have not reached temperatures sufficient to reset these systems partially or fully since Mesoproterozoic time and thus provide the maximum temperature bound on the exploration boxes (highlighted with orange in Fig. 7B). Although the Ar chronometers are not built into HeFTy, any diffusion-based chronometer can be modeled in HeFTy by manually inputting published or preferred kinetic values for a system.

By inverse modeling the He alone and Ar data alone, with no geologic constraints (Fig. 7A), we can investigate the independent contributions of each data set. The apatite and zircon He data alone do not resolve the Precambrian history of these rocks, whereas the Ar ages—specifically the ca. 1350 Ma biotite Ar ages—require that rocks have been colder than ~250 °C since ca. 1300 Ma, with heating events as warm as ~285 °C permissible on timescales of tens of millions of years (Fig. S1). These temperatures directly overlap with the temperature sensitivity of the zircon He system, which according to current kinetic models requires temperatures >~140–220 °C to reset the He content and

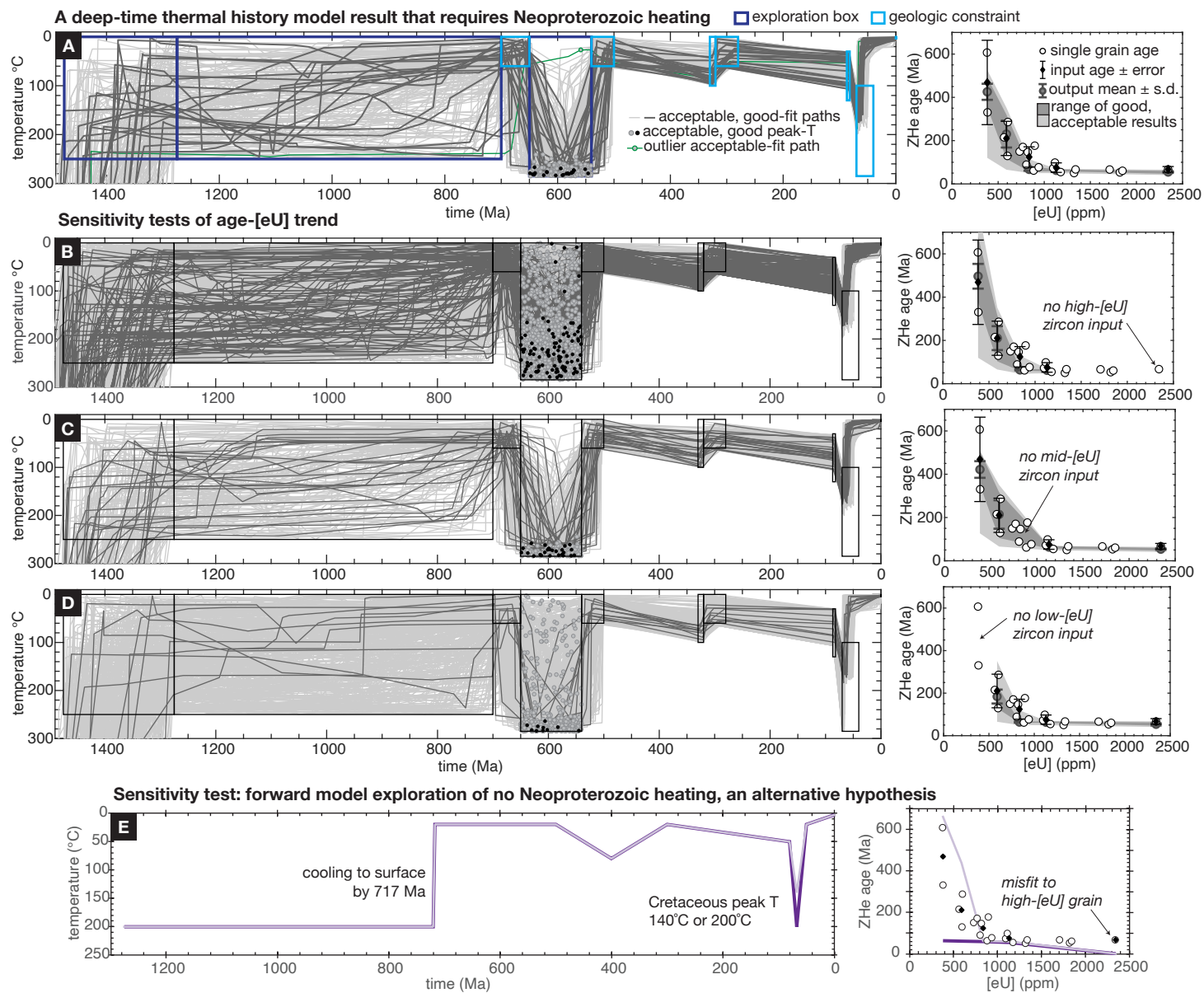


Figure 6. (A) An example of a deep-time thermal history model result, which requires a Neoproterozoic reheating event with a narrow range of peak T between ~235 °C and 285 °C. (B–D) The sensitivity of this model result to the zircon He data inputs that capture the first-order shape of the age-[eU] trend was tested. (E) Forward models of the zircon He age-[eU] trend confirm the surprising sensitivity of the deep-time record to the young, high-[eU], most damaged zircon grains and not the grains with Neoproterozoic–Cambrian He ages. Forward models also illustrate why these data do not support an alternative thermal history for these rocks (Flowers et al., 2020).

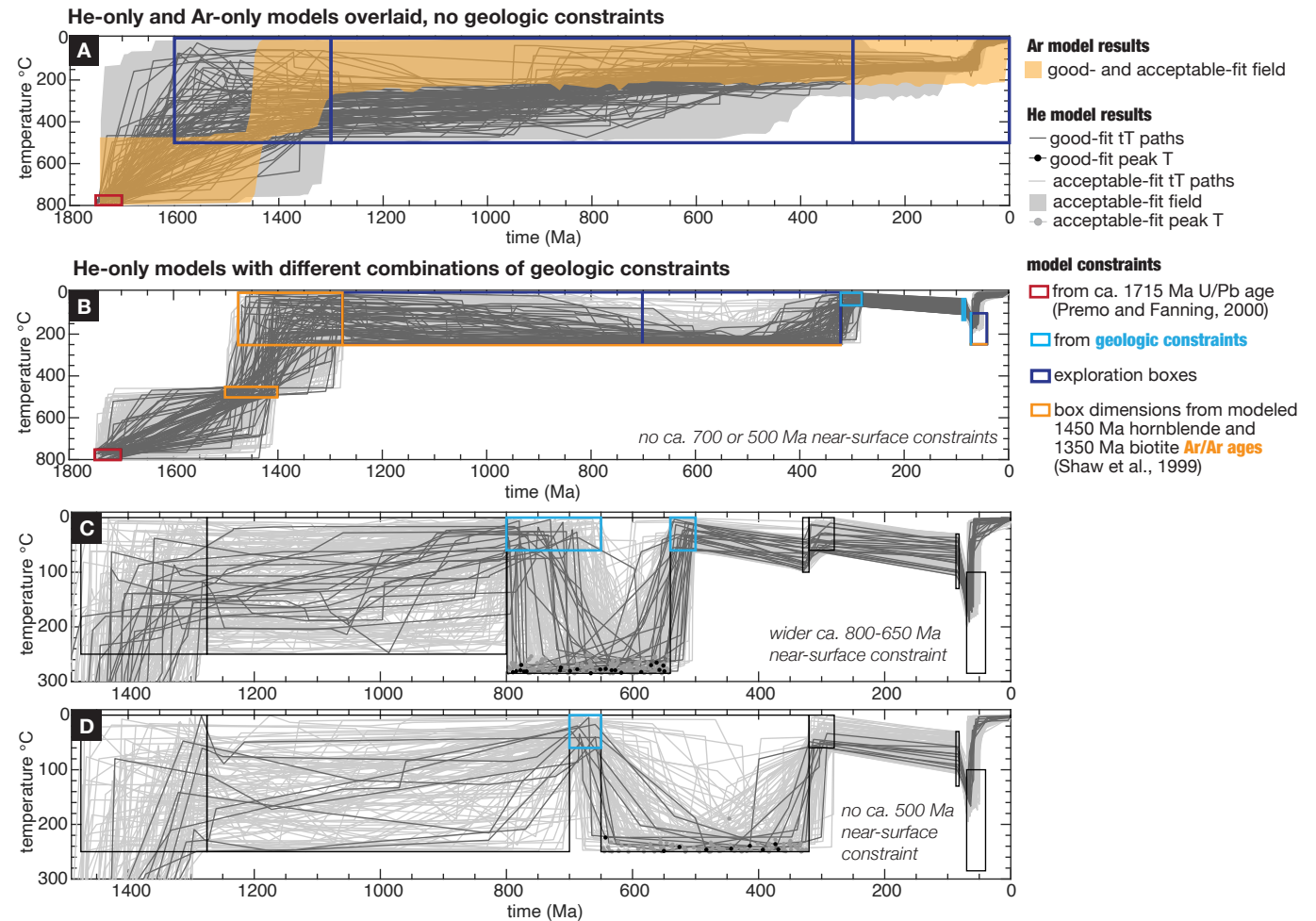


Figure 7. Additional sensitivity tests of the deep-time example. (A) Ar-only and He-only model results with no geologic constraints are overlaid to demonstrate what the data alone resolve about the thermal history of these rocks. (B) Example of an alternative He-only model (compare with Fig. 6A), with constraint boxes colored by their origins. (C–D) Additional examples of He-only models, which are similar to the preferred result (Fig. 6A) but have different implementations of the ca. 700 Ma and 500 Ma near-surface constraints, which are highlighted in blue boxes. These models demonstrate that although the required reheating temperatures are well constrained, the timing of this heating and cooling is poorly constrained. Sensitivity testing of tT constraints derived from the geologic record is important because other geologists may translate geologic information into tT space in different ways (for this data set, cf. Flowers et al., 2020).

temperatures $> \sim 260\text{--}350\text{ }^{\circ}\text{C}$ to reset the radiation damage content of zircon crystals, respectively (Guenther et al., 2013). Therefore, when considered together, the biotite Ar and zircon He chronometers narrow the range of possible deep-time thermal histories substantially (Figs. 7A–7B).

Thus, sensitivity testing has revealed that the narrow range of Neoproterozoic peak temperatures in the preferred model result is bounded by the Mesozoic biotite Ar age at high temperature ($< 285\text{ }^{\circ}\text{C}$) and the requirement of radiation damage annealing in zircon crystals at low temperature ($> 235\text{ }^{\circ}\text{C}$). In fact, an outlier acceptable-fit tT path in the preferred model result (Fig. 6A, green line), which does not require Neoproterozoic reheating, supports this relationship between the two chronometric systems. This path sits at $T = 245\text{ }^{\circ}\text{C}$ from ca. 1400–700 Ma; this history produces minimal radiation damage accumulation in zircon crystals and therefore does not require heating in Neoproterozoic time. However, it would also result in a biotite Ar age that is hundreds of millions of years younger than the measured age and is therefore not consistent with what is known about these rocks. The current kinetic models for the He and Ar systems require a narrow temperature range of reheating, but this may change as new kinetic models for these chronometers are produced or our understanding of damage annealing in the high-[eU] zircons improves (i.e., Guenther, 2021, and references therein).

Finally, we test the impact of the constraints derived from geological field relationships on the preferred model result (Fig. 6A) by designing a generic, no-geologic-constraints model (Fig. 7A) and systematically adding, removing, and adjusting how geologic information is implemented in tT space (e.g., Figs. 7B–7D). It is clear from this approach that the Neoproterozoic heating in the preferred model is not a product of the data alone (Fig. 7A). In particular, the timing of the heating required to reset the ZHe system depends on if and how the ca. 700 Ma and 500 Ma near-surface constraints are implemented in the model (Figs. 6A and 7B–7D). In all cases, the rocks reach a temperature of $\sim 250\text{ }^{\circ}\text{C}$ for tens of millions of years in Neoproterozoic–Paleozoic time because, as previously demonstrated, these rocks had to get hot enough for long enough to completely reset the ZHe system but not perturb the biotite Ar system. However, *when* heating and cooling occurred during this ca. 500 m.y. time period depends upon how narrowly the modeler defines the geologic constraints (compare, for example, Figs. 6A and 7D). This is a critical insight; it emphasizes the powerful role of independent geologic observations, and the boxes that represent them in tT space, in HeFTy model results. Without strong evidence for the timing of this heating and cooling, it could be difficult to interpret this model result and argue, for example, that heating and cooling is the result of one particular geologic process or event.

In summary, there are three principal reasons that the preferred deep-time model (Fig. 6A) requires a Neoproterozoic heating event. First, the high-[eU] zircon crystals retain He today, so the ZrDAAM requires that these crystals have been cold and accumulating radiation damage for no longer than $\sim 600\text{--}700$ m.y. (Figs. 6B–6E). Second, Mesoproterozoic biotite Ar ages require that the rocks have been cold enough to accumulate radiation damage in zircon crystals for much of the last 1300 m.y. (Fig. 7A). Third, the geologic record

places these rocks in the near-surface up to three times prior to Cenozoic time and for much of the Phanerozoic (Fig. 6A). These three conditions are reconciled by a reheating event that reset the He content and radiation damage in zircon crystals without perturbing the Ar chronometers, but the timing of this reheating depends on if and how we choose to translate independent geologic observations into model space (Figs. 7B–7D). A sensitivity analysis like this builds a clear framework for a geologic interpretation of a preferred model result and an accompanying discussion.

Sensitivity Testing Take-Aways

Sensitivity testing guides the modeler, and those who review and read their work, to critically examine the parts of a model that matter most to a preferred result. This can be particularly valuable when sensitivity testing reveals model behaviors that are not otherwise obvious. For example, in the Path 5 assessment of the He data inputs (Fig. 5C), it is intuitive that the ca. 5 Ma burial heating event is documented by the ca. 5 Ma apatite He age. In contrast, a similar assessment of the zircon He data inputs in the deep-time example suggests the opposite: it is the youngest, and not the oldest, zircon crystals that are most sensitive to the Proterozoic thermal history (Fig. 6). In this way, sensitivity tests can facilitate productive comparisons to other studies, focus discussion of how confident we are in the parts of our data and model set-up that produce a preferred result, and direct future research to address key knowledge gaps that can continue to narrow thermal history model results.

There is no one right way to model thermochronologic data—and, as modelers are fond of reminding us, all models are wrong (Vermeesch and Tian, 2014, 2018, 2020; Gallagher and Ketcham, 2018, 2020). So, how do we know when the sensitivity testing of a thermal history model is sufficient? We argue that one is simply obligated to clearly document their modeling process and demonstrate the “why” of their preferred model, allowing others to replicate their results or even in some cases arrive at alternative interpretations for clearly articulated reasons.

Codifying Heating and Cooling Trends from Thermal History Models

As has been demonstrated throughout this text, thermal histories provide information about the timing and rate of heating and cooling events that are not always clearly identified by cooling ages alone. Extracting thermal history patterns from a set of forward or inverse models can reveal trends that are critical to recognizing the thermal signatures of geologic processes (Ketcham, 2005; Gallagher, 2012). We encourage modelers to begin this process by asking themselves the following questions. Is a heating or cooling event defined by its rate, timing of initiation, or duration? Should individual paths, a mean path, or a path envelope be used to define the timing of a thermal event? Once timing is assigned to a heating or cooling event, is there a way to quantify

uncertainty? The answers to these questions may vary depending on the goals of a study. In all cases, workers should establish a set of objective and consistent criteria to identify and define thermal trends and strive to quantify uncertainty associated with these definitions.

Extracting *tT* Trends from Synthetic Inverse Model Results

To illustrate the value of extracting trends from HeFTy, consider Path 5 (Fig. 3H). In this *tT* Path, long-lived heating from 100 Ma to 5 Ma is followed by relatively rapid cooling beginning at 5 Ma (Fig. 2A). Both parts of this thermal history are poorly documented by a single AHe age or even multiple cooling ages at face value (Fig. 2B); however, Path 5 is clearly represented among the range of good-fit *tT* paths identified in the inverse model of the Path 5-generated age-[eU] trend (Fig. 3H). Defining and then extracting thermal history features from model results when the true *tT* path is unknown is an important part of data analysis and interpretation for which there are several potential strategies. HeFTy inverse model results can be saved in a text file format that enables the modeler to evaluate these quantitative metrics in their software of choice.

A study designed to define the duration of heating using the Path 5 inverse model result (Fig. 3H) might decide to define heating by the earliest possible start of heating at ca. 100 Ma and the last possible instance of reheating at ca. 4 Ma (Fig. 8A). Although not every individual path shows reheating between these periods—some paths stay at surface temperatures for tens of millions of years before reheating, whereas others stop heating and begin cooling by as early as ca. 13 Ma—the criteria are consistent and can be applied objectively to more than one sample or model result. Additionally, the range of possible initial timing of heating start or end could be used to approximate uncertainty in the model results. For example, in the Path 5 inversion result, heating starts

between ca. 100 Ma and 75 Ma and ends between ca. 13 Ma—the earliest time an individual path stops cooling and starts heating—and 4 Ma, the time when all paths record cooling (Fig. 8A).

Alternatively, one might decide to define the duration of heating or the onset of cooling by identifying a temperature threshold and the time that each *tT* path passes through that threshold (e.g., Murray et al., 2016). For example, we calculate the fraction of paths in the Path 5 model result that heat to temperatures above 10 °C, 20 °C, 30 °C, 40 °C, and 50 °C between 100 Ma and 20 Ma (Fig. 8B). Because the inverse results from Path 5 are constrained by AHe data only, to then interpret this information we choose a temperature threshold of 40 °C, the lower threshold of the AHe closure temperature window. Using this metric, by 80 Ma, ~11% of good-fit paths are >40 °C, and by 20 Ma, ~89% of good-fit paths are >40 °C, which supports that heating occurred during this time window (Figs. 8A–8B). Overall, this description of the duration of heating is consistent with the strategy for defining heating described above. The hallmark of both strategies is the generation of reproducible and quantitative metrics for extracting heating duration from an inverse model result.

A modeler interested in the Path 5 cooling history, rather than the heating history, may want to objectively define the range of possible cooling start times in the inverse model result. One approach is to use HeFTy's good-fit path envelope, which encompasses the range of all individual good-fit paths, to identify both the earliest and latest possible start times of the required cooling event. For the inverse results of Path 5, this time span, between ca. 7 Ma and 3.5 Ma, provides a range of uncertainty for the initiation of cooling and includes the true start of cooling at 5 Ma (Fig. 8C).

These are just a few of many ways to define and quantify heating and cooling trends using HeFTy inverse model results. Because they are objectively defined, such model interpretations can be effectively compared among many independently modeled samples to identify the presence and distribution of regional geologic processes that drive rock heating and/or cooling.

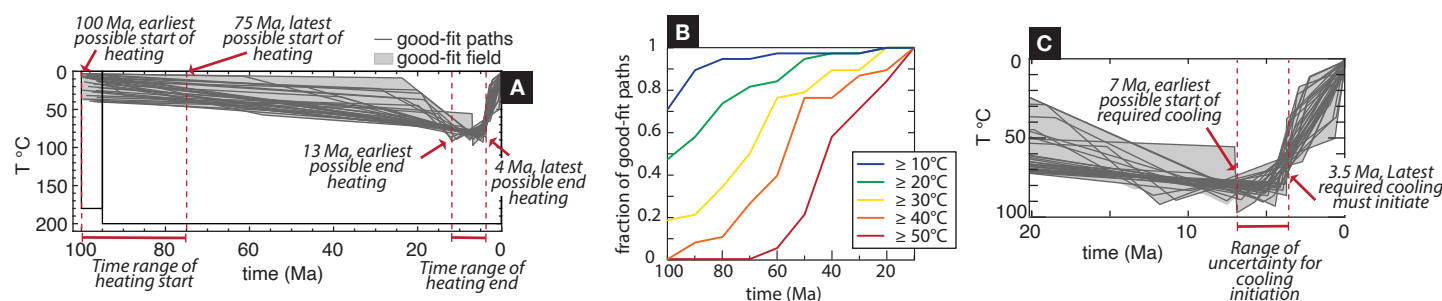


Figure 8. (A) Example of one objective strategy to quantify the duration of a long-lived heating event using the inverse model results for Path 5. This strategy uses individual paths to identify the earliest and latest start of heating and the earliest and latest end of heating. (B) Example of one objective strategy for quantifying the onset of a long-lived heating event using the inverse model results for Path 5. This strategy identifies the fraction of *tT* paths that have heated across five temperature thresholds between 100 Ma and 20 Ma (after Murray et al., 2016). (C) Example of one objective strategy for quantifying the initiation of a cooling event using the inverse model results for Path 5. This strategy uses the shape of the good-fit field, not individual *tT* paths, to identify the time span that required cooling starts.

Extracting *tT* Trends from a Regional Thermochronologic Data Set

Developing an objective set of criteria to codify thermal history trends allows one to compare trends among many independently modeled samples and identify spatiotemporal trends that are not represented in cooling ages alone. We highlight an example from recent work in the southern Patagonian Andes that codified inverse thermal history model outputs from HeFTy to resolve rapid regional Miocene cooling with a distinct spatiotemporal signature that had been previously unrecognized (Stevens Goddard and Fosdick, 2019a, 2019b) despite the existence of abundant thermochronologic data sets (Thomson et al., 2010; Fosdick et al., 2013; Herman et al., 2013; Herman and Brandon, 2015; Christeleit et al., 2017).

Inverse thermal history models of new and existing data by Stevens Goddard and Fosdick (2019a) used a multi-chronometer approach, which required that each independently modeled sample included data from two or three, thermochronometric systems—a combination of ZHe, AFT, or AHe data. The ZHe and AHe cooling ages exhibited no clear age-[eU] or age-size trends, and AFT data yielded insufficient track lengths for analysis. The model design was simple, as it was intended to interrogate a relatively recent event (Neogene) over a short period of geologic time with ZHe, AFT, and AHe cooling ages that are Late Cretaceous or younger in a region where ZFT dates all reflect local batholith cooling primarily in Late Cretaceous–Paleogene time (Thomson et al., 2001). The final model design included a broad initial condition between 80 Ma and 50 Ma of 100–300 °C temperatures (Fig. 9A), which is consistent with the geologic history of the samples, including thrust belt and/or sediment burial (Thomson et al., 2001; Fosdick et al., 2013; Stevens Goddard and Fosdick, 2019a) and final cooling to surface temperatures of $10^{\circ} \pm 5^{\circ} \text{C}$. A suite of sensitivity tests was run using alternate model designs with constraint

boxes that allowed or required reheating; however, no good-fit paths with non-monotonic solutions were identified, which suggests that the preferred model design was appropriate and included a comprehensive set of good-fit paths that were consistent with the measured data.

The inverse modeling results required accelerated rock cooling in Neogene time (Fig. 9A), the timing of which was not evident from individual thermochronologic ages. The cooling was visually evident in a qualitative examination of thermal history model results in *tT* space (e.g., Fig 9A) and is described specifically as “required” to indicate that although the cooling could have started earlier, it is tightly constrained by a temporal threshold. In this study, required cooling was identified at different temperatures for different samples (Stevens Goddard and Fosdick, 2019a). However, meaningfully extracting information about this rock cooling from model results required consideration of the geologic context. In this case, the study prioritized resolving the timing of rapid cooling regardless of the temperature threshold for which this cooling was observed as long as the temperature threshold was within the thermal sensitivity window of the thermochronometers used to model that particular sample. The authors of this study inspected the good-fit path envelope of each inverse model result to identify periods of required accelerated cooling and recorded the earliest and latest possible required cooling (Fig. 9A). The mean of these two times was used as the timing of required cooling, and the range of times was used to reflect uncertainty in the timing (Fig. 9A). Once the timing was extracted from many independent model results across the region, a clear spatial trend emerged: the required cooling—constrained at different temperature levels across samples—varied systematically across all samples according to latitude, with well-constrained rapid cooling progressively younging from south to north (Fig. 9B).

The latitudinal migration of a Miocene cooling signal (Fig. 9B) was a new observation with implications for tectonic plate-scale geologic processes, and

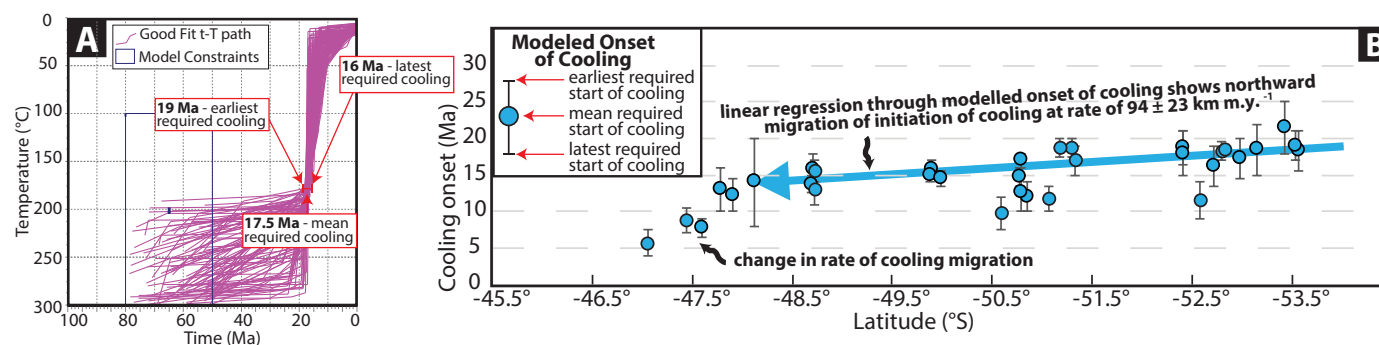


Figure 9. An example illustrating how to quantify and interpret the timing of rock cooling from HeFTy inverse model results. Modified from Stevens Goddard and Fosdick (2019a). (A) HeFTy inverse model result from a single sample. The good-fit path envelope is used to objectively extract the initiation of rock cooling (see also Fig. 8C). Stevens Goddard and Fosdick (2019a) modeled many samples individually and quantified the timing of rock cooling for each. Then, (B) the timing of cooling onset was plotted against sample latitude to reveal a spatiotemporal pattern of rock cooling in the southern Patagonian Andes.

it was revealed by codifying trends observed in inverse thermal history models. This required development of a set of consistent rules to define and assign both a date and an uncertainty to the cooling. In this case, a temperature threshold was not used as a rule because required cooling was observed across variable temperature ranges for different samples; however, in other studies, incorporating a temperature threshold may be useful and appropriate (e.g., Willett, 1997; Fox and Shuster, 2014; Fraser et al., 2021). Importantly, the rules used in this study can be used to codify future thermal history models developed from samples in this region. Although geologic interpretations of the signal may vary (Husson et al., 2019; Stevens Goddard and Fosdick, 2019a, 2019b), the existence of this trend is robust.

A Path Family Approach to Identifying Trends in Inverse Model Results

Ideally, inverse thermal history modeling is used to determine the full range of plausible tT paths that are consistent with the available thermochronologic data and geologic observations. However, in cases for which there is significant tT path diversity in an inverse model result, it can be difficult to interpret a geologic history because the results vary widely and may appear inconclusive. There is no one single way to handle these challenges. Here, we present a strategy for handling a diverse set of HeFTy inverse model results we call the Path Family Approach, which can be used to evaluate such model results from a geologic process-based perspective.

An Example Sample with a Non-Unique Inverse Model Result

To illustrate the Path Family Approach, we use an inverse model result for a real sample (Fig. 10A) with a measured ZFT age of 49.9 ± 2.0 Ma and three ZHe grains with ages of 20.5 ± 1.5 Ma, 21.6 ± 0.5 Ma, and 30.0 ± 0.5 Ma (Tables S3–S4, see footnote 1). We know from field relationships and previous analytical work that this sample is a ca. 160 Ma volcanic/volcaniclastic unit, which is in depositional contact with an overlying sedimentary unit deposited between ca. 150 Ma and 130 Ma. Given these geologic relationships, our inverse model has an initial geologic constraint box between 160 Ma and 130 Ma and 0–100 °C (Fig. 10A, Box 1). A second geologic constraint box between 120 Ma and 70 Ma and 300 °C and 400 °C reflects the timing of observed greenschist grade metamorphism (Fig. 10A, Box 2). We also impose a present-day surface thermal condition of 10 ± 5 °C. Otherwise, little is known about the sample's post-70 Ma thermal history. The goal of this inverse model is to use the thermochronologic data to fill this knowledge gap.

We first performed sensitivity testing of several different model designs (not shown) to find a configuration of post-70 Ma exploration boxes (Boxes 3 and 4 in Fig. 10A) that ask HeFTy to explore a wide range of thermal history behaviors during this unconstrained time. This testing reminds us that if there are no boxes between Box 2 and the modern surface temperature, or just one

box with a maximum temperature bound colder than 400 °C, then no matter what path behavior we select, HeFTy will only attempt monotonic cooling histories from 70 Ma to 0 Ma. For HeFTy to also explore nonmonotonic histories, and thus consider more than one episode of cooling since 70 Ma, two side-by-side exploration boxes are required (also see the model design of the synthetic examples in Fig. 3). Because the geologic record offers us no justification to only consider continuously cooling paths after 70 Ma, we choose a model design that explores a range of monotonic, variable, and episodic paths.

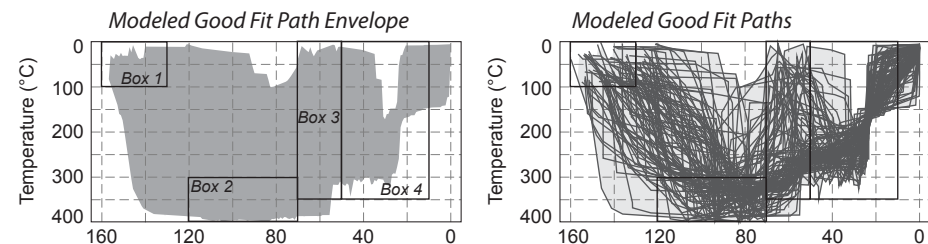
The resulting inverse model result (Fig. 10A) shows that all tT paths converge on monotonic cooling between ca. 50 Ma and 20 Ma. However, the period between 70 Ma and 35 Ma, the time period of greatest interest in this case, is not well constrained by the model results. When model results do not reveal a distinctive or narrow set of good-fit paths in a time period of interest, it is challenging to interpret the results geologically.

One general reason why parts of a thermal history model result can be poorly constrained is that cooling ages represent the cumulative thermal history of a sample, so it is common for different parts of good-fit thermal histories to trade off particular thermal history features (for example, the data could be consistent with either being hot in the Eocene and cold in the Miocene, or cold in the Eocene and hot in the Miocene). These relationships are commonly identified through sensitivity testing exercises (e.g., Fox and Shuster, 2014; Murray et al., 2019) and can be thought of as a kind of correlation amongst thermal history features (Fox and Carter, 2020). When an inverse model result is visualized in HeFTy, all of these possible scenarios are stacked on top of each other (i.e., Figs. 3A, 10A). However, all the good- and acceptable-fit tT paths can be exported from HeFTy and assessed using a spreadsheet or a script written in one's programming language of choice.

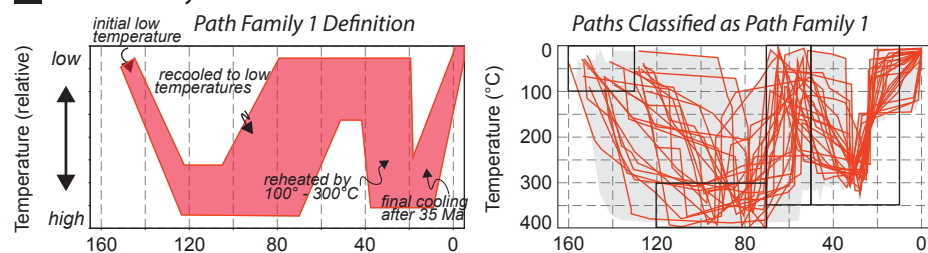
Many workers using HeFTy, as well as other modeling programs, have recognized the challenge of teasing out distinct thermal histories from a complete set of inversion results (Fox and Carter, 2020). For example, Willett (1997) used a convergence sequence to find paths with similar features and then proposed post-modeling analysis of results with similar characteristics, such as peak temperature. Other workers have iterated between inverse model results and forward models to test the relationship between thermal behaviors in a time-temperature history (e.g., Fox and Shuster, 2014; DeLucia et al., 2018; Stevens Goddard et al., 2018). The Path Family Approach we suggest here shares commonalities with these previous studies.

Our Path Family Approach categorizes each good-fit tT path into one of several groups, which are referred to as path families. By grouping paths based on thermal behaviors or features, one can evaluate path families from a geologic processes-based perspective. To define a path family, a modeler must designate objective criteria that yield reproducible classifications. The criteria used to define a path family will depend on the objectives of the study and may focus on characteristics such as monotonic versus non-monotonic path behavior, the timing or rate of heating or cooling, and/or residence temperatures. The Path Family Approach works well with HeFTy, which uses a Monte Carlo method to generate independent tT paths without automatically optimizing

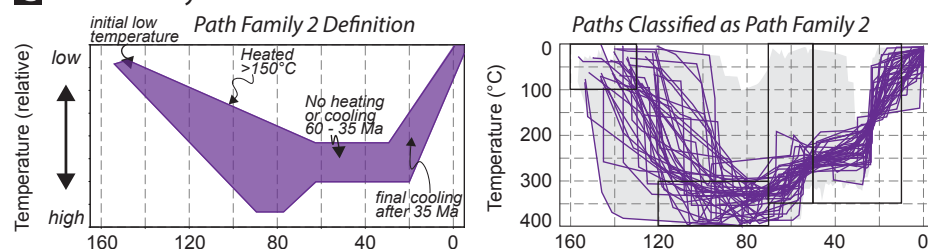
A Inverse Thermal History Model Results



B Path Family 1



C Path Family 2



D Path Family 3

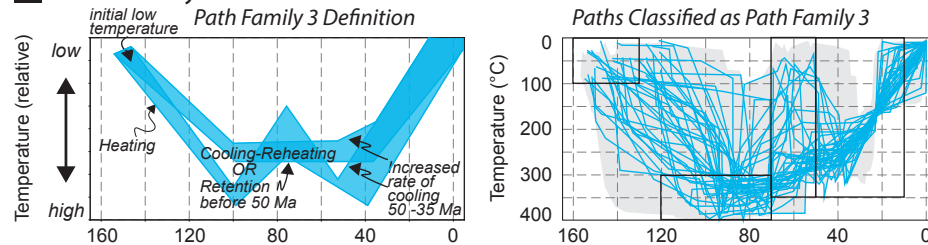


Figure 10. A demonstration of the Path Family Approach. (A) Inversion results show 100 good-fit paths. Boxes 1 and 2 represent information from the geologic record, whereas Boxes 3 and 4 are exploration boxes that ask HeFTy to explore a wide range of thermal histories from 70 Ma to present, during which time little is known about the thermal history of this sample independent of the thermochronologic data. **(B–D)** Three path families were defined and used for classification of individual paths for inversion results. The left panel illustrates the schematic definitions of these path families, which are described in detail in the text. The right panel shows the classification of individual paths from the inversion results, overlaid in color on the light gray good-fit path envelope.

the search of tT space. This inversion approach facilitates the search for individual paths with different thermal behaviors within a single model run and avoids convergence on paths with similar thermal behaviors (Ketcham, 2005).

We use the Path Family Approach to identify three path families with distinct thermal behaviors within the inversion result in Figure 10A. In this example, we categorize the paths based on their behavior during the time period of interest between 70 Ma and 0 Ma. The path behavior during this time is classified according to path monotonicity and timing of initial monotonic cooling. Path Family 1 contains non-monotonic paths that, during the time period of interest, cool to temperatures of <200 °C and are then reheated by 100–300 °C with the initiation of final cooling beginning sometime after 35 Ma (Fig. 10B). Path Family 2 contains monotonic cooling paths that exhibit little to no heating or cooling during the time period of interest with final cooling beginning sometime after 35 Ma (Fig. 10C). Path Family 3 includes both monotonic and non-monotonic cooling paths during the time period of interest, but final cooling starts earlier, between 50 Ma and 35 Ma (Fig. 10D).

These three path families exhibit thermal behaviors that represent quite different geologic histories. The magnitude of cooling and reheating required by Path Family 1 is likely to be corroborated by independent evidence of the geologic processes needed to produce these thermal histories—for example, evidence of several kilometers of erosion followed by evidence of several kilometers of sedimentary or tectonic burial (or an alternative heating mechanism). In contrast, the minor change in temperature experienced by Path Family 2 during this time suggests relative quiescence of erosion, burial, or geothermal activity (Fig. 10C). Path Family 3 has the greatest heterogeneity, because it includes both monotonic and non-monotonic tT paths. Additionally, although not a condition for Path Family 3 classification, the magnitude of initial cooling and reheating observed in the non-monotonic paths in Path Family 3 is lower than the magnitude of heating and cooling observed in the paths in Path Family 1, which provides further thermal distinction (Fig. 10). The timing of final monotonic cooling in Path Family 3 also occurs earlier than in Path Families 1 and 2.

It is important to remember that from a purely thermochronologic perspective, it is equally likely that each path in each family represents the true thermal history; however, when considered in the context of the regional geology, one may be able to reject one or more path families or advocate for a preferred path family. Even if no path families can be rejected on the basis of independent geologic evidence, path family classification allows workers to evaluate process-based geologic hypotheses, constrain plausible interpretations, and even design future work to discriminate among the path family solutions using new geologic observations.

Path Family Approach Take-Aways

The Path Family Approach is a useful alternative or companion exercise to a forward modeling approach that tests the sensitivity of thermochronologic data to specific path behaviors, because it does not require the modeler to

manually predict the full set of plausible thermal histories in a forward sense. Instead, the Path Family Approach relies on HeFTy's inversion algorithm to explore many possible solutions. This method pairs particularly well with HeFTy's non-learning Monte Carlo search of tT space and emphasis on reporting all of the individual tT paths that fit the input data, but it could be adapted to inversion results generated by other modeling programs. The Path Family Approach cannot, however, overcome the fundamental limitations of thermochronologic data. For example, path variability at temperatures higher than the closure temperature of the modeled systems is unlikely to be effectively categorized using this approach. Additionally, the Path Family Approach is also less likely to be useful for inverse models with limited data or geologic inputs. For example, the results of inverse modeling a single 40 Ma apatite grain (Fig. 3A) could be classified according to path families, but the number of path families required to fully categorize this model result may be quite large and difficult to evaluate geologically.

■ FEATURES OF ROBUST THERMAL HISTORY MODELS IN HEFTY

Thermal history models offer powerful geological insights with implications for a broad range of Earth scientists who may never produce thermal history models in the course of their own research. Although much of this contribution is dedicated to guiding modelers in the design and interpretation of thermal history model results in HeFTy, key principles identified in previous work (Ketcham, 2005; Flowers et al., 2015; Fox and Carter, 2020; Flowers et al., 2022) and expanded upon here can be distilled for non-expert consumers keen to evaluate published interpretations of thermal history models. First, the model design must be clearly documented so that it can be replicated (Flowers et al., 2015). Second, any codification or classification of thermal history trends and associated uncertainties should be objective and reproducible. Finally, we argue that inverse thermal history models should be accompanied by an explicit interrogation of what factors control the model results (e.g., model design, geologic constraints, selected diffusion model, and data inputs) such that the “why” of the result is clear. We also recommend the routine use of sensitivity testing to explore alternative solutions. In HeFTy, such sensitivity tests could include (1) additional inverse models that assess the role of constraint boxes in restricting the attempted thermal histories, with particular attention to the power of geologic constraints and whether exploration boxes are required to avoid the assumption of monotonicity and (2) additional forward or inverse models that examine how the inclusion or binning of data impacts a model result.

■ CONCLUSION

Thermal history modeling allows us to leverage the richness of low-temperature thermochronologic data in a wide variety of applications. As new

workers generate and model these data, an explicit and ongoing discussion of modeling approaches and interpretation strategies is needed to highlight the opportunities and limitations of thermal history modeling to resolve geologic problems. Although this contribution features examples using HeFTy, many of the strategies are transferable to forward and inverse thermal history modeling exercises using alternative programs. Our contribution is designed to:

- (1) Cultivate an intuitive approach to thermal history modeling. The synthetic data sets used in this study are “perfect” data with known true thermal histories and thus provide an opportunity to develop an intuitive understanding of the capacity of thermochronologic data sets to distinguish among distinct thermal histories using both forward and inverse modeling. This awareness is useful both during research planning—for example, in deciding which types of data will need to be collected to test competing hypotheses—as well as in the modeling and data-interpretation stages of a project.
- (2) Suggest best practices for testing the sensitivity of inverse model results to model design choices in HeFTy. We use both synthetic data and examples from real data sets to demonstrate strategies for sensitivity testing thermal history models. Sensitivity modeling can be used to avoid “black box” solutions by interrogating what factor(s) control model behavior including, but not limited to, the model design or inversion algorithm, a geologic constraint, a data trend, the spatial relationship between samples, or the choice of kinetic model.
- (3) Demonstrate interpretation strategies well-suited to HeFTy’s thermal history model results. Describing, interrogating, and interpreting thermal history model results can be challenging. We provide examples of reproducible methods to define and/or quantify thermal trends among non-unique thermal history results and introduce a strategy called the Path Family Approach to interpret inverse model results that display variable thermal behaviors.

ACKNOWLEDGMENTS

We thank the participants of the HeFTy and QTQt workshops at the 17th International Conference on Thermochronology for feedback on the tutorial provided in the supplementary materials. The authors acknowledge funding from the France-Berkeley Fund that supported this workshop. We thank Matthew Fox and one anonymous reviewer for comments that improved this work. The authors welcomed Charlie (M. Wildman), Theodore (K.E. Murray), and Timothy (A. Stevens Goddard) into the world during the submission, review, and revision of this manuscript. We thank the editors and reviewers for their patience and accommodation through the review process.

REFERENCES CITED

- Abbey, A.L., Wildman, M., Stevens Goddard, A.L., and Murray, K.E., 2022, Thermal history modeling techniques and interpretation strategies: Applications using QTQt: *Geosphere*, <https://doi.org/10.1130/GES02528.1> (in press).
- Aslanian, C., Jonckheere, R., Wauschkuhn, B., Ratschbacher, L., Universität, T., and Freiberg, B., 2022, Short communication: Experimental factors affecting fission-track counts in apatite: *Geochronology*, v. 4, p. 109–119, <https://doi.org/10.5194/gchron-4-109-2022>.
- Ault, A.K., Gautheron, C., and King, G.E., 2019, Innovations in (U-Th)/He, fission-track, and trapped-charge thermochronometry with applications to earthquakes, weathering, surface-mantle connections, and the growth and decay of mountains: *Tectonics*, v. 38, p. 3705–3739, <https://doi.org/10.1029/2018TC005312>.
- Braun, J., 2003, Pecube: A new finite-element code to solve the 3-D heat transport equation including the effects of a time-varying, finite amplitude surface topography: *Computers & Geosciences*, v. 29, p. 787–794, [https://doi.org/10.1016/S0098-3004\(03\)00052-9](https://doi.org/10.1016/S0098-3004(03)00052-9).
- Carlson, W.D., Donelick, R.A., and Ketcham, R.A., 1999, Variability of apatite fission-track annealing kinetics: I. Experimental results: *The American Mineralogist*, v. 84, p. 1213–1223, <https://doi.org/10.2138/am-1999-0901>.
- Christeleit, E.C., Brandon, M.T., and Shuster, D.L., 2017, Miocene development of Alpine glacial relief in the Patagonian Andes, as revealed by low-temperature thermochronometry: *Earth and Planetary Science Letters*, v. 460, p. 152–163, <https://doi.org/10.1016/j.epsl.2016.12.019>.
- DeLucia, M.S., Guenther, W.R., Marshak, S., Thomson, S.N., and Ault, A.K., 2018, Thermochronology links denudation of the Great Unconformity surface to the supercontinent cycle and snowball Earth: *Geology*, v. 46, p. 167–170, <https://doi.org/10.1130/G39525.1>.
- Dodson, M.H., 1973, Closure temperature in cooling geochronological and petrological systems: *Contributions to Mineralogy and Petrology*, v. 40, p. 259–274, <https://doi.org/10.1007/BF00373790>.
- Flowers, R.M., and Kelley, S.A., 2011, Interpreting data dispersion and “inverted” dates in apatite (U-Th)/He and fission-track datasets: An example from the US midcontinent: *Geochimica et Cosmochimica Acta*, v. 75, p. 5169–5186, <https://doi.org/10.1016/j.gca.2011.06.016>.
- Flowers, R.M., Ketcham, R.A., Shuster, D.L., and Farley, K.A., 2009, Apatite (U-Th)/He thermochronometry using a radiation damage accumulation and annealing model: *Geochimica et Cosmochimica Acta*, v. 73, p. 2347–2365, <https://doi.org/10.1016/j.gca.2009.01.015>.
- Flowers, R.M., Farley, K.A., and Ketcham, R.A., 2015, A reporting protocol for thermochronologic modeling illustrated with data from the Grand Canyon: *Earth and Planetary Science Letters*, v. 432, p. 425–435, <https://doi.org/10.1016/j.epsl.2015.09.053>.
- Flowers, R.M., Farley, K.A., and Ketcham, R.A., 2016, Response to comment on “A reporting protocol for thermochronologic modeling illustrated with data from the Grand Canyon”: *Earth and Planetary Science Letters*, v. 441, p. 213, <https://doi.org/10.1016/j.epsl.2016.02.024>.
- Flowers, R.M., Macdonald, F.A., Siddoway, C.S., and Havranek, R., 2020, Diachronous development of Great Unconformities before Neoproterozoic Snowball Earth: *Proceedings of the National Academy of Sciences of the United States of America*, v. 117, p. 10,172–10,180, <https://doi.org/10.1073/pnas.1913131117>.
- Flowers, R.M., Ketcham, R.A., Enkelmann, E., Gautheron, C., Reiners, P.W., Metcalf, J.R., Danišik, M., Stockli, D.F., and Brown, R.W., 2022, (U-Th)/He chronology: Part 2. Considerations for evaluating, integrating, and interpreting conventional individual aliquot data: *Geological Society of America Bulletin*, <https://doi.org/10.1130/b36268.1>.
- Fosdick, J.C., Grove, M., Hourigan, J.K., and Calderón, M., 2013, Retroarc deformation and exhumation near the end of the Andes, southern Patagonia: *Earth and Planetary Science Letters*, v. 361, p. 504–517, <https://doi.org/10.1016/j.epsl.2012.12.007>.
- Fox, M., and Carter, A., 2020, Heated topics in thermochronology and paths towards resolution: *Geosciences*, v. 10, 375, <https://doi.org/10.3390/geosciences10090375>.
- Fox, M., and Shuster, D., 2014, The influence of burial heating on the (U-Th)/He system in apatite: Grand Canyon case study: *Earth and Planetary Science Letters*, v. 397, p. 174–183, <https://doi.org/10.1016/j.epsl.2014.04.041>.
- Fox, M., and Shuster, D.L., 2020, Lazed and diffused: Untangling noble gas thermochronometry data for exhumation rates: *Elements*, v. 16, p. 337–342, <https://doi.org/10.2138/elements.16.5.337>.
- Fox, M., Herman, F., Willett, S.D., and May, D.A., 2014, A linear inversion method to infer exhumation rates in space and time from thermochronometric data: *Earth Surface Dynamics*, v. 2, p. 47–65, <https://doi.org/10.5194/esurf-2-47-2014>.
- Fraser, K.I., Enkelmann, E., Jess, S., Gilbert, H., and Grieco, R., 2021, Resolving the Cenozoic history of rock exhumation along the Central Rocky Mountain Trench using apatite low-temperature thermochronology: *Tectonics*, v. 40, no. 10, <https://doi.org/10.1029/2021TC006847>.
- Gallagher, K., 1995, Evolving temperature histories from apatite fission-track data: *Earth and Planetary Science Letters*, v. 136, p. 421–435, [https://doi.org/10.1016/0012-821X\(95\)00197-K](https://doi.org/10.1016/0012-821X(95)00197-K).
- Gallagher, K., 2012, Transdimensional inverse thermal history modeling for quantitative thermochronology: *Journal of Geophysical Research: Solid Earth*, v. 117, <https://doi.org/10.1029/2011JB008825>.
- Gallagher, K., 2016, Comment on “A reporting protocol for thermochronologic modeling illustrated with data from the Grand Canyon” by Flowers, Farley and Ketcham: *Earth and Planetary Science Letters*, v. 441, p. 211–212, <https://doi.org/10.1016/j.epsl.2016.02.021>.
- Gallagher, K., and Ketcham, R.A., 2018, Earth-Science Reviews Comment on “Thermal history modelling: HeFTy vs. QTQt” by Vermeesch and Tian, *Earth-Science Reviews* (2014), 139,

- 279–290: *Earth-Science Reviews*, v. 176, p. 387–394, <https://doi.org/10.1016/j.earscirev.2017.11.001>.
- Gallagher, K., and Ketcham, R.A., 2020, Earth-Science Reviews Comment on the reply to the Comment on “Thermal history modelling: HeFTy vs. QTQt” by Vermeesch and Tian, *Earth-Science Reviews* (2014), 139, 279–290: *Earth-Science Reviews*, v. 203, <https://doi.org/10.1016/j.earscirev.2019.102878>.
- Gautheron, C., and Zeitler, P.K., 2020, Noble gases deliver cool dates from hot rocks: *Elements*, v. 16, p. 303–309, <https://doi.org/10.2138/gselements.16.5.303>.
- Gautheron, C., Tassan-Got, L., Barbarand, J., and Pagel, M., 2009, Effect of alpha-damage annealing on apatite (U-Th)/He thermochronology: *Chemical Geology*, v. 266, p. 157–170, <https://doi.org/10.1016/j.chemgeo.2009.06.001>.
- Green, P., and Duddy, I., 2021, Discussion: Extracting thermal history from low temperature thermochronology. A comment on recent exchanges between Vermeesch and Tian and Gallagher and Ketcham: *Earth-Science Reviews*, v. 216, <https://doi.org/10.1016/j.earscirev.2020.103197>.
- Green, P.F., Duddy, I.R., Gleadow, A.J.W., Tingate, P.R., and Laslett, G.M., 1986, Thermal annealing of fission tracks in apatite. 1. A qualitative description: *Chemical Geology. Isotope Geoscience Section*, v. 59, p. 237–253, [https://doi.org/10.1016/0168-9622\(86\)90074-6](https://doi.org/10.1016/0168-9622(86)90074-6).
- Green, P.F., Duddy, I.R., Laslett, G.M., Hegarty, K.A., Gleadow, A.J.W., and Lovering, J.F., 1989, Thermal annealing of fission tracks in apatite 4. Quantitative modelling techniques and extension to geological timescales: *Chemical Geology. Isotope Geoscience Section*, v. 79, p. 155–182, [https://doi.org/10.1016/0168-9622\(89\)90018-3](https://doi.org/10.1016/0168-9622(89)90018-3).
- Guenther, W.R., 2021, Implementation of an alpha damage annealing model for zircon (U-Th)/He thermochronology with comparison to a zircon fission track annealing model: *Geochemistry, Geophysics, Geosystems*, v. 22, <https://doi.org/10.1029/2019GC008757>.
- Guenther, W.R., Reiners, P.W., Ketcham, R.A., Nasdala, L., and Giester, G., 2013, Helium diffusion in natural zircon: Radiation damage, anisotropy, and the interpretation of zircon (U-Th)/He thermochronology: *American Journal of Science*, v. 313, p. 145–198, <https://doi.org/10.2475/03.2013.01>.
- Hager, C., and Stockli, D.F., 2009, A new Matlab[®]-based Helium Modeling Package (“HeMP”) for thermal history recovery from single and multi-thermochronometer (U-Th)/He data and data arrays: *Geological Society of America Abstracts with Programs*, v. 41, no. 7, p. 487.
- Herman, F., and Brandon, M., 2015, Mid-latitude glacial erosion hotspot related to equatorial shifts in southern Westerlies: *Geology*, v. 43, p. 987–990, <https://doi.org/10.1130/G37008.1>.
- Herman, F., Seward, D., Valla, P.G., Carter, A., Kohn, B., Willett, S.D., and Ehlers, T.A., 2013, Worldwide acceleration of mountain erosion under a cooling climate: *Nature*, v. 504, p. 423–426, <https://doi.org/10.1038/nature12877>.
- Husson, L., Guillaume, B., and Martinod, J., 2019, Multichronometer thermochronologic modeling of migrating spreading ridge subduction in southern Patagonia: *COMMENT: Geology*, v. 47, e483, <https://doi.org/10.1130/G46629C.1>.
- Ketcham, R.A., 2005, Forward and inverse modeling of low-temperature thermochronometry data: *Reviews in Mineralogy and Geochemistry*, v. 58, p. 275–314, <https://doi.org/10.2138/rmg.2005.58.11>.
- Ketcham, R.A., and Tamer, M.T., 2021, Confined fission-track revelation in apatite: How it works and why it matters: *Geochronology*, v. 3, p. 433–464, <https://doi.org/10.5194/gchron-3-433-2021>.
- Ketcham, R.A., Donelick, R.A., and Donelick, M.B., 2000, AFTSolve: A program for multi-kinetic modeling of apatite fission-track data: *Geological Materials Research*, v. 2, p. 1–32.
- Ketcham, R.A., Carter, A., Donelick, R.A., Barbarand, J., and Hurford, A.J., 2007, Improved modeling of fission-track annealing in apatite: *American Mineralogist*, v. 92, p. 799–810, <https://doi.org/10.2138/am.2007.2281>.
- Ketcham, R.A., Donelick, R.A., Balestrieri, M.L., and Zattin, M., 2009, Reproducibility of apatite fission-track length data and thermal history reconstruction: *Earth and Planetary Science Letters*, v. 284, p. 504–515, <https://doi.org/10.1016/j.epsl.2009.05.015>.
- McDannell, K.T., and Flowers, R.M., 2020, Vestiges of the ancient: Deep-time noble gas thermochronology: *Elements*, v. 16, p. 325–330, <https://doi.org/10.2138/gselements.16.5.325>.
- Min, K., and Reiners, P.W., 2007, High-temperature Mars-to-Earth transfer of meteorite ALH84001: *Earth and Planetary Science Letters*, v. 260, p. 72–85, <https://doi.org/10.1016/j.epsl.2007.05.019>.
- Murray, K.E., Niemi, N.A., and Clark, M.C., 2022, Apatite and zircon (U-Th)/He thermochronometry data from the Colorado Front Range, Boulder County, Colorado, USA [Data set]: University of Michigan - Deep Blue Data, <https://doi.org/10.7302/g4eg-vr84>.
- Murray, K.E., Orme, D.A., and Reiners, P.W., 2014, Effects of U-Th-rich grain boundary phases on apatite helium ages: *Chemical Geology*, v. 390, p. 135–151, <https://doi.org/10.1016/j.chemgeo.2014.09.023>.
- Murray, K.E., Reiners, P.W., and Thomson, S.N., 2016, Rapid Pliocene–Pleistocene erosion of the central Colorado Plateau documented by apatite thermochronology from the Henry Mountains: *Geology*, v. 44, p. 483–486, <https://doi.org/10.1130/g37733.1>.
- Murray, K.E., Reiners, P.W., Thomson, S.N., Robert, X., and Whipple, K.X., 2019, The thermochronologic record of erosion and magmatism in the Canyonlands region of the Colorado Plateau: *American Journal of Science*, v. 319, p. 339–380, <https://doi.org/10.2475/05.2019.01>.
- Orme, D.A., Guenther, W.R., Laskowski, A.K., and Reiners, P.W., 2016, Long-term tectono-thermal history of Laramide basement from zircon–He age-[eU] correlations: *Earth and Planetary Science Letters*, v. 453, p. 119–130, <https://doi.org/10.1016/j.epsl.2016.07.046>.
- Peyton, S.L., Reiners, P.W., Carrapa, B., and Decelles, P.G., 2012, Low-temperature thermochronology of the northern Rocky Mountains, western U.S.A.: *American Journal of Science*, v. 312, p. 145–212, <https://doi.org/10.2475/02.2012.04>.
- Premo, W.R., and Fanning, C.M., 2000, SHRIMP U-Pb zircon ages for Big Creek gneiss, Wyoming and Boulder Creek batholith, Colorado: Implications for timing of Paleoproterozoic accretion of the northern Colorado province: *Rocky Mountain Geology*, v. 35, p. 31–50, <https://doi.org/10.2113/35.1.31>.
- Reiners, P.W., and Farley, K.A., 2001, Influence of crystal size on apatite (U-Th)/He thermochronology: An example from the Bighorn Mountains, Wyoming: *Earth and Planetary Science Letters*, v. 188, p. 413–420, [https://doi.org/10.1016/S0012-821X\(01\)00341-7](https://doi.org/10.1016/S0012-821X(01)00341-7).
- Shaw, C.A., Snee, L.W., Selverstone, J., and Reed, J.C., 1999, ⁴⁰Ar/³⁹Ar thermochronology of Mesoproterozoic metamorphism in the Colorado Front Range: *The Journal of Geology*, v. 107, p. 49–67, <https://doi.org/10.1086/314335>.
- Stevens Goddard, A.L., and Fosdick, J.C., 2019a, Multichronometer thermochronologic modeling of migrating spreading ridge subduction in southern Patagonia: *Geology*, v. 47, p. 555–558, <https://doi.org/10.1130/G46091.1>.
- Stevens Goddard, A.L., and Fosdick, J.C., 2019b, Multichronometer thermochronologic modeling of migrating spreading ridge subduction in southern Patagonia: *REPLY: Geology*, v. 47, e484, <https://doi.org/10.1130/G46788Y.1>.
- Stevens Goddard, A.L., Larrovere, M.A., Carrapa, B., Aciar, R.H., and Alvarado, P., 2018, Reconstructing the thermal and exhumation history of the Sierras Pampeanas through low-temperature thermochronology: A case study from the Sierra de Velasco: *Geological Society of America Bulletin*, v. 130, p. 1842–1858, <https://doi.org/10.1130/B31935.1>.
- Thomson, S.N., Hervé, F., and Stöckert, B., 2001, Mesozoic–Cenozoic denudation history of the Patagonian Andes (southern Chile) and its correlation to different subduction processes: *Tectonics*, v. 20, p. 693–711, <https://doi.org/10.1029/2001TC900013>.
- Thomson, S.N., Brandon, M.T., Tomkin, J.H., Reiners, P.W., Vásquez, C., and Wilson, N.J., 2010, Glaciation as a destructive and constructive control on mountain building: *Nature*, v. 467, p. 313–317, <https://doi.org/10.1038/nature09365>.
- Vermeesch, P., and Tian, Y., 2014, Earth-Science Reviews thermal history modelling: HeFTy vs. QTQt: *Earth-Science Reviews*, v. 139, p. 279–290, <https://doi.org/10.1016/j.earscirev.2014.09.010>.
- Vermeesch, P., and Tian, Y., 2018, Earth-Science Reviews Reply to Comment on “Thermal history modelling: HeFTy vs. QTQt” by K: *Earth-Science Reviews*, v. 176, p. 395–396, <https://doi.org/10.1016/j.earscirev.2017.11.015>.
- Vermeesch, P., and Tian, Y., 2020, Reply to the Comment on the Reply to the Comment on Vermeesch and Tian (2014): *Earth-Science Reviews*, v. 203, <https://doi.org/10.1016/j.earscirev.2019.102879>.
- Willett, C.D., Fox, M., and Shuster, D.L., 2017, A helium-based model for the effects of radiation damage annealing on helium diffusion kinetics in apatite: *Earth and Planetary Science Letters*, v. 477, p. 195–204, <https://doi.org/10.1016/j.epsl.2017.07.047>.
- Willett, S., 1997, Inverse modeling of annealing of fission tracks in apatite 1: A controlled random search method: *American Journal of Science*, v. 297, p. 939–969, <https://doi.org/10.2475/ajs.297.10.939>.
- Wolf, R.A., Farley, K.A., and Kass, D.M., 1998, Modeling of the temperature sensitivity of the apatite (U-Th)/He thermochronometer: *Chemical Geology*, v. 148, p. 105–114, [https://doi.org/10.1016/S0009-2541\(98\)00024-2](https://doi.org/10.1016/S0009-2541(98)00024-2).
- Zeitler, P.K., 2004, Avert 4.0.1. Inversion of ⁴⁰Ar/³⁹Ar age spectra: User’s Manual: https://eesarchive.lehigh.edu/EESdocs/geochron/downloads/arvert/arvert_41_manual.pdf.