

A mechanistic approach to unmixing detrital geochronologic data using Bayesian nonparametric mixture models

John Tipton · Glenn Sharman · Samuel Johnstone

Received: date / Accepted: date

1 **Abstract** Sedimentary deposits constitute the primary record of changing
2 environmental conditions that have acted on Earth's surface over geologic
3 time. Clastic sediment is eroded from source locations (parents) in sediment
4 routing systems and deposited at sink locations (children). Both parents and
5 children have characteristics that vary across many different dimensions, in-
6 cluding grain size, chemical composition, and the geochronologic age of con-
7 stituent detrital minerals. During transport, sediment from different parents
8 is mixed together to form a child, which in turn may serve as the parent for
9 other sediment further down system or later in time when buried sediment is
10 exhumed. To the extent that parent sources produce sediment with distinguish-
11 able geochronologic ages, the distribution of detrital mineral ages observed in
12 child sediments allows for investigation of the proportions of each parent in
13 the child sediment which ultimately reflects properties of the sediment routing
14 system, such as the relative sediment flux. To model the proportion of dates
15 in a child sample that comes from each of the parent distributions, we use a
16 Bayesian mixture of Dirichlet processes. This model allows for estimation of
17 the mixing proportions with associated uncertainty while making minimal as-
18 sumptions. We also present an extension to the model whereby we reconstruct
19 unobserved parent distributions from multiple observed child distributions us-
20 ing mixtures of Dirichlet processes, accounting for uncertainty in both the
21 number of parent distributions and the mixing proportions.

John R. Tipton
Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR, USA
E-mail: jrtipton@uark.edu

Glenn R. Sharman
Department of Geosciences, University of Arkansas, Fayetteville, AR, USA

Samuel A. Johnstone
U.S. Geological Survey, Geoscience and Environmental Change Science Center, Denver, USA

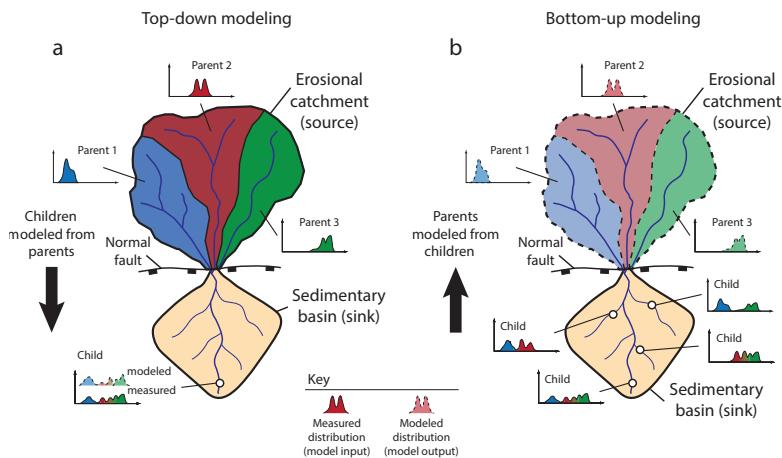


Fig. 1: Schematic depiction of a sediment routing system with an erosional source region characterized by three parents (black, red, and green) and an associated sedimentary basin (yellow). Call out plots represent the age of detrital minerals from the parents and how they are mixed to form child distributions. (a) Top-down modeling (sensu [Sharman and Johnstone(2017)]) where one or more children are modeled as a mixture of two or more parents. (b) Bottom-up modeling where multiple children are used to reconstruct end-member sources, or parents.

22 1 Introduction

23 To understand the origins of modern and ancient physical geography one must
 24 understand how erosional landscapes and associated sedimentary basins evolve
 25 through time [Romans et al.(2016)Romans, Castelltort, Covault, Fildani, and Walsh].
 26 As clastic sediment is generated by weathering and erosion, it is subsequently
 27 transported downstream, mixed, and ultimately deposited into a depositional
 28 sink. Modeling sediment mixing allows inference about these processes that
 29 generated the modern landscape. The ability to decipher the relative propor-
 30 tions of sources that eroded to produce sediment informs understanding of the
 31 underlying geologic processes controlling the evolution of the Earth's surface
 32 [Stock et al.(2006)Stock, Ehlers, and Farley, Sharman et al.(2019)Sharman, Sylvester, and Covault,
 33 Mason et al.(2017)Mason, Fildani, Gerber, Blum, Clark, and Dykstra, Kimbrough et al.(2015)Kimbrough, Grove, Gehre-

34 One of the most common ways to characterize the provenance of sediment is
 35 detrital geochronology – dating the time at which the individual minerals that
 36 make up a rock formed or cooled. These mineralization events typically reflect
 37 the timing of igneous rock forming events or metamorphic alteration of previ-
 38 ously existing rocks [Gehrels(2014)]. In other cases mineral ages reflect the his-
 39 tory of mineral cooling (e.g., ‘thermochronology’, [Reiners and Brandon(2006)]).

40 Detrital geochronologic ages are most commonly determined from measurements of radiogenic isotopes contained within individual mineral crystals.
41 The decay of uranium (U) to lead (Pb) within zircon, a relatively robust
42 mineral, makes this approach ideally suited for tracking sedimentary mixing
43 [Amidon et al.(2005b) Amidon, Burbank, and Gehrels, Sundell and Saylor(2017),
44 Sharman and Johnstone(2017)].

45 We will follow the convention that sediment sources are called *parents* and
46 sink locations are called *children*. Using this language, the manuscript aims to
47 address two questions. First, can we estimate the proportion of each parent age
48 distribution in a child age distribution with associated uncertainty? Second,
49 can we estimate the marginal age distributions for unobserved parents given
50 a set of child age distributions? These questions are answered using “top-
51 down” and “bottom-up” approaches to sediment unmixing, respectively (see
52 Fig. 1; [Sharman and Johnstone(2017)]). The top-down approach models one
53 or more child samples as mixtures of specified parent samples (Fig. 1). The
54 bottom-up approach uses multiple child samples to model likely parents which
55 are more generally referred to as end-members in mixture modeling efforts.
56 [Sharman and Johnstone(2017)].

57 Bayesian mixture modeling of geochronology data, including detrital data,
58 has numerous advantages when addressing single samples [Jasra et al.(2006) Jasra, Stephens, Gallagher, and Holmes]
59 including allowing inference and uncertainty estimates for the number and
60 value of true ages characterized by observed mineral dates. Here we extend this
61 concept to consider the geologic mixing of sediments derived from source ar-
62 eas containing minerals recording different crystallization events. The Bayesian
63 nonparametric statistical model presented herein has a number of advantages
64 over previously used approaches, including being able to derive direct, proba-
65 bilistic estimates of uncertainty associated with the mixture model. We demon-
66 strate the utility of this approach using both a synthetic dataset and a well-
67 constrained, natural case study in central California, USA [Sickmann et al.(2016) Sickmann, Paull, and Graham].
68 The top-down mixing approach is able to successfully reconstruct parent con-
69 tributions in both synthetic and natural datasets. Although the bottom-up
70 unmixing model is able to successfully reconstruct parents in the synthetic
71 dataset, there is evidence of non-identifiability – where parents cannot be
72 uniquely characterized from the children – when applied to the natural dataset.
73 More generally, the framework we present can also give guidance about other
74 scientific questions that relate to mixing of non-parametric sum-to-one data
75 in Earth sciences and other disciplines (e.g., unmixing sediment grain size
76 distribution; [Weltje(1997)] and references within)

78 2 Model Overview

79 To define the statistical model, we follow the convention that letters repre-
80 sent data and Greek symbols represent parameters. A plaintext symbol (y)
81 represents a scalar, a bold lowercase symbol represents a vector (\mathbf{y}), and a
82 bold uppercase symbol is a matrix (\mathbf{Z}) whose columns are vectors written as

⁸³ $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_p)$. We use the notation $[y]$ to represent the probability distribution/mass function (pdf/pmf) and let $[y|\theta]$ represent the conditional pdf/pmf of the random variable y given θ .

⁸⁶ Following [Berliner(2003)], the statistical model described below is divided
⁸⁷ into three components: the data model, the process model, and the prior model.
⁸⁸ In general, the data model defines probability distributions that describe the
⁸⁹ variability in the data due to the observation process. The data model can be
⁹⁰ modified to account for non Gaussian measurement processes like counts, out-
⁹¹ liers, spatial/temporal errors, etc [Tipton et al.(2017)Tipton, Hooten, and Goring,
⁹² Hefley et al.(2017)Hefley, Brost, and Hooten]. Process models describe the best
⁹³ scientific understanding of the process of interest. For example, process models
⁹⁴ have been used to describe the monthly response of trees to climate [Tipton et al.(2016)Tipton, Hooten, Pederson, Tingley,
⁹⁵ the relationship between climate and pollen in sediments [Tipton et al.(2019)Tipton, Hooten, Nolan, Booth, and McLachlan,
⁹⁶ and the movement of ice sheets in Antarctica [Chang et al.(2016)Chang, Haran, Applegate, and Pollard,
⁹⁷ Guan et al.(2018)Guan, Haran, and Pollard]. The prior model describes the
⁹⁸ range of parameter values that are plausible. Sometimes the prior model is
⁹⁹ used as regularization to improve the generalization of the model to unob-
¹⁰⁰ served data [Hooten and Hobbs(2015)].

¹⁰¹ 3 Top-down mixing model

¹⁰² The model framework presented below, which is appropriate for situations
¹⁰³ where the parent and children sediment have been independently character-
¹⁰⁴ ized, will answer the first research question: can one estimate the proportion
¹⁰⁵ of each parent that comprises the child sediment?

¹⁰⁶ 3.1 Top-down mixing data model

¹⁰⁷ Let the n_y observed age measurements of a single child of interest be \mathbf{y} and
¹⁰⁸ let the observed date measurements for each of the $b = 1, \dots, B$ parents be
¹⁰⁹ given by the n_b -dimensional vector \mathbf{z}_b . Because the observed ages are mea-
¹¹⁰ sured with uncertainty reported as a dating error standard deviation, we
¹¹¹ explicitly account for this source of uncertainty in the data model. In the
¹¹² case of U-Pb dating of detrital zircon grains, dates are most commonly de-
¹¹³ termined using laser ablation-inductively coupled plasma-mass spectrometry
¹¹⁴ [Gehrels(2012)]. Such date measurements typically have relative 2σ analytical
¹¹⁵ precision of 1-4%, with relative uncertainty increasing for younger analyses
¹¹⁶ [Puetz et al.(2018)Puetz, Ganade, Zimmermann, and Borchardt]. For each de-
¹¹⁷ trital mineral, the estimate of dating measurement uncertainty is reported as
¹¹⁸ a n_y -vector of standard deviations $\boldsymbol{\sigma}_y$ for the child and B n_b -vectors of stan-
¹¹⁹ dard deviations $\boldsymbol{\sigma}_{z_b}$. We assume the date measurement uncertainty follows a
¹²⁰ Gaussian distribution where the observed sediment particle date is

$$\begin{aligned} \mathbf{y}|\tilde{\mathbf{y}}, \boldsymbol{\sigma}_y^2 &\sim N(\tilde{\mathbf{y}}, \text{diag}(\boldsymbol{\sigma}_y^2)), \\ \mathbf{z}_b|\tilde{\mathbf{z}}_b, \boldsymbol{\sigma}_{z_b}^2 &\sim N(\tilde{\mathbf{z}}_b, \text{diag}(\boldsymbol{\sigma}_{z_b}^2)). \end{aligned} \quad (1)$$

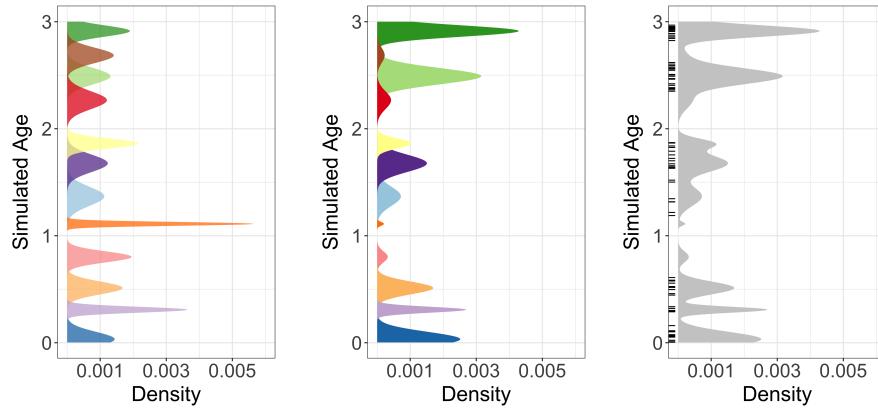
We break the variable naming convention and let $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}_b$ be latent parameters that represent the true, unobserved age of the sediments where \mathbf{y} (\mathbf{z}_b) will be close to $\tilde{\mathbf{y}}$ ($\tilde{\mathbf{z}}_b$) because the dating uncertainty is small relative to the variability in the data (i.e., the average coefficient of variation of measured dates, defined as the dating standard deviation divided by the date, is about 0.02-0.03). To represent more uncertainty in the data or to account for asymmetric measurement errors a Student's-t, log-normal, or other appropriate distribution could be used instead of the normal distribution.

3.2 Top-down mixing process model

The process model addresses two scientific questions: what are the estimates of the true, unobserved detrital mineral age distributions at the parent and child locations? and what proportions of those detrital minerals did each parent source contribute to the child? There are many different methods available to model the true geochronological age distributions from the sample data, including kernel density estimates [Vermeesch(2012)], non-negative matrix factorization methods [Saylor et al.(2019)Saylor, Sundell, and Sharman], and Bayesian nonparametric models of mineral fomation event mixing [Jasra et al.(2006)Jasra, Stephens, Gallagher, and H Tye et al.(2019)Tye, Wolf, and Niemi]; however, Bayesian methods have yet to be applied to the problem of sediment mixing. We develop a Bayesian nonparametric model that approximates our geologic understanding.

Over geologic time, individual minerals may be repeatedly recycled into sedimentary rocks by of erosion, transport, deposition, and exhumation. However, in many cases the dates recorded by individual minerals contained in these deposits are distinctive and unaffected by these recycling processes (e.g., excluding burial reheating of thermochronometers [Fosdick et al.(2015)Fosdick, Grove, Graham, Hourigan, Lovera, and R... We assume that minerals created by the same geologic event share an age distribution that is relatively homogeneous with only small variability. Furthermore, episodes of rock and mineral formation (typically lasting 10^5 to 10^7 years [Chen and Moore(1982), Irwin and Wooden(1999), Wotzlaw et al.(2013)Wotzlaw, Schaltegger, Frick, Dungan, Gerdes, and ... are nearly discrete events relative to geologic time (4.5×10^9 years). While minerals often show overgrowths of different ages, this provides a useful approximation. Under the conceptual model (Fig. 1), sediment is formed by the decomposition of rocks containing minerals created at different times, and sediment at every child location is composed of an unknown number of mineral formation events that are also present at source locations.

Consider the latent, unknown age of a single mineral grain from either the child $\tilde{\mathbf{y}}$ or one of the $b = 1, \dots, B$ parents $\tilde{\mathbf{z}}_b$. We assign the range of mineral ages for the k th formation event the base probability distribution



(a) Distribution of simulated mineral formation events. Each color represents a different formation event. Notice that some formation events have wider standard deviations (i.e., resulting from longer durations of mineral formation), while other formation events are shorter.

(b) The mineral formation events from Fig. 2a are re-weighted to account for relative abundance of potentially replaced by gray because observable mineral ages provided by each parent.

(c) Discrete, colored formation events in the parent distribution in Fig. 2b are replaced by gray because the formation events are unknown. The observed data are shown as a rug plot along the y-axis.

Fig. 2: The mixing model over the mineral formation events. The y-axis of each plot is the age of formation and the x-axis is the probability density of the hypothetical parent distribution.

159 $G(\theta_{bk})$ which depends on parameters θ_{bk} . There are many possible choices
160 for the base distribution $G(\theta_{bk})$; we assume a normal distribution $N(\mu_{bk}, \sigma_{bk}^2)$
161 with mean μ_{bk} and variance σ_{bk}^2 , therefore $\theta_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$. Other possible
162 choices include a log-normal or gamma distribution that enforces a positive
163 support on the observed age dates. Because we assume the age distribution
164 of a single mineral formation event is relatively short with respect to geologic
165 time, the variance parameters σ_{bk}^2 will be small relative to geologic time.

166 In most cases the true number of mineral formation events K recorded
167 by a detrital sample is unknown. Under our model, the set of all K mineral
168 formation events is a mixture of K normal distributions, as shown in Fig. 2a
169 where each mineral formation event is shown in a different color. The centers of
170 each age distribution in Fig. 2a are given by the values of μ_{bk} and the spreads
171 of each age distribution are given by the variances σ_{bk}^2 . Aerial extent, dif-
172 ferential erosion, the abundance of minerals of different ages within different
173 rocks, and other factors can influence the proportion of minerals of a given
174 age in rock at a site [Amidon et al.(2005a)Amidon, Burbank, and Gehrels,
175 Amidon et al.(2005b)Amidon, Burbank, and Gehrels]. Figure 2b shows the age
176 distributions in Fig. 2a that have been re-weighted to account for all of the

177 factors that determine the distribution of ages in a parent source rock. We
 178 do not observe the individual mineral formation event labels ($k = 1, \dots, K$)
 179 directly, we only observe the parent age distributions after influence from the
 180 relative contributions of minerals of different formation ages (Fig. 2c). Notice
 181 that in Fig. 2c, the colors from Figs. 2a and 2b are removed, representing
 182 the fact that the mineral event labels are not observed. In addition, we only
 183 observe a finite sample from the distribution in Fig. 2c, shown as a rug plot
 184 with each tick representing the observed detrital mineral grain date. Thus, the
 185 number of mineral formation events is potentially challenging to extract from
 186 the data.

187 Consider a detrital mineral date \tilde{z}_{ib} from a parent sediment source b from
 188 $i = 1, \dots, n_b$ measurements. The single sediment grain comes from a mineral
 189 formation event implying the mixture distribution

$$\tilde{z}_{ib} | \boldsymbol{\mu}_b, \boldsymbol{\sigma}_b^2, \gamma_{ib} \sim \begin{cases} N(\tilde{z}_{ib} | \mu_{1b}, \sigma_{1b}^2) & \text{if } \gamma_{ib} = 1 \\ \vdots & \vdots \\ N(\tilde{z}_{ib} | \mu_{Kb}, \sigma_{Kb}^2) & \text{if } \gamma_{ib} = K, \end{cases}$$

190 where γ_{ib} is a random variable whose value indicates which formation event
 191 k the detrital mineral comes from. We define the probability of a detrital
 192 mineral coming from formation event k as $p_{bk} \equiv P(\gamma_{ib} = k)$. Then, we write
 193 the joint distribution over all mineral grains from parent b as

$$\tilde{\mathbf{z}}_b | \boldsymbol{\mu}_b, \boldsymbol{\sigma}_b^2, \boldsymbol{\gamma}_b \sim \prod_{i=1}^{n_b} N(\tilde{z}_{ib} | \mu_{1b}, \sigma_{1b}^2)^{I\{\gamma_{ib}=1\}} N(\tilde{z}_{ib} | \mu_{2b}, \sigma_{2b}^2)^{I\{\gamma_{ib}=2\}} \cdots N(\tilde{z}_{ib} | \mu_{Kb}, \sigma_{Kb}^2)^{I\{\gamma_{ib}=K\}}$$

194 where $I\{\gamma_{ib} = k\}$ is an indicator function that takes the value 1 if $\gamma_{ib} = k$
 195 and 0 otherwise. Because there are a large number of indicator functions, we
 196 integrate them out of the process model to improve mixing and model fit. The
 197 integrated age distribution model is

$$\tilde{\mathbf{z}}_b | \boldsymbol{\mu}_b, \boldsymbol{\sigma}_b^2, \mathbf{p}_b \sim \prod_{i=1}^{n_b} \sum_{k=1}^K p_{bk} N(\tilde{z}_{ib} | \mu_{kb}, \sigma_{kb}^2)$$

198 where $\mathbf{p}_b = (p_{b1}, \dots, p_{bK})'$ is a vector of mixing probabilities with $\sum_{k=1}^K p_{bk} =$
 199 1. When the number of formation events K is potentially infinite, the concep-
 200 tual model can be described with the Dirichlet process model described in
 201 detail in Sect. 3.3.

202 In addition to modeling the age distribution of the parents, the process
 203 model specifies the proportion of each parent distribution in the child distri-
 204 bution. We represent the mixing proportions of the B parent distributions with
 205 the parameter $\boldsymbol{\phi} = (\phi_1, \dots, \phi_B)'$, with $\sum_{b=1}^B \phi_b = 1$. The parameter ϕ_b is the
 206 proportion of the child distribution that comes from parent \$b\$ and accounts

for differential mixing of parents. For parents that are comprised of bedrock, ϕ is a function of each parent's relative aerial extent in the drainage catchment, average erosion rate, and average concentration of the detrital mineral of interest [Amidon et al.(2005a)Amidon, Burbank, and Gehrels]. If parents are sediment inputs (e.g., rivers), then ϕ is a function of each parent's relative sediment supply and the average concentration of the detrital mineral of interest within the sediment.

For a single child sediment mineral date \tilde{y}_i , that sediment grain comes from only one parent. We define the categorical random variable δ_i to represent which parent distribution the child sediment came from. Using the categorical variable, the distribution of the child sediment grain is

$$\tilde{y}_i | \mu, \sigma^2, \delta_i \sim \begin{cases} \sum_{k=1}^K p_{1k} N(\tilde{y}_i | \mu_{k1}, \sigma_{k1}^2) & \text{if } \delta_i = 1 \\ \vdots & \vdots \\ \sum_{k=1}^K p_{Bk} N(\tilde{y}_i | \mu_{kB}, \sigma_{kB}^2) & \text{if } \delta_i = B. \end{cases}$$

Then, the probability that the sediment grain came from parent b is $\phi_b \equiv P(\delta_i = b)$. Defining the indicator variable $I\{\delta_i = b\}$ where $P(\delta_i = b) = E(I\{\delta_i = b\})$, we can write the age distribution over all sediment grains as

$$\tilde{y}_i | \mu, \sigma^2, \delta \sim \prod_{i=1}^{n_y} \left(\sum_{k=1}^K p_{1k} N(\tilde{y}_i | \mu_{k1}, \sigma_{k1}^2) \right)^{I\{\delta_i=1\}} \times \left(\sum_{k=1}^K p_{2k} N(\tilde{y}_i | \mu_{k2}, \sigma_{k2}^2) \right)^{I\{\delta_i=2\}} \times \cdots \times \left(\sum_{k=1}^K p_{Bk} N(\tilde{y}_i | \mu_{kB}, \sigma_{kB}^2) \right)^{I\{\delta_i=B\}}$$

where, like the parent mixing model, we integrate out the component indicator variables δ . After integrating out the parent component membership indicators, the child sediment grains have the age distribution

$$\tilde{y}_i | \mu, \sigma^2, \phi \sim \prod_{i=1}^{n_y} \sum_{b=1}^B \phi_b \sum_{k=1}^K p_{bk} N(\tilde{y}_i | \mu_{kb}, \sigma_{kb}^2),$$

where the the posterior distribution of ϕ_b is used to estimate the proportion of child sediment from parent b .

Combining the above results, the full process model is

$$\tilde{\mathbf{y}}, \tilde{\mathbf{Z}} | \phi, \mathbf{p}, \mu, \sigma^2 \sim \left(\prod_{i=1}^{n_y} \sum_{b=1}^B \phi_b \sum_{k=1}^K p_{bk} N(\tilde{y}_i | \mu_{kb}, \sigma_{kb}^2) \right) \times \left(\prod_{b=1}^B \prod_{j=1}^{n_b} \sum_{k=1}^K p_{bk} N(\tilde{z}_{jb} | \mu_{kb}, \sigma_{kb}^2) \right).$$

227 3.3 Top-down mixing prior model

228 The conceptual process model assumes the number of mineral formation events
229 K is known. In practice, the number of formation events is unknown and is a
230 parameter to be estimated. In fact, it is likely that the different parent sites
231 will have different numbers of mineral formation events based on site-specific
232 history. The prior model addresses the fundamental question of estimating the
233 number of mineral formation events.

234 There are a number of potential approaches to model the unknown number
235 of formation events. First, one can treat the number of formation events as
236 a fixed parameter, perform a grid search over the different number of formation
237 events, and choose the model that best fits the data [Miller and Harrison(2018)].
238 A second approach is to model the number of formation events using a re-
239 versible jump algorithm [Green(1995)]. The third approach is to assign a
240 Dirichlet process prior over the number of formation events. The Dirichlet
241 process estimates an unknown number of components without *a priori* speci-
242 fying the number.

243 The Dirichlet process is an infinite dimensional stochastic process which is
244 a distribution over distributions [Ferguson(1973)]. We use the stick-breaking
245 representation of a Dirichlet process

$$\sum_{k=1}^{\infty} p_{bk} G(\boldsymbol{\theta}_{bk}), \quad (2)$$

246 where $G(\cdot)$ is the base distribution with parameters $\boldsymbol{\theta}_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$ and
247 mixing weights p_{bk} with $\sum_{k=1}^{\infty} p_{bk} = 1$. In practice, $p_{bk} \approx 0$ for large k , there-
248 fore, the infinite sum is well approximated by the finite sum $\sum_{k=1}^K p_{bk}$ for a
249 large enough K (for most problems $K=10$ or $K=20$ is sufficiently large). The
250 stick-breaking representation for \mathbf{p}_b is constructed by transforming auxiliary
251 variables $\tilde{\mathbf{p}}_b = (\tilde{p}_{b1}, \dots, \tilde{p}_{bK-1})'$ using the stick-breaking representation

$$p_{bk} = \begin{cases} \tilde{p}_{b1} & \text{for } k = 1, \\ \tilde{p}_{bk} \prod_{k'=1}^{k-1} (1 - \tilde{p}_{bk'}) & \text{for } k = 2, \dots, K-1, \\ 1 - \prod_{k'=1}^{K-1} (1 - \tilde{p}_{bk'}) & \text{for } k = K. \end{cases}$$

252 Priors on the \tilde{p}_{bk} are assigned exchangeable Beta($1, \alpha_b$) priors giving rise
253 to the stick-breaking Dirichlet process. The hyperparameters α_b are given ex-
254 changeable Gamma($1, 1$) priors that control the Dirichlet process concentra-
255 tion (i.e., smaller α_b give fewer formation events, larger α_b give more for-
256 mation events). Because our study site is constrained geographically, the par-
257 ent and child sites contain mineral grains derived from common formation
258 events, we follow [Lock and Dunson(2015)] and used shared kernels by letting
259 $\boldsymbol{\theta}_{bk} \equiv \boldsymbol{\theta}_k = (\mu_k, \sigma_k^2)'$ for all $b = 1, \dots, B$.

260 The standard deviations for the ages of formation are assigned truncated
261 half-Cauchy priors $\sigma_k \sim \text{Cauchy}^+(0, s)I\{0 < \sigma_k < \omega\}$, where we choose s to

262 be small relative to the range of dates observed and ω provides an upper limit
 263 to the duration of formation events. For the case study where the ages span
 264 the range of 0 to about 300 Millions of years (Ma), we set s to be 25 Ma and
 265 set ω to be 50 Ma years. The truncation is important to prevent the Dirichlet
 266 process mixture from generating unrealistically long formation events which
 267 does not match our *a priori* geologic knowledge.

268 The mixing parameter ϕ is assigned a $\text{Dirichlet}(\alpha_\phi \mathbf{1})$ prior where $\mathbf{1}$ is a vec-
 269 tor of ones and the hyperparameter α_ϕ is assigned a $\text{Gamma}(1, 1)$ prior. When
 270 α_ϕ is small the mixing proportions concentrate with a large probability on a
 271 single parent component, when α_ϕ is one ϕ will be uniformly distributed over
 272 all possible mixing proportions, and when α_ϕ is large the mixing proportion
 273 will be concentrated at equal mixing proportions $(\frac{1}{B}, \dots, \frac{1}{B})$.

274 **3.4 Top-down mixing posterior distribution**

275 The top-down mixing model posterior distribution is

$$\begin{aligned} [\phi, \phi, \mu, \sigma^2, b | \mathbf{y}, \mathbf{Z}] &\propto [\mathbf{y} | \tilde{\mathbf{y}}, \sigma_y] \prod_{b=1}^B [\mathbf{z}_b | \tilde{\mathbf{z}}_b, \sigma_b] \times \\ &[\tilde{\mathbf{y}} | \phi, \mathbf{p}_b, \mu_b, \sigma_b] \prod_{b=1}^B [\tilde{\mathbf{z}}_b | \mathbf{p}_b, \mu_b, \sigma_b] \times \\ &[\phi | \alpha_\phi] [\alpha_\phi] \left(\prod_{b=1}^B [\mathbf{p}_b | \alpha_b] [\alpha_b] [\mu_b] [\sigma_b] \right), \end{aligned} \quad (3)$$

276 where the first three line on the right hand side of the proportional symbol
 277 are the data, process, and the prior model, respectively. We estimate the pos-
 278 terior using Markov Chain Monte Carlo (MCMC) with the *R* package *NIMBLE*
 279 [de Valpine et al.(2017)de Valpine, Turek, Paciorek, Anderson-Bergman, Temple Lang, and Bodik]
 280 using an adaptive block Metropolis-Hastings algorithm [Haario et al.(2001)Haario, Saksman, and Tamminen].
 281 The constrained auxiliary variable $\tilde{\mathbf{p}}$ and standard deviation σ are transformed
 282 to unconstrained support (logit- and log-scale transformations) for tuning
 283 the Metropolis-Hastings block proposals, with corresponding Jacobian adjust-
 284 ments to the acceptance probabilities. The sampling of the sum-to-one mixing
 285 proportion ϕ is performed by introducing auxiliary variables $\tilde{\phi}$, assigning a
 286 stick-breaking prior on $\tilde{\phi}$, then sampling on a logit-scale after correcting for
 287 the transformation using the Jacobian to induce a $\text{Dirichlet}(\alpha_\phi \mathbf{1})$ prior on ϕ .

288 **4 Bottom-up unmixing model**

289 The second research question is: can we reconstruct unobserved parent age
 290 distributions from multiple child observations? In previous work, this analysis

291 has been variably termed “end-member mixing analysis”, “end-member modeling”, or “end-member analysis” as applied to unmixing grain size or detrital
 292 age distributions [Sharman and Johnstone(2017), Saylor et al.(2019)Saylor, Sundell, and Sharman].
 293 The end-member unmixing analysis has two components. First, the number
 294 of parents B is unknown and needs to be estimated. Second, given the num-
 295 ber of parents B , what are the unobserved mineral formation age distribu-
 296 tions for the B parents? For this paper, we assume the number of parents
 297 B is known. There are a number of criteria for selecting the number of par-
 298 ents including using Bayesian information criteria, reversible jump MCMC
 299 [Jasra et al.(2006)Jasra, Stephens, Gallagher, and Holmes], assuming a Dirich-
 300 let process over the number of parents, or fitting a mixture of finite mix-
 301 tures [Miller and Harrison(2018)]. Rather than explore these ideas, we devote
 302 our effort on developing the unmixing model for a fixed number of parents
 303 [Miller and Harrison(2018)]. The end-member model uses the same general
 304 framework presented in the mixture of Gaussians model (3) with some modi-
 305 fications.

307 4.1 Bottom-up unmixing data model

308 Let $d = 1, \dots, D$ index the D child sediments that are each composed of $i =$
 309 $1, \dots, n_d$ samples. As before, we assume a Gaussian dating error distribution
 310 for child d given by

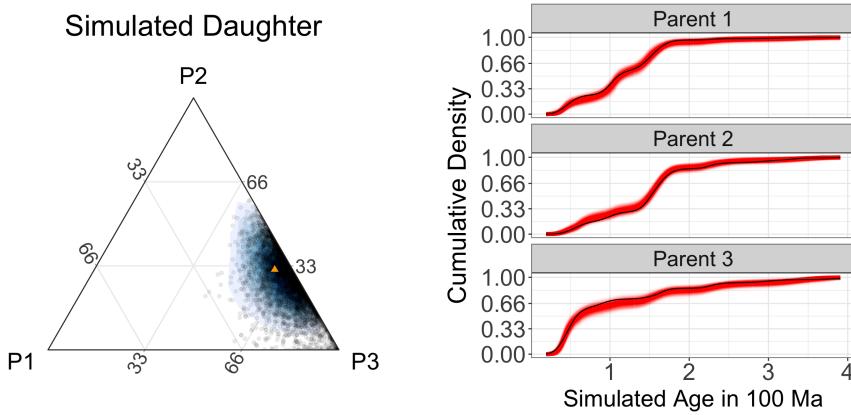
$$\mathbf{y}_d | \tilde{\mathbf{y}}_d, \boldsymbol{\sigma}_d^2 \sim N(\tilde{\mathbf{y}}_d, \text{diag}(\boldsymbol{\sigma}_d^2)),$$

311 where $\tilde{\mathbf{y}}_d$ is the true, unobserved n_d -vector of sediment dates and $\boldsymbol{\sigma}_d$ is a
 312 n_d -vector of dating uncertainty standard deviations.

313 Unlike in the top-down mixing model above, none of the parent **zs** are
 314 observed. Hence, the parent distributions are estimated entirely using child
 315 sediment observations. The end-member process model for a fixed given num-
 316 ber of parents B is

$$\tilde{y}_{id} | \boldsymbol{\phi}_d, \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2 \sim \boldsymbol{\phi}_d \sum_{b=1}^B \sum_{k=1}^K p_{bk} N(\mu_k, \sigma_k^2), \quad (4)$$

317 where, like before, we assume a Gaussian mixing distribution using shared
 318 kernels across the B parents. Like the top-down mixing model, these equations
 319 can be derived by introducing categorical random variables then marginalizing
 320 out the latent indicator variables from the model. The b th unknown parent
 321 age distribution is $\sum_{k=1}^K p_{bk} N(\mu_k, \sigma_k^2)$ and the posterior estimate ϕ_{bd} is the
 322 proportion of child d that comes from parent b . The prior model for the bottom-
 323 up unmixing model is the same as for the top-down mixing model, except for
 324 the dimension of different variables.



(a) Ternary plot showing posterior mixing proportion estimates as black circles and the simulated true mixing proportions as an orange triangle.

(b) Plot of simulated parents with fitted posterior CDF estimates in red and the simulated true CDF in black. Each red line represents a posterior sample of the cumulative parent age density.

Fig. 3: Simulation study results for the top-down mixture modeling approach.

325 4.2 Bottom-up unmixing posterior distribution

326 The posterior distribution that we estimate with the end member unmixing
327 model is

$$[\phi_1, \dots, \phi_D, \mu, \sigma^2, p | \mathbf{Y}] \propto \prod_{d=1}^D [\mathbf{y}_d | \tilde{\mathbf{y}}_d, \sigma_d] \times \quad (5)$$

$$\prod_{d=1}^D [\tilde{\mathbf{y}}_d | \phi_d, \mathbf{p}_d, \mu, \sigma] \times$$

$$\left(\prod_{d=1}^D [\phi_d | \alpha_{\phi d}] [\alpha_{\phi d}] \right) \left(\prod_{d=1}^D [\mathbf{p}_d | \alpha_b] [\alpha_b] [\mu_b] [\sigma_b] \right),$$

328 where the priors and MCMC algorithm are the same as those in (3) ex-
329 cept for a change in dimensionality. Code and data for replication of results
330 presented in this manuscript can be found freely available under the permissive
331 MIT license on GitHub at <https://github.com/jtipton25/mixing-manuscript>.

332 5 Simulation of synthetic detrital age distributions

333 We explore the performance of the model using a simulation of synthetic de-
334 trital age distributions. The aim of the simulation study is to understand how

335 the model performs using realistic data and verify the model is capable of recovering
 336 the simulated parameters. The simulation study framework can also
 337 be used to understand how uncertainty in estimation varies with respect to
 338 sample size, variability in the data, and other questions of interest.

339 First, we create synthetic data using the top-down mixing model in (3)
 340 for $B = 3$ parents and a single child (1a). The parent distributions were
 341 composed of 200, 250, and 150 simulated grain dates, respectively, and the
 342 child distribution was composed of 150 grain dates. In simulation, we used age
 343 dating uncertainties (σ_y, σ_z) that were about 1-3% of the total range of the
 344 age distribution. These are similar to measurement uncertainties in the case
 345 study and demonstrate the model is capable of accounting for measurement
 346 errors.

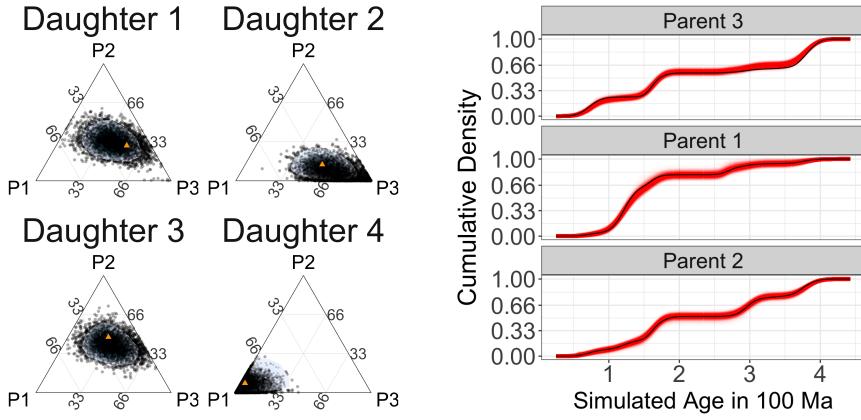
347 The posterior samples for the mixing proportion ϕ are shown in Fig.
 348 3a as black dots with a smoothed posterior density shown shaded in blue
 349 [Hamilton(2018)], and the simulated true mixing proportion is represented by
 350 the orange triangle, demonstrating that the model is accurately estimating
 351 the mixing proportion as the orange triangle is in the bulk of the posterior
 352 samples. Figure 3b shows the estimated CDFs with posterior samples in red
 353 and the simulated CDF in black. The results in Fig. 3 demonstrate that the
 354 model is accurately estimating the simulated mixing proportions ϕ as well
 355 as the parent age distributions, validating the effectiveness of the top-down
 356 mixing model to recover simulated parameters of interest.

357 The second simulation generated data from the bottom-up, end-member
 358 unmixing model (Fig. 1b) to test how well the proposed framework can re-
 359 construct unobserved parent distributions. For the simulation, we used $B = 3$
 360 parents and $D = 20$ children where each child consisted of 250 measured sed-
 361 iment grain dates following the model in (5). The dating uncertainties (σ_y)
 362 were set at about 1-3% of the total range of the age distribution.

363 Figure 4a shows the posterior samples for the mixing proportion of each
 364 child as black dots with the simulated mixing proportion plotted using an
 365 orange triangle. In general, the model can recover the mixing proportions in
 366 this simulation example with high precision. Even though the model uses none
 367 of the data from the parents, the end-member unmixing model produces rea-
 368 sonable end-member parent age distribution estimates. Figure 4b shows the
 369 estimated cumulative distribution function produced by the end-member un-
 370 mixing model which shows the bottom-up unmixing model is capturing the
 371 unobserved parent distributions. However, and not surprisingly, the accuracy
 372 for the bottom-up unmixing model is not as good as the fit that uses observa-
 373 tions from the parent distribution (i.e., the top-down mixture model) as can
 374 be seen in the slightly larger uncertainty estimates of the CDF.

375 6 Application to a Natural Case Study

376 We apply the model presented here to a well-constrained modern dataset from
 377 the central California coast [Sickmann et al.(2016) Sickmann, Paull, and Graham]



(a) Posterior estimates of mixing proportions for 4 of the 20 children from the unmixing model shown. The posterior samples are black circles and the simulated true mixing proportions are shown as orange triangles.

(b) Posterior estimates of the unobserved parent cumulative distribution functions in red. The simulated parent CDF is shown in black.

Fig. 4: Simulation study results for the bottom-up, end-member unmixing model. The bottom-up mixing model does a good job of estimating the true, unobserved parent age cumulative distribution functions despite the model not using any of the parent data.

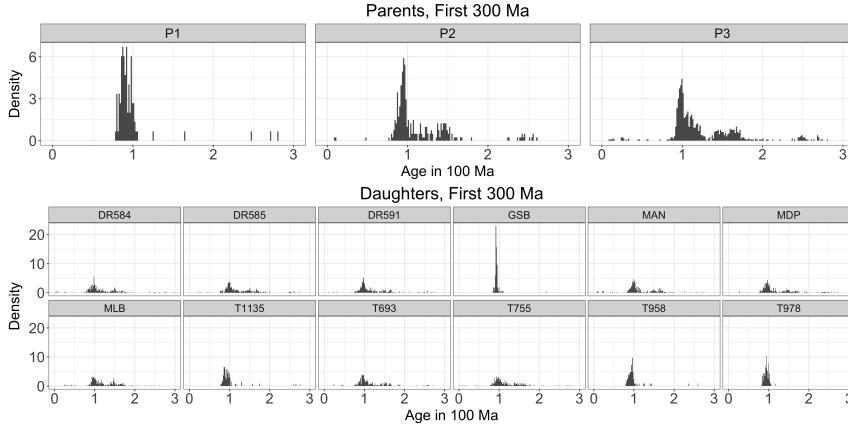


Fig. 5: The sediment age data used for the mixing and unmixing models. The three parent age distributions are shown in the top plot and the 12 child age distributions are shown in the bottom plot. The x -axes represent the measured age in 100 Ma and the y -axes shows the empirical density.

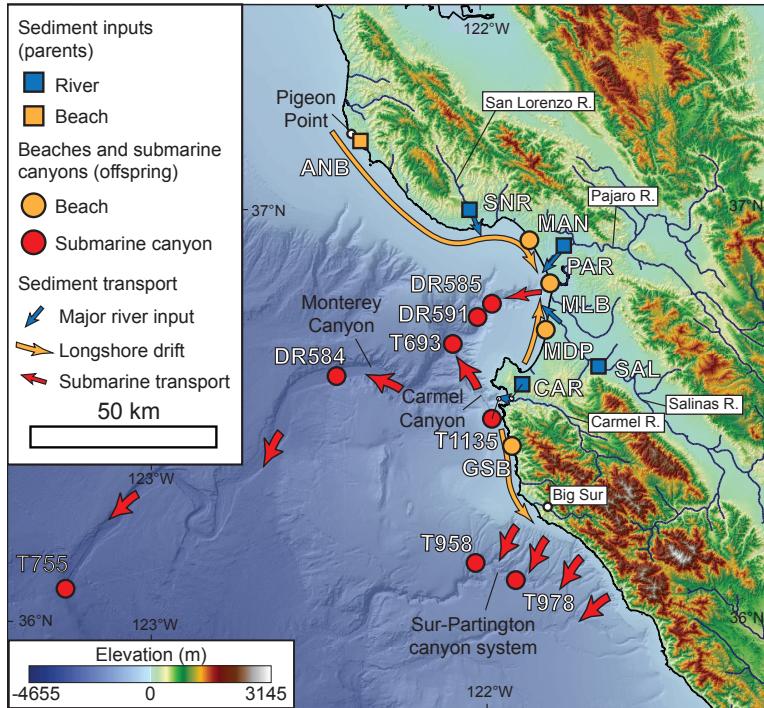
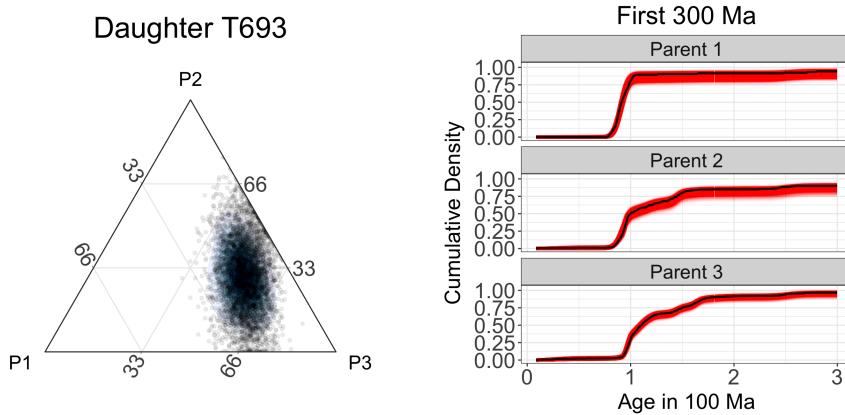


Fig. 6: Locations of the parents and children data for the study region in California, USA.

378 shown in Fig. 5 where we focus on sediments dating to the most recent 300 Ma.
 379 Following the same mixing framework presented in [Sharman and Johnstone(2017)],
 380 there are five samples (river and beach sediment) used to characterize three
 381 distinct sediment inputs (parents) to the region, each with a distinct detrital
 382 age distribution (Fig. 5). Parents 1 and 2 (P1 and P2) are comprised of river
 383 samples (CAR and SAR, respectively) that represent sediment sources along
 384 the Big Sur coastline and Salinas River drainage, respectively. Parent 3 (P3) is
 385 comprised of two river samples (SNR and PAR) and one beach sample (ANB)
 386 that represent northern sediment sources in the Santa Cruz Mountains and
 387 western Diablo Range [Sickmann et al.(2016) Sickmann, Paull, and Graham,
 388 Sharman and Johnstone(2017)]. Twelve child samples (beach and submarine
 389 canyon sediment) are used to characterize how these parents are mixed in lit-
 390 toral and marine environments. In total, this dataset (Fig. 6) consists of 4,026
 391 individual detrital zircon U-Pb analyses, with individual samples having 82 to
 392 316 analyses each (median of 290 analyses per sample) [Sickmann et al.(2016) Sickmann, Paull, and Graham].

393 We first examine the top-down mixture model (Fig. 1a). Figure 7a shows
 394 the reconstruction of the mixing proportions for a sample from a submarine
 395 canyon (T693) modeled as a mixture of the three specified parent distributions
 396 (P1-P3). Visual inspection of the histograms of the data (Fig. 5) would suggest



(a) Ternary plot showing posterior predictive density estimates of mixing proportions. Each dot represents one of 500 MCMC samples. The black shading is proportional to the estimated posterior probability density.

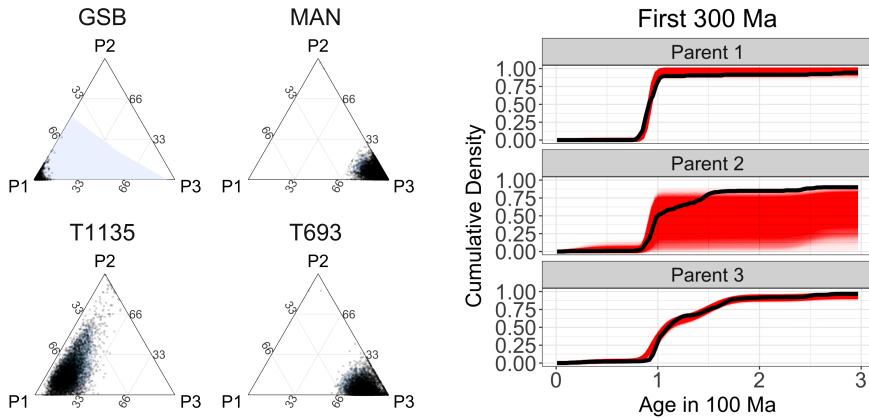
(b) Posterior estimates of the parent and child CDFs shown in red. The empirical CDFs calculated from the raw data are shown in black.

Fig. 7: Results from the top-down mixing applied to sample T693 show that the mixing model is able to accurately reconstruct the parent and child distributions and produce estimates of the mixing proportions with associated uncertainty.

397 that this child sample is a mixture composed mostly of P3 (a combination
 398 of the samples ANB, SNR, and PAR). The posterior estimates of the mixing
 399 proportion of each parent for child T693 confirms that the primary component
 400 of the mixture is from the P3 (Fig. 7a). Figure 7b shows the model is capturing
 401 the basic patterns in the parent CDFs as the estimated CDFs are very close
 402 to the empirical CDFs for the parents.

403 Application of the bottom-up, end-member unmixing model to the data set
 404 in [Sickmann et al.(2016) Sickmann, Paull, and Graham] results in non-identifiability
 405 issues as the simulated age distributions as evidenced by the larger posterior
 406 uncertainty in Fig. 8b. Posterior probability estimates are obtained for the
 407 proportion of each child that comes from the modeled end-member (Fig. 8a),
 408 but based on the differences in estimates from the bottom-up mixing model
 409 estimates, the distributions should be interpreted cautiously. The estimated
 410 CDFs shown in red in Fig. 8b with the empirical CDFs from the parent data
 411 shown in black suggest that the unmixing model is performing well despite
 412 the fact that the model is unaware of the parent data although there is much
 413 uncertainty about the CDF for P2, likely due to the similarity in distribution
 414 to P1 (Fig. 5).

415 A large overlap in the distribution of parent ages is a feature that oc-
 416 curs, particularly in detrital zircon geochronology studies. The preservation
 417 of zircons through multiple cycles of erosion and re-sedimentation means that
 418 overlapping zircon ages will be present in many rocks, for example the prepon-



(a) Posterior estimates for the mixing proportions of each parent for four child sediments. Notice that without observing the parents, the posterior distribution of mixing proportions for child T693 is generally similar to the top-down mixing model in Fig. 7a but has a slightly different shape.

(b) Posterior estimates for the unobserved parent cumulative distribution functions shown in red over 0-300 Ma. The black lines show the empirical cumulative distribution functions.

Fig. 8: Results of end-member unmixing model fit to real data. These figures show that the end-member unmixing model is estimating the parameters of interest, but with some inaccuracies due to a lack of identifiability.

derance of Grenville zircons in a host of sedimentary formations of varying age. For parent age distributions that are quite similar to one another, the reconstruction of the unknown parent distributions suffers from weak identifiability. In these situations, the estimated parents jointly contain all of the correct formation events, but the model is unable to attribute the formation events to the correct parents. In other words, while the model identifies the correct age components, the model sometimes struggles to correctly group these components into the correct parent distributions. This is not an unexpected result because Bayesian nonparametric models are well understood to suffer from non-identifiability issues in the context of the Bayesian nonparametric framework [Ferguson(1983), Diebolt and Robert(1994), Richardson and Green(1997), Frühwirth-Schnatter(2006)]. Non-identifiability is not a weakness of the particular proposed model framework; the non-identifiability applies to end-member unmixing models in general [Weltje and Prins(2007)]. To overcome the non-identifiability, a potential solution is to impose constraints on the end-members and initialization conditions [Donoho and Stodden(2004), Miao and Qi(2007), Chen and Guillaume(2012)]. Therefore, any end-member unmixing model that uses only child age distributions will have issues in accurately reconstructing the parent distributions if the assumption of the constraints is not met (i.e., the parent age distributions are structurally similar). Bottom-up unmixing models provide a useful way to explore large detrital datasets with unknown sedimentary sources. Providing

440 a way to identify those datasets that either are or are not susceptible to non-
441 identifiability, and thus not amenable to bottom-up unmixing, is critical to
442 success. An advantage of our framework is the end-member unmixing model
443 produces uncertainty estimates that are larger when the model is weakly iden-
444 tifiable. Thus, the uncertainty intervals can be used as a diagnostic to check
445 for identifiability.

446 **7 Conclusion**

447 Starting from a conceptual model of how sediments mix over a landscape, we
448 developed a generative Bayesian nonparametric statistical model for detrital
449 mineral age data. This model allows us to characterize the uncertainty in the
450 age distributions of parents and children and the mixing coefficients while ex-
451 plicitly accounting for the uncertainties in measured dates [Tye et al.(2019)Tye, Wolf, and Niemi].
452 Because the model can simulate sediment age distributions, we can directly
453 explore the assumptions of the model by simulating synthetic data. Running
454 a simulation experiment demonstrated the model is capable of recovering sim-
455 ultated distributions supporting the usefulness of the models when applied to
456 the observed data.

457 We proposed two frameworks to model the sediment mixing mechanisms:
458 the top-down mixing model where mineral dates are measured for both par-
459 ent and child sediments and a bottom-up unmixing framework where mineral
460 dates are only measured for the children. The top-down model estimated the
461 parent and child distributions and the mixing proportions with high precision
462 and accuracy. The bottom-up model occasionally demonstrated evidence of
463 non-identifiability, suggesting the inference for the bottom-up model is less
464 precise than for the top-down mixing model; however, the variances of these
465 estimates are larger in our bottom-up unmixing model, recognizing the chal-
466 lenges in reconstructing unobserved parent age distributions while simultane-
467 ously providing feedback to the user about the potential pitfalls in being overly
468 confident about the parent distributions.

469 Using data collected along the central California coast data in [Sickmann et al.(2016)Sickmann, Paull, and Graham],
470 we produced estimates of the mixing proportions of child sediments with both
471 the mixing and unmixing models. Other studies have produced similar esti-
472 mates [Amidon et al.(2005a)Amidon, Burbank, and Gehrels, Amidon et al.(2005b)Amidon, Burbank, and Gehrels];
473 our contribution is the mechanistic model framework that produces estimates
474 of mixing with associated uncertainty. We account for dating uncertainty di-
475 rectly in the model and by modifying the model statement to remove dating
476 uncertainty (i.e., removing the data models in Eq. (1)) we can examine the
477 effects on inference by not accounting for the dating uncertainty.

478 Direct, probabilistic estimates of uncertainty and the ability to calculate
479 derived quantities with uncertainty is a benefit of the Bayesian methodol-
480 ogy. Thus, we can answer questions like what is the probability that at least
481 50% of child sediment T693 comes from parent P3? The answer is calcu-
482 lated directly from the posterior samples using the Monte Carlo approxima-

tion $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3}^{(\ell)} \geq 0.5\} = 0.672$, where $\ell = 1, \dots, L$ are the indices of the MCMC samples and $\phi_{P3}^{(\ell)}$ is the estimated mixing proportion for the ℓ th MCMC iteration. The probability that at least 50% of the child sediment comes from parent P3 and at least 25% comes from parent P2 is $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3} \geq 0.5\} \times I\{\phi_{P2} \geq 0.25\} = 0.283$. Because the model produces a posterior probability, any such probabilistic questions can be calculated as derived quantities. For example, we can ask questions like: what proportion of a given sample contains grains older than a given age? or what is the probability that an unobserved parent contains grains with a particular age range. Once the posterior samples have been calculated, any such questions about derived quantities are answerable using posterior samples.

In addition, the ability to include prior information in the Bayesian framework is a useful tool that can be used to improve estimation and test geologic hypotheses. For example, certain geologic events, such as the Grenville orogeny, produced large amounts of zircon that have since been broadly dispersed and recycled in sedimentary rocks. Priors that account for the likelihood of observing zircons of Grenville-age (or other known zircon-producing events) can be introduced into this model framework to improve model performance. In addition, our framework can accommodate a variety of detrital data with different magnitudes of uncertainty. As analytic techniques for dating minerals improve, it is important to account for dating uncertainties that might have orders of magnitude difference making our method more robust to future improvements in analytic laboratory techniques.

506 8 Declarations

507 The authors declare that they have no known competing financial interests or
508 personal relationships that could have appeared to influence the work reported
509 in this paper. Support for SAJ came from the FEDMAP component of the
510 US Geological Survey National Cooperative Geologic Mapping Program. This
511 draft manuscript is distributed solely for purposes of scientific peer review.
512 Its content is deliberative and predecisional, so it must not be disclosed or
513 released by reviewers. Because the manuscript has not yet been approved for
514 publication by the U.S. Geological Survey (USGS), it does not represent any
515 official USGS finding or policy. Any use of trade, firm, or product names is
516 for descriptive purposes only and does not imply endorsement by the U.S.
517 Government.

518 Code and data for replication of results presented in this manuscript can be
519 found freely available under the permissive MIT license on GitHub at <https://github.com/jtipton25/mixing-manuscript>.

521 8.1 CRediT authorship contribution statement

522 **John R. Tipton:** conceptualization, model development, software, writing -
523 original draft. **Glenn R. Sharman:** conceptualization, model development,
524 writing - review and editing. **Samuel A. Johnstone:** conceptualization, writ-
525 ing - review and editing.

526 **Keywords** Detrital sediment age distributions · Sediment unmixing ·
527 Bayesian nonparametrics · Uncertainty quantification

528 **References**

- 529 Amidon et al.(2005a)Amidon, Burbank, and Gehrels. Amidon WH, Burbank DW, Gehrels
530 GE (2005a) Construction of detrital mineral populations: insights from mixing of U-Pb
531 zircon ages in Himalayan rivers. Basin Research 17(4):463–485
532 Amidon et al.(2005b)Amidon, Burbank, and Gehrels. Amidon WH, Burbank DW, Gehrels
533 GE (2005b) U–Pb zircon ages as a sediment mixing tracer in the Nepal Himalaya. Earth
534 and Planetary Science Letters 235(1-2):244–260
535 Berliner(2003). Berliner LM (2003) Physical-statistical modeling in geophysics. Journal of
536 Geophysical Research: Atmospheres 108(D24)
537 Chang et al.(2016)Chang, Haran, Applegate, and Pollard. Chang W, Haran M, Applegate
538 P, Pollard D (2016) Calibrating an ice sheet model using high-dimensional binary spatial
539 data. Journal of the American Statistical Association 111(513):57–72
540 Chen and Moore(1982). Chen JH, Moore JG (1982) Uranium-lead isotopic ages from the
541 Sierra Nevada batholith, California. Journal of Geophysical Research: Solid Earth
542 87(B6):4761–4784
543 Chen and Guillaume(2012). Chen W, Guillaume M (2012) HALS-based NMF with flexi-
544 ble constraints for hyperspectral unmixing. EURASIP Journal on Advances in Signal
545 Processing 2012(1):54
546 de Valpine et al.(2017)de Valpine, Turek, Paciorek, Anderson-Bergman, Temple Lang, and Bodik.
547 de Valpine P, Turek D, Paciorek C, Anderson-Bergman C, Temple Lang D, Bodik R
548 (2017) Programming with models: writing statistical algorithms for general model struc-
549 tures with NIMBLE. Journal of Computational and Graphical Statistics 26:403–413,
550 DOI 10.1080/10618600.2016.1172487
551 Diebolt and Robert(1994). Diebolt J, Robert CP (1994) Estimation of finite mixture dis-
552 tributions through Bayesian sampling. Journal of the Royal Statistical Society Series B
553 (Methodological) pp 363–375
554 Donoho and Stodden(2004). Donoho D, Stodden V (2004) When does non-negative matrix
555 factorization give a correct decomposition into parts? In: Advances in Neural Informa-
556 tion Processing Systems, pp 1141–1148
557 Ferguson(1973). Ferguson TS (1973) A Bayesian analysis of some nonparametric problems.
558 The Annals of Statistics pp 209–230
559 Ferguson(1983). Ferguson TS (1983) Bayesian density estimation by mixtures of normal
560 distributions. In: Recent Advances in Statistics, Elsevier, pp 287–302
561 Fosdick et al.(2015)Fosdick, Grove, Graham, Hourigan, Lovera, and Romans. Fosdick JC,
562 Grove M, Graham SA, Hourigan JK, Lovera O, Romans BW (2015) Detrital ther-
563 mochronologic record of burial heating and sediment recycling in the Magallanes fore-
564 land basin, Patagonian Andes. Basin Research 27(4):546–572
565 Frühwirth-Schnatter(2006). Frühwirth-Schnatter S (2006) Finite mixture and Markov
566 switching models. Springer Science & Business Media
567 Gehrels(2012). Gehrels G (2012) Detrital zircon U-Pb geochronology: Current methods and
568 new opportunities. Tectonics of Sedimentary Basins: Recent Advances pp 45–62
569 Gehrels(2014). Gehrels G (2014) Detrital zircon U-Pb geochronology applied to tectonics.
570 Annual Review of Earth and Planetary Sciences 42:127–149

- 571 Green(1995). Green PJ (1995) Reversible jump Markov chain Monte Carlo computation
572 and Bayesian model determination. *Biometrika* 82(4):711–732
- 573 Guan et al.(2018)Guan, Haran, and Pollard. Guan Y, Haran M, Pollard D (2018) Inferring
574 ice thickness from a glacier dynamics model and multiple surface data sets. *Environ-
575 metrics* 29(5-6):e2460
- 576 Haario et al.(2001)Haario, Saksman, and Tamminen. Haario H, Saksman E, Tamminen J
577 (2001) An adaptive Metropolis algorithm. *Bernoulli* 7(2):223–242
- 578 Hamilton(2018). Hamilton N (2018) ggtern: An Extension to ggplot2, for the creation of
579 ternary diagrams. URL <https://CRAN.R-project.org/package=ggtern>, r package ver-
580 sion 2.2.2
- 581 Hefley et al.(2017)Hefley, Brost, and Hooten. Hefley TJ, Brost BM, Hooten MB (2017)
582 Bias correction of bounded location errors in presence-only data. *Methods in Ecology
583 and Evolution* 8(11):1566–1573
- 584 Hooten and Hobbs(2015). Hooten MB, Hobbs N (2015) A guide to Bayesian model selection
585 for ecologists. *Ecological Monographs* 85(1):3–28
- 586 Irwin and Wooden(1999). Irwin WP, Wooden JL (1999) Plutons and accretionary episodes
587 of the Klamath Mountains, California and Oregon. Tech. rep., US Geological Survey
- 588 Jasra et al.(2006)Jasra, Stephens, Gallagher, and Holmes. Jasra A, Stephens DA, Gal-
589 lagher K, Holmes CC (2006) Bayesian mixture modelling in geochronology via Markov
590 chain Monte Carlo. *Mathematical Geology* 38(3):269–300
- 591 Kimbrough et al.(2015)Kimbrough, Grove, Gehrels, Dorsey, Howard, Lovera, Aslan, House, and Pearthree.
592 Kimbrough DL, Grove M, Gehrels GE, Dorsey RJ, Howard KA, Lovera O, Aslan A,
593 House PK, Pearthree PA (2015) Detrital zircon U-Pb provenance of the Colorado
594 River: A 5 my record of incision into cover strata overlying the Colorado Plateau and
595 adjacent regions. *Geosphere* 11(6):1719–1748
- 596 Lock and Dunson(2015). Lock EF, Dunson DB (2015) Shared kernel Bayesian screening.
597 *Biometrika* 102(4):829–842
- 598 Mason et al.(2017)Mason, Fildani, Gerber, Blum, Clark, and Dykstra. Mason CC, Fildani
599 A, Gerber T, Blum MD, Clark JD, Dykstra M (2017) Climatic and anthropogenic
600 influences on sediment mixing in the Mississippi source-to-sink system using detrital
601 zircons: Late Pleistocene to recent. *Earth and Planetary Science Letters* 466:70–79
- 602 Miao and Qi(2007). Miao L, Qi H (2007) Endmember extraction from highly mixed data us-
603 ing minimum volume constrained nonnegative matrix factorization. *IEEE Transactions
604 on Geoscience and Remote Sensing* 45(3):765–777
- 605 Miller and Harrison(2018). Miller JW, Harrison MT (2018) Mixture models with a prior
606 on the number of components. *Journal of the American Statistical Association*
607 113(521):340–356
- 608 Puetz et al.(2018)Puetz, Ganade, Zimmermann, and Borchardt. Puetz SJ, Ganade CE,
609 Zimmermann U, Borchardt G (2018) Statistical analyses of global U-Pb database 2017.
610 *Geoscience Frontiers* 9(1):121–145
- 611 Reiners and Brandon(2006). Reiners PW, Brandon MT (2006) Using thermochronology to
612 understand orogenic erosion. *Annual Review Earth Planetary Sciences* 34:419–466
- 613 Richardson and Green(1997). Richardson S, Green PJ (1997) On Bayesian analysis of mix-
614 tures with an unknown number of components (with discussion). *Journal of the Royal
615 Statistical Society: series B (statistical methodology)* 59(4):731–792
- 616 Romans et al.(2016)Romans, Castelltort, Covault, Fildani, and Walsh. Romans BW,
617 Castelltort S, Covault JA, Fildani A, Walsh J (2016) Environmental signal propagation
618 in sedimentary systems across timescales. *Earth-Science Reviews* 153:7–29
- 619 Saylor et al.(2019)Saylor, Sundell, and Sharman. Saylor JE, Sundell K, Sharman G (2019)
620 Characterizing sediment sources by non-negative matrix factorization of detrital
621 geochronological data. *Earth and Planetary Science Letters* 512:46–58
- 622 Sharman and Johnstone(2017). Sharman GR, Johnstone SA (2017) Sediment unmixing us-
623 ing detrital geochronology. *Earth and Planetary Science Letters* 477:183–194
- 624 Sharman et al.(2019)Sharman, Sylvester, and Covault. Sharman GR, Sylvester Z, Covault
625 JA (2019) Conversion of tectonic and climatic forcings into records of sediment supply
626 and provenance. *Scientific reports* 9(1):4115
- 627 Sickmann et al.(2016)Sickmann, Paull, and Graham. Sickmann ZT, Paull CK, Graham SA
628 (2016) Detrital-zircon mixing and partitioning in fluvial to deep marine systems, Central
629 California, USA. *Journal of Sedimentary Research* 86(11):1298–1307

- 630 Stock et al.(2006)Stock, Ehlers, and Farley. Stock GM, Ehlers TA, Farley KA (2006)
631 Where does sediment come from? Quantifying catchment erosion with detrital apatite
632 (U-Th)/He thermochronometry. *Geology* 34(9):725–728
- 633 Sundell and Saylor(2017). Sundell K, Saylor JE (2017) Unmixing detrital geochronology
634 age distributions. *Geochemistry, Geophysics, Geosystems*
- 635 Tipton et al.(2017)Tipton, Hooten, and Goring. Tipton J, Hooten M, Goring S (2017) Re-
636 construction of spatio-temporal temperature from sparse historical records using robust
637 probabilistic principal component regression. *Advances in Statistical Climatology, Me-
638 teorology and Oceanography* 3(1):1–16
- 639 Tipton et al.(2016)Tipton, Hooten, Pederson, Tingley, and Bishop. Tipton JR, Hooten
640 MB, Pederson N, Tingley MP, Bishop D (2016) Reconstruction of late Holocene climate
641 based on tree growth and mechanistic hierarchical models. *Environmetrics* 27(1):42–54
- 642 Tipton et al.(2019)Tipton, Hooten, Nolan, Booth, and McLachlan. Tipton JR, Hooten
643 MB, Nolan C, Booth RK, McLachlan J (2019) Predicting paleoclimate from composi-
644 tional data using multivariate Gaussian process inverse prediction. *Annals of Applied
645 Statistics* 13(4):2363–2388, DOI 10.1214/19-AOAS1281
- 646 Tye et al.(2019)Tye, Wolf, and Niemi. Tye A, Wolf A, Niemi N (2019) Bayesian population
647 correlation: A probabilistic approach to inferring and comparing population distribu-
648 tions for detrital zircon ages. *Chemical Geology* 518:67–78
- 649 Vermeesch(2012). Vermeesch P (2012) On the visualisation of detrital age distributions.
Chemical Geology 312:190–194
- 650 Weltje(1997). Weltje GJ (1997) End-member modeling of compositional data: Numerical-
651 statistical algorithms for solving the explicit mixing problem. *Mathematical Geology*
652 29(4):503–549
- 653 Weltje and Prins(2007). Weltje GJ, Prins MA (2007) Genetically meaningful decomposition
654 of grain-size distributions. *Sedimentary Geology* 202(3):409–424
- 655 Wotzlaw et al.(2013)Wotzlaw, Schaltegger, Frick, Dungan, Gerdes, and Günther. Wotzlaw
656 JF, Schaltegger U, Frick DA, Dungan MA, Gerdes A, Günther D (2013) Tracking the
657 evolution of large-volume silicic magma reservoirs from assembly to supereruption.
Geology 41(8):867–870