

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

4 John R. Tipton^{a,*}, Glenn R. Sharman^b, Samuel A. Johnstone^c

5 ^a*University of Arkansas, Department of Mathematical Sciences, Fayetteville, AR, USA*

6 ^b*University of Arkansas, Department of Geosciences, Fayetteville, AR USA*

7 ^c*U.S. Geological Survey, Geosciences and Environmental Change Science Center, Denver,*

8 *USA*

9 **Abstract**

Sedimentary deposits constitute the primary record of changing environmental conditions that have acted on Earth's surface over geologic time. Clastic sediment is eroded from source locations (parents) in sediment routing systems and deposited at sink locations (children). Both parents and children have characteristics that vary across many different dimensions, including grain size, chemical composition, and the geochronologic age of constituent detrital minerals. During transport, sediment from different parents is mixed together to form a child, which in turn may serve as the parent for other sediment further down system or later in time when buried sediment is exhumed. To the extent that parent sources produce sediment with distinguishable geochronologic ages, the distribution of detrital mineral ages observed in child sediments allows for investigation of the proportions of each parent in the child sediment which ultimately reflects properties of the sediment routing system, such as the relative sediment flux. To model the proportion of dates in a child sample that comes from each of the parent distributions, we use a Bayesian mixture of Dirichlet processes. This model allows for estimation of the mixing proportions with associated uncertainty while making minimal assumptions. We also present an extension to the model whereby we reconstruct unobserved parent distributions from multiple observed child distributions using mixtures of Dirichlet processes, accounting for uncertainty in both the number of parent distributions and the mixing proportions.

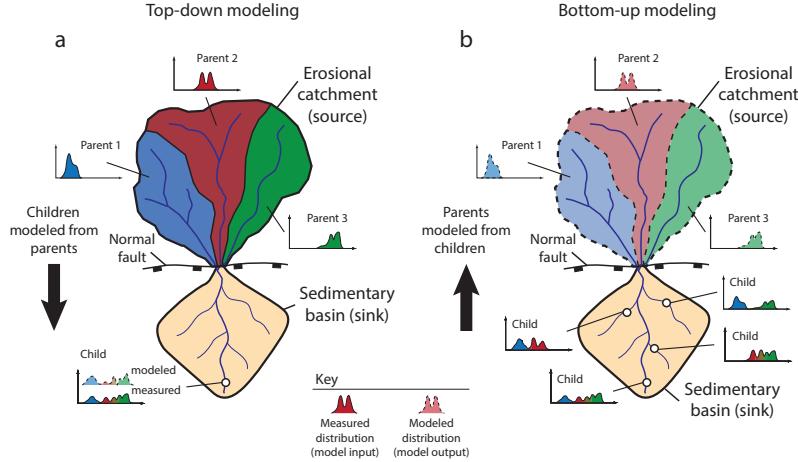


Figure 1: Schematic depiction of a sediment routing system with an erosional source region characterized by three parents (black, red, and green) and an associated sedimentary basin (yellow). Call out plots represent the age of detrital minerals from the parents and how they are mixed to form child distributions. (a) Top-down modeling (*sensu* [34]) 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.

1. Introduction

To understand the origins of modern and ancient physical geography one must understand how erosional landscapes and associated sedimentary basins evolve through time [32]. As clastic sediment is generated by weathering and erosion, it is subsequently transported downstream, mixed, and ultimately deposited into a depositional sink. Modeling sediment mixing allows inference about the processes that generated the modern landscape. The ability to decipher the relative proportions of sources that were eroded to produce sediment may thus inform understanding of the underlying geologic processes that control the evolution of the Earth's surface [37, 35, 26, 24].

One of the most common ways to characterize the provenance of sediment

*Corresponding Author
Email address: jrtipton@quark.edu (John R. Tipton)

is through detrital geochronology, dating the timing at which the individual minerals that make up a rock formed or cooled. These mineralization events typically reflect the timing of igneous rock forming events or metamorphic alteration of previously existing rocks [15]. In other cases mineral ages reflect the history of mineral cooling (e.g. ‘thermochronology’, [30]). Detrital geochronologic ages are most commonly determined from measurements of radiogenic isotopes contained within individual mineral crystals. The decay of uranium (U) to lead (Pb) within zircon, a relatively robust mineral, makes this approach ideally suited for tracking sedimentary mixing [2, 38, 34].

In the following text, we will follow the convention that the sources of sediment are called *parents* and the sink locations are called *children*. Using this language, the manuscript aims to address two questions. First, can we estimate the proportion of each parent age distribution in a child age distribution with associated uncertainty? Second, can we estimate the marginal age distributions for unobserved parents given a set of child age distributions? These questions may be answered using “top-down” and “bottom-up” approaches to sediment unmixing, respectively (see Figure 1; [34]). The top-down approach models one or more child samples as mixtures of specified parent samples (Figure 1). The bottom-up approach uses multiple child samples to model likely parents which are more generally referred to as end-members in mixture modeling efforts. [34].

We present a model framework that is capable of answering the two questions above using top-down and bottom-up approaches (Figure 1). Bayesian mixture modeling of geochronology data, including detrital data, has numerous advantages when addressing single samples [23] including allowing inference and uncertainty estimates of the number and value of true ages characterized by observed mineral dates. Here we extend this concept to consider the geologic mixing of sediments derived from source areas containing minerals recording different

48 crystallization events. The Bayesian nonparametric statistical model presented
49 herein has a number of advantages over previously used approaches, including
50 being able to derive direct, probabilistic estimates of uncertainty associated with
51 the mixture model. We demonstrate the utility of this approach using both a
52 synthetic dataset and a well-constrained, natural case study in central California,
53 USA [36]. The top-down mixing approach is able to successfully reconstruct
54 parent contributions in both synthetic and natural datasets. Although the
55 bottom-up unmixing model is able to successfully reconstruct parents in the
56 synthetic dataset, there is evidence of non-identifiability, where parents cannot
57 be uniquely characterized from the children data, when applied to the natural
58 dataset. More generally, the model framework we present can also give guidance
59 about other scientific questions that relate to mixing of non-parametric sum-to-
60 one data in Earth sciences and other disciplines (e.g. unmixing sediment grain
61 size distribution; [44] and references within)

62 2. Model Overview

63 To define the statistical model, we follow the convention that letters represent
64 data and Greek symbols represent parameters. A plaintext symbol (y) represents
65 a scalar, a bold lowercase symbol represents a vector (\mathbf{y}), and a bold uppercase
66 symbol is a matrix (\mathbf{Z}) whose columns are vectors written as $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_p)$.
67 We use the notation $[y]$ to represent the probability distribution/mass function
68 (pdf/pmf) and let $[y|\theta]$ represent the conditional pdf/pmf of the random variable
69 y given θ .

70 Following [3], the statistical model described below is divided into three
71 components: the data model, the process model, and the prior model. In general,
72 the data model defines probability distributions that describe the variability in the
73 data due to the observation process. The data model can be modified to account

74 for non Gaussian measurement processes like counts, outliers, spatial/temporal
75 errors, etc [39, 20]. Process models describe the best scientific understanding of
76 the process of interest. For example, process models have been used to model
77 monthly response of trees to climate [41], the relationship between climate and
78 pollen in sediments [40], and the movement of ice sheets in Antarctica [4, 17]. The
79 prior model describes the range of parameter values that are plausible. Sometimes
80 the prior model is used as regularization to improve the generalization of the
81 model to unobserved data [21]. Following these conventions, we introduce the
82 data models for the sediment age distributions, process models for the mixing of
83 parents, and prior models for model selection, regularization, and mathematical
84 completeness.

85 **3. Top-down mixing model**

86 The model framework presented below, which is appropriate for situations
87 where the parent and children sediment have been independently characterized,
88 will answer the first research question: can one estimate the proportion of each
89 parent that comprises the child sediment?

90 *3.1. Top-down mixing data model*

91 Let the n_y observed age measurements of a single child of interest be \mathbf{y} and
92 let the observed date measurements for each of the $b = 1, \dots, B$ parents be given
93 by the n_b -dimensional vector \mathbf{z}_b . Because the observed ages are measured with
94 uncertainty reported as a dating error standard deviation, we explicitly account
95 for this source of uncertainty in the data model.

96 In the case of U-Pb dating of detrital zircon grains, dates are most commonly
97 determined using laser ablation-inductively coupled plasma-mass spectrometry
98 [14]. Such date measurements typically have relative 2σ analytical precision of
99 1-4%, with relative uncertainty increasing for younger analyses [29].

100 For each detrital mineral, the estimate of dating measurement uncertainty is
 101 reported as a n_y -vector of standard deviations σ_y for the child and B n_b -vectors
 102 of standard deviations σ_{z_b} . We assume the date measurement uncertainty follows
 103 a Gaussian distribution where the observed sediment particle date is

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

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

112 *3.2. Top-down mixing process model*

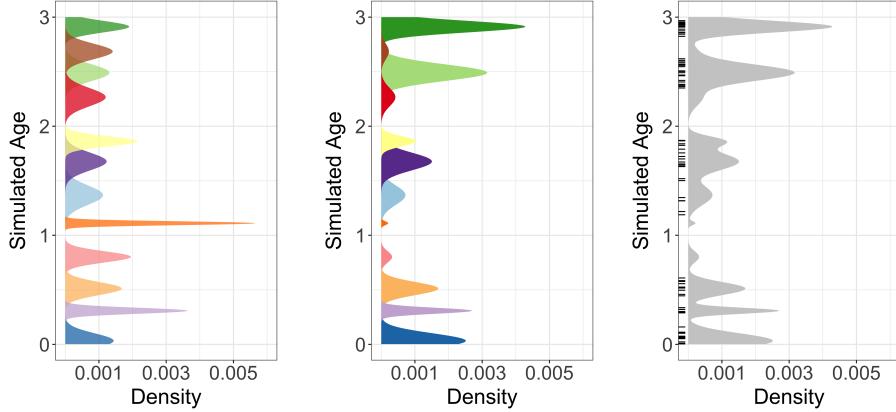
113 The process model addresses two scientific questions: what are the estimates
 114 of the true, unobserved detrital mineral age distributions at the parent and child
 115 locations? and what proportions of those detrital minerals did each parent source
 116 contribute to the child? There are many different methods available to model
 117 the true geochronological age distributions from the sample data, including
 118 kernel density estimates [43], non-negative matrix factorization methods [33],
 119 and Bayesian nonparametric models of mineral fomation event mixing [23, 42];
 120 however, Bayesian methods have yet to be applied to the problem of sediment
 121 mixing. We develop a Bayesian nonparametric model below that approximates

122 the geologic understanding of the data generating process.

123 Over geologic time, individual minerals may be repeatedly recycled into
124 sedimentary rocks by cycles of erosion, transport, deposition, and exhumation.
125 However in many cases the dates recorded by individual minerals contained
126 in these deposits are distinctive and unaffected by these recycling processes
127 (e.g., excluding burial reheating of thermochronometers [12]). We assume that
128 minerals created by the same geologic event share an age distribution that is
129 relatively homogeneous with only small variability. Furthermore, episodes of
130 rock and mineral formation (typically lasting 10^5 to 10^7 years [5, 22, 46]) are
131 nearly discrete events relative to geologic time (4.5×10^9 years). While minerals
132 often show overgrowths of different ages, this provides a useful approximation.
133 Under the conceptual model (Figure 1), sediment is formed by the decomposition
134 of rocks containing minerals created at different times, and sediment at every
135 child location is composed of an unknown number of mineral formation events
136 that are also present at source locations.

137 Consider the latent, unknown age of a single mineral grain from either the
138 child $\tilde{\mathbf{y}}_b$ or one of the $b = 1, \dots, B$ parents $\tilde{\mathbf{z}}_b$. We assign the range of mineral
139 ages for the k th formation event the base probability distribution $G(\boldsymbol{\theta}_{bk})$ which
140 depends on parameters $\boldsymbol{\theta}_{bk}$. There are many possible choices for the base
141 distribution $G(\boldsymbol{\theta}_{bk})$; we assume a normal distribution $N(\mu_{bk}, \sigma_{bk}^2)$ with mean
142 μ_{bk} and variance σ_{bk}^2 , therefore $\boldsymbol{\theta}_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$. Other possible choices include
143 a log-normal or gamma distribution that enforces a positive support on the
144 observed age dates. Because we assume the age distribution of a single mineral
145 formation event is relatively short with respect to geologic time, the variance
146 parameters σ_{bk}^2 will be small relative to the age of the mineral formation event.

147 In most cases the true number of mineral formation events K recorded by a
148 detrital sample is unknown. Under our model, the set of all K mineral formation



(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 Figure 2a are events in the parent distribution in Figure 2b are re-weighted to account for all of the factors that determine the distribution of potentially observable mineral ages provided by each parent.

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

Figure 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.

events is a mixture of K normal distributions, as shown in Figure 2a where each mineral formation event is shown in a different color. The centers of each age distribution in Figure 2a are given by the values of μ_{bk} and the spreads of each age distribution are given by the variances σ_{bk}^2 . Aerial extent, differential erosion, and the abundance of minerals of different ages within different rocks, and other factors can influence the proportion of minerals of a given age in rock at a site [1, 2]. Figure 2b shows the age distributions in Figure 2a that have been re-weighted to account for all of the factors that determine the distribution of ages in a parent source rock. We do not observe the individual mineral formation event labels ($k = 1, \dots, K$) directly, we only observe the parent age distributions after influence from the relative contributions of minerals of different formation ages (Figure 2c). Notice that in Figure 2c, the colors from Figures 2a and 2b are removed, representing the fact that the mineral event labels are not observed. As

¹⁶² a futher complication, we only observe a finite sample from the distribution in
¹⁶³ Figure 2c which is shown by the rug plot with each tick representing the observed
¹⁶⁴ date of each detrital mineral grain. Thus, the number of mineral formation
¹⁶⁵ events is potentially challenging to extract from the data and is modeled using a
¹⁶⁶ Dirichlet process mixture described in Section 3.3.

¹⁶⁷ Consider a detrital mineral date \tilde{z}_{ib} from a parent sediment source b out of
¹⁶⁸ a set of $i = 1, \dots, n_b$ measurements. The single sediment grain comes from a
¹⁶⁹ single mineral formation event implying the origin of the mineral is described by
¹⁷⁰ the mixture distribution

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

¹⁷¹ where γ_{ib} is a random variable whose value acts as a label indicating which
¹⁷² formation event k the detrital mineral comes from. We define the probability
¹⁷³ of a detrital mineral coming from age of formation event k as $p_{bk} \equiv P(\gamma_{ib} = k)$.
¹⁷⁴ Then, we write the joint distribution over all mineral grains from parent b as

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

¹⁷⁵ where $I\{\gamma_{ib} = k\}$ is an indicator function that takes the value 1 if $\gamma_{ib} = k$
¹⁷⁶ and 0 otherwise. Because there are a large number of indicator functions, we
¹⁷⁷ integrate them out of the process model to improve mixing and model fit. The
¹⁷⁸ integrated age distribution model is

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

179 where $\mathbf{p}_b = (p_{b1}, \dots, p_{bK})'$ is a mixing vector of mixing probabilities with
 180 $\sum_{k=1}^K p_{bk} = 1$. When the number of formation events K is potentially infinite,
 181 the conceptual model described above can be described with a Dirichlet process.
 182 The Dirichlet process model is described in detail in Section 3.3.

183 In addition to modeling the age distribution of the parents, the process
 184 model specifies how we infer the proportion of each parent distribution in the
 185 child age distribution. We represent the mixing proportions of the B parent
 186 distributions with the parameter $\phi = (\phi_1, \dots, \phi_B)'$, with $\sum_{b=1}^B \phi_b = 1$. The
 187 parameter ϕ_b is the proportion of the child distribution that comes from parent
 188 b . The parameter ϕ accounts for differential mixing of parents. For parents that
 189 are comprised of bedrock, ϕ is a function of each parent's relative aerial extent
 190 in the drainage catchment, average erosion rate, and average concentration of
 191 the detrital mineral of interest [1]. If parents are sediment inputs (e.g., rivers),
 192 then ϕ is a function of each parent's relative sediment supply and the average
 193 concentration of the detrital mineral of interest within the sediment.

194 For a single child sediment mineral date \tilde{y}_i out of $i = 1, \dots, n_y$ observations,
 195 that sediment grain can only be from one parent. We define the categorical
 196 random variable δ_i to represent which parent distribution the child sediment
 197 came from. Using the categorical variable, the distribution of the child sediment
 198 grain is

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

199 Then, the probability that the sediment grain came from parent b is $\phi_b \equiv$
200 $P(\delta_i = b)$. Defining the indicator variable $I\{\delta_i = b\}$ where $P(\delta_i = b) =$
201 $E(I\{\delta_i = b\})$, we can write the age distribution over all sediment grains as

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

202 where, like the parent mixing model, we integrate out the component indicator
203 variables $\boldsymbol{\delta}$. After integrating out the parent component membership indicators,
204 the child sediment grains have the age distribution

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

205 where the parameter $\phi_b = P(\delta_i = b)$ is the probability that any individual
206 grain comes from the b th parent source and the posterior distribution of ϕ_b
207 estimate the proportion of child sediment from parent b .

208 Combining the above results, the full process model is

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

209 3.3. *Top-down mixing prior model*

210 The conceptual process model assumed that we knew the number of mineral
211 formation events K . In practice, the number of formation events is unknown and
212 is a parameter to be estimated in the model. In fact, it is likely that the different
213 parent sites will have different numbers of mineral formation events based on
214 site-specific history. The prior model addresses the fundamental question of
215 estimating the number of mineral formation events.

216 There are a number of potential approaches to model the unknown number
217 of formation events. First, one can treat the number of formation events as a
218 fixed parameter, perform a grid search over the different number of formation
219 events, and choose the model that best fits the data [28]. A second approach
220 is to model the number of formation events using a reversible jump algorithm
221 [16]. The third approach is to assign a Dirichlet process prior over the number
222 of formation events. The Dirichlet process model estimates an unknown number
223 of components without *a priori* specifying the number of components.

224 The Dirichlet Process is an infinite dimensional stochastic process that is a
225 distribution over distributions [10]. We use the stick-breaking representation of
226 a Dirichlet process

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

227 where $G(\cdot)$ is the base distribution that has parameters $\boldsymbol{\theta}_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$
 228 and mixing weights p_{bk} with $\sum_{k=1}^{\infty} p_{bk} = 1$. In practice, $p_{bk} \approx 0$ for large k .
 229 Therefore, the infinite sum is well approximated by the finite sum $\sum_{k=1}^K p_{bk}$ for
 230 a large enough K (for most problems $K=10$ or $K=20$ is sufficiently large). The
 231 stick-breaking representation for \mathbf{p}_b is constructed by transforming the variables
 232 $\tilde{\mathbf{p}}_b = (\tilde{p}_{b1}, \dots, \tilde{p}_{bK-1})'$ according to the stick-breaking representation

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

233 Priors on the \tilde{p}_{bk} are assigned exchangeable Beta($1, \alpha_b$) priors giving rise to
 234 the stick-breaking Dirichlet process model. The hyperparameters α_b are given
 235 exchangeable Gamma($1, 1$) priors that control the Dirichlet process concentration
 236 (i.e. smaller α_b give fewer formation events, larger α_b give more formation events).
 237 We chose a Gaussian base measure $G(\cdot)$ with parameters $\boldsymbol{\theta}_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$
 238 representing the mean and variance. Because the parent and child sites contain
 239 mineral grains derived from common formation events, we follow [25] and used
 240 shared kernels by letting $\boldsymbol{\theta}_{bk} \equiv \boldsymbol{\theta}_k = (\mu_k, \sigma_k^2)'$ for all $b = 1, \dots, B$.

241 The prior on the standard deviations for the ages of formation is the truncated
 242 half-Cauchy prior $\sigma_k \sim \text{Cauchy}^+(0, s)I\{0 < \sigma_k < \omega\}$, where we choose s to be
 243 small relative to the range of dates observed and ω provides an upper limit
 244 to the duration of formation events. For the age distributions being modeled
 245 that span the range of 0 to about 300 Millions of years (Ma), we set s to be 25
 246 Ma and set ω to be 50 Ma years. The truncation is important to prevent the
 247 Dirichlet process mixture from generating unrealistically long formation events
 248 which does not match our *a priori* knowledge.

249 The mixing parameter ϕ is assigned a $\text{Dirichlet}(\alpha_\phi \mathbf{1})$ prior where $\mathbf{1}$ is a vector
 250 of ones and the hyperparameter α_ϕ is assigned a $\text{Gamma}(1, 1)$ prior. When
 251 α_ϕ is small the mixing proportions concentrate with a large probability on a
 252 single parent component, when α_ϕ is one ϕ will be uniformly distributed over
 253 all possible mixing proportions, and when α_ϕ is large the mixing proportion will
 254 be concentrated at equal mixing proportions $(\frac{1}{B}, \dots, \frac{1}{B})$.

255 *3.4. Top-down mixing posterior distribution*

256 The top-down mixing model posterior distribution is

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

257 where the first line on the right hand side of the proportional symbol is the
 258 data model, the second line is the process model, and the third line is the prior
 259 model. We estimate the posterior using Markov Chain Monte Carlo (MCMC)
 260 using the *R* package *NIMBLE* [7]. The model is fit using an adaptive block
 261 Metropolis-Hastings algorithm [18]. The constrained parameters $\tilde{\mathbf{p}}$ and $\boldsymbol{\sigma}$ are
 262 transformed to unconstrained support (logit- and log-scale transformations) for
 263 tuning the Metropolis-Hastings block proposals, with corresponding Jacobian
 264 adjustments to the acceptance probabilities. The sampling of the sum-to-one
 265 mixing proportion ϕ is performed by introducing auxiliary variables $\tilde{\phi}$. We
 266 assign a stick-breaking prior on $\tilde{\phi}$ and sample on a logit-scale, correcting for the
 267 transformation using the Jacobian to induce a $\text{Dirichlet}(\alpha_\phi \mathbf{1})$ prior on ϕ .

268 **4. Bottom-up unmixing model**

269 The second research question is: can we reconstruct unobserved parent
270 age distributions from multiple observations of children? In previous work,
271 this analysis has been variably termed “end-member mixing analysis”, “end-
272 member modeling”, or “end-member analysis” as applied to unmixing grain size
273 or detrital age distributions [34, 33]. The end-member unmixing analysis has
274 two components. First, the number of parents B is unknown and needs to be
275 estimated. Second, given the number of parents B , what are the unobserved
276 mineral formation age distributions for the B parents? For this paper, we assume
277 the number of parents B is known. There are a number of criteria for selecting
278 the number of parents including using Bayesian information criteria, reversible
279 jump MCMC [23], assuming a Dirichlet process over the number of parents,
280 or fitting a mixture of finite mixtures [28]. Rather than explore these ideas,
281 we devote our effort on developing the unmixing model for a fixed number of
282 parents [28]. The end-member analysis model uses the same general framework
283 presented in the mixture of Gaussians model (3) with some modifications.

284 *4.1. Bottom-up unmixing data model*

285 Let $d = 1, \dots, D$ index the D child sediments that are each composed of
286 $i = 1, \dots, n_d$ samples. As before, we assume a Gaussian dating error distribution
287 for child d given by

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

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

Unlike in the top-down mixing model above, none of the parent \mathbf{z} s are observed in the bottom-up unmixing model. Hence, the parent distributions are estimated entirely using child sediment observations. The end-member process model for a fixed given number of parents B is

$$\tilde{y}_{id} | \phi_d, \mathbf{p}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2 \quad (4)$$

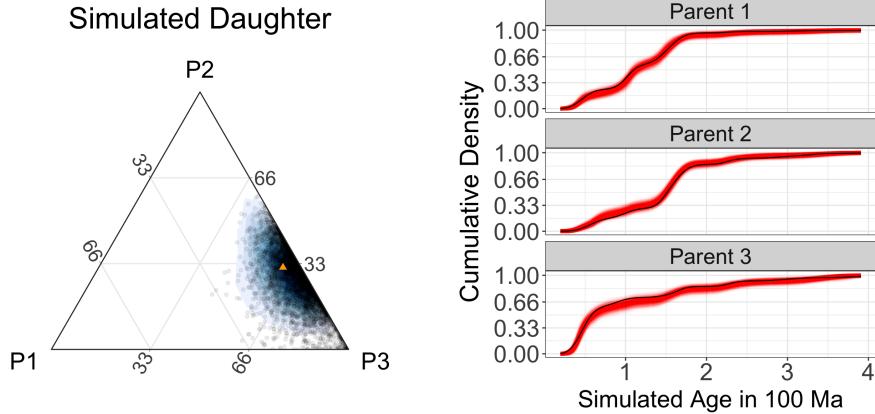
$$= \phi_d \sum_{b=1}^B \sum_{k=1}^K p_{bk} N(\mu_k, \sigma_k^2), \quad (5)$$

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

4.2. Bottom-up unmixing posterior distribution

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

$$\begin{aligned} [\phi_1, \dots, \phi_D, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{p} | \mathbf{Y}] &\propto \prod_{d=1}^D [y_d | \tilde{\mathbf{y}}_d, \boldsymbol{\sigma}_d] \times \\ &\quad \prod_{d=1}^D [\tilde{\mathbf{y}}_d | \phi_d, \tilde{\mathbf{p}}_d, \boldsymbol{\mu}, \boldsymbol{\sigma}] \times \\ &\quad \left(\prod_{d=1}^D [\phi_d | \alpha_{\phi d}] [\alpha_{\phi d}] \right) \left(\prod_{d=1}^D [\tilde{\mathbf{p}}_d | \alpha_b] [\alpha_b] [\boldsymbol{\mu}_b] [\boldsymbol{\sigma}_b] \right), \end{aligned} \quad (6)$$



(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.

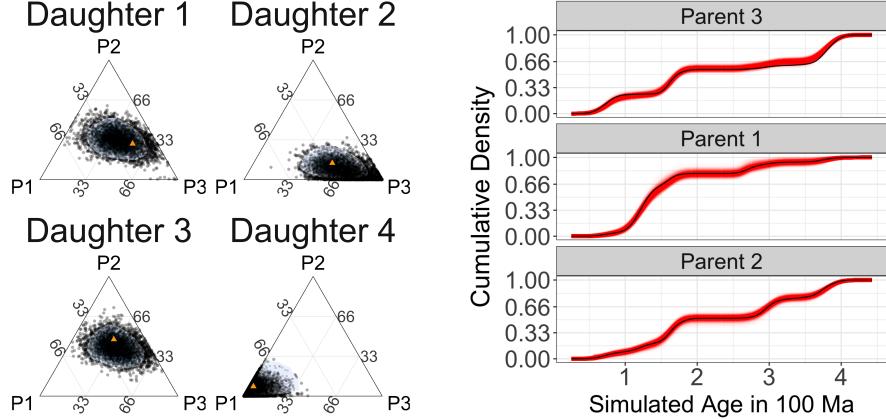
Figure 3: Simulation study results for the top-down mixture modeling approach.

305 where the priors and MCMC algorithm are the same as those in (3) except for a
306 change in dimensionality. Code and data for replication of results presented in this
307 manuscript can be found at <https://github.com/jtipton25/mixing-manuscript>.

308 5. Simulation of synthetic detrital age distributions

309 We explore the performance of the model using a simulation of synthetic
310 detrital age distributions. The aim of the simulation study is to understand
311 how the model performs using realistic data and verify the model is capable
312 of recovering the simulated parameters. The simulation study framework can
313 also be used to understand how uncertainty in estimation varies with respect to
314 sample size, variability in the data, and other questions of interest.

315 First, we create synthetic data using the mixing model in (3) for $B = 3$
316 parents and a single child. The parent distributions were composed of 200,
317 250, and 150 simulated grain dates, respectively, and the child distribution was
318 composed of 150 grain dates. In simulation, we used age dating uncertainties
319 (σ_y, σ_z) that were about 1-3% of the total range of the age distribution. These



(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.

Figure 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.

320 are similar to measurement uncertainties in the data and demonstrate the model
321 is capable of accounting for measurement errors.

322 The posterior mixing proportion estimates ϕ in Figure 3a are plotted as
323 black dots, a smoothed density shaded in blue [19], and the simulated mixing
324 proportion as an orange triangle, demonstrating that the model is accurately
325 estimating the simulated true mixing proportion. Figure 3b shows the estimated
326 CDFs with posterior samples in red and the simulated CDF in black. The
327 results shown in Figure 3 demonstrate that the model is accurately estimating
328 the simulated mixing proportions ϕ as well as the parent age distributions.
329 These results validate the effectiveness of the top-down mixing model to recover
330 simulated parameters of interest given simulated data.

331 We also simulated data from the bottom-up, end-member unmixing model
332 to test how well the proposed framework can reconstruct unobserved parent
333 distributions (assuming the number of parents is known, see Figure 1b). For

334 the simulation, we used $B = 3$ parents and $D = 20$ children where each child
335 consisted of 250 measured sediment grain dates following the model in (6). The
336 dating uncertainties (σ_y) were set at about 2.5% of the total range of the age
337 distribution.

338 Using simulated data, the model is able to reconstruct unobserved parent age
339 distributions. Figure 4a shows the posterior samples for the mixing proportion
340 of each child as black dots with the simulated mixing proportion plotted using
341 an orange triangle. In general, the model can recover the mixing proportions in
342 this simulation example with high precision. Even though the model uses none of
343 the data from the parents, the end-member unmixing model produces reasonable
344 end-member parent age distribution estimates. Figure 4b shows the estimated
345 cumulative distribution function produced by the end-member unmixing model
346 which shows the bottom-up unmixing model is capturing the unobserved parent
347 distributions. However, and not surprisingly, the accuracy for the bottom-up
348 unmixing model is not as good as the fit that uses observations from the parent
349 distribution (i.e. the top-down mixture model) as can be seen in the slightly
350 larger uncertainty estimates of the CDF.

351 A large overlap in the distribution of parent ages is a feature that occurs,
352 particularly in detrital zircon geochronology studies. The preservation of zircons
353 through multiple cycles of erosion and re-sedimentation means that overlapping
354 zircon ages will be present in many rocks, for example the preponderance of
355 Grenville zircons in a host of sedimentary formations of varying age. For parent
356 age distributions that are quite similar to one another, the reconstruction of
357 the unknown parent distributions suffers from weak identifiability. In these
358 situations, the estimated parents jointly contain all of the correct formation
359 events, but the model is unable to attribute the formation events to the correct
360 parents. In other words, while the model identifies the correct age components,

361 the model sometimes struggles to correctly group these components into the
362 correct parent distributions. This is not an unexpected result because Bayesian
363 nonparametric models are well understood to suffer from non-identifiability
364 issues in the context of the Bayesian nonparametric framework [11, 8, 31, 13].
365 Non-identifiability is not a weakness of the particular proposed model framework;
366 the non-identifiability applies to end-member unmixing models in general [45].
367 To overcome the non-identifiability, a potential solution is to impose constraints
368 on the end-members and initialization conditions [9, 27, 6]. Therefore, any
369 end-member unmixing model that uses only child age distributions will have
370 issues in accurately reconstructing the parent distributions if the assumption
371 of the constraints is not met (i.e. the parent age distributions are structurally
372 similar).

373 Bottom-up unmixing models provide a useful way to explore large detrital
374 datasets with unknown sedimentary sources. Providing a way to identify those
375 datasets that either are or are not susceptible to non-identifiability, and thus not
376 amenable to bottom-up unmixing, is critical to success. An advantage of our
377 framework is the end-member unmixing model produces uncertainty estimates
378 that are larger when the model is weakly identifiable. Thus, the uncertainty
379 intervals can be used as a diagnostic to check for identifiability.

380 **6. Model Application to a Natural Case Study**

381 We apply the model presented here to a well-constrained modern dataset from
382 the central California coast [36] shown in Figure 5 where we focus on sediments
383 dating to the most recent 300 Ma. Following the same mixing framework
384 presented in [34], there are five samples (river and beach sediment) used to
385 characterize three distinct sediment inputs (parents) to the region, each with a
386 distinct detrital age distribution (Figure 5). Parents 1 and 2 (P1 and P2) are

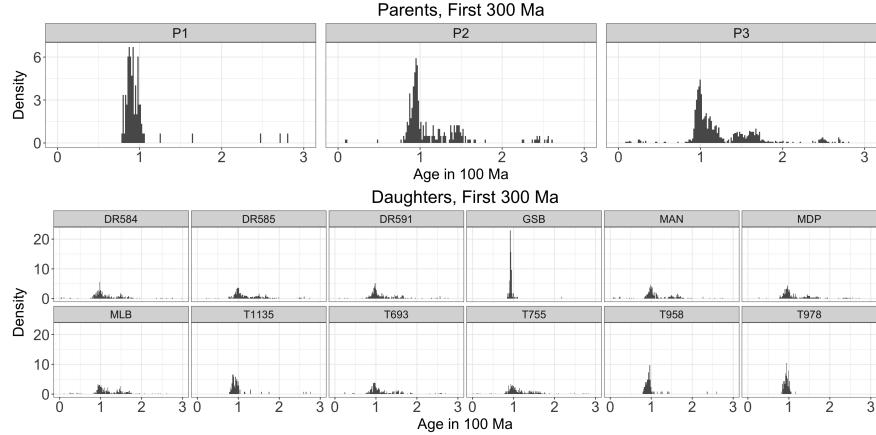


Figure 5: The sediment age data used for the mixing and unmixing models. The three parent age distributions are shown in the top plot and the 12 child age distributions are shown in the bottom plot. The x -axes represent the measured age in 100 Ma and the y -axes shows the empirical density.

387 comprised of river samples (CAR and SAR, respectively) that represent sediment
 388 sources along the Big Sur coastline and Salinas River drainage, respectively.
 389 Parent 3 (P3) is comprised of two river samples (SNR and PAR) and one beach
 390 sample (ANB) that represent northern sediment sources in the Santa Cruz
 391 Mountains and western Diablo Range [36, 34]. Twelve child samples (beach
 392 and submarine canyon sediment) are used to characterize how these parents
 393 are mixed in littoral and marine environments. In total, this dataset (Figure
 394 6) consists of 4,026 individual detrital zircon U-Pb analyses, with individual
 395 samples having 82 to 316 analyses each (median of 290 analyses per sample) [36].

396 We first examine the top-down mixture model (Figure 1a). Figure 7a shows
 397 the reconstruction of the mixing proportions for a sample from a submarine
 398 canyon (T693) modeled as a mixture of the three specified parent distributions
 399 (P1-P3). Visual inspection of the histograms of the data (Figure 5) would suggest
 400 that this child sample is a mixture composed mostly of P3 (a combination of the
 401 samples ANB, SNR, and PAR). The posterior estimates of the mixing proportion
 402 of each parent for child T693 confirms that the primary component of the

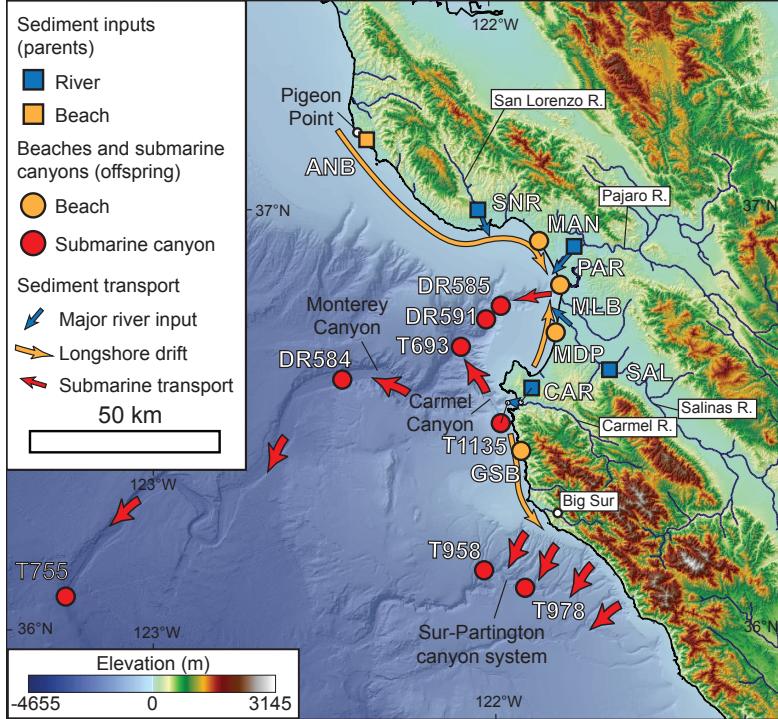
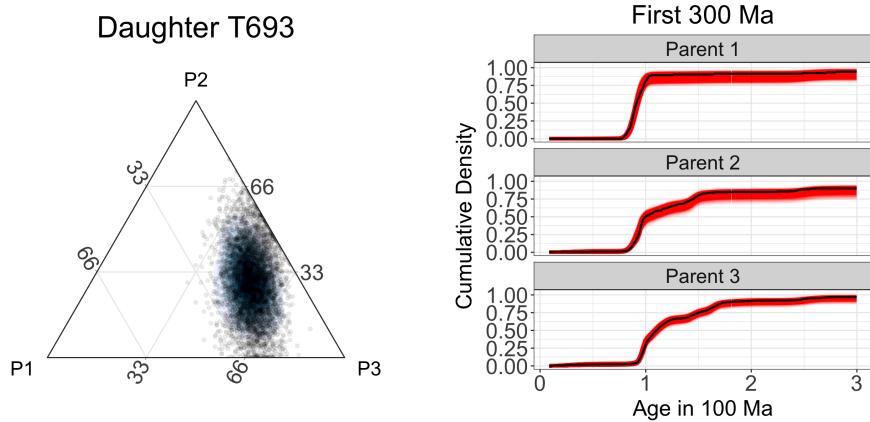


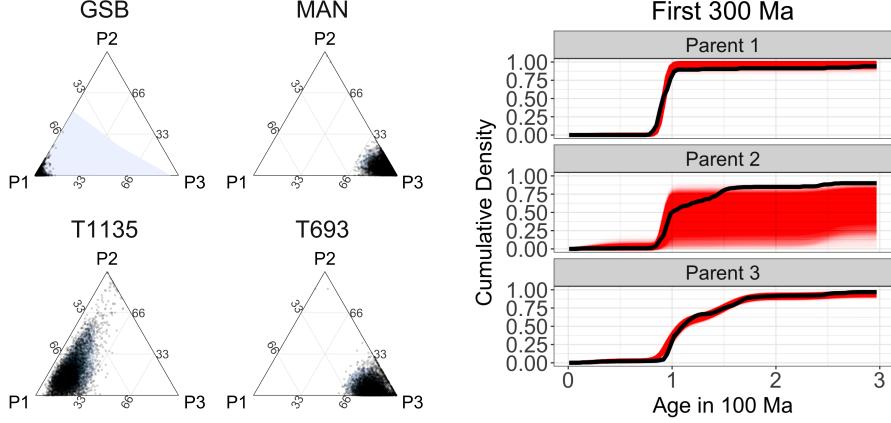
Figure 6: Locations of the parents and children data for the study region in California, USA.



(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.

Figure 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.



(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 Figure 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.

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

403 mixture is from the P3 (Figure 7a). Figure 7b shows the model is capturing the
 404 basic patterns in the parent CDFs as the estimated CDFs are very close to the
 405 empirical CDFs for the parents.

406 Application of the bottom-up, end-member unmixing model to the data set
 407 in [36] results in non-identifiability issues as the simulated age distributions as
 408 evidenced by the larger posterior uncertainty in Figure 8b. Posterior probability
 409 estimates are obtained for the proportion of each child that comes from the
 410 modeled end-member (Figure 8a), but based on the differences in estimates from
 411 the bottom-up mixing model estimates, the distributions should be interpreted
 412 cautiously. Some of these differences are due to the fact that some children are
 413 nearly “pure” representations of a parent (e.g. GSB is nearly a “pure” child of
 414 P1). As more and more children are included, especially those that are close
 415 to “pure” types, they provide a soft constraint on the identifiability problem.
 416 The estimated CDFs shown in red in Figure 8b with the empirical CDFs from

417 the parent data shown in black suggest that the unmixing model is performing
418 well despite the fact that the model is unaware of the parent data although
419 there is much uncertainty about the CDF for P2, likely due to the similarity in
420 distribution to P1 (Figure 5).

421 **7. Conclusion**

422 Starting from a conceptual model of how sediments mix over a landscape,
423 we developed a generative Bayesian nonparametric statistical model for detrital
424 mineral age data. This model allows us to characterize the uncertainty we
425 have in age distributions of parents and children and the mixing coefficients,
426 accounting for the uncertainties in measured dates [42]. The statistical models
427 proposed can be understood as abstractions to the mixing of sediments on the
428 landscape while also estimating probabilistic uncertainty in the distribution
429 estimates. Because the model can simulate sediment age distributions, we
430 can directly explore the assumptions of the model by simulating synthetic data.
431 Running a simulation experiment demonstrated the model is capable of recovering
432 simulated distributions supporting the usefulness of the models when applied to
433 the observed data.

434 We proposed two frameworks to model the sediment mixing mechanisms: the
435 top-down mixing model where mineral dates are measured for both parent and
436 child sediments and a bottom-up unmixing framework where mineral dates are
437 only measured for the children. The top-down model estimated the parent and
438 child distributions and the mixing proportions with high precision and accuracy.
439 The bottom-up model occasionally demonstrated evidence of non-identifiability,
440 suggesting the inference for the bottom-up model is less precise than for the
441 top-down mixing model; however, the variances of these estimates are larger
442 in our bottom-up unmixing model, recognizing the challenges in reconstructing

443 unobserved parent age distributions.

444 Applying the proposed methods to central California coast data in [36],
445 we produced estimates of the mixing proportions of child sediments under
446 both the mixing and unmixing models. Other studies have produced similar
447 estimates [1, 2]; our contribution is the mechanistic model framework that
448 produces estimates of mixing with associated uncertainty. We account for dating
449 uncertainty directly in the model and by modifying the model statement to
450 remove dating uncertainty (i.e., removing the data models in Eq. (1)) we can
451 examine the effects on inference by not accounting for the dating uncertainty.

452 Direct, probabilistic estimates of uncertainty and the ability to calculate de-
453 rived quantities with uncertainty is a benefit of the Bayesian methodology. Thus,
454 we can answer questions like what is the probability that at least 50% of child
455 sediment T693 comes from parent P3? The answer is calculated directly from the
456 posterior samples using the Monte Carlo approximation $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3}^{(\ell)} \geq 0.5\} =$
457 0.672, where $\ell = 1, \dots, L$ are the indices of the MCMC samples and $\phi_{P3}^{(\ell)}$ is the
458 estimated mixing proportion for the ℓ th MCMC iteration. The probability that
459 at least 50% of the child sediment comes from parent P3 and at least 25% comes
460 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
461 model produces a posterior probability, any such probabilistic questions can be
462 calculated as derived quantities. For example, the Bayesian model framework we
463 introduce can be used to ask questions like: what proportion of a given sample
464 contains grains older than a given age? or what is the probability that an unob-
465 served parent contains grains with a particular age range. Once the posterior
466 samples have been calculated, any such questions about derived quantities are
467 answerable using posterior samples.

468 In addition, the ability to include prior information in the Bayesian frame-
469 work is a useful tool that can be used to improve estimation and test geologic

470 hypotheses. For example, certain geologic events, such as the Grenville orogeny,
471 produced large amounts of zircon that have since been broadly dispersed and
472 recycled in sedimentary rocks. Priors that account for the likelihood of observ-
473 ing zircons of Grenville-age (or other known zircon-producing events) can be
474 introduced into this model framework to improve model performance. In addi-
475 tion, our framework can accommodate a variety of detrital data with different
476 magnitudes of uncertainty. As analytic techniques for dating minerals improve,
477 it is important to account for dating uncertainties that might have orders of
478 magnitude difference making our method more robust to future improvements
479 in analytic laboratory techniques.

480 **8. Acknowledgements**

481 Support for SAJ came from the FEDMAP component of the US Geological
482 Survey National Cooperative Geologic Mapping Program. This draft manuscript
483 is distributed solely for purposes of scientific peer review. Its content is delib-
484 erative and predecisional, so it must not be disclosed or released by reviewers.
485 Because the manuscript has not yet been approved for publication by the U.S.
486 Geological Survey (USGS), it does not represent any official USGS finding or
487 policy. Any use of trade, firm, or product names is for descriptive purposes only
488 and does not imply endorsement by the U.S. Government.

489 **References**

- 490 [1] Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005a. Construction of
491 detrital mineral populations: insights from mixing of U-Pb zircon ages in
492 Himalayan rivers. Basin Research 17, 463–485.
- 493 [2] Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005b. U–pb zircon ages

- 494 as a sediment mixing tracer in the nepal himalaya. Earth and Planetary
495 Science Letters 235, 244–260.
- 496 [3] Berliner, L.M., 2003. Physical-statistical modeling in geophysics. Journal
497 of Geophysical Research: Atmospheres 108.
- 498 [4] Chang, W., Haran, M., Applegate, P., Pollard, D., 2016. Calibrating an
499 ice sheet model using high-dimensional binary spatial data. Journal of the
500 American Statistical Association 111, 57–72.
- 501 [5] Chen, J.H., Moore, J.G., 1982. Uranium-lead isotopic ages from the sierra
502 nevada batholith, california. Journal of Geophysical Research: Solid Earth
503 87, 4761–4784.
- 504 [6] Chen, W., Guillaume, M., 2012. HALS-based NMF with flexible constraints
505 for hyperspectral unmixing. EURASIP Journal on Advances in Signal
506 Processing 2012, 54.
- 507 [7] de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple
508 Lang, D., Bodik, R., 2017. Programming with models: writing statistical
509 algorithms for general model structures with nimble. Journal of Computational
510 and Graphical Statistics 26, 403–413. doi:10.1080/10618600.2016.
511 1172487.
- 512 [8] Diebolt, J., Robert, C.P., 1994. Estimation of finite mixture distributions
513 through Bayesian sampling. Journal of the Royal Statistical Society. Series
514 B (Methodological) , 363–375.
- 515 [9] Donoho, D., Stodden, V., 2004. When does non-negative matrix factorization
516 give a correct decomposition into parts?, in: Advances in neural information
517 processing systems, pp. 1141–1148.

- 518 [10] Ferguson, T.S., 1973. A Bayesian analysis of some nonparametric problems.
519 The annals of statistics , 209–230.
- 520 [11] Ferguson, T.S., 1983. Bayesian density estimation by mixtures of normal
521 distributions, in: Recent advances in statistics. Elsevier, pp. 287–302.
- 522 [12] Fosdick, J.C., Grove, M., Graham, S.A., Hourigan, J.K., Lovera, O., Ro-
523 mans, B.W., 2015. Detrital thermochronologic record of burial heating
524 and sediment recycling in the magallanes foreland basin, patagonian andes.
525 Basin Research 27, 546–572.
- 526 [13] Frühwirth-Schnatter, S., 2006. Finite mixture and Markov switching models.
527 Springer Science & Business Media.
- 528 [14] Gehrels, G., 2012. Detrital zircon U-Pb geochronology: Current methods
529 and new opportunities. Tectonics of sedimentary basins: Recent advances ,
530 45–62.
- 531 [15] Gehrels, G., 2014. Detrital zircon U-Pb geochronology applied to tectonics.
532 Annual Review of Earth and Planetary Sciences 42, 127–149.
- 533 [16] Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation
534 and Bayesian model determination. Biometrika 82, 711–732.
- 535 [17] Guan, Y., Haran, M., Pollard, D., 2018. Inferring ice thickness from a
536 glacier dynamics model and multiple surface data sets. Environmetrics 29,
537 e2460.
- 538 [18] Haario, H., Saksman, E., Tamminen, J., et al., 2001. An adaptive Metropolis
539 algorithm. Bernoulli 7, 223–242.
- 540 [19] Hamilton, N., 2018. ggtern: An Extension to ggplot2, for the creation of
541 ternary diagrams. URL: <https://CRAN.R-project.org/package=ggtern>. r
542 package version 2.2.2.

- 543 [20] Hefley, T.J., Brost, B.M., Hooten, M.B., 2017. Bias correction of bounded
544 location errors in presence-only data. *Methods in Ecology and Evolution* 8,
545 1566–1573.
- 546 [21] Hooten, M.B., Hobbs, N., 2015. A guide to Bayesian model selection for
547 ecologists. *Ecological Monographs* 85, 3–28.
- 548 [22] Irwin, W.P., Wooden, J.L., 1999. Plutons and accretionary episodes of
549 the Klamath Mountains, California and Oregon. Technical Report. US
550 Geological Survey.
- 551 [23] Jasra, A., Stephens, D.A., Gallagher, K., Holmes, C.C., 2006. Bayesian
552 mixture modelling in geochronology via markov chain monte carlo. *Mathematical Geology* 38, 269–300.
- 554 [24] Kimbrough, D.L., Grove, M., Gehrels, G.E., Dorsey, R.J., Howard, K.A.,
555 Lovera, O., Aslan, A., House, P.K., Pearthree, P.A., 2015. Detrital zircon
556 u-pb provenance of the colorado river: A 5 my record of incision into cover
557 strata overlying the colorado plateau and adjacent regions. *Geosphere* 11,
558 1719–1748.
- 559 [25] Lock, E.F., Dunson, D.B., 2015. Shared kernel Bayesian screening.
560 *Biometrika* 102, 829–842.
- 561 [26] Mason, C.C., Fildani, A., Gerber, T., Blum, M.D., Clark, J.D., Dykstra,
562 M., 2017. Climatic and anthropogenic influences on sediment mixing in the
563 mississippi source-to-sink system using detrital zircons: Late pleistocene to
564 recent. *Earth and Planetary Science Letters* 466, 70–79.
- 565 [27] Miao, L., Qi, H., 2007. Endmember extraction from highly mixed data
566 using minimum volume constrained nonnegative matrix factorization. *IEEE
567 Transactions on Geoscience and Remote Sensing* 45, 765–777.

- 568 [28] Miller, J.W., Harrison, M.T., 2018. Mixture models with a prior on the
569 number of components. *Journal of the American Statistical Association*
570 113, 340–356.
- 571 [29] Puetz, S.J., Ganade, C.E., Zimmermann, U., Borchardt, G., 2018. Statistical
572 analyses of global U-Pb database 2017. *Geoscience Frontiers* 9, 121–145.
- 573 [30] Reiners, P.W., Brandon, M.T., 2006. Using thermochronology to understand
574 orogenic erosion. *Annual Review Earth Planetary Sciences* 34, 419–466.
- 575 [31] Richardson, S., Green, P.J., 1997. On Bayesian analysis of mixtures with
576 an unknown number of components (with discussion). *Journal of the Royal
577 Statistical Society: series B (statistical methodology)* 59, 731–792.
- 578 [32] Romans, B.W., Castelltort, S., Covault, J.A., Fildani, A., Walsh, J., 2016.
579 Environmental signal propagation in sedimentary systems across timescales.
580 *Earth-Science Reviews* 153, 7–29.
- 581 [33] Saylor, J.E., Sundell, K., Sharman, G., 2019. Characterizing sediment
582 sources by non-negative matrix factorization of detrital geochronological
583 data. *Earth and Planetary Science Letters* 512, 46–58.
- 584 [34] Sharman, G.R., Johnstone, S.A., 2017. Sediment unmixing using detrital
585 geochronology. *Earth and Planetary Science Letters* 477, 183–194.
- 586 [35] Sharman, G.R., Sylvester, Z., Covault, J.A., 2019. Conversion of tectonic and
587 climatic forcings into records of sediment supply and provenance. *Scientific
588 reports* 9, 4115.
- 589 [36] Sickmann, Z.T., Paull, C.K., Graham, S.A., 2016. Detrital-zircon mixing
590 and partitioning in fluvial to deep marine systems, central California, USA.
591 *Journal of Sedimentary Research* 86, 1298–1307.

- 592 [37] Stock, G.M., Ehlers, T.A., Farley, K.A., 2006. Where does sediment
593 come from? quantifying catchment erosion with detrital apatite (u-th)/he
594 thermochronometry. *Geology* 34, 725–728.
- 595 [38] Sundell, K., Saylor, J.E., 2017. Unmixing detrital geochronology age
596 distributions. *Geochemistry, Geophysics, Geosystems* .
- 597 [39] Tipton, J., Hooten, M., Goring, S., 2017. Reconstruction of spatio-temporal
598 temperature from sparse historical records using robust probabilistic princi-
599 pal component regression. *Advances in Statistical Climatology, Meteorology*
600 and *Oceanography* 3, 1–16.
- 601 [40] Tipton, J.R., Hooten, M.B., Nolan, C., Booth, R.K., McLachlan, J., 2019.
602 Predicting paleoclimate from compositional data using multivariate Gaus-
603 sian process inverse prediction. *Annals of Applied Statistics* 13, 2363–2388.
604 doi:10.1214/19-AOAS1281.
- 605 [41] Tipton, J.R., Hooten, M.B., Pederson, N., Tingley, M.P., Bishop, D.,
606 2016. Reconstruction of late Holocene climate based on tree growth and
607 mechanistic hierarchical models. *Environmetrics* 27, 42–54.
- 608 [42] Tye, A., Wolf, A., Niemi, N., 2019. Bayesian population correlation: A
609 probabilistic approach to inferring and comparing population distributions
610 for detrital zircon ages. *Chemical Geology* 518, 67–78.
- 611 [43] Vermeesch, P., 2012. On the visualisation of detrital age distributions.
612 *Chemical Geology* 312, 190–194.
- 613 [44] Weltje, G.J., 1997. End-member modeling of compositional data: Numerical-
614 statistical algorithms for solving the explicit mixing problem. *Mathematical*
615 *Geology* 29, 503–549.

- 616 [45] Weltje, G.J., Prins, M.A., 2007. Genetically meaningful decomposition of
617 grain-size distributions. *Sedimentary Geology* 202, 409–424.
- 618 [46] Wotzlaw, J.F., Schaltegger, U., Frick, D.A., Dungan, M.A., Gerdes, A.,
619 Günther, D., 2013. Tracking the evolution of large-volume silicic magma
620 reservoirs from assembly to supereruption. *Geology* 41, 867–870.