

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.

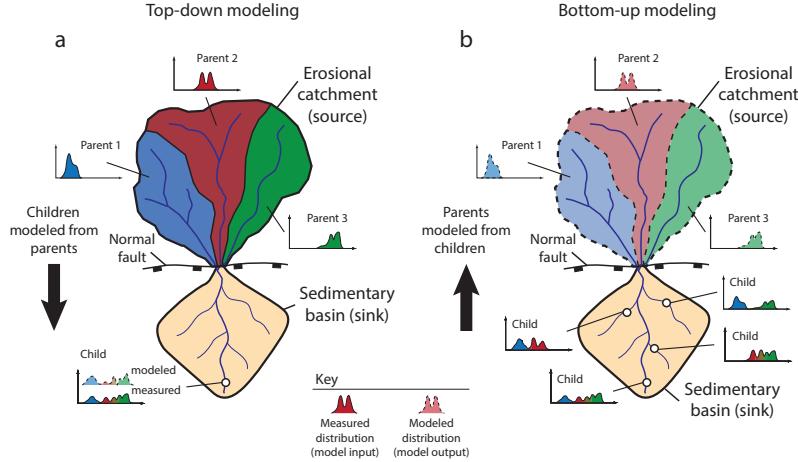


Figure 1: Schematic depiction of a sediment routing system with an erosional source region characterized by three parents (black, red, and green) and an associated sedimentary basin (yellow). Call out plots represent the age of detrital minerals from the parents and how they are mixed to form child distributions. (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 these processes that generated the modern landscape. The ability to decipher the relative proportions of sources that eroded to produce sediment informs understanding of the underlying geologic processes controlling 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 detrital geochronology – dating the time 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].

We will follow the convention that sediment sources are called *parents* and 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 are 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].

Bayesian mixture modeling of geochronology data, including detrital data, has numerous advantages when addressing single samples [23] including allowing inference and uncertainty estimates for 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 crystallization events. The Bayesian nonparametric statistical model presented herein has a number of advantages over previously

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

60 **2. Model Overview**

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

68 Following [3], the statistical model described below is divided into three
69 components: the data model, the process model, and the prior model. In general,
70 the data model defines probability distributions that describe the variability in the
71 data due to the observation process. The data model can be modified to account
72 for non Gaussian measurement processes like counts, outliers, spatial/temporal
73 errors, etc [39, 20]. Process models describe the best scientific understanding of

74 the process of interest. For example, process models have been used to describe
75 the monthly response of trees to climate [41], the relationship between climate
76 and pollen in sediments [40], and the movement of ice sheets in Antarctica
77 [4, 17]. The prior model describes the range of parameter values that are
78 plausible. Sometimes the prior model is used as regularization to improve the
79 generalization of the model to unobserved data [21].

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

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

85 *3.1. Top-down mixing data model*

86 Let the n_y observed age measurements of a single child of interest be \mathbf{y} and
87 let the observed date measurements for each of the $b = 1, \dots, B$ parents be given
88 by the n_b -dimensional vector \mathbf{z}_b . Because the observed ages are measured with
89 uncertainty reported as a dating error standard deviation, we explicitly account
90 for this source of uncertainty in the data model. In the case of U-Pb dating of
91 detrital zircon grains, dates are most commonly determined using laser ablation-
92 inductively coupled plasma-mass spectrometry [14]. Such date measurements
93 typically have relative 2σ analytical precision of 1-4%, with relative uncertainty
94 increasing for younger analyses [29]. For each detrital mineral, the estimate of
95 dating measurement uncertainty is reported as a n_y -vector of standard deviations
96 $\boldsymbol{\sigma}_y$ for the child and B n_b -vectors of standard deviations $\boldsymbol{\sigma}_{z_b}$. We assume the
97 date measurement uncertainty follows a Gaussian distribution where the observed
98 sediment particle date is

$$\mathbf{y}|\tilde{\mathbf{y}}, \boldsymbol{\sigma}_y^2 \sim N(\tilde{\mathbf{y}}, \text{diag}(\boldsymbol{\sigma}_y^2)), \quad (1)$$

$$\mathbf{z}_b|\tilde{\mathbf{z}}_b, \boldsymbol{\sigma}_{z_b}^2 \sim N(\tilde{\mathbf{z}}_b, \text{diag}(\boldsymbol{\sigma}_{z_b}^2)).$$

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

107 *3.2. Top-down mixing process model*

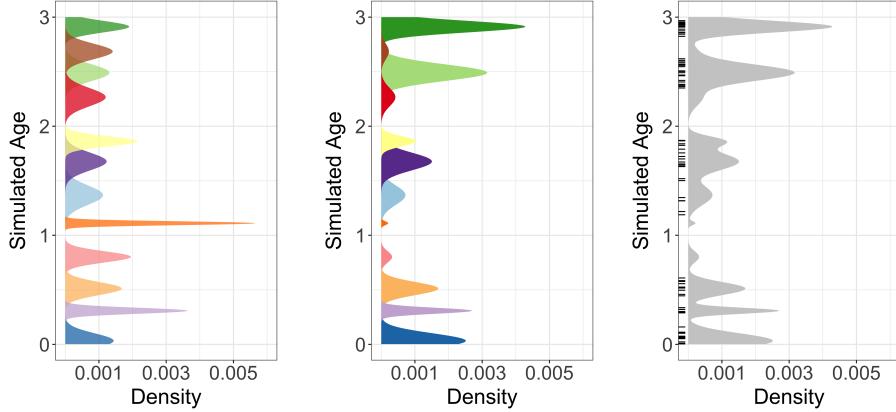
108 The process model addresses two scientific questions: what are the estimates
109 of the true, unobserved detrital mineral age distributions at the parent and child
110 locations? and what proportions of those detrital minerals did each parent source
111 contribute to the child? There are many different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

118 Over geologic time, individual minerals may be repeatedly recycled into sedi-
119 mentary rocks by of erosion, transport, deposition, and exhumation. However, in
120 many cases the dates recorded by individual minerals contained in these deposits

121 are distinctive and unaffected by these recycling processes (e.g., excluding burial
122 reheating of thermochronometers [12]). We assume that minerals created by the
123 same geologic event share an age distribution that is relatively homogeneous
124 with only small variability. Furthermore, episodes of rock and mineral formation
125 (typically lasting 10^5 to 10^7 years [5, 22, 46]) are nearly discrete events relative
126 to geologic time (4.5×10^9 years). While minerals often show overgrowths of
127 different ages, this provides a useful approximation. Under the conceptual model
128 (Figure 1), sediment is formed by the decomposition of rocks containing minerals
129 created at different times, and sediment at every child location is composed of
130 an unknown number of mineral formation events that are also present at source
131 locations.

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

142 In most cases the true number of mineral formation events K recorded by a
143 detrital sample is unknown. Under our model, the set of all K mineral formation
144 events is a mixture of K normal distributions, as shown in Figure 2a where
145 each mineral formation event is shown in a different color. The centers of each
146 age distribution in Figure 2a are given by the values of μ_{bk} and the spreads of
147 each age distribution are given by the variances σ_{bk}^2 . Aerial extent, differential



(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 relative abundance 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.

erosion, 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. In addition, we only observe a finite sample from the distribution in Figure 2c, shown as a rug plot with each tick representing the observed detrital mineral grain date. Thus, the number of mineral formation events is potentially challenging to extract from the data.

161 Consider a detrital mineral date \tilde{z}_{ib} from a parent sediment source b from
 162 $i = 1, \dots, n_b$ measurements. The single sediment grain comes from a mineral
 163 formation event implying the mixture distribution

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

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

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

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

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

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

175 Section 3.3.

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

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

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

191 Then, the probability that the sediment grain came from parent b is $\phi_b \equiv$
192 $P(\delta_i = b)$. Defining the indicator variable $I\{\delta_i = b\}$ where $P(\delta_i = b) =$
193 $E(I\{\delta_i = b\})$, we can write the age distribution over all sediment grains as

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

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

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

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

199 Combining the above results, the full process model is

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

200 *3.3. Top-down mixing prior model*

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

206 number of mineral formation events.

207 There are a number of potential approaches to model the unknown number
208 of formation events. First, one can treat the number of formation events as a
209 fixed parameter, perform a grid search over the different number of formation
210 events, and choose the model that best fits the data [28]. A second approach
211 is to model the number of formation events using a reversible jump algorithm
212 [16]. The third approach is to assign a Dirichlet process prior over the number
213 of formation events. The Dirichlet process estimates an unknown number of
214 components without *a priori* specifying the number.

215 The Dirichlet process is an infinite dimensional stochastic process which is a
216 distribution over distributions [10]. We use the stick-breaking representation of
217 a Dirichlet process

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

218 where $G(\cdot)$ is the base distribution with parameters $\boldsymbol{\theta}_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$ and
219 mixing weights p_{bk} with $\sum_{k=1}^{\infty} p_{bk} = 1$. In practice, $p_{bk} \approx 0$ for large k , therefore,
220 the infinite sum is well approximated by the finite sum $\sum_{k=1}^K p_{bk}$ for a large
221 enough K (for most problems $K=10$ or $K=20$ is sufficiently large). The stick-
222 breaking representation for \mathbf{p}_b is constructed by transforming auxiliary variables
223 $\tilde{\mathbf{p}}_b = (\tilde{p}_{b1}, \dots, \tilde{p}_{bK-1})'$ using the stick-breaking representation

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

224 Priors on the \tilde{p}_{bk} are assigned exchangeable Beta($1, \alpha_b$) priors giving rise
 225 to the stick-breaking Dirichlet process. The hyperparameters α_b are given
 226 exchangeable Gamma($1, 1$) priors that control the Dirichlet process concentration
 227 (i.e. smaller α_b give fewer formation events, larger α_b give more formation events).
 228 Because our study site is constrained geographically, the parent and child sites
 229 contain mineral grains derived from common formation events, we follow [25]
 230 and used shared kernels by letting $\boldsymbol{\theta}_{bk} = \boldsymbol{\theta}_k = (\mu_k, \sigma_k^2)'$ for all $b = 1, \dots, B$.

231 The standard deviations for the ages of formation are assigned truncated
 232 half-Cauchy priors $\sigma_k \sim \text{Cauchy}^+(0, s)I\{0 < \sigma_k < \omega\}$, where we choose s to be
 233 small relative to the range of dates observed and ω provides an upper limit to
 234 the duration of formation events. For the case study where the ages span the
 235 range of 0 to about 300 Millions of years (Ma), we set s to be 25 Ma and set ω
 236 to be 50 Ma years. The truncation is important to prevent the Dirichlet process
 237 mixture from generating unrealistically long formation events which does not
 238 match our *a priori* geologic knowledge.

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

245 3.4. Top-down mixing posterior distribution

246 The top-down mixing model posterior distribution is

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

where the first three line on the right hand side of the proportional symbol are
 the data, process, and the prior model, respectively. We estimate the posterior
 using Markov Chain Monte Carlo (MCMC) with the *R* package *NIMBLE* [7]
 using an adaptive block Metropolis-Hastings algorithm [18]. The constrained
 auxiliary variable $\tilde{\mathbf{p}}$ and standard deviation σ are transformed to unconstrained
 support (logit- and log-scale transformations) for tuning the Metropolis-Hastings
 block proposals, with corresponding Jacobian adjustments to the acceptance
 probabilities. The sampling of the sum-to-one mixing proportion ϕ is performed
 by introducing auxiliary variables $\tilde{\phi}$, assigning a stick-breaking prior on $\tilde{\phi}$,
 then sampling on a logit-scale after correcting for the transformation using the
 Jacobian to induce a $\text{Dirichlet}(\alpha_\phi \mathbf{1})$ prior on ϕ .

4. Bottom-up unmixing model

The second research question is: can we reconstruct unobserved parent age
 distributions from multiple child observations? In previous work, this analysis has
 been variably termed “end-member mixing analysis”, “end-member modeling”,
 or “end-member analysis” as applied to unmixing grain size or detrital age
 distributions [34, 33]. The end-member unmixing analysis has two components.
 First, the number of parents B is unknown and needs to be estimated. Second,
 given the number of parents B , what are the unobserved mineral formation

266 age distributions for the B parents? For this paper, we assume the number of
 267 parents B is known. There are a number of criteria for selecting the number of
 268 parents including using Bayesian information criteria, reversible jump MCMC
 269 [23], assuming a Dirichlet process over the number of parents, or fitting a mixture
 270 of finite mixtures [28]. Rather than explore these ideas, we devote our effort
 271 on developing the unmixing model for a fixed number of parents [28]. The
 272 end-member model uses the same general framework presented in the mixture
 273 of Gaussians model (3) with some modifications.

274 *4.1. Bottom-up unmixing data model*

275 Let $d = 1, \dots, D$ index the D child sediments that are each composed of
 276 $i = 1, \dots, n_d$ samples. As before, we assume a Gaussian dating error distribution
 277 for child d given by

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

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

280 Unlike in the top-down mixing model above, none of the parent **zs** are
 281 observed. Hence, the parent distributions are estimated entirely using child
 282 sediment observations. The end-member process model for a fixed given number
 283 of parents B is

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

284 where, like before, we assume a Gaussian mixing distribution using shared
 285 kernels across the B parents. Like the top-down mixing model, these equations

286 can be derived by introducing categorical random variables then marginalizing
 287 out the latent indicator variables from the model. The b th unknown parent
 288 age distribution is $\sum_{k=1}^K p_{bk} N(\mu_k, \sigma_k^2)$ and the posterior estimate ϕ_{bd} is the
 289 proportion of child d that comes from parent b . The prior model for the bottom-
 290 up unmixing model is the same as for the top-down mixing model, except for
 291 the dimension of different variables.

292 4.2. Bottom-up unmixing posterior distribution

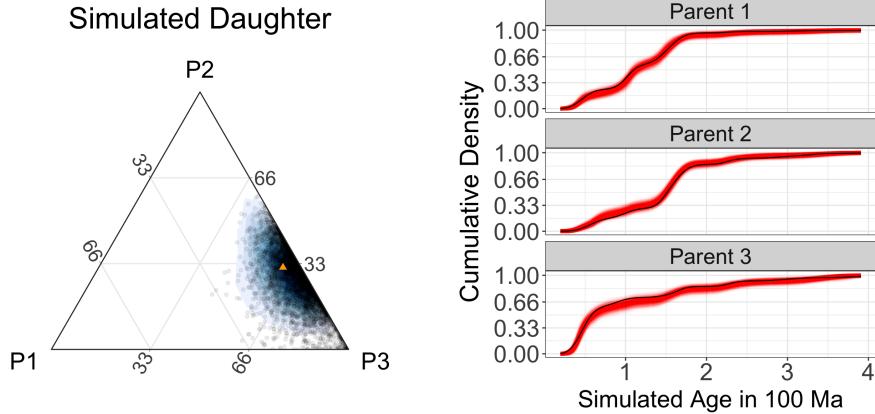
293 The posterior distribution that we estimate with the end member unmixing
 294 model is

$$\begin{aligned}
 [\phi_1, \dots, \phi_D, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{p} | \mathbf{Y}] \propto & \prod_{d=1}^D [\mathbf{y}_d | \tilde{\mathbf{y}}_d, \boldsymbol{\sigma}_d] \times & (5) \\
 & \prod_{d=1}^D [\tilde{\mathbf{y}}_d | \phi_d, \mathbf{p}_d, \boldsymbol{\mu}, \boldsymbol{\sigma}] \times \\
 & \left(\prod_{d=1}^D [\phi_d | \alpha_{\phi d}] [\alpha_{\phi d}] \right) \left(\prod_{d=1}^D [\mathbf{p}_d | \alpha_b] [\alpha_b] [\boldsymbol{\mu}_b] [\boldsymbol{\sigma}_b] \right),
 \end{aligned}$$

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

298 5. Simulation of synthetic detrital age distributions

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

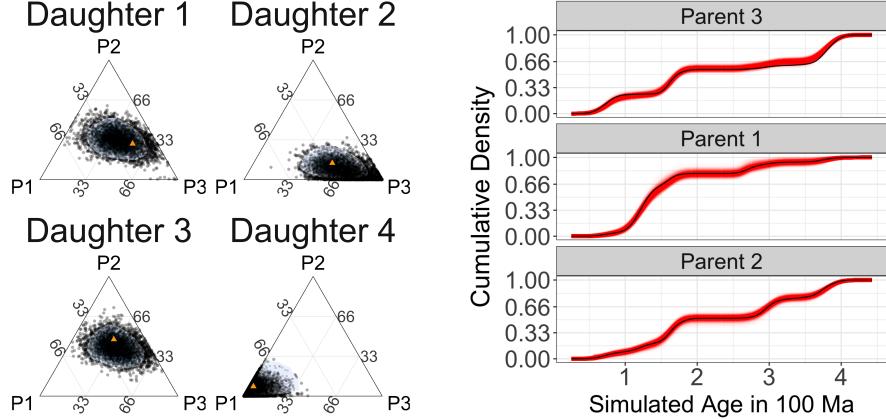


(a) Ternary plot showing posterior mixing proportion estimates as black circles and the simulated true mixing proportions as an orange triangle.
(b) Plot of simulated parents with fitted posterior CDF estimates in red and the simulated true CDF in black. Each red line represents a posterior sample of the cumulative parent age density.

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

305 First, we create synthetic data using the top-down mixing model in (3) for
306 $B = 3$ parents and a single child (1a). The parent distributions were composed of
307 200, 250, and 150 simulated grain dates, respectively, and the child distribution
308 was composed of 150 grain dates. In simulation, we used age dating uncertainties
309 (σ_y, σ_z) that were about 1-3% of the total range of the age distribution. These
310 are similar to measurement uncertainties in the case study and demonstrate the
311 model is capable of accounting for measurement errors.

312 The posterior samples for the mixing proportion ϕ are shown in Figure 3a
313 as black dots with a smoothed posterior density shown shaded in blue [19],
314 and the simulated true mixing proportion is represented by the orange triangle,
315 demonstrating that the model is accurately estimating the mixing proportion as
316 the orange triangle is in the bulk of the posterior samples. Figure 3b shows the
317 estimated CDFs with posterior samples in red and the simulated CDF in black.
318 The results in Figure 3 demonstrate that the model is accurately estimating the
319 simulated mixing proportions ϕ as well as the parent age distributions, validating
320 the effectiveness of the top-down mixing model to recover simulated parameters



(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.

321 of interest.

322 The second simulation generated data from the bottom-up, end-member
 323 unmixing model (Figure 1b) to test how well the proposed framework can
 324 reconstruct unobserved parent distributions. For the simulation, we used $B = 3$
 325 parents and $D = 20$ children where each child consisted of 250 measured sediment
 326 grain dates following the model in (5). The dating uncertainties (σ_y) were set
 327 at about 1-3% of the total range of the age distribution.

328 Figure 4a shows the posterior samples for the mixing proportion of each child
 329 as black dots with the simulated mixing proportion plotted using an orange
 330 triangle. In general, the model can recover the mixing proportions in this
 331 simulation example with high precision. Even though the model uses none of the
 332 data from the parents, the end-member unmixing model produces reasonable
 333 end-member parent age distribution estimates. Figure 4b shows the estimated
 334 cumulative distribution function produced by the end-member unmixing model

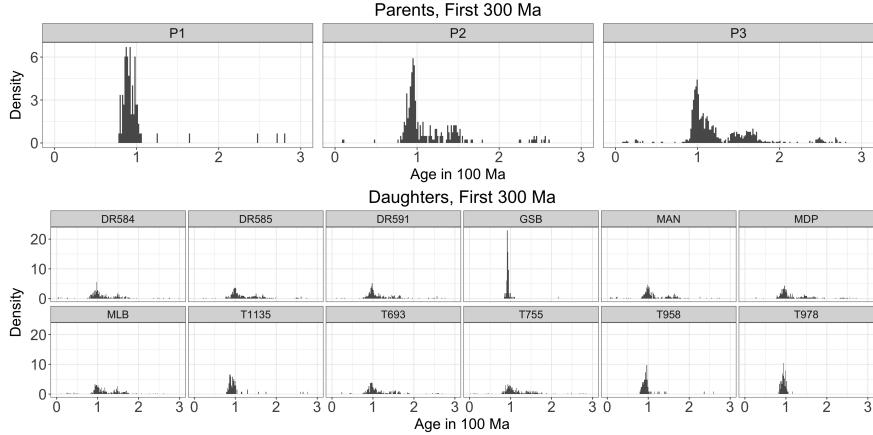


Figure 5: The sediment age data used for the mixing and unmixing models. The three parent age distributions are shown in the top plot and the 12 child age distributions are shown in the bottom plot. The x -axes represent the measured age in 100 Ma and the y -axes shows the empirical density.

which shows the bottom-up unmixing model is capturing the unobserved parent distributions. However, and not surprisingly, the accuracy for the bottom-up unmixing model is not as good as the fit that uses observations from the parent distribution (i.e. the top-down mixture model) as can be seen in the slightly larger uncertainty estimates of the CDF.

6. Application to a Natural Case Study

We apply the model presented here to a well-constrained modern dataset from the central California coast [36] shown in Figure 5 where we focus on sediments dating to the most recent 300 Ma. Following the same mixing framework presented in [34], there are five samples (river and beach sediment) used to characterize three distinct sediment inputs (parents) to the region, each with a distinct detrital age distribution (Figure 5). Parents 1 and 2 (P1 and P2) are comprised of river samples (CAR and SAR, respectively) that represent sediment sources along the Big Sur coastline and Salinas River drainage, respectively. Parent 3 (P3) is comprised of two river samples (SNR and PAR) and one beach

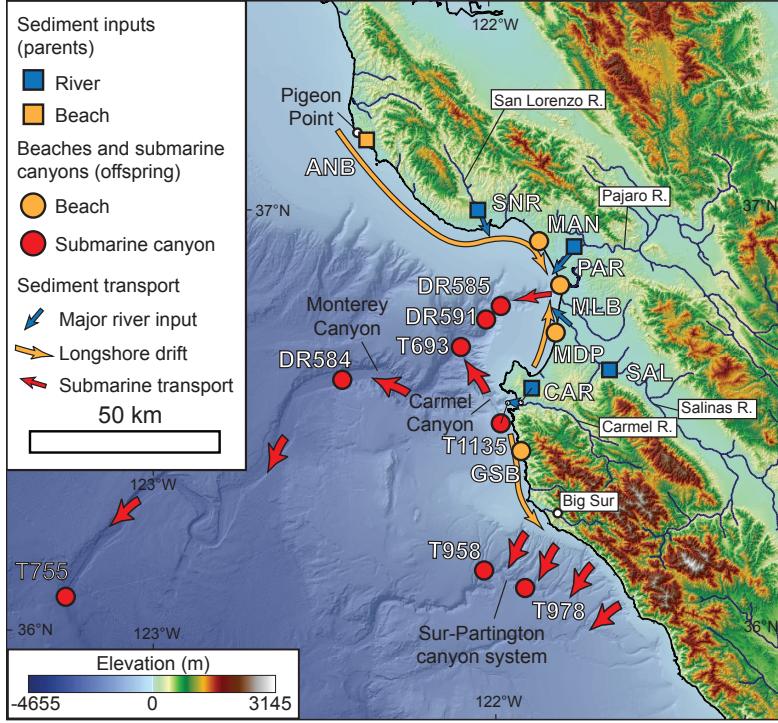
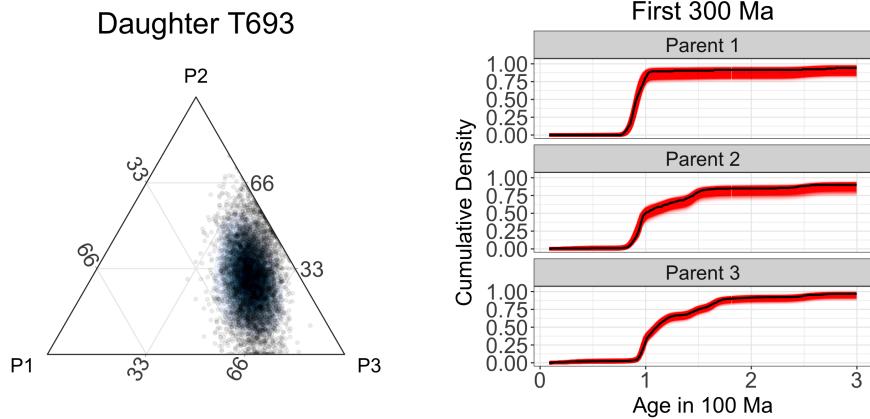


Figure 6: Locations of the parents and children data for the study region in California, USA.

sample (ANB) that represent northern sediment sources in the Santa Cruz Mountains and western Diablo Range [36, 34]. Twelve child samples (beach and submarine canyon sediment) are used to characterize how these parents are mixed in littoral and marine environments. In total, this dataset (Figure 6) consists of 4,026 individual detrital zircon U-Pb analyses, with individual samples having 82 to 316 analyses each (median of 290 analyses per sample) [36].

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



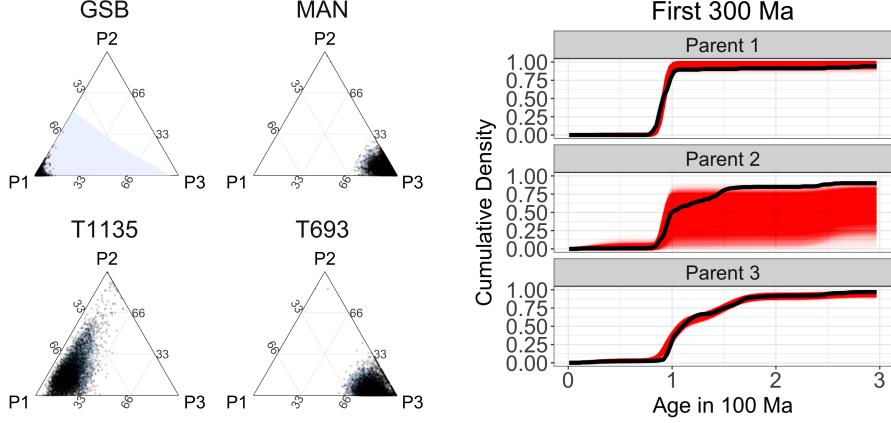
(a) Ternary plot showing posterior predictive density estimates of mixing proportions. Each dot represents one of 500 MCMC samples. The black shading is proportional to the estimated posterior probability density.

(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.

362 of each parent for child T693 confirms that the primary component of the
 363 mixture is from the P3 (Figure 7a). Figure 7b shows the model is capturing the
 364 basic patterns in the parent CDFs as the estimated CDFs are very close to the
 365 empirical CDFs for the parents.

366 Application of the bottom-up, end-member unmixing model to the data set
 367 in [36] results in non-identifiability issues as the simulated age distributions as
 368 evidenced by the larger posterior uncertainty in Figure 8b. Posterior probability
 369 estimates are obtained for the proportion of each child that comes from the
 370 modeled end-member (Figure 8a), but based on the differences in estimates from
 371 the bottom-up mixing model estimates, the distributions should be interpreted
 372 cautiously. The estimated CDFs shown in red in Figure 8b with the empirical
 373 CDFs from the parent data shown in black suggest that the unmixing model is
 374 performing well despite the fact that the model is unaware of the parent data
 375 although there is much uncertainty about the CDF for P2, likely due to the



(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.

³⁷⁶ similarity in distribution to P1 (Figure 5).

³⁷⁷ A large overlap in the distribution of parent ages is a feature that occurs,
³⁷⁸ particularly in detrital zircon geochronology studies. The preservation of zircons
³⁷⁹ through multiple cycles of erosion and re-sedimentation means that overlapping
³⁸⁰ zircon ages will be present in many rocks, for example the preponderance of
³⁸¹ Grenville zircons in a host of sedimentary formations of varying age. For parent
³⁸² age distributions that are quite similar to one another, the reconstruction of
³⁸³ the unknown parent distributions suffers from weak identifiability. In these
³⁸⁴ situations, the estimated parents jointly contain all of the correct formation
³⁸⁵ events, but the model is unable to attribute the formation events to the correct
³⁸⁶ parents. In other words, while the model identifies the correct age components,
³⁸⁷ the model sometimes struggles to correctly group these components into the
³⁸⁸ correct parent distributions. This is not an unexpected result because Bayesian
³⁸⁹ nonparametric models are well understood to suffer from non-identifiability

390 issues in the context of the Bayesian nonparametric framework [11, 8, 31, 13].
391 Non-identifiability is not a weakness of the particular proposed model framework;
392 the non-identifiability applies to end-member unmixing models in general [45].
393 To overcome the non-identifiability, a potential solution is to impose constraints
394 on the end-members and initialization conditions [9, 27, 6]. Therefore, any
395 end-member unmixing model that uses only child age distributions will have
396 issues in accurately reconstructing the parent distributions if the assumption
397 of the constraints is not met (i.e. the parent age distributions are structurally
398 similar). Bottom-up unmixing models provide a useful way to explore large
399 detrital datasets with unknown sedimentary sources. Providing a way to identify
400 those datasets that either are or are not susceptible to non-identifiability, and
401 thus not amenable to bottom-up unmixing, is critical to success. An advantage
402 of our framework is the end-member unmixing model produces uncertainty
403 estimates that are larger when the model is weakly identifiable. Thus, the
404 uncertainty intervals can be used as a diagnostic to check for identifiability.

405 7. Conclusion

406 Starting from a conceptual model of how sediments mix over a landscape,
407 we developed a generative Bayesian nonparametric statistical model for detrital
408 mineral age data. This model allows us to characterize the uncertainty in the age
409 distributions of parents and children and the mixing coefficients while explicitly
410 accounting for the uncertainties in measured dates [42]. Because the model can
411 simulate sediment age distributions, we can directly explore the assumptions of
412 the model by simulating synthetic data. Running a simulation experiment demon-
413 strated the model is capable of recovering simulated distributions supporting
414 the usefulness of the models when applied to the observed data.

415 We proposed two frameworks to model the sediment mixing mechanisms: the

416 top-down mixing model where mineral dates are measured for both parent and
417 child sediments and a bottom-up unmixing framework where mineral dates are
418 only measured for the children. The top-down model estimated the parent and
419 child distributions and the mixing proportions with high precision and accuracy.
420 The bottom-up model occasionally demonstrated evidence of non-identifiability,
421 suggesting the inference for the bottom-up model is less precise than for the
422 top-down mixing model; however, the variances of these estimates are larger
423 in our bottom-up unmixing model, recognizing the challenges in reconstructing
424 unobserved parent age distributions while simultaneously providing feedback to
425 the user about the potential pitfalls in being overly confident about the parent
426 distributions.

427 Using data collected along the central California coast data in [36], we
428 produced estimates of the mixing proportions of child sediments with both the
429 mixing and unmixing models. Other studies have produced similar estimates [1,
430 2]; our contribution is the mechanistic model framework that produces estimates
431 of mixing with associated uncertainty. We account for dating uncertainty directly
432 in the model and by modifying the model statement to remove dating uncertainty
433 (i.e., removing the data models in Eq. (1)) we can examine the effects on inference
434 by not accounting for the dating uncertainty.

435 Direct, probabilistic estimates of uncertainty and the ability to calculate de-
436 rived quantities with uncertainty is a benefit of the Bayesian methodology. Thus,
437 we can answer questions like what is the probability that at least 50% of child
438 sediment T693 comes from parent P3? The answer is calculated directly from the
439 posterior samples using the Monte Carlo approximation $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3}^{(\ell)} \geq 0.5\} =$
440 0.672, where $\ell = 1, \dots, L$ are the indices of the MCMC samples and $\phi_{P3}^{(\ell)}$ is the
441 estimated mixing proportion for the ℓ th MCMC iteration. The probability that
442 at least 50% of the child sediment comes from parent P3 and at least 25% comes

443 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
444 model produces a posterior probability, any such probabilistic questions can be
445 calculated as derived quantities. For example, we can ask questions like: what
446 proportion of a given sample contains grains older than a given age? or what
447 is the probability that an unobserved parent contains grains with a particular
448 age range. Once the posterior samples have been calculated, any such questions
449 about derived quantities are answerable using posterior samples.

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

462 8. Acknowledgements

463 Support for SAJ came from the FEDMAP component of the US Geological
464 Survey National Cooperative Geologic Mapping Program. This draft manuscript
465 is distributed solely for purposes of scientific peer review. Its content is delib-
466 erative and predecisional, so it must not be disclosed or released by reviewers.
467 Because the manuscript has not yet been approved for publication by the U.S.
468 Geological Survey (USGS), it does not represent any official USGS finding or

⁴⁶⁹ policy. Any use of trade, firm, or product names is for descriptive purposes only
⁴⁷⁰ and does not imply endorsement by the U.S. Government.

⁴⁷¹ **References**

- ⁴⁷² [1] Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005a. Construction of
⁴⁷³ detrital mineral populations: insights from mixing of U-Pb zircon ages in
⁴⁷⁴ Himalayan rivers. *Basin Research* 17, 463–485.
- ⁴⁷⁵ [2] Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005b. U–pb zircon ages
⁴⁷⁶ as a sediment mixing tracer in the nepal himalaya. *Earth and Planetary
477 Science Letters* 235, 244–260.
- ⁴⁷⁸ [3] Berliner, L.M., 2003. Physical-statistical modeling in geophysics. *Journal
479 of Geophysical Research: Atmospheres* 108.
- ⁴⁸⁰ [4] Chang, W., Haran, M., Applegate, P., Pollard, D., 2016. Calibrating an
⁴⁸¹ ice sheet model using high-dimensional binary spatial data. *Journal of the
482 American Statistical Association* 111, 57–72.
- ⁴⁸³ [5] Chen, J.H., Moore, J.G., 1982. Uranium-lead isotopic ages from the sierra
⁴⁸⁴ nevada batholith, california. *Journal of Geophysical Research: Solid Earth*
⁴⁸⁵ 87, 4761–4784.
- ⁴⁸⁶ [6] Chen, W., Guillaume, M., 2012. HALS-based NMF with flexible constraints
⁴⁸⁷ for hyperspectral unmixing. *EURASIP Journal on Advances in Signal
488 Processing* 2012, 54.
- ⁴⁸⁹ [7] de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple
⁴⁹⁰ Lang, D., Bodik, R., 2017. Programming with models: writing statistical
⁴⁹¹ algorithms for general model structures with nimble. *Journal of Computational
492 and Graphical Statistics* 26, 403–413. doi:10.1080/10618600.2016.
⁴⁹³ 1172487.

- 494 [8] Diebolt, J., Robert, C.P., 1994. Estimation of finite mixture distributions
495 through Bayesian sampling. *Journal of the Royal Statistical Society. Series*
496 *B (Methodological)* , 363–375.
- 497 [9] Donoho, D., Stodden, V., 2004. When does non-negative matrix factorization
498 give a correct decomposition into parts?, in: *Advances in neural information*
499 *processing systems*, pp. 1141–1148.
- 500 [10] Ferguson, T.S., 1973. A Bayesian analysis of some nonparametric problems.
501 *The annals of statistics* , 209–230.
- 502 [11] Ferguson, T.S., 1983. Bayesian density estimation by mixtures of normal
503 distributions, in: *Recent advances in statistics*. Elsevier, pp. 287–302.
- 504 [12] Fosdick, J.C., Grove, M., Graham, S.A., Hourigan, J.K., Lovera, O., Ro-
505 mans, B.W., 2015. Detrital thermochronologic record of burial heating
506 and sediment recycling in the magallanes foreland basin, patagonian andes.
507 *Basin Research* 27, 546–572.
- 508 [13] Frühwirth-Schnatter, S., 2006. Finite mixture and Markov switching models.
509 Springer Science & Business Media.
- 510 [14] Gehrels, G., 2012. Detrital zircon U-Pb geochronology: Current methods
511 and new opportunities. *Tectonics of sedimentary basins: Recent advances* ,
512 45–62.
- 513 [15] Gehrels, G., 2014. Detrital zircon U-Pb geochronology applied to tectonics.
514 *Annual Review of Earth and Planetary Sciences* 42, 127–149.
- 515 [16] Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation
516 and Bayesian model determination. *Biometrika* 82, 711–732.

- 517 [17] Guan, Y., Haran, M., Pollard, D., 2018. Inferring ice thickness from a
518 glacier dynamics model and multiple surface data sets. Environmetrics 29,
519 e2460.
- 520 [18] Haario, H., Saksman, E., Tamminen, J., et al., 2001. An adaptive Metropolis
521 algorithm. Bernoulli 7, 223–242.
- 522 [19] Hamilton, N., 2018. ggtern: An Extension to ggplot2, for the creation of
523 ternary diagrams. URL: <https://CRAN.R-project.org/package=ggtern>. r
524 package version 2.2.2.
- 525 [20] Hefley, T.J., Brost, B.M., Hooten, M.B., 2017. Bias correction of bounded
526 location errors in presence-only data. Methods in Ecology and Evolution 8,
527 1566–1573.
- 528 [21] Hooten, M.B., Hobbs, N., 2015. A guide to Bayesian model selection for
529 ecologists. Ecological Monographs 85, 3–28.
- 530 [22] Irwin, W.P., Wooden, J.L., 1999. Plutons and accretionary episodes of
531 the Klamath Mountains, California and Oregon. Technical Report. US
532 Geological Survey.
- 533 [23] Jasra, A., Stephens, D.A., Gallagher, K., Holmes, C.C., 2006. Bayesian
534 mixture modelling in geochronology via markov chain monte carlo. Mathe-
535 matical Geology 38, 269–300.
- 536 [24] Kimbrough, D.L., Grove, M., Gehrels, G.E., Dorsey, R.J., Howard, K.A.,
537 Lovera, O., Aslan, A., House, P.K., Pearthree, P.A., 2015. Detrital zircon
538 u-pb provenance of the colorado river: A 5 my record of incision into cover
539 strata overlying the colorado plateau and adjacent regions. Geosphere 11,
540 1719–1748.

- 541 [25] Lock, E.F., Dunson, D.B., 2015. Shared kernel Bayesian screening.
542 Biometrika 102, 829–842.
- 543 [26] Mason, C.C., Fildani, A., Gerber, T., Blum, M.D., Clark, J.D., Dykstra,
544 M., 2017. Climatic and anthropogenic influences on sediment mixing in the
545 mississippi source-to-sink system using detrital zircons: Late pleistocene to
546 recent. Earth and Planetary Science Letters 466, 70–79.
- 547 [27] Miao, L., Qi, H., 2007. Endmember extraction from highly mixed data
548 using minimum volume constrained nonnegative matrix factorization. IEEE
549 Transactions on Geoscience and Remote Sensing 45, 765–777.
- 550 [28] Miller, J.W., Harrison, M.T., 2018. Mixture models with a prior on the
551 number of components. Journal of the American Statistical Association
552 113, 340–356.
- 553 [29] Puetz, S.J., Ganade, C.E., Zimmermann, U., Borchardt, G., 2018. Statistical
554 analyses of global U-Pb database 2017. Geoscience Frontiers 9, 121–145.
- 555 [30] Reiners, P.W., Brandon, M.T., 2006. Using thermochronology to understand
556 orogenic erosion. Annual Review Earth Planetary Sciences 34, 419–466.
- 557 [31] Richardson, S., Green, P.J., 1997. On Bayesian analysis of mixtures with
558 an unknown number of components (with discussion). Journal of the Royal
559 Statistical Society: series B (statistical methodology) 59, 731–792.
- 560 [32] Romans, B.W., Castelltort, S., Covault, J.A., Fildani, A., Walsh, J., 2016.
561 Environmental signal propagation in sedimentary systems across timescales.
562 Earth-Science Reviews 153, 7–29.
- 563 [33] Saylor, J.E., Sundell, K., Sharman, G., 2019. Characterizing sediment
564 sources by non-negative matrix factorization of detrital geochronological
565 data. Earth and Planetary Science Letters 512, 46–58.

- 566 [34] Sharman, G.R., Johnstone, S.A., 2017. Sediment unmixing using detrital
567 geochronology. *Earth and Planetary Science Letters* 477, 183–194.
- 568 [35] Sharman, G.R., Sylvester, Z., Covault, J.A., 2019. Conversion of tectonic and
569 climatic forcings into records of sediment supply and provenance. *Scientific
570 reports* 9, 4115.
- 571 [36] Sickmann, Z.T., Paull, C.K., Graham, S.A., 2016. Detrital-zircon mixing
572 and partitioning in fluvial to deep marine systems, central California, USA.
573 *Journal of Sedimentary Research* 86, 1298–1307.
- 574 [37] Stock, G.M., Ehlers, T.A., Farley, K.A., 2006. Where does sediment
575 come from? quantifying catchment erosion with detrital apatite (u-th)/he
576 thermochronometry. *Geology* 34, 725–728.
- 577 [38] Sundell, K., Saylor, J.E., 2017. Unmixing detrital geochronology age
578 distributions. *Geochemistry, Geophysics, Geosystems* .
- 579 [39] Tipton, J., Hooten, M., Goring, S., 2017. Reconstruction of spatio-temporal
580 temperature from sparse historical records using robust probabilistic princi-
581 pal component regression. *Advances in Statistical Climatology, Meteorology
582 and Oceanography* 3, 1–16.
- 583 [40] Tipton, J.R., Hooten, M.B., Nolan, C., Booth, R.K., McLachlan, J., 2019.
584 Predicting paleoclimate from compositional data using multivariate Gaus-
585 sian process inverse prediction. *Annals of Applied Statistics* 13, 2363–2388.
586 doi:10.1214/19-AOAS1281.
- 587 [41] Tipton, J.R., Hooten, M.B., Pederson, N., Tingley, M.P., Bishop, D.,
588 2016. Reconstruction of late Holocene climate based on tree growth and
589 mechanistic hierarchical models. *Environmetrics* 27, 42–54.

- 590 [42] Tye, A., Wolf, A., Niemi, N., 2019. Bayesian population correlation: A
591 probabilistic approach to inferring and comparing population distributions
592 for detrital zircon ages. *Chemical Geology* 518, 67–78.
- 593 [43] Vermeesch, P., 2012. On the visualisation of detrital age distributions.
594 *Chemical Geology* 312, 190–194.
- 595 [44] Weltje, G.J., 1997. End-member modeling of compositional data: Numerical-
596 statistical algorithms for solving the explicit mixing problem. *Mathematical
597 Geology* 29, 503–549.
- 598 [45] Weltje, G.J., Prins, M.A., 2007. Genetically meaningful decomposition of
599 grain-size distributions. *Sedimentary Geology* 202, 409–424.
- 600 [46] Wotzlaw, J.F., Schaltegger, U., Frick, D.A., Dungan, M.A., Gerdes, A.,
601 Günther, D., 2013. Tracking the evolution of large-volume silicic magma
602 reservoirs from assembly to supereruption. *Geology* 41, 867–870.