

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

John Tipton · Glenn Sharman · Samuel Johnstone

Received: date / Accepted: date

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

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

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

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

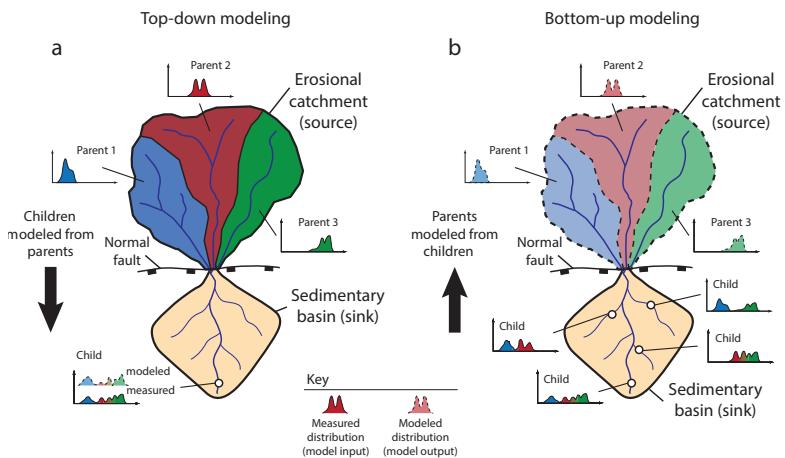


Fig. 1: Schematic depiction of a sediment routing system with an erosional source region characterized by three parents (black, red, and green) and an associated sedimentary basin (yellow). Call out plots represent the age of detrital minerals from the parents and how they are mixed to form child distributions.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

22 1 Introduction

23 To understand the origins of modern and ancient physical geography one must
 24 understand how erosional landscapes and associated sedimentary basins evolve
 25 through time (Romans et al., 2016). 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
 26 inference about these processes that generated the modern landscape. The
 27 ability to decipher the relative proportions of sources that eroded to produce
 28 sediment informs understanding of the underlying geologic processes controlling
 29 the evolution of the Earth’s surface (Stock et al., 2006; Sharman et al.,
 30 2019; Mason et al., 2017; Kimbrough et al., 2015).

31 One of the most common ways to characterize the provenance of sediment
 32 is detrital geochronology – dating the time at which the individual minerals
 33 that make up a rock formed or cooled. These mineralization events typically
 34 reflect the timing of igneous rock forming events or metamorphic alteration of
 35 previously existing rocks (Gehrels, 2014). In other cases mineral ages reflect
 36 the history of mineral cooling (e.g., ‘thermochronology’, (Reiners and Bran-
 37 don, 2006)). Detrital geochronologic ages are most commonly determined from
 38

40 measurements of radiogenic isotopes contained within individual mineral crystals.
41 The decay of uranium (U) to lead (Pb) within zircon, a relatively robust
42 mineral, makes this approach ideally suited for tracking sedimentary mixing
43 (Amidon et al., 2005b; Sundell and Saylor, 2017; Sharman and Johnstone,
44 2017).

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

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

77 2 Model Overview

78 To define the statistical model, we follow the convention that letters repre-
79 sent data and Greek symbols represent parameters. A plaintext symbol (y)
80 represents a scalar, a bold lowercase symbol represents a vector (\mathbf{y}), and a
81 bold uppercase symbol is a matrix (\mathbf{Z}) whose columns are vectors written as
82 $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_p)$. We use the notation $[y]$ to represent the probability distribu-

tion/mass function (pdf/pmf) and let $[y|\theta]$ represent the conditional pdf/pmf of the random variable y given θ .

Following Berliner (2003), the statistical model described below is divided into three components: the data model, the process model, and the prior model. In general, the data model defines probability distributions that describe the variability in the data due to the observation process. The data model can be modified to account for non Gaussian measurement processes like counts, outliers, spatial/temporal errors, etc (Tipton et al., 2017; Hefley et al., 2017). Process models describe the best scientific understanding of the process of interest. For example, process models have been used to describe the monthly response of trees to climate (Tipton et al., 2016), the relationship between climate and pollen in sediments (Tipton et al., 2019), and the movement of ice sheets in Antarctica (Chang et al., 2016; Guan et al., 2018). The prior model describes the range of parameter values that are plausible. Sometimes the prior model is used as regularization to improve the generalization of the model to unobserved data (Hooten and Hobbs, 2015).

3 Top-down mixing model

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

3.1 Top-down mixing data model

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

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

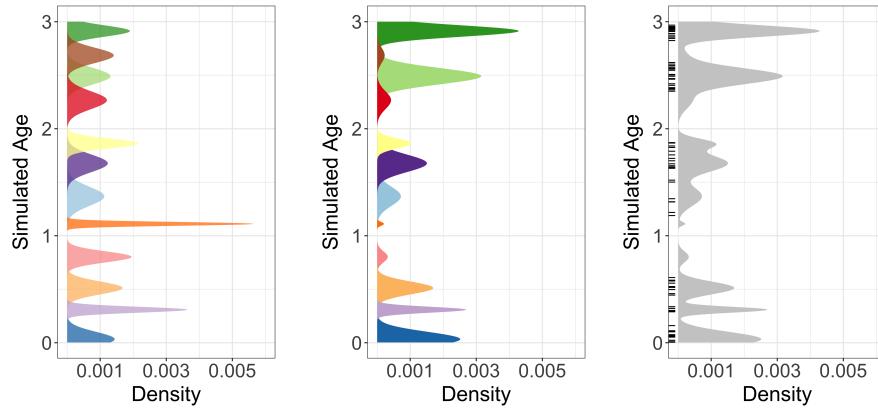
We break the variable naming convention and let $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{z}}_b$ be latent parameters that represent the true, unobserved age of the sediments where \mathbf{y} (\mathbf{z}_b) will be close to $\tilde{\mathbf{y}}$ ($\tilde{\mathbf{z}}_b$) because the dating uncertainty is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

3.2 Top-down mixing process model

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

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

Consider the latent, unknown age of a single mineral grain from either the child $\tilde{\mathbf{y}}$ or one of the $b = 1, \dots, B$ parents $\tilde{\mathbf{z}}_b$. We assign the range of mineral ages for the k th formation event the base probability distribution $G(\theta_{bk})$ which depends on parameters θ_{bk} . There are many possible choices for the base distribution $G(\theta_{bk})$; we assume a normal distribution $N(\mu_{bk}, \sigma_{bk}^2)$ with mean μ_{bk} and variance σ_{bk}^2 , therefore $\theta_{bk} = (\mu_{bk}, \sigma_{bk}^2)'$. Other possible choices include a log-normal or gamma distribution that enforces a positive support on the observed age dates. Because we assume the age distribution



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

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

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

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

of a single mineral formation event is relatively short with respect to geologic time, the variance parameters σ_{bk}^2 will be small relative to geologic time.

In most cases the true number of mineral formation events K recorded by a detrital sample is unknown. Under our model, the set of all K mineral formation events is a mixture of K normal distributions, as shown in Fig. 2a where each mineral formation event is shown in a different color. The centers of each age distribution in Fig. 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, 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 (Amidon et al., 2005a,b). Figure 2b shows the age distributions in Fig. 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 (Fig. 2c). Notice that in Fig. 2c, the colors from Figs. 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

180 distribution in Fig. 2c, shown as a rug plot with each tick representing the
 181 observed detrital mineral grain date. Thus, the number of mineral formation
 182 events is potentially challenging to extract from the data.

183 Consider a detrital mineral date \tilde{z}_{ib} from a parent sediment source b from
 184 $i = 1, \dots, n_b$ measurements. The single sediment grain comes from a mineral
 185 formation event implying the mixture distribution

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

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

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

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

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

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

198 In addition to modeling the age distribution of the parents, the process
 199 model specifies the proportion of each parent distribution in the child distri-
 200 bution. We represent the mixing proportions of the B parent distributions with
 201 the parameter $\boldsymbol{\phi} = (\phi_1, \dots, \phi_B)'$, with $\sum_{b=1}^B \phi_b = 1$. The parameter ϕ_b is the
 202 proportion of the child distribution that comes from parent \$b\$ and accounts
 203 for differential mixing of parents. For parents that are comprised of bedrock,
 204 $\boldsymbol{\phi}$ is a function of each parent's relative aerial extent in the drainage catch-
 205 ment, average erosion rate, and average concentration of the detrital mineral
 206 of interest (Amidon et al., 2005a). If parents are sediment inputs (e.g., rivers),
 207 then $\boldsymbol{\phi}$ is a function of each parent's relative sediment supply and the average
 208 concentration of the detrital mineral of interest within the sediment.

209 For a single child sediment mineral date \tilde{y}_i , that sediment grain comes from
 210 only one parent. We define the categorical random variable δ_i to represent

211 which parent distribution the child sediment came from. Using the categorical
 212 variable, the distribution of the child sediment grain is

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

213 Then, the probability that the sediment grain came from parent b is $\phi_b \equiv$
 214 $P(\delta_i = b)$. Defining the indicator variable $I\{\delta_i = b\}$ where $P(\delta_i = b) =$
 215 $E(I\{\delta_i = b\})$, we can write the age distribution over all sediment grains as

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

216 where, like the parent mixing model, we integrate out the component in-
 217 dicator variables $\boldsymbol{\delta}$. After integrating out the parent component membership
 218 indicators, the child sediment grains have the age distribution

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

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

221 Combining the above results, the full process model is

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

222 3.3 Top-down mixing prior model

223 The conceptual process model assumes the number of mineral formation events
 224 K is known. In practice, the number of formation events is unknown and is a
 225 parameter to be estimated. In fact, it is likely that the different parent sites
 226 will have different numbers of mineral formation events based on site-specific

227 history. The prior model addresses the fundamental question of estimating the
 228 number of mineral formation events.

229 There are a number of potential approaches to model the unknown number
 230 of formation events. First, one can treat the number of formation events as a
 231 fixed parameter, perform a grid search over the different number of formation
 232 events, and choose the model that best fits the data (Miller and Harrison,
 233 2018). A second approach is to model the number of formation events using
 234 a reversible jump algorithm (Green, 1995). The third approach is to assign a
 235 Dirichlet process prior over the number of formation events. The Dirichlet pro-
 236 cess estimates an unknown number of components without *a priori* specifying
 237 the number.

238 The Dirichlet process is an infinite dimensional stochastic process which is
 239 a distribution over distributions (Ferguson, 1973). We use the stick-breaking
 240 representation of a Dirichlet process

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

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

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

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

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

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

269 **3.4 Top-down mixing posterior distribution**

270 The top-down mixing model posterior distribution is

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

271 where the first three line on the right hand side of the proportional symbol
 272 are the data, process, and the prior model, respectively. We estimate
 273 the posterior using Markov Chain Monte Carlo (MCMC) with the *R* pack-
 274 age *NIMBLE* (de Valpine et al., 2017) using an adaptive block Metropolis-
 275 Hastings algorithm (Haario et al., 2001). The constrained auxiliary variable $\tilde{\mathbf{p}}$
 276 and standard deviation $\boldsymbol{\sigma}$ are transformed to unconstrained support (logit- and
 277 log-scale transformations) for tuning the Metropolis-Hastings block proposals,
 278 with corresponding Jacobian adjustments to the acceptance probabilities. The
 279 sampling of the sum-to-one mixing proportion ϕ is performed by introducing
 280 auxiliary variables $\tilde{\boldsymbol{\phi}}$, assigning a stick-breaking prior on $\tilde{\boldsymbol{\phi}}$, then sampling on a
 281 logit-scale after correcting for the transformation using the Jacobian to induce
 282 a Dirichlet($\alpha_\phi \mathbf{1}$) prior on $\boldsymbol{\phi}$.

283 **4 Bottom-up unmixing model**

284 The second research question is: can we reconstruct unobserved parent age
 285 distributions from multiple child observations? In previous work, this analysis
 286 has been variably termed “end-member mixing analysis”, “end-member mod-
 287 eling”, or “end-member analysis” as applied to unmixing grain size or detrital
 288 age distributions (Sharman and Johnstone, 2017; Saylor et al., 2019). The end-
 289 member unmixing analysis has two components. First, the number of parents
 290 B is unknown and needs to be estimated. Second, given the number of parents
 291 B , what are the unobserved mineral formation age distributions for the B
 292 parents? For this paper, we assume the number of parents B is known. There

293 are a number of criteria for selecting the number of parents including using
 294 Bayesian information criteria, reversible jump MCMC (Jasra et al., 2006), as-
 295 suming a Dirichlet process over the number of parents, or fitting a mixture of
 296 finite mixtures (Miller and Harrison, 2018). Rather than explore these ideas,
 297 we devote our effort on developing the unmixing model for a fixed number of
 298 parents (Miller and Harrison, 2018). The end-member model uses the same
 299 general framework presented in the mixture of Gaussians model (3) with some
 300 modifications.

301 4.1 Bottom-up unmixing data model

302 Let $d = 1, \dots, D$ index the D child sediments that are each composed of $i =$
 303 $1, \dots, n_d$ samples. As before, we assume a Gaussian dating error distribution
 304 for child d given by

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

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

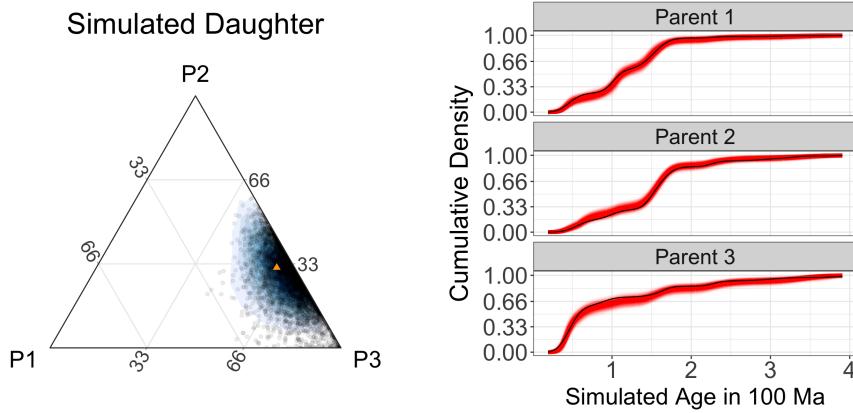
307 Unlike in the top-down mixing model above, none of the parent **zs** are
 308 observed. Hence, the parent distributions are estimated entirely using child
 309 sediment observations. The end-member process model for a fixed given num-
 310 ber of parents B is

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

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

319 4.2 Bottom-up unmixing posterior distribution

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



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

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

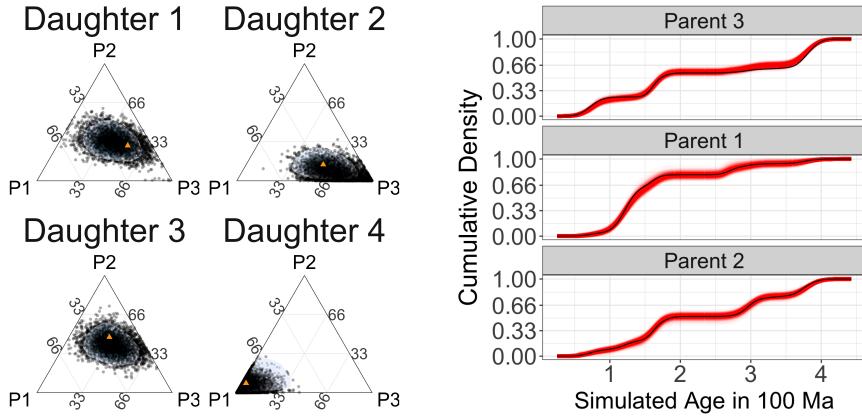
Fig. 3: Simulation study results for the top-down mixture modeling approach.

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

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

326 5 Simulation of synthetic detrital age distributions

327 We explore the performance of the model using a simulation of synthetic de-
 328 trital age distributions. The aim of the simulation study is to understand how
 329 the model performs using realistic data and verify the model is capable of re-
 330 covering the simulated parameters. The simulation study framework can also
 331 be used to understand how uncertainty in estimation varies with respect to
 332 sample size, variability in the data, and other questions of interest.



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

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

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

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

341 The posterior samples for the mixing proportion ϕ are shown in Fig. 3a as
 342 black dots with a smoothed posterior density shown shaded in blue (Hamilton,
 343 2018), and the simulated true mixing proportion is represented by the
 344 orange triangle, demonstrating that the model is accurately estimating the
 345 mixing proportion as the orange triangle is in the bulk of the posterior sam-
 346 ples. Figure 3b shows the estimated CDFs with posterior samples in red and
 347 the simulated CDF in black. The results in Fig. 3 demonstrate that the model
 348 is accurately estimating the simulated mixing proportions ϕ as well as the
 349 parent age distributions, validating the effectiveness of the top-down mixing
 350 model to recover simulated parameters of interest.

351 The second simulation generated data from the bottom-up, end-member
 352 unmixing model (Fig. 1b) to test how well the proposed framework can re-
 353 construct unobserved parent distributions. For the simulation, we used $B = 3$
 354 parents and $D = 20$ children where each child consisted of 250 measured sed-

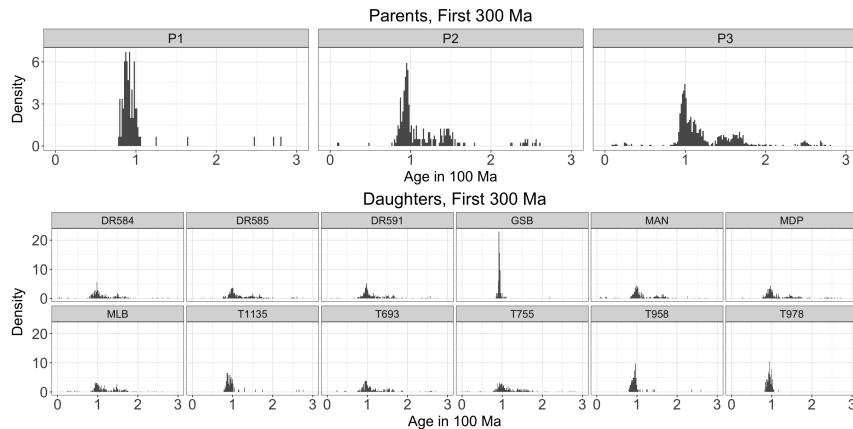


Fig. 5: The sediment age data used for the mixing and unmixing models. The three parent age distributions are shown in the top plot and the 12 child age distributions are shown in the bottom plot. The x -axes represent the measured age in 100 Ma and the y -axes shows the empirical density.

355 iment grain dates following the model in (5). The dating uncertainties (σ_y)
 356 were set at about 1-3% of the total range of the age distribution.

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

369 6 Application to a Natural Case Study

370 We apply the model presented here to a well-constrained modern dataset from
 371 the central California coast (Sickmann et al., 2016) shown in Fig. 5 where we
 372 focus on sediments dating to the most recent 300 Ma. Following the same
 373 mixing framework presented in Sharman and Johnstone (2017), there are five
 374 samples (river and beach sediment) used to characterize three distinct sediment
 375 inputs (parents) to the region, each with a distinct detrital age distribution
 376 (Fig. 5). Parents 1 and 2 (P1 and P2) are comprised of river samples (CAR and
 377 SAR, respectively) that represent sediment sources along the Big Sur coastline

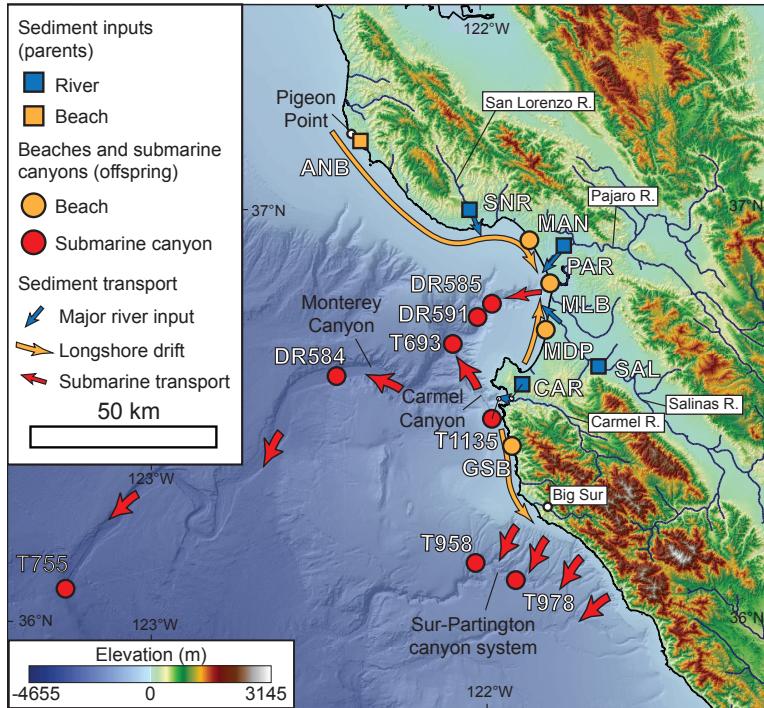
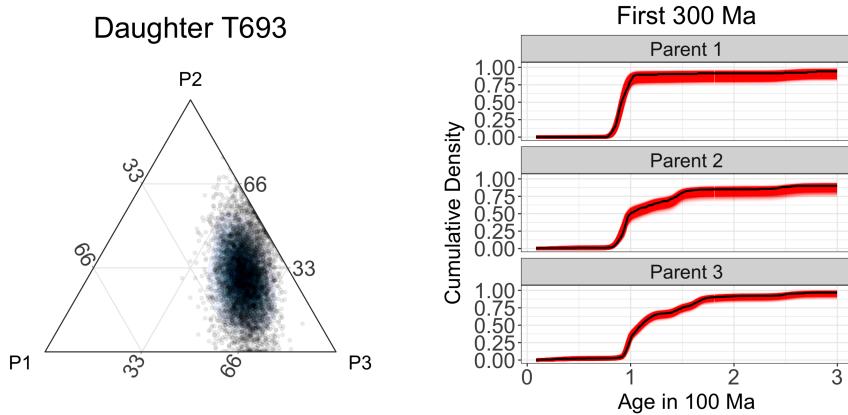


Fig. 6: Locations of the parents and children data for the study region in California, USA.

and Salinas River drainage, respectively. Parent 3 (P3) is comprised of two river samples (SNR and PAR) and one beach sample (ANB) that represent northern sediment sources in the Santa Cruz Mountains and western Diablo Range (Sickmann et al., 2016; Sharman and Johnstone, 2017). 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 (Fig. 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) (Sickmann et al., 2016).

We first examine the top-down mixture model (Fig. 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 (Fig. 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 of each parent for child T693 confirms that the primary component of the mixture is from the P3 (Fig. 7a). Figure 7b shows the model is capturing the basic patterns in the parent CDFs as the estimated CDFs are very close to the empirical CDFs for the parents.



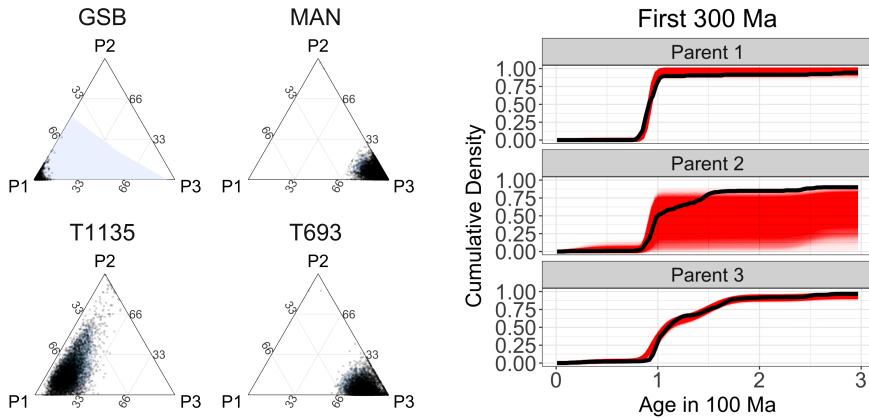
(a) Ternary plot showing posterior predictive density estimates of mixing proportions. Each dot represents one of 500 MCMC samples. The black shading is proportional to the estimated posterior probability density.

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

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

397 Application of the bottom-up, end-member unmixing model to the data
 398 set in Sickmann et al. (2016) results in non-identifiability issues as the sim-
 399 ultated age distributions as evidenced by the larger posterior uncertainty in
 400 Fig. 8b. Posterior probability estimates are obtained for the proportion of
 401 the differences in estimates from the bottom-up mixing model estimates, the
 402 distributions should be interpreted cautiously. The estimated CDFs shown in
 403 red in Fig. 8b with the empirical CDFs from the parent data shown in black
 404 suggest that the unmixing model is performing well despite the fact that the
 405 model is unaware of the parent data although there is much uncertainty about
 406 the CDF for P2, likely due to the similarity in distribution to P1 (Fig. 5).

407 A large overlap in the distribution of parent ages is a feature that oc-
 408 curs, particularly in detrital zircon geochronology studies. The preservation
 409 of zircons through multiple cycles of erosion and re-sedimentation means that
 410 overlapping zircon ages will be present in many rocks, for example the pre-
 411 ponderance of Grenville zircons in a host of sedimentary formations of varying
 412 age. For parent age distributions that are quite similar to one another, the
 413 reconstruction of the unknown parent distributions suffers from weak identi-
 414 fiability. In these situations, the estimated parents jointly contain all of the
 415 correct formation events, but the model is unable to attribute the formation
 416 events to the correct parents. In other words, while the model identifies the cor-
 417 rect age components, the model sometimes struggles to correctly group these



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

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

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

419 components into the correct parent distributions. This is not an unexpected
 420 result because Bayesian nonparametric models are well understood to suffer
 421 from non-identifiability issues in the context of the Bayesian nonparametric
 422 framework (Ferguson, 1983; Diebolt and Robert, 1994; Richardson and Green,
 423 1997; Frühwirth-Schnatter, 2006). Non-identifiability is not a weakness of the
 424 particular proposed model framework; the non-identifiability applies to end-
 425 member unmixing models in general (Weltje and Prins, 2007). To overcome
 426 the non-identifiability, a potential solution is to impose constraints on the end-
 427 members and initialization conditions (Donoho and Stodden, 2004; Miao and
 428 Qi, 2007; Chen and Guillaume, 2012). Therefore, any end-member unmixing
 429 model that uses only child age distributions will have issues in accurately re-
 430 constructing the parent distributions if the assumption of the constraints is
 431 not met (i.e., the parent age distributions are structurally similar). Bottom-up
 432 unmixing models provide a useful way to explore large detrital datasets with
 433 unknown sedimentary sources. Providing a way to identify those datasets that
 434 either are or are not susceptible to non-identifiability, and thus not amenable
 435 to bottom-up unmixing, is critical to success. An advantage of our framework
 436 is the end-member unmixing model produces uncertainty estimates that are
 437 larger when the model is weakly identifiable. Thus, the uncertainty intervals
 438 can be used as a diagnostic to check for identifiability.

⁴³⁹ **7 Conclusion**

⁴⁴⁰ Starting from a conceptual model of how sediments mix over a landscape, we
⁴⁴¹ developed a generative Bayesian nonparametric statistical model for detrital
⁴⁴² mineral age data. This model allows us to characterize the uncertainty in the
⁴⁴³ age distributions of parents and children and the mixing coefficients while ex-
⁴⁴⁴ plicitly accounting for the uncertainties in measured dates (Tye et al., 2019).
⁴⁴⁵ Because the model can simulate sediment age distributions, we can directly
⁴⁴⁶ explore the assumptions of the model by simulating synthetic data. Running
⁴⁴⁷ a simulation experiment demonstrated the model is capable of recovering sim-
⁴⁴⁸ ultated distributions supporting the usefulness of the models when applied to
⁴⁴⁹ the observed data.

⁴⁵⁰ We proposed two frameworks to model the sediment mixing mechanisms:
⁴⁵¹ the top-down mixing model where mineral dates are measured for both par-
⁴⁵² ent and child sediments and a bottom-up unmixing framework where mineral
⁴⁵³ dates are only measured for the children. The top-down model estimated the
⁴⁵⁴ parent and child distributions and the mixing proportions with high precision
⁴⁵⁵ and accuracy. The bottom-up model occasionally demonstrated evidence of
⁴⁵⁶ non-identifiability, suggesting the inference for the bottom-up model is less
⁴⁵⁷ precise than for the top-down mixing model; however, the variances of these
⁴⁵⁸ estimates are larger in our bottom-up unmixing model, recognizing the chal-
⁴⁵⁹ lenges in reconstructing unobserved parent age distributions while simultane-
⁴⁶⁰ ously providing feedback to the user about the potential pitfalls in being overly
⁴⁶¹ confident about the parent distributions.

⁴⁶² Using data collected along the central California coast data in (Sickmann
⁴⁶³ et al., 2016), we produced estimates of the mixing proportions of child sed-
⁴⁶⁴ iments with both the mixing and unmixing models. Other studies have pro-
⁴⁶⁵ duced similar estimates (Amidon et al., 2005a,b); our contribution is the mech-
⁴⁶⁶ anistic model framework that produces estimates of mixing with associated
⁴⁶⁷ uncertainty. We account for dating uncertainty directly in the model and by
⁴⁶⁸ modifying the model statement to remove dating uncertainty (i.e., removing
⁴⁶⁹ the data models in Eq. (1)) we can examine the effects on inference by not
⁴⁷⁰ accounting for the dating uncertainty.

⁴⁷¹ Direct, probabilistic estimates of uncertainty and the ability to calculate
⁴⁷² derived quantities with uncertainty is a benefit of the Bayesian methodol-
⁴⁷³ ogy. Thus, we can answer questions like what is the probability that at least
⁴⁷⁴ 50% of child sediment T693 comes from parent P3? The answer is calcu-
⁴⁷⁵ lated directly from the posterior samples using the Monte Carlo approxima-
⁴⁷⁶ tion $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3}^{(\ell)} \geq 0.5\} = 0.672$, where $\ell = 1, \dots, L$ are the indices of
⁴⁷⁷ the MCMC samples and $\phi_{P3}^{(\ell)}$ is the estimated mixing proportion for the
⁴⁷⁸ ℓ th MCMC iteration. The probability that at least 50% of the child sed-
⁴⁷⁹ iment comes from parent P3 and at least 25% comes from parent P2 is
⁴⁸⁰ $\frac{1}{L} \sum_{\ell=1}^L I\{\phi_{P3} \geq 0.5\} \times I\{\phi_{P2} \geq 0.25\} = 0.283$. Because the model produces
⁴⁸¹ a posterior probability, any such probabilistic questions can be calculated as
⁴⁸² derived quantities. For example, we can ask questions like: what proportion of

483 a given sample contains grains older than a given age? or what is the probability
484 that an unobserved parent contains grains with a particular age range.
485 Once the posterior samples have been calculated, any such questions about
486 derived quantities are answerable using posterior samples.

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

499 **8 Declarations**

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

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

514 **Keywords** Detrital sediment age distributions · Sediment unmixing ·
515 Bayesian nonparametrics · Uncertainty quantification

516 **References**

- 517 Amidon WH, Burbank DW, Gehrels GE (2005a) Construction of detrital mineral populations: insights from mixing of U-Pb zircon ages in Himalayan
518 rivers. *Basin Research* 17(4):463–485
519
520 Amidon WH, Burbank DW, Gehrels GE (2005b) U-Pb zircon ages as a sediment mixing tracer in the Nepal Himalaya. *Earth and Planetary Science Letters* 235(1-2):244–260
521
522

- 523 Berliner LM (2003) Physical-statistical modeling in geophysics. *Journal of*
524 *Geophysical Research: Atmospheres* 108(D24)
- 525 Chang W, Haran M, Applegate P, Pollard D (2016) Calibrating an ice sheet
526 model using high-dimensional binary spatial data. *Journal of the American*
527 *Statistical Association* 111(513):57–72
- 528 Chen JH, Moore JG (1982) Uranium-lead isotopic ages from the Sierra
529 Nevada batholith, California. *Journal of Geophysical Research: Solid Earth*
530 87(B6):4761–4784
- 531 Chen W, Guillaume M (2012) HALS-based NMF with flexible constraints for
532 hyperspectral unmixing. *EURASIP Journal on Advances in Signal Process-*
533 *ing* 2012(1):54
- 534 de Valpine P, Turek D, Paciorek C, Anderson-Bergman C, Temple Lang D,
535 Bodik R (2017) Programming with models: writing statistical algorithms
536 for general model structures with NIMBLE. *Journal of Computational and*
537 *Graphical Statistics* 26:403–413, DOI 10.1080/10618600.2016.1172487
- 538 Diebolt J, Robert CP (1994) Estimation of finite mixture distributions
539 through Bayesian sampling. *Journal of the Royal Statistical Society Series*
540 *B (Methodological)* pp 363–375
- 541 Donoho D, Stodden V (2004) When does non-negative matrix factorization
542 give a correct decomposition into parts? In: *Advances in Neural Information*
543 *Processing Systems*, pp 1141–1148
- 544 Ferguson TS (1973) A Bayesian analysis of some nonparametric problems. *The*
545 *Annals of Statistics* pp 209–230
- 546 Ferguson TS (1983) Bayesian density estimation by mixtures of normal distri-
547 *butions*. In: *Recent Advances in Statistics*, Elsevier, pp 287–302
- 548 Fosdick JC, Grove M, Graham SA, Hourigan JK, Lovera O, Romans BW
549 (2015) Detrital thermochronologic record of burial heating and sediment re-
550 cycling in the Magallanes foreland basin, Patagonian Andes. *Basin Research*
551 27(4):546–572
- 552 Frühwirth-Schnatter S (2006) Finite mixture and Markov switching models.
553 Springer Science & Business Media
- 554 Gehrels G (2012) Detrital zircon U-Pb geochronology: Current methods and
555 new opportunities. *Tectonics of Sedimentary Basins: Recent Advances* pp
556 45–62
- 557 Gehrels G (2014) Detrital zircon U-Pb geochronology applied to tectonics.
558 *Annual Review of Earth and Planetary Sciences* 42:127–149
- 559 Green PJ (1995) Reversible jump Markov chain Monte Carlo computation and
560 Bayesian model determination. *Biometrika* 82(4):711–732
- 561 Guan Y, Haran M, Pollard D (2018) Inferring ice thickness from a glacier dy-
562 namics model and multiple surface data sets. *Environmetrics* 29(5-6):e2460
- 563 Haario H, Saksman E, Tamminen J (2001) An adaptive Metropolis algorithm.
564 *Bernoulli* 7(2):223–242
- 565 Hamilton N (2018) ggtern: An Extension to ggplot2, for the creation of ternary
566 diagrams. URL <https://CRAN.R-project.org/package=ggtern>, r pack-
567 age version 2.2.2

- 568 Hefley TJ, Brost BM, Hooten MB (2017) Bias correction of bounded location
569 errors in presence-only data. *Methods in Ecology and Evolution* 8(11):1566–
570 1573
- 571 Hooten MB, Hobbs N (2015) A guide to Bayesian model selection for ecolo-
572 gists. *Ecological Monographs* 85(1):3–28
- 573 Irwin WP, Wooden JL (1999) Plutons and accretionary episodes of the Kla-
574 math Mountains, California and Oregon. Tech. rep., US Geological Survey
- 575 Jasra A, Stephens DA, Gallagher K, Holmes CC (2006) Bayesian mixture
576 modelling in geochronology via Markov chain Monte Carlo. *Mathematical
577 Geology* 38(3):269–300
- 578 Kimbrough DL, Grove M, Gehrels GE, Dorsey RJ, Howard KA, Lovera O,
579 Aslan A, House PK, Pearthree PA (2015) Detrital zircon U-Pb provenance
580 of the Colorado River: A 5 my record of incision into cover strata overlying
581 the Colorado Plateau and adjacent regions. *Geosphere* 11(6):1719–1748
- 582 Lock EF, Dunson DB (2015) Shared kernel Bayesian screening. *Biometrika*
583 102(4):829–842
- 584 Mason CC, Fildani A, Gerber T, Blum MD, Clark JD, Dykstra M (2017)
585 Climatic and anthropogenic influences on sediment mixing in the Missis-
586 sippi source-to-sink system using detrital zircons: Late Pleistocene to recent.
587 *Earth and Planetary Science Letters* 466:70–79
- 588 Miao L, Qi H (2007) Endmember extraction from highly mixed data using min-
589 imum volume constrained nonnegative matrix factorization. *IEEE Transac-
590 tions on Geoscience and Remote Sensing* 45(3):765–777
- 591 Miller JW, Harrison MT (2018) Mixture models with a prior on the number of
592 components. *Journal of the American Statistical Association* 113(521):340–
593 356
- 594 Puetz SJ, Ganade CE, Zimmermann U, Borchardt G (2018) Statistical anal-
595 yses of global U-Pb database 2017. *Geoscience Frontiers* 9(1):121–145
- 596 Reiners PW, Brandon MT (2006) Using thermochronology to understand oro-
597 genic erosion. *Annual Review Earth Planetary Sciences* 34:419–466
- 598 Richardson S, Green PJ (1997) On Bayesian analysis of mixtures with an
599 unknown number of components (with discussion). *Journal of the Royal
600 Statistical Society: series B (statistical methodology)* 59(4):731–792
- 601 Romans BW, Castelltort S, Covault JA, Fildani A, Walsh J (2016) Environ-
602 mental signal propagation in sedimentary systems across timescales. *Earth-
603 Science Reviews* 153:7–29
- 604 Saylor JE, Sundell K, Sharman G (2019) Characterizing sediment sources by
605 non-negative matrix factorization of detrital geochronological data. *Earth
606 and Planetary Science Letters* 512:46–58
- 607 Sharman GR, Johnstone SA (2017) Sediment unmixing using detrital
608 geochronology. *Earth and Planetary Science Letters* 477:183–194
- 609 Sharman GR, Sylvester Z, Covault JA (2019) Conversion of tectonic and cli-
610 matic forcings into records of sediment supply and provenance. *Scientific
611 reports* 9(1):4115
- 612 Sickmann ZT, Paull CK, Graham SA (2016) Detrital-zircon mixing and parti-
613 tioning in fluvial to deep marine systems, Central California, USA. *Journal*

- of Sedimentary Research 86(11):1298–1307
- Stock GM, Ehlers TA, Farley KA (2006) Where does sediment come from? Quantifying catchment erosion with detrital apatite (U-Th)/He thermochronometry. *Geology* 34(9):725–728
- Sundell K, Saylor JE (2017) Unmixing detrital geochronology age distributions. *Geochemistry, Geophysics, Geosystems*
- Tipton J, Hooten M, Goring S (2017) Reconstruction of spatio-temporal temperature from sparse historical records using robust probabilistic principal component regression. *Advances in Statistical Climatology, Meteorology and Oceanography* 3(1):1–16
- Tipton JR, Hooten MB, Pederson N, Tingley MP, Bishop D (2016) Reconstruction of late Holocene climate based on tree growth and mechanistic hierarchical models. *Environmetrics* 27(1):42–54
- Tipton JR, Hooten MB, Nolan C, Booth RK, McLachlan J (2019) Predicting paleoclimate from compositional data using multivariate Gaussian process inverse prediction. *Annals of Applied Statistics* 13(4):2363–2388, DOI 10.1214/19-AOAS1281
- Tye A, Wolf A, Niemi N (2019) Bayesian population correlation: A probabilistic approach to inferring and comparing population distributions for detrital zircon ages. *Chemical Geology* 518:67–78
- Vermeesch P (2012) On the visualisation of detrital age distributions. *Chemical Geology* 312:190–194
- Weltje GJ (1997) End-member modeling of compositional data: Numerical-statistical algorithms for solving the explicit mixing problem. *Mathematical Geology* 29(4):503–549
- Weltje GJ, Prins MA (2007) Genetically meaningful decomposition of grain-size distributions. *Sedimentary Geology* 202(3):409–424
- Wotzlaw JF, Schaltegger U, Frick DA, Dungan MA, Gerdes A, Günther D (2013) Tracking the evolution of large-volume silicic magma reservoirs from assembly to supereruption. *Geology* 41(8):867–870