Earth and Planetary Science Letters 512 (2019) 46–58

Contents lists available at ScienceDirect

Earth and Planetary Science Letters

www.elsevier.com/locate/epsl

Characterizing sediment sources by non-negative matrix factorization of detrital geochronological data

J.E. Saylor [a,*], K.E. Sundell [b], G.R. Sharman [c]

[a] Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA
[b] Department of Geosciences, University of Arizona, Tucson, AZ, USA
[c] Department of Geosciences, University of Arkansas, Fayetteville, AR, USA

ARTICLE INFO

Article history:
Received 26 December 2018
Accepted 23 January 2019
Available online xxxx
Editor: A. Yin

Keywords:
U–Pb
Book Cliffs
Grand Canyon
Gulf of Mexico
low rank approximation
mixture modeling

ABSTRACT

This paper explores an inverse approach to the problem of characterizing sediment sources' ("source" samples) age distributions based on samples from a particular depocenter ("sink" samples) using non-negative matrix factorization (NMF). It also outlines a method to determine the optimal number of sources to factorize from a set of sink samples (i.e., the optimum factorization rank). We demonstrate the power of this method by generating sink samples as random mixtures of known sources, factorizing them, and recovering the number of known sources, their age distributions, and the weighting functions used to generate the sink samples. Sensitivity testing indicates that similarity between factorized and known sources is positively correlated to 1) the number of sink samples, 2) the dissimilarity among sink samples, and 3) sink sample size. Specifically, the algorithm yields consistent, close similarity between factorized and known sources when the number of sink samples is more than ~3 times the number of source samples, sink data sets are internally dissimilar (cross-correlation coefficient range >0.3, Kuiper *V* value range >0.35), and sink samples are well-characterized (>150–225 data points). However, similarity between known and factorized sources can be maintained while decreasing some of these variables if other variables are increased.

Factorization of three empirical detrital zircon U–Pb data sets from the Book Cliffs, the Grand Canyon, and the Gulf of Mexico yields plausible source age distributions and weights. Factorization of the Book Cliffs data set yields five sources very similar to those recently independently proposed as the primary sources for Book Cliffs strata; confirming the utility of the NMF approach. The Grand Canyon data set exemplifies two general considerations when applying the NMF algorithm. First, although the NMF algorithm is able to identify source age distribution, additional geological details are required to discriminate between primary or recycled sources. Second, the NMF algorithm will identify the most basic elements of the mixed sink samples and so may subdivide sources that are themselves heterogeneous mixtures of more basic elements into those basic elements. Finally, application to a large Gulf of Mexico data set highlights the increased contribution from Appalachian sources during Cretaceous and Holocene time, potentially attributable to drainage reorganization. Although the algorithm reproduces known sources and yields reasonable sources for empirical data sets, inversions are inherently non-unique. Consequently, the results of NMF and their interpretations should be evaluated in light of independent geological evidence. The NMF algorithm is provided both as MATLAB code and a stand-alone graphical user interface for Windows and macOS (.exe and .app) along with all data sets discussed in this contribution.

1. Introduction

One of the primary goals of detrital geochronology research is to identify sediment provenance and characterize sediment sources. The discipline is undergoing a revolution as quantitative methods are increasingly brought to bear to compare and interpret data sets. These methods have been used primarily in a descriptive sense to characterize the degree of similarity among detrital age distributions (e.g., DeGraaff-Surpless et al., 2003; Satkoski et al., 2013; Vermeesch, 2013; Horton et al., 2015; Kimbrough et al., 2015; Vermeesch et al., 2016; Andersen et al., 2018). Parallel research has focused on quantifying the contributions of well-characterized sediment sources ("source" samples, a.k.a "parent" samples) to equally well-characterized samples from a particular depocenter ("sink" samples, a.k.a. "daughter" samples)

* Corresponding author.
E-mail address: jesaylor@uh.edu (J.E. Saylor).

https://doi.org/10.1016/j.epsl.2019.01.044
0012-821X/

(Amidon et al., 2005a, 2005b; Saylor et al., 2013; Kimbrough et al., 2015; Licht et al., 2016; Sickmann et al., 2016; Capaldi et al., 2017; Mason et al., 2017; Sharman and Johnstone, 2017; Sundell and Saylor, 2017). Despite recent advances in quantitative analysis of detrital geochronology data, most current methods either cannot characterize source samples' age distribution, or require *a priori* knowledge of the sources in order to determine their relative contributions to sink samples.

In the geologic record, sediment sources are typically poorly preserved, whereas sediment sinks, by comparison, are typically well-preserved and can be easily characterized. Few studies have taken advantage of this feature of the geologic record to characterize unknown source samples based on known sink samples (Sharman and Johnstone, 2017). However, several critical components of the methods discussed by Sharman and Johnstone (2017) warrant further exploration to promote their widespread implementation. In order for factorized sources to accurately reflect known sources there is a trade-off between the number of sink samples, the number of analyses per sink sample, and the dissimilarity among sink samples. However, the relationship between these variables, and their combined effect on the success of the factorization has not been explored. More importantly, there is currently no method for determining the optimal number of source samples to be characterized based on a particular sink data set.

In this contribution we explore these questions using a modification of the non-negative matrix factorization (NMF) inverse unmixing approach discussed by Sharman and Johnstone (2017), which has been successfully applied to other geological questions (e.g., Paterson and Heslop, 2015). This method requires no prior knowledge of the sources, no supervision, and no training. Using synthetic data sets we explore the characteristics of sink data sets that yield a close correlation between factorized and known sources. We also develop a method of identifying the optimal number of sources that can be determined given a particular set of sink data. These methods are then applied to empirical data sets from the Book Cliffs, Grand Canyon, and circum-Gulf of Mexico. Algorithms are provided both as MATLAB code and stand-alone graphical user interfaces (GUIs) as an executable (.exe) file for Windows and application (.app) file for macOS (Supplemental Files 1).

2. Methods

2.1. Mixture distributions

The unmixing algorithm discussed below takes mixture distributions as its input. These can be kernel density estimates (KDEs), probability density plots (PDPs), or kernel functional estimates (KFE) (Jessberger et al., 1980; Hurford et al., 1984; Dodson et al., 1988; Brandon, 1996; Sircombe and Hazelton, 2004; Gehrels, 2012; Vermeesch, 2012). Details of how KDEs and PDPs are calculated is presented in the Supplemental Text.

2.2. Algorithm

We adopt the NMF approach widely used in signal processing and image analysis (Lee and Seung, 1999; Smaragdis and Brown, 2003; Ozerov and Fevotte, 2010) and recently applied to detrital geochronology (Sharman and Johnstone, 2017). Given a matrix V (composed of m features from n samples), NMF seeks to develop two matrices W (m-by-k) and H (k-by-n) such that

$$V \approx WH \tag{1}$$

(using matrix multiplication notation). In detail, however,

$$V = WH + E, \tag{2}$$

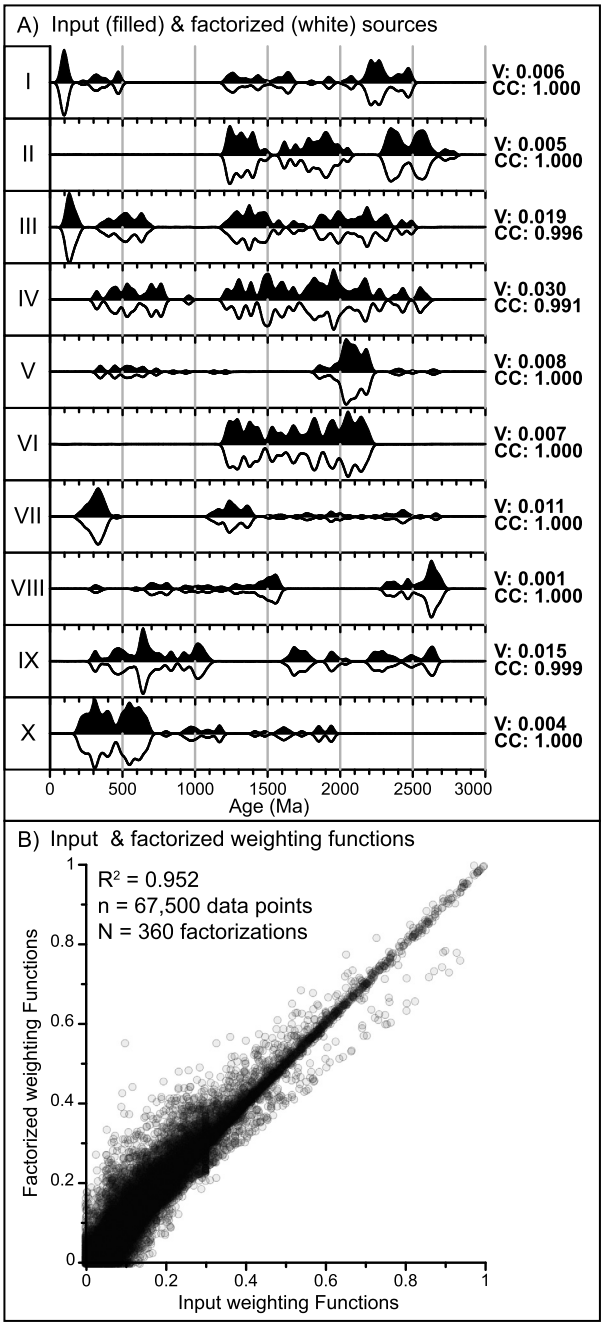

Fig. 1. Factorization of KDEs of sink samples produced from known sources (black), successfully reproduces both the age distributions and weighting functions used to create the sink samples. A) Input (black) sources and factorized (white) sources from data set 10S presented as KDEs using a Gaussian kernel with a 20 Myr bandwidth. Factorized source samples based on 40 sink samples randomly mixed from the ten input source samples. B) Input and factorized weighting functions use to create or factorize the 40 sink samples. Data includes all trials of 10S with $\geq$3.5 times as many sink samples as sources (i.e., $N \geq 35$). CC and V respectively indicate the Cross-correlation coefficient and Kuiper V value between the known source and the subjacent factorized source. Data for this figure are in Supplemental Table 1.

where E is the final residual (calculated as the matrix norm of $V - WH$). In terms of detrital geochronology, V is composed of m probabilities at equal spacing along the mixture distribution (PDP, KDE, or KFE) for each of n samples. W then is composed of m probabilities at the same spacing for each of k factorized components. Finally, H is composed of the k weighting functions for each of the factorized components in the n samples. W and H are estimated by iterative update while seeking to minimize the er-

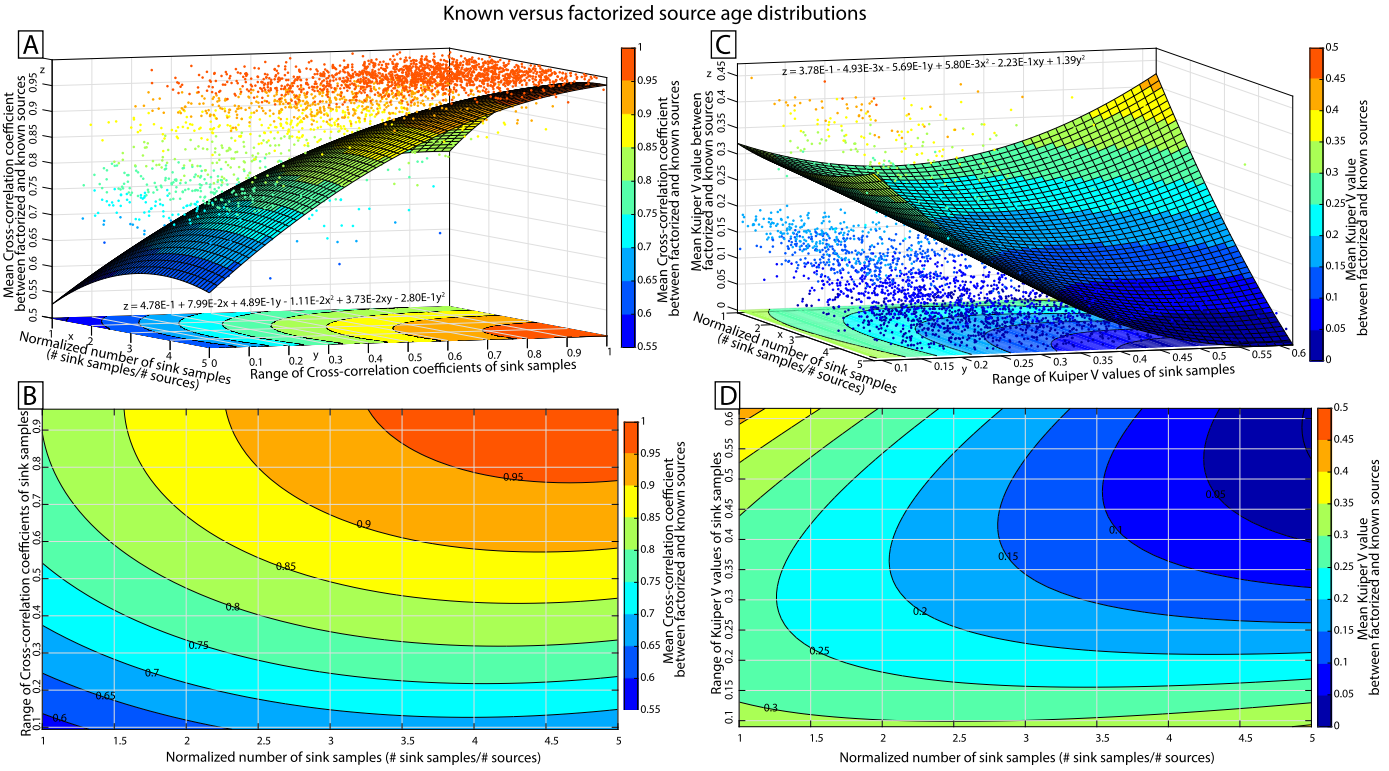

Fig. 2. The similarity between known and factorized sources (z) increases with increasing size (x) and internal dissimilarity (y) of the sink data set. A) Scatter plot of the mean Cross-correlation coefficient between factorized and known sources as a function of number of sink samples and dissimilarity among sink samples. The number of sink samples is normalized to the number of known sources to facilitate comparison to other figures. The internal dissimilarity of the sink data set is shown as the range of Cross-correlation coefficients. The surface is a second order polynomial fit in which ~95% of the data points fall above the surface (i.e., better fit). B) Contour plot of the surface in A. C) As with A, but using the Kuiper V value rather that Cross-correlation. The surface is a second order polynomial fit in which ~95% of the data points fall below the surface (i.e., better fit). D) Contour plot of the surface in C. Data sets 5S, 6S, and 10S are included in this analysis and data are compiled in Supplemental Table 21. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

ror in the approximation $V \approx WH$ (i.e., minimize E). The distance between WH and V can be estimated using multiple functions including, for example, a Euclidean distance

$$d_{Euc}(W, H) = \frac{1}{2}\|V - WH\|^2 \tag{3}$$

which is subject to the constraints that $W \geq 0$ and $H \geq 0$. Following the alternating non-negative least squares (NNLS) example of Kim and Park (2008), this problem can be iteratively solved starting from random initial estimates for W and H and alternatively updating W and H until a stopping criterion is satisfied while minimizing the distance between V and WH. Alternatively solving for W and H guarantees that the solution converges to a stationary point.

In the following examples and the associated software application, we adapted the NNLS algorithm of Li and Ngom (2013) and Van Benthem and Keenan (2004) for geological application. We selected this algorithm because the NNLS optimization that it implements requires fewer input samples than some other optimization routines (Kim and Park, 2007, 2008; Li and Ngom, 2013). Adaptations include normalizing the weighting functions so that the sum of the weighting functions for each group of samples is one, and normalizing the factorized mixture distributions so that the integral of each sample mixture distribution is one.

2.3. Comparison metrics

We use two statistical measures to compare the factorized sources (W) to the initial sources and the reconstructed sink samples (WH) to input sink samples (V): the Cross-correlation coefficient and the Kuiper V value. Cross-correlation is a widely used method in signal processing (Lewis, 1995), template matching (Briechle and Hanebeck, 2001), image matching (Zhao et al., 2006; Pan et al., 2009), and geophysics (Troyan, 2010; Debella-Gilo and Kääb, 2011) and is essentially the equivalent of the normalized cross-correlation function with a zero lag. It was initially applied to detrital geochronology by Saylor et al. (2012) and is calculated as,

$$\frac{\sum_a^b (h_i - h)(j_i - \bar{j})}{\sqrt{\sum_a^b (h_i - \bar{h})^2}\sqrt{\sum_a^b (j_i - \bar{j})^2}} \tag{4}$$

where h and j are the mixture distributions and $[a, b]$ is the age interval under consideration. We include it here because it is sensitive to differences between samples (Saylor and Sundell, 2016). The Kuiper test (Kuiper, 1960; Press et al., 2007) is a modification of the Kolmogorov–Smirnov (KS) test. Unlike the KS test, the Kuiper test is equally sensitive across the range $[a, b]$ and so is preferable to the KS test in detrital geochronology applications (e.g., Wissink et al., 2018). The Kuiper V value is calculated as,

$$V(x) = \max_{-\infty<\infty}\left[F_1(x) - F_2(x)\right] + \max_{-\infty<\infty}\left[F_2(x) - F_1(x)\right] \tag{5}$$

where F_1 and F_2 are the empirical cumulative distribution functions (ECDFs) of sample 1 and 2, respectively.

2.4. Inputs

In our tests we consider five synthetic data sets where the source samples are known. The synthetic data sets are based on mixtures of empirical data published by Laskowski et al. (2013)

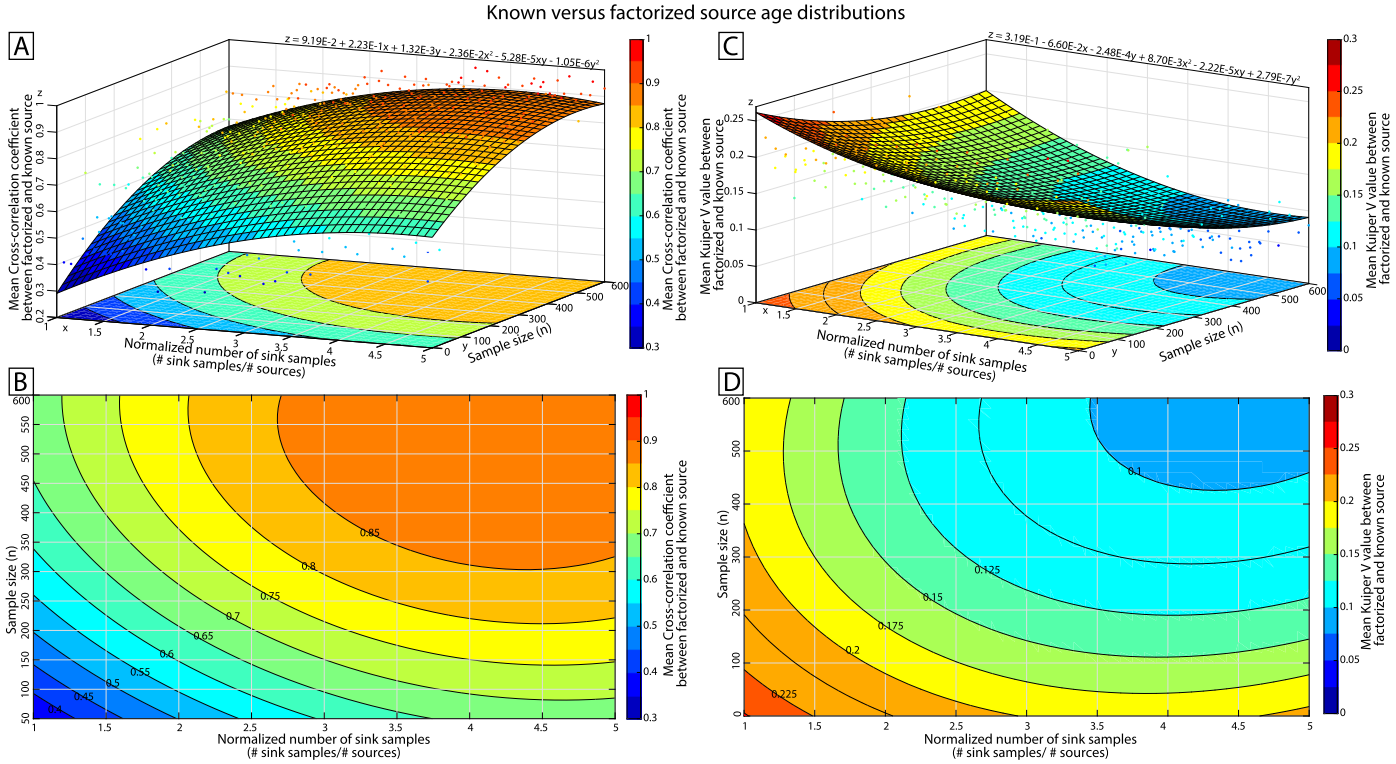

Fig. 3. The similarity between known and factorized sources (z) increases with increasing number of sink samples (x) and increasing sink sample size (y). A) Scatter plot of the mean Cross-correlation coefficient between factorized and known sources as a function of the number of sink samples and sink sample size. The number of sink samples is normalized to the number of known sources to facilitate comparison to other figures. The surface is a second order polynomial fit in which ~85% of the data points fall above the surface (i.e., better fit). B) Contour plot of the surface in A. C) As with A, but using the Kuiper V value rather that Cross-correlation. The surface is a second order polynomial fit in which ~85% of the data points fall below the surface (i.e., better fit). D) Contour plot of the surface in C. Data set 8S is included in this analysis and data are compiled in Supplemental Table 22. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

and Pullen et al. (2014), and synthetic data sets published by Saylor and Sundell (2016), and Sundell and Saylor (2017). For data set 5S, we divided up an existing data set from Pullen et al. (2014) into five potential sources (Fig. S1). For data set 6S we randomly subsampled 500 ages from each of the six of the characteristic North American provenance groups identified by Laskowski et al. (2013) including the Mesozoic eolianites, U.S. passive margin, Canadian passive margin, Mogollon highlands, Cordilleran arc, and Yavapai–Mazatzal provenance groups (Fig. S2). For data set 10S we used the "complex synthetic" data set of Sundell and Saylor (2017, their Fig. 4), comprised of 10 source samples with multimodal, overlapping age populations (black curves in Fig. 1A). In the latter data set, each of the 10 samples has 100 ages with uncertainty between 2 and 12% at the 1σ level. For data set 8S, we used the following samples from Saylor and Sundell (2016): 5 Peaks, 11 Peaks, 22 Peaks 2, 22 Peaks 4, 22 Peaks bimodal, 22 Peaks central, 33 Peaks, and 44 Peaks. We used the same data for 12S but also added the following samples: 4 Peaks, 25 Peaks, 30 Peaks, and 49 Peaks. All data sets can be found in the Supplemental Material (Supplemental Table 1).

We also consider three data sets for which the source samples are unknown. The first of these includes 24 samples from the Book Cliffs data set published by Bartschi et al. (2018). The second is 25 samples from the Grand Canyon published by Gehrels et al. (2011). The final data set is a compilation of 111 Late Cretaceous–modern samples from the circum-Gulf of Mexico compiled from Iizuka et al. (2005), Craddock and Kylander-Clark (2013), Blum and Pecha (2014), Wahl et al. (2016), Mason et al. (2017), Blum et al. (2017), and Xu et al. (2017) (Supplemental Table 2).

3. Model testing-methods

3.1. Model efficiency: number, dissimilarity, and size of sink data sets

The goals of Test 1 and 2 were to determine how the number of sink samples and dissimilarity between sink samples, respectively, impacts the similarity between factorized and known sources. Tests 1 and 2 use the same data sets: 5S, 6S, and 10S. In each case, sources were assigned random weightings to each source to produce up to 25–40 sink samples (Supplemental Tables 3–20). We selected weighting functions (H from Equation (1)) using both the randfixedsum algorithm (Stafford, 2006) and the randomization algorithm developed by Sundell and Saylor (2017). For this test the number of factors (i.e., sources) was always equal to the known number of sources. After running the NMF algorithm we calculated Cross-correlation and V values to compare the factorized sources to the known sources. For each trial, we calculated the mean and standard deviation of these metrics to assess the quality of the factorization. This resulted in a total of 3,050 independent tests of the NMF approach (Fig. 2, Supplemental Table 21).

In Test 3 we evaluated the effect of sink sample size on the similarity between factorized and known sources using data set 8S. From this large data set ($n = 10^6$ per sample) we drew between 25 and 1,000 ages and associated uncertainties per sample and produced between 8 and 40 sink samples using the randomization algorithm of Sundell and Saylor (2017). These were then factorized into eight sources which were compared to the eight known sources (Fig. 3, Supplemental Table 22).

We evaluated the effect of the number of sink data sets, sink data set size, and internal dissimilarity, by fitting a 3-dimensional surface to the factorized sources that are most dissimilar from the known sources. This surface encompasses the majority of the data,

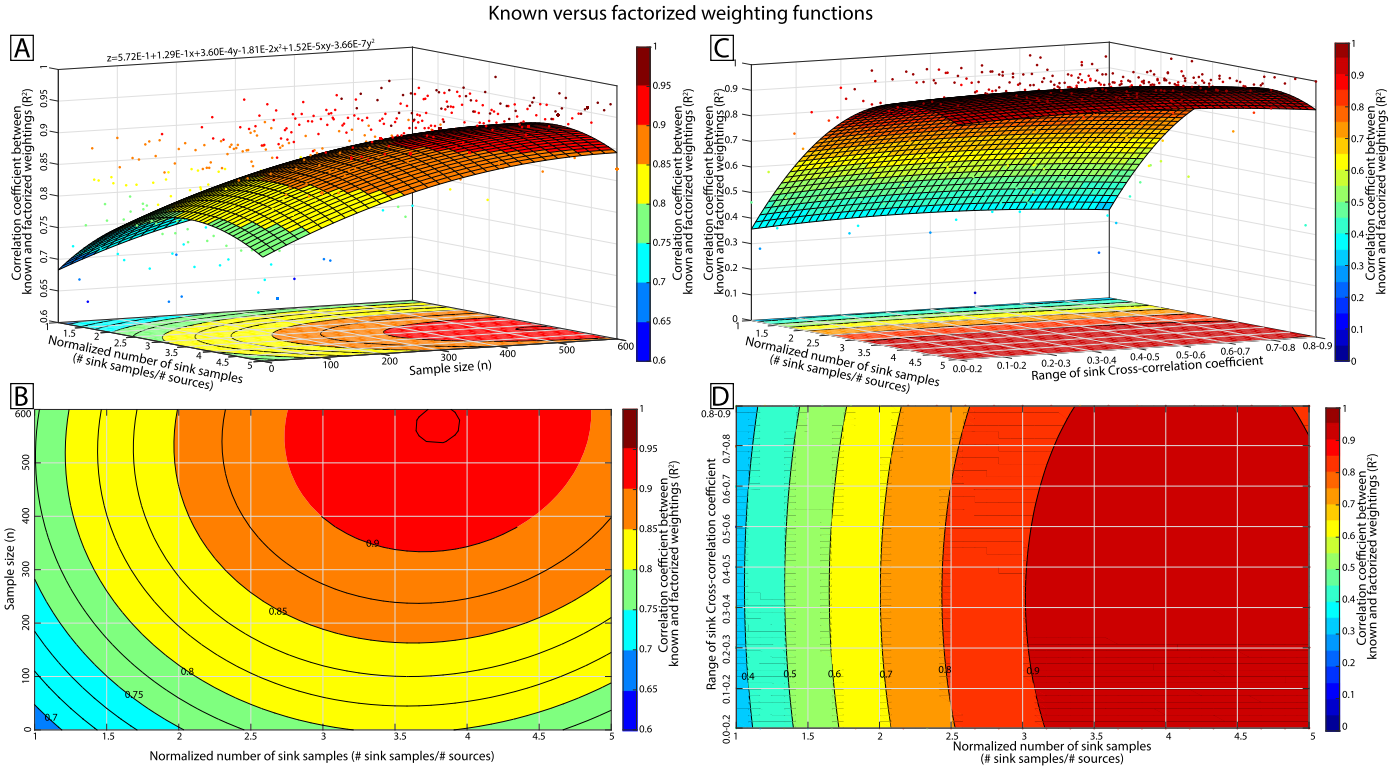

Fig. 4. Correlation (R^2) between known and factorized weightings increases with increasing number of sink samples but is not significantly affected by the sink sample dissimilarity. A) Scatter plot of the mean correlation coefficient between factorized and known weights as a function of the number of sink samples and sink sample size. The surface is a second order polynomial fit in which ~85% of the data points fall above the surface (i.e., better fit). B) Contour plot of the surface in A. C) Scatter plot of the mean correlation coefficient between factorized and known weights as a function of the number of sink sample size and dissimilarity among sink samples. D) Contour plot of the surface in C. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

and we treat it as an envelope of the worst case (i.e., most dissimilar) fits between known and factorized sources. In order to compare results from 5S, 6S, and 10S we normalized the number of sink samples to the number of known source samples for each data set. We selected the worst fits by considering matrices of results for discrete normalized numbers of sink samples (i.e., matrices of y and z values for each x interval in Fig. 2). For Tests 1 and 2, matrices were comprised of mean Cross-correlation coefficients or V values (z in Fig. 2) sorted by the range of coefficients or values of the input sink samples (y in Fig. 2). For Test 3 matrices were comprised of mean Cross-correlation coefficients or V values sorted by the size of the sink samples. From each matrix we subsampled the data by selecting the local extrema in mean Cross-correlation coefficient or V value between factorized and known sources (i.e., local minima in z from Fig. 2) using the peakfind MATLAB algorithm. For Tests 1 and 2 this yielded 277 and 258 data points of the original 3,050 trials for Cross-correlation and Kuiper, respectively (Supplemental Table 21). We fit a second-order 3-dimensional surface to these data (Fig. 2). This surface encompasses ~95% of the data points, and, as noted above, we treat it as an envelope of the worst case fits. For Test 3 this yielded 126 and 117 of the original 462 trials for Cross-correlation and Kuiper, respectively. Due to the lower data density, compared to Tests 1 and 2, the surface for Test 3 encompasses ~85% of the data points. For comparison of input to factorized weighting, we fit a polynomial curve to 115 of 3,456 total data points (Fig. 4A) or 112 of 417 data points (Fig. 4B).

3.2. Optimal number of factorized sources

In Test 4 we developed a method of determining the optimal number of source samples for a given sink data set. Selection of an appropriate number of source samples (i.e., the rank of the factorization) critically affects both the success and potential accuracy of the factorization. We approached this problem by calculating the final residual between input and reconstructed sink samples over a wide range of ranks. Because NMF is intended to provide a low-rank approximation of the input data sets, our criterion is that the optimal number of source samples is the rank above which there is little or no decrease in final residual. This should yield two approximately linear segments: a lower rank segment over which there is rapid decrease in the residual and a higher rank segment over which there is little decrease in residual (Fig. 5A, Supplemental Table 25). The point at which these two segments meet is the optimal rank. This approach is similar to the more familiar "scree" plots applied to determine the optimum number of dimensions in multidimensional scaling (Borg and Groenen, 1997; Hair et al., 1998; Steyvers, 2006). In order to identify the optimum breakpoint we plotted the final residual versus the rank and applied a segmented linear regression with one break. We assumed that the optimal breakpoint, and therefore the optimal number of sources, will be the point which optimizes the linear fit of both the higher and lower rank segments, and calculated the sum of squared residuals (SSR, Draper and Smith, 1998) for each segment as

$$SSR_1 = \sum_{r=2}^{r=x_b} \left(R_r - f(x_r)\right)^2 \tag{6}$$

and

$$SSR_2 = \sum_{r=x_b}^{r=n} \left(R_r - \left(g(x_r)\right)^2. \tag{7}$$

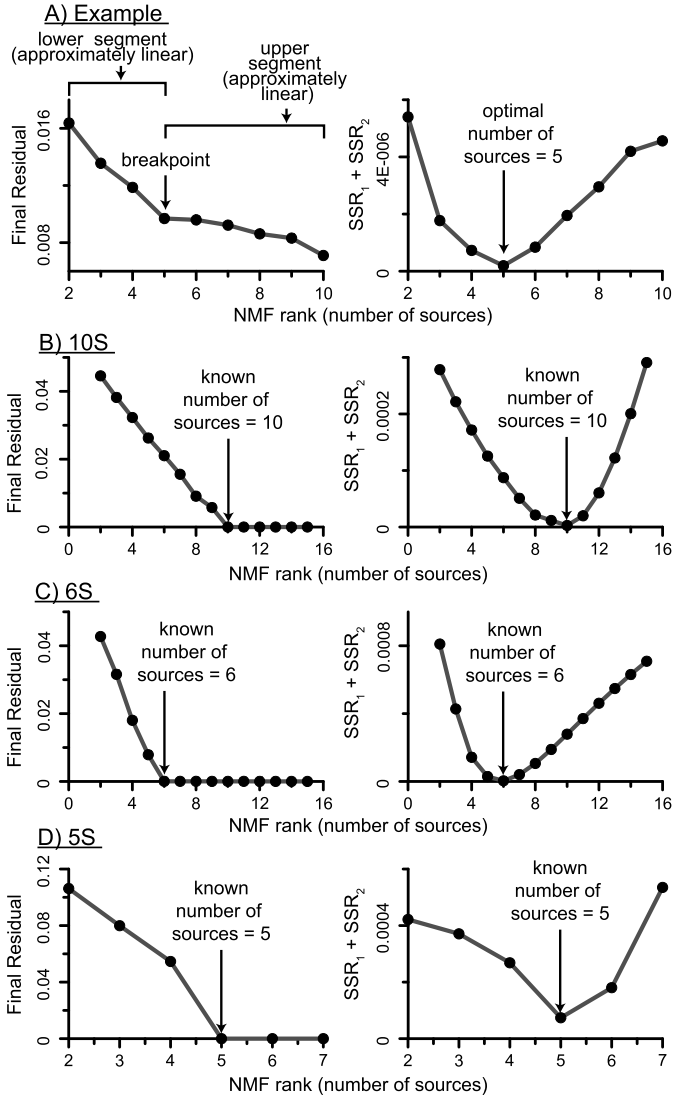

Fig. 5. Segmented linear regression (left) and summed sum of squared residuals (right) for A) a hypothetical sample, B) 10S, C) 6S, and D) 5S demonstrates that the optimum breakpoint coincides with the known number of source samples. On the left in each panel, the NMF final residual versus rank (number of sources) shows two approximately linear segments separated by a clear breakpoint. On the right in each panel, the plot of the summed sum of the squared residuals for the two segments versus the rank for the breakpoint shows a minimum at the optimum number of source samples for each data set. Data for this figure are in Supplemental Table 25.

In these equations x_b is the breakpoint, which is calculated for all integers between 2 and n. n is the maximum rank being considered. R_r is the final residual for each rank (r). $f(x_r)$ and $g(x_r)$ are the expected final residual calculated by linear regression over the lower rank or higher rank segments, respectively. For $x_b = 2$ and $x_b = n$, $f(x_b)$ and $g(x_b)$ are equal to the final residual yielding SSR_1 and $SSR_2 = 0$, respectively. We identified the optimal rank as the x_b which minimizes the sum $SSR_1 + SSR_2$ (Fig. 5A).

We first applied this test to the four data sets with known numbers of sources and then to data sets with an unknown number of sources. We initially considered the ability of this method to reconstruct the known number of sources when the mixture distributions are fully characterized (i.e., zero uncertainty in either source or sink mixture distributions) using data sets 5S, 6S, and 10S. We then considered the effect uncertainty in the sink samples on the ability to reconstruct a known number of sources by varying sink sample sizes using data set 12S. We randomly selected between 2 and 10 sources from the 12 potential sources in the data set. From these sources we selected between 50 and 5,000 ages from the 10^6 ages. From these ages, we constructed between 2 and 24 sink samples. We determined the optimum number of sources to factorize for each of these trials as outlined above and compared this to the known number of sources.

4. Model testing-results

The NMF method yields close matches between known and factorized source distributions (Fig. 1A, also Supplemental Figs. S1A, S2A, and S3A) and weighting functions (Fig. 1B, also Supplemental Figs. S1B, S2B, and S3B) when applied to finite mixture distributions (i.e., PDPs or KDEs) of detrital geochronology. Our experiments implementing the NMF algorithm on ECDFs were unsuccessful, as the resulting distributions were not cumulative with age. Therefore in the following section we focus on the results of factorization of finite mixture distributions.

4.1. Model efficiency: number, dissimilarity, and size of sink data sets

Our first test yielded mean Cross-correlation coefficients between known and factorized sources >0.75 and Kuiper V values <0.15 when the number of sink samples is more than 3 times the number of sources (Fig. 2B and 2D) except where sink samples were very similar. Some trials yielded Cross-correlation coefficients >0.75 and V values <0.15 with fewer sink samples but results were inconsistent (Fig. 2A and 2C). Factorized weights also become better correlated to known weights with increasing numbers of sink samples (Fig. 4B and 4D). The NMF algorithm yielded a close match between factorized and known source samples regardless of the elements constrained or the randomization algorithm used (Supplemental Tables 3–20). However, when sink samples were very similar to one another, as indicated by a low range of Cross-correlation coefficients or Kuiper V values between sink samples, the similarity between factorized and known sources was low (Fig. 2). We therefore investigated the relationship between the similarity of the input sink samples and the similarity between factorized and known sources in Test 2.

The second test indicates that similarity between the sink samples is a second control on the ability to successfully factorize known sources, but not a significant control on the correlation between known and factorized weightings. Trials with a range of Cross-correlation coefficients >∼0.3 yielded mean Cross-correlation coefficients between known sources and factorized sources >0.75 (Fig. 2). Similarly, trials with a range of V values >∼0.35 yielded mean V values between factorized and known sources <0.15. In contrast, a range of Cross-correlation coefficients <0.3, or a range of V values <0.35 resulted in a wide range of similarity between known and factorized sources. Correlation between known and factorized weightings are independent of sink sample dissimilarity and depend only on the number of sink samples (Fig. 4C and 4D).

The results of Test 3 indicate that sample size plays an additional role in the ability to reliably factorize sources and weightings. In addition to ∼3 times the number of sink samples as sources to be factorized, sink samples need to have >∼225 analyses per sample to yield Cross-correlation coefficients >0.75 between known sources and factorized sources (Fig. 3). This criterion can be relaxed if either a lower Cross-correlation coefficient is acceptable or more sink samples are available. For example, a Cross-correlation coefficient of ∼0.7 can be achieved with 2.5 times the number of sink samples as sources and sink samples that are ∼200 analyses each. However, the same Cross-correlation coefficient can be achieved with 4 times as many sink samples as sources even if those sink samples have <100 analyses per sample (Fig. 3). Test

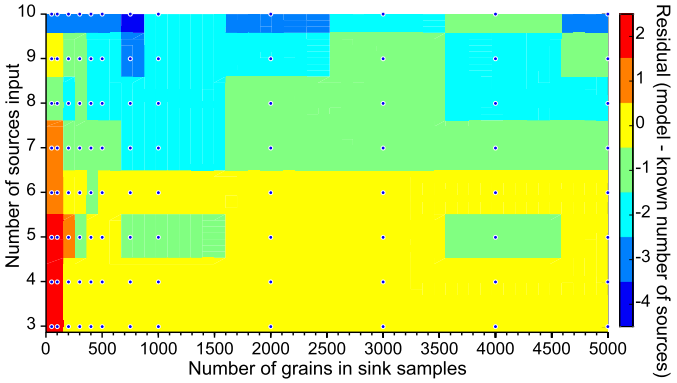

Fig. 6. Increasing the size of each sink sample results in a better fit between the reconstructed number of sources and the known number of sources (i.e., residual closer to zero). Note that the linear segmented regression approach appears to systematically underestimate the number of sources for more than six sources. Data for this figure are in Supplemental Table 26. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

3 also yields a positive correlation between sink sample size and the correlation coefficient between known and factorized weighting functions (Fig. 4A and 4B).

4.2. Optimal number of sources

Results of Test 4 using synthetic data sets (10S, 6S, and 5S) indicate that the minimum in $SSR_1 + SSR_2$ coincides with the known number of sources (Fig. 5, Supplemental Table 25). As expected, when plotting the final residual against the NMF rank we see two linear segments which meet at the known number of sources.

Selecting a finite number of ages to construct sink age distributions introduces additional uncertainty into the ability to identify the optimal number of sources (Fig. 6, Supplemental Table 26). Reconstructing the number of sources from small sink samples (<200 analyses per sample) resulted in residuals (model – known number of sources) of up to two. Larger sample sizes yielded a better fit (residual ≈0) for between three and six sources. Above six sources, the technique appears to systematically underestimate the number of sources, regardless of the sink sample size.

When applying this test to the empirical data sets we found that although the two linear segments in the final residual plots are not as clear as with the control data sets, there were still clear minima in the $SSR_1 + SSR_2$ plots (Fig. 7). These minima indicate that the optimal number of samples for the Book Cliffs data set is five (Fig. 7A). The Book Cliffs data set has a range of Cross-correlation coefficients of 0.97 and a range of Kuiper V values of 0.50. Using the Grand Canyon data set, the lowest $SSR_1 + SSR_2$ values occur at a rank of six (Fig. 7B). The Grand Canyon data set has a slightly lower range of Cross-correlation coefficients of 0.96 and a range of V values of 0.71. The range of Cross-correlation coefficients for the Gulf of Mexico sink data set is 0.81 and the range of V values is 0.85. Finally, the Gulf of Mexico data set has a nadir in $SSR_1 + SSR_2$ at a rank of six (Fig. 7C).

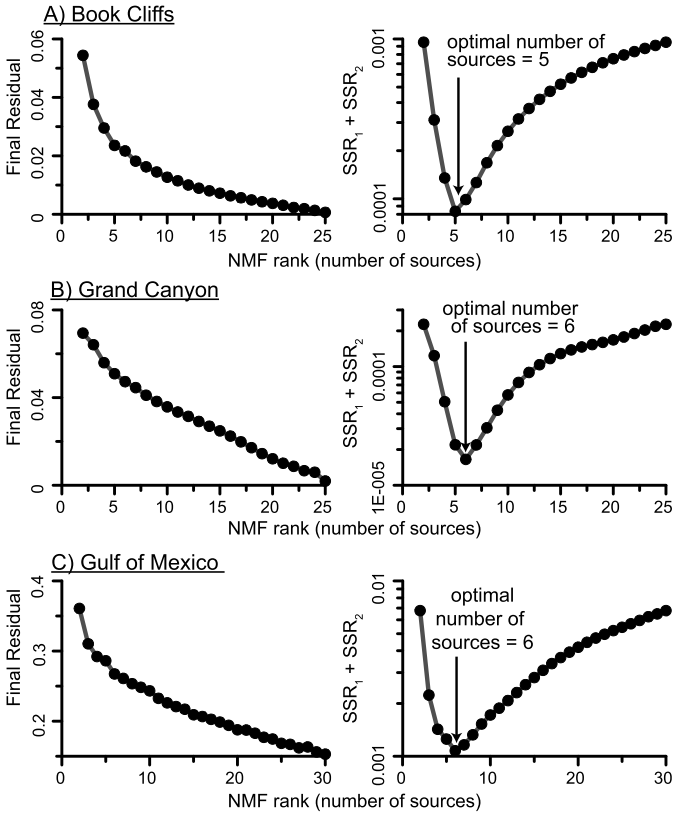

Fig. 7. Segmented linear regression (left) and summed sum of squared residuals (right) for the three empirical data sets (Book Cliffs, Grand Canyon, and Gulf of Mexico) indicate that the optimum number of sources that can be modeled from these data sets is five, six, and six, respectively. Data for this figure are in Supplemental Table 27.

5. Discussion

5.1. Sensitivity testing using synthetic data sets

Based on our analysis of samples mixed from known sources (Fig. 2 and Fig. 3) we conclude that the three factors which contribute to a successful decomposition of sink distributions into a low rank approximation of the source samples are 1) the number of sink samples considered, 2) the degree of similarity between the sink samples, and 3) the size of the sink samples. The synthetic data sets yield close matches between known and factorized source age distributions when the size (N) of the input sink data set is more than 3 times greater than the number of sources. With fewer sink samples, results are inconsistent (e.g. see scatter in Fig. 2) and we therefore caution against use of small N sink data sets. The conclusion above holds for all data sets except those that are very similar (i.e., range of Cross-correlation coefficients <0.3 or V values <0.35) (Fig. 2). This latter observation indicates that the dissimilarity between the input sink samples is an important control on the similarity between factorized and known source age distributions. The relationship shown in Fig. 2 indicates that a range of Cross-correlation coefficients >0.3 or V value >0.35 is needed to ensure a mean Cross-correlation coefficient >0.75 or mean V value <0.15 between known and factorized sources. In contrast, a range of Cross-correlation coefficients <0.3, or a range of V values <0.35 yields a wide range of similarity between known and factorized sources, meaning that one cannot have confidence, *a piori*, in the ability to factorize sources based on a data set with such similar age distributions. However, sink data set similarity appears to be less of a control on the similarity between known and factorized weighting functions (Fig. 4). Sample size (n) appears to be an important third control on the ability to achieve close correlation between known and factorized source samples. Our analysis indicates that ideally the sink data set should have at least 3 times the number of sources, and each sample should be well characterized (n > 150–225). However, there is a trade-off between the number of sink samples and the sink sample size. Comparable similarity between factorized and known sources can be achieved by increasing the number of sink samples or sink sample size, even if the other variable decreases slightly (Fig. 3). Although the specific values determined above are dependent on the comparison metrics employed, the similarity in

behavior suggests that the three factors that contribute to a successful factorization are fundamental behaviors of NMF as applied to detrital geochronology.

5.2. Selection of optimum rank

Comparison of the factorized and known number of sources indicates that the segmented linear regression method (Test 4) is able to reconstruct a known number of sources (Fig. 5). However, reconstruction is degraded by uncertainty in the sink samples (Fig. 6). Although the method involves some uncertainty, we recommend it as a guide, as it provides a first order sense of how many sources can reasonably be factorized from a given data set. This approach also maximizes the user's ability to interpret the results of segmented linear regression in light of independent geological information.

5.3. Application

In the section below we compare the factorized source age distributions for the Book Cliffs, Grand Canyon, and Gulf of Mexico to empirical potential sediment sources for each of these basins. The goal is not to definitively identify sediment sources for any of these regions; such an undertaking would require detailed consideration of the geologic context including at a minimum the depositional environments, tectonic setting, paleocurrent measurements, and paleogeographic linkages between the proposed sources and sinks. Rather, our goal is to show that the factorized sediment sources are consistent with existing potential sediment source data. This in itself is a remarkable accomplishment, given the fact that the algorithm is completely unsupervised and has no training data sets.

The success in factorizing the Book Cliffs data set into five sources provides an important intermediate step between completely controlled and completely unknown situations, indicating that the algorithm factorized this data set into geologically meaningful factors. Previous work by Bartschi et al. (2018) used six provenance groups to model the data from the Book Cliffs using DZmix (Sundell and Saylor, 2017). These included the Mogollon highlands, Magmatic arc, Mesozoic eolianites, US Paleozoic passive margin, Canadian Paleozoic passive margin, and Late Cretaceous foreland basin composite (TSA3) from Laskowski et al. (2013) and May et al. (2013). However, Bartschi et al. (2018) found that there was consistently little input from the Canadian Paleozoic passive margin provenance group. Our findings confirm this result, indicating that only five provenance groups are needed to successfully account for the majority of the variability in the Book Cliffs data set (Fig. 7A). Our results further confirm that the provenance group that is not included in the factorization of this data set is the Canadian Paleozoic passive margin (Fig. 8A).

Although several Book Cliffs factorized sources have clear correlatives in known provenance groups, others have non-unique potential sources or are best matched by subsets of the provenance groups proposed by Laskowski et al. (2013). For example, the first, second, and third factorized sources have clear correlatives in the Late Cretaceous foreland basin composite, Mesozoic eolianites, US Paleozoic passive margin provenance groups, respectively (Fig. 8B, C, D). On the other hand, the fourth factorized source is poorly correlated to the full Mogollon highlands sources (Fig. 8Eii), but matches well to either a US Paleozoic passive margin source (Fig. 8Eiii), or to a subset of the Mogollon highland source that does not include the <250 Ma ages (Fig. 8Ei). Similarly, the fifth source is poorly correlated to the full southern US Magmatic arc source (Fig. 8Fii), but is well correlated to the <250 Ma component from the Mogollon highlands (Fig. 8Fi). However, the division of the Mogollon highlands source into <250 Ma

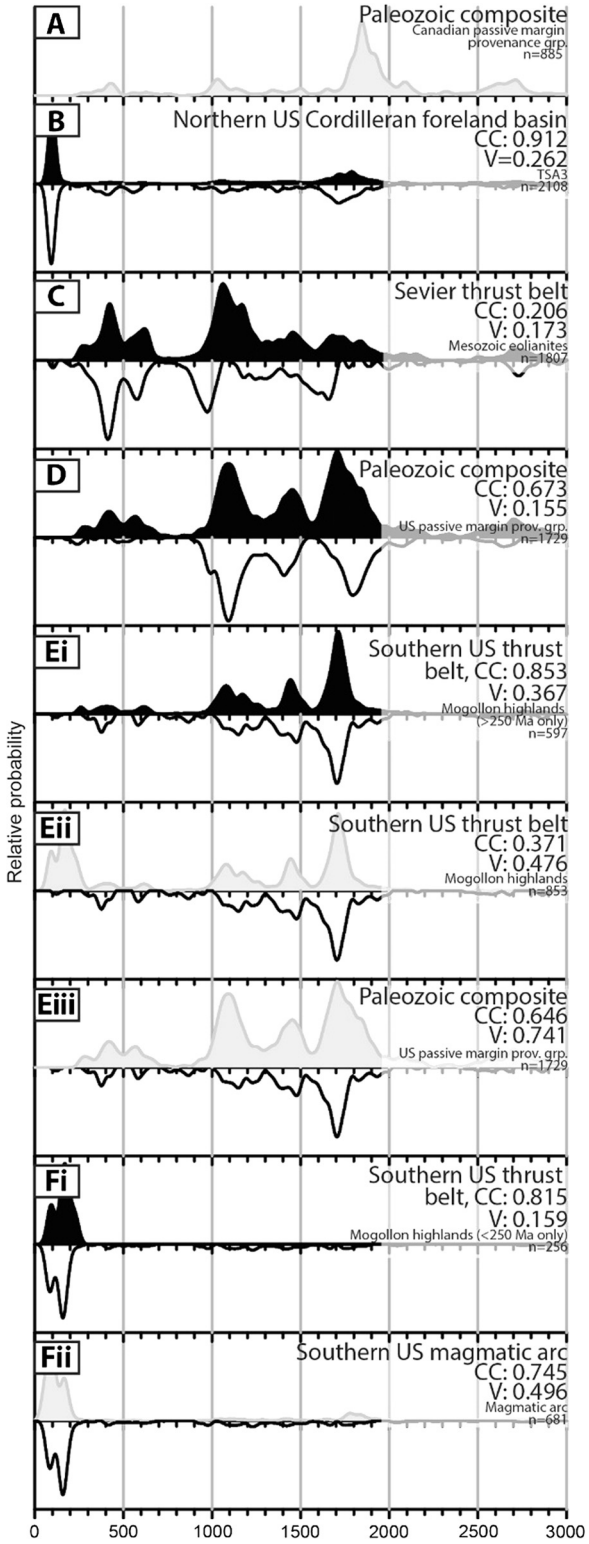

Fig. 8. Comparison of potential sediment sources (black or gray filled KDEs) to factorized sources (inverted unfilled KDEs) for the Book Cliffs data set of Bartschi et al. (2018). Black-filled PDPs are our preferred matches to factorized sources. Potential sources are all from Laskowski et al. (2013, Canadian passive margin, Mesozoic eolianites, US passive margin, Mogollon highlands, and Magmatic arc provenance groups) and May et al. (2013, TSA3 Late Cretaceous composite). Note that, consistent with the findings of Bartschi et al. (2018) the Canadian Paleozoic Passive Margin source is not included in the factorization of the Book Cliffs data set. CC and *V* respectively indicate the Cross-correlation coefficient and Kuiper *V* value between the known source and the subjacent factorized source. KDEs are constructed with a Gaussian kernel with a bandwidth of 20 Myr. Data for this figure are in Supplemental Table 28.

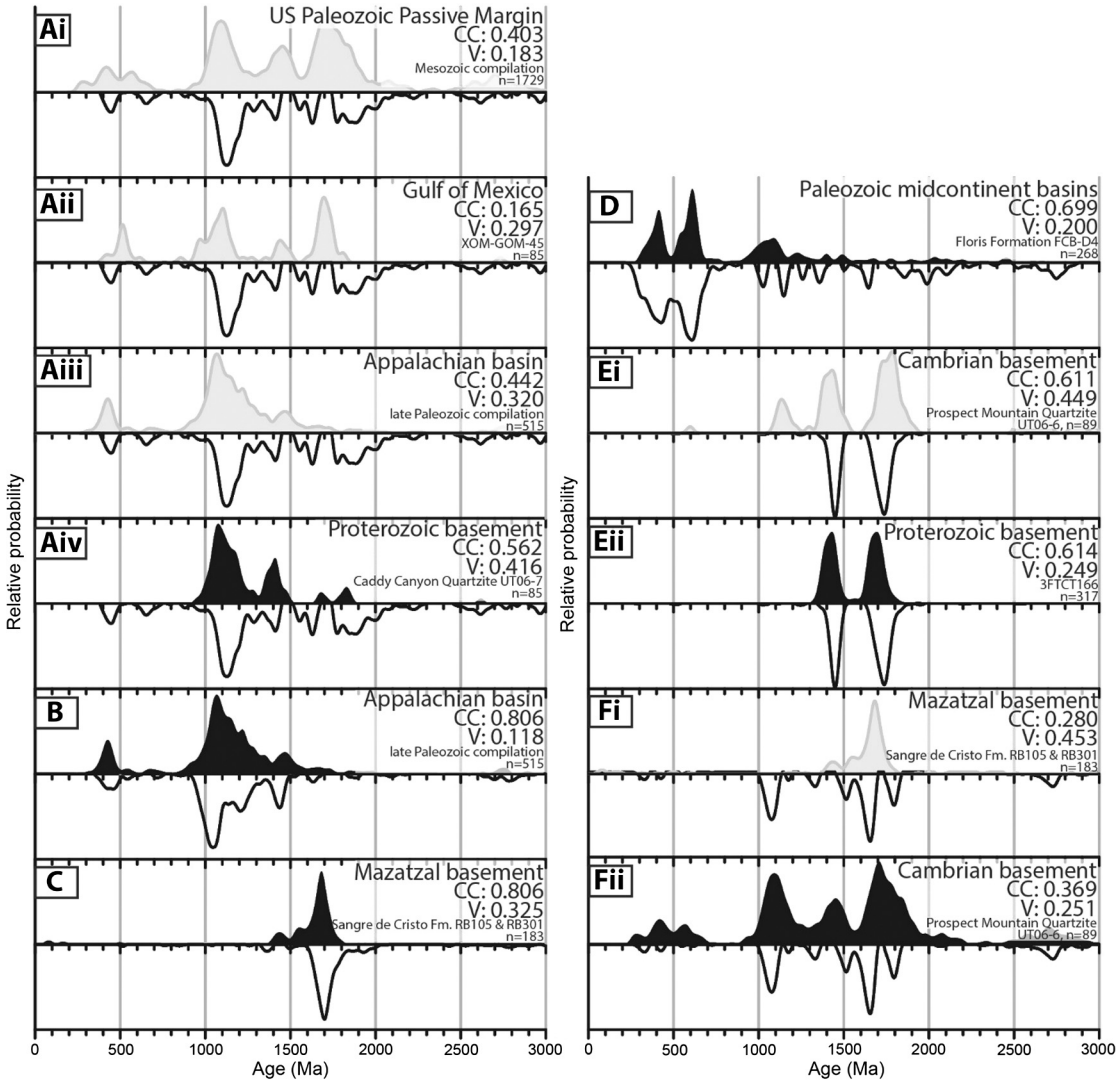

Fig. 9. Comparison of potential sediment sources (black or gray filled KDEs) to factorized sources (inverted unfilled KDEs) for the Grand Canyon data set. Black-filled KDEs are our preferred matches to factorized sources. References for potential sources include (A1) Laskowski et al. (2013), (A2) Blum et al. (2017), (A3 and B) Becker et al. (2005, Pottsville Formation), Becker et al. (2006, Washington and Greene formations), Park et al. (2010, Price, Stony Gap, and Princeton formations), A4) Lawton et al. (2010, Caddy Canyon Quartzite), C) Bush et al. (2016, RB105 and RB301 Sangre de Cristo Formation), D) Kissock et al. (2018, FCB-D4 Floris Formation), F1 and G2) Lawton et al. (2010, UT06-6 Prospect Mountain Quartzite), F2) this study, G1) Laskowski et al. (2013, US Passive Margin provenance group). CC and *V* respectively indicate the Cross-correlation coefficient and Kuiper *V* value between the known source and the subjacent factorized source. KDEs are constructed with a Gaussian kernel with a bandwidth of 20 Myr. Data for this figure are in Supplemental Table 28.

and >250 Ma components yields potential sources that are consistent with the subdivision of the Yavapai–Mazatzal provenance group by Laskowski et al. (2013), suggesting that this provenance group is itself a mixed group.

The Grand Canyon data set is best modeled with six provenance groups (Fig. 7B). The first source is best matched by a sample of Proterozoic Caddy Canyon Quartzite exposed in the Sevier thrust belt (Lawton et al., 2010) (Fig. 9Aiv). Alternatives, including a Paleozoic composite comprised of multiple samples from the Appalachian Basin (Becker et al., 2005, 2006; Park et al., 2010), a representative sample from the Gulf of Mexico (Blum et al., 2017), or the US Paleozoic passive margin (Laskowski et al., 2013) have lower *V* values but also lower Cross-correlation coefficients and clear mismatches in modal ages (Fig. 9Ai–Aiii). The second factorized source is most consistent with sources in the Appalachian basin (Fig. 9B). The third factorized source is most consistent with Proterozoic basement current exposed in the Sangre de Cristo mountains (Bush et al., 2016) (Fig. 9C). The fourth source is most consistent with samples from Paleozoic strata from mid-continent basins (Fig. 9D), which may be derived ultimately from basement sources in New England or Newfoundland (McLennan et al., 2001; Pollock et al., 2007; Fyffe et al., 2009). The fifth source is closely correlated to both Cambrian metasedimentary rocks characterized by the Prospect Mountain Quartzite or Proterozoic basement of the Uncompahgre Uplift (Fig. 9Ei and Eii). The sixth source is equally well matched by either a Paleozoic composite source indicating local recycling and mixing (Fig. 9Fi) or Cambrian metasedimentary rocks potentially resulting from mixture during transcontinental sediment transport (Fig. 9Fii) (Gehrels et al., 2011).

The Gulf of Mexico data set is successfully factorized into six provenance groups which have possible correlatives in Laurentian detrital zircon data sets (Fig. 10). The first provenance group is dominated by 1.25–0.95 Ga and 0.5–0.3 Ga ages that broadly overlap with Grenville, peri-Gondwanan, and Appalachian sources (hereafter termed 'Appalachian affinity' Fig. 10A). This provenance group is consistent with detrital zircon ages derived from the modern Appalachian Mountains region and recycled from older sedimentary sequences in the western U.S. (e.g., Mesozoic units

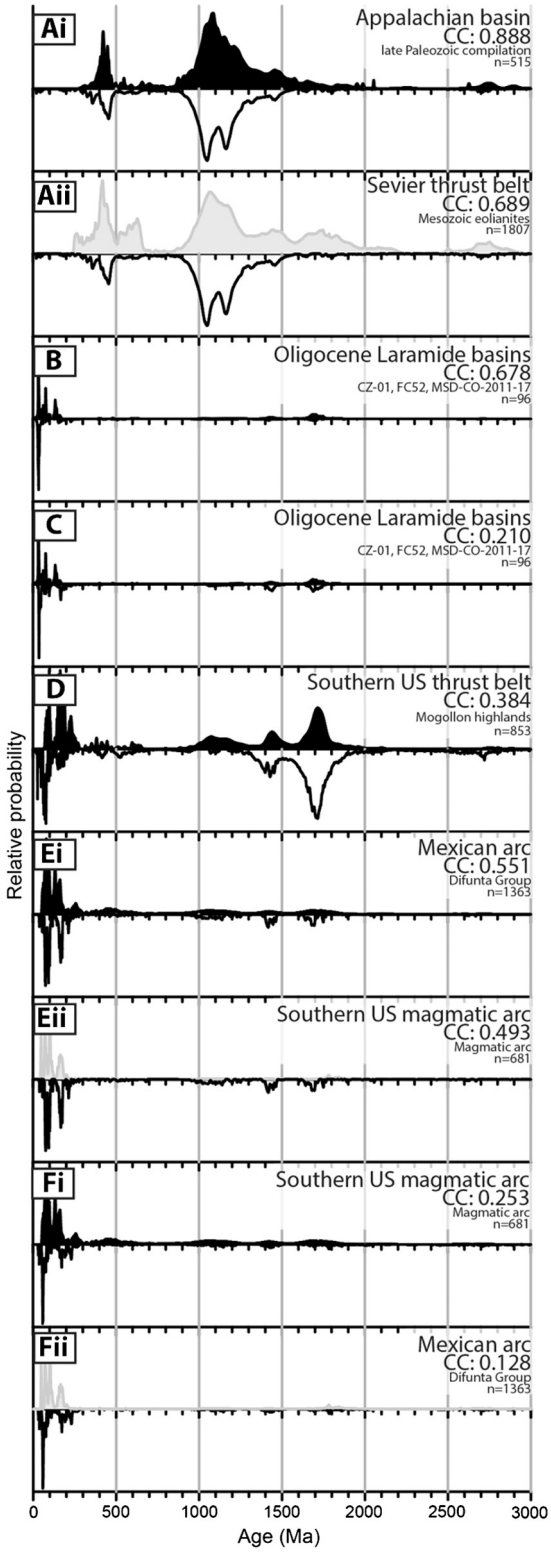

Fig. 10. Comparison of potential sediment sources (black or gray filled PDPs) to factorized sources (inverted unfilled PDPs) for the Gulf of Mexico data set. Black-filled PDPs are our preferred matches to factorized sources. References for potential sources include A1) Becker et al. (2005, Pottsville Formation), Becker et al. (2006, Washington and Greene formations), Park et al. (2010, Price, Stony Gap, and Princeton formations), A2) Laskowski et al. (2013, Mesozoic eolianite provenance group), B1) Escalona-Alcázar et al. (2016, CZ-01), Copeland et al. (2011, FC52), Donahue (2016, MSD-CO-2011-17), B2) May et al. (2013, TSA3 Late Cretaceous composite); D, F1, and G1) Laskowski et al. (2013, Mogollon highlands provenance group), and F2 and G2) Lawton et al. (2009, Difunta Group). CC indicates the Cross-correlation coefficient between the known source and the subjacent factorized source. Data for this figure are in Supplemental Table 28.

in the Sevier fold-thrust belt; Fig. 10). Four of the provenance groups (Fig. 10B, C, E, and F) are dominated by age distributions <300 Ma that correspond with major pulses of Cordilleran arc volcanism during middle Jurassic (175–160 Ma), Late Cretaceous (100–70 Ma), Eocene (58–57 Ma), and Oligocene (33–35 Ma) time. Such ages are widespread in western North America, including in the U.S. Cordilleran foreland basin, U.S. Sevier fold-thrust belt and magmatic arc, and in Mexican arc terranes (Fig. 10). Finally, a sixth provenance group is dominated by ages 1.8–1.35 Ga with peaks at 1.71 and 1.43 Ga that suggests derivation from Yavapai–Mazatzal and/or Midcontinent sources (Fig. 10D). This group also includes a significant <300 Ma age component reflecting Cordilleran arc magmatism with a dominant peak from 80–60 Ma.

The Gulf of Mexico data set also indicates that proportions of potential sources (as indicated by the factorized weighting functions) vary systematically in both space and time, and can be correlated to long-term tectonic changes and geological events (Fig. 11). Middle Cretaceous samples are dominated by Source 1, consistent with an interpretation of a dominantly Appalachian source to these units (Blum and Pecha, 2014; Blum et al., 2017) (Fig. 11B). Paleocene–lower Eocene samples show a dominance of Sources 2–6 (western North American affinity) in the western Gulf of Mexico which correlates broadly with peak Laramide exhumation, and recycling of Sevier foreland basin strata (Lawton, 2008). This mixes with Source 1 (Appalachian affinity) in the eastern Gulf of Mexico with a well-defined mixing zone extending from the Alabama–Mississippi border to the Louisiana–Texas border (Fig. 11B). This trend continues during continued tectonic activity in the Oligocene and lower Miocene. Furthermore, the pronounced increase in abundance of Sources 2 and 3 in Oligocene samples and subsequent decrease in lower Miocene samples likely reflects the Oligocene ignimbrite flare-up event within the western U.S. and Mexico (Fig. 11B). Samples from rivers that drain into the northern Gulf of Mexico also display similar contributions from potential sources as has been noted in previous research (e.g., Blum and Pecha, 2014; Blum et al., 2017; Sharman and Johnstone, 2017; Xu et al., 2017; Fig. 11B). These modern rivers display notably higher proportions of Source 1 than their older, Cenozoic counterparts (Fig. 11B). We speculate that this change could reflect increased recycling of older sedimentary sequences (e.g.,. Fig. 11Ai) and decreased primary derivation from western U.S. Cordilleran arc and Precambrian sources. Although the driver of this change remains uncertain this may point to significant drainage reorganization since the Miocene, and potentially linked with integration of the Rio Grande River (Repasch et al., 2017; Fan et al., 2018).

Two general conclusions can be drawn from the analysis of the Book Cliffs, Grand Canyon, and Gulf of Mexico data sets. The first is that although the NMF algorithm is able to identify detrital zircon source signatures, it cannot discriminate between primary or recycled sources (Dickinson et al., 2009). This is exemplified in the case of the Grand Canyon data set, where possible sources include either sediment recycled from the Appalachian basin to the east and that sourced directly from Proterozoic metasedimentary sources (Fig. 9Aiii and Aiv) or between local recycling or transported admixture (Fig. 9Fi and Fii). The second observation is that the NMF algorithm will identify the most basic elements of the mixed sink samples and so may subdivide sources that are themselves heterogeneous mixtures of more basic elements into those basic elements. This conclusion is demonstrated in the Grand Canyon data set where sources that may have been mixed in a local basin or during transport to the Grand Canyon depocenter are divided into their constituent components including a late Proterozoic–early Paleozoic component and ~1000 Ma components (Fig. 9B and D). It is also observed in the Book Cliffs data set where the closest matches for sources five and six are the <250 Ma and >250 Ma

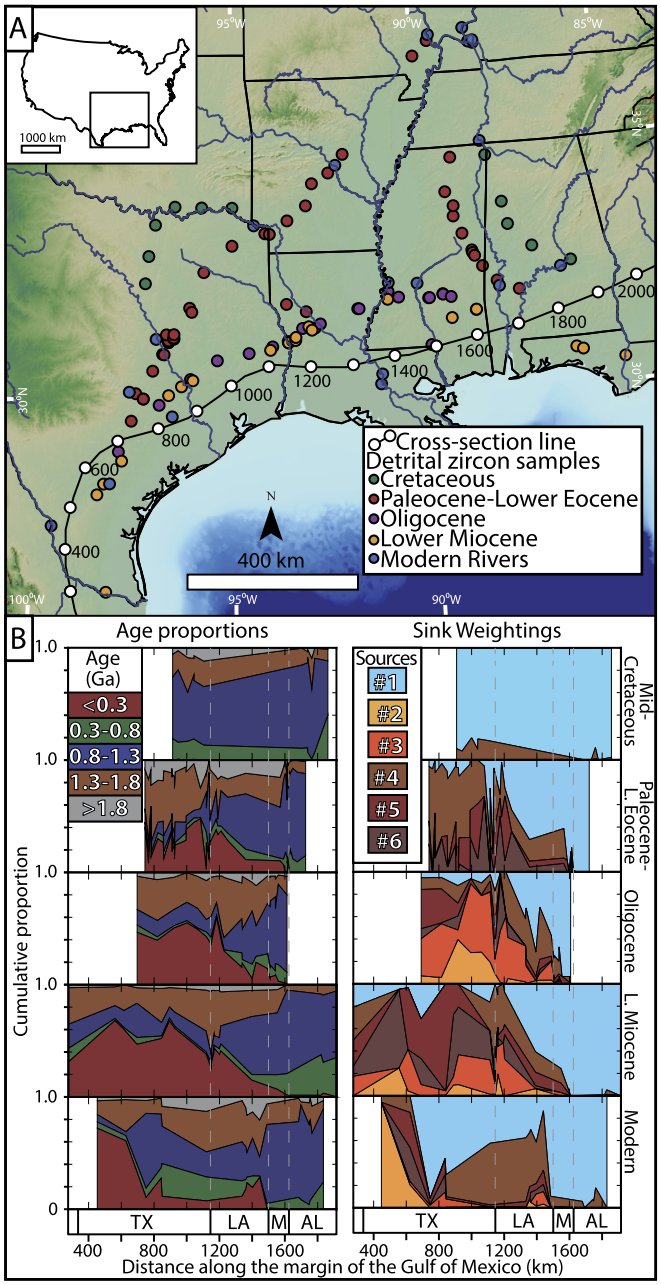

Fig. 11. A) Location and age of 111 samples used in the Gulf of Mexico case study. The dark black line shows the approximate location of the late Paleocene shelf edge (Galloway et al., 2011) and white circles indicate 100 km increments of distance measured along this paleo-shelf edge from southwest to northeast. B) Proportions of detrital zircon U–Pb ages (left) and modeled provenance sources (right) for the 111 samples shown in (i). Source PDPs are shown in Fig. 10A. The *x*-axis corresponds to the projected location of each sample along the late Paleocene shelf edge as shown in (i), used to normalize sample position along the rim of the northern Gulf of Mexico. Abbreviations: AL – Alabama; L. – Lower; LA – Louisiana; M – Mississippi; TX – Texas. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

subsets of the Mogollon highlands source (Fig. 8Ei and Fi). Finally, sources identified for the Gulf of Mexico are not recycled Cretaceous strata from the Sevier foreland basin but rather end-member components from the circum-Gulf of Mexico region, including the Appalachian Basin, Laramide basins, Mogollon highlands, US magmatic arc, and Mexican magmatic arc (Fig. 10). Interestingly, although factorization of the Book Cliffs data set yielded sources that were similar to subdivisions of the southern US magmatic arc source (Fig. 9Ei and Fi), factorization of Gulf of Mexico data set the does not subdivide the Mogollon highlands source in this manner. This may point to thorough mixing and homogenization of this source, potentially in Sevier or Laramide basins, prior to its introduction into the Gulf of Mexico sedimentary system.

6. Conclusions

We introduce a new implementation of non-negative matrix factorization (NMF) to detrital geochronology to characterize sediment sources based on mixed ("sink") samples. The algorithm is unsupervised and requires no training data. Application of this algorithm to samples mixed from known sources indicates that the algorithm can fully recover the source components given only the sink samples. This is true even with complex, multi-modal data sets with overlapping age modes. Sensitivity testing indicates that analyses conducted with 3 or more times as many sink samples as source samples yield a close match between known and factorized sources. However, factorized and known sources become less similar with increasing sink sample similarity. We conclude that a range of Cross-correlation coefficients of >0.3 and/or a range of V values $>\sim 0.35$ consistently yields close similarity between known and factorized sources. Finally, sink samples must be well-characterized (i.e., $n > 150$–225 per sample) or the number of samples in the sink data set must be high (i.e., $N > 3$ times the number of sources) to yield close similarity between known sources and factorized sources.

We also introduce an approach to determine the optimum number of sources based on minimizing the summed sum of squared residuals of two linear segmented fits to the NMF final residual. Application of this approach to samples with a known number of sources confirms that the summed sum of squared residuals is minimized when the breakpoint between the linear segments coincides with the known number of source samples. Application of this approach to the Book Cliffs data set confirms independent previous research that the Book Cliffs data set can be successfully modeled using five source provenance groups and that the Canadian Paleozoic passive margin strata are not significant contributors to Cretaceous Book Cliffs strata. The method's success in this intermediate step between a fully controlled and fully unknown application gives confidence that when applied to empirical data sets the results are robust.

Application of the NMF algorithm to empirical data sets from the Book Cliffs, Grand Canyon, and Gulf of Mexico confirm that the algorithm recovers plausible source age distributions from each of these data sets. However, while the algorithm can factorize the sink samples into plausible source samples, it cannot discriminate between identical or very similar sources. Furthermore, it may factorize heterogeneous mixed sources into their most basic elements, potentially obscuring an intermediate mixing and recycling step. However, it may be "blind" to mixing if the mixed sources are themselves relatively homogeneous. Finally, although the algorithm reproduces known sources and yields reasonable sources for empirical data sets, inversions are inherently non-unique. These caveats highlight the importance of a thorough understanding of the geological context of both sink samples and potential sources prior to drawing conclusions based on NMF.

Acknowledgements

This research was supported by grants from the National Science Foundation (EAR-1550097 and -1742952) to Saylor. The authors thank Dr. Noah McLean for extensive discussions and input which helped to shape this manuscript. We acknowledge Dr. P. Vermeesch who reviewed a previous version of this paper, and are extremely grateful to an anonymous reviewer for several helpful

reviews. Finally, we thank Dr. A. Yin for help in editorial handling of this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.epsl.2019.01.044.

References

Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005a. Construction of detrital mineral populations: insights from mixing of U–Pb zircon ages in Himalayan rivers. Basin Res. 17, 463–485.

Amidon, W.H., Burbank, D.W., Gehrels, G.E., 2005b. U–Pb zircon ages as a sediment mixing tracer in the Nepal Himalaya. Earth Planet. Sci. Lett. 235, 244–260.

Andersen, T., Kristoffersen, M., Elburg, M.A., 2018. Visualizing, interpreting and comparing detrital zircon age and Hf isotope data in basin analysis – a graphical approach. Basin Res. 30, 132–147.

Bartschi, N.C., Saylor, J.E., Lapen, T.J., Blum, M.D., Pettit, B.S., Andrea, R.A., 2018. Tectonic controls on Late Cretaceous sediment provenance and stratigraphic architecture in the Book Cliffs, Utah. Geol. Soc. Am. Bull. 130 (11–12), 1763–1781.

Becker, T.P., Thomas, W.A., Gehrels, G.E., 2006. Linking late Paleozoic sedimentary provenance in the Appalachian basin to the history of Alleghanian deformation. Am. J. Sci. 306, 777–798.

Becker, T.P., Thomas, W.A., Samson, S.D., Gehrels, G.E., 2005. Detrital zircon evidence of Laurentian crustal dominance in the lower Pennsylvanian deposits of the Alleghanian clastic wedge in eastern North America. Sediment. Geol. 182, 59–86.

Blum, M., Pecha, M., 2014. Mid-Cretaceous to Paleocene North American drainage reorganization from detrital zircons. Geology 42, 607–610.

Blum, M.D., Milliken, K.T., Pecha, M.A., Snedden, J.W., Frederick, B.C., Galloway, W.E., 2017. Detrital-zircon records of Cenomanian, Paleocene, and Oligocene Gulf of Mexico drainage integration and sediment routing: implications for scales of basin-floor fans. Geosphere 13, 2169–2205.

Borg, I., Groenen, P.J., 1997. Modern Multidimensional Scaling: Theory and Applications. Springer, New York. 471 p.

Brandon, M.T., 1996. Probability density plot for fission-track grain-age samples. Radiat. Meas. 26, 663–676.

Briechle, K., Hanebeck, U.D., 2001. Template matching using fast normalized cross correlation. Proc. SPIE, 95–102.

Bush, M.A., Horton, B.K., Murphy, M.A., Stockli, D.F., 2016. Detrital record of initial basement exhumation along the Laramide deformation front, southern Rocky Mountains. Tectonics 35, 2117–2130.

Capaldi, T.N., Horton, B.K., McKenzie, N.R., Stockli, D.F., Odlum, M.L., 2017. Sediment provenance in contractional orogens: the detrital zircon record from modern rivers in the Andean fold-thrust belt and foreland basin of western Argentina. Earth Planet. Sci. Lett. 479, 83–97.

Copeland, P., Murphy, M.A., Dupré, W.R., Lapen, T.J., 2011. Oligocene Laramide deformation in southern New Mexico and its implications for Farallon plate geodynamics. Geosphere 7, 1209–1219.

Craddock, W.H., Kylander-Clark, A.R., 2013. U–Pb ages of detrital zircons from the Tertiary Mississippi River Delta in central Louisiana: insights into sediment provenance. Geosphere 9, 1832–1851.

Debella-Gilo, M., Kääb, A., 2011. Sub-pixel precision image matching for measuring surface displacements on mass movements using normalized cross-correlation. Remote Sens. Environ. 115, 130–142.

DeGraaff-Surpless, K., Mahoney, J.B., Wooden, J.L., McWilliams, M.O., 2003. Lithofacies control in detrital zircon provenance studies: insights from the Cretaceous Methow basin, southern Canadian Cordillera. Geol. Soc. Am. Bull. 115, 899–915.

Dickinson, W.R., Lawton, T.F., Gehrels, G.E., 2009. Recycling detrital zircons: a case study from the Cretaceous Bisbee Group of southern Arizona. Geology 37, 503–506.

Dodson, M.H., Compston, W., Williams, I.S., Wilson, J.F., 1988. A search for ancient detrital zircons in Zimbabwean sediments. J. Geol. Soc. Lond. 145, 977–983.

Donahue, M.M.S., 2016. Episodic Uplift of the Rocky Mountains: Evidence From U–Pb Detrital Zircon Geochronology and Low-Temperature Thermochronology with a Chapter on Using Mobile Technology for Geoscience Education. The University of New Mexico.

Draper, N.R., Smith, H., 1998. Applied Regression Analysis, 3rd edition. Wiley and Sons, New York.

Escalona-Alcázar, F.d.J., Solari, L., García y Barragán, J.C., Carrillo-Castillo, C., Bluhm-Gutiérrez, J., García-Sandoval, P., Nieto-Samaniego, Á.F., Núñez-Peña, E.P., 2016. The Palaeocene–early Oligocene Zacatecas conglomerate, Mexico: sedimentology, detrital zircon U–Pb ages, and sandstone provenance. Int. Geol. Rev. 58, 826–848.

Fan, M., Brown, E., Li, L., 2018. Cenozoic drainage evolution of the Rio Grande paleoriver recorded in detrital zircons in South Texas. Int. Geol. Rev., 1–15. https://doi.org/10.1080/00206814.2018.1446368.

Fyffe, L.R., Barr, S.M., Johnson, S.C., McLeod, M.J., McNicoll, V.J., Valverde-Vaquero, P., van Staal, C.R., White, C.E., 2009. Detrital zircon ages from Neoproterozoic and Early Paleozoic conglomerate and sandstone units of New Brunswick and coastal Maine: implications for the tectonic evolution of Ganderia. Atlantic Geol. 45, 110–144.

Galloway, W.E., Whiteaker, T.L., Ganey-Curry, P., 2011. History of Cenozoic North American drainage basin evolution, sediment yield, and accumulation in the Gulf of Mexico basin. Geosphere 7, 938–973.

Gehrels, G., 2012. Detrital zircon U–Pb geochronology: current methods and new opportunities. In: Busby, C., Azor, A. (Eds.), Tectonics of Sedimentary Basins. John Wiley & Sons, Ltd., pp. 45–62.

Gehrels, G.E., Blakey, R., Karlstrom, K.E., Timmons, J.M., Dickinson, B., Pecha, M., 2011. Detrital zircon U–Pb geochronology of Paleozoic strata in the Grand Canyon, Arizona. Lithosphere 3, 183–200.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L., 1998. Multivariate Data Analysis. Prentice Hall, Upper Saddle River, NJ.

Horton, B.K., Anderson, V.J., Caballero, V., Saylor, J.E., Nie, J., Parra, M., Mora, A., 2015. Application of detrital zircon U–Pb geochronology to surface and subsurface correlations of provenance, paleodrainage, and tectonics of the Middle Magdalena Valley Basin of Colombia. Geosphere 11, 1790–1811.

Hurford, A.J., Fitch, F.J., Clarke, A., 1984. Resolution of the age structure of the detrital zircon populations of two Lower Cretaceous sandstones from the Weald of England by fission track dating. Geol. Mag. 121, 269–277.

Iizuka, T., Hirata, T., Komiya, T., Rino, S., Katayama, I., Motoki, A., Maruyama, S., 2005. U–Pb and Lu–Hf isotope systematics of zircons from the Mississippi River sand: implications for reworking and growth of continental crust. Geology 33, 485–488.

Jessberger, E.K., Dominik, B., Staudacher, T., Herzog, G.F., 1980. 40Ar·39Ar ages of Allende. Icarus 42, 380–405.

Kim, H., Park, H., 2007. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. Bioinformatics 23, 1495–1502.

Kim, H., Park, H., 2008. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. SIAM J. Matrix Anal. Appl. 30, 713–730.

Kimbrough, D.L., Grove, M., Gehrels, G.E., Dorsey, R.J., Howard, K.A., Lovera, O., Aslan, A., House, P.K., Pearthree, P.A., 2015. Detrital zircon U–Pb provenance of the Colorado River: a 5 m.y. record of incision into cover strata overlying the Colorado Plateau and adjacent regions. Geosphere 11 (6), 1719–1748.

Kissock, J., Finzel, E., Malone, D., Craddock, J., 2018. Lower–Middle Pennsylvanian strata in the North American midcontinent record the interplay between erosional unroofing of the Appalachians and eustatic sea-level rise. Geosphere 14, 141–161.

Kuiper, N.H., 1960. Tests concerning random points on a circle. Proc. K. Ned. Akad. Wet. 63, 38–47.

Laskowski, A.K., DeCelles, P.G., Gehrels, G.E., 2013. Detrital zircon geochronology of Cordilleran retroarc foreland basin strata, western North America. Tectonics 32, 1027–1048.

Lawton, T.F., 2008. Laramide sedimentary basins. In: Miall, A.D. (Ed.), Sedimentary Basins of the World. Elsevier, pp. 429–450 (Chapter 12).

Lawton, T.F., Bradford, I.A., Vega, F.J., Gehrels, G.E., Amato, J.M., 2009. Provenance of Upper Cretaceous–Paleogene sandstones in the foreland basin system of the Sierra Madre Oriental, northeastern Mexico, and its bearing on fluvial dispersal systems of the Mexican Laramide Province Sierra Madre Oriental sandstone provenance. Geol. Soc. Am. Bull. 121, 820–836.

Lawton, T.F., Hunt, G.J., Gehrels, G.E., 2010. Detrital zircon record of thrust belt unroofing in Lower Cretaceous synorogenic conglomerates, central Utah. Geology 38, 463–466.

Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. Nature 401, 788–791.

Lewis, J.P., 1995. Fast normalized cross-correlation. Vis. Interface, 120–123.

Li, Y., Ngom, A., 2013. The non-negative matrix factorization toolbox for biological data mining. Source Code Biol. Med. 8, 10.

Licht, A., Pullen, A., Kapp, P., Abell, J., Giesler, N., 2016. Eolian cannibalism: reworked loess and fluvial sediment as the main sources of the Chinese Loess Plateau. Geol. Soc. Am. Bull. 128, 944–956.

Mason, C.C., Fildani, A., Gerber, T., Blum, M.D., Clark, J.D., Dykstra, M., 2017. Climatic and anthropogenic influences on sediment mixing in the Mississippi source-to-sink system using detrital zircons: Late Pleistocene to recent. Earth Planet. Sci. Lett. 466, 70–79.

May, S.R., Gray, G.G., Summa, L.L., Stewart, N.R., Gehrels, G.E., Pecha, M.E., 2013. Detrital zircon geochronology from the Bighorn Basin, Wyoming, USA: implications for tectonostratigraphic evolution and paleogeography. Geol. Soc. Am. Bull. 125 (9–10), 1403–1422.

McLennan, S., Bock, B., Compston, W., Hemming, S., McDaniel, D., 2001. Detrital zircon geochronology of Taconian and Acadian foreland sedimentary rocks in New England. J. Sediment. Res. 71, 305–317.

Ozerov, A., Fevotte, C., 2010. Multichannel nonnegative matrix factorization in convolutive mixtures for audio source separation. IEEE Trans. Audio Speech Lang. Process. 18, 550–563.

Pan, B., Qian, K., Xie, H., Asundi, A., 2009. Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. Meas. Sci. Technol. 20, 062001.
Park, H., Barbeau Jr, D.L., Rickenbaker, A., Bachmann-Krug, D., Gehrels, G., 2010. Application of foreland basin detrital-zircon geochronology to the reconstruction of the Southern and Central Appalachian Orogen. J. Geol. 118, 23–44.
Paterson, G.A., Heslop, D., 2015. New methods for unmixing sediment grain size data. Geochem. Geophys. Geosyst. 16, 4494–4506.
Pollock, J., Wilton, D., Van Staal, C., Morrissey, K., 2007. U–Pb detrital zircon geochronological constraints on the Early Silurian collision of Ganderia and Laurentia along the Dog Bay Line: the terminal Iapetan suture in the Newfoundland Appalachians. Am. J. Sci. 307, 399–433.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes: The Art of Scientific Computing, third ed. Cambridge University Press, New York.
Pullen, A., Ibañez-Mejia, M., Gehrels, G.E., Ibañez-Mejia, J.C., Pecha, M., 2014. What happens when $n = 1000$? Creating large-n geochronological datasets with LA-ICP-MS for geologic investigations. J. Anal. At. Spectrom. 29, 971–980.
Repasch, M., Karlstrom, K., Heizler, M., Pecha, M., 2017. Birth and evolution of the Rio Grande fluvial system in the past 8 Ma: progressive downward integration and the influence of tectonics, volcanism, and climate. Earth-Sci. Rev. 168, 113–164.
Satkoski, A.M., Wilkinson, B.H., Hietpas, J., Samson, S.D., 2013. Likeness among detrital zircon populations – an approach to the comparison of age frequency data in time and space. Geol. Soc. Am. Bull. 125, 1783–1799.
Saylor, J.E., Knowles, J.N., Horton, B.K., Nie, J.S., Mora, A., 2013. Mixing of source populations recorded in detrital zircon U–Pb age spectra of modern river sands. J. Geol. 121, 17–33.
Saylor, J.E., Stockli, D.F., Horton, B.K., Nie, J., Mora, A., 2012. Discriminating rapid exhumation from syndepositional volcanism using detrital zircon double dating: implications for the tectonic history of the Eastern Cordillera, Colombia. Geol. Soc. Am. Bull. 124, 762–779.
Saylor, J.E., Sundell, K.E., 2016. Quantifying comparison of large detrital geochronology data sets. Geosphere 12, 203–220.
Sharman, G.R., Johnstone, S.A., 2017. Sediment unmixing using detrital geochronology. Earth Planet. Sci. Lett. 477, 183–194.
Sickmann, Z.T., Paull, C.K., Graham, S.A., 2016. Detrital-zircon mixing and partitioning in fluvial to deep marine systems, Central California, U.S.A. J. Sediment. Res. 86, 1298–1307.
Sircombe, K.N., Hazelton, M.L., 2004. Comparison of detrital zircon age distributions by kernel functional estimation. Sediment. Geol. 171, 91–111.
Smaragdis, P., Brown, J.C., 2003. Non-negative matrix factorization for polyphonic music transcription. In: 2003 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. IEEE, pp. 177–180.
Stafford, R., 2006. Matlab File Exchange: Random Vectors with Fixed Sum (v. 1.0).
Steyvers, M., 2006. Multidimensional Scaling, Encyclopedia of Cognitive Science. John Wiley & Sons, Ltd.
Sundell, K.E., Saylor, J.E., 2017. Unmixing detrital geochronology age distributions. Geochem. Geophys. Geosyst. 18 (8), 2874–2886.
Troyan, V., 2010. Statistical Methods of Geophysical Data Processing. World Scientific.
Van Benthem, M.H., Keenan, M.R., 2004. Fast algorithm for the solution of large-scale non-negativity-constrained least squares problems. J. Chemom. 18, 441–450.
Vermeesch, P., 2012. On the visualisation of detrital age distributions. Chem. Geol. 312, 190–194.
Vermeesch, P., 2013. Multi-sample comparison of detrital age distributions. Chem. Geol. 341, 140–146.
Vermeesch, P., Resentini, A., Garzanti, E., 2016. An R package for statistical provenance analysis. Sediment. Geol. 336, 14–25.
Wahl, P.J., Yancey, T.E., Pope, M.C., Miller, B.V., Ayers, W.B., 2016. U–Pb detrital zircon geochronology of the Upper Paleocene to Lower Eocene Wilcox Group, east-central Texas. Geosphere 12, 1517–1531.
Wissink, G.K., Wilkinson, B.H., Hoke, G.D., 2018. Pairwise sample comparisons and multidimensional scaling of detrital zircon ages with examples from the North American platform, basin, and passive margin settings. Lithosphere 10 (3), 478–491.
Xu, J., Snedden, J.W., Stockli, D.F., Fulthorpe, C.S., Galloway, W.E., 2017. Early Miocene continental-scale sediment supply to the Gulf of Mexico Basin based on detrital zircon analysis. Geol. Soc. Am. Bull. 129, 3–22.
Zhao, F., Huang, Q., Gao, W., 2006. Image matching by normalized cross-correlation. In: 2006 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006 Proceedings, vol. 2. IEEE. pp. II-729–II-732.