

## Application of end-member modelling to grain-size data: Constraints and limitations

MICHAEL DIETZE\*·† , PHILIPP SCHULTE‡ and ELISABETH DIETZE§·¶

\*GFZ German Research Centre for Geosciences, Section 4.6 Geomorphology, Potsdam, 14473, Germany (E-mail: mdietze@gfz-potsdam.de)

†Department of Geography, University of Bonn, Bonn, D-53115, Germany

‡Department of Geography, RWTH (Rheinisch-Westfälische Technische Hochschule) Aachen University, Aachen, 52056, Germany

§Alfred Wegener Institute for Polar and Marine Research, Research Unit Potsdam, Potsdam, 14473, Germany

¶GFZ German Research Centre for Geosciences, Section 3.2 Organic Geochemistry, Potsdam, 14473, Germany

Associate Editor – Subhasish Dey

### ABSTRACT

End-member modelling analysis (EMMA) is a statistical approach to unmixing multimodal grain-size distributions to identify and quantify processes of sediment generation, transport and deposition. While the different computational implementations have been extensively benchmarked and show similarly high reliability characteristics, there is a series of unknowns regarding the applicability, quality and limitations of the method from a practical point of view. This study explores these important unknowns using both empirical and synthetic samples along with Monte Carlo tests. Under ideal conditions (all available samples, randomly mixed components, 116 grain-size classes), EMMA is able to model the grain-size distributions of input end-members (loadings) with  $R^2$  between 0.63 and 0.98 and their relative contributions to each sample (scores) with  $R^2$  between 0.71 and 0.81, thus setting the baseline for model quality. Inappropriate model parameter settings cause severe drops in  $R^2$ . EMMA is able to detect an end-member even if it is present in only one sample or when it contributes less than 10 vol.%. With 20 to 40 samples or more, stable, high quality model results are possible. With 15 or more grain-size classes, model results also reach such stable high reproducibility levels. EMMA can depict originally multimodal end-members ( $R^2$  between 0.78 and 0.99). End-members with identical relative grain-size distribution shape can overlap significantly without causing quality drops;  $R^2$  of identical distributions are invariantly high until mode positions are less than three grain-size classes apart from each other. Gradually widening end-member distributions do not affect the results significantly. However, shifting mode positions have a severe impact. Post-depositional mixing causes drastic deviations of the modelled scores, whereas the loadings are virtually unaffected. In light of these tests, EMMA is a reliable, mostly unbiased tool to identify and quantify sediment generation/transport/deposition regimes from mixed sediment deposits, given that it is used in a geoscientifically meaningful context.

**Keywords** End-member modelling analysis, grain-size modelling, sensitivity test, synthetic data, validation.

## INTRODUCTION

End-member modelling analysis (EMMA) of grain-size data is an established way to identify and quantify the grain-size imprint of distinct sediment sources, transport processes and pathways from the mixed, multimodal grain-size distributions of deposited sediments (Flemming, 2007; Hartmann, 2007; Weltje & Prins, 2007; Dietze *et al.*, 2012, 2014; Vandenberghe, 2013; van Hateren *et al.*, 2018). EMMA uses a measured data set ( $X$ ), consisting of  $m$  grain-size distributions, each described by  $n$  grain-size classes, and produces a modelled data set ( $X'$ ) as linear combination of end-member loadings ( $V$ ) and end-member scores ( $M$ ) that equals the input data set plus an error matrix ( $E$ ), i.e.  $X' = MV^T = X - E$ . Thereby,  $E$  is usually very small so that the  $R^2$  between  $X$  and  $X'$  is at the order of 0.9 and higher. Loadings are the individual grain-size distributions that form a mixed sample and scores are the contribution of each loading to each sample. For a more elaborated introduction and background discussion see Dietze *et al.* (2012) and Dietze & Dietze (2019). While EMMA is most commonly used for grain-size data analysis, the method is more generic and in principle allows to model a wide range of categorical data.

Despite the increasing popularity of EMMA since its introduction almost 25 years ago (Weltje, 1997) there is still notable hesitation in its application. This may be partly due to the 'black box' impression but can also be explained by the lack of a profound validation with natural data, including questions of practical applicability. The most obvious question is certainly how 'well' EMMA reproduces original end-members as well as their contribution to each sample. To test this, Weltje & Prins (2007) used two to four artificial grain-size distributions, defined by normal and Weibull distributions, which were mixed in known proportions. They found in general good agreement but also underestimation and overestimation by several percent. Dietze & Dietze (2019) compared all available EMMA implementations and found overall similar results among these, with each implementation having its particular minor strengths and weaknesses. The comparison used four distinct empirical grain-size end-members, which were randomly mixed to create a reference data set. While this approach was suitable for EMMA implementation benchmarking, it did not account for further sources of uncertainty and

model performance, which arise from an application point of view.

Uncertainty not only arises from model-inherent constraints. One primary source contributes even before that: variability of a sample's grain-size distribution due to bulk sample heterogeneity and measurement device effects. The effects of these two sources of uncertainty are typically addressed by measuring subsequent aliquots of the same bulk sample (accounting for sample heterogeneity) and repeat measurements of the same aliquot (accounting for measurement device effects). However, the analytical accuracy is essentially impossible to determine for samples of natural sediment because the actual size of a particle cannot be expressed by a single number (Pye & Blott, 2004; Roberson & Weltje, 2014). Roberson & Weltje (2014) used ten aliquots of four natural sediment samples, respectively, to estimate the analytical precision of ten different instruments. Differences occur between all particle-size analyzers, even between instruments using the same technique but by different manufacturers. Scatter also varies depending on the size distribution of the analyzed material. Based on quadruple measurements of 1485 silty to loamy samples, Miller & Schaetzl (2012) found a relative precision maximum for silt-dominated samples, and thus higher scatter in the clay and sand fraction. Similarly, measurement precision is higher for better sorted samples, mostly due to smaller deviations from the original sample composition when taking aliquots (Pye & Blott, 2004; Schulte *et al.*, 2016). Hence, in order to properly address uncertainties arising from EMMA, the uncertainty due to sample heterogeneity and the measurement approach need to be considered, as well. However, such a test cannot include the question concerning the representativeness of a sample regarding the depositional dynamics.

With respect to model uncertainty due to parameterization, robust EMMA (Dietze & Dietze, 2019) inherits a statistically optimal parameter estimation. The routine creates as many model scenarios as possible based on the ranges of input parameters and identifies only those that occur persistently. However, especially the determination of the final number of end-members ( $q$ ) is a crucial step that requires a researcher to understand the depositional environment of the analyzed samples. This subjective decision on the supposed number of end-members, but also further parameters of the individual approach implementations (for an

overview see Dietze & Dietze, 2019), is especially important for all deterministic approaches. A simple empiric test of model runs using the correct and incorrect number of end-members can provide insight to the effects of inappropriate parameter settings, at least for the specific data set under consideration.

The number of grain-size classes that are used to describe the grain-size distribution is a further factor of uncertainty. This number is of course determined by the measurement technique. Laser diffraction data sets may easily reach between 40 and 116 classes, while optical imaging approaches yield up to 300 classes, distributed over arbitrarily narrow grain-size limits. At the other end of the spectrum, sieve-sedimentation approaches provide as little as seven classes. Mathematically or computationally, one needs at least as many grain-size classes as there are end-members contributing to the data set, but beyond that, the end-member's grain-size distributions must also be described with sufficient distinctness by the number of grain-size classes to be successfully identified. This discretization problem arises because grain-size data measurements essentially contain binned data and the bins must be narrow enough to adequately represent the structure of the data. Previous EMMA tests (Weltje & Prins, 2007; van Hateren *et al.*, 2018; Dietze & Dietze, 2019) have always used decisively high numbers of grain-size classes to suppress effects of this source of uncertainty, and to focus on other questions of interest.

In a similar way, the number of samples – and the variance in end-member mixing expressed by them – should have a significant impact on the modelling result. Again, from a mathematical perspective one needs at least as many samples as end-members to be modelled, but the samples must also reflect sufficient variability of end-member contribution to the entire data set (Dietze *et al.*, 2012). According to Weltje & Prins (2007) it is important to evaluate how much an end-member must contribute to a sample to identify it properly by EMMA. However, this suggestion has not been quantitatively tested.

It has been repeatedly reported that the different EMMA implementations tend to underestimate low and overestimate high end-member abundances (Paterson & Heslop, 2015; Dietze & Dietze, 2019). However, a quantification of this tendency has not yet been attempted, despite the fact that ultimately scientists will use end-member contributions (scores) as proxy for

palaeoenvironmental interpretations and will thus inevitably need to know the size of this bias.

Post-depositional mixing may introduce blurring of end-member properties. Essentially, one of the fundamental assumptions of EMMA is that the measured grain-size distributions are not affected by any post-depositional alteration, i.e. that the measured data only reflects the underlying transport and depositional dynamics, and not any subsequent modifications. Therefore, it crucially depends on the scientist's judgement to decide whether or not EMMA can be applied to a case study or not. Nevertheless, it is relevant to investigate the actual effect of post-depositional mixing on the model outcomes. Major effects may be rather obvious for the sample composition (scores) but not so obvious for the model results concerning the inferred end-member shapes (loadings). Apart from pure mixing, post-depositional processes can also alter the grain-size of sediments, for example through neoformation of clay minerals or physical weathering. While subsequent addition of neoformation products with a specific grain-size distribution can actually be included in EMMA and ultimately enhance the interpretation of a data set (Dietze *et al.*, 2016), this is not the case for a gradually (for example, down-core) increasing impact of weathering, which results in a shift of affected grain-size distributions. Such a non-stationarity in the properties of the end-members contributing to a sample violates the fundamental assumption of EMMA (Weltje, 1997; Dietze *et al.*, 2012). In a similar way, non-stationarity can also arise from gradually changing transport pathways (shifting distance to source), transport processes (wind or flowing water velocity) or source area conditions (gradual winnowing of selected grain-size classes), but regardless of the cause, the actual impact on modelling results has not been explored, yet.

From a process view, end-members should be unimodal in grain-size distribution, since a given transport regime tends to move (winnow) and deposit (concentrate) a specific grain size. Yet, there are cases when multimodal distributions are indeed characteristic for one end-member, for example reworked loess (Meszner *et al.*, 2013), desert dust (Sweeny *et al.*, 2013) and far travelled dust groups (Vandenberghe, 2013). Weltje & Prins (2003) discuss this possibility and propose that multimodal end-members may not necessarily be decomposed into elementary

sub-populations. However, a quantitative assessment of the ability of EMMA to reproduce primarily multimodal end-members is still lacking. It should be emphasized though that EMMA is not sensitive to the order of the data's classes because it does, by definition, handle the data as categorical. Literally, one could randomly change the order of grain-size classes before unmixing and then re-establish the initial order without any difference of the results. Hence, one would expect that multimodal end-members must be unmixed by EMMA. That said, the test of EMMA to handle multimodal data remains an interesting and relevant one from a geoscientific application perspective.

It is straightforward to identify different contributing end-members as long as they have very distinct grain-size distributions (for example, mixtures of loess and fluvial sediment) that overlap only marginally. However, as the distributions of contributing components become more and more similar to one another, the visual impression of their existence becomes less and less obvious. EMMA has been used decisively to unmix samples with such highly overlapping end-members of a similar shape (e.g. Vandenberghe, 2013) but a systematic evaluation of the reliability of the model results is still pending. In essence, this topic is comparable to the issue of data discretization due to grain-size class bin width, which has been discussed above. However, most grain-size measurement devices come with a fixed maximum number of grain-size classes and boundaries. Thus, to gain insight for practical application cases, this discretization problem needs to be assessed.

In summary, despite successful applications of existing EMMA implementations across many different sediment transport systems, there is a series of open questions regarding constraints that come along with many real world data sets. Currently, these constraints are implicitly assumed to have negligible effects on the model results, however without robust empirical support. This study uses EMMA as available through the R package *EMMAgeo* (M. Dietze & Dietze, 2016) to address the following list of research questions, especially relevant for practical applications of the method in general:

- How much uncertainty in grain-size distributions is caused by the measurement?
- What is the effect of inappropriate model parameter settings?

- How does the number of grain-size classes affect model results?
- How does the number of samples affect model results?
- How does the relative contribution of an end-member to a sample affect the success of EMMA to identify it?
- How does post-depositional mixing affect end-member properties?
- How do gradually widening distributions of the contributing grain-size distributions affect model results?
- How do gradually shifting distribution modes of the contributing grain-size distributions affect model results?
- To which extent is EMMA capable of identifying multimodal end-members?
- How does similarity of end-members affect the model results?

These questions are approached here with empirical measurements and synthetic data, and Monte Carlo methods are utilized to robustly explore effects especially for small sample sizes. The Appendix S1 contains all data sets necessary to reproduce the results reported in this article, as well as the R-scripts used to perform all calculations and generate the corresponding figures.

## MATERIALS AND METHODS

### Sampling, preparation and measurement

The reference data set provided by Dietze & Dietze (2019), which is composed of four natural end-members collected at deposits in the area around Dresden, Germany, is used in this study. Each deposit has been formed by one distinct sediment transport process. North of Dresden, a large Pleistocene alluvial fan covers an area of approximately 2 km<sup>2</sup>. This  $EM_{nat1}$  has been sampled in an open sand pit where the original bedding structures were still visible. Strong winds blew out fine particles of the alluvial landform and deposited them to the east where they have formed the barchan dunes of the Dresdener Heide.  $EM_{nat2}$  has been taken from one of these dunes at a depth of 30 cm, above the illuvial horizons of the Podzol developed in the medium coarse sands.  $EM_{nat3}$  has been taken from a loess section near Ostrau, 50 km north-west of Dresden (Meszner *et al.*, 2013).  $EM_{nat4}$  was taken from an oxbow segment of the River Elbe in the

eastern part of the city of Dresden. The loamy floodplain deposit contained abundant fine roots, which were removed by hand. From each deposit about 2 to 3 kg of sediment have been sampled, air-dried, dry sieved to <2000  $\mu\text{m}$  and homogenized.

From each of the four natural end-members, 150 g material has been sampled. Calcium carbonate and organic material has been removed with hydrochloric acid (10%) and hydrogen peroxide (15%). From each sample, three parallel samples (0.3 to 2.0 g) have been separated for natural end-member grain-size distribution measurements. To keep particles dispersed, the samples were treated with 1.25 mL  $\text{Na}_4\text{P}_2\text{O}_7$  for 12 h in an overhead shaker (ISO 11277, 2002; Pye & Blott, 2004). Particle size distributions have been measured with a Laser Diffraction Particle Size Analyzer (Beckman Coulter LS 13 320; Beckman Coulter, Brea, CA, USA) at RWTH Aachen, delivering 116 classes within a size range of 0.04 to 2000  $\mu\text{m}$ . Between 7 and 16 aliquots per sample have been investigated in triplicate by an 'auto-prep' station enabling equal measuring conditions. Each aliquot has been measured twice while in the measuring bath. To calculate the grain-size distribution the Mie theory has been used with the following parameters: Fluid RI: 1.33; Sample RI: 1.55; Imaginary RI: 0.1 (Buurman *et al.*, 1997; Özer *et al.*, 2010; ISO 13320, 2009). Proportions of the natural end-members were mixed by weight, whereas the particle size distributions measured with laser diffraction are expressed as volume percent. However, assuming that the predominant particle densities of the different natural end-members are similar, the respective weight percentages were considered as equivalent to volume percentage.

### Model test approaches

The flowchart in Fig. 1 provides a summary of the involved natural and artificial end-members along with their combinations using different mixing ratios to generate grain-size distribution data sets to pursue the different research questions. Artificial end-members were created and used when either very large data sets (>100 samples) or data sets with specific needs (for example, bimodal end-members) were required.

For the hand-mixed data set  $X_{\text{nat}}$ , mixing ratios were defined by random numbers generated from a uniform distribution and their sums were normalized to 10 g. According to these

numbers, the respective sediments were amalgamated and dry-mixed by overhead shaking for 24 h. In general, three sets were created: (i) 50 samples with all four end-members; (ii) 25 samples without  $EM_{\text{nat}1}$ ; and (iii) 25 samples without  $EM_{\text{nat}4}$ . This approach accounted for more than just random mixing. It ensured that in a significant part of the data set at least one end-member was not present by definition. This strategy of sample mixing allowed testing EMMA with a more variable data set, especially when drawing random subsamples. It also allowed testing EMMA on essentially three different data sets. The reader may perform further tests of own interest with these data sets. To assess the reproducibility of the hand-mixing approach, three parallel sample sets with identical mixing ratios were included, i.e. groups of sample IDs 7–23–37–42, 51–63 and 84–90.

Parallel to the hand-made mixing and measurements to generate  $X_{\text{nat}}$ , numeric mixing was performed. Therefore, all four  $EM_{\text{nat}}$  were multiplied with the respective mixing ratios and classwise summed to create  $X_{\text{num}}$ . This allowed to estimate the uncertainty related to laboratory mixing and subsequent measurement.

To investigate the relevance of correct model settings in terms of number of end-members  $q$ , a deterministic EMMA (Dietze & Dietze, 2019) was run with the subsample of  $X_{\text{nat}}$  that only contained  $EM_{\text{nat}1}$ ,  $EM_{\text{nat}2}$  and  $EM_{\text{nat}3}$ . There,  $q$  was set to 3 and 4. A further test aimed at the effect of a presence of an additional end-member in only one sample. For this test, one sample containing four EM (i.e. sample 71) was added to the previous data set and EMMA was run again, with  $q=3$  and  $q=4$ .

To test the influence of sample size,  $X_{\text{nat}}$  was subsampled (random sampling without replacement) with stepwise increasing sample sizes. This would in principle yield 97 data sets containing between 4 and 100 samples. To account for the effect of different scenarios, each of the 97 potential data sets was created 1001 times by drawing different random samples from  $X_{\text{nat}}$ . To all of these 97 097 data sets, deterministic EMMA was applied with  $q=4$  and the weight transformation limit set to zero (i.e.  $l=0$ ). The minimum sample size 4 was determined by the number of  $q$ . Robust EMMA (Dietze & Dietze, 2019) was not feasible because of the large number of the data sets and non-automatic determination of some parameters.

To account for the influence of the number of grain-size classes,  $X_{\text{nat}}$  was iteratively re-

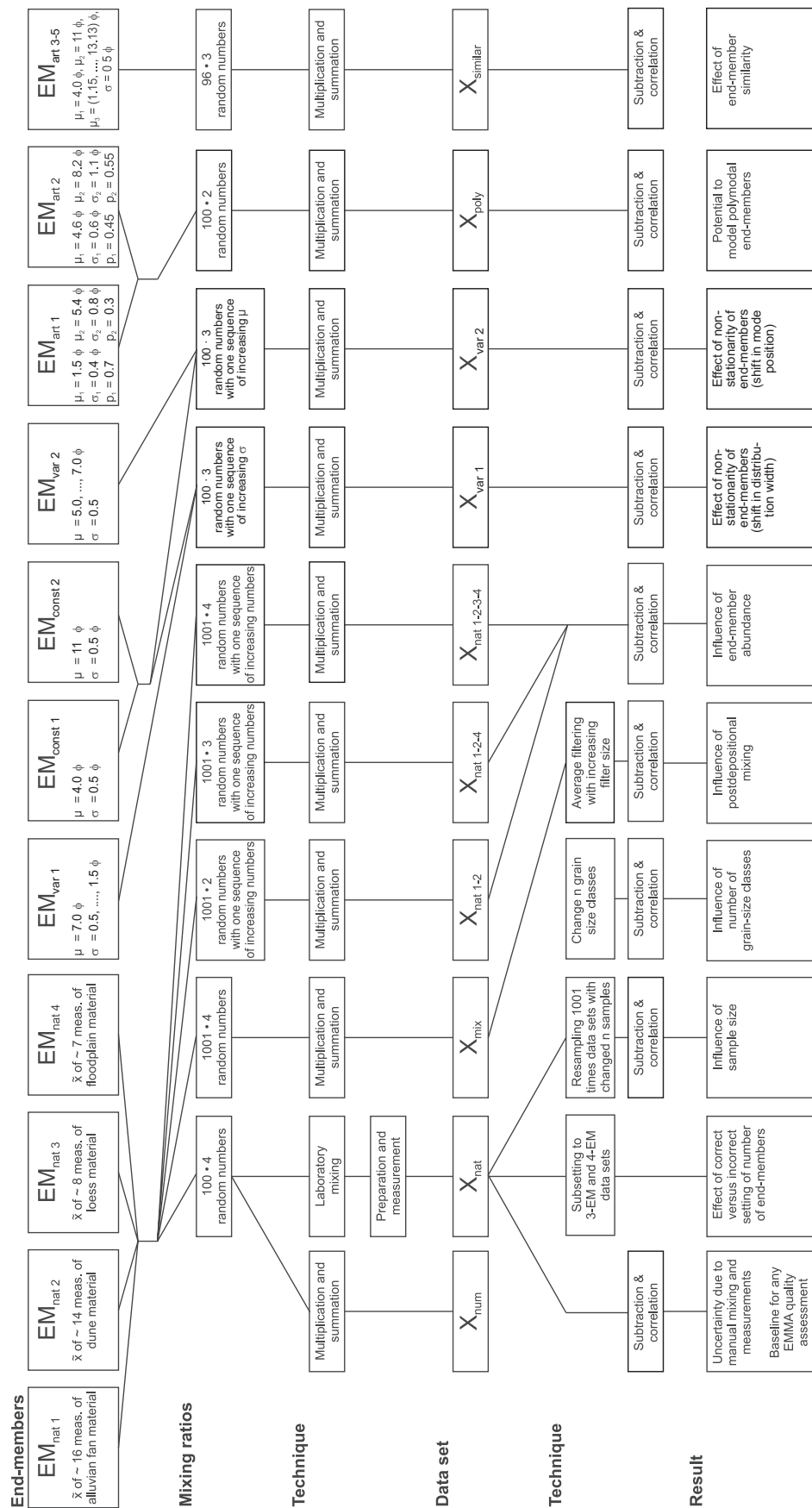


Fig. 1. Flow chart of input end-members, manipulation steps, data sets and resulting questions to explore the capabilities and limitations of EMMA.

interpolated to wider, equally spaced classes (in  $\phi$ -scale). This resulted in data sets ranging from 4 to 116 grain-size classes. Grain-size distributions were obtained by linear interpolation using the EMMAgeo function interpolate.classes. Deterministic EMMA was applied with  $q=4$  and  $l=0$ .

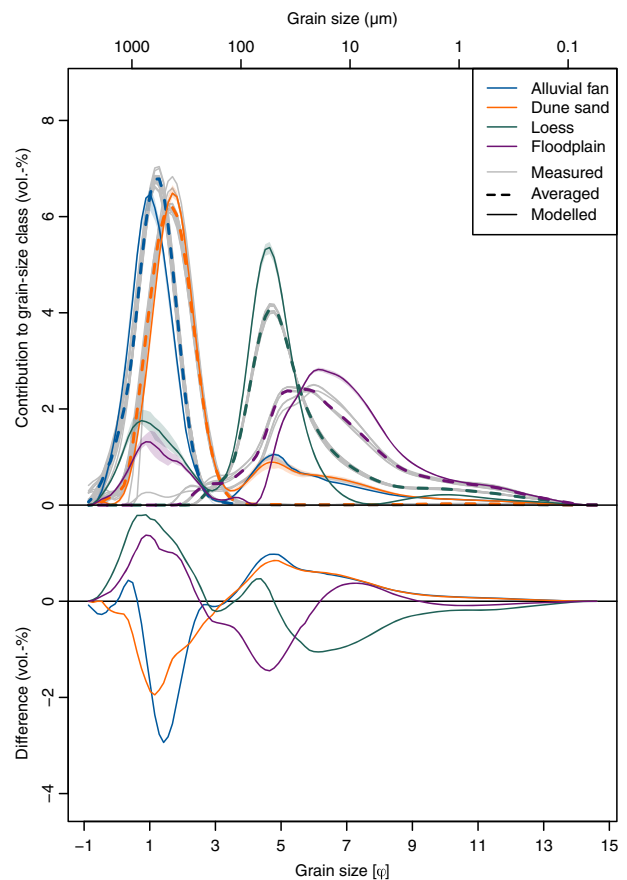
The effect of relative end-member abundance was tested by iteratively increasing the relative contribution of  $EM_{nat1}$  (from 0% to 100% in 1001 equally spaced steps) while assigning random numbers to the remaining end-members. This was done for  $EM_{nat1}$  and  $EM_{nat2}$  (2-EM scenario,  $X_{nat1-2}$ ), for  $EM_{nat1}$ ,  $EM_{nat2}$  and  $EM_{nat4}$  (3-EM scenario,  $X_{nat1-2-4}$ ) as well as for all four end-members (4-EM scenario,  $X_{nat1-2-3-4}$ ). The EMMAgeo function mix.EM was used with a noise level of 0.002 and an autocorrelation value of 5 (i.e. a running mean filter of size 5 is used to smoothen the resulting mixed samples). The noise introduction was necessary to avoid 'perfect' mixing and thus unrealistic results, especially for the 2-EM scenario. Deterministic EMMA was applied with  $q$  according to the respective scenario and  $l=0$ .

To simulate post-depositional mixing, an artificial data set ( $X_{mix}$ ) with 1001 randomly mixed samples was generated and row-wise averaging rectangular filters were applied to the grain-size distributions of adjacent samples. Filter width was increased from 1 to 99. This setting can be regarded as an accreting sediment section, which has been sampled at equal depth intervals 901 times and which has experienced post-depositional mixing of different vertical dimension during accretion. For EMMA, the resulting data sets were truncated to samples 50 to 950 to avoid filter boundary effects. Deterministic EMMA was applied with  $q=4$  and  $l=0$ .

Non-stationarity in the grain-size distribution of end-members was explored by end-members, which were defined by normal distributions:  $EM_{const1}$  with a mean of  $4\phi$  and a standard deviation of 0.5 and  $EM_{const2}$  with a mean of  $11\phi$  and a standard deviation of 0.5. To these two stable end-members a third end-member was added, in two scenarios. The scenario with non-stationary end-member  $EM_{var1}$  implies an end-member with a stable mode at  $7\phi$  and a changing standard deviation, illustrating a gradually widening distribution due to for example a decreasing sorting efficiency. The degree of change was ranged between zero and factor 3 (standard deviation grading from 0.5 to 1.5). The scenario with non-stationary end-member

$EM_{var1}$  included a stable standard deviation of 0.5 but a running mean, a case which is representative of a changing transport energy or source distance case. Here, the modal class position, hence the mean, was allowed to glide by a range from zero (mean constant at  $6\phi$ ) to two (mean gliding from  $5$  to  $7\phi$ ). For each of the 100 realizations of the two scenarios, 1001 data sets were computed and EMMA has been performed with  $q=4$  and  $l=0$ .

Multimodality is already visible in the natural end-members  $EM_{nat3}$  and  $EM_{nat4}$  (cf. Fig. 2). However, for consistent tests two clearly bimodal artificial end-members were created.  $EM_{art1}$  was defined by summing two normal



**Fig. 2.** Comparison of natural and modelled robust end-member loadings. The upper part shows an overlay of natural and modelled grain-size distributions. Individual measurement data (grey lines) and natural end-members (bold dashed lines) are rescaled to 80% for better visual comparison. Modelled end-members are shown as median (solid thin lines) and quartile range (thin polygons). The lower part shows deviation of modelled from natural end-members explicitly (natural end-members not rescaled).



distributions (relative amounts of 70% and 30%) with means of  $1.5 \phi$  and  $5.4 \phi$  and standard deviations of  $0.4 \phi$  and  $0.8 \phi$ , respectively.  $EM_{\text{art2}}$  was created in a similar way with relative amounts of 45% and 55%, means of  $4.6 \phi$  and  $8.2 \phi$  and standard deviations of  $0.6 \phi$  and  $1.1 \phi$ , respectively. The function `create.EM` from the `EMMAgeo`-package (Dietze & Dietze, 2019) was used. These two end-members were randomly mixed to create a data set ( $X_{\text{poly}}$ ) composed of 100 samples. Again, the function `mix.EM` was used with a noise level of 0.002 and an autocorrelation value of 5. Deterministic EMMA was applied with  $q=2$  and  $l=0$ .

The effect of end-member similarity was systematically tested by generating two 'stable' end-members as normal distribution functions (means of  $4 \phi$  and  $11 \phi$  and standard deviations of  $0.5 \phi$  in both cases). Then, another end-member was added with the same standard deviation, but its mean was changed from 1.15 to  $13.13 \phi$  along the size classes as defined by the laser diffraction device. A natural analogue for this test might be the attempt to decipher the contribution of two very similar dune sands to a deposit, simply based on their similar grain-size distribution. All end-members were mixed with random proportions. This resulted in 89 end-member scenarios ( $X_{\text{similar}}$ ), in which two end-members approach each other twice, overlap and then divert again from one another. Each scenario was used for the Monte Carlo approach as described above, using deterministic EMMA with  $q=3$  and  $l=0$ .

### Evaluation of EMMA tests

Comparison of grain-size distributions should ideally be based on robust statistics. This includes the centred log-ratio transform for compositional data (Aitchison, 1986). However, this transform can only be applied to zero-free data. In our case, especially for the natural end-members and modelled loadings there are many classes that contain zeros, which precluded application of the centred log-ratio transform. Accordingly, being aware of this bias, analysis remained with the untransformed data. The measured variability of the four natural end-members and the parallel mixed samples was described by the average classwise 5 to 95 percentile range.

Tests of further parameter influences were based on the comparison of the natural with modelled data sets. Paterson & Heslop (2015) used angular difference as a measure of

deviation. However, this value is not really intuitive and strongly depends on the size of the absolute values of the tested data, which is why the authors decided not to use it. Rather, two other measures of model quality were used: average variance explained by the model (represented by  $R^2$ ) and average model error ( $\tilde{E}$ , i.e. median of the absolute difference between model and input data). Both measures were calculated for  $EM_{\text{nat}}$  versus modelled end-member loadings ( $R_l^2$  and  $\tilde{E}_l$ ) and mixing ratios versus modelled end-member scores ( $R_s^2$  and  $\tilde{E}_s$ ). The explained variance  $R_l^2$  mainly reacts to shifts in mode positions and provides an easily interpretable measure that can be compared among different samples and data sets. Note, however, that while this correlation based method may represent classwise patterns adequately, the final interpretation of the vol-% values in a geoscientific sense may remain biased for high abundances of fine grain-size classes due to the logarithmically scaled class boundaries of many devices.  $\tilde{E}_l$  reacts to both shifts in the modes of individual grain-size distributions and differences in the volume percentages per class. It is well-suited to characterize the relationships of distributions that cover many classes but reaches limitations for narrow distributions (for example, single end-members). To provide a quantitative description of this behaviour the test variable value was indicated (for example, number of samples) when 95% of the maximum of  $R^2$  and 5% of the minimum of  $\tilde{E}$  were reached. Note that  $R^2$  and  $\tilde{E}$  as introduced above are different from  $R^2$  and  $E$  as delivered by the function `EMMA` from the R-package `EMMAgeo` (Dietze & Dietze, 2016) as standard output by comparing the input data set  $X$  with the modelled data set  $X'$ . The measures introduced above can only be used when the natural end-members and their mixing ratios are known.

## RESULTS

### Natural end-members and their mixtures

The four natural end-members show distinct, characteristic distributions (Fig. 2). Clearly unimodal are the first two deposits, alluvial fan ( $EM_{\text{nat1}}$ ) with a mode at  $1.3 \phi$  and dune ( $EM_{\text{nat2}}$ ) with a mode at  $1.6 \phi$ . The loess sample ( $EM_{\text{nat3}}$ ) shows a primary mode in the silt fraction ( $4.8 \phi$ ) and a secondary one in the clay fraction ( $10 \phi$ ). The floodplain deposit ( $EM_{\text{nat4}}$ ) is multimodal



with two dominant modes in the silt fraction ( $5.1 \phi$  and  $5.7 \phi$ ) and suppressed modes in the sand ( $2.8 \phi$ ) and clay ( $11 \phi$ ) fraction. Average classwise 5 to 95 percentile range in the individual measurements is 0.002, 0.010, 0.030 and 0.100% for  $EM_{nat1}$  to  $EM_{nat4}$ , respectively.

The numerically mixed data set  $X_{num}$ , composed of the four natural end-members, showed an overall high similarity to the hand-mixed and measured data set  $X_{nat}$  (Fig. 3). The average  $R^2$  is above 0.96 and average classwise deviation  $\tilde{E}$  is below 0.06 vol.-%. Further analyses were only performed with data from which outliers were removed. The latter were defined as being below the 5 percentile threshold of  $R^2$  or above the 95 percentile of  $\tilde{E}$  (blue and orange horizontal lines in Fig. 3). This resulted in rejection of five samples (IDs 29, 39, 53, 56 and 88). Mixed parallel samples yielded an average classwise 5 to 95 percentile range of 0.159 (IDs 7, 23, 37, 42), 0.074 (IDs 51, 63) and 0.004 (IDs 84, 90) %, respectively. Without an obvious outlier (ID 7), variability of the first parallel set reduced to 0.033.

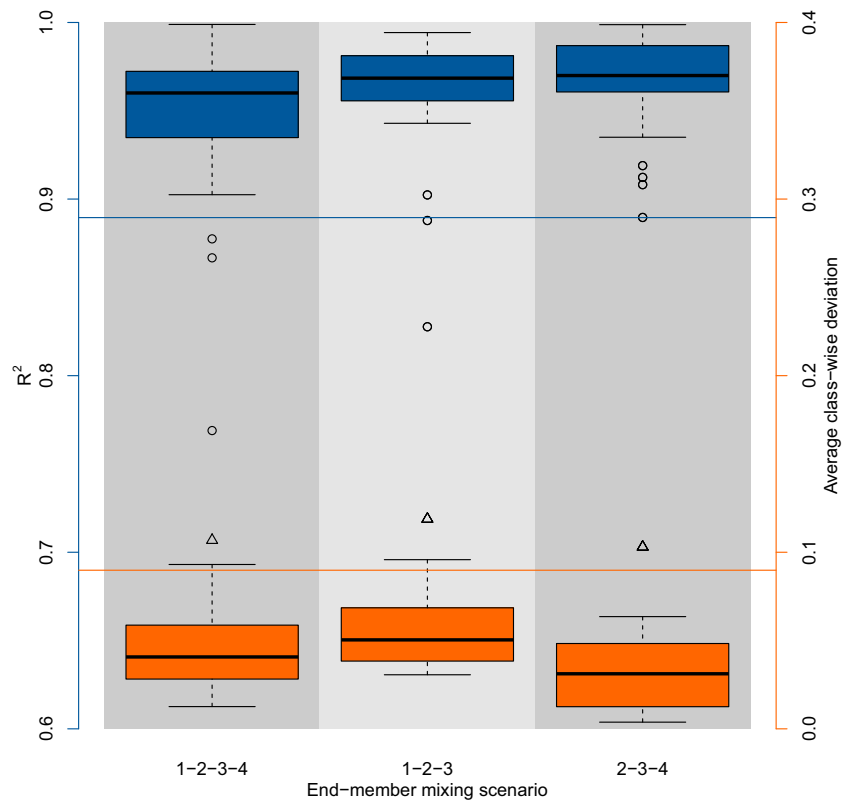
### Modelled end-members

Using the deterministic EMMA protocol with four end-members and no weight limit

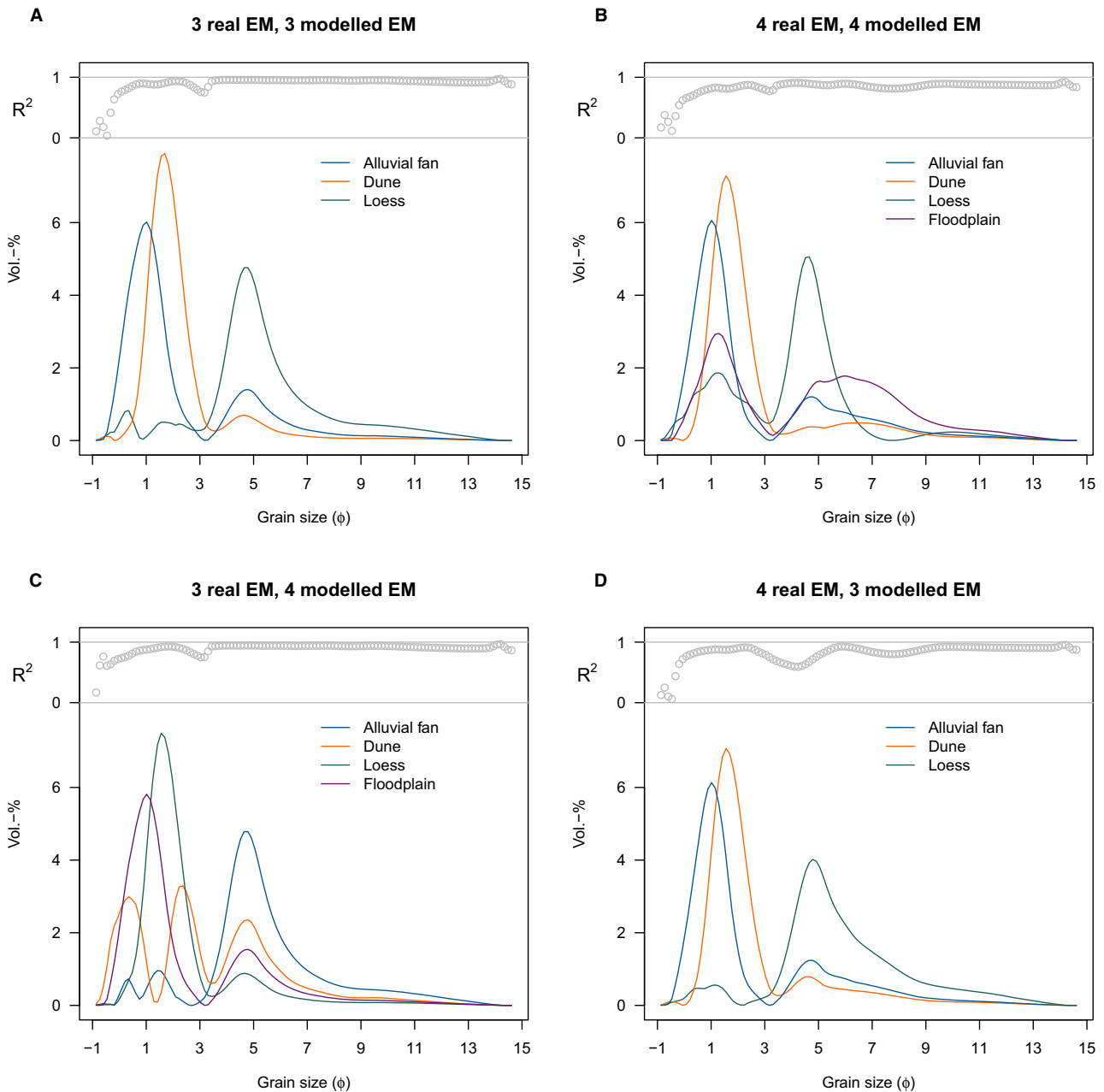
transformation,  $EM_{nat1}$  and  $EM_{nat2}$  show a minor shift of the mode position to coarser sizes by one grain-size class (Fig. 2, Appendix S1).  $EM_{nat4}$  shows a mode deviation of  $0.3 \phi$  (i.e. two grain-size classes). The average model  $R^2$  is 0.93. As can be already seen from the shapes of the grain-size data (Fig. 2),  $EM_{nat1}$  to  $EM_{nat3}$  are reproduced well whereas  $EM_{nat4}$  shows obvious deviation from its original shape. All modelled end-members show artificial secondary modes, preferentially below primary modes of other end-members. This results in underestimation of the volume percentage of respective primary modes by up to 3.2 vol.-% (cf. lower part of Fig. 2).

### Model parameterization

The effects of different choices of the number of end-members to model  $q$  (Fig. 4) are diverse. In the case of three existing components, which are modelled by three end-members, the resulting loadings are nearly identical with the EMMA results of the full data set (Fig. 2). However, if the same data set is modelled with four end-members (Fig. 4C) the additional end-member shows three individual modes ( $-0.2, 2.1, 4.7 \phi$ ) with relevant scores in only two



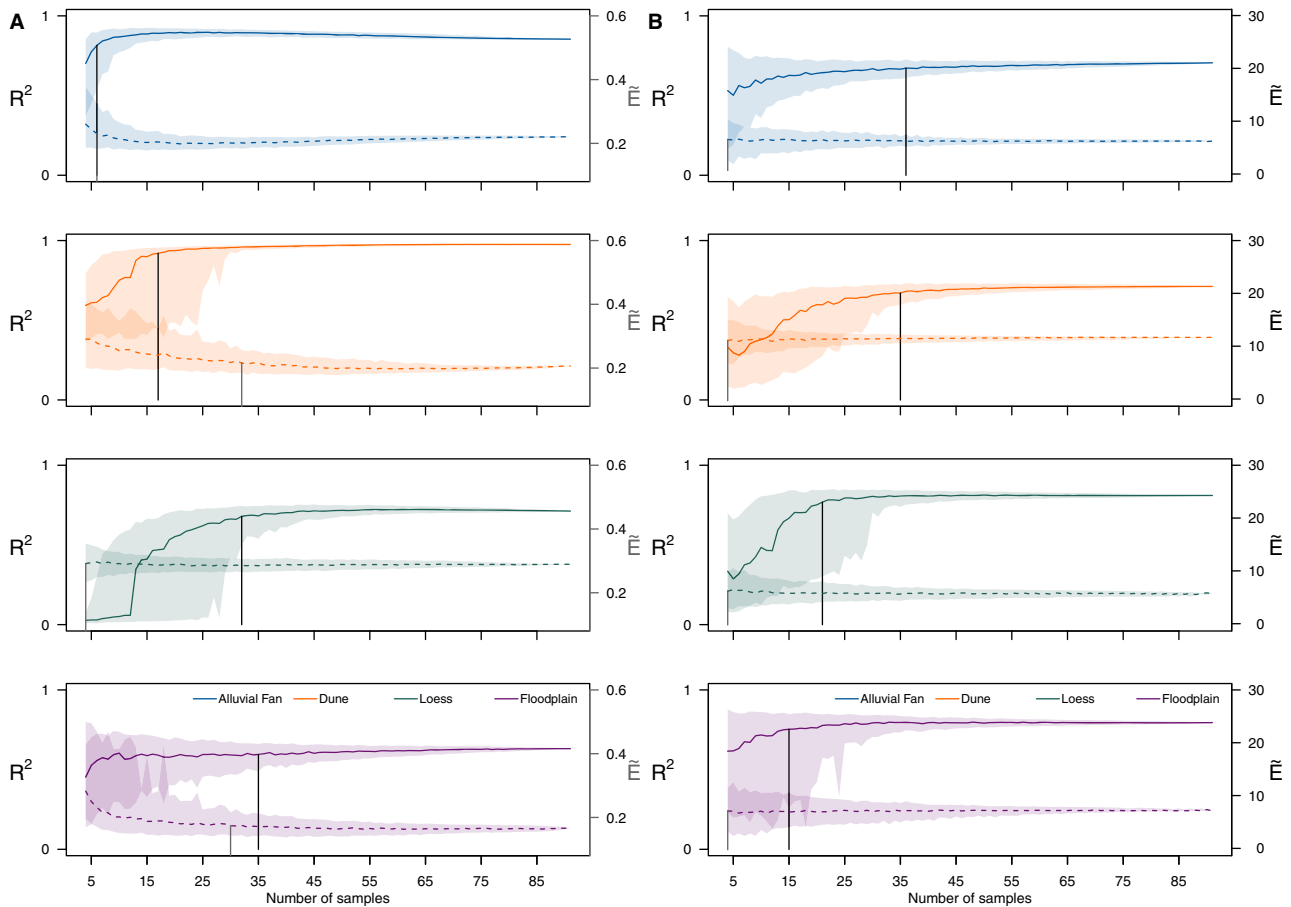
**Fig. 3.** Comparison of manually and numerically mixed samples. For all three subsets ( $X_{nat1-2-3-4}$ ,  $X_{nat2-3-4}$ ,  $X_{nat1-2-3}$ ) the explained variance (blue) and average classwise deviation (orange) is shown. Blue and orange horizontal lines show 5 and 95 percentile thresholds.



**Fig. 4.** Influence of correct versus incorrect parameterization. The four plots show loadings for correct (A, B) and incorrect (C, D) numbers of end-members. Grey dots indicate classwise explained variances of the models. Wrong parameters introduce either spurious end-members or deform end-members so that they are no longer geoscientifically meaningful.

samples (IDs 74, 83, not shown). In the case of four existing end-members this fourth end-member is correctly identified by EMMA (Fig. 4 B), although it is present in only one sample. However, the additional end-member exists not only in the affected sample, but causes scores up to 25% also in other samples (not shown). Describing a data set consisting of four end-

members by a model with only three end-members (Fig. 4D) introduces a bias to the existing end-members; specifically to the one end-member that is closest to the missing one (i.e. EM 3, dark green line in Fig. 4D). This end-member receives a broad shoulder between  $6\phi$  and  $8\phi$  and an elevated slope between  $9\phi$  and  $14\phi$ .



**Fig. 5.** Influence of number of samples on end-member loadings (A) and scores (B). Explained model variance  $R^2$  (solid line) and average error  $\bar{E}$  (dashed line) evolve towards stable values with increasing number of included samples. Vertical lines indicate number of samples required to reach 95% of maximum  $R^2$  or 5% of minimum  $\bar{E}$ . From 20 to 40 samples on both loadings and scores can be described appropriately.

### Data set dimensions

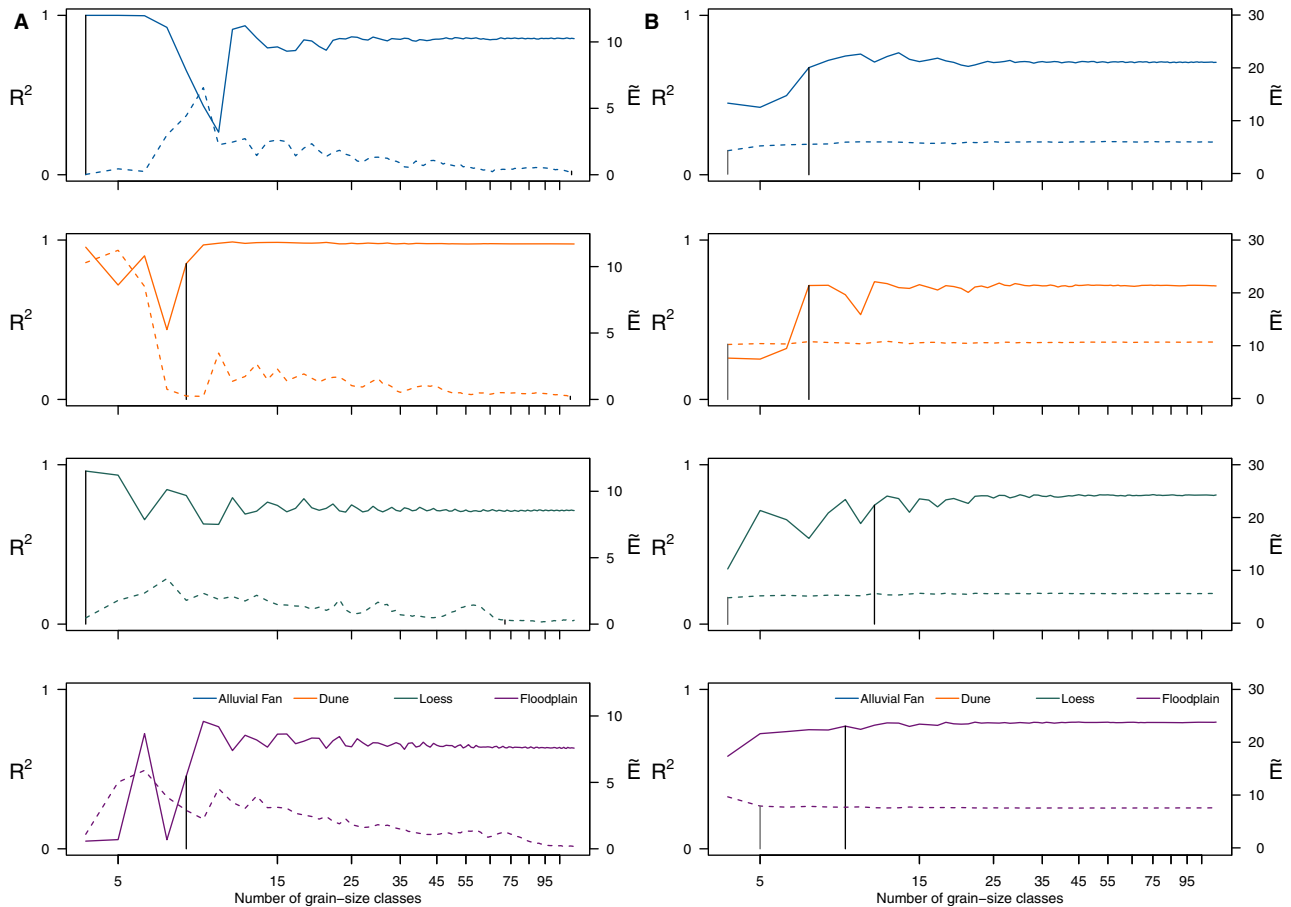
The influence of the number of samples, which are used for EMMA (Fig. 5) is high for small numbers but loses relevance as the size of the data set increases.  $R^2$  and  $\bar{E}$  as well as their associated scatters converge towards stable values. These stable regions are typically reached, when at least 20 to 40 samples are used for EMMA (vertical lines in Fig. 5). Numeric instabilities that were reported for an insufficient sample-to-class-ratio (Dietze *et al.*, 2012) were not encountered with the R package EMMAgeo. Performing the same test but using the numerically mixed data set  $X_{\text{num}}$  (see Appendix S1) reveals very similar patterns, except for very low numbers of samples. The approach of iteratively drawing more and more samples from a global data set to create a subset that eventually contains as many samples as the global data set

introduces a bias towards smaller scatter due to iteratively less variability in the individual elements of the subset. However, this effect is hard to quantify given that  $X_{\text{num}}$  showed less deviation overall and points at only marginal scatter for numerically mixed data sets in general.

The number of grain-size classes does not affect EMMA results in a dramatic way, except for very small numbers (Fig. 6). Explained model variance  $R^2$  (solid line) and average error  $\bar{E}$  (dashed line) typically reach stable values after less than 15 grain-size classes. This stability does not increase any more after 30 to 40 classes. Only  $\bar{E}_1$  decreases slightly with increasing number of classes.

### End-member abundance and mixing

The relative abundance of an end-member in a sample (Fig. 7) mainly has an impact at very



**Fig. 6.** Influence of number of classes on end-member loadings (A) and scores (B). Note logarithmic x-axis for better illustration of changes in small numbers of classes. Vertical lines indicate number of samples required to reach 95% of maximum  $R^2$  or 5% of minimum  $\tilde{E}$ . Solid lines depict  $R^2$ , dashed lines show  $\tilde{E}$ .

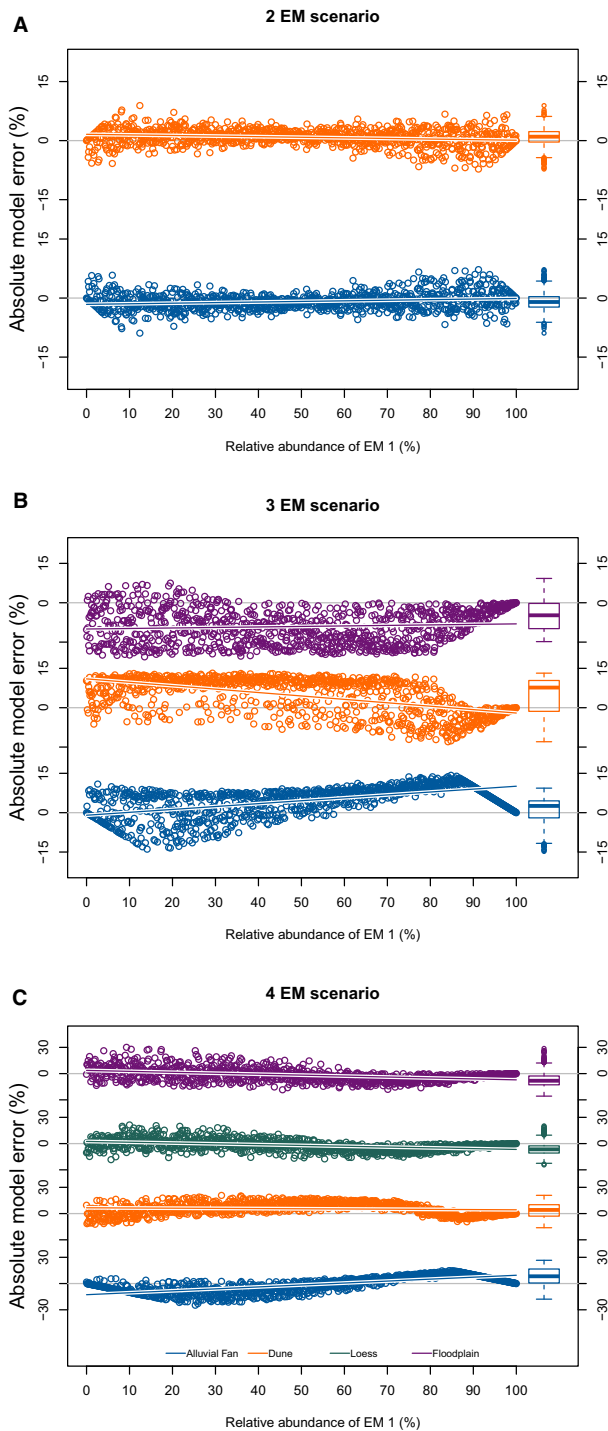
low or high contributions. As  $EM_{\text{nat}1}$  was systematically increased, the diagonal patterns and the white linear trend line of model deviation (scores) indicate that this end-member was underestimated for low abundances and overestimated for high abundances. The quartile ranges of modelled score errors (boxplots in Fig. 7) are around  $\pm 5\%$  for the 2-EM scenario,  $\pm 9\%$  for the 3-EM scenario and between  $\pm 8\%$  and  $\pm 15\%$  for the 4-EM scenario, which is comparable with the overall scatter in scores for  $X_{\text{nat}}$ .

Post-depositional mixing has no general influence on end-member loadings (Fig. 8A);  $R_l^2$  and  $\tilde{E}_l$  do not change significantly with increasing size of the mixing integral (i.e. the number of randomly mixed samples). In contrast, scores are severely affected. A mixture of only three samples reduces  $R_s^2$  already by approximately

50% and correlation of input data and model result is virtually zero after not more than 15 mixed samples.

### Non-stationarity of end-members

Gradual changes in the shape of one end-member can be at first order in two ways (and their combinations): distribution width and distribution mode position. Both tests revealed effects on end-member loadings and scores. For the width change range of  $EM_{\text{var}1}$  from 100 to 300%, loadings (Fig. 9A) are affected only marginally in terms of  $R_l$  ( $>0.97$  throughout) and to some extent in terms of  $\tilde{E}_l$  which showed a gradual increase in grain-size classwise deviation with increasing range of the widths that  $EM_{\text{var}1}$  was allowed to receive throughout the data set. A similar pattern is also visible for



**Fig. 7.** Influence of end-member abundance on scores modelling. Boxplots provide a summary of all 1001 individual model runs. Lines through data clouds are linear model fits.

end-member scores (Fig. 9B), with similar stable  $R_s$  values throughout and also gently increasing  $\tilde{E}_l$  values for all three end-members.

Non-stationarity in terms of mode position (Fig. 10) has more prominent consequences for the model results. For the two stable end-members  $EM_{\text{const1}}$  and  $EM_{\text{const2}}$  the effects are similar to the case above, although the decrease in  $R_l$  is more prominent, from 0.99 to 0.89. However,  $EM_{\text{var2}}$  decreases severely, from an  $R_l > 0.99$  to values around 0.23, in combination with a systematically rising  $\tilde{E}_l$  evolution. Almost the same general behaviour applies to the evolution of end-member scores (Fig. 10B), as well.

### End-member shape identification

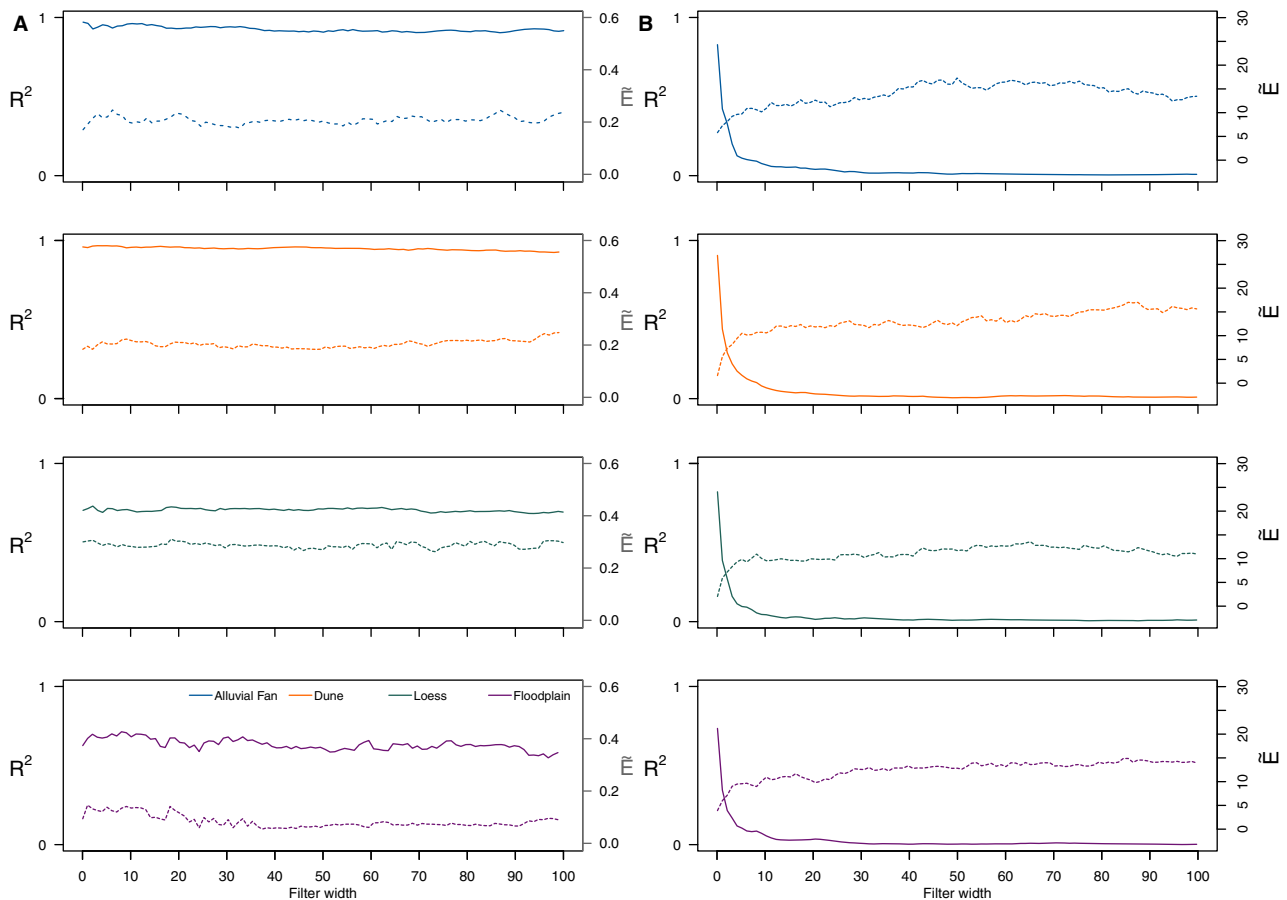
Modelling of multimodal end-members using  $X_{\text{poly}}$  (Fig. 11) yielded nearly optimal results regarding mode recovery and  $R_s^2$ .  $\tilde{E}_l$  is lower than for the mixed natural samples, whereas  $\tilde{E}_s$  and  $R_l^2$  are comparable to natural samples. Drawbacks occur for the modelling of the secondary modes; while the primary mode position is estimated well, the secondary modes deviate from the defined shapes especially where the grain-size curves of  $EM_{\text{art1}}$  and  $EM_{\text{art2}}$  overlap (i.e. 3.5 to 7  $\phi$ ).

End-member similarity test results (Fig. 12) show that, as the shifting end-member approaches the mode position of stable end-members (4  $\phi$  and 11  $\phi$ ),  $R_l^2$  decreases by up to 35%, both for the shifting and the respective stable end-member.  $R_s^2$  shows similar behaviour, although the decrease is more severe. The decrease is gradual, becoming significant for distances of less than three grain-size classes.  $\tilde{E}_l$  and  $\tilde{E}_s$  are more than one order smaller than in other tests.

## DISCUSSION

### Levels of certainty

The representation of sediment transport processes in measured data sets can be assessed by comparing the natural end-members, the mixed samples generated from them and the mixed parallel samples. Given that the errors introduced by amalgamating the samples in the laboratory and the influence of weight-percent versus volume-percent are negligible, all deviations of  $X_{\text{nat}}$  from  $X_{\text{num}}$  can be assigned to the measurement uncertainty, resulting from measurements of the natural end-member samples and the mixed samples. Natural end-members were reproduced with overall classwise scatter



**Fig. 8.** Influence of filter width (i.e. the number of randomly mixed samples) on end-member loadings (A) and scores (B). Solid lines depict  $R^2$ , dashed lines show  $\tilde{E}$ .

of 0.03%, although this value includes scatter in specific grain-size classes one order of magnitude higher and lower. However, with an average classwise deviation of 0.04% the mixed data sets are well in that range. Further insight into the composition of this uncertainty is possible through the parallel sets of mixed samples. These show an overall scatter of 0.04% (excluding the obvious outlier, ID 7) and also point at this apparent limit of certainty. All of these values are of the same magnitude as the measurement uncertainty of the laser diffraction device (2% relative error, applying to values of less than 7 vol.-%) and point at the base level of accuracy, which is passed to EMMA.

The apparently poor  $R_l^2$  values of the model results for the empirical data set can be attributed to the secondary modes introduced by EMMA (Fig. 2). This implies that the heights of primary modes are underestimated by roughly 20%, which needs to be taken into account when comparing EMMA results with other

measurement data. Although it would be possible to remove the secondary modes in the EMMA modelling protocol by manipulating the unscaled loadings, (cf. Appendix S1 in Dietze & Dietze, 2019) the model output has not been changed here to be as generic, transparent and conservative as possible. Likewise, further bias due to logarithmically scaled grain-size class boundaries inherent to many measurement devices needs to be accounted for when interpreting both the modelled vol.-% values and their geoscientific significance. Apart from  $EM_{nat4}$  the coincidence of mode position and shape of the end-members is obvious and points at the general ability of EMMA to correctly describe the grain-size distribution of end-members that constitute a set of mixed samples. Average classwise deviations of less than 0.52% underline this.

Our tests show systematic differences in the  $\tilde{E}$  values for loadings, depending on whether measured or synthetic end-members were used to

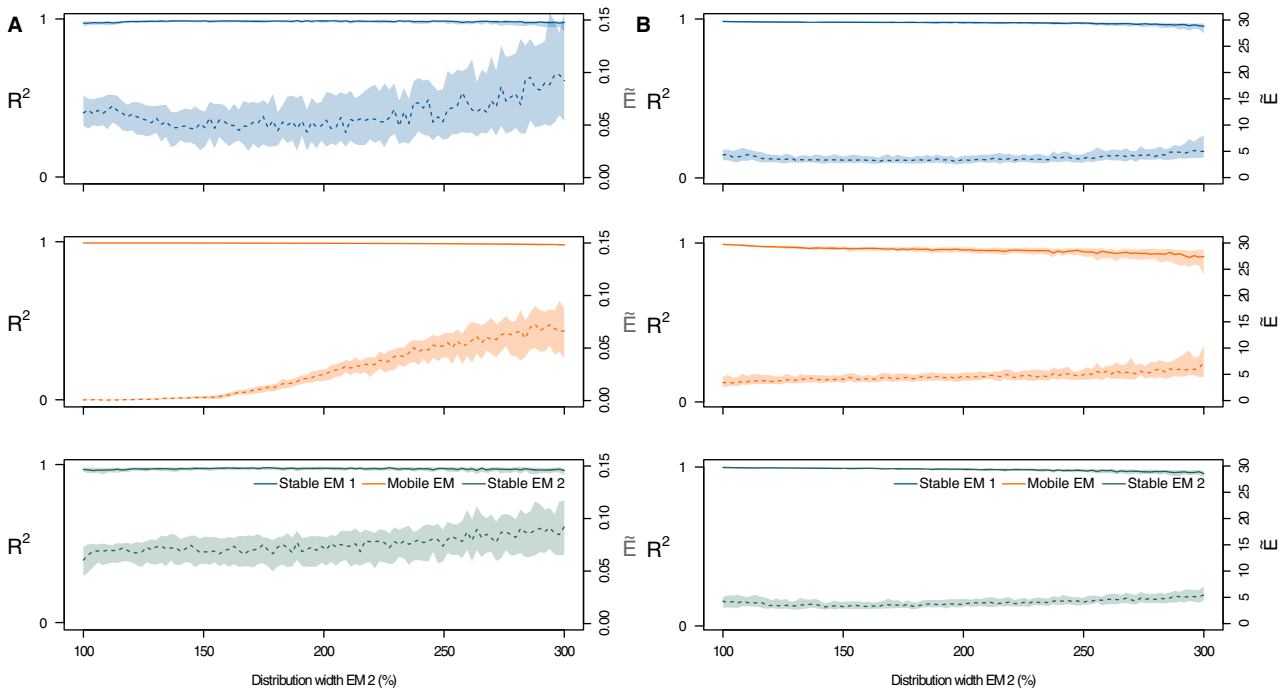


Fig. 9. Effect of one gradually widening end-member on modelled loadings (A) and scores (B).

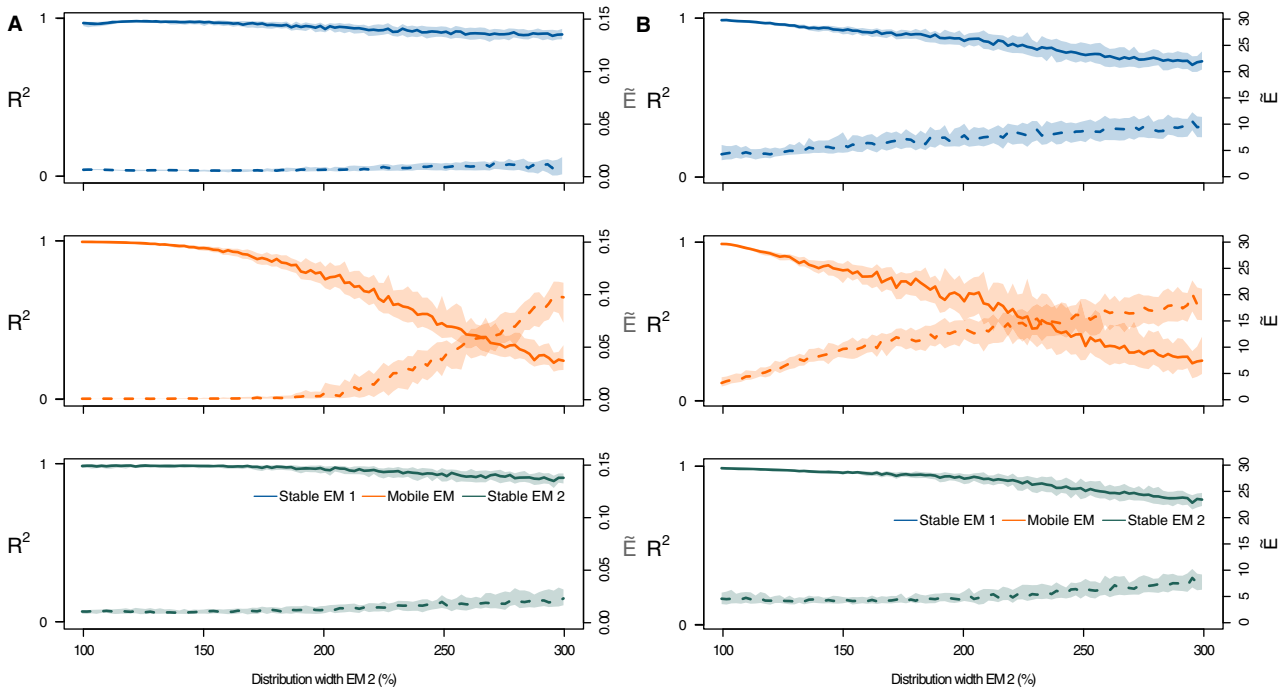
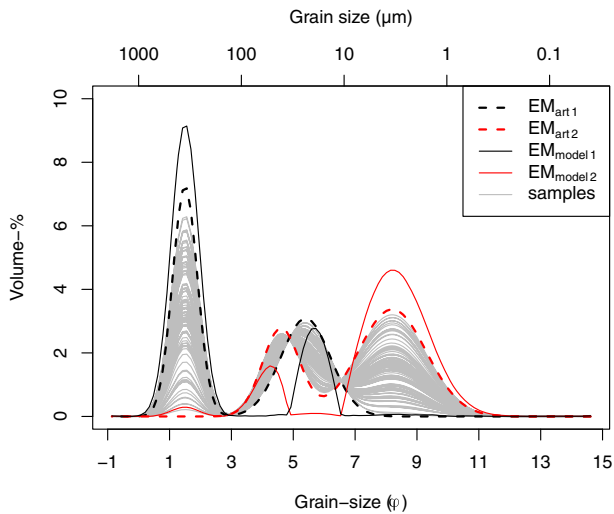


Fig. 10. Effect of one gradually shifting end-member mode position on modelled loadings (A) and scores (B).

create mixed data sets. Data sets based on measured end-members show an overall higher variability because end-members with a primarily

wider distribution were mixed, versus the narrow standard deviations of normal distributions for the synthetic end-members. Accordingly,  $\bar{E}$



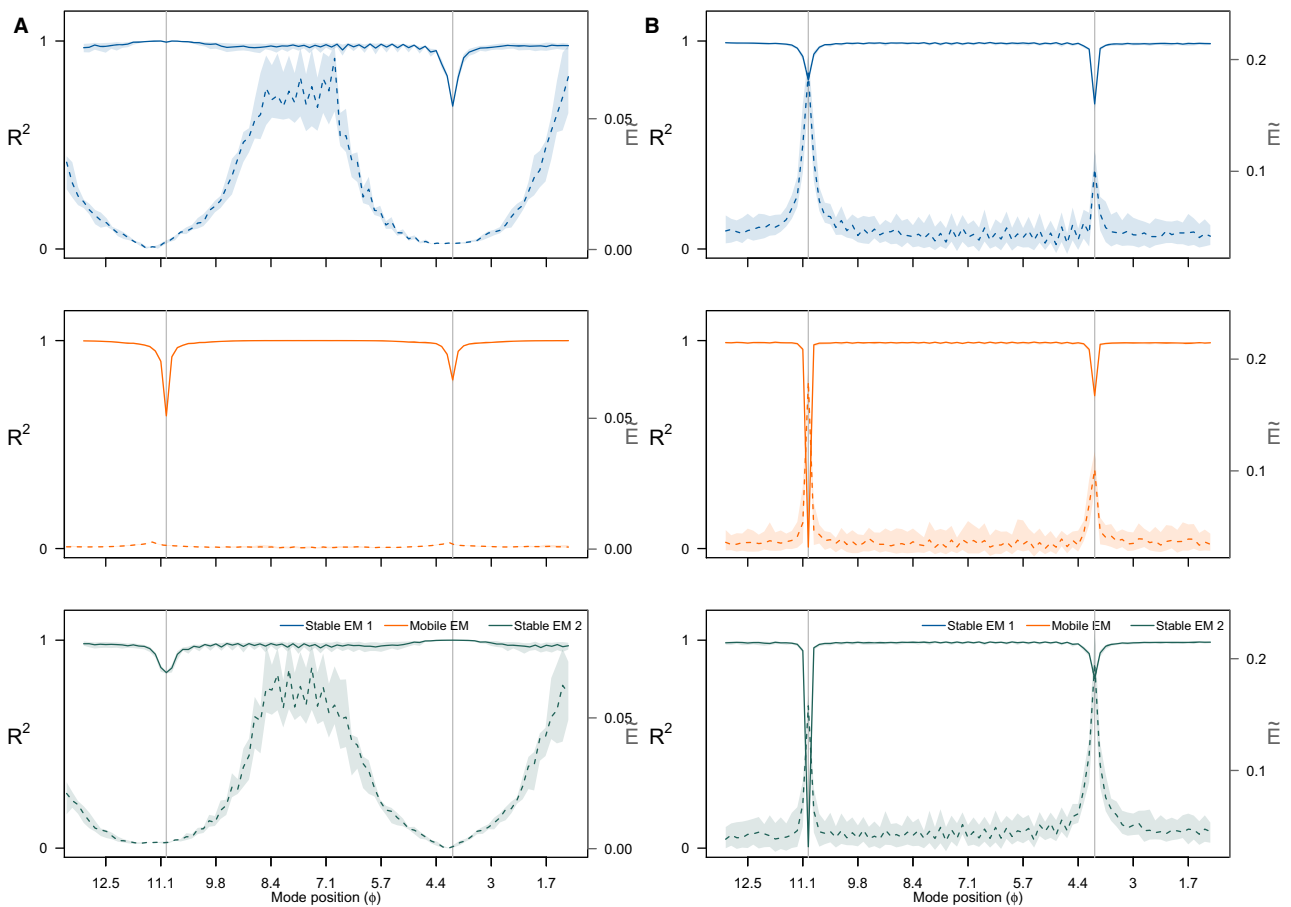


**Fig. 11.** End-member model outputs for two multimodal end-members.

values for loadings reflect the average behaviour of systematically different numbers of grain-size classes and the two categories of data sets cannot be compared directly.

### Model parameterization

Finding a meaningful number of end-members  $q$  is crucial, although almost never a straightforward task. The limits from a mathematical point of view are determined by the number of grain-size samples and classes. EMMA delivers by definition non-unique solutions. Hence, it remains to the scientist to judge the validity of the parameters and respective model outputs, based on a thorough understanding of the sedimentary and geomorphic system and reasonable links between the shape of modelled end-member loadings and supposed underlying



**Fig. 12.** Effect of end-member similarity on end-member loadings (A) and scores (B). Quality measures symmetrically degrade when the mobile end-member approaches the stable end-members (depicted by the grey lines at  $4 \phi$  and  $11 \phi$ ). Note that  $\tilde{E}$  is one order of magnitude lower than for other tests.

processes. One rather obvious effect of a mismatching estimate of  $q$  is shown in Fig. 4C, where the surplus end-member 4 exhibits a multimodal distribution. This is meaningless in a geoscientific context, because there is no process that would tend to consistently generate such a distribution, especially when the three other end-members are already present in the data set. Less obvious is the effect of an insufficient number of end-members for a data set (Fig. 4D), for example when end-member 3 partly covers the apparently necessary floodplain end-member by showing a pronounced shoulder towards finer grain-sizes. It might indeed be possible that the multimodal end-member 3 represents a true process end-member, for example when loess and floodplain sediments are reworked and transported by overland flow to the sampled deposit (Vandenbergh, 2013). The R package EMMAgeo provides a series of tools to investigate such a potential scenario (for example,  $q_{\min}$ -plot,  $q$ - $l$ -image, mode histograms, robust EMMA; cf. (Dietze & Dietze, 2019) and related Appendix S1).

### Data set dimensions

Converging measures of model quality for increasing numbers of samples underline the necessity to apply EMMA only to data sets with a sufficient sample size: 95% of the maximum  $R^2$  and minimum  $\tilde{E}$  values are reached between less than 10 and up to 47 samples. This indicates that there is no universal threshold of a minimum sample size that can be recommended. Presumably, the number of necessary samples depends on the number of inherent end-members and on how well the samples represent different mixing proportions of these end-members.

The number of available grain-size classes has comparatively little influence. For the four sediment samples used in this study, less than 15 classes are sufficient to achieve stable results in terms of  $R^2$ , whereas the more classes that are available the smaller the average absolute deviation of end-member loadings becomes. End-member scores are virtually unaffected. Hence, for the samples used in this data set, classic sieve-sedimentation analysis with perhaps a few more than the seven standard classes would already have been sufficient to successfully apply EMMA. However, similar to the appropriate number of samples, also the grain-size class limits must resolve the grain-size ranges, where

the distributions of different end-members intersect and overlap. For example, when primary modes of  $EM_{\text{nat1}}$  and  $EM_{\text{nat2}}$  differ between 1  $\phi$  and 2  $\phi$ , sieve intervals need to resolve this range with greater detail.

### End-member abundance and mixing

EMMA is able to detect an end-member if it is present in only one sample (Fig. 4B). Furthermore, the presence of this end-member is already sufficient to introduce a significant bias to the model. Hence, any data set should be inspected carefully for potential outliers in terms of samples that include end-members, which are not part of the sedimentary environment to be modelled. This task can be easily performed by checking the end-member scores: a suspicious end-member only appears in one or very few samples and the explained variance of each end-member for the data set ( $M_{\text{qsvr}}$ ) of the suspicious end-member is marginal. If it is reasonable to assume that the outlier is part of the sedimentary regime, the deposit should be resampled (if possible) to increase sampling resolution. Otherwise, the sample should be excluded from the data set.

End-members are identified by EMMA even if they are present with only a few vol.-%. However, the model tends to underestimate it in such cases (Fig. 7 left parts), whereas at the other end, i.e. when the end-member dominates a sample, EMMA overestimates its contribution. Weltje & Prins (2007), Paterson & Heslop (2015) and van Hateren *et al.* (2018) found similar tendencies in their exercises with artificial end-members.

Post-depositional mixing can introduce a significant bias to end-member scores, even for low mixing intensities (i.e. number of adjacent samples used for random mixing). Hence, if post-depositional mixing can be expected, for example based on descriptive field data such as ice wedges, cryoturbation features or crotovina, any process quantification based on end-member scores may be flawed. On the other hand, end-member loadings appear not to be seriously affected by post-depositional mixing. Thus, interpretation of the formerly involved sediment transport and deposition processes is still possible.

### Non-stationarity in end-member properties

End-member stationarity (i.e. stability in the distribution shape of all underlying end-members

throughout) is one of the key assumptions of EMMA (Dietze *et al.*, 2012) and its validity should therefore be considered in any real-world data set. Our tests show that a gradual widening of an end-member throughout the data set (i.e. for example a gradual change in the efficiency of a process to sort the grain-size of a given end-member as the deposit is being formed) affects the model results only marginally, most obviously expressed by  $\tilde{E}$  of loadings, which reflect not only shifts in mode positions but also changes in the overall shape (Fig. 9). Nevertheless, gradually widening distributions of contributing components do not render the interpretation of both loadings and scores flawed.

In contrast, non-stationary in the form of gradually shifting mode positions (Fig. 10) does have significant consequences for model results, and thus for the interpretation of the data in real application scenarios.  $R^2$  decreases to values as low as 0.23 for both loadings and scores as the mode is allowed to change within a data set by as much as  $2\phi$  units, i.e. from 7.8 to 31.3  $\mu\text{m}$  (15 measured grain-size classes). This range is certainly dramatic for archives such as loess deposits and beyond the level of variability in transport mechanisms one would expect (e.g. Vandenberghe, 2013). Nevertheless, other systems with more variable transport mechanisms may be subject to such wide ranges in end-member modes.

Apart from these two fundamental cases, non-stationarity can also arise from further mechanisms, such as the emergence of secondary modes due to increasing weathering and mineral neof ormation downcore. Their effects were however not tested here, but similar test scenarios can be constructed and evaluated in order to gain insight to more specific cases. See the Appendix S1 for detailed information on how to prepare such tests.

### End-member shape identification

Originally multimodal end-members were modelled by EMMA in the natural as well as in the artificial data set (Figs 4B and 11). Both artificial end-members were appropriately recognized and modelled, apart from the mismatching range, where the secondary modes overlap.

As original end-members become more and more similar to one another, the chance to adequately depict them and quantify their contribution to samples by EMMA decreases. However, the effect is not relevant until the mode positions of end-members with the same shape are

less than three grain-size classes apart from each other. Hence, EMMA is able to detect even small changes in the sediment transport regime as expressed by slightly changed grain-size distributions.

### CONCLUSIONS

End-member modelling analysis (EMMA) is a robust, reliable tool to identify and quantify sediment transport regimes from mixed sedimentary deposits. Thereby, the number of parameters is minimal. Virtually, the user will mainly adjust the number of end-members  $q$  and the weight transformation limit  $l$ . In this study, all tests were based on four natural end-members, which represent typical sediment types found in terrestrial systems and which show extensively overlapping grain-size distributions. Nevertheless, mathematically robust decomposition reassurance does not replace the need for geoscientific expert knowledge when it comes to interpreting the meaningfulness of the results in the context of source area, transport pathway and depositional environment.

EMMA can model end-member loadings at the level of measurement uncertainty. The task of appropriately defining the input parameters remains in the responsibility of the scientist who should make use of ample available tools and tests. Although EMMA is able to deliver results close to the optimum, well below the dimensions of the original data set (100 samples, represented by 116 grain-size classes), there is no general recommendation regarding sample size and resolution in terms of grain-size classes. The emergence of secondary modes limits direct comparison of modelled end-members with original data (for example, sampled alluvial fan or dune sediment samples) in terms of absolute vol.-% per class, but these modes have a limited effect on the estimates of end-member scores. EMMA tends to underestimate low end-member abundances and overestimate high abundances, but detects an end-member even if it is present in only one sample. Post-depositional mixing has severe consequences for end-member scores interpretation (process quantification) whereas loadings (process identification) appear to be nearly unaffected. EMMA is able to model primarily multimodal end-members and decipher end-members of identical shape as soon as their modes are separated by more than three grain-size classes.

## ACKNOWLEDGMENTS

We thank Gert Jan Weltje and an anonymous referee for their encouraging and improving thoughts. We thank Thomas Hösel and Claudia Ziener for preparing the mixed samples and Sascha Meszner for discussions and the loess sediment samples. Jens Turowski is thanked for valuable comments on an earlier version of the manuscript. Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

The underlying data sets are provided under <https://doi.org/10.5880/GFZ.4.6.2019.002>. The analysis scripts are provided in the Appendix S1.

## REFERENCES

- Aitchison, J. (1986) *The Statistical Analysis of Compositional Data*. Chapham and Hall, London.
- Buurman, P., Pape, T. and Muggler, C. (1997) Laser grain-size determination in soil genetic studies. 1. Practical problems. *Soil Sci.*, **162**, 211–218.
- Dietze, E. and Dietze, M. (2019) Grain-size distribution unmixing using the R package EMMAgeo. *E&G Quatern. Sci. J.*, **68**, 29–46.
- Dietze, E., Hartmann, K., Diekmann, B., Imker, J., Lehmkuhl, F., Opitz, S. and Borchers, A. (2012) An end-member algorithm for deciphering modern detrital processes from lake sediments of Lake Donggi Cona, NE Tibetan Plateau, China. *Sed. Geol.*, **243–244**, 169–180.
- Dietze, E., Maussion, F., Ahlborn, M., Diekmann, B., Hartmann, K., Henkel, K. and Haberzettl, T. (2014) Sediment transport processes across the Tibetan Plateau inferred from robust grain-size end members in lake sediments. *Clim. Past.*, **10**, 191–106.
- Dietze, M. and Dietze, E. (2016) EMMAgeo: End-member modelling of grain-size data [Computer software manual]. R Package Version 0.9.6.
- Dietze, M., Dietze, E., Lomax, J., Fuchs, M., Kleber, A. and Wells, S.G. (2016) Environmental history recorded in Aeolian deposits under stone pavements, Mojave Desert, USA. *Quatern. Res.*, **85**, 4–16.
- Flemming, B.W. (2007) The influence of grain-size analysis methods and sediment mixing on curve shapes and textural parameters: Implications for sediment trend analysis. *Sed. Geol.*, **202**, 425–435.
- Hartmann, D. (2007) From reality to model: Operationalism and the value chain of particle-size analysis of natural sediments. *Sed. Geol.*, **202**, 383–401.
- van Hateren, J.A., Prins, M.A. and van Balen, R.T. (2018) On the genetically meaningful decomposition of grain-size distributions: A comparison of different end-member modelling algorithms. *Sed. Geol.*, **375**, 49–71.
- International Organization for Standardization (2002) *Soil Quality – Determination of Particle Size Distribution in Mineral Soil Material -- Method by Sieving and Sedimentation Soil (Volume 2002; Standard)*. International Organization for Standardization, Geneva.
- International Organization for Standardization (2009) *Particle Size Analysis – Laser Diffraction Methods (Vol. 2009; Standard)*. International Organization for Standardization, Geneva.
- Meszner, S., Kreutzer, S., Fuchs, M. and Faust, D. (2013) Late Pleistocene landscape dynamics in Saxony, Germany: Paleoenvironmental reconstruction using loess-paleosol sequences. *Quatern. Int.*, **296**, 94–107.
- Miller, B. and Schaetzl, R. (2012) Precision of soil particle size analysis using laser diffractometry. *Soil Sci. Soc. Am. J.*, **76**, 1719.
- Özer, M., Orhan, M. and Isik, N.S. (2010) Effect of particle optical properties on size distribution of soils obtained by laser diffraction. *Environ. Eng. Geosci.*, **16**, 163–173.
- Paterson, G.A. and Heslop, D. (2015) New methods for unmixing sediment grain size data New methods for unmixing sediment grain size data. *Geochem. Geophys. Geosyst.*, **16**, 4494–4506.
- Pye, K. and Blott, S. (2004) Particle size analysis of sediments, soils and related particulate materials for forensic purposes using laser granulometry. *Forensic Sci. Int.*, **144**, 19–27.
- Roberson, S. and Weltje, G. (2014) Inter-instrument comparison of particle-size analysers Inter-instrument comparison of particle-size analysers. *Sedimentology*, **61**, 1157–1174.
- Schulte, P., Lehmkuhl, F., Steininger, F., Loibl, D., Lockot, G., Protze, J. and Stauch, G. (2016) Influence of HCl pretreatment and organo-mineral complexes on laser diffraction measurement of loess-paleosol-sequences. *Catena*, **137**, 392–405.
- Sweeny, M.R., McDonald, E.V. and Markley, C.E. (2013) Alluvial sediment or playas: what is the dominant source of sand and silt in desert soil vesicular A horizons, southwest USA. *J. Geophys. Res., Series F*, **118**, 1–19.
- Vandenberghe, J. (2013) Grain size of fine-grained windblown sediment: A powerful proxy for process identification. *Earth-Sci. Rev.*, **121**, 18–30.
- Weltje, G.J. (1997) End-member modeling of compositional data: Numerical-statistical algorithms for solving the explicit mixing problem. *Math. Geol.*, **29**, 503–549.
- Weltje, G.J. and Prins, M.A. (2003) Muddled or mixed? Inferring palaeoclimate from size distributions of deep-sea clastics. *Sed. Geol.*, **162**, 39–62.
- Weltje, G.J. and Prins, M.A. (2007) Genetically meaningful decomposition of grain-size distributions. *Sed. Geol.*, **202**, 409–424.

Manuscript received 31 March 2021; revision 13 July 2021; revision accepted 29 July 2021

## Supporting Information

Additional information may be found in the online version of this article:

Appendix S1. EMMA evaluation.