This Letter to the Editor is to entice continued interest in Bayesian sediment fingerprinting technology including the assessment of various statistical test and uncertainty quantification methods [1]. Recent contribution is discussed since the pioneering work of Fox and Papanicolaou [2] on the application of Bayesian statistical model using model parameters represented by multiple probability distribution functions.

Excessive sediment delivery control programs require reliable estimation of relative sediment contributions from different sources. Nosrati et al., [3] used a modified Bayesian mixing model to assess the contributions of sources to sediment erosion by using composite signature-based source fingerprinting procedure. They quantified the relative uncertainties associated with sediment contribution estimations from the sources. The study was carried out in a mountainous terrain located at north-eastern Iran by collecting geochemical tracers. Nosrati et al., [3] determined the significantly non-conservative tracers by applying a standard bracket (range) test following Foster and Lees [4]. Tracers whose concentrations at downstream target samples fell outside their ranges in upstream sources were disregarded (or failed the test).

The use of statistical measures to identify composite fingerprints to estimate source contributions is well-known [5-8] and regarded as a useful method for source prediction. Use of multiple statistical method is advantageous over using the same test multiple times to facilitate multidimensional analysis and considered more informative. Three different statistical methods used by Nosrati et al., [3] to identify composite signatures for source discriminations involved:

(1) The Kruskal-Wallis H test (KW-H)
(2) A combination of KW-H and Discriminant Function Analysis (DFA) and
(3) A combination of KW-H and Principal Components and Classification Analysis (PCCA).

The KW-H test is a rank-based test which is similar to one-way ANOVA test [9]. This can compare multiple groups based on the null hypothesis: different groups are drawn from the same distribution/median. Statistically significant tracers as determined through KW-K test were passed in DFA analysis (and also in the PCCA analysis) to determine set of weights for source group discrimination. The weight provides information about an individual that can possibly come from a potential source. The statistical significance of discriminant functions were evaluated through the use of eigenvalue, canonical correlation, Wilks' lambda, and squared Mahalanobis...
tests. The methods considered the sources as dependent and tracers as independent variables. In the PCAA analysis, Principal Components (PC) that had eigenvalues >1 were considered for analysis and varimax rotation technique was used to minimize the tracer numbers with high loading on each PC. Weight values were estimated for tracers in each PC and was a measure of the contribution of the tracer in the PC composition. The tracers with high weight values were considered important and chosen for final composite tracer if they were not highly correlated. For highly correlated tracers, the tracer with highest absolute PC loading value was considered in the final composite tracer. The significant difference check of different sources was done using one-way ANOVA (F-test) and Tukey HSD post hoc tests.

A proposed modified Bayesian model [10] termed modified MixSIR quantified the uncertainties associated with relative contributions of sediment sources by using composite tracers. Making use of multiple composite tracers is important as it can estimate the uncertainties arising from different fingerprinting sets [2,11]. The three steps involved in modified MixSIR are: 1) Generation of prior probability distribution of model parameters; 2) Estimation of likelihood function of the statistical model and 3) Adjustment of prior distributions to derive posterior distributions of model parameters.

According to the Bayesian rule, the posterior distributions of all relative contributions $f_i$ from each source $i$ is proportional to the product of prior distribution and likelihood divided by their sum.

$$P(f_i | \text{data}) = \frac{L(\text{data} | f_i) \times p(f_i)}{\sum_j L(\text{data} | f_j) \times p(f_j)} \quad (1)$$

In equation 1, $L(\text{data} | f_j)$ is the likelihood of data given the relative source contribution $f_j$ of $q$ vectors. The prior probability is expressed as $p(f_j)$ of the true given state base on prior information.

The mean and standard deviation of the model are measure of relative source contributions. For each source $i$ and the final sets of composite fingerprints $j$

The mean is $\bar{f}_j = \sum_{i=1}^{n} (f_{ji} \times m_{sourcei}) \quad (2)$

The standard deviation is $\sigma_j = \sqrt{\sum_{i=1}^{n} (f_{ji}^2 \times S_{sourcei}^2)} \quad (3)$

In equation 3, $S_{sourcei}^2$ is the variance of the $i^{th}$ sediment in the $j^{th}$ source.

The likelihood of data given the proposed sediment mixture based on final composite fingerprint mean and standard deviation is expressed as

$$L(\mathbf{x} | \bar{f}_j, \sigma_j) = \prod_{k=1}^{n} \prod_{l=1}^{q} \frac{1}{\sigma_j \sqrt{2\pi}} \times \exp \left(-\frac{(x_{kl} - \bar{f}_j)^2}{2\sigma_j^2}\right) \quad (4)$$

In equation 4, $X_{kl}$ is the property of $j$th tracer from the $k^{th}$ sediment sample.

Nosrati et al., [3] generated random samples from the posterior distribution of the estimated mixture following Sampling-Importance-Resampling (SIR) method [12] to create a cut-off acceptance value before sampling and to resample. A two sample Kolmogorov-Smirnov test confirm the statistically significant differences in the posterior distributions of relative contributions [13].

The relative contributions are also estimated using the Root Mean Square Difference (RMSD) method for comparison purpose:

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{n} (Y_{i1} - Y_{i2})^2}{n}} \quad (5)$$

In equation 5, $Y_{i1}$ and $Y_{i2}$ are relative contributions from a source, $i$ that is based on composite tracers.

The results show that the relative contributions estimated by KW-H, DFA and PCAA are comparable. Kolmogorov-Smirnov pairwise comparison test confirmed statistically different posteriors of predicted relative contributions from the sources. The advancement in sediment fingerprinting methods described by Nosrati et al., [3] can be used in other study areas for land management and erosion control purposes [1].

References

1. Ahmed I, Karim A, Boutton TW, Strom KB (2019) On the Prediction of Uncertainty in a Sediment Provenance Model. Res Adv Environ Sci 2019: 45-60.
2. Fox JF, Papanicolaou AN (2008) An un-mixing model to study watershed erosion processes. Adv Water Resour 31: 96-108.
3. Nosrati K, Collins AL, Madankan M (2018) Fingerprinting sub-basin spatial sediment sources using different multivariate statistical techniques and the Modified MixSIR model. Catena 164: 32-43.
4. Foster ID, Lees JA (2000) Tracers in Geomorphology: theory and applications in tracing fine particulate sediments. In: Foster IID (Ed). Tracers in Geomorphology, Wiley & Sons, Chichester, UK. Pg no: 3-20.
5. Collins AL, Pulley S, Foster IDL, Gellis A, Porto P, et al. (2017) Sediment source fingerprinting as an aid to catchment management: A review of the current state of knowledge and a methodological decision-tree for end-users. J Environ Manage 194: 86-108.

6. Gellis AC, Noe GB (2013) Sediment source analysis in the Linganore Creek watershed, Maryland, USA, using the sediment fingerprinting approach: 2008 to 2010. J Soils Sediments 13: 1735-1753.

7. Owens PN, Blake WH, Gasper L, Gateuille D, Koiter AJ, et al. (2017) Fingerprinting and tracing the sources of soils and sediments: Earth and ocean science, geo-archaeological, forensic, and human health applications. Earth Sci Rev 162: 1-23.

8. Walling D (2013) The evolution of sediment source fingerprinting investigations in fluvial systems. J Soils Sediments 13: 1658-1675.

9. Dytham C (2011) Choosing and Using Statistics: A biologist’s Guide. John Wiley & Sons, New Jersey, USA.

10. Nosrati K, Govers G, Semmens BX, Ward EJ (2014) A mixing model to incorporate uncertainty in sediment fingerprinting. Geoderma 217: 173-180.

11. Collins AL, Zhang Y, McChesney D, Walling DE, Haley SM, et al. (2012) Sediment source tracing in a lowland agricultural catchment in southern England using a modified procedure combining statistical analysis and numerical modelling. Sci Total Environ 414: 301-317.

12. Moore JW, Semmens BX (2008) Incorporating uncertainty and prior information into stable isotope mixing models. Ecol Lett 11: 470-480.

13. Collins AL, Walling DE, McMellin GK, Zhang Y, Gray J, et al. (2010) A preliminary investigation of the efficacy of riparian fencing schemes for reducing contributions from eroding channel banks to the siltation of salmonid spawning gravels across the south west UK. J Environ Manag 91: 1341-1349.