TECHNICAL COMMENT

Improving Bayesian isotope mixing models: a response to Jackson *et al.* (2009)

Brice X. Semmens,[1]* Jonathan W. Moore[2] and Eric J. Ward[1]

[1]*Northwest Fisheries Science Center, National Marine Fisheries Service, Seattle, WA 98112, USA*

[2]*Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95064, USA*

Correspondence: E-mail: brice.semmens@noaa.gov

Abstract

We recently described a Bayesian framework for stable isotope mixing models and provided a software tool, MixSIR, for conducting such analyses (*Ecol. Lett.*, 2008; **11**:470). Jackson *et al.* (*Ecol. Lett.*, 2009; 12:E1) criticized the performance of our software based on tests using simulated data. However, their simulation data were flawed, rendering claims of erroneous behaviour inaccurate. A re-evaluation of the MixSIR source code did, however, uncover two minor coding errors, which we have fixed. When data are correctly simulated according to eqns (1)–(4) in Jackson *et al.* (2009), MixSIR consistently and accurately estimated the proportional contribution of prey to a predator diet, and was surprisingly robust to additional unquantified error. Jackson *et al.* (2009) also suggested we use a Dirichlet prior on the source proportion parameters, which we agree with. Finally, Jackson *et al.* (2009) propose adding additional error parameters to our mixing model framework. We caution that such increases in model complexity should be evaluated based on data support.

Keywords

Mixing model, MixSIR, model selection, stable isotope.

Ecology Letters (2009) **12**: E6–E8

INTRODUCTION

In a recent article (Moore & Semmens 2008) we outlined a Bayesian framework for incorporation of prior information and uncertainty into stable isotope mixing models. This mixing model allows the estimate of proportional contributions of different sources (e.g. prey) to a mixture (e.g. consumer). To demonstrate the efficacy of this approach we developed an open source software tool, MixSIR (Semmens & Moore 2008). Recently, Jackson *et al.* (2009) criticized this software. First, Jackson *et al.* (2009) proposed the use of a Dirichlet prior on the source proportion parameters, rather than the individual Beta priors we originally proposed. Additionally, through extensive model testing, Jackson *et al.* (2009) noted three problems with the performance of MixSIR: (i) a propensity for the software to fail, or 'crash', (ii) bias in the median estimates of source contributions and (iii) posterior 95% credible intervals that encompassed the true parameter values considerably less than 95% of the time. Jackson *et al.* (2009) suggested that these problems were likely due to coding problems, or 'bugs', in the MixSIR code. Finally, Jackson *et al.* (2009) suggested that the mixing model formulation would be improved by including error terms to account for additional unquantified error (hereafter referred to as residual error).

We appreciate the efforts of Jackson *et al.* (2009) to improve upon the Bayesian mixing model framework we proposed, and are grateful for their efforts at testing the performance of the MixSIR software package. However, Jackson *et al.* (2009) incorrectly simulated data during their testing of MixSIR. Thus, their claims of erroneous model performance are inaccurate. A re-evaluation of the MixSIR source code did, however, uncover two minor coding errors, which we have fixed. Using correctly simulated data according to eqns (1)–(4) in Jackson *et al.* (2009), and following the resolution of these coding errors, we show that MixSIR performs appropriately and is surprisingly robust to residual error (see online Appendix for the corrected data simulation and MixSIR code). We also provide a brief critical evaluation of the residual error term proposed by Jackson *et al.* (2009) Finally, we briefly caution that the residual error parameters proposed by Jackson *et al.* (2009) represent added model complexity that should be evaluated based on data support.

PRIOR FORMULATION

We agree with Jackson *et al.* (2009) that the Dirichlet distribution (multinomial Beta) is a more appropriate formulation for the prior distributions than the individual Beta distributions we originally proposed. Accordingly, we have implemented a Dirichlet prior formulation in the most recent version of MixSIR (version 1.04), available at http://www.ecologybox.org.

SOFTWARE CRASHES AND BIASED PARAMETER ESTIMATES

The software crashes and biased estimates of source contributions identified by Jackson *et al.* (2009) were due to two simple bugs in the MixSIR source code that we have now fixed.

These coding errors were not inherent deficiencies in the modelling framework we proposed. The interested reader can see specific descriptions of the bug fixes and evaluate differences among MixSIR source code versions on the MixSIR project page at http://www.ecologybox.org. Once the errors were fixed, MixSIR no longer crashed or produced bias in the median estimates of source contributions (Table 1). We note here that the bias identified by Jackson *et al.* (2009) was negligible in the context of interpreting model output (e.g. the contribution of prey items to a predator diet); even without the bug fixes MixSIR always identified the median source contributions within 3% of the true contribution (see left-hand column of panels in fig. 2 of Jackson *et al.* 2009).

CREDIBLE INTERVAL COVERAGE

The exceptionally poor credible interval coverage problem documented by Jackson *et al.* (2009) was principally the manifestation of incorrectly developed simulation data. Specifically, the simulation code they used did not include sufficient dimensionality in the random number generation

Table 2 Percentage of source contribution posteriors from 1000 simulations whose 95% credible intervals contained the true underlying dietary proportion for a given food source ($P_1 = 0.7$, $P_2 = 0.2$ and $P_3 = 0.1$). Simulations were carried out after the resolution of two minor coding errors in MixSIR

Standard deviation of residual error	Dietary proportion		
	P_1	P_2	P_3
0	94.9	93.9	95.15
0.1	95.3	95.0	94.9
0.5	73.0	73.1	73.6
1.0	49.6	53.4	48.4

step, such that for each isotope, the same random number was used to draw values from each of the source distributions (see digital Appendix S1, Jackson *et al.* 2009). Thus if in generating a predator data point the program draws from, say, the 97th percentile of the source 1 $\delta^{13}C$ isotope distribution, it would do the same for source 2 and source 3. Based on their eqn (2), these draws should be independent rather than identical. The over-representation of extreme values in their simulated data caused MixSIR to produce credible intervals with poor coverage three to four times more often compared with intervals resulting from correctly simulated data.

Using code that correctly simulates data according to eqns (1)–(4) of Jackson *et al.* (2009), and following the MixSIR bug fixes, we found that the 95% posterior credible interval coverage for the estimated proportional source contributions included the 'true' proportions > 95% of the time when no residual error was included in the simulation (Table 2). Even with a considerable amount of residual error (SD of 0.1, representing a 20% increase over the error associated strictly with prey isotope signatures), MixSIR credible interval posteriors included the true parameter values ~95% of the time. Thus, MixSIR appears to be surprisingly robust to the inclusion of residual error. Predictably, with larger amounts of such error the accuracy

Table 1 Average MixSIR posterior median source contribution percentages based on 1000 simulated datasets. Numbers in parentheses represent standard deviations of the 1000 posterior median values. Simulations were carried out following Jackson *et al.* (2009) after the resolution of two minor errors in the MixSIR source code. Note that posterior estimates of percentage source contributions are essentially unbiased regardless of the amount of residual error

Standard deviation of the residual error	Source 1 (true contribution = 70) MixSIR	Source 2 (true contribution = 20) MixSIR	Source 3 (true contribution = 10) MixSIR
0	69.99 (0.68)	19.98 (0.66)	9.96 (0.68)
0.1	70.03 (0.70)	19.96 (0.71)	9.92 (0.70)
0.5	70.13 (1.09)	19.96 (1.15)	9.89 (1.11)
1.0	70.46 (1.94)	19.80 (1.87)	9.63 (1.99)

of credible intervals lessened – one would expect increasing failure rates as simulation data are generated with increasing unquantified error.

RESIDUAL ERROR

Jackson et al. (2009) proposed that our model framework requires inclusion of residual error terms. To demonstrate this they developed artificial data infused with residual error, and showed that MixSIR's dietary proportion estimates became increasingly inaccurate as the magnitude of the error increased. This is an intuitive and unsurprising finding that belies two broader questions: (i) how, explicitly, should residual error be incorporated? and (ii) should residual error always be incorporated into mixing model formulations using 'real-world' data? Jackson et al. (2009) have failed to properly address either of these questions.

The specific residual error parameterization and prior formulation advocated by Jackson et al. (2009) cannot be discerned from their manuscript. For instance, the only explicit model formulation provided by Jackson et al. (2009) describes their data simulation method rather than their statistical model formulation. Additionally, they provided no information regarding the priors used for the residual error terms they propose. In essence, they claimed to have achieved superior performance from a model without any explicit description. We are therefore in the unfortunate position of not having enough information to carry out a formal model comparison. We can, however, comment briefly on the prudence of explicitly comparing competing models based on data.

A principal challenge in fitting mixing models to empirical data lies in choosing the most appropriate parameterization from a constellation of candidate models. At a basic level, Jackson et al. (2009) have proposed adding parameters to our model framework, rather than a fundamental revision of the framework. We have no doubt that additional model complexity may be warranted in certain circumstances in order to represent specific biological phenomena. For example, one possible extension of our framework would be the incorporation of concentration dependence (Phillips & Koch 2002). Alternatively, the model we proposed could be parameterized in order to account for the fact that certain populations exhibit more inter-individual variation than others (Urton & Hobson 2005; Layman et al. 2007). Given that such added complexity in statistical models requires data support, it is naive to assume that a single model is most appropriate. Apart from philosophical arguments for or against a specific parameterization, how does one go about justifying the use of additional terms? Future efforts to extend our mixing model framework should employ model selection methods in order to compare model parameterizations and evaluating model parsimony (Burnham & Anderson 2002). Jackson et al. (2009) asserted that residual error terms should be included in isotopic mixing models, yet such error terms likely are not appropriate in many circumstances. The prudent scientist would evaluate support for such complexity before incorporating it into a model (Burnham & Anderson 2002).

ACKNOWLEDGEMENTS

REFERENCES

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York, 488 pp.

Jackson, A.L., Inger, R., Bearhop, S. & Parnell, A. (2009). Erroneous behavior of MixSIR, a recently published Bayesian mixing model: a discussion of Moore & Semmens, Ecology Letters, 2008. *Ecol. Lett.*, 12, E1.

Layman, C.A., Arrington, D.A., Montaña, C.G. & Post, D.M. (2007). Can stable isotope ratios provide for community-wide measures of trophic structure? *Ecology*, 88, 42–48.

Moore, J.W. & Semmens, B.X. (2008). Incorporating uncertainty and prior information into stable isotope mixing models. *Ecol. Lett.*, 11, 470–480.

Phillips, D.L. & Koch, P.L. (2002). Incorporating concentration dependence in stable isotope mixing models. *Oecologia*, 130, 114–125.

Semmens, B.X. & Moore, J.W. (2008). *MixSIR: A Bayesian Stable Isotope Mixing Model*, Version 1.04. http://www.ecologybox.org.

Urton, E.J.M. & Hobson, K.A. (2005). Intrapopulation variation in gray wolf isotope (δ^{15}N and δ^{13}C) profiles: implications for the ecology of individuals. *Oecologia*, 145, 317–326.