Vol. 514: 1–12, 2014
doi: 10.3354/meps11017

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published November 6

FREE ACCESS

FEATURE ARTICLE

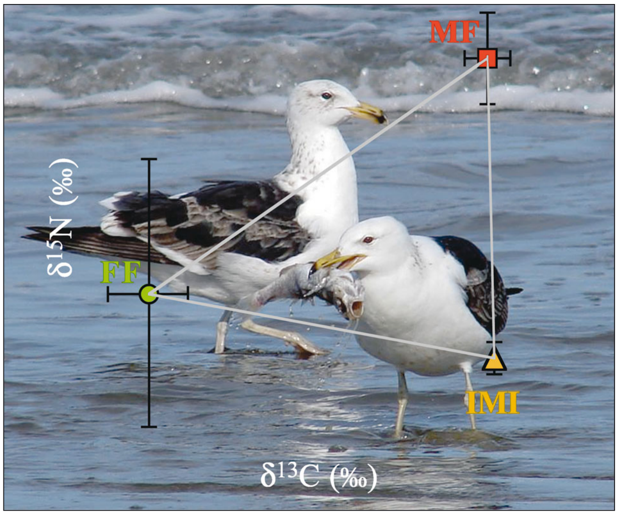
Resource polygon geometry predicts Bayesian stable isotope mixing model bias

Michael T. Brett*

Civil & Environmental Engineering, University of Washington, Seattle, Washington 98105, USA

ABSTRACT: Bayesian mixing model analyses of resource and consumer stable isotope composition are commonly used to infer elemental, energetic, and trophic pathways in aquatic and terrestrial food webs. However, the outputs of these models may be biased towards prior or null generalist assumptions, but the magnitude of this potential bias is unknown. I conducted a series of experiments to determine how this bias is affected by the geometry and end-member uncertainty of resource polygons. These experiments showed that bias is mostly due to isotopic overlap between resources and is very strongly correlated in a sigmoid manner with the normalized surface area of stable isotope resource polygons. The normalized surface area, a classic signal to noise ratio in bivariate space, is calculated by scaling the *x* and *y* ordinates by the mean standard deviations (SD) for $\delta^{13}C$ and $\delta^{15}N$, respectively. When equilateral 3-resource polygons have a surface area <3.4 SD^2, the outputs of Bayesian mixing models primarily reflect the prior generalist assumption. The back-calculated bias for 85 recently published triangular polygons averaged 50 ± 28% (± SD). Analyses of regular resource polygons with 4 to 6 resources required 3.1 to 8.0 times larger normalized surface areas to constrain bias. Furthermore, polygons with 4 or more resources gave poor outcomes for minor diet components. There was a strong bias for resources similar, and against resources dissimilar, to the dominant resource. Overall, Bayesian methods applied to underdetermined models and poorly resolved data very often give results that are highly biased towards centrist and generalist solutions.

KEY WORDS: Bayesian mixing models · Stable isotopes · Trophic pathways · MixSIR · SIAR



A white croaker (*Micropogonias furnieri*) being eaten by a kelp gull (*Larus dominicanus*), the consumer in the superimposed resource polygon from Silva-Costa & Bugoni (2013). This was one of the best-resolved polygons assessed in the present study, with a normalized surface area of 34 SD^2. The polygon is comprised of freshwater fish (FF), marine fish (MF), and intertidal marine invertebrates (IMI).

Image: Fernando Faria

INTRODUCTION

Analyses of resource and consumer stable isotope composition have long been used to infer elemental, energetic, and trophic pathways in aquatic and terrestrial food webs (Peterson & Fry 1987). Typically, the number of potential food resources exceeds the number of stable isotopes that can be used to infer resource utilization. Because these types of mass balance calculations can only resolve n + 1 resources, where n = the number of stable isotopes considered,

*Corresponding author: mtbrett@uw.edu

the majority of food web stable isotope analyses have too many potential resources and not enough tracers and are therefore mathematically underdetermined (Schwarcz 1991, Phillips 2001, Fry 2013a). Natural variability in resource and consumer stable isotope composition exacerbates the underdetermined problem (Phillips & Gregg 2001, Moore & Semmens 2008, Ward et al. 2010). Several 'mixing model' algorithms have been proposed to resolve this problem (Phillips & Gregg 2003, Moore & Semmens 2008, Parnell et al. 2010). Whether the underdetermined problem has been, or even can be, resolved was the subject of a recent debate (Fry 2013b, Semmens et al. 2013). A.W.E. Galloway et al. (unpublished) showed that a fatty acid based approach to these types of problems, which has the advantage of being able to apply 20 or more tracers, gave much more accurate and precise outcomes than did 2 stable isotope analyses. However, A.W.E. Galloway et al. (unpublished) also showed that stable isotope based analyses gave very accurate outcomes when there was very little uncertainty for the stable isotope values.

Based on the results summerized above, I tested the hypothesis that bias for 2 stable isotope cases in Bayesian mixing model outputs is caused by poorly resolved resource polygons. This hypothesis was tested by comparing the outputs for a fixed consumer scenario across a wide gradient of normalized surface areas for triangular resource polygons. Equilateral triangular polygons with uneven end-member uncertainty and acute triangular polygons were also explored. The performance of regular 4, 5, and 6 end-member resource polygons was compared to that of 3 end-member polygons. Finally, the normalized surface area of actual triangular polygons was assessed by carrying out a meta-analysis of papers in the Bayesian mixing model literature that recently reported the configuration of 3 end-member scenarios.

METHODS

Bias in triangular resource Bayesian mixing model analyses

To test whether bias in Bayesian mixing model outputs is related to resource polygon resolution, in an initial experiment I created a fixed triangular resource polygon and a fixed consumer comprised of 80% of the first resource and 10% of the second and third resources. This hypothetical consumer was tested because it was essential to test scenarios that were either much more or much less than the prior, i.e. 33.3% for a triangle. Uncertainty (i.e. standard deviation, SD) values for the end-members were then varied over a geometric gradient to create a series of normalized resource polygons that ranged from being much smaller to much larger than those normally seen in actual field studies (Fig. S1 in the Supplement at www.int-res.com/articles/suppl/m514p001_supp.pdf). The surface areas of the resource polygons were normalized by dividing by the within-polygon mean SD for the δ^{13}C and δ^{15}N values, respectively. The length of each side of these triangles was determined using the Distance Formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

where x and y represent the normalized δ^{13}C and δ^{15}N coordinates, respectively. The normalized surface area of these polygons was then calculated using Heron's Formula:

$$\text{Area} = \sqrt{s(s - ab)(s - bc)(s - ca)} \tag{2}$$

where $ab/bc/ca$ represent the first/second/third side, and s represents the semiperimeter, i.e. one-half the perimeter of the triangle. The untransformed data for these cases were then run through MixSIR (Moore & Semmens 2008) and SIAR (Parnell et al. 2010) based versions of the FASTAR code (Galloway et al. 2014a,b), and the outputs were compared to the normalized surface areas for these cases. These results were also fit to a Boltzmann Sigmoid Function:

$$y = A_2 + \frac{A_1 - A_2}{1 + e^{(x - x_0)/dx}} \tag{3}$$

where A_1 represents the prior lower asymptotic value, A_2 represents the final upper asymptotic value, x represents the natural log transformed surface area, x_0 represents the point of inflection, and dx represents the slope of the function at the point of inflection. These results can also be represented as an expected bias relative to the prior assumption and actual correct value:

$$\text{Bias} = \left(1 - \frac{(\text{mean}_{\text{Boltz}} - \text{assumption})}{(\text{correct} - \text{assumption})}\right) \times 100\% \tag{4}$$

where $\text{mean}_{\text{Boltz}}$ (i.e. y from Eq. 3) represents the average of the equilateral triangle simulations as described by the Boltzmann Sigmoid Function, 'assumption' represents equal contributions from all resources, and 'correct' represents the true value for any resource.

In a second experiment, an equilateral triangular polygon was created, with an 80/10/10 consumer, and the resource SD values were varied for 3 different cases. In the first case, the SD value for the dominant resource was set to one-quarter of the SD values for the minor resources. In the second (null) case, the SD values were the same for all 3 resources. In the third case, the SD value for the dominant resource was set to 4 times the SD value of the minor resources. In all 3 cases, the aggregate average SD values for the 3 resources were the same. These cases were run through the code, and the outputs for the 3 cases were compared.

In a third experiment, an acute isosceles triangle was created with a height that was 4 times its base and with all 3 end-members having the same SD. Then, 2 consumers were created, the first of which was comprised of 80 % of the acute resource and 10 % of the 2 obtuse resources. The second consumer was comprised of 80 % of 1 of the obtuse resources, 10 % of the other obtuse resource, and 10 % of the acute resource.

Bias in 4-, 5-, or 6-resource scenarios

In the fourth, fifth, and sixth experiments, the performance of 2 stable isotope and 4, 5, and 6, respectively, end-member resource polygons were tested. In these experiments, regular (equal-sided) resource polygons with equal variability for each resource and a fixed 80/10/10 consumer were used (Fig. S2 in the Supplement at www.int-res.com/articles/suppl/m514p001_supp.pdf). As was previously done for the first experiment, SD values for the resources were varied over a geometric gradient to create a series of normalized resource polygons that ranged from being much smaller to much larger than typically seen. The normalized surface areas of the regular polygons tested were calculated accordingly:

$$\text{Area} = \frac{r^2 N \sin\left(\frac{2\pi}{N}\right)}{2} \quad (5)$$

where r represents the radius (center to a vertex) of the polygon, N represents the number of sides (or resources), and 'sin' is the sine function in radians.

Consumer sample size influences on mixing model outputs

The original versions of MixSIR and SIAR aggregate outputs for cases where multiple consumers are analyzed simultaneously in a way that reduces dispersion as consumer sample size increases. However, the code used for most of the simulations in this paper aggregated cases separately, so that the dispersion when large samples sizes were considered was no different than when only a single consumer was tested. Because of this difference compared to the conventional codes, I designed a seventh experiment to test whether analyzing multiple consumers with the conventional code would affect model accuracy and precision. Cases with consumer sample sizes of 1, 3, 5, 9, 13, 25, 49, and 101 were tested. All of the consumer datasets considered, with the exception of n = 1, had the same SD values representing the uncertainty typically associated with trophic fractionation in the cases of carbon and nitrogen, and the mean producer SD values for hydrogen. This experiment was based on the data from a poorly resolved polygon in the published literature (surface area = 3.5 SD^2). The resource values used were −29.2 ± 1.5‰, −26.2 ± 2.6‰, and −28.4 ± 4.6‰ for the terrestrial, benthic, and pelagic $\delta^{13}C$ ratios, respectively. Similarly, the $\delta^{15}N$ values used for these resources were −4.6 ± 0.6‰, 2.4 ± 1.7‰, and 5.3 ± 2.3‰, respectively. The $\delta^{2}H$ values used were −129.5 ± 15.2‰, −180.4 ± 18.0‰, and −198.0 ± 8.3‰, respectively. The hypothetical consumer tested was comprised of 10/10/80 terrestrial/benthic/pelagic resources and had a stable isotope composition of −28.3 ± 1.4‰, 4.0 ± 1.0‰, and −189.4 ± 13.8‰ for $\delta^{13}C$, $\delta^{15}N$, and $\delta^{2}H$, respectively.

Meta-analysis of triangular polygons

To determine the size distribution of stable isotope based resource polygons, every paper that cited either Moore & Semmens (2008) or Parnell et al. (2010), n ~ 500, was surveyed for triangular resource polygons. This survey identified 74 cases where the mean ± SD $\delta^{13}C$ and $\delta^{15}N$ values for 3 end-member polygons were either directly provided or could be extracted from bivariate plots using the digitizing software DigitizeIt 1.6.1 (see the Supplement at www.int-res.com/articles/suppl/m514p001_supp.pdf). An additional 11 cases were identified where a third stable isotope was also determined, i.e. 9 cases for $\delta^{2}H$ and 2 cases for $\delta^{34}S$. In a few cases, these triangular polygons differed from the resource polygons originally specified by the authors. For example, Cremona et al. (2009) presented data for 8 aquatic macrophytes, 8 epiphyte biofilms growing on aquatic macrophytes, and 4 suspended particulate matter (SPM) samples. These data were collapsed to only 3 resources, i.e. macrophytes, epiphytes, and SPM. However, in the large majority of cases, the resource polygons extracted were those originally specified by the authors. Once $\delta^{13}C$ and $\delta^{15}N$ coordinates were obtained, they were normalized, and the surface areas of these polygons were calculated as described previously.

RESULTS

Influence of polygon surface area and shape on bias

When using the MixSIR version of the Bayesian mixing model, the first experiment showed bias in outputs that were very strongly related to the normalized surface area of triangular polygons in a sigmoid manner (Fig. 1, Table 1). When the normalized polygon surface area was ≤1.1 SD^2, the mean model output was biased ≥80% towards the prior assumption. For these poorly resolved polygons, the 95% confidence interval for outputs also spanned nearly the entire range of possible values, i.e. 2 to 85% for the dominant (80%) resource. A polygon surface area of 3.4 SD^2 corresponded to 50% bias (Fig. 1). A surface area of ≥10.8 SD^2 reduced bias to ≤20% and considerably reduced the 95% confidence interval for the outputs for the dominant resource, i.e. 46 to 92% (Fig. 1). The experimental outcomes for the 2 minor diet resources were similar to those described above.

When using the SIAR version of the Bayesian mixing model, the first experiment yielded very similar results as for the MixSIR version of the code (Table S1 in the Supplement at www.int-res.com/articles/suppl/m514p001_supp.pdf). However, when well resolved polygons were considered, the SIAR outputs were less accurate and much less precise than MixSIR based outputs. The mean SIAR outputs also had a structural bias that persisted even for the largest polygons tested. Unless otherwise noted, the results presented hereafter are based on MixSIR.

In the second experiment, an equilateral triangle with a normalized surface area of 3.1 SD^2 had its end-member SD values varied between the dominant and minor resources. In the null case, i.e. equal variability for each resource, the mixing model output was 56 ± 20% for the dominant (80%) resource. When the variability for the dominant resource was 4 times larger than that for the minor resources, the mixing model output was 50 ± 16% for the dominant resource. When the variability was one-quarter that for the minor resources, the model output was 64 ± 21% for the dominant resource. In the third experiment, mixing model outputs for acute triangular polygons were examined. These polygons were isosceles triangles with a height that was 4 times greater than the base, equal end-member uncertainty, and a normalized surface area of 3.1 SD^2. When the acute resource was the dominant resource, the algorithm calculated an average contribution of 70 ± 13% from this resource, and contributions of 15 ± 13% for the 2 minor resources. When the dominant resource was obtuse, the model calculated a 47 ± 23% contribution. The model also calculated a contribution of 38 ± 24% for the acute minor resource, and 15 ± 10% for the obtuse minor resource.

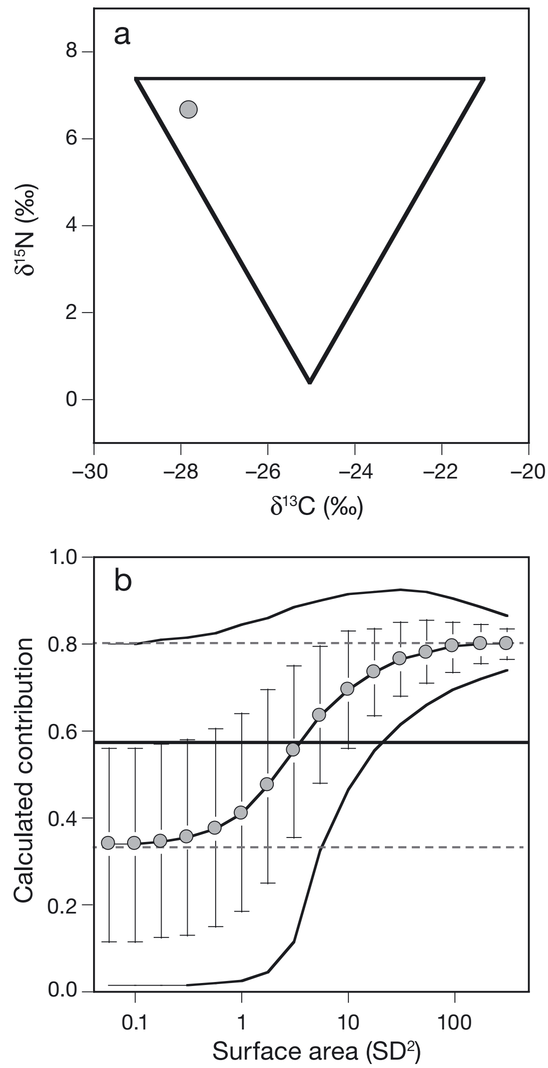


Fig. 1. (a) Triangular resource polygon tested. The grey dot represents a fixed consumer comprised of 80% of the upper left resource, and 10% of the other 2 resources. (b) Calculated resource contribution for the dominant resource (80%) for the 80/10/10 three end-member equilateral scenario. The gradient of normalized surface areas was generated by systematically varying producer SD values, while leaving the producer and consumer stable isotope values fixed. The lower and upper dashed horizontal lines represent the generalist prior assumption and the correct value, respectively, and the solid horizontal line represents the mid-point between these values. The filled circles depict the mean ± SD of the outputs, and the outer envelope represents the 95% confidence interval

Table 1. Boltzmann Sigmoid Functions for the 3-, 4-, 5-, and 6-resource experiments. In the ideal case, the lower asymptotic value (A_1) would correspond to the prior assumption (i.e. 1/N), and the upper asymptotic value (A_2) would correspond to the correct answer (i.e. 0.8 in this case). As is apparent from the A_2 parameter values for the 4, 5, and 6 end-member polygons, the correct value was never achieved even when extremely large polygons (i.e. >1000 SD^2) were considered. These models were fit by using Excel Solver to minimize the model error sum of squares. The experimental outcomes for 3 and 4 end-member resource polygons are shown in Figs. 1 & 2. These model results are for idealized polygons. The average bias for any particular polygon will be similar, but individual cases will vary depending on whether that end-member is obtuse or acute and whether it has greater or lesser variability compared to other end-members of that polygon

Resource parameter	Number of resources 3	4	5	6
A_1	0.334	0.253	0.200	0.167
A_2	0.803	0.750	0.757	0.729
x_0	1.247	2.187	2.872	3.156
dx	0.832	0.851	0.994	1.016
r^2	0.9999	0.9999	0.9999	0.9999

Using equilateral triangles with equal variability for each resource and a surface area of 3.1 SD^2, the influence of the consumer scenario considered was also tested. For this analysis, a series of consumers with varying contributions from the dominant resource (i.e. 100 %, 90 %, 80 %, 70 %, etc.), and even contributions from the 2 minor resources were used. When the dominant resource was 80 % or greater, model output bias consistently averaged ~47 %. As expected, bias decreased somewhat when the hypothetical consumer was less dominated by a single resource, i.e. bias averaged 43 % when the dominant resource contributed 50 %.

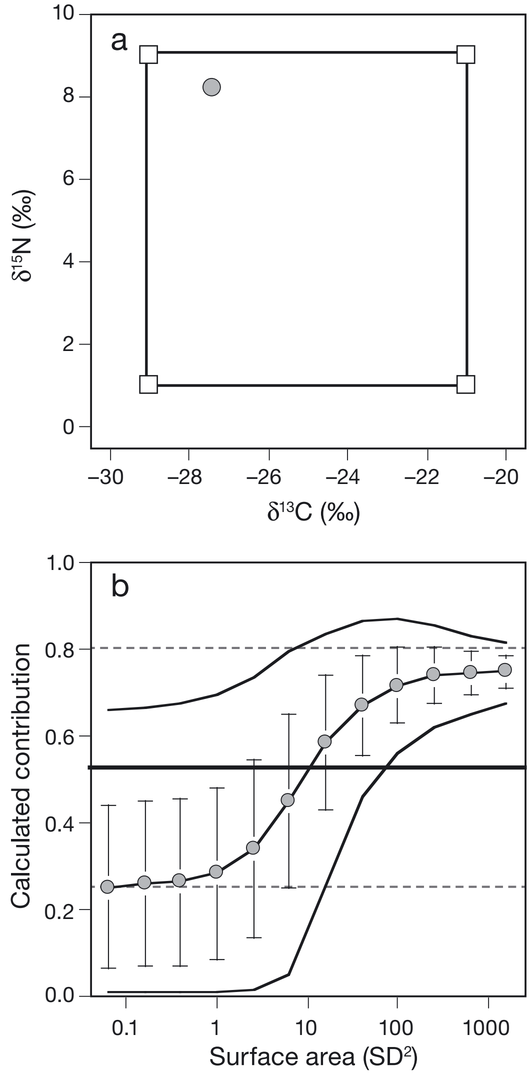


Fig. 2. (a) Rectangular resource polygon tested. The grey dot represents a fixed consumer comprised of 80 % of the upper left resource, 10 % of the next 2 resources in a clockwise direction and 0 % of the fourth resource. (b) Calculated resource contribution for the dominant resource of an 80/10/10/0 4 end-member square scenario. Other details as in Fig. 1

Bias in polygons > n + 1

The results of the fourth through sixth experiments were similar to those for the first experiment in several regards (Fig. 2 and Table 1). For 4, 5, and 6 end-member resource polygons, the calculated average contribution and 95 % confidence interval for a hypothetical dominant resource (i.e. 80 %) depended very strongly on the normalized polygon surface area. However, the polygon surface area required to constrain output bias increased substantially as more resources were considered. As previously noted, a surface area of 3.4 SD^2 corresponded to 50 % bias for an equilateral resource polygon. For a square polygon, this increased to 10.5 SD^2, and for 5 and 6 end-member polygons, surface areas of 20.6 and 30.3 SD^2, respectively, were required (Fig. 3). The fourth through sixth experiments also showed a structural bias of ~10 % against the dominant resource even when very large polygons were considered.

These 4, 5, and 6 end-member cases also indicated poor performance for the minor resources. For example, in the 5 end-member case, the calculated contribution for the minor resource (10 % contribution) that was closest to the dominant resource in a regular polygon (i.e. the second of 5 resources) and converged on 10 % for very large polygons (Fig. 4). The contribution for the third resource in the series,

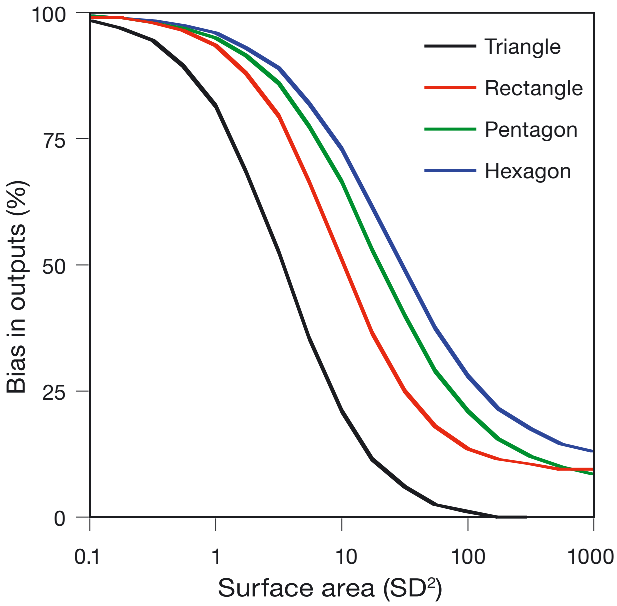


Fig. 3. Polygon surface area versus bias for estimates for the dominant resource obtained from the first, fourth, fifth, and sixth experiments. For 3, 4, 5, and 6 end-member regular polygons, 50% bias was obtained at 3.4, 10.5, 20.6, and 30.3 SD^2, respectively. This figure also shows that in cases where the potential resources exceeded n + 1 there is a persistence structural bias of ~10%, even when the polygon surface area was ≥1000 SD^2

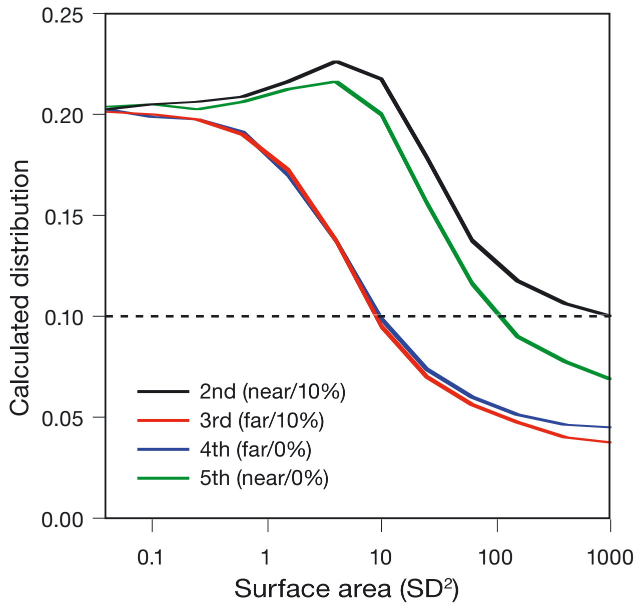


Fig. 4. Calculated contributions for the second, third, fourth, and fifth resource for a 5 end-member regular resource polygon. In this case, the first resource was dominant (i.e. 80%), the second and third resources each contributed 10%, and the fourth and fifth resources did not contribute. The dashed horizontal line represents the correct response for the second and third resources. Very comparable bias in favor of resources similar, and against resources dissimilar, to the dominant resource was also observed for 4 and 6 end-member resource polygons

which should also have been 10%, converged on 5% for very large polygons. This calculated contribution was very similar to that for the fourth potential resource in this case, which should have been 0%. Finally, the fifth resource in this case, which was adjacent to the dominant first resource but should have been 0%, was calculated to be ~7% for very large surface area polygons. In this case, as well as the case for 4 and 6 end-member resource polygons (not shown), physical proximity to the dominant resource in the polygon was more strongly associated with the model outputs than was the actual correct answer (Fig. 4).

Consumer sample size

The model runs with the conventional code for SIAR, which aggregates results when multiple consumers are considered, showed that analyzing multiple consumers simultaneously gave outputs that were much more precise, but no more accurate when a poorly resolved polygon (surface area = 3.5 SD^2) was considered (Fig. 5). This polygon was retested after its resource SD values were multiplied by 0.53, effectively increasing its normalized surface area to 12.1 SD^2, which was the 90th percentile of the 85 published cases summarized below. When this well-resolved polygon was tested, considering more consumers led to moderate improvement in model performance. For example, when only 1 consumer was tested, the output indicated that the pelagic contribution (which should have been 80%) was 61%, and the calculated contribution increased to 71% when a consumer sample size of 13 was used (Fig. 5). As expected, the well-resolved polygon gave much more accurate outcomes irrespective of the number of consumers tested. Although not shown, MixSIR performed worse than SIAR in this experiment. In both cases above, testing more consumers with MixSIR did not improve model accuracy but did greatly reduce dispersion.

Characteristics of published triangular polygons

The 85 triangular polygons surveyed had a median surface area of 3.9 SD^2, and a 10th to 90th percentile

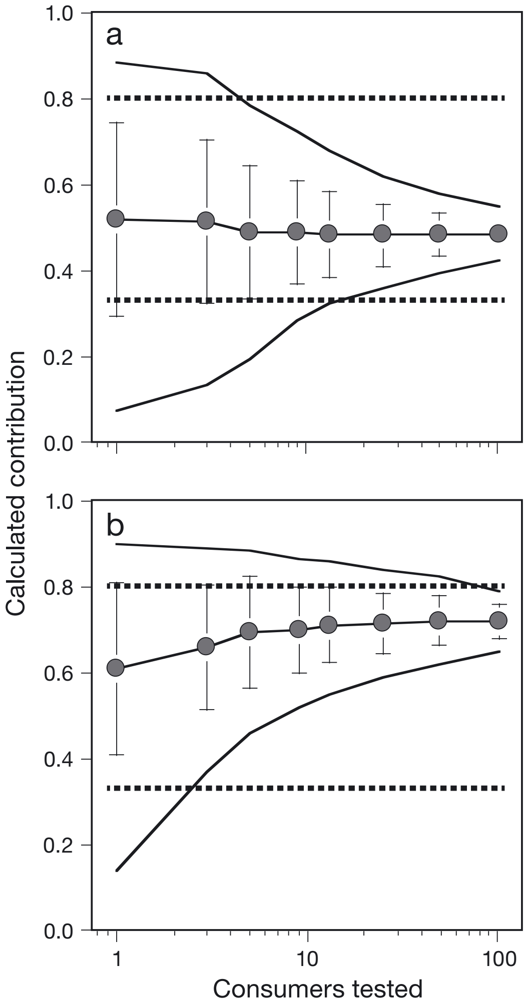


Fig. 5. Calculated resource contribution for the dominant resource of an 80/10/10 consumer for a 3-isotope triangular polygon. This experiment tested a wide range of consumers (i.e. n = 1, 3, 5, 9, 13, 25, 49, and 101) simultaneously using the original version of SIAR which pools results in a manner that tends to reduce dispersion when larger sample sizes are considered. Shown are scenarios where (a) the original polygon SD values were used (surface area = 3.5 SD^2) and (b) the original SD values were reduced to create a well-resolved polygon (surface area = 12.1 SD^2). Other details as in Fig. 1. When using MixSIR, testing more consumers simultaneously did not improve model accuracy

range of 0.5 to 12.1 SD^2 (Fig. 6). According to the sigmoid response from the first experiment (Table 1), this corresponded to an average bias of 50 ± 28%. The δ^{13}C and δ^{15}N ordinates of these polygons had average distances of 6.9 ± 4.2‰ and 4.4 ± 3.1‰, respectively. These ordinates also had average SD values of 1.8 ± 1.0‰ and 1.4 ± 1.0‰, respectively. These dimensions gave a median absolute surface area of 7.4‰^{2}, and a 10th to 90th percentile range of 0.9 to 19.4‰^{2}.

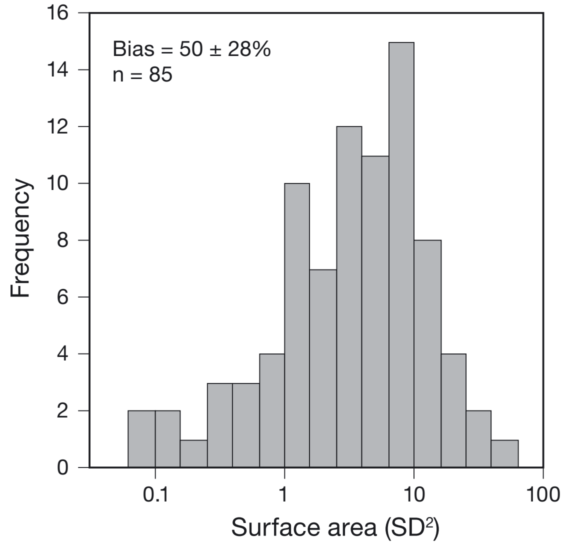


Fig. 6. Normalized surface area distribution for recently published triangular resource polygons (n = 85). These polygons had a median surface area of 3.9 SD^2, and a 10th to 90th percentile range of 0.5 to 12.1 SD^2. When the sigmoid function depicted in Fig. 1, and described in Table 1, is applied to this distribution, an average bias of 50 ± 28% is obtained

DISCUSSION

This analysis shows that the MixSIR Bayesian mixing model (Moore & Semmens 2008) is an efficient algorithm for resolving 2 stable isotope and 3 resource based analyses provided the isotopic ratios for the potential resources are quite distinct. Unfortunately, the meta-analysis of recently published triangular polygons showed that only about 13% of these polygons were, in fact, well resolved (i.e. had less than 20% bias). Overall, bias for triangular polygons averaged 50 ± 28%, indicating there is likely a strong bias towards generalist outcomes in these types of analyses (Fry 2013b). About 30% of the polygons considered were so poorly resolved (>75% bias) that it is unlikely these could provide results that are differentiated from the priors. This bias appears to be much more severe for more complex resource polygon scenarios. When resource polygons exceeded the n + 1 criterion, Bayesian mixing model outputs exhibited a negative structural bias of ~10% for the dominant resource even when extremely large polygons were considered. For polygons with 4 or more resources, there was also a pronounced bias for minor diet constituents in favor of resources with similar stable isotope ratios to the dominant resource. In these cases, the mixing model only appears to correctly characterize the dominant resource, and only when the polygon surface area was large. These outcomes are qualitatively very similar to those reported by Phillips & Gregg (2001), who found that uncertainty in stable isotope analysis is strongly related to how distinct the resources are from each other. However, the results of the present study contrast markedly from the statement of Parnell et al. (2010) that the Bayesian mixing model SIAR works exceptionally well even in underdetermined systems where core assumptions are violated. In fact, both MixSIR, and to a somewhat lesser extent SIAR, were very efficient algorithms for finding the correct answer provided the resource data were distinct and the underdetermined constraint was not violated. Conversely, output quality was poor, and oftentimes completely non-informative, when these conditions were not met.

The bias towards the generalist prior for poorly resolved polygons is probably because when the resources have overlaping stable isotope ratios, a very wide range of solutions is plausible. If these types of data were analyzed using a conventional algebraic and Monte Carlo simulation approach, the wide range of outputs would also include a high proportion of outcomes <0% and >100%, which would be an obvious red flag indicating that the resource polygon being considered does not have a clear solution. However, both MixSIR and SIAR force all results into the >0% to <100% domain, which obscures what would otherwise be a clear indication of model misspecification. These algorithms will always fit a model and provide very detailed outputs with no warnings even if the data are nonsensical (Parnell et al. 2010). Both MixSIR and SIAR exhibited a generalist bias when the polygons were poorly resolved, but in some experiments one algorithm performed better than the other. In general, these algorithms are very similar, with the most obvious difference being that SIAR includes a residual error term whereas MixSIR does not (Parnell et al. 2010).

The normalized surface areas of the 85 triangular polygons (Fig. 6) indicates that on average about half of the outcomes reported in the recent published literature mainly reflect the prior generalist assumptions inherent in most Bayesian stable isotope mixing models. For example, the default prior for SIAR is the uninformative uniform Dirichlet distribution, which starts with all potential resources being equally important (Parnell et al. 2010). However, it is possible to use informative priors if additional information on diet or feeding is available. Only 13% of the 85 triangular polygons had normalized surface areas greater than 10.8 SD^2, which indicated outputs that were less than 20% biased. It should also be noted that mixing model outputs for resources in triangular polygons were more accurate and precise than otherwise expected, for an equilateral case, when that resource was an acute end-member of the polygon or it was less variable than the other end-members. Conversely, outputs for obtuse and more variable end-members were worse. It is also noteworthy that whereas the largest normalized triangular polygon surveyed for this study had a surface area of 43 SD^2, the polygon used as an example by Jackson et al. (2009) when debating the attributes of these types of analyses had a surface area of 693 SD^2. The fact that this hypothetical polygon was 15 times larger than the largest polygon observed in the meta-analysis conducted for this study could indicate that perceptions of what may constitute 'typical' resource mixing polygons are askew.

The real situation may be worse than depicted in Fig. 6 because the actual uncertainty for the resources considered in the 85 polygons may be larger than reported. Some of the studies summarized indicated that resources were sampled at a very limited spatial and temporal scale, meaning these were in effect pseudo-replicated. Many studies only provided a cursory description of the sampling design used to characterize resources. The overall average resource SD values, i.e. 1.8 ± 1.0‰ and 1.4 ± 1.0‰, respectively, for $\delta^{13}C$ and $\delta^{15}N$, were much lower than those reported by Cloern et al. (2002) who collected 870 basal resource stable isotope samples (e.g. terrestrial riparian vegetation, salt marsh vegetation, aquatic macrophytes, and phytoplankton) and found that within a plant group, $\delta^{13}C$ and $\delta^{15}N$ values varied on averaged by ±5 to 10‰. Cloern et al. (2002) identified 3 modes of variability within their dataset, i.e. (1) between plant species and microhabitats, (2) over seasonal cycles of growth and senescence, and (3) between living and decomposing biomass. The latter point is especially germane because many field studies collect basal resources in 1 habitat (e.g. terrestrial vegetation) and assume those stable isotope values are representative of what is processed in another (e.g. leaf litter in streams) without accounting for microbial degradation.

Of the 85 actual polygons assessed, 74 were based on 2 isotopes and 11 were based on 3 isotopes, with deuterium the most common third isotope. Polygons based on 3 isotopes followed the same normalized surface area versus bias response summarized in Fig. 1, but they were somewhat larger and therefore less biased than 2-isotope triangular polygons. The normalized surface area for 10 of the 3-isotope triangular polygons was on average 44% larger than for the corresponding 2-isotope polygon for that dataset, i.e. mean 6.2 ± 27 versus 4.3 ± 2.6 SD^2. However, in the case of Batt et al. (2012), adding deuterium as a third tracer increased the polygon surface area by a factor of 5. The improvement in resolution for these 11 cases resulted in a reduction in mean bias from 53 ± 21 % to 38 ± 18 % with the addition of a third isotope.

The most well-resolved triangular polygons, i.e. those with surface areas greater than 10.8 SD^2, tended to be cases where very different resource types were being compared. For example, the studies of Semmens et al. (2009), Wootton (2012), Cummings & Schindler (2013), and Silva-Costa & Bugoni (2013) explored the contributions of marine and freshwater or terrestrial resources to consumers in spatially subsidized food webs (Polis et al. 1997). Jensen et al. (2012), Xu et al. (2012), and Ruokonen et al. (2012) studied benthic–pelagic or benthic–terrestrial systems. However, studying ecotones does not guarantee a well-resolved resource polygon, as many ecotonal food webs were actually poorly resolved. As previously noted, because the stable isotope ratios of particular basal resources can vary considerably temporally and spatially (Cloern et al. 2002), some resource polygons could appear to be well resolved simply because they are based on a narrow range of conditions.

The experiment conducted with the conventional SIAR code showed that considering multiple consumers simultaneously greatly reduced output dispersion, but did not have a consistent effect on output accuracy (Fig. 5). This is disconcerting because many people conflate precision with accuracy, and when Bayesian stable isotope mixing models are applied in field settings, only output precision will be known. This attribute of the conventional code can be easily addressed by modifying the algorithm so that each consumer case is treated completely independently when reporting model outputs.

Analyses of 4, 5, or 6 end-member resource polygons indicated that much larger surface areas are required to constrain bias. For example, a 4 end-member polygon would have to have a 3.1 times larger surface area (i.e. 10.5 SD^2), than a triangular polygon, to achieve 50 % or less bias. Because any rectangle can also be conceptualized as 2 triangles, tripling the surface area of a triangular polygon by only adding a single point can only be accomplished by adding a very distal point. Conversely, in real applications, the stable isotope values for a fourth resource could easily fall close to one of the original resources or even within the initial triangle. For regular polygons with the same radiuses, 4-, 5-, and 6-sided polygons will have surface areas that are respectively 1.5, 1.8, and 2.0 times larger than a corresponding triangle. In contrast, the results of the experiments with 4, 5, and 6 end-member resource polygons indicate that they would have to be 3.1, 6.0, and 8.8 times larger, respectively, to maintain the same level of bias for the dominant component of the diet. These simulations also indicated that the precision and accuracy for minor diet components declined dramatically when the n + 1 criterion was violated. When considering these cases, minor resources are never ruled out as being important even when they actually contributed 0 % to the consumer tested. Poorly resolved resource isotopic data resulted in higher estimated contributions from minor or unimportant resources in all cases. This indicates that in many field applications, mixing model outputs for minor resources should be interpreted with great caution. Overall, these outcomes indicate that using more than 3 resources for these types of mixing model analyses is only warranted when the additional resources are very distinct from the original resources. Lumping multiple food resources down to 3 will in most cases give less biased outcomes for dominant and minor diet components.

Bayesian mixing models can also be biased towards the prior generalist assumption when fractionation-corrected consumers fall outside of the resource polygon (Parnell et al. 2010, Smith et al. 2013). Consumers falling outside of the resource polygon may indicate that either the potential food sources or the consumer trophic enrichment has been misrepresented. Smith et al. (2013) developed a creative Monte Carlo based approach to quantitatively test whether the point-in-polygon assumption is met and thus whether a Bayesian mixing model analysis is even warranted. In cases where the consumer does not fall within the resource polygon, Bayesian mixing models will provide misleading results.

Fry (2013a) pointed out that in cases where consumers fall in the region adjacent to the polygon centroid, mixing models will inevitably output equal contributions from all potential resources. For example, in a regular hexagonal resource polygon, a consumer that is the mathematical average of all resources could be comprised of 50/50 contributions of any 2 opposite resources in the polygon. Alternatively, they could be comprised of 33/33/33 of the first, third, and fifth resources, etc., as well as an infinite number of permutations between these simple examples and 16.7 % contributions from all 6 resources. An interesting elaboration of Fry's (2013a) point from this study is that when considering a perfectly average (centroid) consumer in a hexagonal resource scenario, specialist consumers (e.g. 50/50 for the first and fourth resources) are actually statistically excluded as plausible outcomes for larger resource polygons. For example, for regular hexagonal cases with surface areas >4 SD^2, a 50% consumer falls outside of the 95% confidence interval. For consumers similar to the centroid, larger polygons assure precise generalist outcomes.

RECOMMENDATIONS

Investigators can use the Boltzmann Sigmoid Functions summarized in Table 1 and Eqs. (3) & (4) to assess whether their particular polygon falls within the well or poorly resolved domains. Investigators should also compare values from conventional algebraic analyses to the means from Bayesian mixing models. The program IsoError can also be used with triangular polygons as an alternative to Bayesian mixing models (Phillips & Gregg 2001). IsoError will yield unbiased mean and error estimates when only 1 or 2 isotopic tracers are available. Even more importantly, all stable isotope mixing model analyses should present a plot of their polygons, and make the mean ± SD values for each resource and consumer readily available. Assumptions regarding consumer isotopic fractionation and trophic level should be explicitly stated. Investigators should also acknowledge that Bayesian mixing models are not a panacea for poorly resolved data.

Every investigator should quantify the tendency of their own resource polygon to misclassify consumers. The simplest test for misclassification error is to quantify the tendency of Bayesian mixing models to misclassify the raw resource data on which they are based. This test is not confounded by uncertainty related to the assumed consumer isotopic fractionation or trophic level. To demonstrate, consider the moderately well resolved (surface area = 7.6 SD^2) 3-isotope triangular polygon reported by Hondula & Pace (2014) (Table 2). When these resource isotopic values were run through MixSIR, an average misclassification of error of 26 ± 4% was obtained. When these data were run through SIAR, the average error increased to 46 ± 6%. For example, MixSIR classified the isotopic values for microalgae as representing 75% microalgae, 13% macroalgae, and 12% macrophytes (Fig. 7). This indicates that any individual outcome below 12% is within the expected background error for this particular polygon. Although the original authors did not consider 8 resources simultaneously, because they reported the mean ± SD isotopic values for all 8 resources used to obtain their triangu-

Table 2. Stable isotope values for microalgae, macroalgae, and macrophytes from Hondula & Pace (2014). These values are based on the mean ± SD values reported for the individual components of these groups. Microalgae: benthic microalgae and phytoplankton; macroalgae: *Agardhiella, Codium, Gracilaria*, and *Ulva*; macrophytes: *Spartina* and *Zostera*

	$\delta^{13}C$	$\delta^{15}N$	$\delta^{2}H$	SD C	SD N	SD H
Microalgae	−23.7	5.4	−141.4	3.0	0.6	32.8
Macroalgae	−18.2	9.1	−137.9	2.9	1.4	46.1
Macrophytes	−11.9	5.8	−94.7	2.0	1.3	11.6

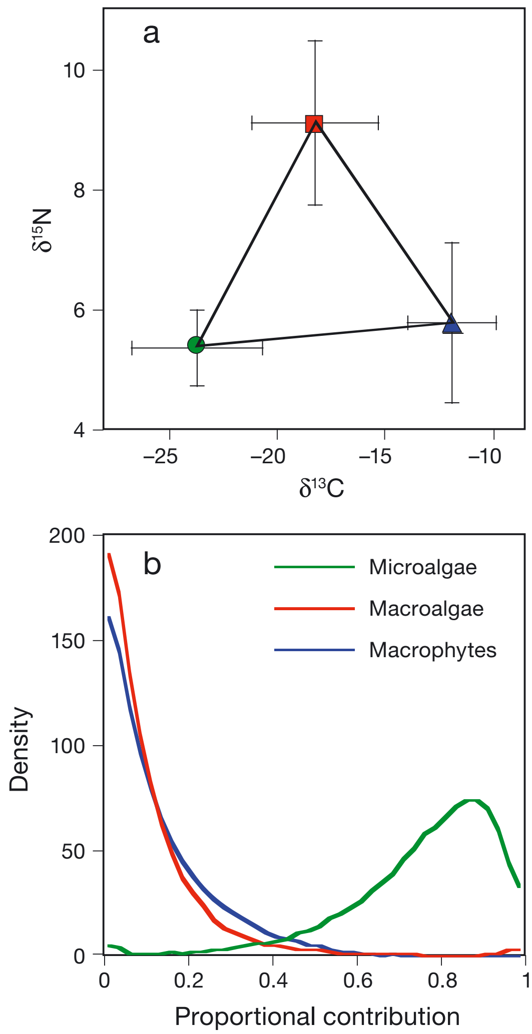


Fig. 7. Triangular polygon based on the data of Hondula & Pace (2014) (see Table 2 for raw data). (a) The $\delta^{13}C$ and $\delta^{14}N$ isotope values; however, $\delta^{2}H$ values were also used for the calculations. For this polygon, most of the differentiation was due to the carbon and nitrogen isotopes because the hydrogen values for microalgae and macroalgae completely overlap. (b) The outcome of a Bayesian mixing model analysis based on the microalgae mean ± SD isotopic values when using 3 isotopes

lar polygon these can be used to assess misclassification error when a polygon is extremely underdetermined, i.e. 8 resources with only 3 isotopes. In this case, the misclassification error averaged 81 ± 1% because the Bayesian mixing model always used all 8 resources in its solutions. In several cases, the particular resource that was being assessed was not even classified as the dominant resource. For example, when the isotopic values for *Spartina* were run though the model, the solution included 38% *Zosteria* and only 19% *Spartina*. When phytoplankton isotopic values were run through the model, *Ulva* classified as 26% compared to only 20% for phytoplankton. This misclassification error was greatly reduced when the resource SD values were changed to extremely low values, i.e. SD/100. Misclassification error was larger when these data were processed with SIAR, and in this case, uncertainty associated with the hypothetical consumers was very influential.

Finally, researchers should consider employing additional dietary tracers. For some systems, fatty acids have proven to be robust dietary tracers that make it possible to explore basal resource contributions at a much finer scale than possible when only using a few isotopes (Galloway et al. 2014b). Most basal resource types synthesize 10 to 20 different fatty acid molecules, some of which are very characteristic for particular plant, algal, or bacterial groups (Galloway et al. 2012, Taipale et al. 2013). Dietary fatty acids have also been shown to leave strong signatures in some consumers (Brett et al. 2006, Galloway et al. 2014a). With additional research, other compounds may prove to be useful tracers, and future research on this topic is very much needed. Fortunately, the Bayesian mixing model framework can be easily modified to accommodate multiple tracers (Nosrati et al. 2014).

CONCLUSIONS

The sigmoidal bias models reported in Table 1 provide a basis to predict whether future, or even past, Bayesian mixing model analyses are markedly affected by bias towards the prior assumption. As should be intuitive, these results show that Bayesian mixing models can only resolve diet contributions when the food resources considered are well resolved and distinct. These results also show that the undetermined constraint creates a structural bias against the dominant resource in cases where the n + 1 criterion is violated, i.e. the resources considered exceed the tracers used by ≥2. In these cases, the Bayesian mixing model also poorly resolves minor food resources and never rules them out as being potentially important. Investigators should quantify the tendency of their own polygons to misclassify data. Due to the mathematical limitations imposed by the underdetermined constraint, employing a greater number of dietary tracers is the most promising means to improve the performance of this class of models (Fry 2013a).

Acknowledgements. Thanks to Eric Ward, Gordon Holtgrieve, and Aaron Galloway for refining the script (from the original MixSIR and SIAR) on which this analysis is based.

LITERATURE CITED

Batt RD, Carpenter SR, Cole JJ, Pace ML, Cline TJ, Johnson RA, Seekell DA (2012) Resources supporting the food web of a naturally productive lake. Limnol Oceanogr 57: 1443–1452

Brett MT, Müller-Navarra DC, Ballantyne AP, Ravet JL, Goldman CR (2006) *Daphnia* fatty acid composition reflects that of their diet. Limnol Oceanogr 51:2428–2437

Cloern JE, Canuel EA, Harris D (2002) Stable carbon and nitrogen isotope composition of aquatic and terrestrial plants of the San Francisco Bay estuarine system. Limnol Oceanogr 47:713–729

Cremona F, Hamelin S, Planas D, Lucotte M (2009) Sources of organic matter and methylmercury in littoral macroinvertebrates: a stable isotope approach. Biogeochemistry 94:81–94

Cummings BM, Schindler DE (2013) Depth variation in isotopic composition of benthic resources and assessment of sculpin feeding patterns in an oligotrophic Alaskan lake. Aquat Ecol 47:403–414

Fry B (2013a) Alternative approaches for solving underdetermined isotope mixing problems. Mar Ecol Prog Ser 472:1–13

Fry B (2013b) Minmax solutions for underdetermined isotope mixing problems: Reply to Semmens et al. (2013). Mar Ecol Prog Ser 490:291–294

Galloway AWE, Britton–Simmons KH, Duggins DO, Gabrielson PW, Brett MT (2012) Fatty acid signatures differentiate marine macrophytes at ordinal and family ranks. J Phycol 48:956–965

Galloway AWE, Eisenlord ME, Dethier MN, Holtgrieve G, Brett MT (2014a) Quantitative estimates of isopod resource utilization using a Bayesian fatty acid mixing model. Mar Ecol Prog Ser 507:219–232

Galloway AWE, Taipale SJ, Hiltunen M, Peltomaa E, Strandberg U, Brett MT, Kankaala P (2014b) Diet–specific biomarkers show that high-quality phytoplankton fuels herbivorous zooplankton in large boreal lakes. Freshw Biol 59:1902–1915

Hondula KL, Pace ML (2014) Macroalgal support of cultured hard clams in a low nitrogen coastal lagoon. Mar Ecol Prog Ser 498:187–201

Jackson AL, Inger R, Bearhop S, Parnell A (2009) Erroneous behaviour of MixSIR, a recently published Bayesian isotope mixing model: a discussion of Moore & Semmens (2008). Ecol Lett 12:E1–E5

Jensen H, Kiljunen M, Amundsen PA (2012) Dietary

ontogeny and niche shift to piscivory in lacustrine brown trout *Salmo trutta* revealed by stomach content and stable isotope analyses. J Fish Biol 80:2448–2462
Moore JW, Semmens BX (2008) Incorporating uncertainty and prior information into stable isotope mixing models. Ecol Lett 11:470–480
Nosrati K, Govers G, Semmens BX, Ward EJ (2014) A mixing model to incorporate uncertainty in sediment fingerprinting. Geoderma 217-218:173–180
Parnell AC, Inger R, Bearhop S, Jackson AL (2010) Source partitioning using stable isotopes: coping with too much variation. PLoS ONE 5:e9672
Peterson BJ, Fry B (1987) Stable isotopes in ecosystem studies. Annu Rev Ecol Syst 18:293–320
Phillips DL (2001) Mixing models in analyses of diet using multiple stable isotopes: a critique. Oecologia 127: 166–170
Phillips DL, Gregg JW (2001) Uncertainty in source partitioning using stable isotopes. Oecologia 127:171–179
Phillips DL, Gregg JW (2003) Source partitioning using stable isotopes: coping with too many sources. Oecologia 136:261–269
Polis GA, Anderson WB, Holt RD (1997) Toward an integration of landscape and food web ecology: the dynamics of spatially subsidized food webs. Annu Rev Ecol Syst 28: 289–316
Ruokonen TJ, Kiljunen M, Karjalainen J, Hämäläinen H (2012) Invasive crayfish increase habitat connectivity: a case study in a large boreal lake. Knowl Manag Aquat Ecosyst 407:08
Schwarcz HP (1991) Some theoretical aspects of isotope paleodiet studies. J Archaeol Sci 18:261–275
Semmens BX, Ward EJ, Moore JW, Darimont CT (2009) Quantifying inter-and intra-population niche variability using hierarchical Bayesian stable isotope mixing models. PLoS ONE 4:e6187
Semmens BX, Ward EJ, Parnell AC, Phillips DL and others (2013) Statistical basis and outputs of stable isotope mixing models: Comment on Fry (2013). Mar Ecol Prog Ser 490:285–289
Silva-Costa A, Bugoni L (2013) Feeding ecology of Kelp Gulls (*Larus dominicanus*) in marine and limnetic environments. Aquat Ecol 47:211–224
Smith JA, Mazumder D, Suthers SM, Taylor MD (2013) To fit or not to fit: evaluating stable isotope mixing models using simulated mixing polygons. Methods Ecol Evol 4:612–618
Taipale S, Strandberg U, Peltomaa E, Galloway AWE, Ojala A, Brett MT (2013) Fatty acid composition as biomarkers of freshwater microalgae: analysis of 37 strains of microalgae in 22 genera and in seven classes. Aquat Microb Ecol 71:165–178
Ward EJ, Semmens BX, Schindler DE (2010) Including source uncertainty and prior information in the analysis of stable isotope mixing models. Environ Sci Technol 44: 4645–4650
Wootton JT (2012) River food web response to large-scale riparian zone manipulations. PLoS ONE 7:e51839
Xu J, Wen Z, Gong Z, Zhang M, Xie P, Hansson LA (2012) Seasonal trophic niche shift and cascading effect of a generalist predator fish. PLoS ONE 7:e49691

Editorial responsibility: Steven Morgan, Bodega Bay, California, USA

Submitted: February 17, 2014; Accepted: August 25, 2014
Proofs received from author(s): October 10, 2014