Vol. 472: 1–13, 2013
doi: 10.3354/meps10168

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published January 9

OPEN ACCESS

FEATURE ARTICLE: REVIEW

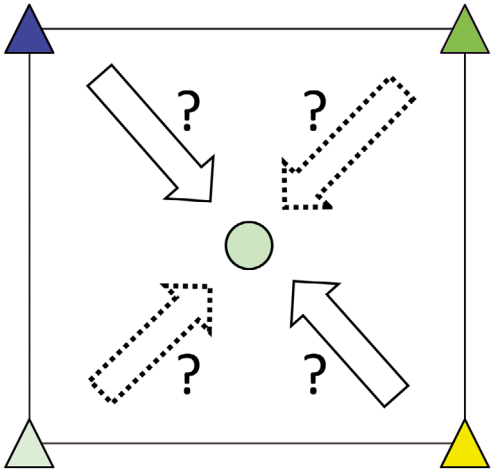
Alternative approaches for solving underdetermined isotope mixing problems

Brian Fry*

Australian Rivers Institute, Griffith University, 170 Kessels Road, Brisbane, Queensland 4111, Australia

ABSTRACT: Statistical mixing models have been developed to help ecologists deal with isotope tracer data and to estimate source contributions in complex systems such as food webs and sediments. However, there are often too few tracer measurements and too many sources, so that unique solutions are not possible in underdetermined mixing models. This review highlights 3 approaches for solving otherwise underdetermined mixing models. The approaches include frequency-based statistics, calculations based on sectors measured in mixing polygons, and linear mixing between central and sidewall points in the mixing polygons. All approaches have some assumptions that allow extrapolation of mean solutions from measured data, with the simplest assumption being that any uncertainty in source contributions is divided in an even-handed manner among sources. A new graphical approach is proposed that allows scientists to critically recognize and separate data-supported aspects of solutions from any assumed aspects of solutions. The data-supported aspects of solutions can be tracked conservatively as the sum of the minimum source contributions, $\sum_{MIN}$, and for the many cases where $\sum_{MIN}$ is low, additional ways to approach mixing problems are summarized from the published literature. Many underdetermined mixing problems do not have robust mean solutions with tracers employed thus far, so that there is a longer-term need for additional tracers and methodologies to really solve these complex ecological problems. This review concludes with several practical steps one can take to interpret isotope tracer information from underdetermined systems.

KEY WORDS: Stable isotopes · Mixing models · Tracers · Food web · IsoSource · SIAR



New mixing-model approaches strengthen the use of stable isotope tracers for solving environmental problems.

Figure: B. Fry

INTRODUCTION

Marine systems are complex, and isotope tracer studies have been used widely during the past decades to understand sources and pathways of organic matter flow in estuaries and oceans (Fry & Sherr 1984, Lee et al. 2012). The isotope studies have made special contributions in the area of food web studies, helping to constrain ideas about energy flow and trophic pathways in many field settings (Demopoulos et al. 2008, Oakes et al. 2010). Several modelling approaches have been developed to deal with a problem that complicates interpretation of the isotope results within mass-balance mixing models, the problem of too many sources and not enough tracers (Minagawa 1992, Phillips & Gregg 2003, Lubetkin & Simenstad 2004, Parnell et al. 2010). Stable isotope

*Email: b.fry@griffith.edu.au

tracers used to generate these mixing models include measurements of δ^2H, $\delta^{13}C$, $\delta^{15}N$, $\delta^{18}O$ and $\delta^{34}S$, among others (Phillips & Gregg 2003). Programs such as IsoSource (www.epa.gov/wed/pages/models/stableIsotopes/isosource/isosource.htm) help evaluate these underconstrained problems and provide summary information about the suite of possible or feasible solutions. This summary information includes means, error estimates and minimum and maximum values. However, how to use this information is not easily understood by many ecologists. Especially, ecologists often use the statistical mean solutions as seemingly the 'most likely', in spite of explicit warnings that minimum and maximum estimates are the robust output from the modelling (Benstead et al. 2006). This review seeks to pinpoint and illustrate the sources of confusion about using the statistical mixing models, especially that using the mean solutions mixes 2 kinds of information, well-constrained minimum source estimates and poorly-constrained estimates where all sources contribute to an assumed solution. This review uses simple geometric examples involving squares and trapezoids to clarify the role of assumption in solving these underdetermined mixing problems, then gives a case study to illustrate the combined use of mean, minimum and maximum feasible results to identify the robust results from statistical model output. Much of this review focuses on the widely used IsoSource program, but comments apply also to other statistical packages such as SIAR (Parnell et al. 2010) that use underconstrained approaches to estimate source contributions. Examples concern food web applications.

SOLUTIONS FOR UNDERDETERMINED MIXING MODELS ARE VIA ASSUMPTION

Many isotope mixing problems are underdetermined, with too many sources and not enough tracers. Fig. 1A gives an example using 2 hypothetical tracer measurements (δ_1 and δ_2 plotted on the *x*- and *y*-axes) that are similar to e.g. $\delta^{13}C$ and $\delta^{15}N$ or δ^2H and $\delta^{34}S$, 4 sources (the apices of the square), and a point in the middle (the data point of interest, also the centroid of the square). The fractional contributions of the sources contributing to the data point can be represented by 3 equations:

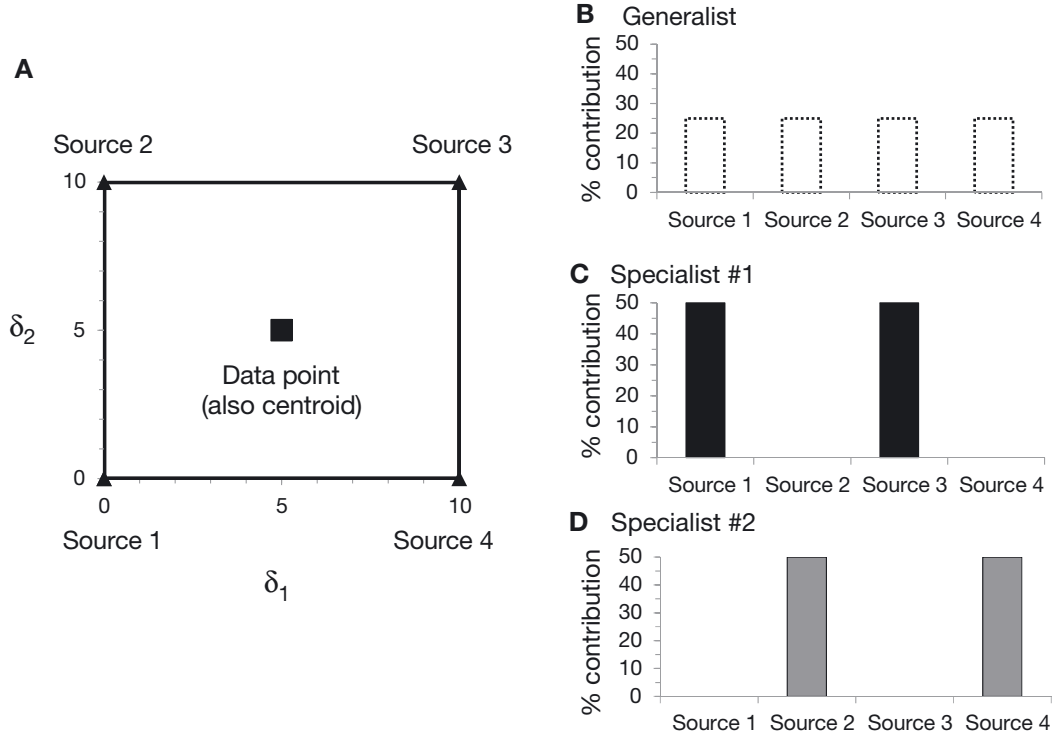


Fig. 1. Mixing model example for 4 sources and 2 tracers. Centroid point in the middle of (A) has no unique source assignment, and (B–D) show 3 of the infinite number of possible feasible solutions for % source contributions. Programs such as IsoSource and SIAR give the generalist solution (B) as the mean solution, but all other feasible solutions such as those shown in (C) and (D) are in fact equally probable compared to the (B) generalist solution

$$f_1 + f_2 + f_3 + f_4 = 1 \quad (1)$$

$$f_1 \times \delta_1\text{Source1} + f_2 \times \delta_1\text{Source2} + f_3 \times \delta_1\text{Source3} + f_3 \times \delta_1\text{Source4} = \delta_1\text{Sample} \quad (2)$$

$$f_1 \times \delta_2\text{Source1} + f_2 \times \delta_2\text{Source2} + f_3 \times \delta_2\text{Source3} + f_4 \times \delta_2\text{Source4} = \delta_2\text{Sample} \quad (3)$$

where f is the fractional contribution of each source. However, because there are 4 unknown fractional contributions and only 3 equations, no solution is reliably indicated. Instead, the data support alternative or feasible solutions rather than a uniquely determined solution.

This underdetermined situation is also evident graphically. Three of the feasible solutions are shown in Fig. 1, the even-handed 'all sources are equal' solution (Fig. 1B) versus 2 solutions where only 2 of the 4 sources contribute (Fig. 1C,D). Many more feasible solutions are possible that are a mix of the solutions shown in Fig. 1, and in fact there are an infinite number of feasible solutions for such underdetermined problems. To 'solve' underdetermined problems and distinguish between alternatives such as those shown in Fig. 1B–D, some additional assumptions are needed.

THREE APPROACHES TO SOLVING UNDERDETERMINED MIXING PROBLEMS

Statistical subsampling

One approach to solving underdetermined mixing problems is to use statistical subsampling of the range of feasible solutions and then calculate averages from the subsample. This is the approach used in the software packages IsoSource and SIAR. The mean source contribution values calculated by IsoSource for the center point in Fig. 1 are the same for each source, 25% contribution, with a reported SD of 15% for each of these sources. The minimum and maximum reported values are also the same for each source, 0 and 50% respectively. This output can be understood graphically, with Fig. 1B showing the mean solution, and Fig. 1C,D showing extremes where sources can contribute as little as 0% (minimum values) or as high as 50% (maximum values). Reporting the mean value and error of 25 ± 15% emphasizes one central solution (Fig. 1B), and relegates the others to error (Fig. 1C,D). The implicit assumption scientists make in reporting mean values from IsoSource is that frequency is equal to probability, but the underlying reality is that all these feasible solutions are equally probable. This may seem a semantic point, so to clarify the distinction between frequency and probability, imagine a shopkeeper who finds that someone is stealing from the store. The shopkeeper consults a colleague who recommends a frequentist solution, i.e. the most frequent customer is the most likely shoplifter. However, the shopkeeper knows the most frequent customer very well because this person is also the best customer, and for these reasons, decides to collect further evidence rather than relying on the assumption that frequency is equal to probability. Because it is only loosely expected that frequency might equal probability in some cases, and with many exceptions in other cases, IsoSource programming discourages the use of frequency-based means, even though means are reported in IsoSource, and also in SIAR. An early user of this frequency-based analysis did not find the approach wholly satisfactory, and in the end employed additional tracers that allowed a strongly determined, unique solution (Minagawa 1992). Ecologists also use observations such as gut content studies (Peterson 1999, Fanelli & Cartes 2010) to help constrain the range of feasible solutions.

The reader may also wish to note a truly semantic point, that the statistical 'means' reported in IsoSource and SIAR programming are used in a different way than scientists normally encounter, i.e. the statistical 'means' do not necessarily represent the mid-point value existing in the real world. This is because each of the feasible solutions considered by IsoSource and SIAR is in fact equally likely and cannot be excluded as a true and valid representation of the mid-point existing in the real world. The statistical 'means' gloss over this uncomfortable fact that all alternative solutions are equally likely, i.e. the 'mean' solution represented by Fig. 1B has no higher mathematical probability than the solution shown in Fig. 1C or Fig. 1D.

Geometric sector approach

A second approach to solving underdetermined mixing problems uses a geometric sector approach. This approach provides feasible source contributions for squares, which can be divided into sectors next to the sources using lines that pass through the data point of interest (Fig. 2). For the square in Fig. 2, the dividing lines are parallel to the sides of the square, resulting in a cross-hairs approach centered on the data point, and the contributions of the sources are inversely proportional to sector areas, i.e. the fractional source contribution for an individual source is calculated as:

Fraction = 1/(sector area for the source)/
{Σ [1/(all sector areas)]} (4)

This inverse approach was developed for mixing triangles (reviewed in Phillips 2001), but was largely replaced by the more flexible statistical programs that easily give solutions for the many types of polygons (Phillips & Gregg 2003). However, the geometric inverse approach follows the intuitively correct logic that the closer a data point is to a source, the more important is that source. Calculation of the source contributions for Fig. 2 with this geometric approach gives the 'all sources are equal' result, i.e. the Fig. 1B solution is indicated, not the Fig. 1C or 1D solutions. The assumption in this approach is that other geometric ways of estimating solutions are not important, an assumption that is questionable since, for example, a diagonal line can be drawn from Source 1 to Source 3 in Fig. 2 to explain the central data point as a 50/50 mix of these 2 sources. However, the geometric approach based on areas gives the 'all sources are equal' answer for Fig. 1, a solution that coincides with the frequentist mean (Fig. 1B). Ecologists can be tempted to recognize this mean as correct from the coincidence of central tendencies in the frequentist and geometric sector approaches.

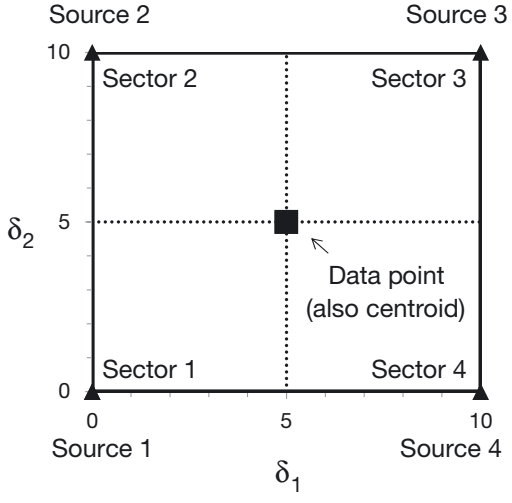


Fig. 2. Same mixing model as Fig. 1, but source contributions to the centroid data point are solved with a geometric approach rather than IsoSource or SIAR. In this approach, dotted lines divide the mixing polygon into sectors adjacent to sources, and the fractional source contribution for each individual source is calculated using inverse sector areas: fraction = [1/(sector area for the source)] / {Σ[1/(each sector area)]}. The 4 sectors have equal areas in this example, and each source contributes equally to the data point, the same generalist result shown in Fig. 1B

Linear mixing

It is also possible to examine these same underdetermined mixing problems with a third approach involving center points and sidewall points, still using a square mixing polygon as the example. The center point is calculated as the (average *x*, average *y*) of the source points, and is the centroid point in the middle in Fig. 3A. A sidewall point is on any line connecting 2 of the apex source points, e.g. the diamond sidewall point in Fig. 3A, is intermediate between apex source Points 1 and 2. In this approach, any data point can be understood as a mixture between the centroid and a sidewall point. The solutions for sidewall points are completely known and determined from 2-source mixing from mass balance equations (Fry 2006), e.g. for the sidewall point of Fig. 3A:

$$f_1 + f_2 = 1 \quad (5)$$

$$f_1 \times \delta_2\text{Source1} + f_2 \times \delta_2\text{Source2} = \delta_2\text{Sidewall point} \quad (6)$$

The solution of these 2 equations for the 2 unknowns f_1 and f_2 is unique:

$f_1 = (\delta_2\text{Sample} - \delta_2\text{Source1})/(\delta_2\text{Source2} - \delta_2\text{Source1})$
and
$f_2 = 1 - f_2$, so that $f_1 = (5–0)/(10–0) = 0.5$,
and
$f_2 = 1 - 0.5 = 0.5$ for the sidewall point in Fig. 3A (7)

In this approach, the centroid point solutions are unknown because they are in fact infinite and all equally likely (Fig. 1). In face of this uncertainty, this approach assumes that the source contributions for the centroid are divided equally among all sources at 25% each for the example in Fig. 3A, in effect dealing with uncertainty in the most even-handed way. The result is that each data point contains determined and undetermined aspects, with the sidewall contributions representing the determined aspects and the centroid contributions representing the evenly partitioned undetermined aspects. This approach cleanly divides the solutions into data-supported aspects (sidewall contributions) and assumed aspects (centroid contributions).

For simple regular polygons like squares of Figs. 1–3 and for 4-sided parallelograms, the underdetermined mixing problems can be viewed from this third perspective in a very routine and reliable way, i.e. this perspective yields the same mean results as the IsoSource and SIAR approaches based on frequencies (more irregular polygons are discussed below). Points that fall closer to the sidewall in regular polygons become more determined until they are fully determined via 2-source mixing at the sidewall itself.

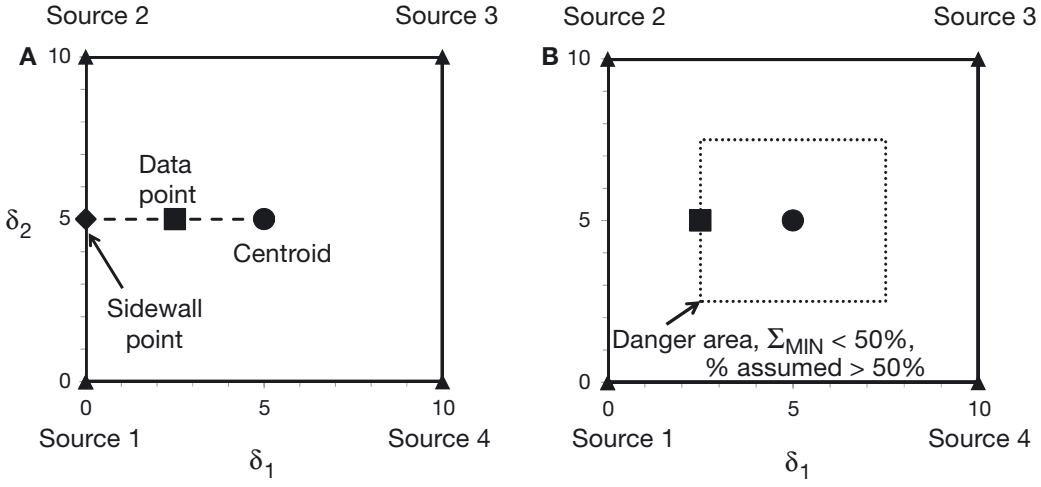


Fig. 3. Source mixing contributions viewed from a different perspective that uses 2-source mixing between a poorly constrained centroid and a uniquely constrained sidewall point. Sidewall points are determined uniquely as 2-source mixtures of nearest sources, the centroid point has an assumed 25/25/25/25 mixture of sources, and data points between the sidewall and centroid are 2-source mixtures of these end points. (A) Example shows 2-source mixing line between sidewall and centroid, with data point in the middle. (B) A dotted polygon halfway between sidewalls and the centroid encloses a central 'danger zone' for interpretations, where source contributions are >50% dominated by the assumed centroid contributions and are little constrained by the measured data. Interpretations are supported more strongly by measurement and less by assumption for points outside the 'danger zone' and closer to the sidewalls

But in the middle space, between the centroid and the sidewall, there is an intermediate level of confidence and solution, with lower confidence and more assumption accorded to points nearer the centroid, and more confidence and less assumption accorded to points nearer the sidewall. The overall confidence can be tracked graphically by examining how close data points fall to the centroid versus sidewalls. Mathematically, the minimum contributions calculated by IsoSource or SIAR for a data point can be summed (Σ_{MIN}) across sources to give the overall strength of solution, with 100% indicating a unique solution. Values for Σ_{MIN} range from 0 to 100% for the centroid point in Fig. 3A versus any sidewall point in Fig. 3A, respectively.

Generally, the Σ_{MIN} strength of solution is good to consider when evaluating the overall reliability of conclusions, especially whether conclusions are well supported by data with Σ_{MIN} > 50% or poorly supported by data with Σ_{MIN} < 50%. For parallelograms like the squares of Figs. 1–3, those points with Σ_{MIN} > 50% will be more than halfway to the sidewalls from the centroid points, and points in the center 1/4 of mixing diagrams have Σ_{MIN} < 50%. A visual representation of this central 'less solved' area can be made by constructing an interior polygon around the centroid, with walls of the interior polygon parallel to the main polygon and halfway to the center (Fig. 3B). Because data points falling in this central region are interpreted more by assumption than by measurement, scientists should be extra-cautious about reporting mean source values for data points from this 'danger zone', the dotted middle box in Fig. 3B.

There is an another metric besides Σ_{MIN} that is valuable for tracking source contributions:

$$\% \text{ resolved} = 100 - (\text{maximum} - \text{minimum}) \quad (8)$$

The % resolved values can be calculated when maximum and minimum source solutions are determined from IsoSource or SIAR programming. This metric identifies the better resolved or better constrained solutions as those where minimum and maximum estimates approach each other, with unique solutions indicated where maximum = minimum. The % resolved values for the centroid, data point and sidewall point in Fig. 3A are 50, 75 and 100%, respectively, and the % resolved values are perfectly correlated ($r^2 = 1$) with Σ_{MIN} in this graph. Generally, this metric helps identify regions where % resolution is >75% and the minima and maxima are only 25% apart, so that mean values are confined to a narrow 25% range and are relatively reliable. This >75% resolution zone is the area outside the dotted perimeter of the danger area given in Fig. 3B, so that both less assumed and more resolved solutions exist in this outer zone.

It is also good to consider an example of more complex, irregular mixing polygons that can be common (Phillips & Gregg 2003). In squares and in regular polygons with equilateral sides and angles and in 4-sided parallelograms, the centroid coincides with

the 'minimum' point, the point where minimum source contributions are lowest and usually equal zero. But in more irregular polygons, the minimum point or area can be relatively distant from the centroid, while the centroid is partially resolved and has non-zero source contributions. Fig. 4A shows such an example for an irregular polygon, a trapezoid, where the minimum point where Σ_{MIN} is 0% is shown as an open circle at the intersection of lines connecting opposite source points, while the centroid where Σ_{MIN} is 35%, shown as a closed circle, is located differently.

Solutions for source contributions in irregular polygons show several similarities and some differences to solutions in regular polygons. First, the determination of the source contributions for the minimum point open circle in Fig. 4A is still by assumption, with statistical programs giving the most even-handed partitioning among the 4 sources rather than involving only 2 sources. This is similar to the generalist versus specialist solutions of Fig. 1. The even-handed % generalist source contributions for the Fig. 4A minimum point are 10/10/40/40 for sources 1/2/3/4, respectively. While these contributions are not equal, they are even-handed in that the source contributions are as similar as possible for any of the feasible solutions. The 10/10/40/40 source mix is also even-handed, because dissimilarity between sources in minimized, i.e. it has the smallest possible total difference between the sources compared to other feasible solutions, such as a 20/0/80/0 mix. Second, both the irregular and regular polygons have a danger zone in the middle of the polygon where assumption dominates solutions. This danger zone with >50% assumed solutions is halfway to the sidewall from the minimum point (Fig. 4A) rather than from the centroid (Fig. 3B), but this zone occupies 25% of the larger mixing polygon in both cases. This is to be expected because mixing is proceeding between an assumed uncertain center to a certain sidewall. Lastly, the asymmetrical shape of irregular polygons gives a more complex differentiation of % resolution among different sources. For a data point in regular polygons, % resolution is the same for all sources, but in irregular polygons, % resolution can vary for the different sources. For the example shown in Fig. 4B, there is a large area where Sources 3 and 4 are poorly resolved (<75%), but in contrast, Sources 1 and 2 are acceptably resolved (>75%) throughout this whole trapezoid. The reason Sources 3 and 4 are poorly resolved in this example stems from the geometry of Fig. 4B, where Sources 3 and 4 lie close together and so can be rather easily substituted for each other in the mixing calculations. In contrast, Sources 1 and 2 are well-separated from each other and from Sources 3 and 4, and so have high % resolved values throughout the polygon. Interestingly, the % resolved values are high for Sources 1 and 2 even in some cases where assumption is strong, near the minimum point in Fig. 4A. This is because the maximum possible Source 1 and 2 contributions in this region are close to zero and show little difference from minimum values that are also near zero. Thus, in this region where Sources 1 and 2 can only possibly make relatively small contributions in the 0–25% range, their contributions are >75% resolved, and mean estimates for these source contributions will be fairly reliable even though the region has high overall assumption (Fig. 4A).

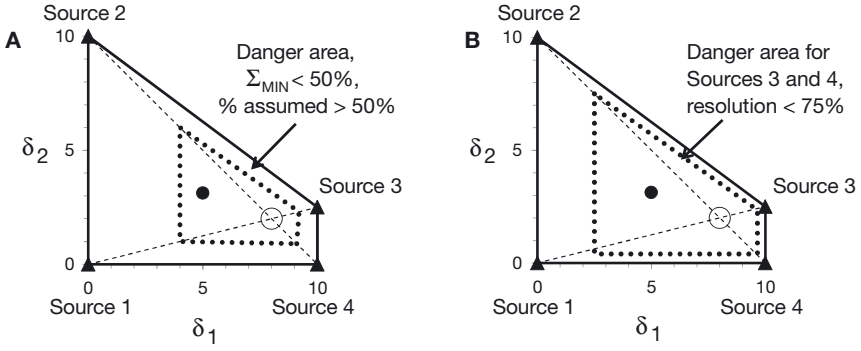


Fig. 4. Source mixing in an irregular polygon, a trapezoid. In this example, the centroid (●) is not co-located with the 'minimum' point (○) where minimum source contributions are zero (Σ_{MIN} = 0) at the intersection of lines connecting opposite sources. Similar to the situation for the center point in Fig. 1, a wide range of generalist and specialist solutions are feasible for this minimum point, and source contributions are given only by assumption. (A) Danger area is outlined by the central polygon, halfway from the minimum point to the sidewall. Solutions are >50% assumed in this area. (B) Some sources are better resolved than others in irregular polygons. The interior dotted trapezoid outlines the relatively large area where resolution is poor (<75%) for Sources 3 and 4, an area that includes the centroid. Contributions from Sources 1 and 2 are relatively well resolved throughout the polygon, as explained in the text

In summary, there are some general features and also nuances to making geometric assessments of mixing model results. For data points in regular polygons like the square of Fig. 3, $\sum_{MIN}$ and % resolved values for source contributions vary together and record the same mixing dynamic between centroid and sidewall points (Fig. 3). But for more irregular polygons, the % resolved values can vary for different sources, reflecting the irregular shape of the polygon. Many polygons are regular or fairly regular, so that the simpler Fig. 3 models apply. But for more irregular polygons where centroids and minimum points are well-separated (e.g. Fig. 4), ecologists may need to use geometry and statistical programming to find the minimum point or area that is the center of the danger zone with $\sum_{MIN}$ <50% and where strong assumption is involved in interpretations. For these irregular polygons, ecologists may also wish to use statistical programming to generate several versions of the same polygons to show the low resolution zones for each of the sources used in the mixing polygon. These low resolution zones can overlap and be combined as shown for Sources 3 and 4 in Fig. 4B, and generally, use of such graphs helps develop a visual understanding of zones where interpretations about source contributions are more and less well-constrained by data. Ecologists can benefit from studying the geometry of their mixing polygons and try to develop a visual understanding of their data to complement any mathematical or statistical interpretations.

The examples of Figs. 1–4 give a sampling of underdetermined tracer problems encountered in mixing polygons, but more complex polygons are possible. Given source values, IsoSource or SIAR will forward-calculate the frequency distribution of solutions for more complex polygons, and it is convenient to use those calculations to find points or areas where minimum contributions are at or near zero. These areas are the danger zones of Figs. 3 & 4, generally where $\sum_{MIN}$ is less than 50% and strong assumption is needed to back-calculate or unmix the source contributions to specific values. More complex mixing polygons can also give results that are not intuitively even-handed, but result from the statistical forward calculations with random resampling. Considering an example having 3 sources with 0, 5 and 10 values and a sample in the middle with a value of 5, the IsoSource-calculated source contributions are 25/50/25 and not the more intuitive one-third from each source. In these cases, the even-handedness arises more from random resampling than intuition, but the main point remains the same, that when samples are near the middle in these underdetermined mixing problems, unmixed actual solutions are only achieved by assumption.

JUSTIFICATIONS FOR SOLVING BY ASSUMPTION

The finding that undetermined models are being solved by assumption is not surprising, and ecologists using mean solutions should state why they have confidence in the assumed solutions. The underlying assumption that becomes apparent in the third approach using polygons with centroid-sidewall mixing (Fig. 3B) or mixing between minimum and sidewall points (Fig. 4A) is an even-handed partitioning of uncertainty among the alternative sources, a common assumption of modelling that is often justified by assuming that the result will 'not be too far wrong'. That is, an even-handed solution is less likely to be wrong than a specialist solution. However, biology works with both means and extremes in processes such as evolution, and some thought needs to go into accepting mean solutions that are underdetermined. These solutions are least likely to be wrong for generalist situations where all sources are in play, and for more integrated ecological situations, for example for high-level carnivores that consume a variety of more specialist low-level consumers in a food web. On the other hand, the assumption of even-handed partitioning of the uncertainty will usually assign some importance to all sources in mixing diagrams so that low-level mean contributions appear for all sources. And to the extent that more and more sources are considered, the solutions become more and more diffuse as these low-level contributions accumulate. These considerations also lead to the realization that the choice of sources also importantly determines mean solutions, with more sources giving more diffuse mean solutions and fewer sources giving more sharply contrasting mean solutions. Experience shows that aggregating sources to a minimum number is advisable, and guidelines have been developed to help with this aggregation (Phillips et al. 2005).

It is perhaps good to close this section about assumptions with the thought that natural abundance isotope analyses are really performed to exclude possible solutions when sources have low maximum importance and graphically are far away from data points. But using underdetermined mean solutions is an attempt to pinpoint actual solutions. This pinpointing is not really possible in science or ecology because of the general objection that other uncon-

sidered factors may be important. This objection is quite realistic for many source-mixing problems where unknown sources may have been missed in sampling, then omitted in the mixing polygons, culminating in a flawed solution. For such reasons, it is often advisable to use mean values with great caution because they extrapolate and come to solutions perhaps too quickly. The more cautious scientist is often satisfied leaving some unknowns for future work.

CASE STUDY

The foregoing does not focus on the several positive features of programs like IsoSource and SIAR. These features include a uniform way to evaluate mixing dynamics, incorporation of error terms (e.g. variation in field replicates, in source isotope values, and in trophic enrichment factors or fractionations), and easy mathematical manipulation of model scenarios and model inputs. With this in mind, results from a published study were re-examined to identify robust results arising from IsoSource models, with the criteria for robustness being that trends found in means would also be evident in minimum and maximum values for the various feasible contributions.

The study chosen concerned crabs sampled from mangrove creeks near a shrimp mariculture facility in Thailand, where the intent was to use C and N isotopes to track any food web alterations for local crabs due to effluent discharged from the facility (Kon et al. 2009). The mariculture ponds discharged into an 'impact' creek sampled at its headwaters next to the pond. Sites farther away in mid-creek and farthest away at the creek mouth were sampled in the 'impact creek', and similar sampling was done at 3 stations in a 'reference' creek system located nearby. Five organic matter sources were considered as possibly important in the crab food web, including shrimp feed, mangrove leaf litter, microphytobenthos (MPB), phytoplankton and epiphytic macroalgae. Only 2 tracers were used, and the IsoSource results indicated an underdetermined outcome with many feasible solutions. Authors reported mean IsoSource values along with minimum and maxima for feasible contributions in their Table 2, and averages of those tabulated values for the various crab species are plotted in Fig. 5. Generally, crabs from the reference creeks showed high food web reliance on mangrove and MPB, while crabs in the headwater reaches of the impact stream next to the shrimp farm showed some contributions and impacts from the farm (Fig. 5).

Considered in more detail, there were several trends that were common in the Kon et al. (2009) results calculated for means, minima and maxima, and these trends could be considered robust. Three common trends visible by close inspection of Fig. 5 were (1) food web inputs from the shrimp farm were highest in the impact creek headwater site nearest the shrimp farm, (2) the normal mangrove contributions to diets were lower at headwater and mid-creek sites near the farm than in the reference creek, and (3) MPB contributed more to crab diets at the mouths of creeks than at the more shaded mid-creek and headwater areas (Fig. 5). Because these 3 results emerge from all 3 ways of viewing the data, they can be considered robust. In contrast, some detailed source assignments visible in the mean solutions were not paralleled in the solutions for minima and maxima and should not be considered robust. In particular, 2 trends were detected in mean and maximum data output (Figs. 5A,C) but not in minimum output (Fig. 5B): (1) low-level but apparently widespread influence of shrimp farm foods in both impact and reference creeks, and (2) moderate importance of epiphyte and phytoplankton foods at all sites (Fig. 5). These 2 mean results probably should be discounted and de-emphasized in final interpretations because they are just those low-level contributions expected to arise from assumed rather than measured parts of the analysis, i.e. the assumption that all sources contribute to solutions will force all sources to acquire importance, a questionable interpretation since it comes from assumption rather than measurement. Overall, $\sum_{MIN}$ averaged slightly more than 50%, and % resolved averaged across individual sources was slightly more than 75%, indicating that the system was moderately solved in spite of being underdetermined (Fig. 5). However, other studies show much lower resolution and $\sum_{MIN}$ in the 3–20% range (e.g. Melville & Connolly 2003). As Melville & Connolly (2003) point out, additional tracers and approaches are likely needed to come to satisfactory evaluation of such strongly underdetermined mixing problems.

OTHER APPROACHES

Underconstrained mixing models can be solved if enough assumption is blended with measured data. This mixture can make it difficult to understand how to evaluate and interpret results, but some general lessons have emerged over the past years that help ecologists deal with underdetermined systems. One important lesson is that using additional qualitative

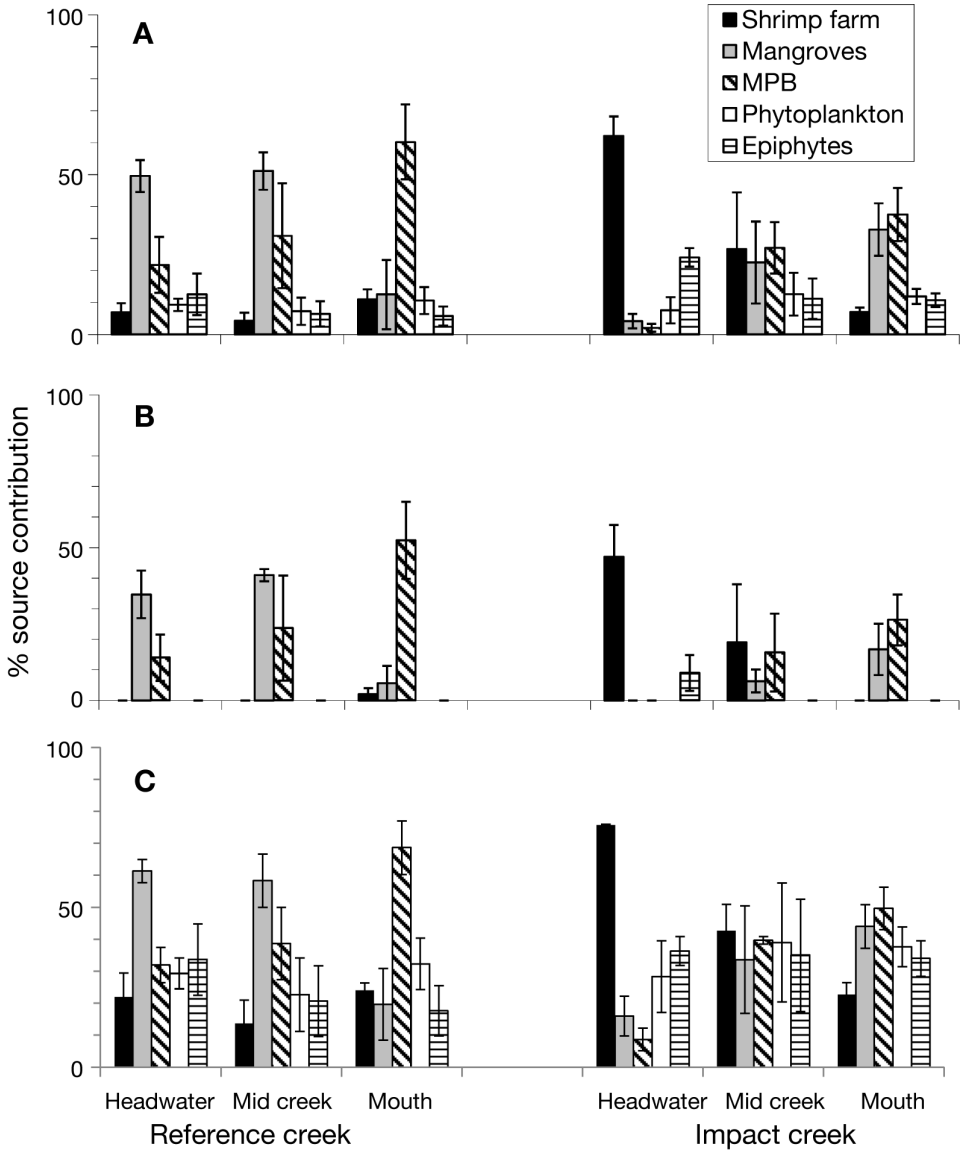


Fig. 5. Example showing how to make robust interpretations using output from IsoSource programming where there are too many sources and not enough tracers. The data are taken from a published study of crabs in a mangrove estuary in Thailand (Kon et al. 2009, averages for crab species from their Table 2). Food source contributions for mangrove crabs were investigated in reference and impact creeks near a shrimp farm that discharged wastewater into the headwater site of the impact creek. Panels show source contributions calculated by IsoSource programming for (A) means, (B) minima and (C) maxima. Careful inspection shows some common, robust results that appear in all 3 panels, but also trends in the mean results (A) that are much less robust and not supported by results for minimum contributions (B), as detailed in the text in 'Case study'. For example, most of the lower-level (5–15%) contributions for phytoplankton, epiphytes and shrimp farm effluents indicated in means (A) do not appear in the minima (B). In this case, the mean results are unsupported by the more conservative view of the data represented by the minimum values (B), and interpretation of these mean values is not robust based on only the isotope measurements. Key: Shrimp farm = shrimp food or shrimp farm ecosystem; Mangrove = mangrove or mangrove ecosystem (mangal); MPB = marine mudflat ecosystem; Epiphytes = epiphytic macroalgae growing on mangroves. Error bars are SE

information can be very powerful in constraining otherwise underdetermined solutions by narrowing the ranges of reported minimum and maximum source contributions (see 'additional constraints' link at IsoSource website, www.epa.gov/wed/pages/models/stableIsotopes/isosource/isosource.htm). The additional information often can be obtained in food web studies where observations of foraging behaviour and gut contents are traditional complements to isotope tracer work (Peterson 1999). For example, ecologists may know that time spent foraging on Source A is more than on Source B, or that sources

A and B are consumed in about equal quantities, but that the nutritional quality of Source A is much greater than that of Source B. With such information, model output can be sorted for only solutions where Source A contributions are greater than those of Source B, and such sorting has the beneficial effect of much reducing the range of feasible solutions, i.e. what was a broad range of solutions becomes narrow, close to a completely determined solution. But in many cases, scientists want to use the isotope tracer data as independent evidence, and so keep a focus on using measured mixing polygons.

Source selection

Experience shows that keeping these mixing polygons as simple as possible is best, with only important sources included (e.g. Fanelli & Cartes 2010). Guidelines exist for aggregating and reducing the number of sources (Phillips & Gregg 2003), but generally, ecologists rely on several lines of evidence to help decide which sources to include and which to exclude. These approaches include gut content studies, feeding observations and experimental results (e.g. Laurand & Riera 2006, Fanelli & Cartes 2010). When considering source inclusion or exclusion for mixing polygons, graphing initial source values and discarding unimportant sources may be acceptable in some cases. It is tempting to add more sources, because including more sources seems to better mimic complex real situations, but experience shows that adding more sources will also inevitably increase internal errors and uncertainty. For these reasons, attention to source selection is important for these mixing diagrams, and generally selecting the simplest source array is highly desirable. Many scientists try to simplify mixing problems to 2-source or 3-source problems that can be solved unambiguously with 1 or 2 tracers, respectively (Phillips et al. 2005, Newsome et al. 2007). In these 2- and 3-source problems, models are no longer underdetermined and unique solutions are possible. For example, with the common situations of only 2 tracer measurements in $\delta^{13}C \times \delta^{15}N$ biplots, aggregating sources to 3 sources to form a mixing triangle will allow unique solutions, in line with the general observations that n tracer measurements allow unique solutions for $n + 1$ sources. For example, aggregating Sources 3 and 4 in Fig. 4A would allow unique solutions for all data points in the mixing polygon, in effect trading off some losses in ecological resolution of particular sources for large gains in mathematical power. However, ecologists often prefer to keep more complex source mixtures in polygons because they are interested in resolving the individual, unaggregated source contributions.

Varieties of tracer approaches

Ecologists working with complex mixtures have developed several approaches, including the procedures recommended with Fig. 5, for finding the common results between calculated mean, minimum and maximum IsoSource solutions. Other approaches consider how consumer isotopes shift in time or space, testing whether consumer isotopes shift to follow isotopes of suspected important food resources (McCutchan & Lewis 2002, Melville & Connolly 2003, Finlay 2010, Rasmussen 2010). Average source values can be regressed against other independent source mixture indicators, such as gut content information, to test for significant correlations expected for converging lines of evidence. This works best if the source mixtures are in fact fairly well known, with the isotopes helping by essentially fine-tuning the source estimates (Rutz et al. 2010). Another approach is to fit triangles closely around measured consumer data points to help identify sources near the apices of the triangle that are most important (Layman 2007). Triangles have the advantage of providing unique solutions in isotope biplots such as $\delta^{34}S$ versus $\delta^{13}C$ (Newsome et al. 2007), although sources are often aggregated at the habitat level in top-down triangles constructed around consumers. Shifting the mixing axis from specific food sources to aggregated sources indicated by certain consumers (e.g. benthic versus pelagic consumers) is often a useful way to aggregate source data and find useful source information in food web data (Demopoulos et al. 2008, Grippo et al. 2011). Where food sources differ strongly in C, N or S element concentrations, alternative concentration-dependent modelling provides important alternative estimates of source contributions for food web problems (Phillips & Koch 2002, Bouillon et al. 2008, www.epa.gov/wed/pages/models/stableIsotopes/isoconc/isoconc1_01.htm). Tracer addition experiments can greatly help to test source contributions that are poorly resolved with natural abundance isotope data (Peterson et al. 1993), but these addition experiments need careful time course sampling, knowledge of turnover dynamics and good modelling for correct interpretation of results (Fry 2006, Oakes et al. 2010, Lee et al. 2012).

Explicit and implicit sources

Ecologists have also proposed using the measured isotope data directly to infer mixing dynamics without using explicit source values (Bearhop et al. 2004, Layman et al. 2007). This approach has been criticized because it is hard to know if a 1‰ range or SD represents a large or small difference until the bracketing source values are considered (Hoeinghaus & Zeug 2008), e.g. a 1‰ SD is a large error if sources are separated by only 2‰, but a much smaller error if sources are separated by 10‰. If the source isotope values are stable across systems, then valid comparisons can still be made without the explicit source evaluation, but unfortunately, source isotope values commonly show a local differentiation related to local complexities of biogeochemical cycling (Fry & Sherr 1984). Ecologists have been frustrated by this situation not only because it requires much local sampling effort but also because it is often very difficult to sample the sources most important to animals. Animals are good at finding and consuming important foods so that ecologists undersample and miss those important foods. In such cases, ecologists still hope that their source sampling will at least provide representative isotope values for foods in the area. This situation obviously is less than ideal, and future isotope tracer work might consider using a 'community referenced' assessment approach that is based on the consumer measurements. For example, consider a fish food web where average values of planktonic, benthic and top carnivore fish are measured and could be used to generate a combined community background SD or reference ellipse (Jackson et al. 2011) of isotope variation at a sampling location. Then the individual and species-level fish isotope variation (Bearhop et al. 2004, Layman et al. 2007) could be assessed against this area to generate a 'community referenced' value. This type of locally referenced metric could aid ecologists who are seeking to understand patterns of trophic change in natural and human-disturbed food webs (Cucherousset et al. 2012).

More tracers

Isotopes alone often are not enough to solve environmental problems (Peterson 1999). Scientists in other fields use many more than 2 tracers to completely solve mixing problems with 4–8 sources (Schauer et al. 1996, Belicka & Harvey 2009, Collins et al. 2010). Employing more tracers to deal with multiple sources (Weiss et al. 2008, Griffith et al. 2009) is probably the long-term solution for currently underconstrained isotope mixing problems.

RECOMMENDATIONS

Underconstrained mixing models force ecologists to think about the balance between assumption and testing in making isotope interpretations and in judging model output. Almost always, models show that a mixture of sources is important in explaining measured data; it is just pinpointing the details of these mixtures and actual % contributions of various sources that is at question. An important overall message from this review is that for many points in underconstrained mixing diagrams, source contributions can be obtained only by assumption. Assuming solutions seems anti-scientific, but modellers often work with underconstrained solutions in formulating hypotheses and in performing sensitivity analysis associated with various scenarios. These are legitimate modelling exercises, but scientists interested primarily in interpretations based on strong empirical testing should probably focus on and interpret data falling closer to sidewalls of mixing diagrams, where source contributions are known with the most confidence. It is because mean solutions for underdetermined mixing models often contain a substantial element of assumption that the following text appears in bold on the IsoSource website and also with every set of solutions calculated by IsoSource: 'To avoid misrepresenting the results, users of this (IsoSource) procedure should report the distribution of feasible solutions rather than focusing on a single value such as the mean' (www.epa.gov/wed/pages/models/stableIsotopes/isosource/isosource.htm). The SIAR programming lacks similar warnings (Parnell et al. 2010), but similar considerations apply to isotope-based calculations from SIAR. But in spite of such caveats, many ecologists use mean values and error estimates given by IsoSource and SIAR. With this use in mind, the following 6 practical steps give a framework for working with these means of feasible solutions (Fig. 1B–D), using the example of an underdetermined isotope food web problem.

(1) In planning your study, recognize that in many cases, source contributions are hard to resolve with isotopes, so that you likely should use other approaches to help test your main hypotheses. Use some preliminary sampling to see if isotopes will be helpful. Go ahead with isotopes if they appear promising, but also make the isotope work just one of the approaches you are using to answer your scientific

questions. The isotope approach may work best with multiple tracers and using comparisons made possible by extensive isotope sampling in time and space as noted above under 'other approaches'. But most studies focus within a single system and just C and N isotopes at natural abundance levels because these measurements are relatively inexpensive and allow much replication. The following steps apply to such dual tracer studies.

(2) At the end of the study with isotopes, plot your data along with the sources and connect the sources with lines to form a mixing polygon. Consult the literature for reasonable estimates of trophic enrichment factors (TEFs) and use those factors to generate a second graph where all data has been corrected for trophic level effects (Post 2002). Work with both the original and TEF-corrected graphs in the following, so that you know the effect trophic corrections have on your interpretations. If TEF effects are large for especially nitrogen, you may want to consider using carbon isotopes alone in mixing models (Sierszen et al. 2006, Grippo et al. 2011). In any event, you may want to publish one or both of these plots along with any output from IsoSource or SIAR, so that it is clearer how data are being interpreted.

(3) Use your data plots to visually check if one or more sources is far from data points and so has low maximum importance. This is a strong and interesting result, with isotopes ruling out or disproving major contributions from those far-away sources. This disproof based on 'low maximum importance' is often the strongest use of isotopes. Examine output from IsoSource or SIAR to obtain the highest numerical values that are the maximum contributions, and record those as needed.

(4) If you have multi-source data with not enough tracers and too many sources in your data plots, see if you can simplify the source mix by deleting sources that Step 3 indicates are not important, or by aggregating sources. Simpler source mixtures can give unique solutions. For example, for C and N isotope plots, a triangle of sources will produce uniquely determined solutions that are 100% solved, and SIAR is an excellent program for propagating errors around those mean values. If you have 3 tracers, you can uniquely solve for 4 sources; if you have 4 tracers, you can resolve 5 sources, etc. The IsoSource and SIAR mixing model programs can work with these increasing numbers of tracers and dimensions, but you need to use unweighted isotope values (not multivariate inputs) that reflect underlying model assumptions of mass balance and simple accounting of (ultimately) numbers of atoms.

(5) If you are still not satisfied with maximum contributions (Step 3) and do not want to aggregate sources (Step 4) but decide to keep multiple sources and work with means in an underdetermined system (e.g. Fig. 1), add a smaller interior polygon enclosing the area where points are >50% assumed. This interior polygon has walls half-way to sidewalls from the centroid for regular and near-regular polygons (Fig. 3B), while for more complex polygons such as the trapezoid of Fig. 4, the point or area with minimum source contributions can be identified with IsoSource or SIAR and used as the center point for this interior polygon (Fig. 4A). Use these interior polygons as visual guides for a first sorting of solutions along a gradient from generally 'very undetermined' near the center of the interior polygon to 'very determined' near sidewalls. This is a gradient that tracks increasing values of $\sum_{MIN}$ towards the sidewalls. For irregular polygons, a second sorting may be needed using the % resolved metric for individual sources, as outlined in the Fig. 4B example and accompanying text. Then think about emphasizing solutions that are most robust, with the highest $\sum_{MIN}$, where the % resolved value is high (>75%) and where means, minima and maxima give parallel results (Fig. 5). If you report solutions for all data, report the $\sum_{MIN}$ and % resolved alongside the mean values, or use colours or darker shading of the means and errors to indicate solutions that are most robust with higher $\sum_{MIN}$ and % resolved.

(6) Finally, think about the other knowledge you have of the system, and examine the data plots and mean solutions from another point of view, asking the question 'Are these data points in the right place?' The points may occur in the center where the math does not give a unique solution, but you may still find that the location of these data is consistent with your other ecological knowledge of the system. Conversely, points near the center may make you think differently about the other types of ecological information that you have. Here you are really using isotopes as checks on other lines of evidence, asking if isotopes do not disprove those other interpretations and so support them. In fact, the most powerful uses of isotopes usually occur in this manner, by combining isotope information with other lines of evidence, even when isotopes alone offer fairly unconstrained solutions.

Acknowledgements. R. Baker, K. Galvan, T. Jardine, K. Rogers and M. Sheaves provided critical comments on initial versions of this manuscript. Comments from anonymous reviewers also helped shape and focus the manuscript.

LITERATURE CITED

Bearhop S, Adams CE, Waldron S, Fuller RA, Macleod H (2004) Determining trophic niche width: a novel approach using stable isotope analysis. J Anim Ecol 73:1007–1012

Belicka LL, Harvey HR (2009) The sequestration of terrestrial organic carbon in Arctic Ocean sediments: a comparison of methods and implications for regional carbon budgets. Geochim Cosmochim Acta 73:6231–6248

Benstead JP, March JG, Fry B, Ewel KC, Pringle CM (2006) Testing IsoSource: stable isotope analysis of a tropical fishery with diverse organic matter sources. Ecology 87: 326–333

Bouillon S, Connolly RM, Lee SY (2008) Organic matter exchange and cycling in mangrove ecosystems: recent insights from stable isotope studies. J Sea Res 59:44–58

Collins AL, Walling DE, Webb L, King P (2010) Apportioning catchment scale sediment sources using a modified composite fingerprinting technique incorporating property weightings and prior information. Geoderma 155:249–261

Cucherousset J, Bouletreau S, Martino A, Roussel JM, Santoul F (2012) Using stable isotope analyses to determine the ecological effects of non-native fishes. Fish Manag Ecol 19:111–119

Demopoulos AWJ, Cormier N, Ewel KC, Fry B (2008) Use of multiple chemical tracers to define habitat use of Indo-Pacific mangrove crab, *Scylla serrata* (Decapoda: Portunidae). Estuaries Coasts 31:371–381

Fanelli E, Cartes JE (2010) Temporal variations in the feeding habits and trophic levels of three deep-sea demersal fishes from the western Mediterranean Sea, based on stomach contents and stable isotope analyses. Mar Ecol Prog Ser 402:213–232

Finlay JC, Doucett RR, McNeely C (2010) Tracing energy flow in stream food webs using stable isotopes of hydrogen. Freshw Biol 55:941–951

Fry B (2006) Stable isotope ecology. Springer, New York, NY

Fry B, Sherr E (1984) $\delta^{13}C$ measurements as indicators of carbon flow in marine and freshwater ecosystems. Contrib Mar Sci 27:13–47

Griffith DR, Barnes RT, Raymond PA (2009) Inputs of fossil carbon from wastewater treatment plants to US rivers and oceans. Environ Sci Technol 43:5647–5651

Grippo M, Fleeger JW, Dubois SF, Condrey R (2011) Spatial variation in basal resources supporting benthic food webs revealed for the inner continental shelf. Limnol Oceanogr 56:841–856

Hoeinghaus DJ, Zeug SC (2008) Can stable isotope ratios provide for community-wide measures of tropic structure? Comment. Ecology 89:2353–2357

Jackson AL, Inger R, Parnell AC, Bearhop S (2011) Comparing isotopic niche widths among and within communities: SIBER—stable isotope Bayesian ellipses in R. J Anim Ecol 80:595–602

Kon K, Kawakubo N, Aoki J, Tongnunui P, Hayashizaki K, Kurokura H (2009) Effect of shrimp farming organic waste on food availability for deposit feeder crabs in a mangrove estuary, based on stable isotope analysis. Fish Sci 75:715–722

Laurand S, Riera P (2006) Trophic ecology of the supralittoral rocky shore (Roscoff, France): a dual stable isotope ($\delta^{13}C$, $\delta^{15}N$) and experimental approach. J Sea Res 56:27–36

Layman C (2007) What can stable isotope ratios reveal about mangroves as fish habitat? Bull Mar Sci 80:513–527

Layman CA, Arrington DA, Montana CG, Post DM (2007) Can stable isotope ratios provide for community-wide measures of trophic structure? Ecology 88:42–48

Lee KM, Lee SY, Connolly RM (2012) Combining process indices from network analysis with structural population measures to indicate response of estuarine trophodynamics to pulse organic enrichment. Ecol Indic 18:652–658

Lubetkin SC, Simenstad CA (2004) Multi-source mixing models to quantify food web sources and pathways. J Appl Ecol 41:996–1008

McCutchan JH Jr, Lewis WM Jr (2002) Relative importance of carbon sources for macroinvertebrates in a Rocky Mountain stream. Limnol Oceanogr 47:742–752

Melville AJ, Connolly RM (2003) Spatial analysis of stable isotope data to determine primary sources of nutrition for fish. Oecologia 136:499–507

Minagawa M (1992) Reconstruction of human diet from $\delta^{13}C$ and $\delta^{15}N$ in contemporary Japanese hair: a stochastic method for estimating multi-source contribution by double isotopic tracers. Appl Geochem 7:145–158

Newsome SD, Martinez del Rio C, Bearhop S, Phillips DL (2007) A niche for isotopic ecology. Front Ecol Environ 5:429–436

Oakes JM, Connolly RM, Revill AT (2010) Isotope enrichment in mangrove forests separates microphytobenthos and detritus as carbon sources for animals. Limnol Oceanogr 55:393–402

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 (1999) Stable isotopes as tracers of organic matter input and transfer in benthic food webs: a review. Acta Oecol 20:479–487

Peterson B, Fry B, Deegan L, Hershey A (1993) The trophic significance of epilithic algal production in a fertilized tundra river ecosystem. Limnol Oceanogr 38:872–878

Phillips DL (2001) Mixing models in analyses of diet using multiple stable isotopes: a critique. Oecologia 127:166–170

Phillips DL, Gregg JW (2003) Source partitioning using stable isotopes: coping with too many sources. Oecologia 136:261–269

Phillips DL, Koch PL (2002) Incorporating concentration dependence in stable isotope mixing models. Oecologia 130:114–125

Phillips DL, Newsome SD, Gregg JW (2005) Combining sources in stable isotope mixing models: alternative methods. Oecologia 144:520–527

Post DM (2002) Using stable isotopes to estimate trophic position: models, methods, and assumptions. Ecology 83: 703–718

Rasmussen JB (2010) Estimating terrestrial contribution to stream invertebrates and periphyton using a gradient-based mixing model for $\delta^{13}C$. J Anim Ecol 79:393–402

Rutz C, Bluff LA, Reed N, Troscianko J and others (2010) The ecological significance of tool use in New Caledonian crows. Science 329:1523–1526

Schauer JJ, Rogge WF, Hildemann LM, Mazurek MA, Cass GR, Simoneit BRT (1996) Source apportionment of airborne particulate matter using organic compounds as tracers. Atmos Environ 30:3837–3855

Sierszen ME, Peterson GS, Trebitz AS, Brazner JC, West CW (2006) Hydrology and nutrient effects on food web structure in ten Lake Superior coastal wetlands. Wetlands 26:951–964

Weiss DJ, Rehkdmper M, Schoenberg R, McLaughlin M and others (2008) Application of nontraditional stable-isotope systems to the study of sources and fate of metals in the environment. Environ Sci Technol 42:655–664

Editorial responsibility: Christine Paetzold, Oldendorf/Luhe, Germany

Submitted: September 3, 2012; Accepted: November 8, 2012
Proofs received from author(s): December 13, 2012