

Donald L. Phillips

Mixing models in analyses of diet using multiple stable isotopes: a critique

Received: 30 January 2000 / Revision: 20 July 2000 / Accepted: 10 October 2000 / Published online: 10 January 2001
 © Springer-Verlag 2001

Stable isotope analysis is used frequently to determine the relative contributions of different food sources to an animal's diet (Hobson 1999). Isotopic ratios for the animal tissues and each of its potential food sources are determined. The similarity of the ratios for the animal tissues with those of individual food sources (after correcting for fractionation during digestion and assimilation) gives an idea of their relative importance in the diet; in other words "you are what you eat" (DeNiro and Epstein 1978). Two food sources can be partitioned using the isotopic ratio for a single element (e.g., $\delta^{13}\text{C}$), or three food sources can be partitioned using isotopic ratios for two elements (e.g., $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$) (Kwak and Zedler 1997). A number of recent papers have used geometric procedures to quantify the contributions of three food sources to the diet using $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ (Ben-David et al. 1997a, 1997b; Kline et al. 1993; Szepanski et al. 1999; Whitlege and Rabeni 1997). However, these methods do not provide correct solutions to this three-endmember mixing problem. The purpose of this paper is to point out the shortcomings of these methods and to propose an alternative procedure which avoids them.

Figure 1 shows a graphical representation of the analytical situation. The dietary isotopic composition is represented by point *D* within the triangle bounded by the points for the adjusted food source isotopic compositions *A'*, *B'*, and *C'*. In the geometric procedures, Euclidean distances are calculated for line segments *DA'*, *DB'*, and *DC'* and are used to compute the dietary contributions. Several variations of this calculation have been utilized.

Kline et al. (1993) used the following equation:

$$\%X \text{ in diet} = \left(1 - \frac{(DA' + DB' + DC') - DX}{DA' + DB' + DC'}\right) \times 100 \quad (1)$$

which simplifies to:

$$\%X \text{ in diet} = \left(\frac{DX}{DA' + DB' + DC'}\right) \times 100 \quad (1a)$$

where *X* represents one of the food sources *A*, *B*, or *C*, and *DX* represents the corresponding line distance *DA'*, *DB'*, or *DC'*.

Whitlege and Rabeni (1997) used reciprocal distances in the formula:

$$\%X \text{ in diet} = \left(\frac{(1/DA' + 1/DB' + 1/DC') - 1/DX}{1/DA' + 1/DB' + 1/DC'}\right) \times 100 \quad (2)$$

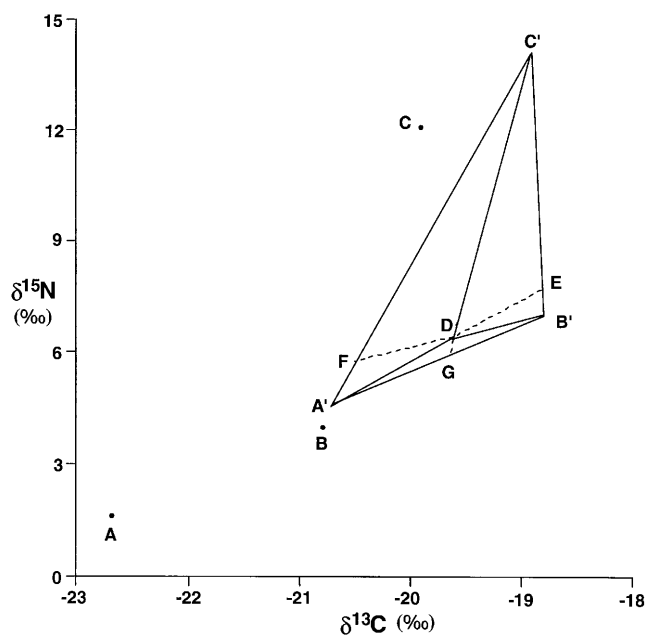


Fig. 1 Plot of dual isotopic compositions of food sources *A*, *B*, and *C*, and predator *D*. *A'*, *B'*, and *C'* represent the food source isotopic composition after adjustment for trophic fractionation. As in Table 1, numerical values for this example were taken from mean isotopic ratios reported by Szepanski et al. (1999) for *A* moose, *B* caribou, *C* salmon, and *D* interior Alaska wolf

D.L. Phillips (✉)
 U.S. Environmental Protection Agency,
 National Health and Environmental Effects Research Laboratory,
 200 SW 35th St., Corvallis, OR 97333, USA
 e-mail: don@mail.cor.epa.gov
 Fax: +1-541-7544799

which simplifies to:

$$\%X \text{ in diet} = \left(1 - \frac{1/DX}{1/DA' + 1/DB' + 1/DC'}\right) \times 100 \quad (2a)$$

Ben-David et al. (1997a, 1997b) and Szepanski et al. (1999) similarly used reciprocal distances but in a slightly different equation:

$$\%X \text{ in diet} = \left(\frac{1/DX}{1/DA' + 1/DB' + 1/DC'}\right) \times 100 \quad (3)$$

While the geometric representation in Fig. 1 is a nice heuristic tool, and it intuitively makes sense that the dietary proportions are somehow related to the lengths of the line segments between the animal tissue and its dietary components, none of the formulas that utilize these Euclidean distances (Eqs. 1, 2, 3) correctly estimate the dietary proportions. Ben-David et al. (1997a, 1997b) and Szepanski et al. (1999) point out that their equation (Eq. 3) assumes that all three types of food sources are utilized, and that the equation underestimates the proportions for commonly used food sources and overestimates the proportions for rarely consumed food sources. For the dual element, three source case, a simple linear mixing model can be formulated from the following mass balance equations (see e.g., Schwarcz 1991):

$$\begin{aligned} \delta J_D &= f_A \delta J_A + f_B \delta J_B + f_C \delta J_C \\ \delta K_D &= f_A \delta K_A + f_B \delta K_B + f_C \delta K_C \\ 1 &= f_A + f_B + f_C \end{aligned} \quad (4)$$

where δJ and δK represent isotopic ratios for two elements (e.g., $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$); A , B , C , and D subscripts represent three food sources and the consumer, respectively; and f represents the fractional contribution of each food source to the consumer's diet (corresponding to the %Xs in the notation of Eqs. 1, 2, 3). The isotopic ratios for food sources A , B , and C may be adjusted (A' , B' , and C') to reflect trophic fractionation factors in the animal. This linear mixing model is a system of three equations in three unknowns (f_A , f_B , and f_C) which can be solved for the unique values of the unknowns. One expression of the solution is:

$$\begin{aligned} f_A &= \frac{(\delta K_C - \delta K_B)(\delta J_D - \delta J_B) - (\delta J_C - \delta J_B)(\delta K_D - \delta K_B)}{(\delta K_C - \delta K_B)(\delta J_A - \delta J_B) - (\delta J_C - \delta J_B)(\delta K_A - \delta K_B)} \\ f_B &= \frac{(\delta J_D - \delta J_C) - (\delta J_A - \delta J_C)f_A}{\delta J_B - \delta J_C} \\ f_C &= 1 - f_A - f_B \end{aligned} \quad (4a)$$

It should be noted that this model, as well as the Euclidean distance methods, makes the implicit assumption that the partitioning of food sources is the same for both elements, e.g., C and N. If, however, the C/N ratios vary considerably among the food sources, then the fraction of the C intake from one source may not equal the fraction of the N intake from that source. A more complex mixing model which incorporates varying C and N concentrations among food sources is needed if this assumption is to be avoided (D.L. Phillips and P.L. Koch, unpublished work). However, this assumption is less likely to be a problem where the food sources are all animal tissues, as in the examples discussed here, than where they are a mix of animal and plant tissues which may have significantly different C/N ratios.

Unlike the Euclidean distance methods, this linear mixing model correctly estimates the proportions for three food sources regardless of whether all sources are utilized. In practice, random measurement errors, variation in isotopic composition among prey individuals, and variation in assimilation, fractionation, and consequently composition among predator individuals will lead to some uncertainty around these proportion estimates (Phillips and Gregg 2001). The proportions calculated should always be treated as estimates of dietary contributions within some confidence limits rather than exact point values. However, among the methods presented here, only the linear mixing model provides mathematically unbiased expected values of these proportion estimates.

Table 1 shows an example of estimates of dietary proportions for three food sources using each of the four equations. Several features should be noted. First, Eq. 2

Table 1 Example estimates of the dietary proportions of moose, caribou, and salmon for interior Alaska wolves, using mean data (*first line*) after trophic fractionation corrections (*second line*) from Szepanski et al. (1999). The three Euclidean distance mixing models (Eqs. 1, 2, 3) and the linear mixing model (Eq. 4a) were used.

	Moose	Caribou	Salmon	Wolf
$\delta^{13}\text{C}$, $\delta^{15}\text{N}$ (‰) – original	–22.7, 1.6	–20.8, 4.0	–19.9, 12.1	–19.6, 6.4
$\delta^{13}\text{C}$, $\delta^{15}\text{N}$ (‰) – corrected	–20.7, 4.6	–18.8, 7.0	–18.9, 14.1	–19.6, 6.4
Euclidean distance (‰)	2.1	1.0	7.7	
Equation 1	19%	9%	71%	–19.2, 11.6
Equation 2	70%	38%	92%	–39.0, 18.8
Equation 2×0.5	35%	19%	46%	–19.5, 9.4
Equation 3 ^a	30%	62%	8%	–19.4, 6.9
Equation 4a	42%	53%	6%	–19.6, 6.4

^a The proportions for Eq. 3 shown here differ slightly from those in Szepanski et al. (1999) Table 1 because their figures are the means of proportions calculated from isotopic ratios for individual

Equation 2 estimates resulted in proportions which summed to 2, rather than 1, so an additional entry was made to rescale these estimates down by a factor of 2. For each method, the predicted isotopic ratios for the wolf tissue, back-calculated from the estimated dietary proportions, are also shown in the right-hand column.

wolves. The proportions above were calculated from the mean isotopic ratios for wolves since the individual data were not available

Table 2 Estimates of wolf dietary proportions (%) from the four methods for a simulated range of diets, using the mean prey isotopic ratio data from Szepanski et al. (1999). As in Table 1, a row

has been added which rescales Eq. 2 estimates by a factor of 2 (*M* moose, *C* caribou, *S* salmon)

Method	Simulated diet Component Simulated %	Moose			Moose/salmon			Mixed 1			Mixed 2			Caribou/salmon		
		M	C	S	M	C	S	M	C	S	M	C	S	M	C	S
		100	0	0	75	0	25	50	25	25	25	50	25	0	75	25
Eq. 1		0	24	76	22	13	65	29	10	61	35	11	54	39	15	46
Eq. 2		–	–	–	67	44	89	76	35	89	79	35	86	77	42	81
Eq. 2×0.5		–	–	–	33	22	44	38	17	44	39	18	43	39	21	40
Eq. 3		–	–	–	33	56	11	24	65	11	21	65	14	23	58	19
Eq. 4a		100	0	0	75	0	25	50	25	25	25	50	25	0	75	25
	Simulated wolf isotopic ratios ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$)	(-20.7, 4.6)			(-20.3, 7.0)			(-19.8, 7.6)			(-19.3, 8.2)			(-18.8, 8.8)		

from Whitley and Rabeni (1997) gives proportions which sum to 2, rather than 1. An additional line in the table has been added to scale these estimates down by a factor of 2. Second, there is considerable variation in the dietary proportion estimates among the four methods. The moose proportion varies from 19% to 42%, caribou varies from 9% to 62%, and salmon varies from 6% to 71% (ignoring Eq. 2 results before halving). Finally, I back-calculated the expected isotopic ratio for the wolf tissue from the isotopic ratios and the proportions estimated for each food source. By mathematical necessity, the linear mixing model of Eq. 4a correctly returns the observed isotopic ratios because it is based on the principle of mass balance. However, none of the other models return the observed isotopic ratios, showing they do not conserve mass balance.

To test the accuracy of the four methods, I also simulated a wide range of example diets (Table 2). For each diet, the simulated proportion of each food source (after correcting for trophic fractionation) was multiplied by the corresponding isotopic ratios, which were then summed to find the isotopic ratios for the wolf. All four methods were used to estimate the dietary proportions, which were then compared with the expected values. As can be seen from Table 2, the linear mixing model (Eq. 4a) correctly returned all the proportions (by mathematical necessity), but none of the other methods were consistent with the preservation of mass balance in this way. In many cases as the dietary proportion of a food source decreased, the estimated proportion went in the other direction. Examination of Eqs. 1, 2 reveals that as the proportion of *A* in the diet decreases, the length $DX'=DA'$ increases, but this leads to increasing estimates of %*A* in the diet. While Eq. 3 does not suffer from this incorrect directionality in its formulation, neither Eq. 2 nor Eq. 3 can come up with estimates of proportions when only one food source is utilized, as shown in the first simulated diet in Table 2. For the other diet scenarios, Eqs. 1, 2 (corrected by halving), and Eq. 3 were as much as 60% (15% calculated vs. 75% actual), 54% (21% calculated vs. 75% actual), and 56% (56% calculated vs. 0% actual) off the correct dietary proportions simulated, respectively.

Several other problems in the application of the Euclidean distance methods should be noted. Ben-David et al. (1997b) and Szepanski et al. (1999) used Eq. 3 to calculate dietary fractions for individual predators whose isotopic ratios fell outside the area bounded by the trophic corrected ratios for their food sources (interior of triangle in Fig. 1). In some cases even the means of the predator populations had $\delta^{13}\text{C}$ and/or $\delta^{15}\text{N}$ ratios which did not overlap those of the food sources at all after trophic correction (e.g., spring riverine mink in Ben-David et al. 1997b). Exceeding these limits is an indication either that there may be other food sources not included, or that the trophic correction factors are not appropriately estimated. In such cases, there does not exist any positive linear combination of food sources which can account for the observed isotopic ratios in the predator. Use of any of the Euclidean distance formulas (Eq. 1, 2, or 3) will always assign a positive fraction for each food source regardless of whether the predator's isotopic ratios overlap those of the food sources. The meaningfulness of such numbers is questionable. The linear mixing model (Eq. 4a) will estimate fractional contributions in such cases, but one or more of them will be negative, indicating that the mixture falls out of bounds. Very small negative proportions might possibly be ignored as due to measurement and sampling error, but large ones indicate potential problems as mentioned above.

The second problem comes from trying to determine the dietary proportions for more than three food sources with dual isotopes (Ben-David et al. 1997b). The Euclidean distance equations (Eqs. 1, 2, 3) will calculate unique values for any number of food sources. However, food sources which do not have significantly different isotopic ratios are often combined and their combined isotopic ratios are used (Ben-David et al. 1997b). The problem is that there is not a unique solution when the number of food sources exceeds the number of elemental isotopic ratios by more than 1. The linear mixing model (Eq. 4a) can be extended to include more than three food sources, but it would result in a system of three equations in more than three unknowns. Solutions may exist and be found, but they are not unique. For example, Ben-David et al. (1997b) presented dietary proportions for

Table 3 Estimates of spring coastal mink dietary proportions using the mean prey (after trophic fractionation correction) and predator isotopic ratios from Ben-David et al. (1997b, their Fig. 2, Tables 2, 3). Estimates from Ben-David et al. (1997b) were calculated using Eq. 3 for 7 variables; linear mixing model estimates were calculated using Eq. 4 for 7 variables. Since there is no unique so-

lution using Eq. 4, all possible combinations of dietary proportions in increments of 2% were evaluated for consistency with the observed mink isotopic ratios, with tolerances of (± 0.1 , 0.1‰) for ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$) (1,069 solutions), and (± 1.7 , 0.6‰) (1,295,499 solutions). The latter tolerance represents the observed-predicted difference using the method of Ben-David et al. (1997b)

Food source	$\delta^{13}\text{C}$, $\delta^{15}\text{N}$ (‰) observed (corrected)	Ben-David et al. (1997b) estimates	Linear mixing model estimates (± 0.1 , 0.1‰)	Linear mixing model estimates (± 1.7 , 0.6‰)
Intertidal fish	-14.23, 14.68	30%	48–66%	6–82%
Blue mussels	-18.51, 9.74	12%	0–18%	0–42%
Crabs	-15.28, 11.2	21%	14–44%	0–60%
Shrimp	-16.9, 11.96	17%	0–28%	0–76%
Rodents	-24.61, 10.07	5%	0–6%	0–24%
Amphipods	-18.69, 14	9%	0–12%	0–52%
Ducks	-21.28, 14.92	6%	0–6%	0–32%
Mink	-15.1, 13.2‰ observed	-16.8, 12.6‰ predicted		
Mink (observed – predicted)		1.7, -0.6‰		

seven food sources for spring coastal mink populations using Eq. 3 extended to seven variables, and this equation resulted in a single solution (Table 3). However, when all possible dietary proportions were calculated in increments of 2%, this resulted in over a thousand combinations which were consistent with the observed mink isotopic ratios within 0.1‰ using the mass balance equations from Eq. 4. The proportions for these combinations covered a range as wide as 30% for a single food source (Table 3). Multiplying the proportions of Ben-David et al. (1997b) by the isotopic ratios for each food source resulted in predicted mink $\delta^{13}\text{C}$ and $\delta^{15}\text{N}$ ratios of -16.8‰ and 12.6‰ , respectively, compared to the observed -15.1‰ and 13.2‰ . If deviations this wide (1.7‰ and 0.6‰) are allowed, there are over 1 million possible combinations of the seven food sources, which vary as much as 76% for a single food source (Table 3). Thus, the Euclidean distance method gives the false impression of a unique solution for partitioning more than three food sources using two isotopes, and this solution results in predicted predator isotopic ratios which may be considerably different from those observed.

The geometric interpretation of distance from the vertices of the triangle (Fig. 1) as a measure of dietary proportions is visually appealing and makes intuitive sense. However, the equations that have been used to quantify this relationship in a number of papers do not appropriately scale these proportions. The simple algebraic system of three equations in three unknowns (Eq. 4) forms a linear mixing model which correctly estimates the proportions of diet due to the three sources, even when all three sources are not utilized. Similar procedures have been used for some time in archaeological studies of human diet (Schwarcz 1991). Ostrom et al. (1997) provide a recent example of the use of this mixing model in animal ecology. If a geometric Euclidean distance interpretation is desired for heuristic purposes, the linear mixing model can be cast in that light as well. First, the line segments from the vertices through the point D are extended to the opposite sides of the triangle

at points E, F, and G as shown in Fig. 1. It can be shown that the estimates for dietary proportions calculated as:

$$\begin{aligned} f_A &= ED / EA' \\ f_B &= FD / FB' \\ f_C &= GD / GC' \end{aligned} \quad (4b)$$

are mathematically equivalent to the estimates from Eq. 4a. This makes intuitive sense as can be seen by examining the lengths of GD and GC' in Fig. 1. If point D (the mixture) lay along the side $A'B'$, the mixture would be entirely composed of A and B , with no contribution from C . In Eq. 4b, GD would be 0 and so would f_C . As the contribution of C increased, GD would represent a larger proportion of the length of GC' . If D were coincident with the vertex C' , GD/GC' would be 1, indicating $f_C=1$, where the diet consisted only of C . In practice, use of the algebraic solution of Eq. 4a is simpler, but the clarity of a visual geometric model need not be abandoned in order to use the linear mixing model. However, the Euclidean distance formulas in Eqs. 1, 2, and 3 are flawed and should not continue to be used.

The situation can be generalized for uniquely partitioning p sources using q isotopic ratios, where $p \leq q+1$ (Schwarcz 1991). The solution should be obtained by algebraic solution of a system of p equations in p unknowns (such as that given in Eq. 4 for $p=3$), rather than by use of Euclidean distances of the mixture from the individual sources in p -space, which do not correctly scale the linear mixing of sources. If less than $p-1$ isotopic ratios are available, the linear mixing model may still be used to find a range of possible solutions, while mathematically the Euclidean distance methods give the false impression of a unique solution.

Acknowledgements The information in this document has been funded by the U.S. Environmental Protection Agency. It has been subjected to the Agency's peer and administrative review, and it has been approved for publication as an EPA document. Constructive reviews were provided by Renee Brooks, Jillian Gregg, Henry Lee, Craig McFarlane, David Myrold, Merav Ben-David, and Paul Koch.

References

- Ben-David M, Flynn RW, Schell DM (1997a) Annual and seasonal changes in diets of martens: evidence from stable isotope analysis. *Oecologia* 111:280–291
- Ben-David M, Hanley TA, Klein DR, Schell DM (1997b) Seasonal changes in diets of coastal and riverine mink: the role of spawning Pacific salmon. *Can J Zool* 75:803–811
- DeNiro M, Epstein S (1978) Influence of the diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta* 42:495–506
- Hobson KA (1999) Tracing origins and migration of wildlife using stable isotopes: a review. *Oecologia* 120:314–326
- Kline TC Jr, Goering JJ, Mathisen OA, Poe PH, Parker PL, Scalan RS (1993) Recycling of elements transported upstream by runs of Pacific salmon. II. $\delta^{15}\text{N}$ and $\delta^{13}\text{C}$ evidence in the Kvichak River watershed, Bristol Bay, southwestern Alaska. *Can J Fish Aquat Sci* 50:2350–2365
- Kwak TJ, Zedler JB (1997) Food web analysis of southern California coastal wetlands using multiple stable isotopes. *Oecologia* 110:262–277
- Ostrom PH, Colunga-Garcia M, Gage SH (1997) Establishing pathways of energy flow for insect predators using stable isotope ratios: field and laboratory evidence. *Oecologia* 109:108–113
- Phillips DL, Gregg JW (2001) Uncertainty in source partitioning using stable isotopes. *Oecologia* DOI 10.1007/s004420000578
- Schwarcz HP (1991) Some theoretical aspects of isotope paleodiet studies. *J Archaeol Sci* 18:261–275
- Szepanski MM, Ben-David M, Van Ballenberghe V (1999) Assessment of anadromous salmon resources in the diet of the Alexander Archipelago wolf using stable isotope analysis. *Oecologia* 120:327–335
- Whitledge GW, Rabeni CF (1997) Energy sources and ecological role of crayfishes in an Ozark stream: insights from stable isotopes and gut analysis. *Can J Fish Aquat Sci* 54:2555–2563