*Ecology Letters*, (2008) **11**: xxx–xxx doi: 10.1111/j.1461-0248.2008.01163.x

LETTER

# Incorporating uncertainty and prior information into stable isotope mixing models

**Jonathan W. Moore**[1,2]*,† and **Brice X. Semmens**[1,]†
[1]*National Marine Fisheries Service, Northwest Fisheries Science Center, 2725 Montlake Boulevard East, Seattle, WA 98112, USA*
[2]*Center for Ocean Health, University of California, 100 Shaffer Road, Santa Cruz, CA 95060, USA*
**Correspondence: E-mail: jwmoore@biology.ucsc.edu*

†*Authors contributed equally to this manuscript.*

## Abstract

Stable isotopes are a powerful tool for ecologists, often used to assess contributions of different sources to a mixture (e.g. prey to a consumer). Mixing models use stable isotope data to estimate the contribution of sources to a mixture. Uncertainty associated with mixing models is often substantial, but has not yet been fully incorporated in models. We developed a Bayesian-mixing model that estimates probability distributions of source contributions to a mixture while explicitly accounting for uncertainty associated with multiple sources, fractionation and isotope signatures. This model also allows for optional incorporation of informative prior information in analyses. We demonstrate our model using a predator–prey case study. Accounting for uncertainty in mixing model inputs can change the variability, magnitude and rank order of estimates of prey (source) contributions to the predator (mixture). Isotope mixing models need to fully account for uncertainty in order to accurately estimate source contributions.

## Keywords

Bayesian, carbon, diet, food web, isotopic fractionation, MixSIR, nitrogen, rainbow trout, salmon, sampling importance resampling.

*Ecology Letters* (2008) 11: 1–11

## INTRODUCTION

Movements of nutrients and matter through food webs and ecosystems are often difficult to quantify, being transient, variable and difficult to observe (Polis 1991). There is growing recognition that stable isotopes of naturally occurring elements, such as carbon, sulphur, nitrogen, oxygen and hydrogen, are a powerful tool to trace these flows (Peterson & Fry 1987; Schindler & Lubetkin 2004; Fry 2006; West *et al.* 2006). As a result, stable isotope studies are increasing exponentially in ecology (Schindler & Lubetkin 2004; Martínez del Rio & Wolf 2005). One of the most common uses of stable isotopes is to quantify contributions of different sources to a mixture. For example, stable isotopes have been used to investigate the potential sources of pollution (Clouquet *et al.* 2006). Alternatively, because organisms 'are what they eat, isotopically' (DeNiro & Epstein 1978), stable isotopes can illuminate many aspects of food web ecology, such as food-web length (e.g. Kling *et al.* 1992; Post 2002), niche space (e.g. Layman *et al.* 2007) and diet compositions (e.g. Benstead *et al.* 2006). In this paper, we focus principally on such food-web analyses in examples and discussion, although the analytic approach we describe is not limited to species interactions or ecological systems in general.

To quantify the relative contribution of different sources to a mixture, stable isotope studies often construct a mixing model using stable isotope signatures from potential sources and the mixture. Because stable isotopes are conserved through time, and change relatively predictably during biological processes, mass-balance mixing models can quantitatively assess the relative contribution of different sources to the mixture of interest (Peterson & Fry 1987; Phillips & Gregg 2001, 2003; Schindler & Lubetkin 2004; Martínez del Rio & Wolf 2005). Like other powerful tools, mixing models have the potential to be misused. In fact, as previous authors have noted (Phillips & Gregg 2001; Fry 2006), stable isotope mixing models are often performed with little or no consideration of the substantial and multiple sources of uncertainty. Given the ongoing proliferation of stable isotope studies, there is a clear need for robust analytical techniques that allow for the estimation of uncertainty surrounding source contributions. A basic mass-balance mixing model assumes that, for a given isotope, the isotopic signature of the mixture ($\delta_M$) is defined as follows:

$$\delta_M = f_1 * (\delta_1 + \gamma_1) + f_2 * (\delta_2 + \gamma_2) \cdots f_n * (\delta_n + \gamma_n), \qquad (1)$$

where $f_i$ is the proportional contribution of the $i$th source to the mixture, $\delta_i$ the isotopic signature of the $i$th source and $\gamma_i$ the isotope-specific fractionation of the $i$th source (often fractionation is assumed to be constant across sources). Stable isotope signatures are generally expressed as a ratio of the heavy to light isotope, relative to a standard. If the number of sources is less than or equal to the number of isotopes + 1, these equations can be solved exactly for the contributions of the different sources. For instance, with data from two isotopes, only the contribution of three or fewer sources can be partitioned exactly. This mixing model is the basis for stable isotope source partitioning.

Mixing models are challenged to incorporate a variety of sources of uncertainty. Here, we present a mixing model that more fully incorporates the following sources of uncertainty.

### Variability in isotope signatures

Stable isotope signatures typically have substantial uncertainty, primarily due to process errors (Phillips & Gregg 2001). Isotopic studies of food webs often investigate how populations of prey (sources) contribute to a population of consumers (mixture). However, within a population, differences in diet lead to differences in the isotopic signatures of individuals (Angerbjörn *et al.* 1994; Matthews & Mazumder 2004; Urton & Hobson 2005; Araújo *et al.* 2007; Layman *et al.* 2007). Within an individual, stable isotope signatures are subject to a small amount of measurement error (Jardine & Cunjak 2005), can vary across tissue types (e.g. Tieszen *et al.* 1983) and be influenced by preservation and sampling techniques (Sarakinos *et al.* 2002; Post *et al.* 2007). Furthermore, given that isotope signatures are temporally and spatially variable, the sampling regime used to collect isotope signatures is of critical importance, especially given that consumers and their prey will likely have different scales of isotopic integration (Cabana & Rasmussen 1996; Vander Zanden & Rasmussen 1999; O'Reilly *et al.* 2002; Post 2002).

### Isotopic fractionation

Ratios of isotopes systematically change as elements are ingested, excreted or catabolized (e.g. trophic fractionation; DeNiro & Epstein 1981; Minagawa & Wada 1984; Martínez del Rio & Wolf 2005). Fractionation can vary depending on the characteristics of the consumer, such as diet composition or feeding rate (Vander Zanden & Rasmussen 2001; Post 2002; McCutchan *et al.* 2003; Vanderklift & Ponsard 2003; Martínez del Rio & Wolf 2005). Despite this substantial variability, studies typically assume constant fractionation rates and ignore the associated uncertainty (but see Vander Zanden & Rasmussen 2001).

### Too many sources

As the number of potential sources included in a mixing model increases, the uncertainty in the contribution of any one source also increases (Phillips & Gregg 2003; Lubetkin & Simenstad 2004). Mixing models cannot deterministically solve mass-balance equations when the number of sources exceeds the number of isotopes + 1, a common occurrence in ecological systems. There have been several statistical treatments to account for mixing models without deterministic solutions (e.g. Phillips & Gregg 2003; Lubetkin & Simenstad 2004), and we build upon these previous models to more fully incorporate uncertainty.

Mixing model estimates can be refined by incorporating additional information. Many predator–prey isotope studies have used gut content or faecal analyses to informally refine the estimates of prey contributions to a consumer (e.g. Kling *et al.* 1992; Vander Zanden *et al.* 1997; Layman *et al.* 2007). However, to our knowledge, there have been limited attempts to establish a formal means to incorporate such prior information into mixing model analyses (but see Phillips & Gregg 2003). Thus, we see a pressing need for a stable isotope mixing model that can partition multiple sources, incorporate multiple sources of uncertainty and provide an explicit framework for using prior information to guide analyses. These challenging aspects of stable isotope analyses can be addressed through Bayesian statistical techniques. In this paper, we describe a Bayesian stable isotope mixing model that achieves these goals. As an example, we use this mixing model to estimate contributions of different prey to rainbow trout (*Oncorhynchus mykiss*) from Alaskan streams and demonstrate that accurately interpreting stable isotope data with mixing models requires addressing the uncertainty associated with these data.

## METHODS

### Statistical model

We developed and implemented a stable isotope mixing model, hereafter referred to as MixSIR, using a Bayesian framework to determine the probability distributions for the proportional contribution ($f_i$) of each source $i$ to the mixture of interest. Bayesian statistics offer a powerful means to interpret data because they can incorporate prior information, integrate across sources of uncertainty and explicitly compare the strength of support for competing models or parameter values (Hilborn & Mangel 1997; Ellison 2004).

For this application, Bayesian techniques allow for the estimation of posterior probability distributions for all $f_i$ through numerical integration. This numerical integration requires randomly generating $q$ proposed vectors of proportional source contributions $\mathbf{f}_q$ representing possible states of nature, where all $f_i$ elements in $\mathbf{f}_q$ sum to unity. Based on Bayes theorem, the probability of each $\mathbf{f}_q$ is then calculated based on data and prior information (Hilborn & Mangel 1997; Ellison 2004) such that:

$$P(\mathbf{f}_q \mid \text{data}) = \frac{L(\text{data} \mid \mathbf{f}_q) * p(\mathbf{f}_q)}{\sum L(\text{data} \mid \mathbf{f}_q) * p(\mathbf{f}_q)}, \tag{2}$$

where $L(\text{data} \mid \mathbf{f}_q)$ is the likelihood of the data given $\mathbf{f}_q$, $p(\mathbf{f}_\text{q})$ represents the prior probability of the given state of nature being true based on prior information and the denominator is a numerical approximation of the marginal probability of the data (a normalizing constant). The numerator $L(\text{data} \mid \mathbf{f}_q) * p(\mathbf{f}_q)$, hereafter referred to as the unnormalized posterior probability (Gelman *et al.* 2003), yields the absolute probability of a given $\mathbf{f}_q$ based on data and prior beliefs.

Suppose we are trying to estimate the contribution of $i$ sources to a mixture of $j$ isotopes. In MixSIR, isotope signatures from the mixture population constitute the data and are assumed to be normally distributed. For instance, if we wish to determine the contribution of prey items to a predator diet, the data would be isotope signatures from individual predators. Uncertainty in source isotope values are factored into the model by defining mean and variance parameters for each $i, j$. Prior beliefs regarding proportional source contributions are defined using beta distributions on the interval [0, 1] and the two non-negative shape parameters, $\alpha$ and $\beta$.

In order to calculate the likelihood of the data given $\mathbf{f}_q$, the proposed proportional contributions are combined with both user-specified source isotope distributions and their associated user-specified fractionation distributions in order to develop resultant proposed isotope distributions for the mixture. The likelihood of this distribution given the mixture data is then determined by calculating the product of the likelihoods of each individual mixture isotope value, given the proposed mixture distribution specific to that isotope. The proposed isotope distributions for the mixture are determined by solving for the proposed means $\hat{\mu}_j$ and standard deviations $\hat{\sigma}_j$ of the mixture based on the randomly drawn $f_i$ values comprising a vector $\mathbf{f}_q$:

$$\hat{\mu}_j = \sum_{i=1}^{n}\left[f_i * \left(m_{j_{\text{source}_i}} + m_{j_{\text{frac}_i}}\right)\right] \tag{3}$$

$$\hat{\sigma}_j = \sqrt{\sum_{i=1}^{n}\left[f_i^2 * \left(s^2_{j_{\text{source}_i}} + s^2_{j_{\text{frac}_i}}\right)\right]}, \tag{4}$$

where $m_{j_{\text{source}_i}}$ is the mean of the $j$th isotope of the $i$th source, $m_{j_{\text{frac}_i}}$ the mean fractionation of the $j$th isotope of the $i$th source, $s^2_{j_{\text{source}_i}}$ the variance of the $j$th isotope of the $i$th source and $s^2_{j_{\text{frac}_i}}$ the variance in fractionation of the the $j$th isotope of the $i$th source. Once the $\hat{\mu}_j$'s and $\hat{\sigma}_j$'s are determined, the likelihood of the data given the proposed mixture is calculated as:

$$L(x \mid \hat{\mu}_j, \hat{\sigma}_j) = \prod_{k=1}^{n}\prod_{j=1}^{n}\left[\frac{1}{\hat{\sigma}_j * \sqrt{2 * \pi}} * \exp\left(-\frac{(x_{kj} - \hat{\mu}_j)^2}{2 * \hat{\sigma}_j^2}\right)\right], \tag{5}$$

where $x_{kj}$ represents the $j$th isotope of the $k$th mixture in the data file. Next, the likelihood of $\mathbf{f}_q$ given prior information (user-specified $\alpha$ and $\beta$ for each source) is calculated according to a beta distribution:

$$L(\mathbf{f}_q \mid \alpha_i, \beta_i) = \prod_{i=1}^{n}\frac{f_i^{\alpha_i - 1} * (1 - f_i)^{\beta_i - 1}}{B(\alpha_i, \beta_i)}. \tag{6}$$

Finally, the likelihood of $\mathbf{f}_q$ given prior information is multiplied by the likelihood of the mixture data given $\mathbf{f}_q$ in order to calculate the unnormalized posterior probability of $\mathbf{f}_q$ given priors and data.

We implemented the Hilborn (after Professor Ray Hilborn) sampling-importance-resampling (SIR) algorithm (Rubin 1988) to examine the posterior probability of a vector of proportional source contributions ($\mathbf{f}_q$) through numerical integration. The Hilborn SIR method is functionally equivalent to a basic SIR algorithm with a uniform importance function such that the resample weight for a given state of nature $\mathbf{w}(\mathbf{f}_q)$ is equal to the unnormalized posterior probability (Rubin 1988; McAllister & Ianelli 1997). However, rather than saving all initial samples in a file and subsequently resampling from this file based on $\mathbf{w}(\mathbf{f}_q)$, the Hilborn SIR method establishes a threshold acceptance value ($T$) prior to sampling and uses it to simultaneously resample, as the unnormalized posterior probabilities for each $\mathbf{f}_q$ sample are calculated. We used the Hilborn method because it is programmatically intuitive, and because it does not require all initial samples to be stored (advantageous for large model runs). The method works as follows:

(1) Use 10% of user-specified model iterations to establish a threshold ($T$):
  (a) set $T$ to 0 before beginning iterations;
  (b) for each threshold iteration, randomly draw values for each $f_i$ in $\mathbf{f}_q$ (e.g. for a three-source model, a contribution parameter draw might be 0.1, 0.1 and 0.8);
    (i) calculate the unnormalized posterior probability ($L$) of the parameter draw based on prior information and data;
    (ii) if $L$ is greater than the current $T$, then $T = L$.

(2) Use all user-specified model iterations to develop samples and simultaneously resample based on $T$ and a cumulative likelihood value ($C$):
  (a) set $C$ to 0 before beginning iterations;
  (b) for each iteration, randomly draw values for each $f_i$ in $\mathbf{f}_q$;
    (i) calculate the $L$ of the parameter draw based on prior information and data;
    (ii) add $L$ to cumulative likelihood, $C = C + L$;
    (iii) If $C$ exceeds $T$ then save the $\mathbf{f}_q$ for that iteration in list of resamples and adjust the cumulative likelihood value, $C = C - T$.

The SIR algorithm is well suited to models having relatively few parameters with well-defined intervals. Because all of the parameters in the model are proportional contributions of each source, models will generally have few parameters. Additionally, as these parameters are proportions, they are bounded in the interval 0–1. Finally, because bounded proportions must all sum to 1, parameter values have cross-dependencies that can result in multi-modal posterior distributions and thus limit the applicability of basic Markov-Chain Monte-Carlo sampling techniques. Given these constraints, a SIR algorithm is an effective method for resampling proportional parameter space in order to develop accurate posterior distributions.

### *Incorporating prior information*

The Bayesian framework allows a user to establish informative priors to guide model estimates. As stated above, the probabilities of source contributions are evaluated against prior information according to the beta distribution. The beta-distributed priors on each $\mathbf{f}_q$ are highly flexible. Informative priors can be established on some or all $\mathbf{f}_q$ so that a user can incorporate varying degrees of belief regarding the contribution of each source to the mixture. Note, however, that because all elements of $\mathbf{f}_q$ must sum to unity, priors on the elements of $\mathbf{f}_q$ are not independent (Connor & Mosiman 1969). Highly informative priors will often 'sharpen' the peaks in the likelihood surface, and the model will consequently require more iterations to develop an appropriate posterior. However, the model may fail to converge if informative priors specify implausible source contributions based on the model formulation and input data. Generally, the more data a user provides the model, the less influential prior information will be on the model. When both $\alpha$ and $\beta$ are set to 1, all source contributions are *a priori* equally likely (uninformative priors).

### *Application*

We developed MixSIR and an associated graphical user interface (GUI) using MATLAB. The MixSIR GUI allows a user to input isotope data and specify graphical and textual outputs of probability distributions for the source contributions. MixSIR takes space delimited text files as inputs that specify the averages and standard deviations of source isotope signatures, the averages and standard deviations of fractionation values, and mixture isotope signature data. Estimates of the fractionation of common isotopes can be found in several recent reviews (e.g. Vander Zanden & Rasmussen 2001; Post 2002; McCutchan *et al.* 2003; Vanderklift & Ponsard 2003). Users can optionally input informative priors by specifying $\alpha$ and $\beta$ parameters that define prior beliefs about the form of source contribution distributions. MixSIR outputs graphically and numerically descriptive posterior probability distribution estimates of the source contributions. MixSIR is available over the Green-Boxes code sharing network (http://conserver.iugo-cafe.org), along with example data and a user guide with more information regarding the model form and function.

### *Alaskan stream food web example*

To demonstrate MixSIR, we used data from stream food webs in Alaska. The model for this example uses isotope data from predatory rainbow trout (*O. mykiss*) and five prey sources (benthic invertebrates, prey fishes, shrews, salmon eggs and terrestrial invertebrates) to estimate the dietary patterns of rainbow trout in streams in the Wood River system of south-western Alaska. Between 2000 and 2006, 90 rainbow trout above 250 mm in total length were non-lethally sampled during the open water season (June–August) using hook and line and small seines. Following anesthetization, gut contents were collected via gastric lavage, identified and dry mass of the different diet items recorded. Prior to fish release, small plugs of dorsal muscle were taken for subsequent isotopic analysis. Based on prior diet analyses, we gathered the representative prey items from the stream or diets for isotopic analyses. Because these prey items likely have more rapid isotopic turnover rates than their consumers (e.g. Cabana & Rasmussen 1996; Vander Zanden & Rasmussen 1999; O'Reilly *et al.* 2002; Post 2002), prey items were collected across multiple sampling dates throughout the open water season. Sampling for all prey types spanned at least three summer months with at least three sampling events, with the exception of shrews and salmon eggs, which were sampled on 2 and 3 dates, respectively, across 3 weeks. Isotopic samples were preserved in ethanol. Subsets of prey and consumer samples were subsequently dried, homogenized and sent to the UC-Davis Stable Isotopes Facility to determine $\delta^{13}$C and $\delta^{15}$N signatures. This sampling was part of a larger effort by the Alaska Salmon Program of the University of Washington to understand the ecological importance of Pacific salmon (*Oncorhynchus* spp.) on lotic food webs (Schindler *et al.* 2003; Scheuerell *et al.* 2007; Moore *et al.* 2008).

We used isotopic data from eight rainbow trouts between 300–400 mm in total length for our mixing model analysis. These trout were relatively enriched in $^{15}$N, as were salmon eggs. Benthic invertebrates were relatively enriched in $^{13}$C, while terrestrial invertebrates were relatively depleted in $^{13}$C (Fig. 1a). Typical of isotope data, consumer and prey isotopes signatures were variable (Fig. 1a). We used the individual isotope signatures of the rainbow trout and the means and SD of prey as inputs to MixSIR. We used previously published fractionation values of 2.3 ± 1.61 for $^{15}$N and 0.4 ± 1.20 for $^{13}$C for aquatic organisms (mean ± 1 SD; McCutchan *et al.* 2003). In addition, using these same data, we ran similar analyses using IsoSource, a popular mixing model that qualitatively accounts for uncertainty through a 'tolerance' parameter (Phillips & Gregg 2003). In order to investigate the influence of model uncertainty, we re-ran MixSIR (Fig. 1) after decreasing (1) the error in source isotope values, (2) error in fractionation, (3) error in mixture isotope values and (4) error in all sources. Finally, we re-ran MixSIR with the same data, but with informative priors on the prey contribution parameters based on gut-content data.

We developed prior beliefs in the form of beta distributions representing the proportional diet contributions of each source using a bootstrap procedure applied to gut-content data. Because a bootstrap routine works best with large sample sizes, we used all diet samples ($n$ = 90). Data were organized into a matrix where rows defined the proportional gut contents of a fish and columns comprised the source contributions to total gut-content dry weight as proportions. For each of $1 \times 10^5$ bootstrapped samples, we resampled eight fish (the number of individuals used in our mixing model) from the data matrix with replacement and subsequently averaged the source proportions in the gut contents of these eight resamples. Finally, we fit beta distributions to these $1 \times 10^5$ averaged bootstrapped gut-content proportions by identifying the maximum likelihood estimates of the $\alpha$ and $\beta$ parameters for each source. We have made the code for this bootstrap procedure available over the GreenBoxes code sharing network (http://con-server.iugo-cafe.org).

### *Model validation*

Because the fundamental construct of our model is based on a mass-balance mixing model (eqn 1), the performance of our model is subject to many of the same constraints described previously for such models (e.g. Phillips & Gregg 2003). Thus, demonstrating that the results of MixSIR converge to the results of existing mixing models as uncertainty is reduced provides a simple means of model validation. Furthermore, to evaluate the performance of MixSIR when model input uncertainty is realistically specified, we developed a MATLAB script that generated artificial data by randomly choosing a number of sources (between 3 and 5), their associated $\delta^{13}$C and $\delta^{15}$N signatures and their proportional contribution to a theoretical mixture. Mixture data points were then generated by drawing values from each of the source isotope distributions, multiplying them by their associated proportional contributions and, finally, summing these values across sources for each isotope to generate a mixture isotope signature.

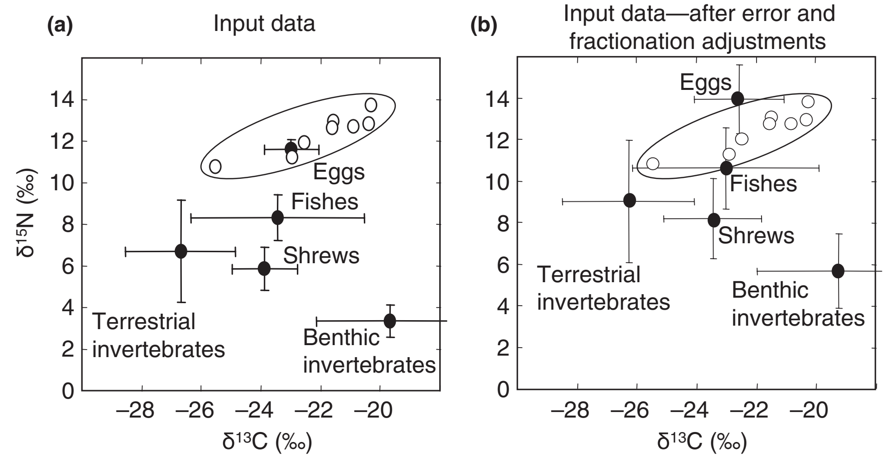

**Figure 1** Model input and diagnostic histogram for stable isotope mixing model: (a) carbon and nitrogen stable isotope data from Alaskan stream food webs used as an input for MixSIR. Open circles represent the stable isotope signatures of individual rainbow trout. Solid points and error bars represent, respectively, the mean and standard deviation of isotope signatures of prey sources for rainbow trout. (b) Input data following adjustment due to fractionation and error incorporation. Data were adjusted based on error propagation and incorporation of fractionation, following eqns 4 and 5. MixSIR performs these calculations as part of the algorithm, and the adjusted data are shown here for demonstrative purposes.

## RESULTS

### Alaskans stream food-web example

Using uninformative priors and estimates of uncertainty associated with mixing model inputs, a MixSIR model run of $5 \times 10^6$ iterations resulted in convergence on the posterior source contributions of the different prey items to the diet of rainbow trout (Figs 1 and 2). The model resampled a total of 24 147 posterior draws with no duplicates in less than 1 min on a laptop with 2 GB of RAM and an Intel® Core™ Duo processor T2500 (Lenovo, Morrisville, NC, USA). The maximum importance ratio (calculated by determining the ratio of the maximum unnormalized posterior probability resample to the sum of all unnormalized posterior probability resamples) was below 0.001, suggesting that our model was effective in estimating the true posterior density (McAllister & Pikitch 1997). Additionally, the distribution of these resampled unnormalized posterior probabilities (Fig. 2a and b) demonstrated that our model placed appropriate weight on the tails of the posterior distribution (McAllister & Ianelli 1997). Based on model results, rainbow trout derive most of their tissue from salmon eggs and prey fishes such as sculpin and juvenile salmon (Fig. 2c). Specifically, the model estimated that eggs contributed a median of 63% (48–76%; this and the following represent the 5 and 95% confidence percentiles) and prey fishes contributed a median of 18% (2–37%) to rainbow trout. The estimated median contribution of the three other prey groups all were 10% or less, with 95% confidence limits all less than 20% and 5% confidence limits less than 1%.

Running the model with informative priors modified the median posterior estimates of source contributions and reduced the variance in the posterior source contribution estimates, especially for prey fishes (Fig. 2d). However, the model results were not dramatically altered from those resulting from uninformative priors, suggesting that our data were informative and that our specified priors generally agreed with the data.

As expected, running the uninformative prior model after decreasing all sources of uncertainty constrained the estimates of the relative contributions of the different prey sources to the diet of rainbow trout (Table 1; Fig. 3). However, we were surprised to find that changes in uncertainty influenced the magnitude and rank order of estimated prey contributions to the diet of trout, sometimes considerably (Table 1). For example, reducing the variation in source isotope signatures by 50% changed the estimate of the contribution of eggs both in terms of the median and range of contributions from 0.63 (0.48 to 0.76) to 0.71 (0.31 to 0.83), and of fishes from 0.18 (0.02 to 0.37) to 0.09 (0.01 to 0.62) (Table 1). Reducing the individual sources of uncertainty changed the model output to varying degrees (Table 1; Fig. 3). In the extreme case of reducing all sources of uncertainty, model results changed not only the median and range of source contributions, but also the rank order of median contributions. Specifically, the model that fully incorporated uncertainty identified fish as the second most important prey source, while the model where all uncertainty was reduced identified fish as the third most important prey source with benthic invertebrates becoming the second most important prey source. IsoSource results were virtually identical to the results of MixSIR model run with all sources of uncertainty reduced. On the other hand, output from IsoSource was extremely different from that of the MixSIR model run under fully specified uncertainty.

### Model validation

When we reduced variability in model inputs in our Alaska stream food-web example, our model results converged to those of IsoSource (Table 1). Reducing the variability in only one of the model inputs (source signatures, fractionation, mixture data) did not necessarily bring the results of MixSIR and IsoSource closer together, highlighting the complex influence of multiple sources of variation on model results. Our analysis of artificial data sets demonstrated that with fewer sources, the model was more likely to correctly identify contributions. Source geometry also influenced model performance, such that the degree of separation between source isotope signatures strongly influenced the ability of the model to correctly identify source contributions (Phillips & Gregg 2003). Moreover, in certain instances when sources had similar isotope signatures, the posterior distributions of source contributions exhibited strong multimodality, reflecting support for alternative states of nature (see Fig. S1 in Supplementary material). In these instances, establishing informative prior distributions had the potential to dramatically influence posterior probabilities (Fig. S1), highlighting the importance of prior specification when source geometry and the variability in inputs are insufficient to identify unimodal posterior distributions.

## DISCUSSION

Mixing model studies typically use messy isotope data and poorly described fractionation values. Here, we described and demonstrated a Bayesian stable isotope mixing model that fully integrates across this uncertainty to characterize the posterior probability distributions of source contributions to a mixture. Furthermore, our Bayesian approach can incorporate prior information, so that additional sources of information can be used to increase the power to discern source contributions. While we focus our analyses on an example using two isotopes and multiple sources, this model

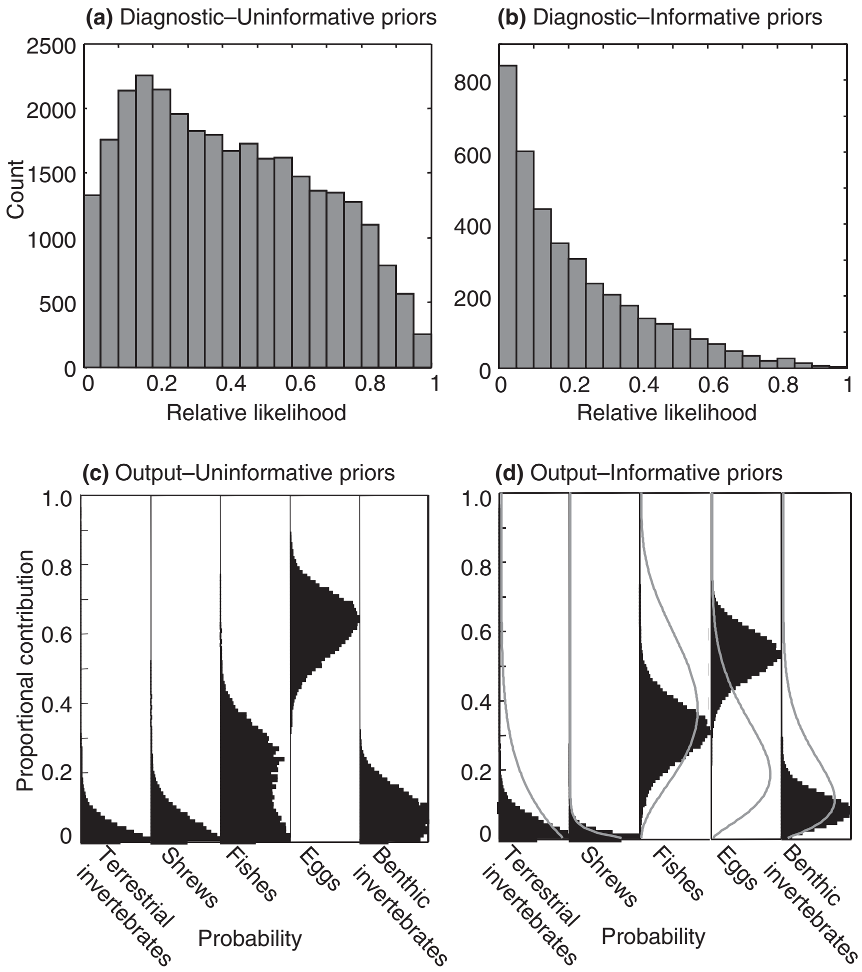

**Figure 2** Estimation of source contributions using data from an Alaskan stream food web. (a) Diagnostic histogram for model run based on this example using uninformative priors. This histogram represents the resampled unnormalized posterior probability values relative to the single largest value in set of posterior draws. If the Hilborn SIR function places too little weight in the tails of the posterior distribution, it may be inefficient in approximating posterior distributions. Thus, the graph should not be heavily right skewed (the majority of the posterior draws are at or very near the 'best' draw based on the posterior likelihoods; McAllister & Ianelli 1997). (b) Diagnostic histogram for model run based on informative priors. (c) Posterior estimates of proportional contributions of prey sources to rainbow trout, based on MixSIR, using uninformative priors. These histograms represent the distributions of posterior probabilities of source contributions. (d) Distributions of posterior estimates of proportional contributions using informative priors. The superimposed grey lines represent the specified prior distributions for each source contribution based on previous diet studies. The prior distributions were defined by $\alpha$ parameter values of 2.08, 2.48, 3.11, 1 and 1, and the $\beta$ parameter values of 9.27, 7.41, 4.56, 43.87 and 6.57 for benthic invertebrates, salmon eggs, fishes, shrews and terrestrial invertebrates, respectively. These parameters were estimated using a bootstrap algorithm for diet data that are described more in the text.

can account for an unlimited number of isotopes and sources.

The MixSIR program incorporates isotope and fractionation uncertainty in the development of posterior probability distributions of source contributions. By incorporating these sources of uncertainty, MixSIR produces source contribution estimates with explicit probability distributions. It is important to carefully examine these probability distributions and report the output conscientiously; for example, if the distributions are flat (e.g. Fig. 3d), then they do not support a certain solution but rather a range of equally feasible solutions and should be reported as such. Not surprisingly, when we reduced uncertainty in all model inputs, the resulting probability distributions became more restricted (Table 1; Figs 2 and 3). However, reducing different sources of error had different impacts, due to the interplay of data distributions and contribution parameter cross-dependencies. Although previous mixing model implementations, such as IsoSource and SOURCE/STEP, can deal with the uncertainty associated with more sources

**Table 1** Model estimates of the contribution of different prey items to the diet of rainbow trout using different amounts of uncertainty associated with mixing model inputs (see Fig. 3)

| | Prey source | | | | |
|---|---|---|---|---|---|
| Model type | Benthic | Eggs | Fish | Shrew | Terrestrial |
| MixSIR | | | | | |
| (no reduction) | 0.08 (0.01–0.20) | 0.63 (0.48–0.76) | 0.18 (0.02–0.37) | 0.05 (0.00–0.16) | 0.03 (0.00–0.11) |
| (reduced mixture*) | 0.09 (0.01–0.26) | 0.44 (0.05–0.69) | 0.15 (0.01–0.51) | 0.10 (0.01–0.43) | 0.11 (0.01–0.47) |
| (reduced source†) | 0.08 (0.01–0.21) | 0.71 (0.31–0.83) | 0.09 (0.01–0.62) | 0.03 (0.00–0.11) | 0.02 (0.00–0.07) |
| (reduced fractionation‡) | 0.02 (0.00–0.14) | 0.56 (0.45–0.79) | 0.39 (0.02–0.50) | 0.01 (0.00–0.05) | 0.01 (0.00–0.06) |
| (reduced all§) | 0.18 (0.17–0.19) | 0.80 (0.79–0.81) | 0.01 (0.00–0.02) | 0.00 (0.00–0.01) | 0.00 (0.00–0.01) |
| IsoSource¶ | 0.18 (0.17–0.19) | 0.79 (0.76–0.82) | 0.01(0.00–0.06) | 0.00(0.00–0.03) | 0.00 (0.00–0.02) |

Shown are the results from our model with different scenarios where uncertainty is alternately maintained in full, reduced in the mixture isotope data, reduced in source isotope estimates, reduced in the estimates of isotope fractionation and reduced in all aspects. Estimates from IsoSource using a 0.1‰ tolerance are included. Shown are the medians and percentiles (unless otherwise noted they are 5th and 95th) of the posterior source contributions.

*We eliminated uncertainty in the isotopes values of the mixture (rainbow trout) by averaging the individual isotope signatures to create a single mixture value for each isotope.

†We reduced the standard deviation values of source isotopes by half.

‡We reduced the standard deviation values of fractionation by half.

§We reduced the standard deviation values by a factor of 10 and used a single average mixture value for each isotope.

¶IsoSource outputs 1 and 99% confidence limits.

than an analytical solution allows (Phillips & Gregg 2003; Lubetkin & Simenstad 2004), they do not formally incorporate the variation in isotope signatures or fractionation. When used to analyse the same data from our case study, IsoSource obtained different results than our model, both in terms of the median and range of those contributions, and their rank order. When all error sources were reduced in MixSIR, the source contribution estimates converged with mixing models that do not incorporate multiple sources of uncertainty. Thus, models that do not explicitly account for variability in source isotope values, fractionation and mixture data may fail to accurately identify the magnitude of source contributions to a mixture and the uncertainty surrounding contribution estimates.

If prior information that appropriately characterizes source contributions to a mixture is available, this information can be used to improve mixing model estimates. In our example, prior information was based on gut-content data. When we incorporated this prior information, the medians of the posterior source contributions were altered and the variances were reduced (Fig. 2). However, these changes were relatively modest both because our isotope data were informative and because these data supported posterior distributions similar to the prior distributions developed from gut-content data. The isotopic signatures of rainbow trout tissue in our study provide information on consumption integrated across an extended period of time, likely months (e.g. MacAvoy *et al.* 2001). Similarly, the gut-content data we used to develop prior information were collected over the open water season (3 months). Nonetheless, the isotope signatures of trout and the gut-content data we collected may integrate across different temporal windows. We urge careful and skeptical consideration of the relevance of prior information (e.g. gut-content data, results from previous isotope studies) before developing and implementing priors. Specifically, prior development should be informed by the rich literature regarding the methods, advantages and pitfalls of prior selection and specification (e.g. Jeffreys 1961; Lindley 1983; Robert 1994; Gelman *et al.* 2003). Informative priors should be used with extreme restraint and developed *a priori*, as they are a tempting way to manipulate output (Robert 1994).

The SIR algorithm we developed is at heart a 'brute force' method of Bayesian analysis, as it draws proposals uniformly over proportional parameter space. The more iterations, the more likely that the model output will reflect the true posteriors of the source contributions. The specific number of iterations required to generate sufficient posterior draws depends on the data, the variances in source isotope signatures and fractionations, and the extent to which the isotope mixture precludes the contribution of specified sources. For instance, the inclusion of implausible sources based on the isotope mixture and fractionation will lower the resampling rate because the model will coincidentally sample implausible (*c.* 0 likelihood) parameter space. A large number of iterations are also important in order to establish an appropriate threshold ($T$), as the more iterations the model uses to develop a $T$ value, the closer this value will be to the true maximum likelihood of the posterior. If too few iterations are used, the threshold establishment phase of the

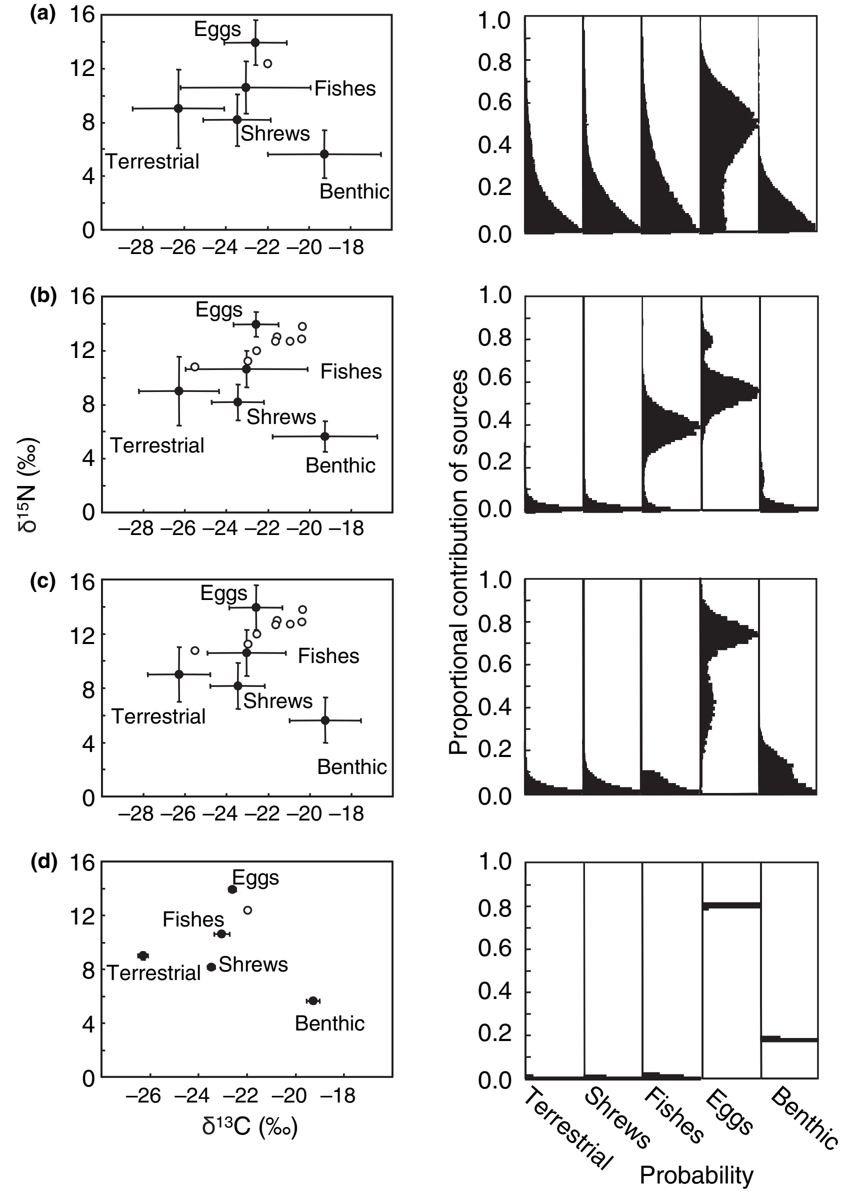

**Figure 3** Plots of the data and model solutions for each of the reduced uncertainty scenarios presented in Table 1. The scatter plots on the left present the input data following adjustment due to fractionation and error incorporation. Data were adjusted based on error propagation and incorporation of fractionation, following eqns 4 and 5. The histograms on the right show the posterior model estimates of prey contributions for each scenario. 'Benthic' and 'terrestrial' refer to invertebrate prey from those habitats. (a) Plots of the data and model results after replacing the mixture data with a single value (mixture data means) for each isotope. (b) Plots of the data and model results following a 50% reduction in fractionation. (c) Plots of the data and model results following a 50% reduction of source uncertainty. (d) Plots of the data and model results following a 10× reduction of source and fractionation uncertainty, and using the single mean values of mixture isotopes.

model run may yield an inappropriately small $T$, and this in turn may cause the SIR algorithm to resample a single $\mathbf{f}_q$ with high likelihood tens or even thousands of times.

The stable isotope mixing model presented here builds upon previously published mixing model methods, but is still limited by some of the basic assumptions of the isotope mixing model approach. For instance, MixSIR assumes that the mixture is constructed exclusively from those sources included as model inputs. In addition, mixing models necessitate sampling prey isotopes over an appropriate time frame, which can be challenging given potential temporal mismatching in isotopic turnover rates between longer-lived consumers and their prey (Cabana & Rasmussen 1996; Vander Zanden & Rasmussen 1999; O'Reilly *et al.* 2002; Post 2002). Other issues such as concentration dependence, tissue compartmentalization, a lack of distinctiveness in source isotope signatures and the correct choice of tissues for isotope analysis are also generic to isotopic mixing models, including our own (Gannes *et al.* 1997; Phillips & Koch 2002; Phillips & Gregg 2003; Martínez del Rio & Wolf 2005; Phillips *et al.* 2005; Fry 2006). Finally, the Bayesian framework of the model presented here establishes distributional assumptions that are unique among published mixing models, namely, the beta-distributed priors on source contributions.

Stable isotope mixing models are a potentially powerful tool for ecologists and can be used to quantify relationships that would otherwise be difficult or impossible to describe.

However, like much ecological data, stable isotopes are often highly variable. This variability contains important information that should be incorporated into analyses. The Bayesian mixing model we outlined here estimates the probability distributions of source contributions by incorporating variable data and prior information. Although the posterior distributions this approach produces might be less appealing than the products of mixing models based on analytical solutions, they present a realistic assessment of the ability to discern source contributions to a mixture based on data and prior knowledge.

## ACKNOWLEDGEMENTS

This paper benefited from discussions with Eric Buhle, Tim Essington, Ray Hilborn, Andre Punt, Anne Salomon, Mark Scheuerell, Daniel Schindler, Heather Tallis and Eric Ward. Comments from Mark Scheuerell, Eric Ward and three anonymous referees on an earlier draft of this manuscript were extremely helpful. Parts of this work were performed while the authors held a National Research Council Research Associateship Award at the Northwest Fisheries Sciences Center, National Marine Fisheries Service, NOAA.

## REFERENCES

Angerbjörn, A., Hersteinsson, P., Lidén, K. & Nelson, E. (1994). Dietary variation in Arctic foxes (*Alopex lagopus*) – an analysis of stable carbon isotopes. *Oecologia*, 99, 226–232.

Araújo, M.S., Bolnick, D.I., Machado, G., Giaretta, A.A. & dos Reis, S.F. (2007). Using $\delta^{13}$C stable isotopes to quantify individual-level diet variation. *Oecologia*, 152, 643–654.

Benstead, J.P., March, J.G., Fry, B., Ewel, K.C. & Pringle, C.M. (2006). Testing IsoSource: stable isotope analysis of a tropical fishery with diverse organic matter sources. *Ecology*, 87, 326–333.

Cabana, G. & Rasmussen, J.B. (1996). Comparison of aquatic food chains using nitrogen isotopes. *Proc. Natl. Acad. Sci. U.S.A.*, 93, 10844–10847.

Clouquet, C., Carignan, J., Libourel, G., Sterckeman, T. & Perdrix, E. (2006). Tracing source pollution in soils using cadmium and lead isotopes. *Environ. Sci. Technol.*, 40, 2525–2530.

Connor, R.J. & Mosiman, J.E. (1969). Concepts of independence for proportions with a generalization of the Dirichlet distibution. *J. Am. Stat. Assoc.*, 64, 194–206.

DeNiro, M.J. & Epstein, S. (1978). Influence of diet on the distribution of carbon isotopes in animals. *Geochim. Cosmochim. Acta*, 42, 495–506.

DeNiro, M.J. & Epstein, S. (1981). Influence of diet on the distribution of nitrogen isotopes in animals. *Geochim. Cosmochim. Acta*, 48, 341–351.

Ellison, A.M. (2004). Bayesian inference in ecology. *Ecol. Lett.*, 7, 509–520.

Fry, B. (2006). *Stable Isotope Ecology*. Springer, New York, NY.

Gannes, L.Z., O'Brien, D.M. & Martínez del Rio, C. (1997). Stable isotopes in animal ecology: assumptions, caveats, and a call for more laboratory experiments. *Ecology*, 78, 1271–1276.

Gelman, A., Carlin, J.B, Sterm, H.S. & Rubin, D.B. (2003). *Bayesian Data Analysis*. CRC Press, Boca Raton, FL.

Hilborn, R. & Mangel, M. (1997). *The Ecological Detective*. Princeton University Press, Princeton, NJ.

Jardine, T.D. & Cunjak, R.A. (2005). Analytical error in stable isotope ecology. *Oecologia*, 144, 528–533.

Jeffreys, H. (1961). *Theory of Probability*, 3rd edn. Clarendon Press, Oxford.

Kling, G.W., Fry, B. & O'Brien, W.J. (1992). Stable isotopes and planktonic trophic structure in Arctic lakes. *Ecology*, 73, 561–566.

Layman, C.A., Quattrochi, J.P., Peyer, C.M. & Allgeier, J.E. (2007). Niche width collapse in a resilient top predator following ecosystem fragmentation. *Ecol. Lett.*, 10, 937–944.

Lindley, D.V. (1983). Theory and practice of Bayesian statistics. *Statistician*, 32, 1–11.

Lubetkin, S.C. & Simenstad, C.A. (2004). Multi-source mixing models to quantify food web sources and pathways. *J. Appl. Ecol.*, 41, 996–1008.

MacAvoy, S.E., Macko, S.A. & Garman, G.C. (2001). Isotopic turnover in aquatic predators: quantifying the exploitation of migratory prey. *Can. J. Fish. Aquat. Sci.*, 58, 923–932.

Martínez del Rio, C. & Wolf, B.O. (2005). Mass-balance models for animal isotopic ecology. In: *Physiological and Ecological Adaptations to Feeding in Vertebrates* (eds Starck, J.M. & Wang, T.). Science Publishers, Enfield, NH, pp. 141–174.

Matthews, B. & Mazumder, A. (2004). A critical evaluation of intrapopulation variation of $\delta^{13}$C and isotopic evidence of individual specialization. *Oecologia*, 140, 361–371.

McAllister, M.K. & Ianelli, J.N. (1997). Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. *Can. J. Fish. Aquat. Sci.*, 54, 284–300.

McAllister, M.K. & Pikitch, E.K. (1997). A Bayesian approach to choosing a design for surveying fishery resources: application to the eastern Bering Sea trawl survey. *Can. J. Fish. Aquat. Sci.*, 54, 301–311.

McCutchan, J.H. Jr, Lewis, W.M. Jr, Kendall, C. & McGrath, C.C. (2003). Variation in trophic shift for stable isotope ratios of carbon, nitrogen, and sulfur. *Oikos*, 102, 378–390.

Minagawa, M. & Wada, E. (1984). Stepwise enrichment of $^{15}$N along food chains: further evidence and the relation between $\delta^{15}$N and animal age. *Geochim. Cosmochim. Acta*, 48, 1135–1140.

Moore, J.W., Schindler, D.E. & Ruff, C.P. (2008). Habitat saturation drives thresholds in stream subsidies. *Ecology*, 89, 306–312.

O'Reilly, C.M., Hecky, R.E., Cohen, A.S. & Plisnier, P.D. (2002). Interpreting stable isotopes in food webs: recognizing the role of time averaging at different trophic levels. *Limnol. Oceanogr.*, 47, 306–309.

Peterson, B.J. & Fry, B. (1987). Stable isotopes in ecosystem studies. *Annu. Rev. Ecol. Syst.*, 18, 290–320.

Phillips, D.L. & Gregg, J.W. (2001). Uncertainty in source partitioning using stable isotopes. *Oecologia*, 127, 171–179 (see also erratum, *Oecologia*, 128, 204).

Phillips, D.L. & Gregg, J.W. (2003). Source partitioning using stable isotopes: coping with too many sources. *Oecologia*, 136, 261–269.

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

Phillips, D.L., Newsome, S.D. & Gregg, J.W. (2005). Combining sources in stable isotope mixing models: alternative methods. *Oecologia*, 144, 520–527.

Polis, G.A. (1991). Complex trophic interactions in deserts: an empirical critique of food-web theory. *Am. Nat.*, 138, 123–155.

Post, D.M. (2002). Using stable isotopes to estimate trophic position: models, methods, and assumption. *Ecology*, 83, 703–718.

Post, D.M., Layman, C.A., Arrington, D.A., Takimoto, G., Quattrochi, J. & Montaña, C.G. (2007). Getting to the fat of the matter: models, methods and assumptions for dealing with lipids in stable isotope analyses. *Oecologia*, 152, 179–189.

Robert, C.P. (1994). *The Bayesian Choice: a Decision-Theoretic Motivation*. Springer-Verlag, New York, NY.

Rubin, D.B. (1988). Using the SIR algorithm to simulate posterior distributions. In: *Bayesian Statistics 3: Proceedings of the Third Valencia International Meeting, June 1–5, 1987* (eds Bernardo, J.M., Degroot, M.H., Lindley, D.V. & Smith, A.M.). Clarendon Press, Oxford, pp. 385–402.

Sarakinos, H.C., Johnson, M.L. & Vander Zanden, M.J. (2002). A synthesis of tissue-preservation effects on carbon and nitrogen stable isotope signatures. *Can. J. Zool.*, 80, 381–387.

Scheuerell, M.D., Moore, J.W., Schindler, D.E. & Harvey, C.J. (2007). Varying effects of anadromous sockeye salmon on the trophic ecology of two species of resident salmonids in southwest Alaska. *Freshw. Biol.*, 52, 1944–1956.

Schindler, D.E. & Lubetkin, S.C. (2004). Using stable isotopes to quantify material transport in food webs. In: *Food Webs at the Landscape Level* (eds Polis, G.A., Power, M.E. & Huxel, G.R.). The University of Chicago Press, Chicago, IL, pp. 25–42.

Schindler, D.E., Scheuerell, M.D., Moore, J.W., Gende, S.M., Francis, T.B. & Palen, W.J. (2003). Pacific salmon and the ecology of coastal ecosystems. *Front. Ecol. Environ.*, 1, 31–37.

Tieszen, L.L., Boutton, T.W., Tesdahl, K.G. & Slade, N.A. (1983). Fractionation and turnover of stable carbon isotopes in animal tissues: implications for $\delta^{13}$C analysis of diet. *Oecologia*, 57, 32–37.

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

Vander Zanden, M.J. & Rasmussen, J.B. (1999). Primary consumer $\delta^{13}$C and $\delta^{15}$N and the trophic position of aquatic consumers. *Ecology*, 80, 1395–1404.

Vander Zanden, M.J. & Rasmussen, J.B. (2001). Variation in $\delta^{15}$N and $\delta^{13}$C trophic fractionation: implications for aquatic food web studies. *Limnol. Oceanogr.*, 46, 2061–2066.

Vander Zanden, M.J., Cabana, G. & Rasmussen, J.B. (1997). Comparing trophic position of freshwater fish calculated using stable nitrogen isotope ratios ($\delta^{15}$N) and literature dietary data. *Can. J. Fish. Aquat. Sci.*, 54, 1142–1158.

Vanderklift, M.A. & Ponsard, S. (2003). Sources of variation in consumer-diet $\delta^{15}$N enrichment: a meta-analysis. *Oecologia*, 136, 169–182.

West, J.B., Bowen, G.J., Cerling, T.E. & Ehleringer, J.R. (2006). Stable isotopes as one of nature's ecological recorders. *Trends Ecol. Evol.*, 21, 408–414.

## SUPPLEMENTARY MATERIAL

The following supplementary material is available for this article:

**Figure S1** Example output from model validation of artificial data sets.

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/full/10.1111/j.1461-0248.2008.01163.x.

Editor, Tim Wootton
Manuscript received 11 October 2007
First decision made 6 November 2007
Second decision made 9 January 2008
Manuscript accepted 16 January 2008