2-9-2010

Stable isotope sourcing using sampling

Erik Erhardt

Erik Barry Erhardt
Candidate

Mathematics and Statistics
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

EJ Bedrick, Chairperson

Ronald Christensen

Gabriel Huerta

Michael Quintana

Wesley O. Johnson

Stable Isotope Sourcing using Sampling

by

Erik Barry Erhardt

B.A., Franklin Pierce College, 1997
M.S., Worcester Polytechnic Institute, 2003

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Statistics

The University of New Mexico

Albuquerque, New Mexico

December, 2009

Dedication

To Shomari

Acknowledgments

Ed Bedrick, UNM Statistics and advisor, for mentoring me in academic and professional practice, for nominating me for the PIBBS fellowship and encouraging me to consider the problem that this dissertation regards, for introductions to people in the statistical community, and for repeatedly getting me unstuck.

Blair Wolf, UNM Biology, for introducing me through the PIBBS program to stable isotope sourcing problems, for contextual expertise in stable isotopes, for connecting me with Jim Ehleringer's SIRFIR lab and securing my participation in the summer course, and for encouragement along the way.

James Brown and Felisa Smith, both UNM Biology, for funding me as a fellow in the Program in Interdisciplinary Biological and Biomedical Sciences (PIBBS).
Shannon McCoy-Hayes, UNM Biology PIBBS, for making the opportunities of being a PIBBS fellowship easy and enjoyable to take advantage of.

Howard Hughes Medical Institute (HHMI) for funding the PIBBS program at UNM, and for hosting the gathering of interfaces scholars in September 2008.

Keith Hobson for lots of data, guidance, and promotion.

Merav Ben-David for mink observational data.

Donald Phillips, EPA, for his initial work on mixing models to solve source partitioning problems and for comments for the work in Chapter 3.

Jim Ehleringer, University of Utah Biology, and his SIRFIR lab for my participation in his summer 2008 Stable Isotopes in Ecology Course multi-instructor lecture (Biology 7473) and laboratory (Biology 7475) short course, for helpful comments for the work in Chapter 3, and for fantastic BBQ and hiking at his cabin.

Spider Vetter, University of California, Davis, a fellow SIRFER (2009) for information using trace elements for mixing models, as well as a terrific contra dance at Glen Echo in June, 2007.

Hal Stern, UC Irvine, for working with me to clarify issues surrounding the use of the Wishart distribution.

Stable Isotope Sourcing using Sampling

by

Erik Barry Erhardt

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Statistics

The University of New Mexico

Albuquerque, New Mexico

December, 2009

Stable Isotope Sourcing using Sampling

by

Erik Barry Erhardt

B.A., Franklin Pierce College, 1997
M.S., Worcester Polytechnic Institute, 2003
Ph.D., Statistics, University of New Mexico, 2009

Abstract

Stable isotope sourcing is used to estimate proportional contributions of sources to a mixture, such as in the analysis of animal diets, plant nutrient use, geochemistry, pollution, and forensics. We focus on animal ecology because of the particular complexities due to the process of digestion and assimilation. Parameter estimation has been a challenge because there are often many sources and few isotopes leading to an underconstrained linear system for the diet probability vector. This dissertation offers three primary contributions to the mixing model community. (1) We detail and provide an R implementation of a better algorithm (SISUS) for representing possible solutions in the underconstrained case (many sources, few isotopes) when no variance is considered (Phillips and Gregg, 2003). (2) We provide general methods for performing frequentist estimation in the perfectly-constrained case using the delta method and the bootstrap, which extends previous work applying the delta method to two- and three-source problems (Phillips and Gregg, 2001). (3) We propose two Bayesian models, the implicit representation model estimating the population mean

diet through the mean mixture isotope ratio, and the explicit representation model estimating the population mean diet through mixture-specific diets given individual isotope ratios. Secondary contributions include (4) estimation using summaries from the literature in lieu of observation-level data, (5) multiple methods for incorporating isotope ratio discrimination (fractionation) in the analysis, (6) the use of measurement error to account for and partition more uncertainty, (7) estimation improvements by pooling multiple estimates, and (8) detailing scenarios when one model is preferred over another. We show that the Bayesian explicit representation model provides more precise diet estimates than other models when measurement error is small and informed by the necessary calibration measurements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The goal of stable isotope sourcing is to estimate the proportional contributions of sources to a mixture. Such models are used in the analysis of animal diets, plant nutrient use, geochemistry, pollution, and forensics. Stable isotope sourcing models are increasingly used to help understand foodwebs, water sources in soils, plants, or water bodies, geological sources for soils or marine systems, decomposition and soil organic matter dynamics, tracing animal migration patterns, and evaluate management scenarios (Phillips, 2001; Phillips and Koch, 2002; Phillips and Gregg, 2003; Martínez del Rio and Wolf, 2005). Furthermore, these models can be applied to other data, such as trace element geochemistry in earth and environmental science, and combined with stable isotope data to solve sourcing problems. Studies of trophic interactions in food webs, the dynamics of nutrient allocation to tissues and reproduction, and the economics of animal behavior represent a few of the many topics that animal biologists are increasingly calling on stable isotope methods to provide quantitative, robust answers (Martínez del Rio and Wolf, 2005). From a modeling standpoint a rich application is in animal ecology because of the additional complexity of the preferential assimilation of elements from given sources into different tissues, so in this dissertation we focus our attention here. Stable isotopes provide

quantitative insight into these interactions because animal tissues carry an isotopic imprint of the processes and resources used for their synthesis (DeNiro and Epstein, 1978).

Stable isotope analyses of a consumer animal's tissues (the "mixture") and their potential prey and diet (the "sources") is a powerful and well-studied means of quantifying relative contributions of isotopically distinct dietary components (Hobson and Wassenaar, 1999). Stable isotope analysis provides several benefits in comparison with traditional methods for quantifying diet, such as the analysis of stomach and fecal contents (Cree et al., 1999). In particular, stable isotope analysis (1) provides information for all the individuals sampled, including those with empty stomachs, (2) avoids the bias resulting from the differential digestion of soft- vs. hard-bodied prey, (3) provides information on the foods assimilated, not just ingested, (4) gives information on diet assimilated over a relatively long time window, rather than just the time "snapshot" provided by stomach and fecal analysis, (5) provides long-term information on habitat use without the logistical difficulties of marking and resighting, (6) can be performed using a variety of tissues depending on the time frame of interest (e.g., bone can reflect life-long diet while blood reflects recent diet), and (7) can be nonlethal (Hobson and Wassenaar, 2008, 1999). In this dissertation, we assume the mixture's diet has been stationary long enough that the isotope ratio of the tissue of interest at the time of measurement represents the current diet of the given sources consumed rather than reflecting what was consumed some time past. Additionally, any processes affecting the observations should be stationary at the time of observation. That is, source isotope ratio measurements should reflect the diet of the mixture at the time of consumption.

Parameter estimation for source contributions has thus far been a challenge because there are often few isotopes to provide information on the many sources considered. In general, the equations used to model the mean structure represent physical or biological processes, and the associated parameters represent constants of

these processes having a meaning outside the model. Also, there is a now wealth of supplementary (prior) information available about the proportion values from studies reported in the literature concerning similar processes under similar conditions. Also, supplemental information, such as measurement error, has traditionally been discarded but can inform the analysis when appropriately incorporated in the model. Since there are often more parameters to estimate than independent sources of data and there is prior information, it is clearly of interest to envision a Bayesian approach.

1.1 Stable isotopes

The isotope ratio, δ, is the fundamental measurement used as data for mixing models. The isotope ratio is a normalized ratio of the number of the rarer to common isotopes in a sample relative to an international standard given in parts per thousand (per mil, ‰), $\delta = 1000(R_{\text{sample}}/R_{\text{standard}} - 1)$, where R_{sample} and R_{standard} are the ratios of the number of rare to common isotopes for the sample and standard, respectively (see Appendix A). Carbon and nitrogen are among the most commonly used elements used for diet sourcing. For example, the approximate isotope ratio ranges for bulk tissues commonly used are -65‰ to -5‰ (V-PDB standard) for $\delta^{13}C$ and -2‰ to $+25$‰ (air standard) for $\delta^{15}N$ (Hobson and Wassenaar, 2008, Tab. 2.1). The stable isotope ratio is sometimes referred to as a "signature" since it may be characteristic of the measured source. The I-dimensional vectors of isotope ratios will be designated by $\underset{\sim}{\beta}$ for the mixture (consumer) and by $\underset{\sim}{\delta}_s$ for source $s = 1, \ldots, S$.

Discrimination, Δ, or fractionation (see Appendix A for the technical difference) is the systematic change of isotope ratios as elements are ingested, excreted, or catabolized (e.g., trophic fractionation, DeNiro and Epstein, 1981; Minagawa and Wada, 1984; Martínez del Rio and Wolf, 2005). Discrimination is defined as the difference of the isotope ratio in the source, $\underset{\sim}{\delta}_s$, and the resulting isotope ratio in the mixture's tissues, $\underset{\sim}{\delta}'_s$, as $\underset{\sim}{\Delta}_s = \underset{\sim}{\delta}'_s - \underset{\sim}{\delta}_s$. Fractionation can vary depending on the characteristics

of the consumer, such as diet composition or feeding rate (Vander Zanden and Rasmussen, 2001; Post, 2002; McCutchan et al., 2003; Vanderklift and Ponsard, 2003; Martínez del Rio and Wolf, 2005). Despite this substantial variability, studies typically assume constant fractionation rates and ignore the associated uncertainty (but see Vander Zanden and Rasmussen, 2001). While fractionation is the correct and accurate method for measuring the isotope ratio offset between two substances, it relies on an equilibrium constant for the associated reaction, thus the discrimination (or isotope separation) is often used in practice and is usually sufficiently accurate (Hobson and Wassenaar, 2008, ch. 3.II).

Concentration refers to the proportional amount of each element a source contains, such as carbon (C) and nitrogen (N). For example, some carnivores, such as piscivores, eat just one class of foods (fish) which may exhibit a fairly restricted range of C and N concentrations on a whole body basis. At the other extreme, omnivores consume both plant and animal food sources which may differ greatly in C and N concentrations. One might expect that the proportion of N derived from a N-rich food source such as meat might be higher than from plant material.

Assimilation efficiency refers to the proportional amount of each element from a source synthesized into the mixture's tissues. When dietary sources provide an element in just one macromolecular form then that element may be assimilated and metabolized in a uniform fashion. For example, dietary N is supplied by protein, which may be digested in a broadly similar fashion regardless of diet type. In contrast, dietary C can exist as carbohydrate, lipid, or protein, each with a distinct $\delta^{13}C$ value (Schoeller et al., 1984). Dietary proteins and lipids may be preferentially "routed" to synthesis of body proteins or lipids, respectively (Krueger and Sullivan, 1984; Ambrose and Norr, 1993; Tieszen and Fagre, 1993). In such a situation, C isotopes in body proteins would be disproportionately labeled by dietary proteins, leading to an over-estimate of the fraction of protein-rich foods in the consumer's diet. Similarly, dietary lipids may be routed to synthesis of body fat (Stott et al.,

1997).

Although physiological ecologists estimate the assimilation efficiency for food types and even specific nutrients routinely, there are few accounts of the efficiency with which different elements are assimilated. Incorporating food stoichiometry in the models described in this dissertation requires more empirical work. It requires analyzing (or at least estimating) the food's elemental composition and may require determining the efficiency with which different elements in each diet are assimilated. Field researchers may understandably complain that models incorporating concentration and assimilation efficiency require additional data to the basic mixing model (Robbins et al., 2002), but the simple linear mixing models that dominate the literature are misleading if the elemental composition of diet components differs substantially (Martínez del Rio and Wolf, 2005). Considering the potential effect of a source's elemental composition and differential assimilation on isotopic incorporation adds realism to mixing models.

1.1.1 Variation affecting stable isotope ratios

Observed stable isotope ratios may have substantial variability (Phillips and Gregg, 2001). Additional sources of variation to consider in the sampling design and analysis of diet include tissue type differences in isotope ratio (Post et al., 2007), discrimination (Tieszen et al., 1983), concentration (Phillips and Koch, 2002), assimilation efficiency (Martínez del Rio and Wolf, 2005), sample preservation (Sarakinos et al., 2002), and temporal and spatial variablility (Cabana and Rasmussen, 1996; Vander Zanden and Rasmussen, 1999; O'Reilly et al., 2002; Post, 2002; Hobson and Wassenaar, 2008). Here we do not consider tissue type, sample preservation, temporal, or spatial variability. We assume the tissues of the mixture (consumer) have reached equilibrium given the current diet of known sources, and that the source isotope ratio and discrimination distributions are invariant.

1.2 Model components and formulations

The models described in this dissertation have three distinct components: the mean defining equation, the data likelihoods, and the parameter priors. The **mean defining equation** specifies how the mixture population mean (or a subject-specific mixture mean) isotope ratio relates to the diet-weighted average of the source isotope ratio population means. The **likelihood**, or data sampling distribution, describes probabilistic distributions of the data components. The **prior**, used in the Bayesian formulation of the model, encapsulates prior belief about the parameters in the model. Notational conventions use Roman characters for data and Greek characters for parameters to be estimated, except those with a 0 subscript which are specified as prior information.

Two model formulations are presented, differing only in the modeling of the diet proportions. The primary parameter of interest is the population mean diet, $\underset{\sim}{\pi}$, and this is estimated by implicitly or explicitly modeling the subject-specific mixture diets, $\underset{\sim}{\pi}_j$, $j = 1, \ldots, J$. The **implicit representation model** does not model the subject-specific diets directly, but treats the mixture isotope ratio observations as drawn from a common population of mixture isotope ratios. This is what has been done to date (Phillips and Gregg, 2001). The **explicit representation model** hierarchically models the subject-specific isotope ratios conditional on subject-specific diets and takes these diets as drawn from a distribution centered at the population mean diet, $\underset{\sim}{\pi}$. This is a new approach.

1.2.1 Mean defining equation

The ecological literature refers to the mean defining equation as a **mixing model** and a great deal of effort has gone into the correct formulation of this equation that relates the mean isotope ratios of the sources to that of the mixture through the dry-

weight biomass of the mixture diet. Phillips (2001), in his critique of mixing models, argues that the linear mass-balance mixing model provides mathematically unbiased expected values of the proportional contributions of dietary sources to a mixture. Thus far, mixing models have a common linear structure that can be compactly represented using matrix notation $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$. The I-by-S matrix $\mathbf{A}$ includes population mean isotope ratios for the sources, with I isotopes measured on S sources. The vector $\underset{\sim}{\beta}$ includes the population mean isotope ratio for the mixture. Data informs the contents of matrix $\mathbf{A}$ and vector $\underset{\sim}{\beta}$ and we are interested to estimate the mixture population mean diet vector $\underset{\sim}{\pi}$ given the specified sources, where $\underset{\sim}{\pi}$ represents population mean dry-weight biomass proportion of each source consumed.

The implicit representation model does not model the subject-specific mixture diets directly, but implicitly uses the individual diets to get an estimate of the population mean diet through $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$. The explicit representation model does model the subject-specific mixture diets and everything in this chapter can be modified to include them by, $\underset{\sim}{\beta}_j = \mathbf{A}\underset{\sim}{\pi}_j$, $j = 1, \ldots, J$, where J is the number of individual mixtures in the sample. It is then the role of the statistical model to estimate the population mean diet, $\underset{\sim}{\pi}$, from the subject-specific diets, $\underset{\sim}{\pi}_j$.

1.2.1.1 Basic mixing model

The basic mixing model, BMM, is the simplest mass-balance mixing model assuming that the mean isotope ratio of the mixture equals the diet-weighted average of the mean discrimination-corrected isotope ratio composition of the sources (DeNiro and Epstein, 1978; Schwarcz, 1991; Phillips, 2001). Each coefficient δ'_{is} is the observed isotope ratio value for isotope i from source s, δ_{is}, corrected by the addition of the isotope ratio discrimination (or fractionation), Δ_{is}, $\delta'_{is} = \delta_{is} + \Delta_{is}$. Coefficient β_i is the observed isotope ratio for isotope i in the mixture. It is sometimes convenient to

represent the defining equation in matrix notation,

$$\underset{\sim}{\beta} \;=\; \mathbf{A}\underset{\sim}{\pi} \;=\; \sum_{s=1}^{S}(\underset{\sim}{\delta}_s + \underset{\sim}{\Delta}_s)\pi_s \;=\; \sum_{s=1}^{S}\underset{\sim}{\delta}'_s\pi_s \tag{1.1}$$

$$\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_I \end{bmatrix}_{I\times 1} = \begin{bmatrix} \delta'_{11} & \delta'_{12} & \cdots & \delta'_{1S} \\ \delta'_{21} & \delta'_{22} & \cdots & \delta'_{2S} \\ \vdots & \vdots & \vdots & \vdots \\ \delta'_{I1} & \delta'_{I2} & \cdots & \delta'_{IS} \end{bmatrix}_{I\times S} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_S \end{bmatrix}_{S\times 1} . \tag{1.2}$$

Mean diet vector $\underset{\sim}{\pi}$ is constrained to be a probability vector restricted to the **simplex**, that is, the sum of the source proportions is 1 and each proportion is a number between 0 and 1,

$$1 \;=\; \sum_{s=1}^{S}\pi_s, \quad 0 \le \pi_s \le 1, \quad s = 1, \ldots, S. \tag{1.3}$$

In certain situations it is convenient to augment $\underset{\sim}{\beta}$ and $\mathbf{A}$ with a row of ones to include the simplex constraint forcing the diet proportions to sum to one, indicated by an x subscript,

$$\underset{\sim}{\beta}_{\mathrm{x}} \;=\; \begin{bmatrix} \underset{\sim}{\beta} \\ 1 \end{bmatrix} \qquad \text{and} \qquad \mathbf{A}_{\mathrm{x}} \;=\; \begin{bmatrix} \mathbf{A} \\ \underset{\sim}{1}^{\top} \end{bmatrix} \qquad \text{for} \qquad \underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}. \tag{1.4}$$

The model can also be written in the form $\underset{\sim}{0} = \mathbf{B}\underset{\sim}{\pi}$, where $\mathbf{B} = \mathbf{A} - \underset{\sim}{\beta}\underset{\sim}{1}^{\top}$, or alternatively

$$\underset{\sim}{0} \;=\; \mathbf{B}\underset{\sim}{\pi} \tag{1.5}$$

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \delta'_{11} - \beta_1 & \cdots & \delta'_{1S} - \beta_1 \\ \delta'_{21} - \beta_2 & \cdots & \delta'_{2S} - \beta_2 \\ \vdots & \vdots & \vdots \\ \delta'_{I1} - \beta_I & \cdots & \delta'_{IS} - \beta_I \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_S \end{bmatrix}. \tag{1.6}$$

If we incorporate the simplex constraint into the left and right hand sides of (1.6) we get

$$\begin{bmatrix} \underset{\sim}{0} \\ 1 \end{bmatrix} = \mathbf{B}_{\mathrm{x}}\underset{\sim}{\pi} \qquad \text{where} \qquad \mathbf{B}_{\mathrm{x}} = \begin{bmatrix} \mathbf{B} \\ \underset{\sim}{1}^{\top} \end{bmatrix}. \tag{1.7}$$

1.2.1.2 Extended mixing model

The concentration mixing model(CMM) makes more realistic modeling possible by accounting for differences in the elemental concentrations in each source. Concentration describes the dry-weight biomass proportion of a source that is composed of a given atomic element, such as certain terrestrial plants being about 0.44 carbon and 0.01 nitrogen, and certain terrestrial meat being about 0.52 carbon and 0.14 nitrogen. The CMM assumes that a source's elemental contribution is proportional to the contributed biomass times the elemental concentration in that source (Phillips and Koch, 2002). Elemental concentrations of dietary sources can vary widely. Among plants in terrestrial environments, for example, C:N ratios, the relative concentration of carbon to nitrogen, of leaves can vary by two orders of magnitude and animal tissues may be an order of magnitude more enriched than plant materials with the lowest C:N ratios (Sterner and Elser, 2002). For omnivores, such as a brown bear, feeding on salmon as well as grasses and fruit, the nitrogen concentration of these materials can vary by an order of magnitude or more. As a consequence, using a mixing model neglecting these differences in elemental concentrations can potentially produce large errors in estimates of source contributions to the bear's diet (see Phillips and Koch, 2002). Elemental concentrations of resources are easy to incorporate in the model by measuring the mass of the sample before combusting in the elemental analyzer and isotope ratio mass spectrometer.

The extended mixing model (EMM) further increases the realism and accuracy of our mixing model by recognizing that the digestive efficiencies of consumers for

different food types can vary considerably (Martínez del Rio and Wolf, 2005). Ignoring the differences in the relative assimilation of nutrients derived from different sources can produce significant errors in the estimation of source proportions (Koch and Phillips, 2002). Using the brown bear example again, the assimilation efficiency of the bear's digestive system for salmon is much higher than that of grasses (around 1.00 versus 0.35) and may differ for carbon and nitrogen. The EMM acknowledges that "you are what you assimilate", and accounts for the large differences in the assimilation efficiencies observed in omnivores eating both plant and animal materials (DeNiro and Epstein, 1978; Koch and Phillips, 2002; Martínez del Rio and Wolf, 2005).

The EMM estimates both the proportions of biomass consumed, $\underset{\sim}{\pi}$, and the proportions of each assimilated element from each source, ϖ_{is}. Equation (1.8) extends (1.1) by multiplying the proportions π_s by the population mean concentration of element i in source s, κ_{is}, and by the population mean mixture's assimilation efficiency of element i from source s, η_{is},

$$\beta_i = \frac{\sum_{s=1}^{S} \delta'_{is}\kappa_{is}\eta_{is}\pi_s}{\sum_{s=1}^{S} \kappa_{is}\eta_{is}\pi_s}, \quad i = 1, \ldots, I. \tag{1.8}$$

To write this in the matrix notation of (1.5), the form of (1.8) must first be rewritten as

$$0 = \sum_{s=1}^{S} (\delta'_{is} - \beta_i)\kappa_{is}\eta_{is}\pi_s, \quad i = 1, \ldots, I, \tag{1.9}$$

then we have

$$\underset{\sim}{0} = \begin{bmatrix} (\delta'_{11} - \beta_1)\kappa_{11}\eta_{11} & \cdots & (\delta'_{1S} - \beta_1)\kappa_{1S}\eta_{1S} \\ (\delta'_{21} - \beta_2)\kappa_{21}\eta_{21} & \cdots & (\delta'_{2S} - \beta_2)\kappa_{2S}\eta_{2S} \\ \vdots & \vdots & \vdots \\ (\delta'_{I1} - \beta_I)\kappa_{I1}\eta_{I1} & \cdots & (\delta'_{iS} - \beta_i)\kappa_{IS}\eta_{IS} \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_S \end{bmatrix}. \tag{1.10}$$

Given biomass proportions π_s, the proportion of each element contributed by each source, ϖ_{is}, may also be of interest, given by

$$\varpi_{is} = \frac{\kappa_{is}\eta_{is}\pi_s}{\sum_{s=1}^{S}\kappa_{is}\eta_{is}\pi_s}. \tag{1.11}$$

The intermediate CMM, when assimilation efficiency is not considered, has $\eta_{is} \equiv 1$ for all $s = 1, \ldots, S$ and $i = 1, \ldots, I$. If the assimilation efficiency of an element is the same for all sources, and this is true for each element, i.e., $\eta_{is} = \eta_{is^*}$ for all $s, s^* = 1, \ldots, S$ and $i = 1, \ldots, I$, then the EMM and CMM are equivalent. Additionally, if the concentration of an element is the same for all sources, and this is true for each element, i.e., $\kappa_{is} = \kappa_{is^*}$ for all $s, s^* = 1, \ldots, S$ and $i = 1, \ldots, I$, then the CMM and BMM are equivalent (Phillips and Koch, 2002).

1.2.1.3 Model assumptions

The results of the mixing models rely on the assumptions listed below. Some of these assumptions regard the biological and biochemical nature of the data and some regard statistical issues. Specifically, the assumptions regarding no variability or uncertainty are for use in Chapter 3, and these assumptions are relaxed in later chapters where variation is accounted for.

The BMM provides a base set of assumptions, with the CMM and EMM adding and replacing some assumptions by incorporating additional information. To the degree that these assumptions are met, the model will accurately predict feasible proportional contributions of the sources to the mixture. However, inaccuracies may be great for even small departures from the assumptions. The absence of variability in the model will clearly underestimate the uncertainty in the estimates. Also, for example, poor estimates of the multiplicative factors of concentration and assimilation efficiency when they are near zero can introduce large errors. Nitrogen

concentrations in plant material are often near zero and vary widely among plants, while nitrogen concentrations from meat is often nearer one and varies little among animals. Phillips and Koch (2002) provide a sensitivity analysis when specifying concentration in the model.

The BMM has the following assumptions (b=basic). (b1) All nonignorable sources contributing to the mixture have been included in the model. (b2) The contribution of a source to the mixture is proportional to the consumed dry-weight biomass. (b3) Elemental concentrations of each element is the same for all sources, that is, the sources are stoichiometrically identical (e.g., C:N is constant for all sources, concentrations may vary in the CMM). (b4) Equal proportions of each element are derived from a source (relaxed in CMM). (b5) Elemental assimilation efficiencies of each element into the tissues of the mixture is the same for all sources, (assimilation efficiencies may vary in the EMM). (b6) Elemental concentrations from all dietary sources are completely homogenized in the mixture prior to tissue synthesis (this assumption is not satisfied when, for example, a dietary shift has occurred more recently than the tissue turnover time (Phillips and Eldridge, 2006)). (b7) Isotope ratios for each element are the same for each source tissue type, be it carbohydrate, lipid, protein, or other so that the biomass consumed does not depend on type and proportion of each tissue consumed (this is not satisfied when, for example, dietary proteins in the sources are preferentially routed to synthesis of body proteins in the mixture) (Schoeller et al., 1984). (b8) There is no variation in the model inputs of isotope ratio measurements, isotope ratio discriminations, elemental concentrations in sources, and elemental assimilation of source tissues into the mixture (we will address these variances in later chapters). (b9) The biological needs of the mixture are met with the given solutions, for example, protein requirements and energy demands are satisfied, when considering animals (Minagawa, 1992).

The CMM has the following assumptions (c=concentration). (c1) For each element, the contribution of a source to the mixture is proportional to the consumed

biomass times the elemental concentration in that source, replacing (b2). (c2) The contribution of a source to the mixture's elemental pool depends on how much of each element that source contains, replacing (b3, b4). The remaining assumptions are the same as for the BMM, namely (b1, b5–b9).

The EMM has the following assumptions (e=extended). (e1) For each element, the contribution of a source to the mixture is proportional to the consumed biomass times the elemental concentration in that source times the assimilation efficiency of that element from that source, replacing (c1). (e2) The contribution of a source to the mixture's elemental pool depends on how much of each element that source contains and how efficiently the element is assimilated into the mixture, replacing (c2, b5). The remaining assumptions are the same as for the CMM, namely (b1, b6–b9).

1.2.2 Alaskan bear example

The summaries in Table 1.1, plotted in the convex hull plot in Figure 1.1, reconstructs a data subset from a two-isotope, three-source unique-solution example reused from Koch and Phillips (2002, Table 1). Although these summaries are estimates we will treat them as fixed for this illustration. The objective of the original study was to model the biomass contribution of salmon, terrestrial meat, and fruit to the diet of an average brown bear (*Ursus arctos*) from the Kenai Peninsula, Alaska, at a particular time of year (Jacoby et al., 1999). Concentration for carbon is the proportion of carbon in dry matter, and concentration for nitrogen is the proportion crude protein in dry matter times the proportion nitrogen in protein. Assimilation for carbon is the digestible proportion of dry matter, and assimilation for nitrogen is the digestible proportion of dry matter times the protein digestibility proportion. Note that the product of these concentration and assimilation values are what Koch and Phillips (2002) report as Digest C and Digest N. Our results differ from theirs because they

use Digest [C] and Digest [N] in their calculations, which is the Digest X divided by the proportion digestible dry matter.

Table 1.1: Alaskan bear example, brown bear hair as a mixture of salmon, meat, and fruit, using isotopes of carbon ($i = 1$) and nitrogen ($i = 2$) (Jacoby et al., 1999).

	Isotope Ratios		Discrimination		Concentration		Assim. Effic.	
Mixture	β_1	β_2						
Brown bear	−20.3	10.9						
Sources	δ_{1s}	δ_{2s}	Δ_{1s}	Δ_{2s}	κ_{1s}	κ_{2s}	η_{1s}	η_{2s}
Salmon ($s = 1$)	−20.5	13.2	1.2	2.3	0.548	0.118	1.00	1.00
Meat ($s = 2$)	−21.5	3.9	4.9	4.0	0.515	0.141	1.00	1.00
Fruit ($s = 3$)	−26.6	−0.9	3.3	4.1	0.45	0.0126	0.634	0.571

For the BMM in (1.1) the defining equation is

$$\begin{bmatrix} -20.3 \\ 10.9 \\ 1 \end{bmatrix} = \begin{bmatrix} -19.3 & -16.6 & -23.3 \\ 15.5 & 7.9 & 3.2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}, \tag{1.12}$$

where, for example, $\delta'_{12} = \delta_{12} + \Delta_{12} = (-21.5 + 4.9) = -16.6$. Concentration and assimilation efficiencies are not used in the BMM. The first row represents the equation for carbon, the second row is for nitrogen, and the third row is for the simplex. The first column is for salmon, the second column is for meat, and the third column is for fruit.

In the isotope ratio data space, Figure 1.1, each axis represents values of discrimination-corrected isotope measurements for each pair of isotopes. Figure 1.1 plots the first two rows of (1.12) as carbon/nitrogen pairs, $(\delta'_{1s}, \delta'_{2s})$. In particular, with carbon on the horizontal axis and nitrogen on the vertical axis, salmon is $(-19.3, 15.5)$ and brown bear is $(-20.3, 10.9)$. The **convex hull** is the set of lines connecting the "outside" sources. The hull is called "convex" because any line connecting two points on the hull stays entirely inside or on the hull. A solution to (1.12) exists in the solution space, Figure 3.1, if and only if the mixture brown bear lies inside all pairwise isotopic convex hulls. Two convex hulls are shown in Figure 1.1. The thin straight-lined

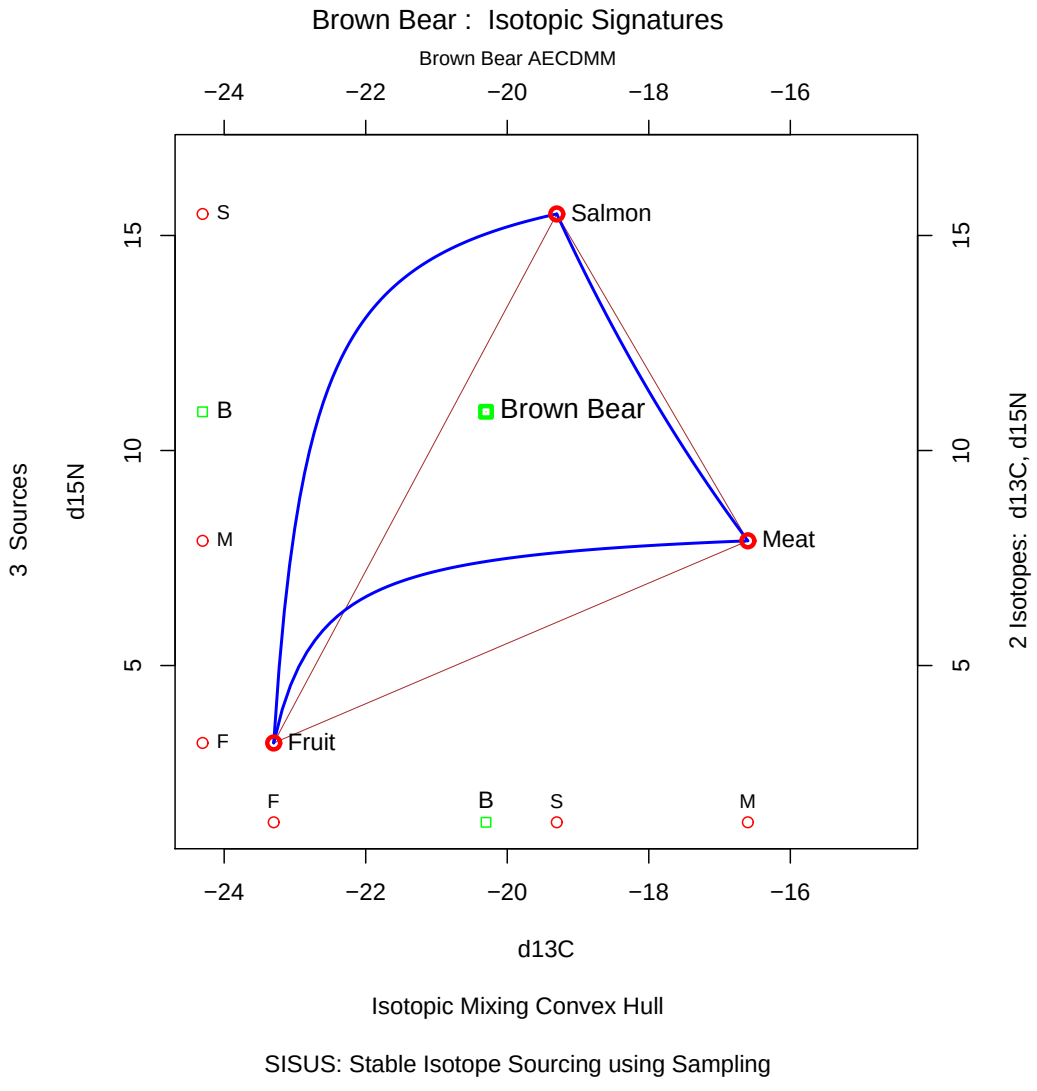

Figure 1.1: Convex hull plot of the Alaskan bear example in Table 1.1. This plot represents the isotope ratio data space of discrimination-corrected carbon and nitrogen isotope ratios. The thin straight-lined convex hull is from the BMM, and the thick curved-lined hull is from the EMM.

convex hull is from the BMM, and the thick curved-lined hull shows the effect of the concentration and assimilation efficiency in the EMM to curve the lines in proportion to the concentration-assimilation ratio, $\kappa_{is}\eta_{is}/\kappa_{i^*s}\eta_{i^*s}$, between sources. Typically, the closer the isotope ratio values of the mixture are to a source's discrimination-corrected isotope ratio values, the more similar the mixture is isotopically to that source, and the larger the contribution of that source to the mixture is able to be.

The matrix $\mathbf{A}_{\mathrm{x}}$ in (1.12) is invertible, thus the solution $\underset{\sim}{\pi}$ to (1.12) is unique for the BMM. The EMM also has a unique solution. Table 1.2 reports the unique solutions for both the BMM and EMM. The first row gives proportions of biomass from each of the three sources constituting the mixture using the BMM estimating that brown bear hair was sourced from 0.59 salmon, from 0.10 meat, and from 0.31 fruit. In contrast, the EMM estimates 0.22, 0.20, and 0.58. Thus, the incorporation of concentration and assimilation efficiency reverses the estimated importance of salmon and fruit. Additionally, the EMM estimates that, in spite of the 0.20 contribution of meat, 0.49 of the nitrogen present in the brown bear hair was sourced from the meat.

Table 1.2: Unique solutions for the Alaskan bear example in Table 1.1 using the BMM and the EMM.

Brown bear		Salmon	Meat	Fruit
Model	Proportions	$s=1$	$s=2$	$s=3$
BMM	**Biomass**	$\pi_1 = 0.59$	$\pi_2 = 0.10$	$\pi_3 = 0.31$
EMM	**Biomass**	$\pi_1 = 0.22$	$\pi_2 = 0.20$	$\pi_3 = 0.58$
EMM	Carbon $i=1$	$\varpi_{11} = 0.31$	$\varpi_{12} = 0.27$	$\varpi_{13} = 0.43$
EMM	Nitrogen $i=2$	$\varpi_{21} = 0.44$	$\varpi_{22} = 0.49$	$\varpi_{23} = 0.07$

1.3 Goals and outcomes

The bear example in this chapter ignores the estimation of $\mathbf{A}$ and $\underset{\sim}{\beta}$ and, in general, the solution of $\underset{\sim}{\pi}$ may not be unique. The aim of this dissertation is to address estimation in both the frequentist and Bayesian paradigms using both the implicit and explicit representation of the mean defining equation. Chapter 2 introduces the mean defining equation representations and their justifications. Chapter 3 illustrates a better algorithm for providing solutions to mixing models without variation when there are many more sources than isotopes. Chapter 4 makes rigorous the frequentist approaches for the basic mixing model, and provides general approaches using the delta method and bootstrap. Chapter 5 introduces flexible Bayesian modeling for the basic mixing model. Chapter 6 provides extentions, sets the stage for future work, and concludes this dissertation.

Chapter 2

Statistical framework for stable isotope sourcing

2.1 Mean defining equation

The mean defining equation for the BMM, $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$, in Chapter 1 provides a relationship between the mixture isotope ratio mean, $\underset{\sim}{\beta}$, and the discrimination-corrected isotope ratio means of the sources, $\mathbf{A}$, through the mean mixture diet, $\underset{\sim}{\pi}$. Under what conditions might it be reasonable that $\underset{\sim}{\beta}$ and $\underset{\sim}{\pi}$ represent mean values for the mixture population? If we consider $\mathbf{A}$ to be fixed for each individual in the mixture population, then it is reasonable to posit that a subject-specific diet, $\underset{\sim}{\pi}_j$, satisfies $\underset{\sim}{\beta}_j = \mathbf{A}\underset{\sim}{\pi}_j$, where the resulting individual isotope ratio, $\underset{\sim}{\beta}_j$, results from the diet of the individual. Individuals within the population will have different $\underset{\sim}{\pi}_j$ values, consequently there is a distribution for the $\underset{\sim}{\pi}_j$ values and a corresponding distribution for $\underset{\sim}{\beta}_j$. Without specifying these distributions we can say through linearity, given that $\mathbf{A}$ is fixed, that $\mathrm{E}[\underset{\sim}{\beta}_j] = \mathrm{E}[\mathbf{A}\underset{\sim}{\pi}_j] = \mathbf{A}\mathrm{E}[\underset{\sim}{\pi}_j]$, where $\underset{\sim}{\beta} \equiv \mathrm{E}[\underset{\sim}{\beta}_j]$ and $\underset{\sim}{\pi} \equiv \mathrm{E}[\underset{\sim}{\pi}_j]$. There is evidence supporting modeling of subject-specific diets because differences in diet lead to differences in the isotope ratios of individuals within a population

(Angerbjörn et al., 1994; Matthews and Mazumder, 2004; Urton and Hobson, 2005; Araújo et al., 2007; Layman et al., 2007).

Depending on the number of isotopes and sources considered, three situations regarding $\underset{\sim}{\pi}$ are possible. The **underconstrained** situation where $S > I + 1$ is common, such as the mink example in Chapter 1, where the number of sources considered is many greater than the number of isotopes. Because the dimension of $\underset{\sim}{\pi}$ is greater than the number of equations this will typically lead to an infinite set of solutions. Chapters 3 and 5 discuss this case. The **perfectly constrained** situation where $S = I+1$ is ideal, such as the Alaskan bear example in Chapter 1, where the $\mathbf{A}$ matrix augmented with a row of ones is square and nonsingular providing a unique solution of $\underset{\sim}{\pi}$. Chapters 4 and 5 discuss this case when a unique solution exists. The **overconstrained** situation, where $S < I + 1$, is less common but typically requently occurs when there are two sources and carbon and nitrogen isotopes are both measured. No solution is available using all the isotopes simultaneously, but we discuss methods to provide a single "best" solution. This is discussed in Chapter 6 among other extentions.

Let $\underset{\sim}{\pi}_j$ have a density on the simplex with population mean $\mathrm{E}[\underset{\sim}{\pi}_j] = \underset{\sim}{\pi}$, and consider $\mathbf{A}$ fixed. Because $\underset{\sim}{\pi}_j$ has $1 = \sum_{s=1}^{S} \pi_{js}$, we can write

$$\begin{aligned}
\underset{\sim}{\beta}_j &= \sum_{s=1}^{S} \underset{\sim}{\delta}'_s \pi_{js} \\
&= \sum_{s=1}^{S-1} (\underset{\sim}{\delta}'_s - \underset{\sim}{\delta}'_S)\pi_{js} + \underset{\sim}{\delta}'_S \\
&= \mathbf{A}^* \underset{\sim}{\pi}^*_j + \underset{\sim}{\delta}'_S,
\end{aligned}$$

where $\mathbf{A}^* = [\underset{\sim}{\delta}'_1 - \underset{\sim}{\delta}'_S, \cdots, \underset{\sim}{\delta}'_{S-1} - \underset{\sim}{\delta}'_S]$ is an $I \times (S-1)$-dimensional matrix consisting of the last column $\underset{\sim}{\delta}'_S$ taken from the first $S-1$ columns of $\mathbf{A}$ and $\underset{\sim}{\pi}^*_j$ is an $(S-1)$-dimensional vector consisting of the first $S-1$ rows of $\underset{\sim}{\pi}_j$. The density for $\underset{\sim}{\pi}^*_j$ defines the density for $\underset{\sim}{\pi}_j$. It is now clear that $\underset{\sim}{\beta}_j$ can have a density when $S \geq I+1$ because

$\underset{\sim}{\beta}_j \in \mathbb{R}^I$ and $\underset{\sim}{\pi}_j \in \mathbb{R}^{S-1}$. Furthermore, the first two moments of this distribution, given $\mathbf{A}$, are

$$\mathrm{E}[\underset{\sim}{\beta}_j] \;=\; \mathrm{E}[\mathbf{A}\underset{\sim}{\pi}_j] \;=\; \mathrm{E}[\mathbf{A}^*\underset{\sim}{\pi}^*_j + \underset{\sim}{\delta}'_S] \;=\; \mathbf{A}^*\mathrm{E}[\underset{\sim}{\pi}^*_j] + \underset{\sim}{\delta}'_S \;=\; \mathbf{A}^*\underset{\sim}{\pi}^* + \underset{\sim}{\delta}'_S \;\equiv\; \mathbf{A}\underset{\sim}{\pi}$$

where $\underset{\sim}{\pi}^* = [\pi_1, \ldots, \pi_{S-1}]^\top$, and

$$\mathrm{Cov}[\underset{\sim}{\beta}_j] \;=\; \mathrm{Cov}[\mathbf{A}\underset{\sim}{\pi}_j] \;=\; \mathrm{Cov}[\mathbf{A}^*\underset{\sim}{\pi}^*_j + \underset{\sim}{\delta}'_S] \;=\; \mathbf{A}^*\mathrm{Cov}[\underset{\sim}{\pi}^*_j]\mathbf{A}^{*\top} \;=\; \mathbf{A}^*\Sigma_{\underset{\sim}{\pi}^*}\mathbf{A}^{*\top}.$$

Note that $\Sigma_{\underset{\sim}{\pi}^*}$ is a square $(S-1)$-dimensional matrix and the $\mathrm{Cov}[\underset{\sim}{\beta}_j]$ is a square I-dimensional matrix. In the absense of a parametric model for $\underset{\sim}{\pi}_j$ in the case when $S \geq I+1$ it would be reasonable to allow $\mathrm{Cov}[\underset{\sim}{\beta}_j]$ to be arbitrary, but $\Sigma_{\underset{\sim}{\pi}^*}$ can be recovered from an estimate of $\mathrm{Cov}[\underset{\sim}{\beta}_j]$ only when $S = I+1$. Thus, in the underconstrained case ($S < I+1$) we assume $\underset{\sim}{\beta}_j$ has a density with mean $\mathbf{A}\underset{\sim}{\pi}$ and arbitrary covariance matrix. Further detail is given in Section 2.2.1.

2.1.1 Invertibility of $\mathbf{A}_\mathrm{x}$ and $\mathbf{B}_\mathrm{x}$

In the perfectly constrained case ($S = I+1$) it will be assumed with little loss of generality that square matrix $\mathbf{A}_\mathrm{x}$ will be invertible. The elements of $\mathbf{A}_\mathrm{x}$ are parameters defined over the real numbers. Over the field of real numbers, the set of singular S-by-S matrices, considered as a subset of $\mathbb{R}^{S\times S}$, has Lebesgue measure zero. Thus, almost all S-by-S matrices are invertible. This does not guarantee that $\mathbf{A}_\mathrm{x}$ is invertible, but does make it invertible with probability 1. In the frequentist analyses of Chapters 4 and 6 we assume that $\mathbf{A}_\mathrm{x}$ is invertible, which implies that $\underset{\sim}{\beta}_\mathrm{x} = \mathbf{A}_\mathrm{x}\underset{\sim}{\pi}$ has a unique solution. This is not an issue in the Bayesian analysis because inverting $\mathbf{A}_\mathrm{x}$ is not necessary.

Here we show that if $\mathbf{A}_\mathrm{x}$ is invertible when $S = I+1$, then $\mathbf{B}_\mathrm{x}$ is also invertible. The following are equivalent: (1) $\mathbf{A}_\mathrm{x}$ is invertible, (2) the row rank of $\mathbf{A}_\mathrm{x}$ is S, (3)

$c_i \equiv 0, \forall i = 1, \ldots, S$ is the only set of $c_i \in \mathbb{R}$ such that $c_S \underset{\sim}{1}^\top + \sum_{i=1}^{S-1} c_i \underset{\sim}{\delta}_i^{\text{tr}} = \underset{\sim}{0}^\top$, where $\underset{\sim}{\delta}_i^{\text{tr}}$ is row i of $\mathbf{A}$. We need to show that this implies $c_S \underset{\sim}{1}^\top + \sum_{i=1}^{S-1} c_i (\underset{\sim}{\delta}_i^{\text{tr}} - \beta_i \underset{\sim}{1}^\top) = \underset{\sim}{0}^\top$. Adding β_i to all columns of row i is adding a multiple of $\underset{\sim}{1}^\top$ to each row, thus we have a new constant c_S^* in statement (3) above, $c_S^* \underset{\sim}{1}^\top = (c_S \underset{\sim}{1}^\top - \sum_{i=1}^{S-1} c_i \beta_i \underset{\sim}{1}^\top)$. Thus $\mathbf{B}_{\text{x}}$ is invertible.

2.2 Two representations for the mixture diet

We consider two model formulations. The **implicit representation**, or means representation, models the $\underset{\sim}{\pi}_j$ implicitly through the variation in the estimates of $\underset{\sim}{\beta}$ and $\mathbf{A}$ to estimate $\underset{\sim}{\pi}$. The **explicit representation**, or subject-specific representation, models the $\underset{\sim}{\pi}_j$ directly where each $\underset{\sim}{\beta}_j$ is defined as the function $\mathbf{A}\underset{\sim}{\pi}_j$ and the $\underset{\sim}{\pi}_j$ are distributed on the simplex with mean $\underset{\sim}{\pi}$.

In either representation, $\underset{\sim}{\beta}_j$ is assumed to have a density in the underconstrained or perfectly constrained situations, when $S \geq I + 1$, when $\underset{\sim}{\pi}_j$ has a density on the simplex, as discussed above. It is important to take into account the structure of the relationship between $\underset{\sim}{\beta}_j$ and underlying $\underset{\sim}{\pi}_j$. In the implicit representation the density is directly on $\underset{\sim}{\beta}_j$ and the $\underset{\sim}{\pi}_j$ are not modeled directly. In the explicit representation the density is on $\underset{\sim}{\pi}_j$ which imposes a density on $\underset{\sim}{\beta}_j$. We will develop these models for the BMM, and they are easily extended to the EMM. We will use the frequentist paradigm for the implicit representation, and Bayesian paradigm for both the implicit representation and explicit representation. With multiple measurements on individual mixtures, frequentist methods in principle can also be used effectively with the explicit representation. This view of modeling the mean isotope ratio of the mixture conditional on the source isotope ratios and the subject-specific diet proportion vector suggests the two following hierarchical models.

2.2.1 Implicit representation

In the implicit representation we do not model the individual diet vectors, $\underset{\sim}{\pi}_j$, but consider the variation in the isotope ratio values in the mixture and sources to make inference in the mean population diet, $\underset{\sim}{\pi}$. Let the mixture isotope ratio vectors have $\mathrm{E}[\underset{\sim}{\beta}_j] = \underset{\sim}{\beta}$ and $\mathrm{Cov}[\underset{\sim}{\beta}_j] = \Sigma_{\underset{\sim}{\beta}}$. The covariance matrix for the $\underset{\sim}{\beta}_j$ could conceivably be anything, as noted in Section 2.1. The multivariate normal is a distribution allowing for arbitrary mean vectors and covariances matrices, so as a first approximation we can consider $\underset{\sim}{\beta}_j \sim \mathrm{Normal}(\underset{\sim}{\beta}, \Sigma_{\underset{\sim}{\beta}})$. We can write this as

$$\underset{\sim}{\beta}_j \;=\; \underset{\sim}{\beta} + \underset{\sim}{\psi}_j \;=\; \mathbf{A}\underset{\sim}{\pi} + \underset{\sim}{\psi}_j \qquad \text{with} \qquad \underset{\sim}{\psi}_j \;\sim\; \mathrm{Normal}(\underset{\sim}{0}, \Sigma_{\underset{\sim}{\beta}}). \tag{2.1}$$

The error term $\underset{\sim}{\psi}_j$ represents between-individual variability. The implicit representation provides a framework for the modeling that has been developed to date (Phillips, 2001; Phillips and Koch, 2002; Phillips and Gregg, 2003; Martínez del Rio and Wolf, 2005).

In the perfectly constrained situation ($S = I + 1$), the estimate of $\underset{\sim}{\pi}$ using frequentist large sample theory is performed by substituting estimates for $\mathbf{A}$ and $\underset{\sim}{\beta}$ in $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$ subject to $\underset{\sim}{1}^\top \underset{\sim}{\pi} = 1$, and solving to give $\hat{\underset{\sim}{\pi}} = \hat{\mathbf{A}}_{\mathrm{x}}^{-1}\hat{\underset{\sim}{\beta}}_{\mathrm{x}}$ (Phillips, 2001; Phillips and Gregg, 2001; Martínez del Rio and Wolf, 2005). This is detailed in Chapter 4.

In the underconstrained situation ($S > I + 1$), IsoSource and SISUS are two methods for representing the set of $\underset{\sim}{\pi}$ satisfying $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$. In Chapter 5.2.1 we see that this approximates a Bayesian analysis where the posteriors for $\underset{\sim}{\beta}$ and $\mathbf{A}$ go to point masses, that is the variance terms are effectively zero. Chapter 3 details the no-variance approximation and Chapter 5 illustrates the implicit representation model this method approximates.

In the overconstrained situation ($S < I + 1$), $\underset{\sim}{\beta}_j = \mathbf{A}\underset{\sim}{\pi}_j$ may not have a density. Estimates for $\underset{\sim}{\pi}$ using frequentist large sample theory uses least squares strategies.

The Bayesian paradigm uses the same model as in the other situations, and because of measurement error and within-individual variation in the mixture isotope ratio observations informing $\underset{\sim}{\beta}_j$, both the observations and their means, $\underset{\sim}{\beta}_j$, will have a density.

2.2.2 Explicit representation

The explicit representation models the subject-specific diets, $\underset{\sim}{\pi}_j$, through a subject-specific version of (1.1), leading to an analogous way to make inferences about $\underset{\sim}{\pi}$ through

$$\underset{\sim}{\beta}_j \;=\; \mathbf{A}\underset{\sim}{\pi}_j. \tag{2.2}$$

The $\underset{\sim}{\pi}_j$s are latent variables, as they cannot be directly observed. We will assume a distribution for $\underset{\sim}{\pi}_j$ which induces a distribution on $\underset{\sim}{\beta}_j$. This strategy provides a more natural way to think about the diet problem and a principled approach for estimating $\underset{\sim}{\pi}$, elaborated on in Chapter 5.

Chapter 3

Models without variation

In this chapter we describe an efficient algorithm for providing representative probabilistic exact solutions to mixing models without variation. A common situation is to have many more sources than isotopes, resulting in an underconstrained situation ($S > I + 1$). In this situation there are multiple exact solutions to the defining relationship $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$ in (1.4), or equivalently $[\underset{\sim}{0}^{\top}, 1]^{\top} = \mathbf{B}_{\mathrm{x}}\underset{\sim}{\pi}$ given in (1.7). This set of solution vectors, $\underset{\sim}{\pi}$, is called the **solution polytope**. The strategy that Phillips and Gregg (2003) introduced for providing solutions for $\underset{\sim}{\pi}$ has been to ignore variability and uncertainty in parameter estimation and provide approximate deterministic solutions to (1.4) using the BMM implemented in IsoSource software. With hundreds of papers published using this method, people have tried to increase the computational efficiency (Lubetkin and Simenstad, 2004), and researchers continue to praise IsoSource, and even use "IsoSource" in the title of their articles (Benstead et al., 2006). Here we demonstrate a probabilistic sampling-based method, implemented in the author's SISUS R package, performing in many ways better than the qualitatively similar deterministic methods.

Methods ignoring uncertainty inherent in parameter estimation are discouraged. Their usefulness may be limited to providing quick and overly-precise estimates. For

inference models accounting for uncertainty at all levels of uncertainty should be used, such as those described in Chapter 5.

3.1 Relationship between data and solution spaces

The data consists of I discrimination-corrected isotope ratio means on each of S sources, thus, the isotope ratio data is I-dimensional while the source proportion solutions, $\underset{\sim}{\pi}$, are S-dimensional. Using the Alaskan bear example in Table 1.1, we now describe the relationship defined by (1.4) between the $I = 2$-dimensional isotope ratio data space shown in the convex hull plot in Figure 1.1 and the $S = 3$-dimensional proportion solution space in Figure 3.1. To aid understanding and clarity we use the BMM in this description, however, an analogous description can be given for the EMM. We later use the EMM for the analysis. When the relevant data is available we encourage consideration of the EMM for analyses, since large errors can be corrected by incorporating accurately estimated concentration and assimilation information, as seen by the difference in solutions in Table 1.2.

In the Alaskan bear example, because $S = I + 1$, there is a unique solution given by $\underset{\sim}{\pi} = \mathbf{A}_{\mathrm{x}}^{-1}\underset{\sim}{\beta}_{\mathrm{x}}$, because $\mathbf{A}_{\mathrm{x}}^{-1}$ exists. However, often many sources but few isotopes are considered, $S > I + 1$, and the solution is not unique. Thus, we will use a subset of this small example to view a solution space without a unique solution as an analogy to larger situations.

In the source-proportion solution space each axis represents proportions of biomass contributed for each of S sources, and for $S > 3$ this is not plottable. In the Alaskan bear example, $S = 3$. Figure 3.1 shows the solution space where the axes are labeled for salmon, meat, and fruit. Each of the $I + 1$ linear equations, as rows in (1.12), define an $(S - 1)$-dimensional hyperplane in the S-dimensional proportion solution space. A common 2-dimensional plane is an example of a hyperplane, but in general

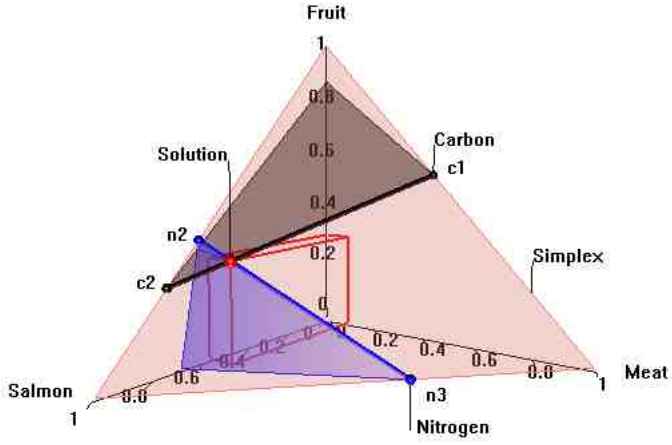

Figure 3.1: Source-proportion solution space for salmon, meat, and fruit in the Alaskan bear example using the BMM. The axes represent biomass contributions of each source, π_1, π_2, and π_3. The data in Table 1.1 assign the values in equations (1.1) to define the planes in this image. The equilateral triangular plane is given by the simplex equation, the plane heading back and up is given by the carbon equation, and the plane heading downwards is given by the nitrogen equation. The intersection of the carbon plane with the simplex is the line with endpoints c1=(0, 0.45, 0.55) and c2=(0.75, 0, 0.25), and the intersection of the nitrogen plane with the simplex is the line with endpoints n2=(0.63, 0, 0.37) and n3=(0.39, 0.61, 0). The solution polytope is the intersection of the three planes, the point labeled "solution" at (0.59, 0.10, 0.31).

a hyperplane is a "flat" subspace of any higher-dimensional space. The data plotted in the isotope ratio data space in Figure 1.1 translate into the three 2-dimensional planes shaded in Figure 3.1, one plane each for carbon, nitrogen, and the simplex. The solution is the intersection of these three hyperplanes.

Next we show how first the simplex and then the isotope ratio data map onto the solution space. The third row of (1.12) defines the simplex, as in (1.3), that the biomass proportions sum to 1, $1 = \pi_1 + \pi_2 + \pi_3$. Three obvious solutions to this equation are $\underset{\sim}{\pi} = [\pi_1, \pi_2, \pi_3]^\top = [1, 0, 0]^\top$, $[0, 1, 0]^\top$, and $[0, 0, 1]^\top$. These points define the extremes of the simplex hyperplane restricted to the unit cube, and will always be an $(S-1)$-dimensional equilateral triangular region. In Figure 3.1, the pink equilateral triangle is the hyperplane corresponding to the simplex. Similarly, the plane for carbon corresponding to the first row of (1.12) extends back and up, and the plane for nitrogen corresponding to the second row of (1.12) extends down. While these planes extend in space without limit, only the most illustrative fragment is plotted, with bold lines representing intersections with the simplex.

If only one isotope ratio is considered with three sources there are likely an infinite number of possible solutions, provided a solution exists. To illustrate this, consider only the carbon isotope ratio, represented by the horizontal axis in isotope ratio data space in Figure 1.1 observing the small points along the horizontal axis. Abbreviate the mixture and sources by their first initial: brown bear by B, and salmon, meat, and fruit by S, M, and F. The mixture B and sources (S, M, F) are ordered from low to high by isotope ratio values of carbon as F, B, S, M. It is clear that if we remove M, by setting its proportion contribution to $\pi_2 = 0$, that B can be represented as a unique weighted average of F and S. This is solved by using the two-equation system of the simplex, $1 = \pi_1 + 0 + \pi_3$, and the carbon equation, $-20.3 = -19.3\pi_1 - 16.6(0) - 23.3\pi_3$, and solving for the two-unknown proportions giving $\pi_1 = 0.75$ and $\pi_3 = 0.25$. In Figure 3.1 this solution for carbon when $\pi_2 = 0$ is labeled c2 $= [0.75, 0.00, 0.25]^\top$. Similarly, if we remove S, by setting its proportion

contribution to $\pi_1 = 0$, then B can be represented as a unique weighted average of F and M by $1 = 0 + \pi_2 + \pi_3$ and $-20.3 = -19.3(0) - 16.6\pi_2 - 23.3\pi_3$ giving $\pi_2 = 0.45$ and $\pi_3 = 0.55$. In Figure 3.1, c3 $= [0.00, 0.45, 0.55]^\top$. However, again referring to the horizontal axis in Figure 1.1, if F is removed, by setting its proportion contribution to $\pi_3 = 0$, B falls outside the interval of S and M, and so can not be represented as a weighted average of S and M where both π_1 and π_2 are between 0 and 1, $0 \leq \pi_s \leq 1, s = 1, 2$. When all three sources are considered together, possible solutions are weighted averages of these two unique solutions. The line connecting endpoints c2 and c3 in Figure 3.1 indicates the infinite number of solutions given by the intersection of the simplex and carbon hyperplanes. If carbon was considered alone there would not be a unique solution, but the solutions would be all the points along this line.

Similarly, the intersection of the nitrogen and simplex hyperplanes can be determined as a line segment with endpoints n3 $= [0.39, 0.61, 0.00]^\top$ and n2 $= [0.63, 0.00, 0.37]^\top$.

When carbon and nitrogen are considered together with the simplex then a unique solution exists, because these three hyperplanes intersect at a point, $\underset{\sim}{\pi} = [0.59, 0.10, 0.31]^\top$. Equivalently, in this case there is a single solution vector $\underset{\sim}{\pi}$ satisfying (1.12), represented geometrically as the intersection point in Figure 3.1.

In principle, this method can find solutions in larger cases, when $S > I + 1$. But the process can be automated in a way that does not rely on pictures such as Figure 3.1.

3.1.1 Underconstrained case, mink

We use mink data from Ben-David et al. (1997) to compare the qualitative results of SISUS and IsoSource. The mink data are discussed more completely in Section 5.4, and our results differ from the analyses in Phillips (2001, Table 3) and Phillips and Gregg (2003, Figure 4) because of discrepancies in their summarized values. For

the BMM in (1.1) the mink data are summarized in (3.1) where the rows are for carbon ($\delta^{13}C + \Delta^{13}C$), nitrogen ($\delta^{15}N + \Delta^{15}N$), and the simplex. The vector $\underset{\sim}{\beta}_{\mathrm{x}}$ is the mixture mink and the columns of matrix $\mathbf{A}_{\mathrm{x}}$ are for the $S = 7$ sources of fish, mussels, crabs, shrimps, rodents, amphipods, and ducks,

$$\begin{bmatrix} -15.11 \\ 13.81 \\ 1 \end{bmatrix} = \begin{bmatrix} -14.23 & -18.51 & -15.28 & -16.90 & -24.61 & -18.69 & -21.38 \\ 14.68 & 10.74 & 12.20 & 12.96 & 10.07 & 15.00 & 14.92 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_7 \end{bmatrix} \quad (3.1)$$

The plot in Figure 3.2 shows the convex hull of the sources encompassing the mink. Because the mink is in the convex hull, there are feasible solutions. We can expect high proportional contributions of fish and crabs because of the proximity of mink to those sources on the boundary.

3.2 Algorithms for feasible solutions

The goal is to represent the **solution polytope**, the feasible solutions defined as the set of all probability vectors $\underset{\sim}{\pi}$ satisfying the defining (1.4), $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$. This applies to the BMM but also the EMM through the equivalent representation (1.7). Equivalently, the solution polytope is defined as the intersection of the simplex and isotopic hyperplanes, as previously illustrated in Figure 3.1. Each additional independent linear constraint reduces the solution polytope by one dimension, where the simplex and each isotopic equation are linear constraints.

A favorable property of the solution polytope is that it is convex, that is, a line connecting any two points in the convex polytope will also be entirely inside the polytope. Descriptively, "convex" indicates the polytope is like a cut diamond rather than a star. The geometric fact used is that the intersection of convex objects, such

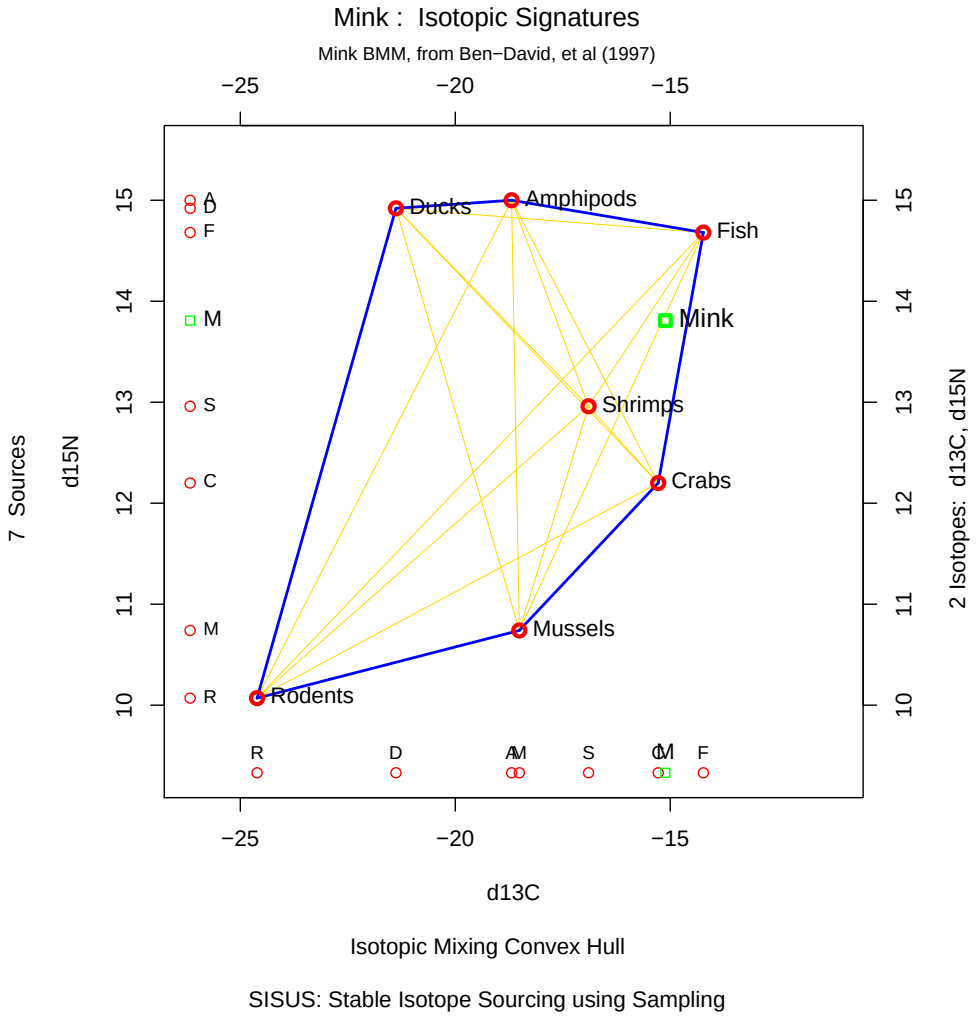

Figure 3.2: Convex hull plot of the mink example using (3.1).

as the proportion hypercube with simplex and isotopic hyperplanes, is necessarily convex.

In the Alaskan bear example, if carbon is considered alone with the simplex there are an infinite number of solutions given geometrically as the intersection of the carbon and simplex hyperplanes. We will use this carbon-only situation to illustrate the ideas often applied to larger examples, such as the mink data, when many more sources are considered relative to isotopes measured, $S > I + 1$.

3.2.1 Deterministic approximate solutions via IsoSource

IsoSource is a popular deterministic algorithm used in stable isotope sourcing to represent the solution polytope from underconstrained linear mixing models where $S > I + 1$ (Phillips and Gregg, 2003; Phillips, 2006). IsoSource evaluates a user-specified uniformly-spaced grid of points on the simplex, labeling a point a solution if it satisfies the linear system in (1.4) within a user-specified tolerance (see Figure 3.3a). These are points on or close to the solution polytope, consistent with all possible solutions being equally likely *a priori*. This is a brute force strategy because no information is used regarding the location of the solution polytope within the simplex. For a fixed tolerance, decreasing the increment of the grid space hyperexponentially increases the number of points evaluated, increasing both the number of solutions returned and the time for the algorithm to execute. For a fixed increment, decreasing the tolerance increases the accuracy of the solutions by excluding points far from the solution polytope. Because the number of approximate solutions depends on the dimension and sub-volume of the solution polytope, the increment grid spacing, and the solution tolerance, it may be challenging to choose settings to balance the desire for many solutions, accurate solutions, and acceptable execution time.

3.2.2 Probabilistic exact solutions via SISUS

Recall, the goal is to represent the infinite number of solutions in the solution polytope with a finite sample of solutions. If solutions are equally likely, a representation of the solution polytope can be given by a sample where points from the solution polytope are drawn with equal probability. If the solution polytope was always a simple object, like a line, then this sampling would be straightforward. Because the solution polytope can be quite complicated when the number of sources S is large, like an asymmetrically cut diamond, a direct sampling strategy may not be available. In the following section we describe a probabilistic algorithm for providing a user-specified number of exact solutions drawn uniformly from the solution polytope regardless of the geometric complexity of the solution polytope. The algorithm consists of two steps: first the vertices of the solution polytope are determined so the exact location of the solution polytope is known (Fukuda and Prodon, 1996), then probabilistic sampling is performed strictly from the polytope (Meeden and Lazar, 2006).

3.2.2.1 Uniform sampling of the solution polytope

The first step is to determine the vertices (and, therefore, the boundaries) of the solution polytope. The method is complex to describe (Fukuda and Prodon, 1996), but implemented in Fukuda (2005). The important point is that because the location of the solution polytope is known, we can sample directly from the set of exact solutions efficiently.

The second step is the probabilistic sampling from the solution polytope. Here we describe some technical details regarding how we propose sampling uniformly from the solution polytope using a Markov chain Monte Carlo (MCMC) algorithm. The random directions symmetric mixing algorithm has three steps (Boneh and Golan, 1979; Smith, 1980). (0) Choose a starting point inside the solution polytope,

$\underset{\sim}{\pi}^{(0)} \in \mathcal{P}$, and set counter $r = 0$. In practice, this is an interior point of the polytope after a short burn-in of the algorithm. (1) This step does all the work: Generate a random direction d uniformly distributed over a direction set inside the solution polytope $D \subseteq \Re^S$, find the line set $L = \mathcal{P} \cap \{\underset{\sim}{\pi}^{(r)} + ld, l \in \Re\}$, and generate a random point $\underset{\sim}{\pi}^{(r+1)}$ uniformly distributed over L. Descriptively, we draw a line segment through the current point $\underset{\sim}{\pi}^{(r)}$ to the edges of the polytope along that direction and in the opposite direction, then generate the next point $\underset{\sim}{\pi}^{(r+1)}$ uniformly at random from that line segment. In this way we move around the polytope collecting a representative sample. (2) If we have the desired number of samples, $r = R$, then stop. Otherwise, increment the counter $r \leftarrow r + 1$ and return to step (1).

The sample generated converges to the uniform distribution over an arbitrary polytope, at worst at an exponentially fast rate, which would be slow (Smith, 1984). However, our convex polytope is a best-case scenario converging almost immediately with small samples.

SISUS uses the implementation of the random directions symmetric mixing algorithm in R package `polyapost`, function `constrppprob` (R Development Core Team, 2006; Lazar and Meeden, 2003; Meeden and Lazar, 2006). The number of sources S and isotopes I may both be large, provided $S \geq I + 1$. A representation of the solution space is quickly available with small sample sizes and running time changes roughly linearly as a function of the number of sources S, later shown in Table 3.1. Sample sizes of $R = 1000$ and 10000 appear reasonable for exploration and publication, respectively. Standard MCMC diagnostics are used to monitor the quality of convergence. The `sisus` R package may be downloaded from CRAN or run online at http://StatAcumen.com/sisus.

3.2.2.2 MCMC convergence diagnostics

Convergence diagnostics are a standard way of determining the quality of a sample drawn using an MCMC procedure, a number of which are implemented in SISUS via CODA (Plummer et al., 2006). SISUS uses the diagnostics to evaluate the sample and provides a global pass/fail status, with specific suggestions if the sample fails any of the diagnostics. Gelman et al. (1995, sec. 11.11) provide a short history regarding simulation and monitoring convergence.

A few diagnostics implemented are trace plots, autocorrelations, Geweke's convergence diagnostic, Heidelberger and Welch's convergence diagnostic, and Raftery and Lewis's diagnostic. Geweke's convergence diagnostic, as implemented, tests whether the mean of first 0.10 of samples is different from the last 0.50, and suggest a minimum length of burn-in to make these two means not statistically different (Geweke, 1992). Heidelberger and Welch's convergence diagnostic uses the Cramer-von-Mises statistic to test whether the sampled values come from a stationary distribution (Heidelberger and Welch, 1981, 1983). Raftery and Lewis's diagnostic determines the sample size required for estimating quantiles with a specified precision (Raftery and Lewis, 1992, 1995).

3.2.3 Illustration of IsoSource and SISUS solution sampling methods

We analyze the Alaskan bear example using the EMM for the carbon isotope only, that is, excluding the nitrogen information. For the EMM using carbon only, (1.9)

defines each element of $\mathbf{B}$, so that $[\underset{\sim}{0}^\top, 1]^\top = \mathbf{B}_{\mathrm{x}}\underset{\sim}{\pi}$ in (1.7) is

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.548 & 1.906 & -0.856 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}, \tag{3.2}$$

where, for example, $1.906 = ((-21.5 + 4.9) + 20.3)(0.515)(1)$. SISUS provides solutions to the EMM directly, while (though undocumented) IsoSource can provide solutions to the EMM by inputting into the software a transposed $\underset{\sim}{\beta}$ vector in the mixture fields and a transposed $\mathbf{A}$ matrix in the sources fields. Figure 3.3 shows the simplex and carbon hyperplanes and their intersection solution polytope as a line. Note that the carbon hyperplane intersects the simplex at a different line than the BMM example in Figure 3.1 due to the concentration and assimilation information in the EMM. Also, the carbon hyperplane now passes through the origin, due to the centering in (1.9). Figure 3.3a shows an example of the IsoSource deterministic sampling strategy, with a grid increment of 0.02 and tolerance of 0.1 where 117 of the 1326 points evaluated are determined approximate solutions. Notice that the points evaluated are uniform over the simplex, but the approximate solutions provided are only roughly uniform near the solution polytope. Figure 3.3b shows one realization of $R = 117$ exact probabilistic solutions from SISUS which are qualitatively similar to the deterministic approximate solutions of IsoSource, yet algorithmic advantages are clear. The exact probabilistic solutions from SISUS converge quickly to a uniform distribution over the solution polytope.

3.2.4 Execution time, solution predictability, and solution accuracy

There are several reasons why the probabilistic approach (SISUS) is preferred over the deterministic approach (IsoSource). These are (1) the relatively short execution

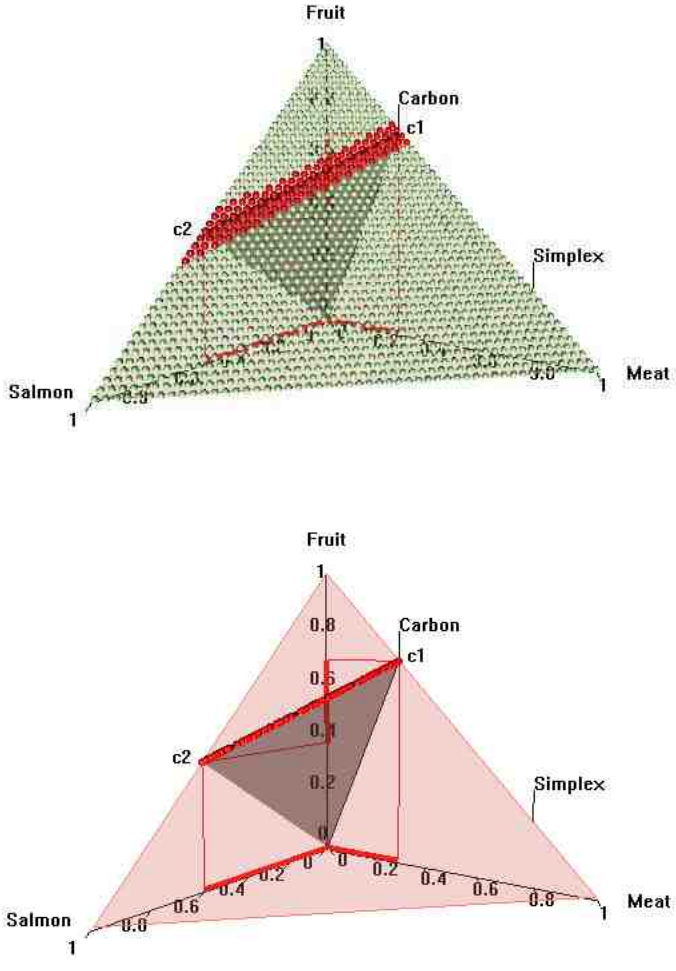

Figure 3.3: EMM Alaskan bear example using carbon (and not nitrogen) finding solutions for (3.2). (a) IsoSource evaluates points on the lattice over the simplex, returning 117 approximate solutions of the 1326 evaluated points. (b) SISUS samples 117 exact solutions uniformly over the solution polytope.

time, (2) the predictability of the number of solutions, and (3) the solution accuracy. Below we use a fabricated example to illustrate the differences in these approaches.

To compare the results of the probabilistic SISUS algorithm to the deterministic IsoSource algorithm, we use subsets of the problem shown in (3.3). The values in this matrix were chosen to provide solutions to the problem sizes we study, but are otherwise essentially random within a range of integers centered at 0. The full problem has $S = 10$ sources and $I = 5$ isotope ratios, as that is the extent of the problem size that IsoSource is programmed to solve. For each example, Table 3.1 specifies a given value for S and I, and the problem is defined by choosing the first S columns and first I rows plus the simplex constraint of (3.3). For example, the subset where $S = 5$ and $I = 2$ uses the first five columns of the matrix (ignoring the last five columns) and uses the first two rows and the last row for the simplex (ignoring the third through fifth rows). Table 3.1 reports the results of an execution time comparison between SISUS and IsoSource running on a PC (Dell Optiplex GX260 with Intel Pentium 4 2.40GHz CPU with 512 MB RAM) without any additional significant processes running.

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & -2 & 2 & 0 & 5 & -6 & 3 & 1 \\ 2 & -1 & 1 & -2 & 0 & 3 & -4 & 2 & 5 & -6 \\ -3 & -1 & 1 & 2 & 3 & -4 & 0 & 5 & -2 & 7 \\ 4 & 2 & -1 & 1 & -2 & -3 & -4 & 6 & 0 & 5 \\ -5 & 1 & 3 & 4 & -5 & 0 & 2 & -2 & 6 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \vdots \\ \pi_{10} \end{bmatrix} \quad (3.3)$$

For SISUS we obtain R exact solutions for each problem size by finding $100 \times R$ solutions and keeping every 100th to increase the independence of the samples (not currently important) and to improve solution polytope coverage (helpful only for small samples).

For IsoSource the number of iterations is given by the choose formula

$$it = \binom{100/inc + S - 1}{S - 1},$$

where it is the number of iterations necessary to cover the simplex at increments of inc in percent values for S sources (Phillips and Gregg, 2003, eq. 3). Observe the hyperexponential growth in iterations in Table 3.1 when either S or inc increases.

In Table 3.1 we make the following observations. The time for SISUS increases up to a few minutes while IsoSource increases to hours. For SISUS a specified number of exact solutions are requested and the time to obtain those solutions grows nearly linearly with the number of sources S and quadratically with R due to time for memory allocation. For IsoSource an increment and tolerance are specified which determines a hyperexponentially growing number of iterations with the number of sources, then approximate solutions are returned which are within the tolerance of an exact solutions. Time scales roughly linearly with the number of iterations. For SISUS, the computation time for $R = 10000$ exact solutions triples from 10 seconds to 30 seconds, increasing roughly linearly with $S = 3$ to 10. Not shown in the table, SISUS provides exact solutions for each of these cases in 2 seconds or less when $R = 1000$. For $R = 20000$ and 30000 exact solutions the times nearly triple from 45 to 120 and 100 to 230 seconds over $S = 3$ to 10. The time reported here is only the time for sampling, and additional time is used to read the data and create plots. On this computer we obtain roughly one million iterations per second for IsoSource, so we can predict the length of time for IsoSource's very large cases. For small problems ($S = 5$), IsoSource takes only a few seconds. But as the problem size increases, the time jumps to hours. A few cases to note: (1) For $S = 5$, as I increases ($I = 1, 2, 3$) and the solution polytope decreases in dimension ($d = 3, 2, 1$) the solutions become more difficult to find for IsoSource resulting in fewer or no approximate solutions. (2) For $S = 8$, $I = 5$, $inc = 2$ with a tolerance of $tol = 0.01$ finds no solutions after almost 6 minutes of computation, but increasing the tolerance to a very wide

margin of 0.10 provides over 3000 approximate solutions. Therefore, the solution polytope is not close enough to the simplex grid to find solutions with an increment of 2 with tolerance 0.01. (3) Further, for $S = 8$, $I = 5$, increasing the increment to 1 with a tolerance of 0.01 finds only 19 approximate solutions after over 8 hours of computation. (4) The largest example, $S = 10$, $I = 5$, $inc = 2$, finds only 1 approximate solution after almost 5 hours of computation, and larger increments fail to find any approximate solutions. Note that setting the increment to 1 will increase the running time to roughly 68 days. In each of these cases, SISUS provides 10000 exact solutions in less than 30 seconds.

Obtaining a predictable number of solutions is valuable for reliable analysis. The initial step in SISUS of identifying the boundaries of the solution polytope allows quick sampling of a selected number of exact solutions. Because IsoSource does not know where the solution polytope is located in the simplex, it iterates over the entire simplex returning approximate solutions when within the tolerance. Thus, the number of approximate solutions is unpredictable through a complex relationship of S, I, inc, and tol, and the model equations. This point is clearly shown in the IsoSource sol column of Table 3.1. Obtaining a large number of solutions is not an indication of the quality of the solutions, though having only a few solutions is unlikely to represent the solution polytope well.

Raftery and Lewis's diagnostic (Raftery and Lewis, 1992) is used in SISUS as a quality check of the MCMC sampling by determining the number of samples required to estimate a quantile of the marginal distribution of the solution polytope to within a given relative accuracy with a prespecified probability. Note that because the solution polytope is bounded and we know the extremes exactly, the number of suggested solutions will be overestimated by this diagnostic. For the mink example, where $S = 7$ and $I = 2$, the disgnostic recommends $R = 1522$ samples to estimate the first quantile, $q = 0.01$, with an accuracy of half a percentage point with probability 0.95. That is, in 95% of samples of size 1522 we expect the estimate of the first

quantile to be within half a percent relative to the true quantile, with accuracy increasing with sample size.

Table 3.1: Comparison of execution time and number of solutions for SISUS and IsoSource. Column labels are number of sources S, number of isotopes I, number of SISUS solutions R, and time in seconds. IsoSource parameters are increment *inc*, tolerance *tol*, number of iterations *it*, number of solutions, and time in seconds.

Size		SISUS		IsoSource				
S	I	sol (R)	time (s)	*inc*	*tol*	*it*	sol	time (s)
5	1	10000	13	2	0.01	316251	4023	1
5	1	20000	44	1	0.01	4598126	89726	4
5	2	10000	13	2	0.01	316251	45	1
5	2	20000	45	1	0.01	4598126	1535	4
5	3	10000	12	2	0.01	316251	0	1
5	3	30000	99	1	0.01	4598126	18	4
7	2	30000	150	1	0.01	1705904746	265021	(27 m) 1631
8	2	10000	22	5	0.01	888030	424	1
8	2	20000	79	2	0.01	264385836	24162	344
8	2	30000	158	1	0.01	26075972546	5747298	(8.2 h)
8	5	10000	20	5	0.01	888030	0	1
8	5	20000	77	2	0.01	264385836	0	344
8	5			2	**0.10**	264385836	3167	344
8	5	30000	177	1	0.01	26075972546	**19**	(8.2 h) **29488**
10	2	10000	26	10	0.01	92378	103	1
10	2	20000	90	5	0.01	10015005	3455	18
10	2	30000	(4 m) 228	2	0.01	12565671261	850563	(5 h) 17990
10	5	10000	25	10	0.01	92378	0	1
10	5	20000	90	5	0.01	10015005	0	18
10	5	30000	203	2	0.01	12565671261	**1**	(4.8 h) **17373**
10	5			1	0.01	4263421511271	?	(68 days)

3.3 Results

3.3.1 Accuracy of SISUS and IsoSource

We use the mink data in Table 5.7 to illustrate that while the results using the probabilistic method of SISUS and the deterministic method of IsoSource are qualitatively similar, there are some important differences that favor the use of the probabilistic method.

Table 3.2 compares the summaries returned by both SISUS and IsoSource. SISUS summaries are from 10000 exact solutions, and using 30000 gives the same results to the precision presented. IsoSource summaries are from 790 approximate solutions using an increment of 1% (the finest resolution allowed) and a tolerance of 0.01. Time to compute the feasible solutions are similar to those given in Table 3.1 for $S = 7$ sources and $I = 2$ isotopes, with SISUS running one-tenth of the time as IsoSource. Minimum, maximum, mean, and standard deviation of feasible source contributions are compared. The minimum and maximum bounds reported by SISUS are exact based on the vertices of the solution polytope. The extremes reported by IsoSource differ because they are approximate. The mean values estimate the center of the solution polytope and are consistent between methods. The standard deviation values of IsoSource are larger than for SISUS. While not obvious from the results here, because IsoSource uses a tolerance around the mixture values, the degree of error associated with each source's proportion values may be different. Also, the effect on the results is minor for centering the values before using IsoSource, as in (1.9), as compared to the results reported using (1.1).

To visualize the solution polytope, projections into one-dimensional lines and two-dimensional planes are made. The samples provide an empirical representation of the solution polytope. If, for example, all points of the solution polytope are assumed *a priori* equally likely, consistent with a uniform distribution over the solution polytope,

Table 3.2: Selected numerical summaries for the mink example based on SISUS and IsoSource methods. Values represent feasible source contributions of biomass to mink.

	SISUS				IsoSource			
Sources	Min	Max	Mean	SD	Min	Max	Mean	SD
Fish	0.52	0.76	0.630	0.038	0.53	0.77	0.626	0.046
Mussels	0.00	0.20	0.040	0.033	0.00	0.20	0.044	0.039
Crabs	0.04	0.37	0.230	0.051	0.03	0.36	0.224	0.062
Shrimps	0.00	0.27	0.054	0.044	0.00	0.25	0.061	0.053
Rodents	0.00	0.06	0.012	0.010	0.00	0.06	0.010	0.012
Amphipods	0.00	0.11	0.022	0.018	0.00	0.10	0.022	0.022
Ducks	0.00	0.07	0.014	0.012	0.00	0.07	0.013	0.014

then the projected one- and two-dimensional histograms approximate the marginal densities of the solution polytope. The solution polytope is an $S-I-1 = 7-2-1 = 4$-dimensional object in $S = 7$-dimensional space. The solution scatterplot using SISUS in Figure 3.4 is comparable to Phillips and Gregg (2003, Figure 5) using IsoSource. The diagonal has 1-dimensional projections shown as marginal histograms of feasible source contributions. The off-diagonals show 2-dimensional projections into planes of paired sources. The upper triangle includes points for each of the 10000 exact solutions of SISUS. Notice that the regions are triangular and reach to the extent of the triangles, while in Phillips and Gregg (2003, Figure 5) the entire solution polytope is not revealed (in particular Crab/Shrimp). Finding solutions at the extremes of the solution polytope appears difficult for IsoSource. The lower triangle includes two-dimensional density histograms to show where the majority of the points are.

3.4 Discussion

The ability to quickly sample a user-specified number of exact solutions accurately representing the solution polytope favors the probabilistic method, which SISUS uses. This approach provides a probabilistic guarantee that the solution polytope will be uniformly represented by a sample of any specified size. The validity of this method for inference is subject to the model assumptions provided in Section 1.2.1.3.

3.4.1 Probabilistic interpretation

It is natural to want to interpret the source proportion solutions probabilistically. The purpose of chapters 5 and 6 of this dissertation is to consider probabilistic methods. With regards to what we have discussed so far, if a proportion p of the solution polytope is between bounds a and b for source s, then we may want to infer that with probability p the contribution of source s is between a and b. This

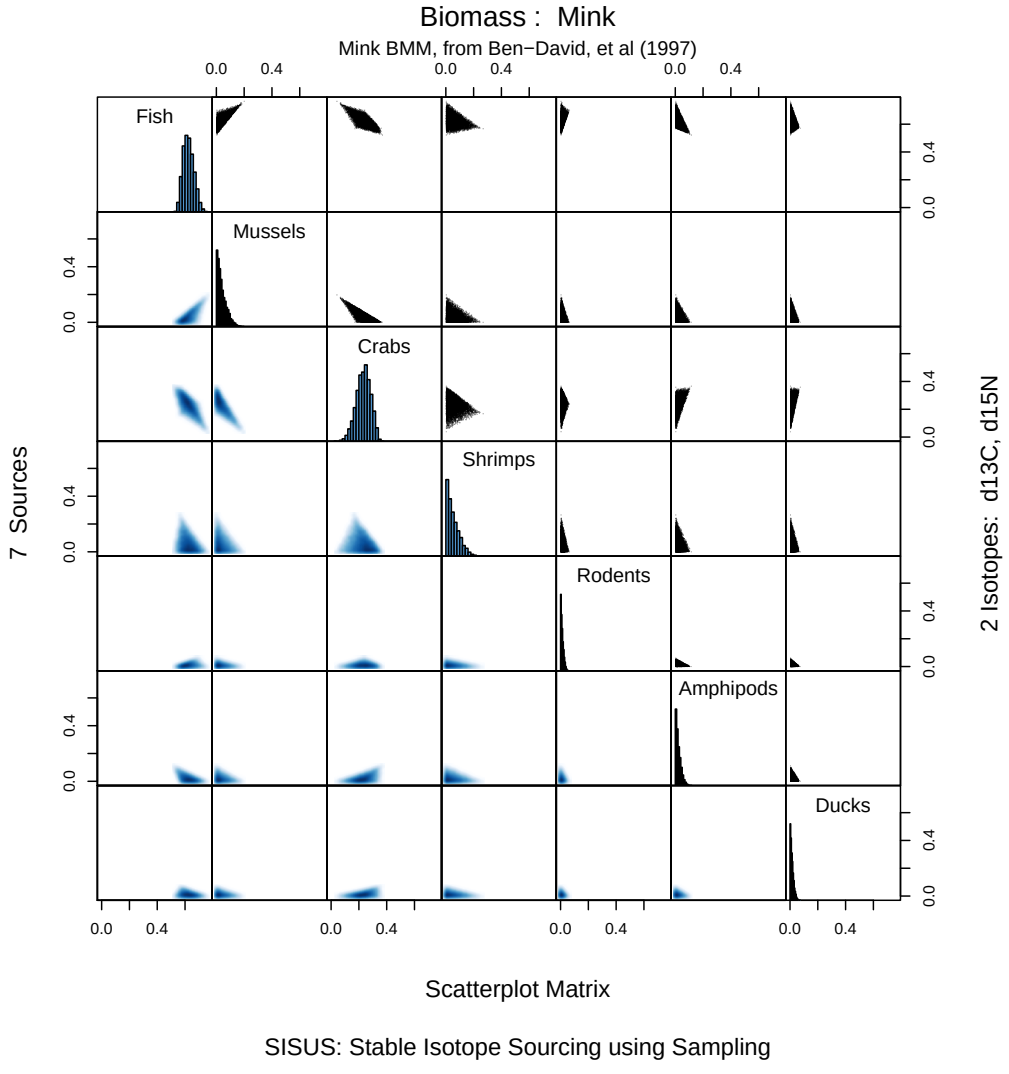

Figure 3.4: Mink BMM example marginal histograms along diagonal, scatterplot of paired source contributions on the upper diagonal, and corresponding two-dimensional density histograms on the lower diagonal. These are equivalently one- and two-dimensional empirical projections of the solution polytope.

interpretation is consistent with the model $\underset{\sim}{\beta}_x = \mathbf{A}_x\underset{\sim}{\pi}$ given in (1.4) with the following two additional assumptions. First, the isotope ratio, discrimination, concentration, and assimilation efficiency parameters are estimated without measurement error and not subject to uncertainty, Thus, the only solutions are those defined by the solution polytope. Second, all proportion values $\underset{\sim}{\pi}$ are equally likely *a priori*, corresponding to a uniform distribution over the simplex. This means that before observing $\underset{\sim}{\beta}$ and $\mathbf{A}$ any set of proportions summing to 1 is as likely as any other. Given the data, the intersection of the linear constraints defines the solution polytope, and this region of feasible solutions inherits the uniform property assumed on the simplex. It is consistent with these model assumptions that we can interpret the solutions probabilistically. This heuristic argument will be made more rigorous in Section 5.2.1.

An obvious point is that it is unreasonable to assume the values of $\underset{\sim}{\beta}$ and $\mathbf{A}$ are known. In practice, the values of $\underset{\sim}{\beta}$ and $\mathbf{A}$ are sums and products of means estimated with error. It is the Bayesian probabilistic method of obtaining feasible solutions that will allow us to implement models including variation and uncertainty in the underconstrained case. The following chapters model these sources of variation to remedy this point.

Chapter 4

Frequentist methods, BMM

Frequentist methods in this chapter apply to the perfectly constrained case ($S = I+1$) of the implicit representation model. We develop the model by starting with the defining equation (1.1) and sequentially adding variation to model componenets. We generalize the delta method-based model by Phillips and Gregg (2001) and provide an easy-to-implement method using the implicit function theorem. We also introduce the use of the bootstrap for stable isotope sourcing problems. Further, estimation procedures of discrimination-corrected isotope ratios using regression is developed. Results are demonstrated on two- and three-source datasets.

The BMM is $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$ where the columns of $\mathbf{A}$ are the discrimination-corrected source isotope ratios. For convenience in the frequentist framework $\underset{\sim}{\beta}$ and $\mathbf{A}$ are appended with a row of $\underset{\sim}{1}^{\top}$ to incorporate the simplex condition, denoted $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$. Phillips and Gregg (2001) apply asymptotic results to calculate approximate variances and confidence intervals for source proportions to account for variability in the isotope ratios of the sources and the mixture. These measures of uncertainty can be used in cases where there is a unique solution, when matrix $\mathbf{A}_{\mathrm{x}}$ is square and full rank with S sources and $I = S - 1$ isotopes. They derive the large sample solutions for two cases, $S = 2$ and $S = 3$, and provide an Excel workbook to perform

the calculations. Their solution can incorporate correlation within the isotope ratios of a source or mixture. We generalize their work and compare the results between frequentist and bootstrap methods.

4.1 Estimation

For the BMM all model components are derived from isotope ratio measurements for mixtures, sources, and diet experiments, where diet experiments give diet and tissue measurements used to estimate discrimination. We use the multivariate method of moment estimates as described in Christensen (2001, Ch. 1) to estimate parameters of the model for isotope ratios and various methods for discriminations. While natural variations in C and N isotope ratios are generated by different biochemical and ecological processes, and Phillips and Koch (2002) claim that an assumption of independence of isotope ratios is probably valid in most cases, Figure 4.2 shows a clear example of correlation. We model correlation between isotopes.

4.1.1 Mixture isotope ratios

For a sample of mixtures, $j = 1, \ldots, J$, in the mixture population we observe isotope ratio measurements, $\underset{\sim}{b}_j$. Commonly, only one measurement per individual is made, but this work can be generalized to include multiple measurements on each individual. We use the one-sample multivariate model to estimate the mixture population mean isotope ratio $\underset{\sim}{\beta}$ as described in Christensen (2001, sec. 1.3). Let the sample of

J observed mixture isotope ratios be

$$\begin{bmatrix} \underset{\sim}{b}_1^\top \\ \underset{\sim}{b}_2^\top \\ \vdots \\ \underset{\sim}{b}_J^\top \end{bmatrix} = \begin{bmatrix} b_{11} & \cdots & b_{1I} \\ b_{21} & \cdots & b_{2I} \\ \vdots & \vdots & \vdots \\ b_{J1} & \cdots & b_{JI} \end{bmatrix}, \tag{4.1}$$

so that each row is a distinct sample of I isotope ratios. Population mean $\underset{\sim}{\beta}$ is estimated with sample mean $\bar{\underset{\sim}{b}}$ where

$$\bar{\underset{\sim}{b}}^\top = \frac{1}{J}\sum_{j=1}^{J} \underset{\sim}{b}_j^\top = [\bar{b}_1, \cdots, \bar{b}_I], \tag{4.2}$$

and population covariance $\mathrm{Cov}[\underset{\sim}{b}_j] = \Sigma_{\underset{\sim}{b}}$ is estimated with

$$\mathbf{S}_{\underset{\sim}{b}} = \frac{1}{J-1}\sum_{j=1}^{J} (\underset{\sim}{b}_j - \bar{\underset{\sim}{b}})(\underset{\sim}{b}_j - \bar{\underset{\sim}{b}})^\top. \tag{4.3}$$

Also, let $\mathbf{S}_{\bar{\underset{\sim}{b}}} = \mathbf{S}_{\underset{\sim}{b}}/J$.

Relative to the implicit representation discussed in Section 2.2.1, we are marginalizing out the subject-specific isotope ratio vectors $\underset{\sim}{\beta}_j$. That is, if $\mathrm{E}[\underset{\sim}{\beta}_j] = \underset{\sim}{\beta}$ and $\mathrm{Cov}[\underset{\sim}{\beta}_j] = \Sigma_{\underset{\sim}{\beta}}$, then $\underset{\sim}{b}_j = \underset{\sim}{\beta}_j + \underset{\sim}{\varepsilon}_j$ has $\mathrm{E}[\underset{\sim}{b}_j] = \underset{\sim}{\beta}$ and $\mathrm{Cov}[\underset{\sim}{b}_j] = \Sigma_{\underset{\sim}{\beta}} + \Sigma_{\underset{\sim}{\varepsilon}} \equiv \Sigma_{\underset{\sim}{b}}$ when $\underset{\sim}{\varepsilon}_j$ has mean zero and is independent of $\underset{\sim}{\beta}_j$. Assuming $\underset{\sim}{\beta}_j$ and $\underset{\sim}{\varepsilon}_j$ are normally distributed would lead to a normal distribution for the $\underset{\sim}{b}_j$s.

4.1.2 Source isotope ratios

Estimating source isotope ratios, $\underset{\sim}{\delta}_s$, is the same as for mixtures, and each observed source individual is likely to have only a single measurement. For each independent

source, $s = 1, \ldots, S$, we observe a sample of isotope ratio measurements, $\underset{\sim}{d}_{sk}$, $k = 1, \ldots, K_s$. Let the K_s observed isotope ratios for source s be

$$\begin{bmatrix} \underset{\sim}{d}_{s1}^{\top} \\ \underset{\sim}{d}_{s2}^{\top} \\ \vdots \\ \underset{\sim}{d}_{sK_s}^{\top} \end{bmatrix} = \begin{bmatrix} d_{1s1} & \cdots & d_{Is1} \\ d_{1s2} & \cdots & d_{Is2} \\ \vdots & \vdots & \vdots \\ d_{1sK_s} & \cdots & d_{IsK_s} \end{bmatrix} \tag{4.4}$$

with mean $\underset{\sim}{\delta}_s$ estimated with $\underset{\sim}{\bar{d}}_s$ where

$$\underset{\sim}{\bar{d}}_s^{\top} = \frac{1}{K_s} \sum_{k=1}^{K_s} \underset{\sim}{d}_{sk}^{\top} = [\bar{d}_{1s}, \cdots, \bar{d}_{Is}], \tag{4.5}$$

and population covariance $\mathrm{Cov}[\underset{\sim}{d}_s] = \Sigma_{\underset{\sim}{d}_s}$ is estimated with

$$\mathbf{S}_{\underset{\sim}{d}_s} = \frac{1}{K_s - 1} \sum_{k=1}^{K_s} \left(\underset{\sim}{d}_{sk} - \underset{\sim}{\bar{d}}_s \right) \left(\underset{\sim}{d}_{sk} - \underset{\sim}{\bar{d}}_s \right)^{\top}. \tag{4.6}$$

Also, let $\mathbf{S}_{\underset{\sim}{\bar{d}}_s} = \mathbf{S}_{\underset{\sim}{d}_s} / K_s$.

4.1.3 Discrimination

Discrimination, $\underset{\sim}{\Delta}_s$, is an offset reflecting the difference in isotope ratio of the source and the resulting isotope ratio in the mixture due to the digestive process. Discriminations can be derived for each source separately or derived for a set of sources simultaneously via regression. The strategy in both cases is to feed a captive sample of consumers (mixtures) a single-source diet until the consumer's tissues of interest reaches equilibrium, then the isotope ratio difference of the diet and the consumer's tissues can be compared. If this is done for a range of diets, then a regression can be applied to the range of diets. Caut et al. (2009) recommend the use of a

diet-dependent discrimination factor (DDDF) because of the differences in discrimination for $\Delta^{13}C$ and $\Delta^{15}N$ among different taxonomic classes of animals, among tissue types, and among diets. Using DDDFs can make large differences ($> 2‰$) in discrimination for carbon and nitrogen compared to using a discrimination calculated from a different species or tissue, or applying a common discrimination for all sources.

In the single-diet case, the discrimination is the isotope ratio difference between the mixture tissue (T) and the source diet (D), $\underset{\sim}{\Delta}_s = \underset{\sim}{\delta}_{s\mathrm{T}} - \underset{\sim}{\delta}_{s\mathrm{D}}$. As in Section 4.1.2, these quantities are estimated by

$$\begin{aligned} \bar{\underset{\sim}{d}}_{s\mathrm{T}} &= \frac{1}{K_{s\mathrm{T}}} \sum_{k=1}^{K_{s\mathrm{T}}} \underset{\sim}{d}_{s\mathrm{T}k} \quad \text{and} \\ \bar{\underset{\sim}{d}}_{s\mathrm{D}} &= \frac{1}{K_{s\mathrm{D}}} \sum_{k=1}^{K_{s\mathrm{D}}} \underset{\sim}{d}_{s\mathrm{D}k} \end{aligned}$$

with estimated covariances

$$\begin{aligned} \mathbf{S}_{\underset{\sim}{d}_{s\mathrm{T}}} &= \frac{1}{K_{s\mathrm{T}} - 1} \sum_{k=1}^{K_{s\mathrm{T}}} \left(\underset{\sim}{d}_{s\mathrm{T}k} - \bar{\underset{\sim}{d}}_{s\mathrm{T}} \right) \left(\underset{\sim}{d}_{s\mathrm{T}k} - \bar{\underset{\sim}{d}}_{s\mathrm{T}} \right)^{\top} \quad \text{and} \\ \mathbf{S}_{\underset{\sim}{d}_{s\mathrm{D}}} &= \frac{1}{K_{s\mathrm{D}} - 1} \sum_{k=1}^{K_{s\mathrm{D}}} \left(\underset{\sim}{d}_{s\mathrm{D}k} - \bar{\underset{\sim}{d}}_{s\mathrm{D}} \right) \left(\underset{\sim}{d}_{s\mathrm{D}k} - \bar{\underset{\sim}{d}}_{s\mathrm{D}} \right)^{\top}. \end{aligned}$$

Also, let $\mathbf{S}_{\bar{\underset{\sim}{d}}_{s\mathrm{T}}} = \mathbf{S}_{\underset{\sim}{d}_{s\mathrm{T}}}/K_{s\mathrm{T}}$ and $\mathbf{S}_{\bar{\underset{\sim}{d}}_{s\mathrm{D}}} = \mathbf{S}_{\underset{\sim}{d}_{s\mathrm{D}}}/K_{s\mathrm{D}}$. The data from the diet experiment provide estimated discrimination $\hat{\underset{\sim}{\Delta}}_s = \bar{\underset{\sim}{d}}_{s\mathrm{T}} - \bar{\underset{\sim}{d}}_{s\mathrm{D}}$ with estimated covariance $\mathbf{S}_{\hat{\underset{\sim}{\Delta}}_s} = \mathbf{S}_{\bar{\underset{\sim}{d}}_{s\mathrm{T}}} + \mathbf{S}_{\bar{\underset{\sim}{d}}_{s\mathrm{D}}}$. The method based on regression over a range of diets is illustrated as part of the bear example in Section 4.3.1.3.

4.1.4 Sources ensemble

The columns of $\mathbf{A}$ are estimated by plugging in the estimates for the source isotope ratios and discriminations. $\underset{\sim}{\delta}'_s = \underset{\sim}{\delta}_s + \underset{\sim}{\Delta}_s$ is estimated with $\hat{\underset{\sim}{\delta}}'_s = \bar{\underset{\sim}{d}}_s + \hat{\underset{\sim}{\Delta}}_s$, with

estimated covariance $\mathbf{S}_{\underset{\sim}{\hat{\delta}'}_s} = \mathbf{S}_{\underset{\sim}{\bar{d}}_s} + \mathbf{S}_{\underset{\approx}{\hat{\Delta}}_s}$. The estimates of $\mathbf{A}_{\mathrm{x}}$ and $\underset{\sim}{\beta}_{\mathrm{x}}$ are therefore $\hat{\mathbf{A}}_{\mathrm{x}} = [\hat{\mathbf{A}}^{\top}, \underset{\sim}{1}]^{\top}$ and $\underset{\sim}{\bar{b}}_{\mathrm{x}} = [\underset{\sim}{\bar{b}}^{\top}, 1]^{\top}$.

4.1.5 Dunlin two-source example

Evans Ogden et al. (2005) quantify the proportional use that *Calidris alpina pacifica* (dunlin) made of terrestrial farmland vs. marine estuarine resources on the Fraser River Delta, British Columbia, using stable isotope analysis ($\delta^{13}C$, $\delta^{15}N$) of blood from 268 dunlin over four winters, 1997 through 2000. These shorebirds spend the winter in temperate areas and frequently use estuarine and supratidal (upland) feeding habitats (Figure 4.1). This has management implications since loss or reduction of agricultural habitat adjacent to estuaries may negatively impact shorebird fitness.

We analyze 234 dunlin (of the 268) captured between October and April over the winters of 1997–1998 through 1999–2000, and in November 2000, using mistnets and a floating clap net placed around the Fraser Delta (Figure 4.1). The data was provided by K. A. Hobson from Evans Ogden et al. (2005).

As an aside, there are two primary photosynthetic pathways in plants and these produce two distinct isotope ratio ranges (Cerling et al., 1997). The carbon isotopic composition of CO_2 in the atmosphere, which is what the plants use in photosynthesis, is about -6 to -7‰. Most plants are C3 plants, so-called because the first organic carbon compound made in photosynthesis contains three carbon atoms. The carbon isotopic fractionation of such C3 plants is very large, so that plant material from such plants (including rice, wheat, soybeans and potatoes) has $\delta^{13}C$ values between -33 and -23‰, with an average of about -26‰. By far most of the plants in the world are C3 plants. C3 plants must keep the openings in their leaves (stomata) wide open through which they take up CO2 during photosynthesis, which means that they lose water through these openings. Plants that live in dry regions (for which water loss is a problem) or in salty water (where loss of fresh water is a prob-

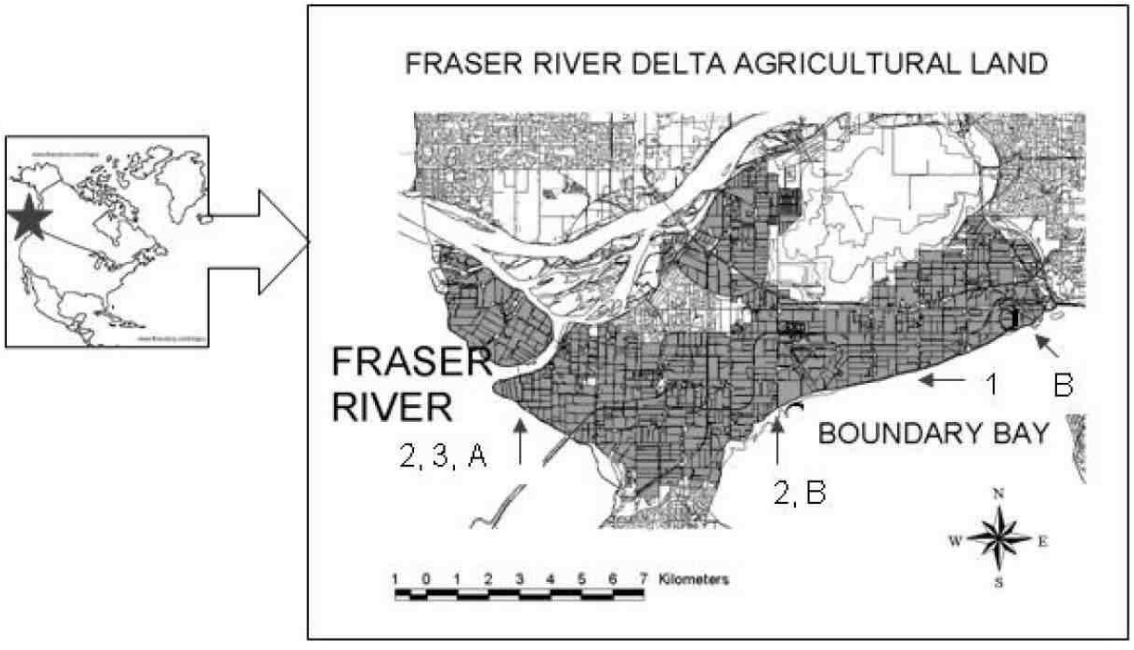

Figure 4.1: Agricultural lands in the Fraser River Delta (gray shading). Arrows indicate locations of dunlin capture and invertebrate sampling. Numbers indicate year of sampling for dunlin commencing 1997–1998. Letters indicate type of source invertebrates sampled: A = terrestrial invertebrates, B = estuarine invertebrates.

lem) have adapted by using a different pathway during the photosynthetic reaction, which allows them to open the stomata sparingly. This pathway is called the C4 reaction because the first organic compound formed contains four carbon atoms. C4 plants show less isotopic fractionation of carbon than C3 plants, and have $\delta^{13}C$ values ranging from -16 to -9‰, with an average about -13‰. Plants using the C4 reaction pathway include most tropical grasses and salt marsh grasses, and corn is an important C4 crop plant.

The relative contribution of terrestrial vs. marine-derived protein in the diets of birds has been estimated using stable-carbon ($\delta^{13}C$) and nitrogen ($\delta^{15}N$) isotope ratios because foodweb components differ significantly and predictably between marine and C3 terrestrial ecosystems (reviewed by Hobson, 1999). The evidence suggests that isotopic incorporation into body tissues does not vary with an organism's age (e.g., Minagawa and Wada, 1984). Differences in isotopic discrimination factors between individuals are small (e.g., Hobson and Clark, 1993), so individual differences in isotope ratio values can usually be ascribed to individual differences in diet. Dunlin were consistently observed flying into and feeding in terrestrial habitats during diurnal and nocturnal habitat surveys at high tide (Evans Ogden, 2002), and the few alternate terrestrial habitats available are not recognized as frequently used by dunlin, e.g., suburban and industrial areas, wooded areas, a golf course, and a bog (Butler and Campbell, 1987; Shepherd, 2001). Thus, we assumed that terrestrial isotope ratio values reflected diet assimilated via feeding in agricultural fields. Our isotopic model was based, then, on delineating protein inputs from a terrestrial C3 ecosystem (about -26‰) versus estuarine mudflats that experienced mixed marine and freshwater inputs (about -13‰). Using the data, Evans Ogden et al. (2005) were not able to distinguish between the different types of agricultural fields in which birds fed, but data on selection of field type are presented elsewhere (Evans Ogden, 2002).

4.1.5.1 Mixture isotope ratios

Summary isotope ratio values for the $J = 234$ dunlin blood measurements are

$$\underset{\sim}{\bar{b}} = \begin{bmatrix} -16.81 \\ 12.13 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{\underset{\sim}{b}} = \begin{bmatrix} 10.27 & 2.48 \\ 2.48 & 1.99 \end{bmatrix}$$

and summarized in Table 4.1. Figure 4.2 illustrates the correlation of the carbon and nitrogen isotope ratios for blood ($r = 0.549$, p-value $< 2.2 \times 10^{-16}$), and the data are clearly not normal, especially carbon, with an apparent truncation of larger values.

Table 4.1: Dunlin blood summaries for carbon ($i = 1$) and nitrogen ($i = 2$).

Isotope ratios	J	$\bar{b}_1$	SD	SE($\bar{b}_1$)	$\bar{b}_2$	SD	SE($\bar{b}_2$)	r
Dunlin blood	234	−16.81	3.205	0.2095	12.13	1.409	0.0921	0.549

4.1.5.2 Source isotope ratios

Summary isotope ratio values for the $K_1 = 16$ terrestrial and $K_2 = 21$ marine source measurements are

$$\underset{\sim}{\bar{d}}_1 = \begin{bmatrix} -25.36 \\ 6.05 \end{bmatrix} \qquad \mathbf{S}_{\underset{\sim}{d_1}} = \begin{bmatrix} 1.72 & 0.74 \\ 0.74 & 1.59 \end{bmatrix}$$

$$\underset{\sim}{\bar{d}}_2 = \begin{bmatrix} -13.98 \\ 10.94 \end{bmatrix} \qquad \mathbf{S}_{\underset{\sim}{d_2}} = \begin{bmatrix} 10.55 & -0.47 \\ -0.47 & 3.80 \end{bmatrix}$$

and summarized in Table 4.2. Figure 4.3 also shows that the carbon and nitrogen isotope ratios are not statistically significantly correlated for terrestrial ($r = 0.447$, p-value $= 0.083$) or marine ($r = -0.075$, p-value $= 0.748$) sources.

Table 4.2: Terrestrial and marine summaries for carbon ($i = 1$) and nitrogen ($i = 2$).

Isotope ratios	K	$\bar{d}_{1s}$	SD	SE($\bar{d}_{1s}$)	$\bar{d}_{2s}$	SD	SE($\bar{d}_{2s}$)	r
Terrestrial	16	−25.36	1.313	0.3283	6.05	1.261	0.3152	0.447
Marine	21	−13.98	3.248	0.7087	10.94	1.948	0.4251	−0.075

4.1.5.3 Discrimination

Evans Ogden et al. (2004) conducted a diet experiment consisting entirely of foods of terrestrial C3 origin on four control birds. Food samples were archived weekly throughout the experiment so isotope ratio variation between weeks could be de-

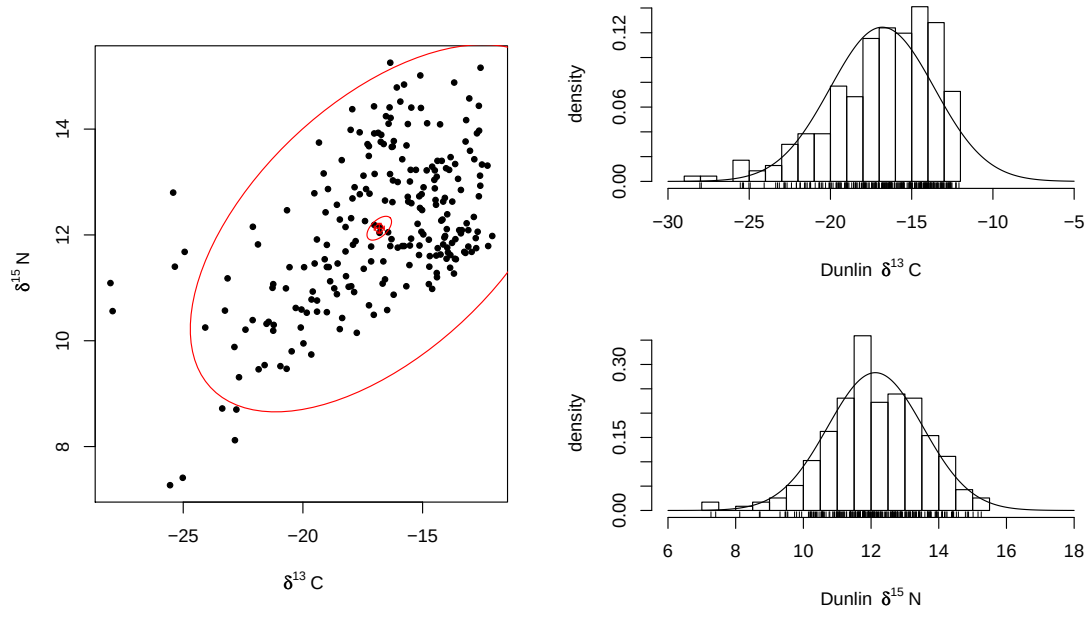

Figure 4.2: Dunlin blood $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate normal 95% confidence and prediction ellipses. Marginal distributions with estimated normal distributions.

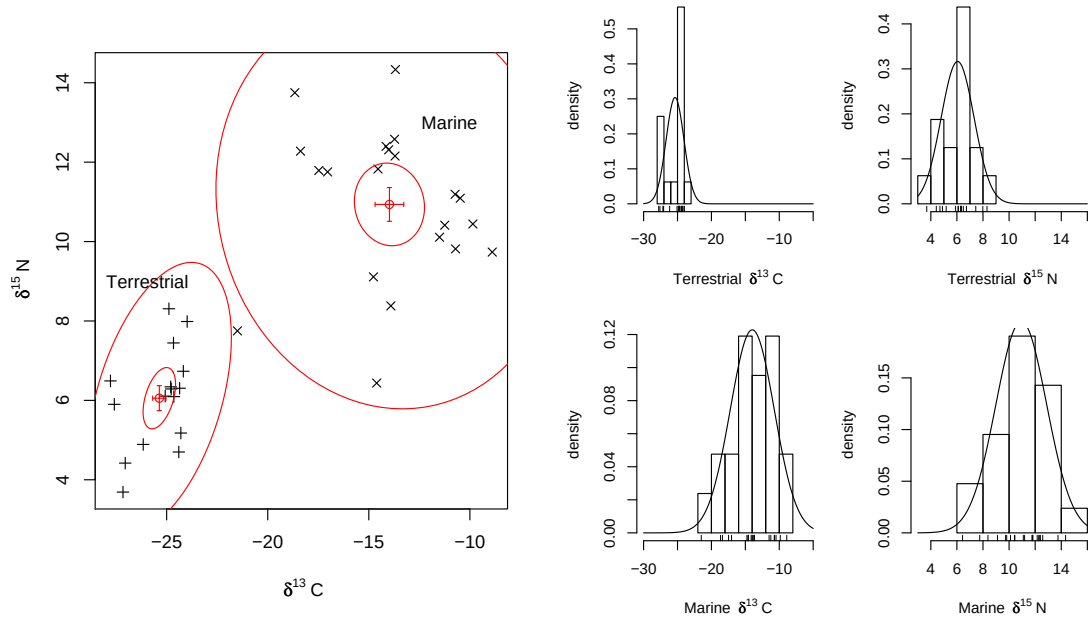

Figure 4.3: Terrestrial and marine $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate normal 95% confidence and prediction ellipses. Marginal distributions with estimated normal distributions.

termined. Sample isotope ratio values for the dunlin blood and diet in a control experiment are shown in Figure 4.4 and summarized in Table 4.3, as in Section 4.1.3. The estimated discrimination, $\hat{\underset{\sim}{\Delta}}_s$, is the difference between the average blood, $\bar{\underset{\sim}{d}}_{\mathrm{T}}$, and diet, $\bar{\underset{\sim}{d}}_{\mathrm{D}}$, isotope ratios, i.e., $\hat{\underset{\sim}{\Delta}}_s = \bar{\underset{\sim}{d}}_{\mathrm{T}} - \bar{\underset{\sim}{d}}_{\mathrm{D}}$. We will assume that the discrimination is the same for both marine and terrestrial sources, that is, $\underset{\sim}{\Delta}_1 = \underset{\sim}{\Delta}_2$. The estimated covariance of $\hat{\underset{\sim}{\Delta}}_s$ is

$$\begin{aligned}\mathbf{S}_{\hat{\underset{\sim}{\Delta}}_s} &= \mathbf{S}_{\underset{\sim}{d}_{\mathrm{T}}}/4 + \mathbf{S}_{\underset{\sim}{d}_{\mathrm{D}}}/28 \\ &= \begin{bmatrix} 0.0262 & 0.0039 \\ 0.0039 & 0.0030 \end{bmatrix} + \begin{bmatrix} 0.0052 & 0.0004 \\ 0.0004 & 0.0033 \end{bmatrix} = \begin{bmatrix} 0.0313 & 0.0043 \\ 0.0043 & 0.0063 \end{bmatrix}.\end{aligned}$$

The isotope ratios are not statistically significantly correlated for blood ($r = 0.087$, p-value $= 0.553$) or diet ($r = 0.447$, p-value $= 0.661$).

Table 4.3: Dunlin blood and diet isotope ratio values for carbon ($i = 1$) and nitrogen ($i = 2$) and the discrimination between the two.

Isotope ratios	K	$\bar{d}_{1s}$	SD	SE($\bar{d}_{1s}$)	$\bar{d}_{2s}$	SD	SE($\bar{d}_{2s}$)	r
Blood	4	−23.28	0.323	0.1617	6.50	0.109	0.0545	0.087
Diet	28	−24.63	0.380	0.0719	3.55	0.306	0.0578	0.447

Isotope ratios	$\hat{\Delta}_1$	SE($\hat{\Delta}_1$)	$\hat{\Delta}_2$	SE($\hat{\Delta}_2$)
Discrimination	1.35	0.1770	2.95	0.0795

4.1.5.4 Mixture and discrimination-corrected source ensemble

It is convenient to combine the source isotope ratios and discrimination estimates to visualize the effective relationships among the sources and the mixture. In Table 4.4 are the combined summaries, where the discrimination-corrected source isotope ratios, $\hat{\delta}'_{is} = \bar{d}_{is} + \hat{\Delta}_{is}$, plotted in Figure 4.5 have $\mathrm{SE}(\hat{\delta}'_{is})^2 = \mathrm{SE}(\bar{d}_{is})^2 + \mathrm{SE}(\hat{\Delta}_{is})^2$.

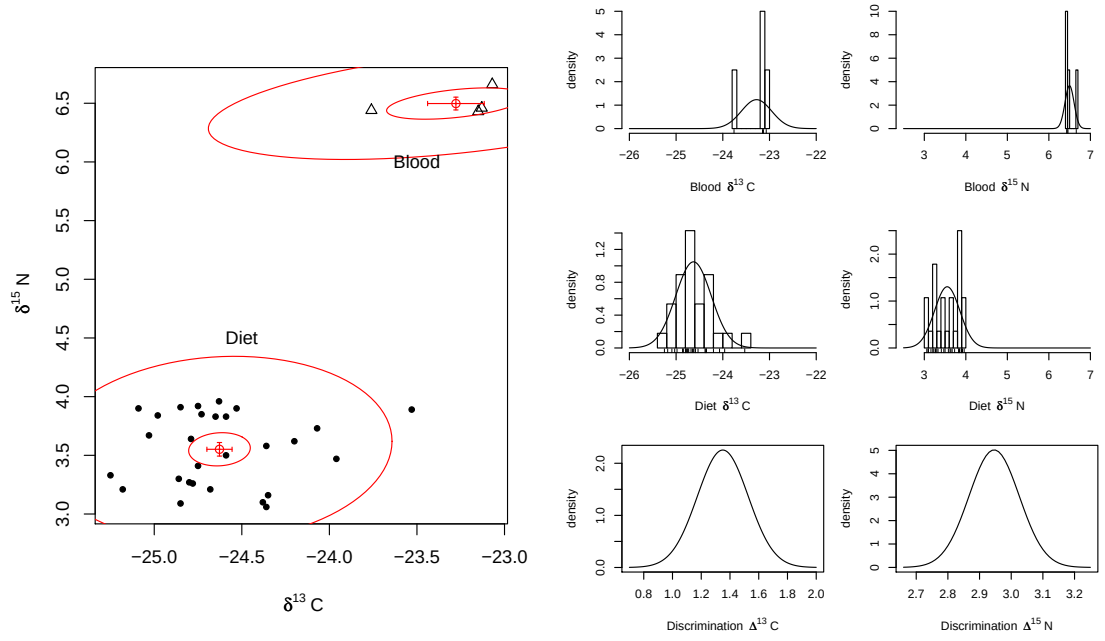

Figure 4.4: Diet experiment terrestrial diet and dunlin blood $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate 95% data and confidence ellipses. Marginal distributions with estimated normal distributions.

4.2 Implicit representation

The implicit representation perfectly constrained case solves for $\underset{\sim}{\pi}$ in $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$, when $\mathbf{A}_{\mathrm{x}}$ is square and nonsingular, with $\underset{\sim}{\pi} = \mathbf{A}_{\mathrm{x}}^{-1}\underset{\sim}{\beta}_{\mathrm{x}}$. The easiest large sample analysis

Table 4.4: Estimated parameters for the population means of mixture and discrimination-corrected source isotope ratios for carbon ($i = 1$) and nitrogen ($i = 2$).

Estimates	$\bar{b}_1$	SE($\bar{b}_1$)	$\bar{b}_2$	SE($\bar{b}_2$)	r
Dunlin	−16.81	0.209	12.13	0.092	0.549

Estimates	$\hat{\delta}'_{1s}$	SE($\hat{\delta}'_{1s}$)	$\hat{\delta}'_{2s}$	SE($\hat{\delta}'_{2s}$)
Terrestrial	−24.01	0.373	9.00	0.325
Marine	−12.63	0.731	13.88	0.432

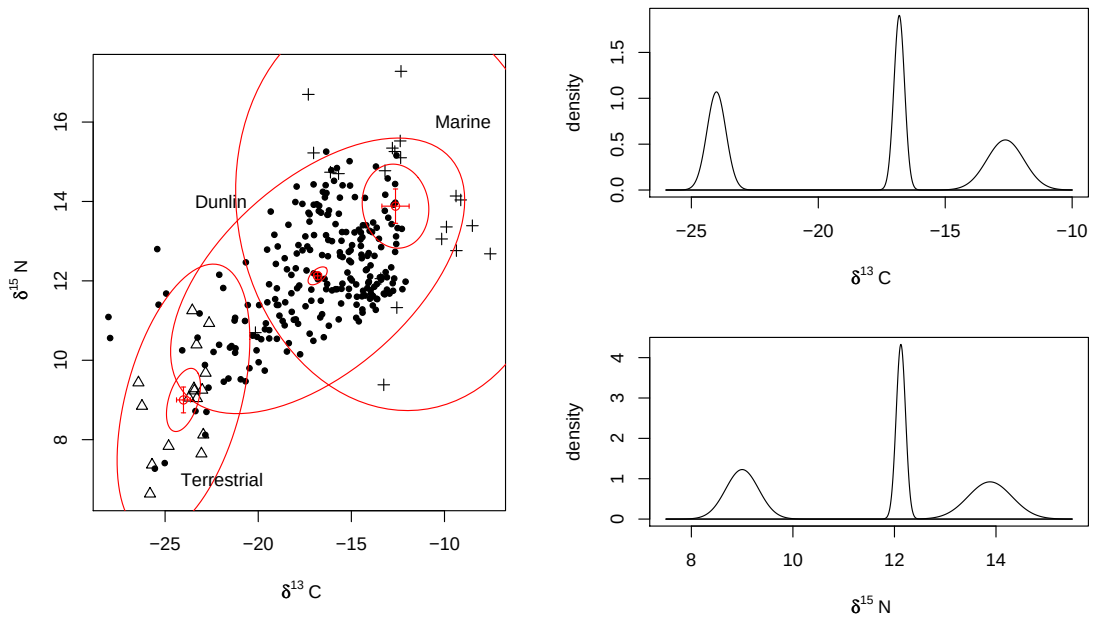

Figure 4.5: Left: Dunlin mixture and terrestrial and marine discrimination-corrected source $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate normal 95% data and confidence ellipses. Right: Estimated normal distributions for the mean isotope ratios from Table 4.4.

assumes $\underset{\sim}{\pi}$ sums to one, but does not force it to be nonnegative, thus the analysis is performed without assuming anything about $\underset{\sim}{\pi}$, not even that $\underset{\sim}{\pi}$ is a probability vector. We are interested in the sampling distribution for $\hat{\underset{\sim}{\pi}}$ when the components of $\mathbf{A}_{\mathrm{x}}$ and $\underset{\sim}{\beta}_{\mathrm{x}}$ are subject to variability. This section discusses the use of large sample and bootstrap methods. Recall that in Chapter 3 we discussed providing solutions to $\underset{\sim}{\beta}_{\mathrm{x}} = \mathbf{A}_{\mathrm{x}}\underset{\sim}{\pi}$ assuming that $\underset{\sim}{\beta}_{\mathrm{x}}$ and $\mathbf{A}_{\mathrm{x}}$ were known. In practice we estimate $\underset{\sim}{\beta}$ with $\bar{\underset{\sim}{b}}$ and $\mathbf{A}$ with $\hat{\mathbf{A}}$, substituting the estimates into the defining equation, and solving for $\hat{\underset{\sim}{\pi}}$ to estimate $\underset{\sim}{\pi}$. In particular $\hat{\underset{\sim}{\pi}} = \hat{\mathbf{A}}_{\mathrm{x}}^{-1}\bar{\underset{\sim}{b}}_{\mathrm{x}}$ where $\hat{\mathbf{A}}_{\mathrm{x}} = [\hat{\mathbf{A}}^{\top}, \underset{\sim}{1}]^{\top}$ and $\bar{\underset{\sim}{b}}_{\mathrm{x}} = [\bar{\underset{\sim}{b}}^{\top}, 1]^{\top}$. We use the dunlin example with carbon ($I = 1$) and the terrestrial and marine sources ($S = 2$) to illustrate the various methods.

Example 4.2.1. Dunlin point estimate using carbon

We substitute the estimates for $\underset{\sim}{\beta}_{\mathrm{x}}$ and $\mathbf{A}_{\mathrm{x}}$ into the defining equation and solve for $\hat{\underset{\sim}{\pi}}$,

$$\hat{\mathbf{A}}_{\mathrm{x}} = \begin{bmatrix} -24.01 & -12.63 \\ 1 & 1 \end{bmatrix} \quad \text{giving} \quad \hat{\mathbf{A}}_{\mathrm{x}}^{-1} = \begin{bmatrix} -0.0879 & -1.110 \\ 0.0879 & 2.110 \end{bmatrix}$$

and

$$\bar{\underset{\sim}{b}}_{\mathrm{x}} = \begin{bmatrix} -16.81 \\ 1 \end{bmatrix} \quad \text{giving} \quad \hat{\underset{\sim}{\pi}} = \hat{\mathbf{A}}_{\mathrm{x}}^{-1}\bar{\underset{\sim}{b}}_{\mathrm{x}} = \begin{bmatrix} 0.367 \\ 0.633 \end{bmatrix}$$

as the estimated proportional contributions of terrestrial and marine sources to the dunlin diet. △

The solution $\underset{\sim}{\pi} = \mathbf{A}_{\mathrm{x}}^{-1}\underset{\sim}{\beta}_{\mathrm{x}}$ makes sense when $\underset{\sim}{\pi}$ is on the simplex, and this happens when $\underset{\sim}{\beta}_{\mathrm{x}}$ is in the convex hull of $\{\delta'_{i1}, \ldots, \delta'_{iS}\}$, as described in Section 3.1. Therefore, for each $i = 1, \ldots, I$, there must exist s and s^* such that $\delta'_{is} \leq \beta_i \leq \delta'_{is^*}$ and that this results in β_i being contained in all pairwise isotopic convex hulls, and these indicies s and s^* may be different for each i. Given this, by the law of large numbers, for large

enough sample sizes the estimates will also satisfy $\hat{\delta}'_{is} \leq \bar{b}_i \leq \hat{\delta}'_{is^*}$, for all i, implying that $\hat{\mathbf{A}}_{\mathrm{x}}$ will eventually be invertible and $\hat{\underset{\sim}{\pi}} = \hat{\mathbf{A}}_{\mathrm{x}}^{-1}\bar{\underset{\sim}{b}}_{\mathrm{x}}$ will be a probability vector.

4.2.1 Variation in mixture only, large samples

Let $\mathbf{A}_{\mathrm{x}}$ be fixed and known and $\underset{\sim}{\beta}_{\mathrm{x}}$ be subject to uncertainty and estimated as in Section 4.1. We use defining equation (1.1) augmented with a row of ones and substitute the estimate $\bar{\underset{\sim}{b}}_{\mathrm{x}}$ to yield $\hat{\underset{\sim}{\pi}} = \mathbf{A}_{\mathrm{x}}^{-1}\bar{\underset{\sim}{b}}_{\mathrm{x}}$. That

$$\bar{\underset{\sim}{b}}_{\mathrm{x}} \;\dot{\sim}\; \text{Normal}\left(\underset{\sim}{\beta}_{\mathrm{x}}, \Sigma_{\underset{\sim}{b}_{\mathrm{x}}}/J\right), \qquad \text{where} \quad \Sigma_{\underset{\sim}{b}_{\mathrm{x}}} = \begin{bmatrix} \Sigma_{\underset{\sim}{b}} & 0 \\ 0 & 0 \end{bmatrix},$$

implies that

$$\hat{\underset{\sim}{\pi}} \;\dot{\sim}\; \text{Normal}\left(\underset{\sim}{\pi},\ \mathbf{A}_{\mathrm{x}}^{-1}(\Sigma_{\underset{\sim}{b}_{\mathrm{x}}}/J)\mathbf{A}_{\mathrm{x}}^{-\top}\right).$$

Example 4.2.2. One-isotope, two-source example ($I = 1, S = 2$)
Assume we have a random sample of size J from the mixture population. In this case, we need only to derive the estimation of π_1, since $\pi_2 = 1 - \pi_1$. The estimated proportion vector is

$$\hat{\underset{\sim}{\pi}} = \mathbf{A}_{\mathrm{x}}^{-1}\bar{\underset{\sim}{b}}_{\mathrm{x}} = \frac{1}{\delta'_{11} - \delta'_{12}} \begin{bmatrix} 1 & -\delta'_{12} \\ -1 & \delta'_{11} \end{bmatrix} \begin{bmatrix} \bar{b}_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\bar{b}_1 - \delta'_{12}}{\delta'_{11} - \delta'_{12}} \\ \frac{\delta'_{11} - \bar{b}_1}{\delta'_{11} - \delta'_{12}} \end{bmatrix}, \tag{4.7}$$

with estimated covariance

$$\begin{aligned} \hat{\Sigma}_{\hat{\underset{\sim}{\pi}}} &= \mathbf{A}_{\mathrm{x}}^{-1}(\mathbf{S}_{\underset{\sim}{b}_{\mathrm{x}}}/J)\mathbf{A}_{\mathrm{x}}^{-\top} \\ &= \frac{1}{J(\delta'_{11} - \delta'_{12})^2} \begin{bmatrix} 1 & -\delta'_{12} \\ -1 & \delta'_{11} \end{bmatrix} \begin{bmatrix} s^2_{\underset{\sim}{b}} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -\delta'_{12} & \delta'_{11} \end{bmatrix} \\ &= \frac{s^2_{\underset{\sim}{b}}}{J(\delta'_{11} - \delta'_{12})^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \end{aligned}$$

implying

$$\widehat{\text{Var}}[\hat{\pi}_1] \;=\; \frac{s^2_{\underset{\sim}{b}}}{J(\delta'_{11}-\delta'_{12})^2}. \tag{4.8}$$

In this case $\mathbf{A}_{\text{x}}^{-1}$ will exist if $\delta'_{11} \neq \delta'_{12}$, when the determinant of $\mathbf{A}_{\text{x}}$ is not zero, which occurs when the source mean discrimination-corrected isotope ratios are distinct from each other. Also, note that the variance increases as the source values are more similar since it is more difficult to distinguish between the sources. △

Example 4.2.3. Mixture variation dunlin carbon
Let $s^2_{\underset{\sim}{b}} = 10.270$ and $J = 234$ as in Table 4.1, then

$$\hat{\Sigma}_{\hat{\underset{\sim}{\pi}}} \;=\; 0.000339 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \qquad \text{and} \quad \text{SD}[\hat{\pi}_1] \;=\; 0.0184.$$

△

4.2.2 Variation in mixture and sources, large samples

Let both $\mathbf{A}_{\text{x}}$ and $\underset{\sim}{\beta}_{\text{x}}$ be subject to uncertainty and estimated as in Section 4.1. We use our defining equation (1.1) and substitute the estimates $\bar{\underset{\sim}{b}}_{\text{x}}$ and $\hat{\mathbf{A}}_{\text{x}}$ to yield $\bar{\underset{\sim}{b}}_{\text{x}} = \hat{\mathbf{A}}_{\text{x}}\hat{\underset{\sim}{\pi}}$. In the following Theorem 4.2.4, elements of $\mathbf{A}$ are discrimination-corrected source isotope ratios, δ'_{is} as in Section 4.1.5.4. Later in the implicit function theorem in Section 4.2.2.1 we separate $\delta'_{is} = \delta_{is} + (\delta_{i\text{T}} - \delta_{i\text{D}})$ into its individual components. Both treatments of δ'_{is} can be implemented using both the delta method and the implicit function theorem.

Theorem 4.2.4. Delta method
Let $\underset{\sim}{\mu}_y = [\text{Vec}[\mathbf{A}]^\top, \underset{\sim}{\beta}^\top]^\top$, the stacked columns of $\mathbf{A}$ and $\underset{\sim}{\beta}$, where δ'_{is} is the discrimination-corrected source isotope ratio as in Section 4.1.4. Let $\mathbf{L}$ be the derivative of $\underset{\sim}{\pi} =$

$\underset{\sim}{\pi}(\mathbf{A}, \underset{\sim}{\beta}) \equiv \underset{\sim}{\pi}(\underset{\sim}{\mu}_y)$ with respect to the elements of $\mathbf{A}$ and $\underset{\sim}{\beta}$,

$$\mathbf{L} = \begin{bmatrix} \frac{\partial \pi_1}{\partial \delta'_{11}} & \cdots & \frac{\partial \pi_1}{\partial \delta'_{I1}} & & \frac{\partial \pi_1}{\partial \delta'_{1S}} & \cdots & \frac{\partial \pi_1}{\partial \delta'_{IS}} & \frac{\partial \pi_1}{\partial \beta_1} & \cdots & \frac{\partial \pi_1}{\partial \beta_I} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial \pi_S}{\partial \delta'_{11}} & \cdots & \frac{\partial \pi_S}{\partial \delta'_{I1}} & & \frac{\partial \pi_S}{\partial \delta'_{1S}} & \cdots & \frac{\partial \pi_S}{\partial \delta'_{IS}} & \frac{\partial \pi_S}{\partial \beta_1} & \cdots & \frac{\partial \pi_S}{\partial \beta_I} \end{bmatrix}_{S \times (S^2-1)}.$$

Let $\underset{\sim}{y}_k = [\text{Vec}[\hat{\mathbf{A}}]^\top, \underset{\sim}{\bar{b}}^\top]^\top$ with common sample size, K, for each column of $\hat{\mathbf{A}}$ and $\underset{\sim}{\bar{b}}$, and let $\sqrt{K}\underset{\sim}{y}_k$ have covariance matrix $\mathbf{\Omega}$. If the central limit theorem holds, that

$$\sqrt{K}\left(\underset{\sim}{y}_k - \underset{\sim}{\mu}_y\right) \xrightarrow{\mathscr{L}} \text{Normal}\left(\underset{\sim}{0}, \mathbf{\Omega}\right), \tag{4.9}$$

then by the Delta method we have

$$\sqrt{K}\left(\hat{\underset{\sim}{\pi}} - \underset{\sim}{\pi}\right) \xrightarrow{\mathscr{L}} \text{Normal}\left(\underset{\sim}{0}, \mathbf{L}\mathbf{\Omega}\mathbf{L}^\top\right). \tag{4.10}$$

Therefore, an approximate distribution for $\hat{\underset{\sim}{\pi}}$ is

$$\hat{\underset{\sim}{\pi}} \dot{\sim} \text{Normal}\left(\underset{\sim}{\pi}, \mathbf{V}\right),$$

where $\mathbf{V} = \mathbf{L}(\mathbf{\Omega}/K)\mathbf{L}^\top$. □

In practice, $\mathbf{\Omega}$ is estimated with $\hat{\mathbf{\Omega}}$ and $\mathbf{L}$ with $\hat{\mathbf{L}}$. Also, it is convenient to include different sample sizes for sources and mixture in $\hat{\mathbf{\Omega}}/K$ as $\hat{\mathbf{\Omega}}_k$, defined later in (4.14). Substitute estimated covariances to give an approximate covariance for $\hat{\underset{\sim}{\pi}}$ as

$$\hat{\mathbf{V}} = \hat{\mathbf{L}}\hat{\mathbf{\Omega}}_k\hat{\mathbf{L}}^\top = \hat{\mathbf{L}}\begin{bmatrix} \mathbf{S}_{\hat{\underset{\sim}{\delta}}'_1} & & & \\ & \ddots & & \\ & & \mathbf{S}_{\hat{\underset{\sim}{\delta}}'_S} & \\ & & & \mathbf{S}_{\underset{\sim}{\bar{b}}} \end{bmatrix}\hat{\mathbf{L}}^\top.$$

The sample sizes in the estimates $\hat{\underset{\sim}{\delta}}'_1, \ldots, \hat{\underset{\sim}{\delta}}'_S$, and $\underset{\sim}{\bar{b}}$ are assumed to increase to infinity at the same rate. See Section 4.2.2.1 for a more precise definition.

Example 4.2.5. One-isotope, two-source example ($I = 1, S = 2$)
As in (4.7), we can write

$$\hat{\underset{\sim}{\pi}} = \begin{bmatrix} (\bar{b}_1 - \hat{\delta}'_{12})(\hat{\delta}'_{11} - \hat{\delta}'_{12})^{-1} \\ (\hat{\delta}'_{11} - \bar{b}_1)(\hat{\delta}'_{11} - \hat{\delta}'_{12})^{-1} \end{bmatrix}. \tag{4.11}$$

For the covariance matrix $\hat{\mathbf{V}}$ we need

$$\hat{\mathbf{\Omega}}_k = \begin{bmatrix} s^2_{\hat{\delta}'_{11}} & & \\ & s^2_{\hat{\delta}'_{12}} & \\ & & s^2_b/J \end{bmatrix}$$

and

$$\begin{aligned} \mathbf{L} &= \begin{bmatrix} \frac{\partial \pi_1}{\partial \delta'_{11}} & \frac{\partial \pi_1}{\partial \delta'_{12}} & \frac{\partial \pi_1}{\partial \beta_1} \\ \frac{\partial \pi_2}{\partial \delta'_{11}} & \frac{\partial \pi_2}{\partial \delta'_{12}} & \frac{\partial \pi_2}{\partial \beta_1} \end{bmatrix} \\ &= \begin{bmatrix} (\delta'_{12} - \beta_1)(\delta'_{11} - \delta'_{12})^{-2} & (\beta_1 - \delta'_{11})(\delta'_{11} - \delta'_{12})^{-2} & (\delta'_{11} - \delta'_{12})^{-1} \\ (\beta_1 - \delta'_{12})(\delta'_{11} - \delta'_{12})^{-2} & (\delta'_{11} - \beta_1)(\delta'_{11} - \delta'_{12})^{-2} & -(\delta'_{11} - \delta'_{12})^{-1} \end{bmatrix} \\ &= (\delta'_{11} - \delta'_{12})^{-1} \begin{bmatrix} -\pi_1 & -\pi_2 & 1 \\ \pi_1 & \pi_2 & -1 \end{bmatrix}. \end{aligned}$$

The natural estimates for δ'_{is} and β_i are substituted to give $\hat{\mathbf{L}}$. The resulting covariance matrix is the 2×2 matrix $\hat{\mathbf{L}}\hat{\mathbf{\Omega}}_k\hat{\mathbf{L}}^\top$ with diagonal elements

$$\mathrm{Var}[\hat{\pi}_1] = \mathrm{Var}[\hat{\pi}_2] = (\hat{\pi}_1^2 s^2_{\hat{\delta}'_{11}} + \hat{\pi}_2^2 s^2_{\hat{\delta}'_{12}} + s_b^2/J)/(\hat{\delta}'_{11} - \hat{\delta}'_{12})^2$$

and off-diagonal elements $\mathrm{Cov}[\hat{\pi}_1, \hat{\pi}_2] = -\mathrm{Var}[\hat{\pi}_1]$. As in (4.8), the denominator scales the variance, and may substantially determine the magnitude. In addition is the additional cost of estimating $\mathbf{A}$, where the variance component for $s^2_{\hat{\delta}'_{11}}$, for example from (4.11), contributes more to the variance of $\hat{\pi}_1$ when $\bar{b}_1$ is further from $\hat{\delta}'_{11}$ than closer. $\triangle$

Example 4.2.6. Mixture and source variation dunlin carbon
$\hat{\mathbf{A}}_{\mathrm{x}}$ and $\underset{\sim}{\bar{b}}_{\mathrm{x}}$ are defined in example 4.2.1, and variances are given in Table 4.4. $\mathrm{Var}[\hat{\delta}'_{11}] = 0.373^2$ and $\mathrm{Var}[\hat{\delta}'_{12}] = 0.731^2$ are substituted for $s^2_{\hat{\delta}'_{11}}$ and $s^2_{\hat{\delta}'_{12}}$ above, and $s^2_{\bar{b}_1} = 0.04389$. Calculation results in a standard deviation for $\hat{\pi}_1$ about 2.5 times larger than with mixture variation alone,

$$\hat{\mathbf{V}} \;=\; 0.002132 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \qquad \text{and} \quad \mathrm{SD}[\hat{\pi}_1] \;=\; 0.0462.$$

Recall that the sample sizes are all less than 30, see Tables 4.2 and 4.3, which suggests that this example may not be a good candidate for this large sample method. △

The implicit-function method that follows provides a general way to perform the calculations of matrix $\mathbf{L}$ in the delta method, and the results will be the same.

4.2.2.1 Implicit-function method

The following theorems and corollary in stable isotope notation is adapted from Benichou and Gail (1989). This general result includes discrimination estimated separately from the isotope ratio values, as in the dunlin example. The formulas simplify if, in place of isotope ratio values, discrimination-corrected isotope ratio values are used, and the formulas expand if the discrimination is decomposed into its constituent diet experiment components. Let $\underset{\sim}{\mu}_y$ be a vector of population means for all elements in matrix $\mathbf{B} = \mathbf{A} - \underset{\sim}{\beta}\underset{\sim}{1}^{\top}$ as in (1.5) for the BMM. The order of the elements is not important, provided the ordering is consistent with the matrices $\mathbf{\Omega}$, $\mathbf{G}$, and $\mathbf{H}$ defined later. We choose the ordering

$$\begin{aligned} \underset{\sim}{\mu}_y \;=\; [&\delta_{11}, \ldots, \delta_{I1}, \delta_{11\mathrm{T}}, \ldots, \delta_{I1\mathrm{T}}, \delta_{11\mathrm{D}}, \ldots, \delta_{I1\mathrm{D}}, \ldots, \\ &\delta_{1S}, \ldots, \delta_{IS}, \delta_{1S\mathrm{T}}, \ldots, \delta_{IS\mathrm{T}}, \delta_{1S\mathrm{D}}, \ldots, \delta_{IS\mathrm{D}}, \beta_1, \ldots, \beta_I], \end{aligned} \tag{4.12}$$

a $1 \times n$ vector, where $n = 3SI + I = (3S + 1)(S - 1)$, since $I = S - 1$. There are three elements in the BMM associated with each source-isotope pair, and one element for each mixture isotope. Let $\underset{\sim}{y}_k$ be the vector of sample means corresponding to population means, $\underset{\sim}{\mu}_y$,

$$\begin{aligned} \underset{\sim}{y}_k &= [\bar{d}_{11}, \ldots, \bar{d}_{I1}, \bar{d}_{11\mathrm{T}}, \ldots, \bar{d}_{I1\mathrm{T}}, \bar{d}_{11\mathrm{D}}, \ldots, \bar{d}_{I1\mathrm{D}}, \ldots, \\ &\qquad \bar{d}_{1S}, \ldots, \bar{d}_{IS}, \bar{d}_{1S\mathrm{T}}, \ldots, \bar{d}_{IS\mathrm{T}}, \bar{d}_{1S\mathrm{D}}, \ldots, \bar{d}_{IS\mathrm{D}}, \bar{b}_1, \ldots, \bar{b}_I]. \end{aligned} \tag{4.13}$$

Let the q distinct sample sizes associated with the elements of $\underset{\sim}{y}_k$, k_m, $m = 1, \ldots, q$, with $k = \sum_{m=1}^{q} k_m$, approach infinity at a common rate such that $k_m/k \xrightarrow{p} p'_m > 0, m = 1, \ldots, q$. Finally, let the covariance matrix, $\mathbf{\Omega}_k$, associated with $\underset{\sim}{y}_k$ be a block diagonal matrix of the covariance matrices associated with the correlated elements of $\underset{\sim}{y}_k$. That is, let $\mathrm{Cov}[\underset{\sim}{\bar{d}}_s] = \Sigma_{\underset{\sim}{\bar{d}}_s}$, $\mathrm{Cov}[\underset{\sim}{\bar{d}}_{s\mathrm{T}}] = \Sigma_{\underset{\sim}{\bar{d}}_{s\mathrm{T}}}$, and $\mathrm{Cov}[\underset{\sim}{\bar{d}}_{s\mathrm{D}}] = \Sigma_{\underset{\sim}{\bar{d}}_{s\mathrm{D}}}$, $s = 1, \ldots, S$, and let $\mathrm{Cov}[\underset{\sim}{\bar{b}}] = \Sigma_{\underset{\sim}{\bar{b}}}$, where each covariance depends on the sample size informing that estimate independent of the others. Then

$$\mathbf{\Omega}_k = \mathrm{BlkDiag}[\Sigma_{\underset{\sim}{\bar{d}}_1}, \Sigma_{\underset{\sim}{\bar{d}}_{1\mathrm{T}}}, \Sigma_{\underset{\sim}{\bar{d}}_{1\mathrm{D}}}, \ldots, \Sigma_{\underset{\sim}{\bar{d}}_S}, \Sigma_{\underset{\sim}{\bar{d}}_{S\mathrm{T}}}, \Sigma_{\underset{\sim}{\bar{d}}_{S\mathrm{D}}}, \Sigma_{\underset{\sim}{\bar{b}}}]. \tag{4.14}$$

Theorem 4.2.7. Implicit Function Theorem

Let O be an open subset of the $(S + n)$-dimensional space with elements $[\underset{\sim}{\pi}^\top, \underset{\sim}{\mu}_y] = (\pi_1, \ldots, \pi_S, \mu_{y1}, \ldots, \mu_{yn})$. Let real functions $g_i,\ i = 1, \ldots, S$, be continuous in O and have continuous first partial derivatives in O satisfying $g_i(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) = 0$ at some point $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$ in O. Define

$$\mathbf{G} = \left[\frac{\partial g_i}{\partial \pi_{i'}}\right]_{S \times S}, \quad i, i' = 1, \ldots, S. \tag{4.15}$$

If the determinant $|\mathbf{G}| \neq 0$ at $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$, then there exists an open rectangular region in O satisfying

$$|\pi_s - \pi_{0s}| < \varepsilon_s, \quad s = 1, \ldots, S, \quad \text{and} \tag{4.16}$$

$$|\mu_{ym} - \mu_{0ym}| < \varepsilon_m, \quad m = 1, \ldots, n, \tag{4.17}$$

and there exists a set of S real functions f_s mapping each element in $\underset{\sim}{\mu}_y$ in region (4.17) to a single $\underset{\sim}{\pi}$ in region (4.16) such that $\underset{\sim}{\pi} = (\pi_1, \ldots, \pi_S)^\top = (f_1(\underset{\sim}{\mu}_y), \ldots, f_S(\underset{\sim}{\mu}_y))^\top$ and $g_i(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) = 0,\ i = 1, \ldots, S$. Moreover, the functions $f_s,\ s = 1, \ldots, S$, are continuous and have continuous first partial derivatives $[\partial f_s / \partial \underset{\sim}{\mu}_y]_{S \times n}$ which are elements of the matrix product $-\mathbf{G}^{-1}\mathbf{H}$, where

$$\mathbf{H} \ = \ \left[\frac{\partial g_i}{\partial \underset{\sim}{\mu}_y} \right]_{S \times n} . \tag{4.18}$$

□

The implicit function theorem thus asserts the existence, in a neighborhood of $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$, of the explicit functions f_s needed to apply Theorem 4.2.4, the delta method (Taylor and Mann, 1983, Chap 8).

Theorem 4.2.7 can be applied to the BMM. Such an open subset O exists because proportions $\pi_s \subset \mathbb{R}^1,\ s = 1, \ldots, S$, and all isotope values are contained in $\mathbb{R}^1$, thus taking $O \equiv \mathbb{R}^{S+n}$ is one possible open subset. Functions g_i are the ith row, $i = 1, \ldots, S$, of the difference of the right-hand side from the left-hand side of (1.5),

$$\begin{bmatrix} \mathbf{B} \\ \underset{\sim}{1}^\top \end{bmatrix} \underset{\sim}{\pi} - \begin{bmatrix} \underset{\sim}{0} \\ 1 \end{bmatrix} \ = \ \underset{\sim}{0},$$

or, alternatively, $\mathbf{B}_{\mathrm{x}}\underset{\sim}{\pi} - [\underset{\sim}{0}^\top, 1]^\top = \underset{\sim}{0}$. Specifically,

$$\begin{aligned} g_i(\underset{\sim}{\pi}^\top, \underset{\sim}{\mu}_y) \ &= \ [\delta_{i1} + (\delta_{i1\mathrm{T}} - \delta_{i1\mathrm{D}}) - \beta_i, \ldots, \delta_{iS} + (\delta_{iS\mathrm{T}} - \delta_{iS\mathrm{D}}) - \beta_i]\, \underset{\sim}{\pi} \\ &= \ 0 \end{aligned} \tag{4.19}$$

for $i = 1, \ldots, I$, and

$$g_S(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) \ = \ \pi_1 + \cdots + \pi_S - 1 \ = \ 0 \tag{4.20}$$

for $i = S$.

The implicit function Theorem 4.2.7 sets $\mathbf{L} = -\mathbf{G}^{-1}\mathbf{H}$ in the delta method Theorem 4.2.4, where $\mathbf{G} = \mathbf{B}_{\mathrm{x}}$ for the BMM and $\mathbf{H}$ is defined in (4.18). The implicit function theorem requires that $\mathbf{G}^{-1}$ exists, which will by assumption when $\mathbf{G} = \mathbf{B}_{\mathrm{x}}$ as shown in Section 2.1.1. The main result relies on the following corollary.

Corollary 4.2.8. Let $\underset{\sim}{\mu}_y$ be as in Theorem 4.2.4. Suppose there exists a unique $S \times 1$ vector $\underset{\sim}{\pi}_0$ satisfying $g_i(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y}) = 0,\ i = 1, \ldots, S$, that the functions f_s are continuous with continuous first partial derivatives in an open set containing $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$, and that $\mathbf{G}$, defined in Theorem 4.2.7, has $|\mathbf{G}| \neq 0$ evaluated at $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$. To each $\underset{\sim}{y}_k$ there corresponds a unique solution $\hat{\underset{\sim}{\pi}}_k = (f_1(\underset{\sim}{y}_k), \ldots, f_S(\underset{\sim}{y}_k))^\top$ with $g_i(\hat{\underset{\sim}{\pi}}_k, \underset{\sim}{y}_k) = 0,\ i = 1, \ldots, S$. As k increases, by the strong law of large numbers, $\underset{\sim}{y}_k$ goes almost surely to $\underset{\sim}{\mu}_{0y}$ and $\hat{\underset{\sim}{\pi}}_k$ goes almost surely to $\underset{\sim}{\pi}_0$. Then,

$$\sqrt{k}(\hat{\underset{\sim}{\pi}}_k - \underset{\sim}{\pi}_0) \xrightarrow{\mathscr{L}} \text{Normal}(\underset{\sim}{0}, \mathbf{G}^{-1}\mathbf{H}\boldsymbol{\Omega}(\mathbf{G}^{-1}\mathbf{H})^\top), \tag{4.21}$$

where $\mathbf{G}$ and $\mathbf{H}$ are evaluated at $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$ (Benichou and Gail, 1989). □

If $\hat{\boldsymbol{\Omega}}$ is a consistent estimator of $\boldsymbol{\Omega}$ evaluated at $\underset{\sim}{y}_k$ then the asymptotic covariance $\mathbf{G}^{-1}\mathbf{H}\boldsymbol{\Omega}(\mathbf{G}^{-1}\mathbf{H})^\top$ can be consistently estimated by substituting $\hat{\boldsymbol{\Omega}}$ for $\boldsymbol{\Omega}$ and by substituting $\hat{\mathbf{G}}^{-1}\hat{\mathbf{H}}$ evaluated at $(\hat{\underset{\sim}{\pi}}_k, \underset{\sim}{y}_k)$ for $\mathbf{G}^{-1}\mathbf{H}$ evaluated at $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$ (Benichou and Gail, 1989). This follows from the continuity of $\mathbf{G}$ and $\mathbf{H}$ in an open space containing $(\underset{\sim}{\pi}_0, \underset{\sim}{\mu}_{0y})$. These substitutions are necessary to evaluate the covariance matrix.

4.2.2.1.1 Implementation Applying the implicit function theorem for the BMM is straightforward. An approximate large sample distribution for $\hat{\underset{\sim}{\pi}}_k$ is

$$\hat{\underset{\sim}{\pi}}_k \ \dot{\sim}\ \text{Normal}\left(\underset{\sim}{\pi}_0,\ \hat{\mathbf{G}}^{-1}\hat{\mathbf{H}}\hat{\boldsymbol{\Omega}}_k(\hat{\mathbf{G}}^{-1}\hat{\mathbf{H}})^\top\right). \tag{4.22}$$

where $\hat{\mathbf{\Omega}}_k$ is an estimate of $\text{Cov}[\underset{\sim}{y}_k] = \mathbf{\Omega}_k$. In (4.22), we need expressions for $\hat{\mathbf{\Omega}}_k, \hat{\mathbf{G}}^{-1}$, and $\hat{\mathbf{H}}$.

The covariance matrix $\hat{\mathbf{\Omega}}_k$ reflects what is known and the assumptions about the relationships between the sample means in $\underset{\sim}{y}_k$. Most covariances will be zero since the mixture and each source are measured independently.

The functions $g_i(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) = 0$ are defined using (4.19) and (4.20) and used in both $\mathbf{G}$ and $\mathbf{H}$, where $\mathbf{G} = \mathbf{B}_{\text{x}}$ and $\mathbf{H} = [\partial g_i / \partial \underset{\sim}{\mu}_y]_{S \times n}$. The rows of matrix $\mathbf{H}$ are given from (4.18) as

$$
\begin{aligned}
\mathbf{H}_i &= \left[\frac{\partial g_i}{\partial \underset{\sim}{\mu}_y}\right]_{1\times n} \qquad (4.23)\\
&\quad [\overbrace{0}^{(1,1)}, \ldots, \overbrace{\pi_{01}}^{(i,1)}, \ldots, \overbrace{0}^{(I,1)}, \overbrace{0}^{(1,1)}, \ldots, \overbrace{\pi_{01}}^{(i,1)}, \ldots, \overbrace{0}^{(I,1)}, \overbrace{0}^{(1,1)}, \ldots, \overbrace{-\pi_{01}}^{(i,1)}, \ldots, \overbrace{0}^{(I,1)},\\
&\quad \overbrace{0}^{(1,2)}, \ldots, \overbrace{\pi_{02}}^{(i,2)}, \ldots, \overbrace{0}^{(I,2)}, \overbrace{0}^{(1,2)}, \ldots, \overbrace{\pi_{02}}^{(i,2)}, \ldots, \overbrace{0}^{(I,2)}, \overbrace{0}^{(1,2)}, \ldots, \overbrace{-\pi_{02}}^{(i,2)}, \ldots, \overbrace{0}^{(I,2)},\\
&= \quad \ldots,\\
&\quad \overbrace{0}^{(1,S)}, \ldots, \overbrace{\pi_{0S}}^{(i,S)}, \ldots, \overbrace{0}^{(I,S)}, \overbrace{0}^{(1,S)}, \ldots, \overbrace{\pi_{0S}}^{(i,S)}, \ldots, \overbrace{0}^{(I,S)}, \overbrace{0}^{(1,S)}, \ldots, \overbrace{-\pi_{0S}}^{(i,S)}, \ldots, \overbrace{0}^{(I,S)},\\
&\quad \overbrace{0}^{(1)}, \ldots, \overbrace{-1}^{(i)}, \ldots, \overbrace{0}^{(I)}]_{1\times n}\,,
\end{aligned}
$$

$i = 1, \ldots, I$, and

$$
\mathbf{H}_S = \left[\frac{\partial g_S}{\partial \underset{\sim}{\mu}_y}\right]_{1\times n} = [0, \ldots, 0]_{1\times n}\,. \qquad (4.24)
$$

Thus, $\mathbf{H}$ has matrix form

$$\begin{aligned}\mathbf{H} &= \begin{bmatrix} \pi_0^\top \otimes \mathbf{I}_I \otimes [1\ 1\ 1] & -1 \otimes \mathbf{I}_I \\ \mathbb{Q}^\top & \mathbb{Q}^\top \end{bmatrix} \\ &= \begin{bmatrix} \pi_{01} & \cdots & 0 & & -1 & \cdots & 0 \\ & \ddots & & \cdots \text{ 3 times for} & & \ddots & \\ 0 & \cdots & \pi_{01} & \text{each } \pi_{0s} \cdots & 0 & \cdots & -1 \\ 0 & \cdots & 0 & & 0 & \cdots & 0 \end{bmatrix}. \end{aligned} \tag{4.25}$$

To calculate the covariance $\hat{\mathbf{V}} = \hat{\mathbf{G}}^{-1}\hat{\mathbf{H}}\hat{\mathbf{\Omega}}_k(\hat{\mathbf{G}}^{-1}\hat{\mathbf{H}})^\top$, substitute the natural estimates of the population parameters.

Example 4.2.9. One-isotope, two-source example ($I = 1, S = 2$)
To estimate $\mathbf{V}$ we first construct $\mathbf{G}^{-1}$, $\mathbf{H}$, and $\mathbf{\Omega}_k$. Now $\mathbf{G}$ takes the value of $\mathbf{B}_{\mathrm{x}}$ as in (1.7), thus

$$\begin{aligned}\mathbf{G} &= \begin{bmatrix} \delta_{11} + (\delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) - \beta_1 & \delta_{12} + (\delta_{12\mathrm{T}} - \delta_{12\mathrm{D}}) - \beta_1 \\ 1 & 1 \end{bmatrix} \\ \mathbf{G}^{-1} &= ((\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) - (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}}))^{-1} \begin{bmatrix} 1 & \beta_1 - (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}}) \\ -1 & (\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) - \beta_1 \end{bmatrix} \end{aligned}$$

and

$$\mathbf{H} = \begin{bmatrix} \pi_{01} & \pi_{01} & \pi_{01} & \pi_{02} & \pi_{02} & \pi_{02} & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

giving

$$\begin{aligned}\mathbf{G}^{-1}\mathbf{H} &= ((\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) - (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}}))^{-1} \\ &\quad \times \begin{bmatrix} \pi_1 & \pi_1 & \pi_1 & \pi_2 & \pi_2 & \pi_2 & -1 \\ -\pi_1 & -\pi_1 & -\pi_1 & -\pi_2 & -\pi_2 & -\pi_2 & 1 \end{bmatrix}. \end{aligned}$$

Also,

$$\boldsymbol{\Omega}_k = \begin{bmatrix} \mathrm{Var}[\bar{d}_{11}] & & & & & & & \\ & \mathrm{Var}[\bar{d}_{11\mathrm{T}}] & & & & & & \\ & & \mathrm{Var}[\bar{d}_{11\mathrm{D}}] & & & & & \\ & & & \mathrm{Var}[\bar{d}_{12}] & & & & \\ & & & & \mathrm{Var}[\bar{d}_{12\mathrm{T}}] & & & \\ & & & & & \mathrm{Var}[\bar{d}_{12\mathrm{D}}] & & \\ & & & & & & \mathrm{Var}[\bar{b}_1] & \end{bmatrix}$$

giving the 2-by-2 matrix

$$\mathbf{V} = \mathbf{G}^{-1}\mathbf{H}\boldsymbol{\Omega}_k(\mathbf{G}^{-1}\mathbf{H})^\top$$

with diagonal elements

$$\begin{aligned} v_{11} = v_{22} = {} & (\pi_{01}^2\mathrm{Var}[\bar{d}_{11}] + \pi_{01}^2\mathrm{Var}[\bar{d}_{11\mathrm{T}}] + \pi_{01}^2\mathrm{Var}[\bar{d}_{11\mathrm{D}}] \\ & + \pi_{02}^2\mathrm{Var}[\bar{d}_{12}] + \pi_{02}^2\mathrm{Var}[\bar{d}_{12\mathrm{T}}] + \pi_{02}^2\mathrm{Var}[\bar{d}_{12\mathrm{D}}] + \mathrm{Var}[\bar{b}_1]) \\ & \times((\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) - (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}}))^{-2} \end{aligned}$$

and off-diagonal elements

$$v_{12} = v_{21} = -v_{11}.$$

Note that $|\mathbf{G}| \neq 0$ implies that $(\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) \neq (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}})$, meaning that the mean discrimination-corrected source isotope ratios are different, which allows us to differentiate between sources. However, when they are close, $(\delta_{11} + \delta_{11\mathrm{T}} - \delta_{11\mathrm{D}}) \doteq (\delta_{12} + \delta_{12\mathrm{T}} - \delta_{12\mathrm{D}})$, then the difficulty of clearly differentiating between sources results in a large variance for $\hat{\underline{\pi}}$. △

The implicit function theorem gives the same result as the Delta method in (4.12), but in a matrix form that is easy to implement.

4.2.2.2 Confidence Intervals

Approximate $100(1-\alpha)\%$ confidence intervals for π_s can be constructed in many ways. One uses the normal distribution,

$$\hat{\pi}_s \pm z_{\alpha/2}\sqrt{\widehat{\mathrm{Var}}[\hat{\pi}_s]}, \quad s = 1, \ldots, S.$$

An alternative is the t interval using the Satterthwaite (1946) approximation for the degrees of freedom.

Example 4.2.10. CI for mixture and source variation dunlin carbon

A normal 95% CI is

$$\begin{array}{rcll} \hat{\pi}_1 & \pm & z_{0.025}\sqrt{\widehat{\mathrm{Var}}[\hat{\pi}_s]} & \\ 0.367 & \pm & 1.960 \times 0.00213 & \\ 0.367 & \pm & 0.091 \qquad \text{giving} & \qquad (0.277, 0.458). \end{array}$$

△

In summary, the implicit function theorem gives variances that match the complicated formulas of Phillips and Gregg (2001) for $S = 2$ and $S = 3$, but the variances are easy to compute for any dimension (S) using matrix multiplication. If sample sizes are not large enough for the asymptotic results to apply, then the large sample analysis may become misleading and the bootstrap or Bayesian analysis discussed later would be more appropriate.

4.2.3 Bootstrap inference

The bootstrap is a resampling method originally proposed by Efron (1979) for estimating properties of an estimator by sampling from an approximating distribution,

such as the empirical distribution of the observed data. The advantage of bootstrapping over analytical methods is its great simplicity. Bootstrap methods can be applied both when there is a completely specified probability model for the data and when there is not. In the case where a set of observations can be assumed to be from a random sample, a nonparametric bootstrap can be implemented by constructing a number of resamples of the observed dataset (and of equal size to the observed dataset), each of which is obtained by random sampling with replacement from the original dataset. A parametric boostrap can be implemented through a similar procedure where the resamples are drawn from parametrically estimated distributions. The bootstrap is often used as an alternative to inference based on parametric assumptions when those assumptions are in doubt, or where parametric inference is impossible or requires very complicated formulas for the calculation of standard errors (Davison and Hinkley, 1997; Efron and Tibshirani, 1993).

The large sample results in the previous sections are approximations relying on limiting arguments that sample sizes increase without bound. The elements of $\mathbf{A}$ and $\underset{\sim}{\beta}$ are estimated using independent samples with potentially different sample sizes. An implementation of the bootstrap will forego the large sample assumptions and use the empirical distributions of the data components or a probability model with sensible resampling for inference. While the bootstrap is often used in small sample situations, one can use it for large samples, as well.

Below we describe the parametric and nonparametric bootstrap in algorithmic terms to make it straightforward to implement. The core idea is to resample from either the estimated parametric distributions or the empirical distributions in accordance with the way the data were collected. For example, multivariate points are sampled as vectors, and this will be done by resampling with replacement. Theoretical considerations are provided elsewhere (Efron and Tibshirani, 1993; Davison and Hinkley, 1997).

4.2.3.1 Parametric Bootstrap

Suppose we have a parametric model for the distribution of the data components. This might be a multivariate normal distribution for the isotope ratio measurements of each source, where isotopes within sources and within the mixture may be correlated. Typically we can substitute the maximum likelihood estimates of the parameters in the parametric model to calculate approximate distributions for the statistic of interest, usually the sample mean, on which to base inferences about their estimands. From these fitted models, we draw a bootstrap repetition of the data simulation with replacement from the fitted models, then calculate the estimate of the parameter of interest R times, $\hat{\underset{\sim}{\pi}}^{(r)}$, $r = 1, \ldots, R$. From the bootstrap repetitions of $\hat{\underset{\sim}{\pi}}^{(r)}$ we can calculate moments, bias, confidence intervals, and other quantities of interest. Our primary interests will be standard error and confidence intervals.

Example 4.2.11. One-isotope, two-source example ($I = 1, S = 2$)
We illustrate the simplest example without discrimination, $\underset{\sim}{\Delta}$, though generalizations are immediate. For the mixture and sources, R bootstrap repetitions of the data simulation will be drawn from the fitted distributions, for example,

$$\begin{aligned} b_{j1}^{(r)} &\sim \text{Normal}(\bar{b}_1, s_b^2),\ j = 1, \ldots, J, \\ d_{11k}^{(r)} &\sim \text{Normal}(\bar{d}_{11}, s_1^2),\ k = 1, \ldots, K_1, \text{ and} \\ d_{12k}^{(r)} &\sim \text{Normal}(\bar{d}_{12}, s_2^2),\ k = 1, \ldots, K_2. \end{aligned}$$

For each $r = 1, \ldots, R$, the mean isotope ratios are calculated

$$\bar{b}_1^{(r)} = \frac{1}{J}\sum_{j=1}^{J} b_{j1}^{(r)}, \quad \bar{d}_{11}^{(r)} = \frac{1}{K_1}\sum_{k=1}^{K_1} d_{11k}^{(r)}, \text{ and } \quad \bar{d}_{12}^{(r)} = \frac{1}{K_2}\sum_{k=1}^{K_2} d_{12k}^{(r)}$$

and substituted into the defining equation

$$\begin{bmatrix} \bar{b}_1^{(r)} \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{d}_{11}^{(r)} & \bar{d}_{12}^{(r)} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \hat{\pi}_1^{(r)} \\ 1 - \hat{\pi}_1^{(r)} \end{bmatrix}, \tag{4.26}$$

and evaluated to give

$$\hat{\underset{\sim}{\pi}}^{(r)} = \hat{\mathbf{A}}_{\mathrm{x}}^{(r)-1}\bar{\underset{\sim}{b}}_{\mathrm{x}}^{(r)}.$$

The $\hat{\underset{\sim}{\pi}}^{(r)}$, $r = 1, \ldots, R$, give an estimate of the bootstrap distribution of $\hat{\underset{\sim}{\pi}}$ which can be used for inference on $\underset{\sim}{\pi}$. The distribution will be roughly centered at the estimated value from the data, $\hat{\underset{\sim}{\pi}}$. The standard error for $\hat{\underset{\sim}{\pi}}$ is calculated as the standard deviation of the bootstrap replicates. △

4.2.3.2 Nonparametric Bootstrap

Suppose we are unwilling to assume a parametric model, but that it is sensible to assume that for each data component the observations are independent and indentically distributed according to an unknown distribution function. We use the empirical distribution (ED) function to estimate the unknown distribution function. The ED puts equal probability on each observation. We draw a bootstrap repetition of the data simulation with replacement from the ED, then calculate the estimate of the parameter of interest, $\hat{\underset{\sim}{\pi}}^{(r)}$, $r = 1, \ldots, R$.

Example 4.2.12. One-isotope, two-source example ($I = 1, S = 2$)
For the mixture and sources, R bootstrap repetitions of the data simulation will be drawn from the empirical distributions,

$$\begin{aligned} b_{j1}^{(r)} &\sim \mathrm{ED}(b_{j1};\ j = 1, \ldots, J), j = 1, \ldots, J, \\ d_{11k}^{(r)} &\sim \mathrm{ED}(d_{11k};\ k = 1, \ldots, K_1), k = 1, \ldots, K_1, \text{ and} \\ d_{12k}^{(r)} &\sim \mathrm{ED}(d_{12k};\ k = 1, \ldots, K_2), k = 1, \ldots, K_2. \end{aligned}$$

The remaining steps proceed as in the parametric bootstrap. △

4.2.3.3 Bootstrap Confidence Intervals

An extensive menu of bootstrap confidence intervals is provided in Davison and Hinkley (1997, chap. 5). Here we use the simplest and most intuitive of all the intervals, the basic bootstrap CI, taking the interval bounds as the bootstrap estimates of the percentiles. For example, the central $100(1-\alpha)\%$ CI is estimated with the R sorted bootstrap repetitions of the data simulation by selecting the bootstrap order statistics in positions $R\alpha/2$ and $R(1-\alpha/2)$, rounded to the extremes to be conservative. Thus, for a 95% interval and for $R = 10000$, the 250th and 9751th sorted bootstrap repetitions of the data simulation would designate the lower and upper confidence bounds.

Example 4.2.13. Parametric bootstrap dunlin carbon
The sample size and mean and variance MLEs are provided in Tables 4.1, 4.2, and 4.3. Plots of the data in Figures 4.2, 4.3, and 4.4 do not severely contradict the use of normal distributions for the data components so the normal distribution will be used for illustrative purposes.

R bootstrap repetitions of the data simulation are drawn from fitted distributions for the mixture, sources, and diet and blood that determine discrimination,

$$
\begin{aligned}
b_{j1}^{(r)} &\sim \text{Normal}(\bar{b}_1, s_b^2) = \text{Normal}(-16.81, 10.27),\ j = 1, \ldots, J = 234, \\
d_{11k}^{(r)} &\sim \text{Normal}(\bar{d}_{11}, s_1^2) = \text{Normal}(-25.36, 1.724),\ k = 1, \ldots, K_1 = 16, \\
d_{12k}^{(r)} &\sim \text{Normal}(\bar{d}_{12}, s_2^2) = \text{Normal}(-13.98, 10.55),\ k = 1, \ldots, K_2 = 21, \\
d_{1\text{T}k}^{(r)} &\sim \text{Normal}(\bar{d}_{1\text{T}}, s_{1\text{T}}^2) = \text{Normal}(-24.63, 0.145),\ k = 1, \ldots, K_\text{T} = 28, \text{ and} \\
d_{1\text{D}k}^{(r)} &\sim \text{Normal}(\bar{d}_{1\text{D}}, s_{1\text{D}}^2) = \text{Normal}(-23.28, 0.105),\ k = 1, \ldots, K_\text{D} = 4.
\end{aligned}
$$

For each $r = 1, \ldots, R$, the mean isotope ratios are calculated

$$\begin{aligned} \bar{b}_1^{(r)} &= \frac{1}{J}\sum_{j=1}^{J} b_{j1}^{(r)}, \quad \bar{d}_{11}^{(r)} = \frac{1}{K_1}\sum_{k=1}^{K_1} d_{11k}^{(r)}, \quad \bar{d}_{12}^{(r)} = \frac{1}{K_2}\sum_{k=1}^{K_2} d_{12k}^{(r)} \\ \bar{d}_{1\mathrm{T}}^{(r)} &= \frac{1}{K_\mathrm{T}}\sum_{k=1}^{K_\mathrm{T}} d_{1\mathrm{T}k}^{(r)}, \text{ and } \quad \bar{d}_{1\mathrm{D}}^{(r)} = \frac{1}{K_\mathrm{D}}\sum_{k=1}^{K_\mathrm{D}} d_{1\mathrm{D}k}^{(r)}. \end{aligned}$$

The discrimination is calculated, $\hat{\Delta}_{11}^{(r)} = \bar{d}_{1\mathrm{T}}^{(r)} - \bar{d}_{1\mathrm{D}}^{(r)}$. We assume that the discrimination is the same for both sources, $\underset{\sim}{\Delta}_1 = \underset{\sim}{\Delta}_2$, thus we use $\hat{\Delta}_{12}^{(r)} \equiv \hat{\Delta}_{11}^{(r)}$. The means $\hat{\delta}_{1s}^{\prime(r)}$ are substituted for $\bar{d}_{1s}^{(r)}$ in the defining relation (4.26) with

$$\begin{aligned} \hat{\delta}_{11}^{\prime(r)} &= \bar{d}_{11}^{(r)} + \bar{d}_{1\mathrm{T}}^{(r)} - \bar{d}_{1\mathrm{D}}^{(r)} \\ \hat{\delta}_{12}^{\prime(r)} &= \bar{d}_{12}^{(r)} + \bar{d}_{1\mathrm{T}}^{(r)} - \bar{d}_{1\mathrm{D}}^{(r)}. \end{aligned}$$

Using $R = 10000$ bootstrap repetitions of the data simulation, we estimate $\mathrm{Var}[\hat{\pi}_1]$ with the sample variance of the $\hat{\pi}_1^{(r)}$, 0.002227, giving a standard deviation of 0.0472. The 95% basic bootstrap CI is $(0.266, 0.452)$. This interval is similar to the one given with the implicit function method. △

Example 4.2.14. Nonparametric bootstrap dunlin carbon

We will resample the observations plotted in Figures 4.2, 4.3, and 4.4. R bootstrap repetitions of the data simulation are drawn from the empirical distributions for the mixture, sources, and diet and blood that determine discrimination. The rest of the procedure is as in the parametric bootstrap. Using $R = 10000$ bootstrap repetitions of the data simulation, the estimate of $\mathrm{Var}[\hat{\pi}_1]$ is 0.002625, and $\mathrm{SD}[\hat{\pi}_1]$ is 0.0512, with a 95% CI of $(0.255, 0.455)$. This nonparametric bootstrap CI is slightly wider than the parametric bootstrap CI, and the estimated sampling distribution for $\hat{\pi}_1$ a little left-skewed. △

4.2.3.3.1 Dunlin discussion Numerical and graphical summaries of the dunlin data of the five methods discussed in this chapter are in Table 4.5 and Figure 4.6. The

implicit function theorem (delta method) normal approximation appears consistent with the bootstrap methods, in spite of skewed data distributions and small sample sizes. The parametric and nonparametric bootstraps perform similarly in this case, though the nonparametric bootstrap suggests slight left-skewness in the sampling distribution of $\hat{\pi}_1$, a result of skewness in the data distribution.

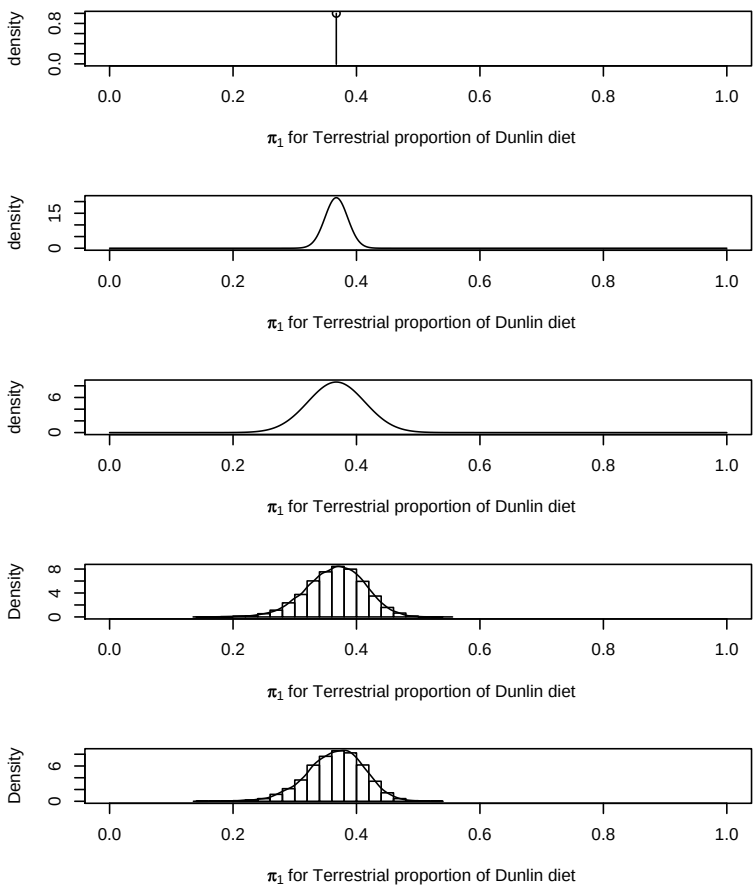

Figure 4.6: Frequentist estimates for proportion contribution of terrestrial sources to dunlin diet, π_1, corresponding to Table 4.5. The first is the point mass estimated without variability. The second and third distributions correspond to large sample normal approximations for variation in the mixture only and in the mixture and sources, respectively. The fourth and fifth distributions are for the parametric and nonparametric bootstraps, respectively.

Table 4.5: Frequentist implicit representation estimates for proportion contribution of terrestrial sources to the dunlin diet with central 95% CI from methods in Section 4.2: no variation, large sample variation in mixture only, large sample variation in mixture and sources, parametric bootstrap variation in mixture and sources, and nonparametric bootstrap variation in mixture and sources.

Method	$\hat{\pi}_1$	$\text{Var}[\hat{\pi}_1]$	$\text{SD}[\hat{\pi}_1]$	95% CI
No variation	0.367			
Mixture var only	0.367	0.000339	0.0184	(0.331, 0.403)
Mixture and sources var	0.367	0.002132	0.0462	(0.277, 0.458)
Par Bootstrap	0.367	0.002227	0.0472	(0.266, 0.452)
NP Bootstrap	0.367	0.002625	0.0512	(0.255, 0.455)

4.3 Bear three-source example

This section introduces the regression-based discrimination-corrected isotope ratio and its incorporation in the model. Additionally, this three-source bear example illustrates several things. There is often very little data on the sources, and those data can have a great deal of variability. The uncertainty of the source estimates can make the estimates of the diet proportion $\underset{\sim}{\pi}$ unreliable. The bootstrap is a practical reality-check for the large sample estimation when sample sizes are small, though bootstrap methods themselves are supported by large sample arguments, in particular that the empirical distribution of the sample captures the important features of the population distribution. While the results of this example are not remarkable, the detailed illustration of the methods will be useful for the practitioner, especially those in Section 4.3.1.3 using regression to estimate discrimination-corrected isotope ratios.

Vulla et al. (2009) discuss seasonal effects of a brown bear's diet in Estonia and central and northern Europe. Because berries, cereals, and insects (ants) have previously been identified as being important diet sources in the autumn when bears grow hair, we look at this three-source subset of the dataset appearing in their paper, though expanded by Keith Hobson (personal communication). Using three sources and two isotopes (carbon and nitrogen) we have a perfectly-constrained case resulting in a unique solution.

4.3.1 Bear Data

4.3.1.1 Mixture isotope ratios

Summary isotope ratio values for the $J = 42$ bear hair measurements are

$$\underset{\sim}{\bar{b}} = \begin{bmatrix} -22.74 \\ 5.36 \end{bmatrix} \quad \text{and} \quad \mathbf{S}_{\underset{\sim}{b}} = \begin{bmatrix} 0.311 & 0.042 \\ 0.042 & 2.740 \end{bmatrix},$$

and summarized in Table 4.6. Figure 4.7 illustrates that the correlation of the carbon and nitrogen isotope ratios for hair is not statistically significantly different from zero ($r = 0.046$, p-value $= 0.775$), and the data are slightly right-skewed for nitrogen.

Table 4.6: Bear hair summaries for carbon ($i = 1$) and nitrogen ($i = 2$).

Isotope ratios	J	$\bar{b}_1$	SD	SE($\bar{b}_1$)	$\bar{b}_2$	SD	SE($\bar{b}_2$)	r
Bear hair	42	-22.74	0.557	0.0860	5.357	1.655	0.2554	0.046

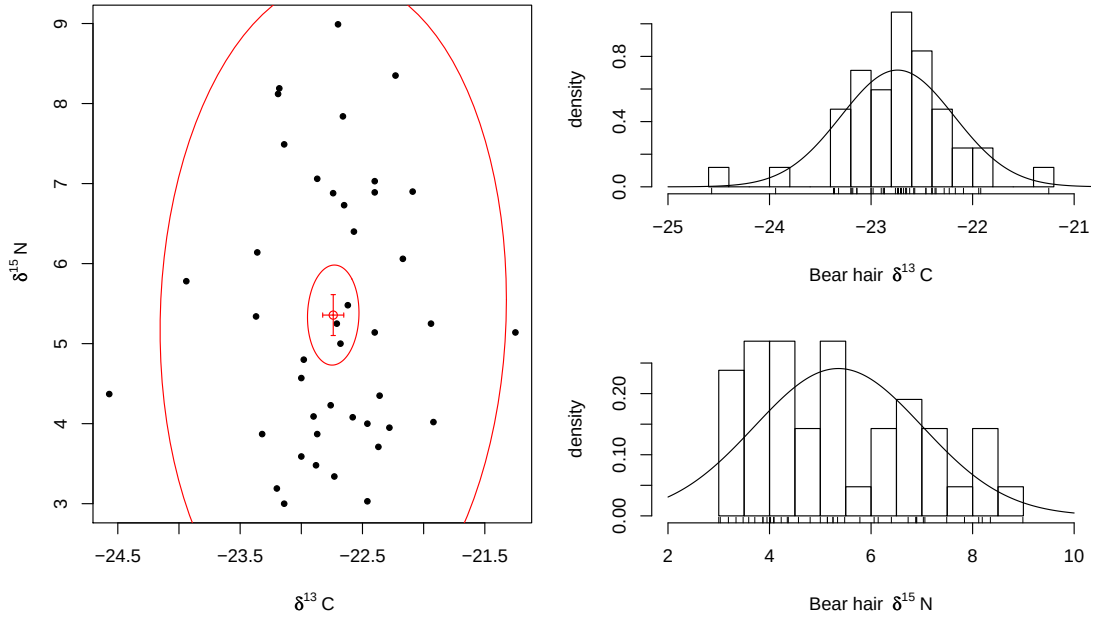

Figure 4.7: Bear hair $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate normal 95% confidence and prediction ellipses with marginal distributions with estimated normal distributions.

4.3.1.2 Source isotope ratios

Sample isotope ratio values for source measurements are summarized in Table 4.7. The only statistically significantly correlated isotope values are for insects (p-value $<$ 0.0001). Figure 4.8 also shows that the carbon and nitrogen isotope ratios for most sources have a great deal of variability and overlap, and the sample sizes are small.

Table 4.7: Source carbon ($i = 1$) and nitrogen ($i = 2$) isotope ratios.

		Carbon			Nitrogen			
Sources	K_s	$\bar{d}_{1s}$	SD	SE($\bar{d}_{1s}$)	$\bar{d}_{2s}$	SD	SE($\bar{d}_{2s}$)	r
Berries	4	-31.93	2.073	1.0363	-3.65	4.188	2.0938	-0.738
Cereals	2	-28.30	0.283	0.2000	3.45	1.202	0.8500	N/A
Insects	9	-24.75	1.478	0.4927	2.95	0.945	0.3149	0.985

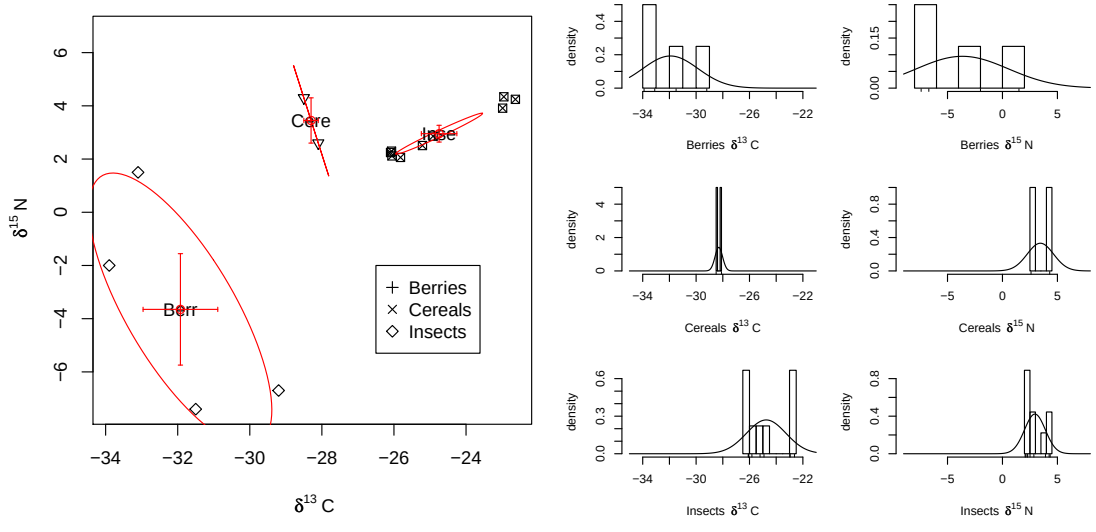

Figure 4.8: Source $\delta^{13}C$ and $\delta^{15}N$ observations with SE intervals of the mean and bivariate normal 95% confidence ellipses with marginal distributions with estimated normal distributions.

4.3.1.3 Discrimination-corrected isotope ratios using regression

If a diet experiment is conducted for a range of sources then it is possible to model the discrimination-corrected isotope ratios for all sources together using a regression. This experiment is conducted independently from the data used for the diet analysis. Let $\underset{\sim}{d}_{k\mathrm{D}}$ be the isotope ratio of the diet and $\underset{\sim}{d}'_{k\mathrm{T}}$ be the isotope ratio of the animal tissue for diet/tissue pair $k = 1, \ldots, K_{i\mathrm{reg}}$. For simplicity assume that the relationship between diet and tissue for isotope i is linear. Then a simple linear regression,

$$d'_{ik\mathrm{T}} \;=\; \gamma_{0i} + \gamma_{1i} d_{ik\mathrm{D}} + \varepsilon_{ik},$$

can be used to estimate the discrimination or directly predict the discrimination-corrected isotope ratio value for a given source/tissue pair. Given the estimated regression coefficients, $\underset{\sim}{\hat{\gamma}}_i = [\hat{\gamma}_{0i}, \hat{\gamma}_{1i}]^\top$, $i = 1, \ldots, I$, the discrimination-corrected isotope ratios, $\delta'_{is} = \delta_{is} + \Delta_{is}$, can be predicted for each source in the analysis using

$$\hat{\delta}'_{is} \;=\; \hat{\gamma}_{0i} + \hat{\gamma}_{1i} \bar{d}_{is}.$$

This relationship defines the predicted discrimination,

$$\hat{\Delta}_{is} \;\equiv\; \hat{\delta}'_{is} - \bar{d}_{is} = \hat{\gamma}_{0i} + (\hat{\gamma}_{1i} - 1)\bar{d}_{is}.$$

While this method has been used previously to estimate the discrimination (Felicetti et al., 2003), the variation associated with this estimate has not previously been incoporated into the analysis.

Using the implicit-function theorem from Section 4.2.2.1, we can estimate the variance for $\hat{\underset{\sim}{\pi}}$ incorporating the isotope ratio variance and regression variance. Using the three-source, two-isotope example, we detail the estimation procedure. In the implicit-function theorem we set the regression-based model matrix from (1.5) equal to zero as $\underset{\sim}{0} = \mathbf{B}_{\mathrm{x}} \underset{\sim}{\pi} - [0, 0, 1]^\top$,

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma_{01}+\gamma_{11}\delta_{11}-\beta_1 & \gamma_{01}+\gamma_{11}\delta_{12}-\beta_1 & \gamma_{01}+\gamma_{11}\delta_{13}-\beta_1 \\ \gamma_{02}+\gamma_{12}\delta_{21}-\beta_2 & \gamma_{02}+\gamma_{12}\delta_{22}-\beta_2 & \gamma_{02}+\gamma_{12}\delta_{23}-\beta_2 \\ 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (4.27)$$

The functions $g_i(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y)$ come from the rows of (4.27),

$$\begin{aligned} g_1(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) &= (\gamma_{01}+\gamma_{11}\delta_{11}-\beta_1)\pi_1 + (\gamma_{01}+\gamma_{11}\delta_{12}-\beta_1)\pi_2 + (\gamma_{01}+\gamma_{11}\delta_{13}-\beta_1)\pi_3 \\ g_2(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) &= (\gamma_{02}+\gamma_{12}\delta_{21}-\beta_2)\pi_1 + (\gamma_{02}+\gamma_{12}\delta_{22}-\beta_2)\pi_2 + (\gamma_{02}+\gamma_{12}\delta_{23}-\beta_2)\pi_3 \\ g_3(\underset{\sim}{\pi}, \underset{\sim}{\mu}_y) &= 0. \end{aligned}$$

Let

$$\underset{\sim}{\mu}_y = [\delta_{11}, \delta_{21}, \delta_{12}, \delta_{22}, \delta_{13}, \delta_{23}, \beta_1, \beta_2, \gamma_{01}, \gamma_{11}, \gamma_{02}, \gamma_{12}],$$

which defines $\mathbf{G} = \mathbf{B}_{\mathrm{x}}$ and

$$\mathbf{H} = \begin{bmatrix} \gamma_{11}\pi_1 & 0 & \gamma_{11}\pi_2 & 0 & \gamma_{11}\pi_3 & 0 & -1 & 0 \\ 0 & \gamma_{12}\pi_1 & 0 & \gamma_{12}\pi_2 & 0 & \gamma_{12}\pi_3 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 & \delta_{11}\pi_1+\delta_{12}\pi_2+\delta_{13}\pi_3 & 0 & 0 \\ 0 & 0 & 1 & \delta_{21}\pi_1+\delta_{22}\pi_2+\delta_{23}\pi_3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The covariance for $\underset{\sim}{y}_k = [\bar{d}_{11}, \bar{d}_{21}, \bar{d}_{12}, \bar{d}_{22}, \bar{d}_{13}, \bar{d}_{23}, \bar{b}_1, \bar{b}_2, \hat{\gamma}_{01}, \hat{\gamma}_{11}, \hat{\gamma}_{02}, \hat{\gamma}_{12}]$ is

$$\mathbf{\Omega}_k = \begin{bmatrix} \mathrm{Cov}[\underset{\sim}{\bar{d}}_1] & & & & & \\ & \mathrm{Cov}[\underset{\sim}{\bar{d}}_2] & & & & \\ & & \mathrm{Cov}[\underset{\sim}{\bar{d}}_3] & & & \\ & & & \mathrm{Cov}[\underset{\sim}{\bar{b}}] & & \\ & & & & \mathrm{Cov}[\underset{\sim}{\hat{\gamma}}_1] & \\ & & & & & \mathrm{Cov}[\underset{\sim}{\hat{\gamma}}_2] \end{bmatrix}.$$

Thus, the estimated covariance for $\underset{\sim}{\hat{\pi}}$ is

$$\mathbf{V} = \mathbf{G}^{-1}\mathbf{H}\mathbf{\Omega}_k(\mathbf{G}^{-1}\mathbf{H})^\top,$$

which is estimated by substituting the natural estimates in the matrices above. This is easily extended to any problem size.

4.3.1.4 Bear Discrimination

Felicetti et al. (2003) performed a feeding trial in six (three male and three female siblings) 16- to 21-month-old grizzley bears housed at the Washington State University Bear Research, Education, and Conservation Facility in Pullman, Washington. Each bear was fed each of 5 diets (whitebark pine nuts, other plant matter, army cutworm moths, ungulates, cutthroat trout) for 21 days and blood-plasma samples were collected at the end of each trial. Plasma isotope ratios equilibrate (i.e., become asymptotic) with the diet within 10–14 days and have the same isotope ratios as hair (Hilderbrand et al., 1996; Hildebrand et al., 1999). Diet samples were collected daily and pooled and homogenized at the end of the trial for isotopic analyses. We used datathief (Tummers, 2006) to extract the paired diet/plasma isotope ratio values of carbon and nitrogen from the regression plots in Felicetti et al. (2003), but we do not know which measurements belong to which bears so we assume the

measurements are independent. The plots in Figure 4.9 show the data used to fit the regressions for carbon and nitrogen from the bear feeding trial experiment. The relationship between the points used in the carbon and nitrogen regressions is unknown, thus a separate estimation is done for carbon and nitrogen. Because the source isotope ratios and bear plasma isotope ratios show a relationship, the estimated discrimination-corrected isotope ratios can be estimated by the regression line. Table 4.8 summarizes the regression relationships, and the regression coefficient covariance matrices for carbon and nitrogen are

$$\widehat{\text{Cov}}[\underset{\sim}{\hat{\gamma}}_1] = \begin{bmatrix} 3.8184 & 0.1679 \\ 0.1679 & 0.0075 \end{bmatrix} \quad \text{and} \quad \widehat{\text{Cov}}[\underset{\sim}{\hat{\gamma}}_2] = \begin{bmatrix} 0.0619 & -0.0063 \\ -0.0063 & 0.0010 \end{bmatrix}.$$

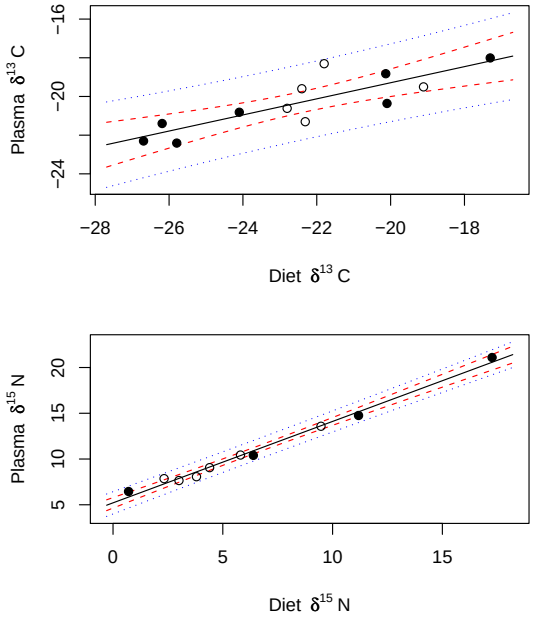

Figure 4.9: Diet experiment results for the relationship between diet source and bear blood plasma isotope ratios (Felicetti et al., 2003) with confidence and prediction intervals, where the difference gives the estimated discrimination. Solid points are single-species diets, and open circles are mixed diets.

Table 4.8: Regression summaries for the relationship between diet and blood plasma in Figure 4.9.

$\delta^{13}C$	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	−10.96675	1.95407	−5.612	0.000224
Diet d13C	0.41617	0.08658	4.807	0.000716
Residual standard error: 0.8436 on 10 degrees of freedom				
Multiple R-squared: 0.6979				

$\delta^{15}N$	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	5.21145	0.24872	20.95	2.83e−08
Diet d15N	0.88966	0.03119	28.53	2.47e−09
Residual standard error: 0.4647 on 8 degrees of freedom				
Multiple R-squared: 0.9903				

4.3.1.5 Mixture and discrimination-corrected source ensemble

For this example we consider the proportional contributions to the bear's diet reflected through the bear hair of the three sources of berries, cereals, and insects. In Table 4.9 are the estimated discrimination-corrected source isotope ratios, $\hat{\delta}'_{is} = \hat{\gamma}_{0i} + \hat{\gamma}_{1i}\bar{d}_{is}$. The discrimination-corrected isotope ratios in Figure 4.10 show that the bear falls inside the convex hull of the three considered sources, thus a unique solution exists.

Table 4.9: European bear example, estimated mixture and regression-based discrimination-corrected source isotope ratios for carbon ($i = 1$) and nitrogen ($i = 2$).

Estimates	$\bar{b}_1$	$\bar{b}_2$
Bear hair	-22.74	5.357

Estimates	$\hat{\delta}'_{1s}$	$\hat{\delta}'_{2s}$
Berries	-24.25	1.964
Cereals	-22.74	8.281
Insects	-21.27	7.839

4.3.2 Analysis

We consider the proportional contributions of berries ($s = 1$), cereals ($s = 2$), and insects ($s = 3$) in the bear's diet.

Example 4.3.1. No variation bear

Assume the estimated values for $\underset{\sim}{\beta}$ and $\mathbf{A}$ are the true values. We can solve for $\underset{\sim}{\pi}$ by plugging in the values

$$\mathbf{A}_{\mathrm{x}} = \begin{bmatrix} -24.25 & -22.74 & -21.27 \\ 1.96 & 8.28 & 7.84 \\ 1 & 1 & 1 \end{bmatrix}$$

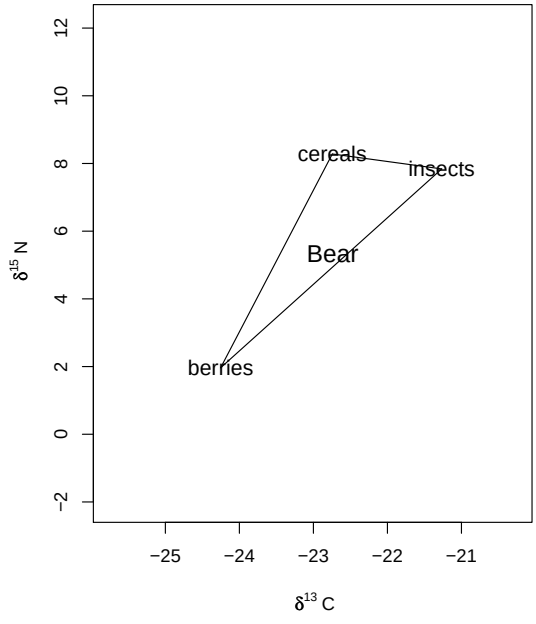

Figure 4.10: European bear example, discrimination-corrected isotope ratios for the three sources. The bear hair is contained in the convex hull of berries, cereals, and insects.

and

$$\underset{\sim}{\beta}_{\mathrm{x}} = \begin{bmatrix} -22.74 \\ 5.36 \\ 1 \end{bmatrix} \qquad \text{giving} \qquad \pi = \mathbf{A}_{\mathrm{x}}^{-1}\underset{\sim}{\beta}_{\mathrm{x}} = \begin{bmatrix} 0.432 \\ 0.124 \\ 0.444 \end{bmatrix}$$

as the true proportional contributions of berries, cereals, and insects bear diet. △

In subsequent examples the estimate of π is identical to the estimate above and the variability of the estimate will depend on the assumption of what is known.

Example 4.3.2. Mixture variation bear

Here we assume that the only source of variation is in the mixture and we assume the sources are fixed as in Section 4.2.1. The estimated mixture covariance matrix is

$$\frac{\underset{\sim}{\Sigma}_{b_{\mathrm{x}}}}{J} = \frac{1}{42}\begin{bmatrix} 0.311 & 0.042 & 0 \\ 0.042 & 2.740 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

giving

$$\hat{\underset{\sim}{\Sigma}}_{\hat{\pi}} = \begin{bmatrix} 0.00145 & -0.00261 & 0.00116 \\ -0.00261 & 0.00801 & -0.00540 \\ 0.00116 & -0.00540 & 0.00424 \end{bmatrix}.$$

Table 4.10 and Figure 4.11 show the estimated proportions with estimated variances and standard deviations. △

Example 4.3.3. Mixture variation parametric bootstrap bear

This example uses the bootstrap similar as in Section 4.2.3.1, but with variation only in the mixture. This bootstrap result is similar to the large sample result, and is given in Table 4.10 and Figure 4.11. △

Example 4.3.4. Mixture and source variation bear

Here we assume uncertainty in both the mixture and the sources. Using the implicit-function method in Section 4.3.1.3 the covariance for $\hat{\underset{\sim}{\pi}}$ is

$$\hat{\mathbf{V}} = \begin{bmatrix} 0.01683 & -0.03280 & 0.01597 \\ -0.03280 & 0.23311 & -0.20032 \\ 0.01597 & -0.20032 & 0.18435 \end{bmatrix}. \tag{4.28}$$

Table 4.10 and Figure 4.11 show the estimated proportions with estimated variances and standard deviations. △

While there are several ways to implement the regression bootstrap (Ch. 6 Davison and Hinkley, 1997), it is clear in this example due to the great variability in the sources that inference about the diet will not be better informed by implementing it here.

4.3.2.0.1 Bear discussion A summary of the bear data is in Table 4.10 and Figure 4.11. The sample sizes are very small for the sources so this is a case where the large sample approximation is unlikely to perform well. When using variation in both the mixture and the sources the estimated standard deviations for $\hat{\underset{\sim}{\pi}}$, especially for sources 2 and 3 (cereals or insects), are both greater than 0.4, making the estimates noninformative for $\underset{\sim}{\pi}$. This example highlights, however, the use of regression in estimating the discrimination-corrected isotope ratios and accounting for variation using the implicit function theorem.

4.4 Discussion

In this chapter we showed that the implicit function theorem is an easy way to implement the delta method for frequentist large sample inference for the estimated

Table 4.10: Frequentist implicit representation estimates for proportion contribution of sources to the bear diet, $[\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3]^\top = [0.432, 0.124, 0.444]^\top$, with central 95% CI, which would be truncated to $[0, 1]$ in practice. The three cases illustrate large sample variation in mixture only, parametric bootstrap with variation in mixture only, and large sample variation in mixture and sources.

Method	s	$\text{Var}[\hat{\pi}_s]$	$\text{SD}[\hat{\pi}_s]$	95% CI
Mixture var only	1	0.00145	0.038	0.357, 0.507
	2	0.00801	0.090	−0.052, 0.299
	3	0.00424	0.065	0.317, 0.572
Mixture var only Par BS	1	0.00144	0.038	0.358, 0.506
	2	0.00782	0.088	−0.050, 0.300
	3	0.00414	0.064	0.318, 0.571
Mixture and sources var	1	0.01683	0.130	0.178, 0.686
	2	0.23311	0.483	−0.822, 1.070
	3	0.18435	0.429	−0.397, 1.286

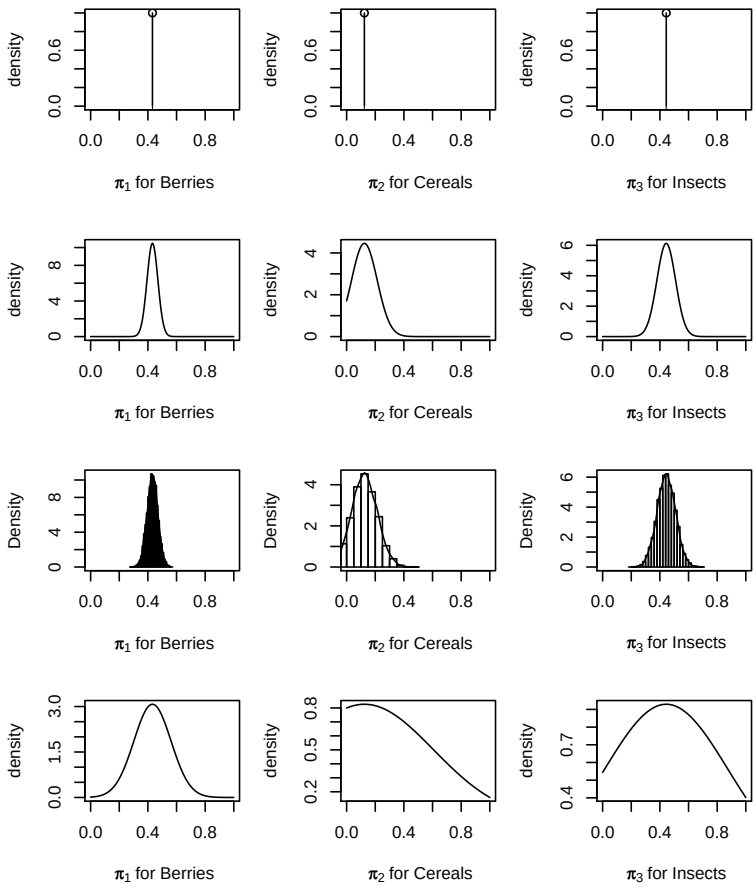

Figure 4.11: Frequentist estimates for proportion contribution of berries, cereals, and insects to the bear diet. The four cases illustrate the point estimate with no variation, large sample variation in mixture only, parametric bootstrap with variation in mixture only, and large sample variation in mixture and sources.

diet proportion parameter vector. Whether or not distributional assumptions are made, the bootstrap can be used for frequentist resampling inference. The nonparametric bootstrap does not require a complete parametric model specification, which is important both as a check of the adequacy of the assumptions in the large sample methods and as an alternative method of inference.

An important criticism of the large sample and bootstrap frequentist methods is that our frequentist estimates of $\underset{\sim}{\pi}$ are not inherently constrained to $[0, 1]$. For example, in the bear example with variation only in the mixture, a small area of the estimated sampling distribution of π_2 is less than zero, and a substantial area is less than zero with variation also in the sources. Therefore, estimates near the boundary may have a potentially substantial amount of sampling distribution outside the support of the parameter. The sampling distribution for $\hat{\underset{\sim}{\pi}}$ can be truncated to the interval $[0, 1]$, for example, with

$$\hat{\underset{\sim}{\pi}}_{\text{new}} \;=\; 0 \times 1_{(\hat{\underset{\sim}{\pi}}<0)} + \hat{\underset{\sim}{\pi}} \times 1_{(0\leq\hat{\underset{\sim}{\pi}}\leq 1)} + 1 \times 1_{(\hat{\underset{\sim}{\pi}}>1)},$$

but this post-hoc correction is not contextually satisfying. Likelihood methods (including Bayesian) restrict the parameter support and may eliminate this difficulty.

Chapter 5

Bayesian methods, BMM

This chapter illustrates how a Bayesian approach based on first principles can be applied to jointly model the mean mixture population diet, subject-specific mixture diets, source and discrimination compositions, measurement error, and other parameters of interest. Here we provide statistical justification for the methods in Chapter 3 previously only heuristically justified, e.g., IsoSource.

Adopting the Bayesian paradigm has several advantages (Gelman et al., 1995, Ch. 9). First, Bayesian methods incorporate all reasonable sources of uncertainty in inferential summaries by incorporating multiple levels of randomness to combine information from different sources. Therefore, for the best inference, the statistician must have access to all available information, rather than data summaries, and work closely with experts to incorporate nonmeasurement data (via elicitation of priors). Second, it is natural for practitioners to interpret interval estimates as Bayesian intervals, that is, as probability statements about the likely values of unknown quantities conditional on the evidience in the data. These probability statements require full probability models, including prior probability specifications for unknown quantities. Third, Bayesian inferences in their attempt to represent complicated real-world relationships are conditional on probability models that invariably contain approxima-

tions. If the Bayesian inferences differ greatly over a range of scientifically reasonable assumptions that are not disputable by the data, then the corresponding range of conclusions must be considered possible.

There are problems in the previous applications of Bayesian methods to mass-balance mixing models. Moore and Semmens (2008) have many notational issues in the description of their MixSIR model, making it a challenge to understand their model precisely, but here is a summary. (0) The name comes from using the diet mixing model with the Hilborn SIR sampling method (Rubin, 1988). (1) For the diet prior they assume independent beta distributions for the individual source proportions, $\pi_s|\alpha_{\underset{\sim}{\pi}s}, \beta_{\underset{\sim}{\pi}s} \sim \text{Beta}(\alpha_{\underset{\sim}{\pi}s}, \beta_{\underset{\sim}{\pi}s})$, $s = 1, \ldots, S$, but they do not overtly incorporate the constraint $1 = \sum_{s=1}^{S} \pi_s$ in their prior. (2) They ignore the uncertainty from estimating $\mathbf{A}$. (3) They fail to detail priors on other parameters and they use sample statistics in place of population parameters for the distribution of the observed mixture isotope ratios. Jackson et al. (2009) have criticized the beta priors Moore and Semmens (2008) specify on $\underset{\sim}{\pi}$, and found in a simulation study that the MixSIR model failed to identify true dietary proportions more than 50% of the time. The authors of the criticism have written their own software as the R package Stable Isotope Analysis in R (SIAR) (Parnell and Jackson, 2008), and while addressing the issue with the prior on $\underset{\sim}{\pi}$, (and modifying the sampling from using Hilborn SIR to the Metropolis-Hastings algorithm (Chib and Greenberg, 1995)) they still seemingly ignore issues (2) and (3).

In a frequentist analysis of the perfectly-constrained case ($S = I + 1$) the delta method has been applied to the mean structure in (2.2) (Chapters 4 and 6). In this case the mean mixture population diet $\underset{\sim}{\pi}$ is being modeled indirectly based on the isotope ratio means of the mixture in $\underset{\sim}{\beta}$ and the sources in $\mathbf{A}$. We believe that a direct approach is more appealing, whereby each mixture's isotope ratio $\underset{\sim}{\beta}_j$ is conditional on the mixture's diet $\underset{\sim}{\pi}_j$ and the common sources $\mathbf{A}$, as introduced in Section 2.1.

In the underconstrained case ($S > I + 1$) exact feasible-solutions of the mean

diet $\underset{\sim}{\pi}$ have been provided with the probabilistic algorithm introduced in Chapter 3. Because no variation is considered, the defining structure for the mean is the only model component, and therefore this strategy can be seen as an approximate Bayes procedure of our implicit representation model below without variation and a uniform prior distribution on $\underset{\sim}{\pi}$. Alternatively, Lubetkin and Simenstad (2004) introduce two models, SOURCE and STEP, which use a linear programming technique to give results similar to IsoSource (Phillips and Gregg, 2003), but instead of describing the volume of the solution polytope, just the outer bounds of the possible diet proportions are given.

Two models for the diet components are presented. The implicit representation is consistent with what has been done thus far, modeling the distribution of mixture isotope ratio observations centered at the mean isotope ratio value conditional on the average diet (recall Section 2.2.1). The explicit representation is new, modeling the mixture isotope ratio means conditional on the distribution of diets centered at the average diet (recall Section 2.2.2). We show that the explicit representation model supplemented with measurement error information provides subject-specific diets and typically more precise estimates of the mean diet, if the model is consistent with the data. The insight was to approach the problem from a random effects perspective, rather than only estimating fixed effects.

The following sections focus on the two models when sufficient data and/or prior information is available to inform all the model components. First, we model the $\mathbf{A}$ matrix of source information including discrimination, develop the Bayesian implicit representation model, and show how the model of the $\mathbf{A}$ matrix can be simplified in various ways. Next, we develop the Bayesian explicit representation model and the incorporation of measurement error. Then we illustrate the robustness of inference of the implicit representation and explicit representation when certain model violations are present. In the last section we illustrate both models using the mink example, a case when data are limited.

5.1 Modeling source data components

Here we provide a model for each of the source data components, the mixtures and diets are modeled later either in the implicit representation or the explicit representation models. We assume that sources are mutually independent. The distributions selected below are for convenience and are subject to refinement for any particular application. In the Bayesian paradigm the distributional assumptions for the data are necessary to inform the parameters in the defining equation $\underset{\sim}{\beta}_j = \mathbf{A}\underset{\sim}{\pi}_j$. Most distributional assumptions can be implemented in WinBUGS, a flexible software for the Bayesian analysis of complex statistical models using Markov chain Monte Carlo (MCMC) methods (Lunn et al., 2000).

5.1.1 Isotope ratio measurements

For each independent source, $s = 1, \ldots, S$, we observe isotope ratio measurements, $\underset{\sim}{d}_{sk}$, $k = 1, \ldots, K_s$, and let $\mathbf{d}_s^* = [\underset{\sim}{d}_{s1}, \ldots, \underset{\sim}{d}_{sK_s}]$. Let isotope ratio measurements from each source have a multivariate normal distribution

$$\underset{\sim}{d}_{sk}|\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s} \overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s}), \qquad k = 1, \ldots, K_{\underset{\sim}{d}_s}, \tag{5.1}$$

where $\underset{\sim}{\delta}_s$ is a mean (column) vector of length I and $\Sigma_{\underset{\sim}{\delta}_s}$ is an $I \times I$ positive definite covariance matrix. The likelihood based on data from the sth source is

$$\prod_{k=1}^{K_s} p(\underset{\sim}{d}_{sk}|\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s}) \propto |\Sigma_{\underset{\sim}{\delta}_s}|^{-K_s/2} \exp\left\{-\frac{1}{2}\sum_{k=1}^{K_s}(\underset{\sim}{d}_{sk} - \underset{\sim}{\delta}_s)^\top \Sigma_{\underset{\sim}{\delta}_s}^{-1}(\underset{\sim}{d}_{sk} - \underset{\sim}{\delta}_s)\right\}. \tag{5.2}$$

For convenience, conjugate prior distributions are defined for the mean and variance,

$$\underset{\sim}{\delta}_s|\underset{\sim}{\delta}_{0s}, \Sigma_{\underset{\sim}{\delta}_s}, \upsilon_{0\underset{\sim}{\delta}_s} \sim \text{Normal}(\underset{\sim}{\delta}_{0s}, \Sigma_{\underset{\sim}{\delta}_s}/\upsilon_{0\underset{\sim}{\delta}_s}) \tag{5.3}$$

$$\Sigma_{\underset{\sim}{\delta}_s}|\Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s} \sim \text{Inv-Wishart}(\Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s}), \tag{5.4}$$

where $\underset{\sim}{\delta}_{0s}$ is the prior mean and $\upsilon_{0\underset{\sim}{\delta}_s}$ is the number of prior measurements on the $\Sigma_{\underset{\sim}{\delta}_s}$ scale, and $\Sigma_{0\underset{\sim}{\delta}_s}$ is the prior covariance with degrees of freedom $\nu_{0\underset{\sim}{\delta}_s}$ (the inverse-Wishart is described in Section 5.1.1.1 below). This corresponds to a joint prior density

$$\begin{aligned} p(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s} | \underset{\sim}{\delta}_{0s}, \upsilon_{0\underset{\sim}{\delta}_s}, \Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s}) & \\ \propto \; |\Sigma_{\underset{\sim}{\delta}_s}|^{-((\nu_{0\underset{\sim}{\delta}_s}+I)/2+1)} & \\ \times \; \exp\left\{-\frac{1}{2}\text{tr}\left(\Sigma_{0\underset{\sim}{\delta}_s}\Sigma_{\underset{\sim}{\delta}_s}^{-1}\right) - \frac{\upsilon_{0\underset{\sim}{\delta}_s}}{2}(\underset{\sim}{\delta}_s - \underset{\sim}{\delta}_{0s})^\top \Sigma_{\underset{\sim}{\delta}_s}^{-1}(\underset{\sim}{\delta}_s - \underset{\sim}{\delta}_{0s})\right\}. & \end{aligned} \tag{5.5}$$

Multiplying the prior density by the normal likelihood gives the "posterior" density for $(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s})$ of the same form as the prior with parameters

$$\begin{aligned} \upsilon_{K_s\underset{\sim}{\delta}_s} &= \upsilon_{0\underset{\sim}{\delta}_s} + K_s \\ \nu_{K_s s} &= \nu_{0s} + K_s \\ \underset{\sim}{\delta}_{K_s s} &= \frac{\upsilon_{0\underset{\sim}{\delta}_s}}{\upsilon_{K_s\underset{\sim}{\delta}_s}}\underset{\sim}{\delta}_{0s} + \frac{K_s}{\upsilon_{K_s\underset{\sim}{\delta}_s}}\bar{\underset{\sim}{d}}_s \\ \Sigma_{K_s\underset{\sim}{\delta}_s} &= \Sigma_{0\underset{\sim}{\delta}_s} + (K_s - 1)\mathbf{S}_{\underset{\sim}{d}_s} + \frac{\upsilon_{0\underset{\sim}{\delta}_s}K_s}{\upsilon_{K_s\underset{\sim}{\delta}_s}}(\bar{\underset{\sim}{d}}_s - \underset{\sim}{\delta}_{0s})(\bar{\underset{\sim}{d}}_s - \underset{\sim}{\delta}_{0s})^\top, \end{aligned}$$

giving

$$\begin{aligned} p(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s} | \mathbf{d}_s^*, \underset{\sim}{\delta}_{0s}, \upsilon_{0\underset{\sim}{\delta}_s}, \Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s}) & \\ \propto \; |\Sigma_{\underset{\sim}{\delta}_s}|^{-((\nu_{K_s s}+I)/2+1)} & \\ \times \; \exp\left\{-\frac{1}{2}\text{tr}\left(\Sigma_{K_s\underset{\sim}{\delta}_s}\Sigma_{\underset{\sim}{\delta}_s}^{-1}\right) - \frac{\upsilon_{K_s\underset{\sim}{\delta}_s}}{2}(\underset{\sim}{\delta}_s - \underset{\sim}{\delta}_{K_s s})^\top \Sigma_{\underset{\sim}{\delta}_s}^{-1}(\underset{\sim}{\delta}_s - \underset{\sim}{\delta}_{K_s s})\right\}. & \end{aligned} \tag{5.6}$$

Note, "posterior" here is in quotes because this is an intermediate result and not the posterior of the full diet model. Alternatively, this important result in (5.6), used later in Section 5.2.2 for the two-step model, can be written hierarchically,

$$\underset{\sim}{\delta}_s | \mathbf{d}_s^*, \underset{\sim}{\delta}_{0s}, \Sigma_{\underset{\sim}{\delta}_s}, \upsilon_{0\underset{\sim}{\delta}_s} \sim \text{Normal}(\underset{\sim}{\delta}_{K_s s}, \Sigma_{\underset{\sim}{\delta}_s}/\upsilon_{K_s\underset{\sim}{\delta}_s}) \tag{5.7}$$

$$\Sigma_{\underset{\sim}{\delta}_s} | \mathbf{d}_s^*, \Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s} \sim \text{Inv-Wishart}(\Sigma_{K_s\underset{\sim}{\delta}_s}, \nu_{K_s\underset{\sim}{\delta}_s}). \tag{5.8}$$

The marginal "posterior" distribution of $\underset{\sim}{\delta}_s$ is multivariate t,

$$\begin{aligned}\underset{\sim}{\delta}_s|\mathbf{d}_s^*, \underset{\sim}{\delta}_{0s}, \upsilon_{0\underset{\sim}{\delta}_s}, \Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0s} \\ \sim t_{\nu_{K_s s}-I+1}\left(\underset{\sim}{\delta}_{K_s s}, \Sigma_{K_s\underset{\sim}{\delta}_s}/(\upsilon_{K_s\underset{\sim}{\delta}_s}(\nu_{K_s s}-I+1))\right).\end{aligned}$$

The distribution in (5.7) and (5.8) are for one source's isotope ratio mean and covariance independent of other sources. In the context of the complete model (later as (5.28)), the posterior for the source parameters $(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s})$ may be different conditional on all the data. Note that it can be a sign of a potential problem if the posterior for one or more of the $\underset{\sim}{\delta}_s$ is very different from the distribution above in (5.7) and (5.8). Recall the discussion in Section 1.2.2 that a solution to the mixing model equation (1.1) exists if and only if the mixture value lies inside all pairwise isotope ratio convex hulls. If the mixture mean is outside the convex hull of the discrimination-corrected source means, then the source means may attempt to relocate to accommodate the mixture mean and this will be reflected by the posterior for the source parameters, $(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s})$, being different from as in (5.7) and (5.8). This issue is discussed more completely in Section 5.3.4.3.

Jeffreys prior For completeness, the noninformative multivariate Jeffreys prior density,

$$p(\underset{\sim}{\delta}_s, \Sigma_{\underset{\sim}{\delta}_s}) \propto |\Sigma_{\underset{\sim}{\delta}_s}|^{-(I+1)/2}, \tag{5.9}$$

is the limit of the conjugate prior density (5.5) as $\upsilon_{0\underset{\sim}{\delta}_s} \to 0$, $\nu_{0\underset{\sim}{\delta}_s} \to -1$, and $|\Sigma_{0\underset{\sim}{\delta}_s}| \to 0$. The corresponding "posterior" can be written,

$$\underset{\sim}{\delta}_s|\mathbf{d}_s^*, \Sigma_{\underset{\sim}{\delta}_s} \sim \text{Normal}(\bar{\underset{\sim}{d}}_s, \Sigma_{\underset{\sim}{\delta}_s}/K_s) \tag{5.10}$$

$$\Sigma_{\underset{\sim}{\delta}_s}|\mathbf{d}_s^* \sim \text{Inv-Wishart}\left((K_s-1)\mathbf{S}_{\underset{\sim}{d}_s}, K_s-1\right). \tag{5.11}$$

This Jeffreys prior is useful when no prior information is available. In the "posterior" in (5.10) and (5.11) the Inv-Wishart distribution is always proper (integral is finite) and the mean is always finite, though a degenerate form occurs (random matrix no longer positive definite) when $K_s < I$. Additionally, Jeffreys prior can not be used directly in WinBUGS as in (5.9), since prior distributions must be proper, but can be implemented using the two-step method (Section 5.2.2) as in (5.10) and (5.11).

5.1.1.1 Wishart distribution

The Wishart distribution and especially the inverse-Wishart distribution are the source of some confusion because they occasionally appear with alternative parameterizations. The definition we use for the Wishart is

$$p(\Sigma_{\underset{\sim}{\delta}_s}^{-1}|\Sigma_{0\underset{\sim}{\delta}_s}^{-1}, \nu_{0\underset{\sim}{\delta}_s}) \quad \propto \quad |\Sigma_{0\underset{\sim}{\delta}_s}|^{\nu_{0\underset{\sim}{\delta}_s}/2}|\Sigma_{\underset{\sim}{\delta}_s}^{-1}|^{(\nu_{0\underset{\sim}{\delta}_s}-I-1)/2}\exp\left\{-\frac{1}{2}\text{tr}(\Sigma_{0\underset{\sim}{\delta}_s}\Sigma_{\underset{\sim}{\delta}_s}^{-1})\right\} \quad (5.12)$$

with $\text{E}[\Sigma_{\underset{\sim}{\delta}_s}^{-1}] = \nu_{0\underset{\sim}{\delta}_s}\Sigma_{0\underset{\sim}{\delta}_s}^{-1}$ and where value of the $(i,j)^{\text{th}}$ element of $\Sigma_{\underset{\sim}{\delta}_s}^{-1}$ is $\nu_{0\underset{\sim}{\delta}_s}(\tau_{0ij}^2 + \tau_{0ii}\tau_{0jj})$, where τ_{0ij} is the $(i,j)^{\text{th}}$ element of $\Sigma_{0\underset{\sim}{\delta}_s}^{-1}$, $\nu_{0\underset{\sim}{\delta}_s}\Sigma_{0\underset{\sim}{\delta}_s}^{-1}$ is the inverse of the prior population covariance, and $\nu_{0\underset{\sim}{\delta}_s}$ can be thought of as the prior sample size. The Wishart distribution can be used to model a covariance matrix or a precision matrix (the inverse of a covariance matrix) (pp. 161–2, 175–6, 180 in Kotz et al., 2004; Dickey et al., 1985). The Wishart is the conjugate prior distribution for the precision matrix (Σ^{-1}) in a multivariate normal distribution. Also, the sample covariance matrix ($\mathbf{S}$) for iid multivariate normal data has a Wishart distribution. Specifically, the scatter matrix, $\mathbf{S}' = \sum_{k=1}^{K}(\underset{\sim}{d}_k - \underset{\sim}{\delta})(\underset{\sim}{d}_k - \underset{\sim}{\delta})^\top$, has $\mathbf{S}'|\Sigma, K \sim \text{Wishart}(\Sigma, K)$, where Σ is the population covariance and K is the sample size informing $\mathbf{S}'$. In both cases, the parameter matrix and the random matrix are on the same scale, that is, they are either both precision matrices or covariance matrices (Appendix A, Gelman et al., 1995). The Wishart distribution is proper (integral is finite) when $\nu_{0\underset{\sim}{\delta}_s} \geq I$, and the mean is finite when $\nu_{0\underset{\sim}{\delta}_s} \geq I + 1$, $s = 1, \ldots, S$.

The inverse-Wishart is the conjugate prior distribution for the multivariate normal covariance matrix. If $\Sigma^{-1} \sim \text{Wishart}(\Sigma_0^{-1}, \nu_0)$, then $\Sigma \sim \text{Inv-Wishart}(\Sigma_0, \nu_0)$. The definition of the Inv-Wishart, implied by the definition of the Wishart above, is

$$p(\Sigma_{\underset{\sim}{\delta}_s}|\Sigma_{0\underset{\sim}{\delta}_s}, \nu_{0\underset{\sim}{\delta}_s}) \propto |\Sigma_{0\underset{\sim}{\delta}_s}|^{\nu_{0\underset{\sim}{\delta}_s}/2}|\Sigma_{\underset{\sim}{\delta}_s}|^{-(\nu_{0\underset{\sim}{\delta}_s}+I+1)/2}\exp\left\{-\frac{1}{2}\text{tr}(\Sigma_{0\underset{\sim}{\delta}_s}\Sigma_{\underset{\sim}{\delta}_s}^{-1})\right\} \quad (5.13)$$

with $\text{E}[\Sigma_{\underset{\sim}{\delta}_s}] = (\nu_{0\underset{\sim}{\delta}_s} - I - 1)^{-1}\Sigma_{0\underset{\sim}{\delta}_s}$ and

$$\text{Var}[\sigma_{ij}] = \frac{(\nu_{0\underset{\sim}{\delta}_s} - I - 1)\sigma_{0ij}^2 + (\nu_{0\underset{\sim}{\delta}_s} - I + 1)\sigma_{0ii}\sigma_{0jj}}{(\nu_{0\underset{\sim}{\delta}_s} - I)(\nu_{0\underset{\sim}{\delta}_s} - I - 1)^2(\nu_{0\underset{\sim}{\delta}_s} - I - 3)},$$

where σ_{ij}^2 is the $(i,j)^{\text{th}}$ element of $\Sigma_{\underset{\sim}{\delta}_s}$ and σ_{0ij}^2 is the $(i,j)^{\text{th}}$ element of $\Sigma_{0\underset{\sim}{\delta}_s}$.

Notice that the expected value for the Wishart (Inv-Wishart) distribution is a scale of the prior precision (covariance) matrix. Therefore, it may be desirable to pre-scale the prior precision (covariance) matrix so that the expected value is what is desired. This is not an issue in the univariate specification below. Also, note that the variance is a function of both the prior precision (covariance) matrix and degrees of freedom.

5.1.1.2 Univariate specification

The Wishart is a multivariate generalization of the scaled-χ^2 distribution, a gamma distribution. That is,

$$\text{scaled-}\chi^2(\tau_{\delta_s}|\sigma_{0\delta_s}^2, \nu_{0\delta_s}) \equiv \text{Gamma}(\tau_{\delta_s}|\nu_{0\delta_s}/2, \nu_{0\delta_s}\sigma_{0\delta_s}^2/2),$$

where $\tau_{\delta_s} \equiv (\sigma_{\delta_s}^2)^{-1}$ is the precision, with density function

$$p(\tau_{\delta_s}|\sigma_{0\delta_s}^2, \nu_{0\delta_s}) \propto (\sigma_{0\delta_s}^2)^{\nu_{0\delta_s}/2}(\tau_{\delta_s})^{\nu_{0\delta_s}/2-1}\exp\left\{-\nu_{0\delta_s}\sigma_{0\delta_s}^2\tau_{\delta_s}/2\right\} \quad (5.14)$$

with $\mathrm{E}[\tau_{\delta_s}] = (\sigma^2_{0\delta_s})^{-1}$ and $\mathrm{Var}[\tau_{\delta_s}] = (\nu_{0\delta_s}(\sigma^2_{0\delta_s})^2/2)^{-1}$. Similarly, the

$$\text{Scaled-inv-}\chi^2(\sigma^2_{\delta_s}|\sigma^2_{0\delta_s},\nu_{0\delta_s}) \;\equiv\; \text{Inv-Gamma}(\sigma^2_{\delta_s}|\nu_{0\delta_s}/2, \nu_{0\delta_s}\sigma^2_{0\delta_s}/2)$$

with density function

$$p(\sigma^2_{\delta_s}|\sigma^2_{0\delta_s},\nu_{0\delta_s}) \;\propto\; (\sigma^2_{0\delta_s})^{\nu_{0\delta_s}/2}(\sigma^2_{\delta_s})^{-(\nu_{0\delta_s}/2+1)}\exp\left\{-\nu_{0\delta_s}\sigma^2_{0\delta_s}/(2\sigma^2_{\delta_s})\right\} \tag{5.15}$$

with $\mathrm{E}[\sigma^2_{\delta_s}] = (\nu_{0\delta_s}/(\nu_{0\delta_s}-2))\sigma^2_{0\delta_s}$ and $\mathrm{Var}[\sigma^2_{\delta_s}] = \{2\nu^2_{0\delta_s}/[(\nu_{0\delta_s}-2)^2(\nu_{0\delta_s}-4)]\}(\sigma^2_{0\delta_s})^2$.

For the univariate case ($I = 1$), let isotope ratio measurements from each source have

$$d_{sk}|\delta_s,\sigma^2_{\delta_s} \overset{\text{ind}}{\sim} \text{Normal}(\delta_s,\sigma^2_{\delta_s}), \qquad k = 1,\ldots,K_{\underset{\sim}{d}s}. \tag{5.16}$$

For convenience, conjugate prior distributions for the mean and variance are

$$\begin{aligned}
\delta_s|\delta_{0s},\sigma^2_{\delta_s},\upsilon_{0\delta_s} &\sim \text{Normal}(\delta_{0s},\sigma^2_{\delta_s}/\upsilon_{0\delta_s}) &(5.17)\\
\sigma^2_{\delta_s}|\sigma^2_{0\delta_s},\nu_{0\delta_s} &\sim \text{Scaled-inv-}\chi^2(\sigma^2_{0\delta_s},\nu_{0\delta_s}). &(5.18)
\end{aligned}$$

The "posterior" can be written hierarchically,

$$\begin{aligned}
\delta_s|\mathbf{d}^*_s,\delta_{0s},\sigma^2_{\delta_s},\upsilon_{0\delta_s} &\sim \text{Normal}(\delta_{K_s s},\sigma^2_{\delta_s}/\upsilon_{K_s\delta_s}) &(5.19)\\
\sigma^2_{\delta_s}|\mathbf{d}^*_s,\sigma^2_{0\delta_s},\nu_{0\delta_s} &\sim \text{Scaled-inv-}\chi^2(\sigma^2_{K_s\delta_s},\nu_{K_s\delta_s}), &(5.20)
\end{aligned}$$

where

$$\begin{aligned}
\upsilon_{K_s\delta_s} &= \upsilon_{0\delta_s} + K_s\\
\nu_{K_s s} &= \nu_{0s} + K_s\\
\delta_{K_s s} &= \frac{\upsilon_{0\delta_s}}{\upsilon_{K_s\delta_s}}\delta_{0s} + \frac{K_s}{\upsilon_{K_s\delta_s}}\bar{d}_s\\
\sigma^2_{K_s\delta_s} &= \frac{\nu_{0s}}{\nu_{K_s s}}\sigma^2_{0\delta_s} + \frac{K_s-1}{\nu_{K_s s}}s^2_{d_s} + \frac{\upsilon_{0\delta_s}K_s}{\nu_{K_s s}\upsilon_{K_s\delta_s}}(\bar{d}_s-\delta_{0s})^2.
\end{aligned}$$

Univariate Jeffreys prior The univariate analog to (5.9), (5.10), and (5.11) are

$$p(\delta_s, \sigma^2_{\delta_s}) \propto (\sigma^2_{\delta_s})^{-1} \tag{5.21}$$

$$\delta_s|\mathbf{d}^*_s, \sigma^2_{\delta_s} \sim \text{Normal}(\bar{d}_s, \sigma^2_{\delta_s}/K_s) \tag{5.22}$$

$$\sigma^2_{\delta_s}|\mathbf{d}^*_s \sim \text{Scaled-inv-}\chi^2(s^2_{d_s}, K_s - 1). \tag{5.23}$$

5.1.2 Discrimination

There are multiple ways to estimate discrimination, as seen in the dunlin and bear examples in Chapter 4, and we can implement the same two methods in the Bayesian model. We illustrate the single-source diet experiment method of estimating discrimination. The regression-based method for estimating discrimination for a set of sources simultaneously, introduced in Section 4.3.1.3, can be easily developed under the Bayesian paradigm.

5.1.2.1 Single-source diet experiment

Discrimination is often estimated as the difference between a controlled diet and tissues of the mixture at isotopic equilibrium, as in Section 4.1.3. In this case, for each source $s = 1, \ldots, S$, the mean isotope ratio is estimated from the single diet source and for a tissue from the mixture, then the difference of these provides inference for the discrimination. Controlled diets representative of the sources, $s = 1, \ldots, S$, are used to calculate the discrimination for all sources in the model.

For source s we observe isotope ratio measurements for the diet source and for the mixture tissue, $\underset{\sim}{d}_{\text{D}sk}$, $k = 1, \ldots, K_{\text{D}s}$, and $\underset{\sim}{d}_{\text{T}sk}$, $k = 1, \ldots, K_{\text{T}s}$, and let $\mathbf{d}^*_{\text{D}s} = [\underset{\sim}{d}_{\text{D}s1}, \ldots, \underset{\sim}{d}_{\text{D}sK_{\text{D}s}}]$ and $\mathbf{d}^*_{\text{T}s} = [\underset{\sim}{d}_{\text{T}s1}, \ldots, \underset{\sim}{d}_{\text{T}sK_{\text{T}s}}]$. We use the same model for the diet source and tissue isotope ratios as we used earlier in Section 5.1.1, therefore the "posterior" distribution derived there for the source isotope ratios apply here for the

diet and tissue isotope ratios as well. For diets,

$$\begin{aligned}
\underset{\sim}{d}_{\mathrm{D}sk}|\underset{\sim}{\delta}_{\mathrm{D}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{D}s}} &\overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{\delta}_{\mathrm{D}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{D}s}}), \quad k = 1, \ldots, K_{\mathrm{D}s} \\
\underset{\sim}{\delta}_{\mathrm{D}s}|\underset{\sim}{\delta}_{0\mathrm{D}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{D}s}}, \upsilon_{0\mathrm{D}s} &\sim \text{Normal}(\underset{\sim}{\delta}_{0\mathrm{D}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{D}s}}/\upsilon_{0\mathrm{D}s}) \\
\Sigma_{\underset{\sim}{\delta}_{\mathrm{D}s}}|\Sigma_{0\underset{\sim}{\delta}_{\mathrm{D}s}}, \nu_{0\underset{\sim}{\delta}_{\mathrm{D}s}} &\sim \text{Inv-Wishart}(\Sigma_{0\underset{\sim}{\delta}_{\mathrm{D}s}}, \nu_{0\underset{\sim}{\delta}_{\mathrm{D}s}}).
\end{aligned}$$

For corresponding tissues,

$$\begin{aligned}
\underset{\sim}{d}_{\mathrm{T}sk}|\underset{\sim}{\delta}_{\mathrm{T}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{T}s}} &\overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{\delta}_{\mathrm{T}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{T}s}}), \quad k = 1, \ldots, K_{\mathrm{T}s} \\
\underset{\sim}{\delta}_{\mathrm{T}s}|\underset{\sim}{\delta}_{0\mathrm{T}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{T}s}}, \upsilon_{0\mathrm{T}s} &\sim \text{Normal}(\underset{\sim}{\delta}_{0\mathrm{T}s}, \Sigma_{\underset{\sim}{\delta}_{\mathrm{T}s}}/\upsilon_{0\mathrm{T}s}) \\
\Sigma_{\underset{\sim}{\delta}_{\mathrm{T}s}}|\Sigma_{0\underset{\sim}{\delta}_{\mathrm{T}s}}, \nu_{0\underset{\sim}{\delta}_{\mathrm{T}s}} &\sim \text{Inv-Wishart}(\Sigma_{0\underset{\sim}{\delta}_{\mathrm{T}s}}, \nu_{0\underset{\sim}{\delta}_{\mathrm{T}s}}).
\end{aligned}$$

Discrimination is defined as $\underset{\sim}{\Delta}_s = \underset{\sim}{\delta}_{\mathrm{T}s} - \underset{\sim}{\delta}_{\mathrm{D}s}$. The prior on $\underset{\sim}{\Delta}_s$ is directly induced by the priors on $\underset{\sim}{\delta}_{\mathrm{T}s}$ and $\underset{\sim}{\delta}_{\mathrm{D}s}$.

5.2 Implicit representation, population mean diet

As in the BMM in (1.1), we model the population mean isotope ratio of the mixture, $\underset{\sim}{\beta}$, as a function of the mean isotope ratios of the sources consumed, $\underset{\sim}{\delta}_s$, $s = 1, \ldots, S$, corrected for the isotope ratio discrimination applying to each source, $\underset{\sim}{\Delta}_s$, weighted by the population mean dry-weight biomass proportion of each source consumed, π_s,

$$\underset{\sim}{\beta} \;=\; \mathbf{A}\underset{\sim}{\pi} \;=\; \sum_{s=1}^{S} \underset{\sim}{\delta}'_s \pi_s \tag{5.24}$$

where, for example, in the case of the single-source diet experiment, the discrimination-corrected source isotope ratio is $\underset{\sim}{\delta}'_s = \underset{\sim}{\delta}_s + (\underset{\sim}{\delta}_{\mathrm{T}s} - \underset{\sim}{\delta}_{\mathrm{D}s})$. This view of modeling the mean isotope ratio of the mixture conditional on the source isotope ratios and the diet proportion vector suggests the following hierarchical model. Assume we have

one isotope ratio observation per individual mixture, $\underset{\sim}{b}_j$, and let these measurements have

$$\underset{\sim}{b}_j|\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta^* \overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{\beta}, \Sigma_{\underset{\sim}{b}}), \quad j = 1, \ldots, J \tag{5.25}$$

$$\Sigma_{\underset{\sim}{b}}|\Sigma_{0\underset{\sim}{b}}, \nu_{0\underset{\sim}{b}} \sim \text{Inv-Wishart}(\Sigma_{0\underset{\sim}{b}}, \nu_{0\underset{\sim}{b}}), \tag{5.26}$$

where $\Sigma_{0\underset{\sim}{b}}$ and $\nu_{0\underset{\sim}{b}}$ specify the prior covariance matrix of the isotope ratio values of the mixtures and an effective prior sample size, and where $\delta^*_{\text{S}} = [\underset{\sim}{\delta}_1, \ldots, \underset{\sim}{\delta}_S]$, $\delta^*_{\text{D}} = [\underset{\sim}{\delta}_{\text{D}1}, \ldots, \underset{\sim}{\delta}_{\text{D}S}]$, and $\delta^*_{\text{T}} = [\underset{\sim}{\delta}_{\text{T}1}, \ldots, \underset{\sim}{\delta}_{\text{T}S}]$, with $\delta^* = [\delta^*_{\text{S}}, \delta^*_{\text{D}}, \delta^*_{\text{T}}]$. The observed mixture values are centered at the population mean isotope ratio, $\underset{\sim}{\beta}$, which depends on the average diet through (5.24). While individual mixtures have their own diets, this specification implicitly models their differences in diet by the variation in their isotope ratio.

The model above has three variance components. From (2.1), the subject-specific mean is $\underset{\sim}{\beta}_j = \underset{\sim}{\beta} + \underset{\sim}{\psi}_j$ with $\underset{\sim}{\psi}_j \sim \text{Normal}(\underset{\sim}{0}, \Sigma_{\underset{\sim}{\beta}})$, where $\Sigma_{\underset{\sim}{\beta}}$ is the between-individual covariance. Furthermore, the mixture observations are

$$\underset{\sim}{b}_j = \underset{\sim}{\beta}_j + \underset{\sim}{\xi}_j + \underset{\sim}{\varepsilon}_j$$

with

$$\underset{\sim}{\xi}_j \sim \text{Normal}(\underset{\sim}{0}, \Sigma_{\underset{\sim}{\xi}}) \quad \text{and} \quad \underset{\sim}{\varepsilon}_j \sim \text{Normal}(\underset{\sim}{0}, \Sigma_\varepsilon)$$

where $\Sigma_{\underset{\sim}{\xi}}$ is the within-individual covariance and Σ_ε is the measurement error covariance (discussed further in Section 5.3.1). Having assumed to have only one observation per mixture and to simplify the modeling, we set $\underset{\sim}{\xi}_j \equiv \underset{\sim}{0}$, $j = 1, \ldots, J$, implying that $\Sigma_{\underset{\sim}{\xi}} \equiv \mathbf{0}$, thus confounding within-individual variability with measurement error variability. Alternatively, $\Sigma_{\underset{\sim}{\xi}}$ can be estimated directly by taking multiple measurements on each mixture, or possibly indirectly in other ways. Note that $\Sigma_{\underset{\sim}{b}} = \Sigma_{\underset{\sim}{\beta}} + \Sigma_\varepsilon$

and in the implicit representation model no attempt is made to separately estimate these variance components. As an aside, we note that in the overconstrained case ($S < I + 1$) that $\underset{\sim}{b}_j$ can have a density because of the between-individual variability. In the absence of between-individual variability the issue of a density for $\underset{\sim}{\beta}$ would be more critical as the conditional distribution in (5.25) would be degenerate.

The mean diet proportion vector, $\underset{\sim}{\pi}$, for the mixture population is the primary parameter of interest. It is up for question what the best distribution is for the diet vector. A natural first approximation for modeling the mean population diet is the Dirichlet distribution, which assigns probability on the simplex given the specification of location and precision. Thus, we assume

$$\underset{\sim}{\pi}|\underset{\sim}{\pi}_0, \alpha_{0\pi}, S \;\sim\; \text{Dirichlet}(\alpha_{0\pi} S \underset{\sim}{\pi}_0), \tag{5.27}$$

where $\alpha_{0\pi}$ specifies the precision of the prior distribution about the prior mean, $\underset{\sim}{\pi}_0$. The prior density for $\underset{\sim}{\pi}$ is more concentratrated about $\underset{\sim}{\pi}_0$ when $\alpha_{0\pi} > 1$ and more disperse when $\alpha_{0\pi} < 1$. A flat "noninformative" prior has $\underset{\sim}{\pi}_0 = S^{-1}\underset{\sim}{1}$ and $\alpha_{0\pi} = 1$, that is, Dirichlet parameter equal to $\underset{\sim}{1}$, where the *a priori* density function is equal for all $\underset{\sim}{\pi}$ on the simplex.

It will be convenient to have notation to group the data and parameters. For data, let $\mathbf{b}^* = [\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_J]$ for mixtures, and let the sources and discrimination be consolidated into $\mathbf{d}^* = [\mathbf{d}^*_1, \ldots, \mathbf{d}^*_S]$, $\mathbf{d}^*_{\mathrm{D}} = [\mathbf{d}^*_{\mathrm{D}1}, \ldots, \mathbf{d}^*_{\mathrm{D}S}]$, $\mathbf{d}^*_{\mathrm{T}} = [\mathbf{d}^*_{\mathrm{T}1}, \ldots, \mathbf{d}^*_{\mathrm{T}S}]$, and $\mathbf{D}^* = [\mathbf{d}^*, \mathbf{d}^*_{\mathrm{D}}, \mathbf{d}^*_{\mathrm{T}}]$. Also, let $\Sigma^*_{\mathrm{S}} = [\Sigma_{\underset{\sim}{\delta}_1}, \ldots, \Sigma_{\underset{\sim}{\delta}_S}]$, $\Sigma^*_{\mathrm{D}} = [\Sigma_{\underset{\sim}{\delta}_{\mathrm{D}1}}, \ldots, \Sigma_{\underset{\sim}{\delta}_{\mathrm{D}S}}]$, $\Sigma^*_{\mathrm{T}} = [\Sigma_{\underset{\sim}{\delta}_{\mathrm{T}1}}, \ldots, \Sigma_{\underset{\sim}{\delta}_{\mathrm{T}S}}]$, and $\Sigma^* = [\Sigma^*_{\mathrm{S}}, \Sigma^*_{\mathrm{D}}, \Sigma^*_{\mathrm{T}}]$. Finally, recall that $\mathbf{A} = [\underset{\sim}{\delta}'_1, \ldots, \underset{\sim}{\delta}'_S] = \delta^*_{\mathrm{S}} + \delta^*_{\mathrm{T}} - \delta^*_{\mathrm{D}}$, and $\underset{\sim}{\beta} = \mathbf{A}\underset{\sim}{\pi}$. The last two parameters, $\underset{\sim}{\pi}$ and $\Sigma_{\underset{\sim}{b}}$, will be kept separate.

The joint posterior is

$$
\begin{aligned}
p(\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta^*, \Sigma^* | \mathbf{b}^*, \mathbf{D}^*) \quad \propto \quad & \left[\prod_{j=1}^{J} f(\underline{b}_j | \underset{\sim}{\beta}, \Sigma_{\underset{\sim}{b}}) \right] p(\Sigma_{\underset{\sim}{b}}) p(\underset{\sim}{\pi}) \\
& \times \prod_{s=1}^{S} \left\{ \left[\prod_{k=1}^{K_s} f(\underline{d}_{sk} | \underline{\delta}_s, \Sigma_{\underline{\delta}_s}) \right] p(\underline{\delta}_s | \Sigma_{\underline{\delta}_s}) p(\Sigma_{\underline{\delta}_s}) \right\} \\
& \times \prod_{s=1}^{S} \left\{ \left[\prod_{k=1}^{K_{\mathrm{D}s}} f(\underline{d}_{\mathrm{D}sk} | \underline{\delta}_{\mathrm{D}s}, \Sigma_{\underline{d}_{\mathrm{D}s}}) \right] p(\underline{\delta}_{\mathrm{D}s} | \Sigma_{\underline{d}_{\mathrm{D}s}}) p(\Sigma_{\underline{d}_{\mathrm{D}s}}) \right\} \\
& \times \prod_{s=1}^{S} \left\{ \left[\prod_{k=1}^{K_{\mathrm{T}s}} f(\underline{d}_{\mathrm{T}sk} | \underline{\delta}_{\mathrm{T}s}, \Sigma_{\underset{\sim}{d}_{\mathrm{T}s}}) \right] p(\underline{\delta}_{\mathrm{T}s} | \Sigma_{\underset{\sim}{d}_{\mathrm{T}s}}) p(\Sigma_{\underset{\sim}{d}_{\mathrm{T}s}}) \right\} \\
= \quad & p(\mathbf{b}^* | \delta^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi}) p(\Sigma_{\underset{\sim}{b}}) p(\underset{\sim}{\pi}) \times p(\mathbf{D}^* | \delta^*, \Sigma^*) p(\delta^*, \Sigma^*).
\end{aligned}
\tag{5.28}
$$

Note that the first line on the right hand side of (5.28) is $p(\mathbf{b}^* | \delta^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi}) p(\Sigma_{\underset{\sim}{b}}) p(\underset{\sim}{\pi})$ and the next three, collectively, are $p(\mathbf{D}^* | \delta^*, \Sigma^*) p(\delta^*, \Sigma^*)$. Also, note that the hyper-parameters are in these distributions but are omitted from the equation display. The posterior density of $\underset{\sim}{\pi}$ may be a challenge to derive analytically, but all posteriors can be estimated computationally using the Bayesian paradigm, for example via Gibbs sampling with WinBUGS or OpenBUGS (Casella and George, 1992; Gelfand and Smith, 1990; Lunn et al., 2000).

When the prior on $\underset{\sim}{\pi}$ is uniform, the posterior for $\underset{\sim}{\pi}$ attains its maximum over the solution polytope (detailed in Chapter 3). Otherwise, the maximum within the solution polytope is determined by the prior on $\underset{\sim}{\pi}$.

5.2.1 IsoSource (or SISUS) as the no-variation limit of the implicit representation

IsoSource (or SISUS), discussed in Chapter 3, can be seen as an approximation to the Bayesian implicit representation model where the variability goes to zero. The following constructed example with three sources and one isotope illustrates this.

Let the mixture have population mean $\beta = 3$ and sources $s = 1, 2, 3$ have population means $\delta_S^* = [0, 5, 10]$. For IsoSource (no variation) the four means are the only data inputs, and solutions of $\underset{\sim}{\pi}$ are sought for

$$\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 10 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix},$$

where the simplex condition is included in the second row. Without variation, IsoSource approximates the solution polytope for $\underset{\sim}{\pi}$ defined by the line segment with endpoints $(0.7, 0.0, 0.3)$ and $(0.4, 0.6, 0.0)$. We consider five scenarios with population variances assumed common to the sources and mixture, $\sigma_0^2 = \{4, 1, 0.1, 0.01, 0.0001\}$, to show variability of all model components going to zero. More generally, many similar constructions where all variances approach zero will suffice. The defining relation using (1.1) is

$$\beta = \delta_S^* \underset{\sim}{\pi} = \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 \end{bmatrix} \begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{bmatrix}.$$

The distributions for the sources are defined as priors in the model,

$$\begin{aligned} \delta_1 | \delta_{01}, \sigma_0^2 &\sim \text{Normal}(\ 0, \sigma_0^2) \\ \delta_2 | \delta_{02}, \sigma_0^2 &\sim \text{Normal}(\ 5, \sigma_0^2) \\ \delta_3 | \delta_{03}, \sigma_0^2 &\sim \text{Normal}(10, \sigma_0^2), \end{aligned}$$

and no data informs $\delta_s, s = 1, \ldots, 3$. The mixture distribution is

$$b_1 | \delta_S^*, \underset{\sim}{\pi}, \sigma_0^2 \sim \text{Normal}(\beta, \sigma_0^2),$$

and a single observation of $b_1 = 3$ was used. The flat diet distribution is

$$\underset{\sim}{\pi}|\underset{\sim}{\pi}_0, \alpha_{0\pi}, S \;\sim\; \text{Dirichlet}([1,1,1]^\top).$$

The joint posterior using (5.28) is

$$p(\underset{\sim}{\pi}, \delta_1, \delta_2, \delta_3 | b_1 = 3, \sigma_0^2) \;\propto\; (\sigma_0^2)^{-1/2} \exp\left\{ -\frac{1}{2\sigma_0^2} \left(3 - \sum_{s=1}^{3} \delta_s \pi_s \right)^2 \right\} \times (\sigma_0^2)^{-3/2} \exp\left\{ -\frac{1}{2\sigma_0^2} \sum_{s=1}^{3} (\delta_s - \delta_{0s})^2 \right\}.$$

The posterior attains its maximum where, simultaneously, $\underset{\sim}{\pi}$ satisfies $3 = \sum_{s=1}^{3} \delta_s \pi_s$, and $\delta_s = \delta_{0s}$, $s = 1, 2, 3$, and decays exponentially from this maximum scaled by the variance, σ_0^2. As σ_0^2 decreases, δ_s becomes more concentrated about δ_{0s}, $s = 1, 2, 3$, and the marginal posterior of $\underset{\sim}{\pi}$ becomes more concentrated on the "solution polytope" that satisfies $3 = \sum_{s=1}^{3} \delta_{0s} \pi_s$. The simulation below qualitatively confirms this heuristic argument.

WinBUGS was given the model above and allowed a 1000 iterate burn-in after which every 10th iterate was retained for 2500 posterior draws for analysis. The plot in Figure 5.1 shows the mixture and source distributions under the five variance scenarios. The plot in Figure 5.2 shows that the implicit representation model posterior has the mode at the solution polytope and as the variation goes to zero the posterior concentrates about the mode. The plot in Figure 5.3 shows the results as IsoSource reports them, as marginal histograms. When variation is included, there is substantial probability outside the no-variation solution histograms. The large variation in the first two of these five scenarios is not unlike what is observed in some natural systems, and some have even more distributional overlap than in Figure 5.1. Therefore, IsoSource reports results with optimistic certainty since it does not include variability.

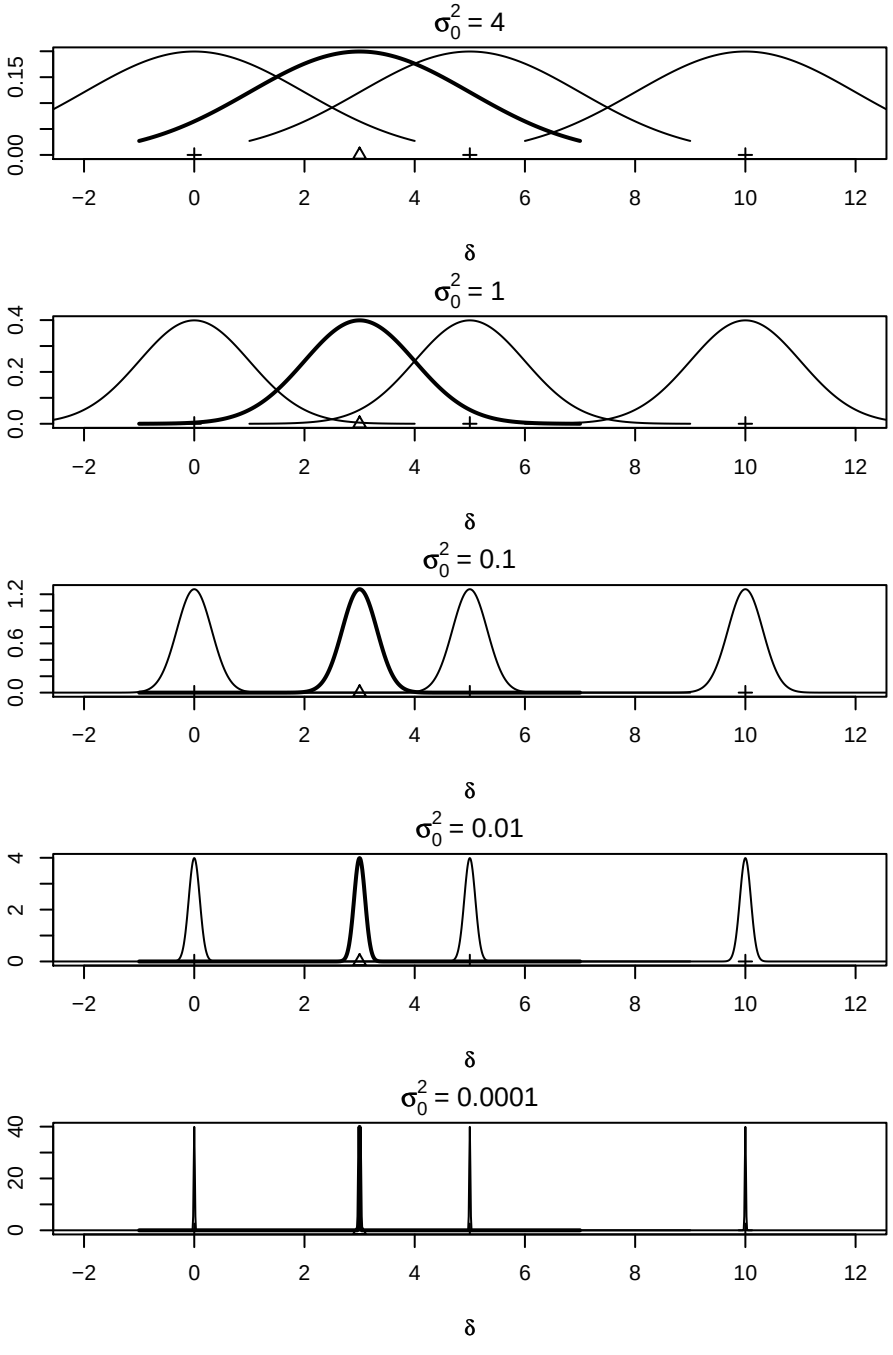

Figure 5.1: Implicit representation IsoSource limiting argument, mixture and source distributions for five variances $\sigma_0^2 = \{4, 1, 0.1, 0.01, 0.0001\}$. The triangle indicates the prior mean for the mixture and the crosses the prior means for the three sources.

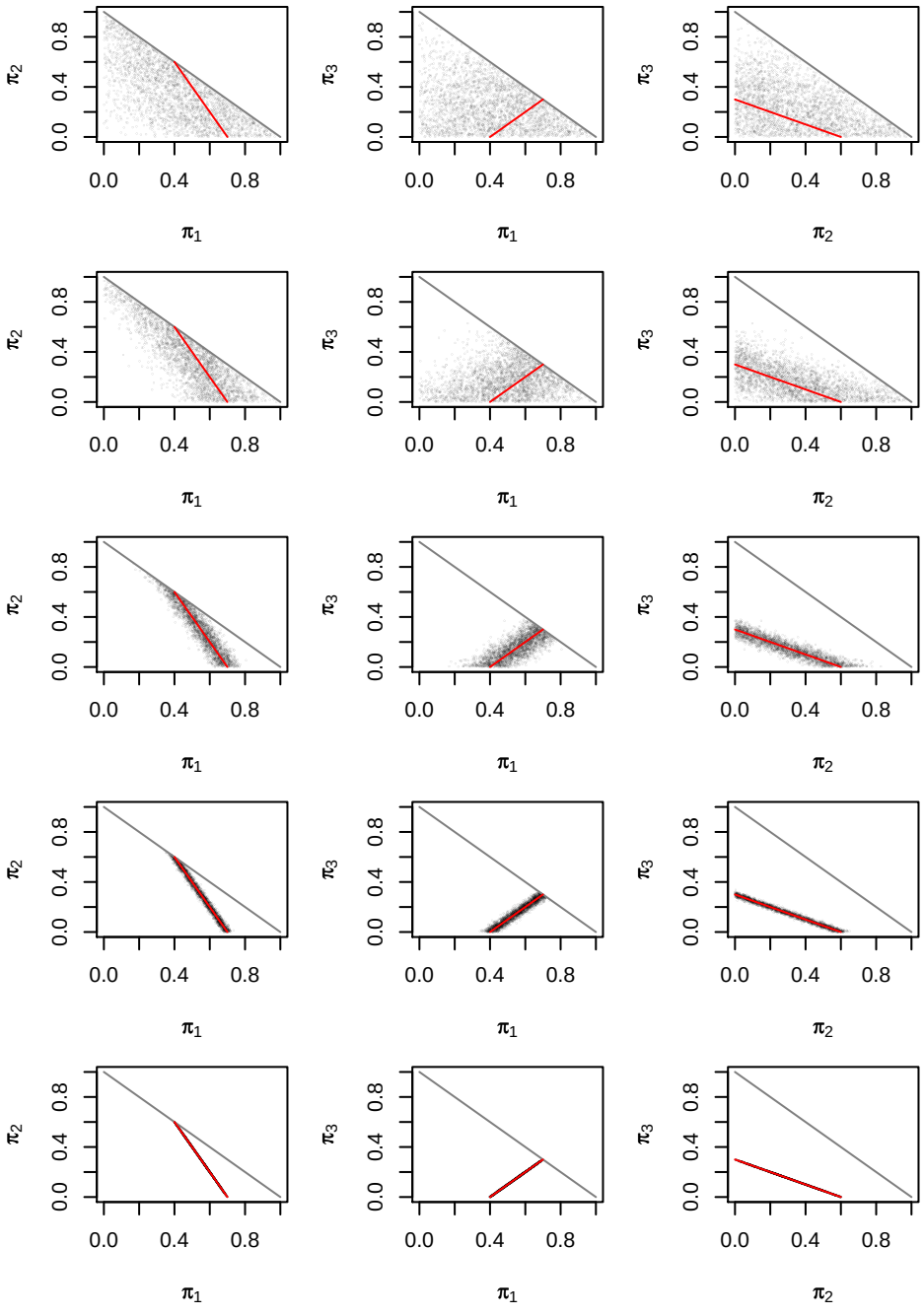

Figure 5.2: Implicit representation IsoSource limiting argument, bivariate distributions of pairs of (π_1, π_2, π_3) with the mixture and source distributions as in Figure 5.1. Each row corresponds, from top to bottom, to the five scenarios for $\sigma_0^2 = \{4, 1, 0.1, 0.01, 0.0001\}$. The red line indicates the no-variation solution approximated by IsoSource.

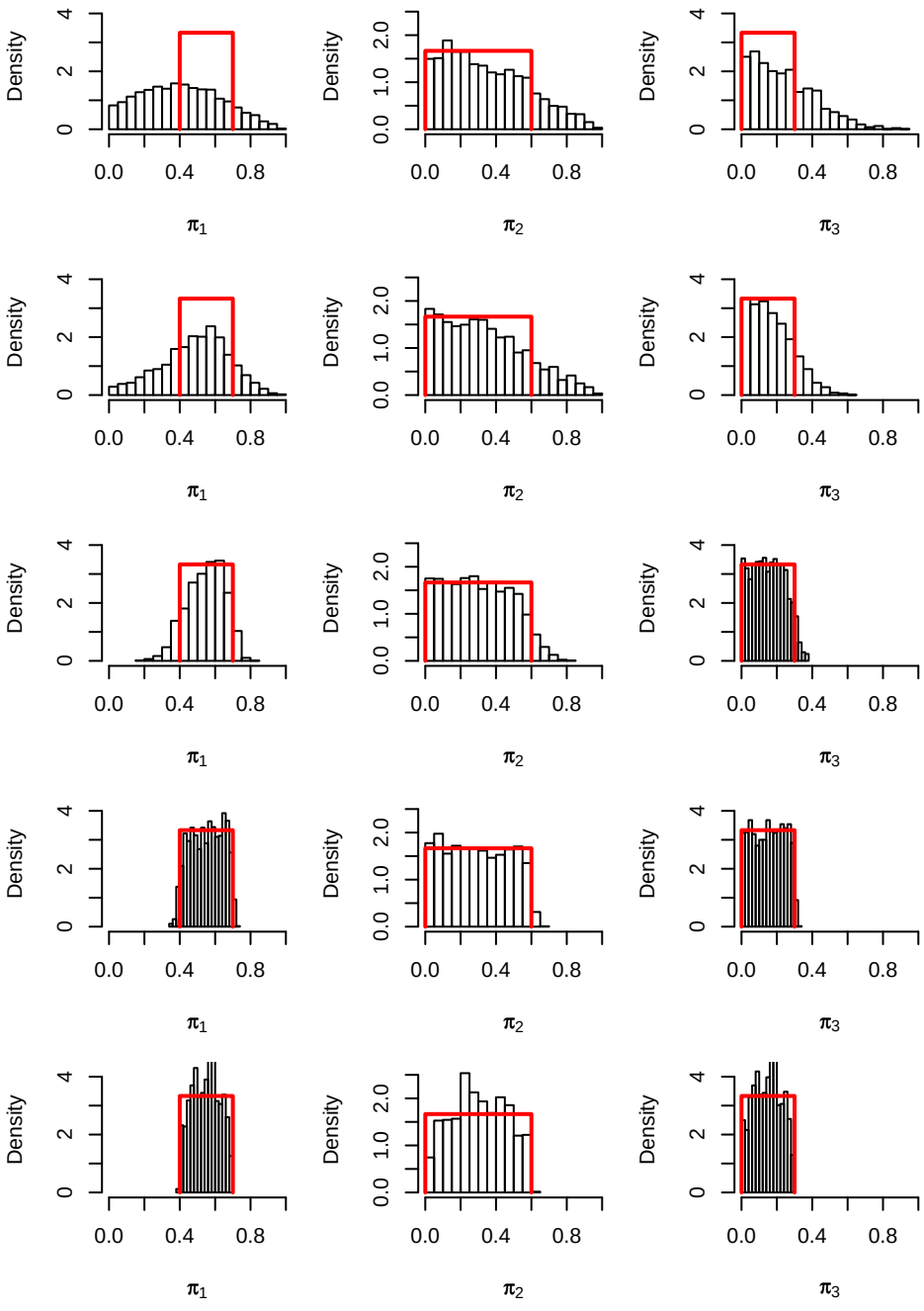

Figure 5.3: Implicit representation IsoSource limiting argument, distributions of (π_1, π_2, π_3) with the mixture and source distributions as in Figure 5.1. Each row corresponds, from top to bottom, to the five scenarios for $\sigma_0^2 = \{4, 1, 0.1, 0.01, 0.0001\}$. The red box indicates the no-variation solution approximated by IsoSource.

5.2.2 Two-step estimation

There are times when it is useful to adopt a two-step approach to parameter estimation, such as to simplify the model to improve convergence or decrease computation time, or when data summaries are the only information available, or to help assess whether the data are consistent with the model. To make this concrete, suppose inference on the isotope ratio means, $\underset{\sim}{\delta}_s$, $s = 1, \ldots, S$, were of interest. Estimation can be done in one step using joint posterior (5.28) using (5.1), (5.3), and (5.4), or in two steps by (1) first obtaining the "posterior" for $\underset{\sim}{\delta}_s$ using (5.7) and (5.8) as step one, (2) then using that as a prior in (5.28) to jointly estimate all model parameters in step two. In particular, note that in (5.28)

$$\begin{aligned} p(\mathbf{D}^*|\delta^*, \Sigma^*)p(\delta^*, \Sigma^*) &= p(\delta^*, \Sigma^*|\mathbf{D}^*)p(\mathbf{D}^*) \\ &= p(\delta^*|\Sigma^*, \mathbf{D}^*)p(\Sigma^*|\mathbf{D}^*)p(\mathbf{D}^*), \end{aligned}$$

so that (5.28) is equivalent to

$$\begin{aligned} p(\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta^*, \Sigma^*|\mathbf{b}^*, \mathbf{D}^*) &\propto p(\mathbf{b}^*|\delta^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi})p(\Sigma_{\underset{\sim}{b}})p(\underset{\sim}{\pi})p(\delta^*, \Sigma^*|\mathbf{D}^*)p(\mathbf{D}^*) \\ &\propto p(\mathbf{b}^*|\delta^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi})p(\Sigma_{\underset{\sim}{b}})p(\underset{\sim}{\pi}) \\ &\quad \times p(\delta^*|\Sigma^*, \mathbf{D}^*)p(\Sigma^*|\mathbf{D}^*). \end{aligned} \tag{5.29}$$

In two-step estimation we first compute $p(\delta^*, \Sigma^*|\mathbf{D}^*)$ and then use it as the "prior" for (δ^*, Σ^*) in conjuction with the joint prior of $(\Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi})$ in (5.29) to compute the posterior for all parameters. The posterior implied by (5.29) is identical to that given in (5.28). This argument holds regardless of the specified distributions, only requiring the independence structure of the model. For all parameters in δ^*, note that the conditional distributions for $\underset{\sim}{\delta}_s$, $\underset{\sim}{\delta}_{\mathrm{D}s}$, and $\underset{\sim}{\delta}_{\mathrm{T}s}$ given δ^* and Σ^* depend on $\mathbf{D}^*$ only through $\mathbf{d}^*_s$, $\mathbf{d}^*_{\mathrm{D}s}$, and $\mathbf{d}^*_{\mathrm{T}s}$, respectively, for $s = 1, \ldots, S$. For our specific distributional assumptions, this reduction is precisely what is given by (5.7) and (5.8) in the normal hierarchical model independently for each source, $s = 1, \ldots, S$, for δ^*_{S},

δ^*_{D}, and δ^*_{T} in δ^*. The same result will hold for the explicit representation given in Section 5.3 since the manipulation only deals with δ^*.

5.2.3 Three-step estimation

An intermediate step can be added to the two-step estimation to combine the three distributions for the source and diet experiment means, $\underset{\sim}{\delta}_s$, $\underset{\sim}{\delta}_{\mathrm{D}s}$, and $\underset{\sim}{\delta}_{\mathrm{T}s}$, into a single distribution for $\underset{\sim}{\delta}'_s$, $s = 1, \ldots, S$. Alternatively, just the diet experiment distributions can be combined for a distribution for discrimination. One benefit of combining the source parameters is to improve convergence of the MCMC chains. In some situations the sum converges, $\underset{\sim}{\delta}'_s = \underset{\sim}{\delta}_s + (\underset{\sim}{\delta}_{\mathrm{T}s} - \underset{\sim}{\delta}_{\mathrm{D}s})$, but the individual components $(\underset{\sim}{\delta}_s, \underset{\sim}{\delta}_{\mathrm{D}s}, \underset{\sim}{\delta}_{\mathrm{T}s})$ may synchronously vacillate between multiple modes (later, see Figure 5.11 for an example of this behavior). Below we show that for the parameters of most interest, inference is the same.

Starting from the two-step estimation posterior (5.29), recognise that $p(\mathbf{b}^*|\delta^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi})$ depends on δ^* only through $\delta'^* = [\underset{\sim}{\delta}'_1, \ldots, \underset{\sim}{\delta}'_S] = \delta^*_{\mathrm{S}} + (\delta^*_{\mathrm{T}} - \delta^*_{\mathrm{D}})$. Thus, through reparameterization (5.29) is transformed into

$$\begin{aligned} p(\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta'^*, \delta^*_{\mathrm{T}}, \delta^*_{\mathrm{D}}, \Sigma^*|\mathbf{b}^*, \mathbf{D}^*) \;\propto\; & p(\mathbf{b}^*|\delta'^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi}) p(\Sigma_{\underset{\sim}{b}}) p(\underset{\sim}{\pi}) \\ & \times p(\delta'^*, \delta^*_{\mathrm{T}}, \delta^*_{\mathrm{D}}|\Sigma^*, \mathbf{D}^*) p(\Sigma^*|\mathbf{D}^*), \end{aligned}$$

and by integrating out δ^*_{T} and δ^*_{D} we have

$$\begin{aligned} p(\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta'^*, \Sigma^*|\mathbf{b}^*, \mathbf{D}^*) \;&\propto\; \int_{\delta^*_{\mathrm{T}}} \int_{\delta^*_{\mathrm{D}}} p(\underset{\sim}{\pi}, \Sigma_{\underset{\sim}{b}}, \delta'^*, \delta^*_{\mathrm{T}}, \delta^*_{\mathrm{D}}, \Sigma^*|\mathbf{b}^*, \mathbf{D}^*) \, \mathrm{d}\delta^*_{\mathrm{D}} \, \mathrm{d}\delta^*_{\mathrm{T}} \\ &\propto\; p(\mathbf{b}^*|\delta'^*, \Sigma_{\underset{\sim}{b}}, \underset{\sim}{\pi}) p(\Sigma_{\underset{\sim}{b}}) p(\underset{\sim}{\pi}) \\ &\quad \times p(\delta'^*|\Sigma^*, \mathbf{D}^*) p(\Sigma^*|\mathbf{D}^*). \end{aligned} \tag{5.30}$$

Use of (5.30) simplifies the two-step model in (5.29) at the loss of separate estimates for δ^*_{S}, δ^*_{T}, and δ^*_{D}. The key point is that inference for $\underset{\sim}{\pi}$ and $\Sigma_{\underset{\sim}{b}}$ depends on δ^*

only through δ'^*, thus with respect to these primary parameters, the two-step and three-step estimation methods are equivalent.

To summarize, the three-step estimation (1) first obtains the "posterior" for $(\delta^*_{\mathrm{S}}, \delta^*_{\mathrm{D}}, \delta^*_{\mathrm{T}}|\mathbf{D}^*, \Sigma^*)$ using (5.7) and (5.8) as step one, (2) then reparameterizes the model in terms of δ'^* as discussed above to give $p(\delta'^*|\Sigma^*, \mathbf{D}^*)$ as step two, and (3) finally uses the reparameterized source components as a prior in (5.30) as step three. In some cases the loss of the ability to estimate δ^*_{S}, δ^*_{D}, and δ^*_{T} as individual components in the three-step estimation may be offset by numerical stability in the MCMC. In other cases the two-step estimation, which we showed was equivalent to the one-step estimation, may be preferred.

5.2.4 Dunlin implicit representation example, two sources and one isotope

We reanalyze the dunlin data with one isotope and two sources introduced in Section 4.1.5. Since this is the first and simplest example, we use this opportunity to illustrate the full model specification in detail. We specify weakly informative source priors and compare the results using a noninformative and an informative prior on the $\underset{\sim}{\pi}$ vector to see the influence of the prior.

5.2.4.1 Source and mixture priors

The defining equation relates the population mean isotope ratio of the dunlin to the terrestrial and marine source isotope ratios, the isotope ratio discriminations, and the mean diet proportion of each source,

$$\beta = (\delta_1 + (\delta_{\mathrm{T}} - \delta_{\mathrm{D}}))\pi_1 + (\delta_2 + (\delta_{\mathrm{T}} - \delta_{\mathrm{D}}))\pi_2.$$

Following the discussion at the beginning of Section 4.1.5, we can be somewhat informative regarding the mean source proportions for the terrestrial C3 ecosystem (about $-26‰$) and estuarine mudflats that experienced mixed marine and freshwater inputs (about $-13‰$). This information can further help us specify the mixture distribution. If we believe with probability 0.95 that the mixture mean is within this 13‰ range, then $\sigma_b = 13/4 = 3.25$ gives a rough estimate of the standard deviation. Thus, we set $\sigma_{0b}^2 = 3.25^2 = 10.5625$ with a noninformative $\nu_{0b} = 0.1$ (the prior specification is worth ν observations) which results in a gamma distribution for the precision τ_b with $\mathrm{E}[\tau_b] = (\sigma_{0b}^2)^{-1} = 0.0947$ and $\mathrm{Var}[\tau_b] = (\nu_{0b}(\sigma_{0b}^2)^2/2)^{-1} = 1.893$.

The sampling distribution for the dunlin isotope ratio is

$$\begin{aligned} b_j|\underset{\sim}{\pi}, \sigma_b^2, \delta_1, \delta_2, \delta_{\mathrm{T1}}, \delta_{\mathrm{D1}} &\overset{\text{ind}}{\sim} \text{Normal}(\beta, \sigma_b^2), \quad j = 1, \ldots, J = 234 \\ (\sigma_b^2)^{-1} = \tau_b|\sigma_{0b}^2, \nu_{0b} &\sim \text{Gamma}(0.05, 0.528). \end{aligned} \tag{5.31}$$

For the terrestrial source, C3 plants are between $-33‰$ and $-23‰$ with a mean of about $\delta_{01} = -26‰$. Using the larger difference between the mean and the endpoints (7‰) and assuming that the true mean is within two standard deviations with probability 0.95, let $\sigma_{0\delta_1} = 3.5$, $\nu_{0\delta_1} = 0.1$, and $\upsilon_{0\delta_1} = 1$. The sampling and prior distributions for the terrestrial source are

$$\begin{aligned} d_{1k}|\delta_1, \sigma_{d_1}^2 &\overset{\text{ind}}{\sim} \text{Normal}(\delta_1, \sigma_{d_1}^2), \quad k = 1, \ldots, K_1 = 16 \\ \delta_1|\delta_{01}, \sigma_{\delta_1}^2, \upsilon_{0\delta_1} &\sim \text{Normal}(-26, \sigma_{d_1}^2/1) \\ (\sigma_{d_1}^2)^{-1} = \tau_{d_1}|\sigma_{0\delta_1}^2, \nu_{0\delta_1} &\sim \text{Gamma}(0.05, 0.6125). \end{aligned}$$

For the marine source, C4 plants are between $-16‰$ and $-9‰$ with a mean of about $\delta_{02} = -13‰$. Using the larger difference between the mean and the endpoints (4‰) and assuming that the true mean is within two standard deviations with probability 0.95, let $\sigma_{0\delta_2} = 2$, $\nu_{0\delta_2} = 0.1$, and $\upsilon_{0\delta_2} = 1$. The sampling and prior distributions for

the marine source are

$$\begin{aligned} d_{2k}|\delta_2, \sigma^2_{d_2} &\overset{\text{ind}}{\sim} \text{Normal}(\delta_2, \sigma^2_{d_2}), \quad k = 1, \ldots, K_2 = 21 \\ \delta_2|\delta_{02}, \sigma^2_{\delta_2}, \upsilon_{0\delta_2} &\sim \text{Normal}(-13, \sigma^2_{d_2}/1) \\ (\sigma^2_{d_2})^{-1} = \tau_{d_2}|\sigma^2_{0\delta_2}, \nu_{0\delta_2} &\sim \text{Gamma}(0.05, 0.2). \end{aligned}$$

Discrimination is estimated from a single-source diet experiment, where we take the sampling and prior distributions for the C3 plant control diet the same as the terrestrial source above,

$$\begin{aligned} d_{\text{D}k}|\delta_\text{D}, \sigma^2_{d_\text{D}} &\overset{\text{ind}}{\sim} \text{Normal}(\delta_\text{D}, \sigma^2_{d_\text{D}}), \quad k = 1, \ldots, K_\text{D} = 28 \\ \delta_\text{D}|\delta_{0\text{D}}, \sigma^2_{\delta_\text{D}}, \upsilon_{0\delta_\text{D}} &\sim \text{Normal}(-26, \sigma^2_{d_\text{D}}/1) \\ (\sigma^2_{d_\text{D}})^{-1} = \tau_{d_\text{D}}|\sigma^2_{0\delta_\text{D}}, \nu_{0\delta_\text{D}} &\sim \text{Gamma}(0.05, 0.6125). \end{aligned}$$

The resulting animal tissue (blood) is expected to differ by about 2‰ (Herrera M et al., 2006), thus we shift the mean by the expected difference and specify the same uncertainty as in the diet,

$$\begin{aligned} d_{\text{T}k}|\delta_\text{T}, \sigma^2_{d_\text{T}} &\overset{\text{ind}}{\sim} \text{Normal}(\delta_\text{T}, \sigma^2_{d_\text{T}}), \quad k = 1, \ldots, K_\text{T} = 4 \\ \delta_\text{T}|\delta_{0\text{T}}, \sigma^2_{\delta_\text{T}}, \upsilon_{0\delta_\text{T}} &\sim \text{Normal}(-24, \sigma^2_{d_\text{T}}/1) \\ (\sigma^2_{d_\text{T}})^{-1} = \tau_{d_\text{T}}|\sigma^2_{0\delta_\text{T}}, \nu_{0\delta_\text{T}} &\sim \text{Gamma}(0.05, 0.6125). \end{aligned}$$

5.2.4.2 Two-step estimation for dunlin example

For two-step estimation as in Section 5.2.2, the data and priors are combined and their updated values are for terrestrial $\delta_{16,1} = -25.40007$, $\upsilon_{16,\delta_1} = 17$, $\sigma_{16,\delta_1} = 1.306338$, and $\nu_{16,1} = 16.1$, for marine $\delta_{21,2} = -13.93551$, $\upsilon_{21,\delta_2} = 22$, $\sigma_{21,\delta_2} = 3.171903$, and $\nu_{21,2} = 21.1$, for diet $\delta_{28,\text{D}} = -24.67483$, $\upsilon_{28,\delta_\text{D}} = 29$, $\sigma_{28,\delta_\text{D}} = 0.4973354$, and $\nu_{28,\text{D}} = 28.1$, and for tissue $\delta_{4,\text{T}} = -23.422$, $\upsilon_{4,\delta_\text{T}} = 5$, $\sigma_{4,\delta_\text{T}} = 0.6907896$, and $\nu_{4,\text{T}} = 4.1$.

5.2.4.3 Three-step estimation for dunlin example

Building on the two-step estimation for the dunlin example above, the three-step estimation in Section 5.2.3 uses the estimated group means for $\underset{\sim}{\delta}'_1$ and $\underset{\sim}{\delta}'_2$ as -24.14725 and -12.68268. The remaining parameters (υ_{16,δ_1}, σ_{16,δ_1}, $\nu_{16,1}$, υ_{21,δ_2}, σ_{21,δ_2}, $\nu_{21,2}$, $\upsilon_{28,\delta_\mathrm{D}}$, $\sigma_{28,\delta_\mathrm{D}}$, $\nu_{28,\mathrm{D}}$, $\upsilon_{4,\delta_\mathrm{T}}$, $\sigma_{4,\delta_\mathrm{T}}$, and $\nu_{4,\mathrm{T}}$) are the same as in Section 5.2.4.2.

5.2.4.4 Noninformative terrestrial diet prior

The noninformative prior for the mean diet vector has $\underset{\sim}{\pi}_0 = (0.5, 0.5)^\top$ and $\alpha_{0\pi} = 1$,

$$\underset{\sim}{\pi}|\underset{\sim}{\pi}_0, \alpha_{0\pi}, S \;\sim\; \text{Dirichlet}(2(0.5, 0.5)^\top).$$

The prior on the terrestrial diet, $\pi_1 = 1 - \pi_2$, is uniform over $[0, 1]$ in this noninformative case, $\pi_1 \sim \text{Beta}(1, 1)$.

5.2.4.5 Informative terrestrial diet prior

We also specify an informative prior on the diet vector. Keith Hobson recalled a guess before the data collection regarding the terrestrial diet being about 0.80 probability between 0.05 and 0.25. A prior for $\underset{\sim}{\pi}$ reflecting this belief can be constructed by defining $\underset{\sim}{\pi}_0 = (0.14, 0.86)^\top$ and $\alpha_{0\pi} = 9$ to give a prior distribution $\pi_1 \sim \text{Beta}(2.52, 15.48)$, with 0.10 probability on either side of the interval, this prior shown in Figure 5.4 has the equivalent information of 18 observations,

$$\underset{\sim}{\pi}|\underset{\sim}{\pi}_0, \alpha_{0\pi}, S \;\sim\; \text{Dirichlet}(18(0.14, 0.86)^\top).$$

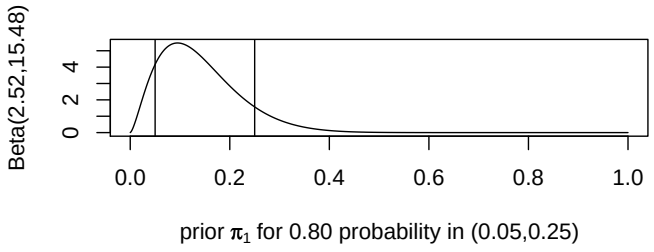

Figure 5.4: Dunlin Bayes implicit representation example, informative terrestrial diet prior.

5.2.4.6 Jeffreys prior

The two-step estimation method from Section 5.2.2 is necessary when using Jeffreys prior in WinBUGS because Jefferys prior is not proper. However, conditional on $K_s > I$, $K_{\mathrm{D}s} > I$, and $K_{\mathrm{T}s} > I$, $s = 1, \ldots, S$, the distribution from the first step in (5.7) and (5.8) will be proper and estimation using the Jeffreys prior may proceed in the second step. Though the results are not shown, the use of Jeffreys prior gives similar results as using the specified priors above because the small degrees of freedom makes the influence of the priors very weak on the posteriors.

5.2.4.7 Results

For both analyses, using the noninformative and informative diet prior, we ran 5 chains for 110000 iterations prethinned every 10. Based on the convergence in traceplots we use iterates 10001–110000 further thinning every 10 to reduce the sample size to 50000. We checked that the Monte Carlo error for each parameter of interest was less than about 5% of the sample standard deviation (p. 55, Spiegelhalter et al., 2003). Results were effectively the same whether one-step, two-step, or three-step estimation was used. Figure 5.5 shows the MCMC chains have converged and are mixing well for a selection of parameters using the noninformative diet prior.

For the model using the noninformative diet prior, Table 5.1 provides posterior summaries for all the parameters and Figure 5.6 shows the posterior distributions with highest posterior density (HPD) interval. Note that the posterior summaries include an equal-tailed interval which is slightly different from the HPD. First, note that the posterior estimates for the isotope ratio parameters are similar to what was given by the method of moment estimates in Section 4.1.5 with a little additional variance due to the noninformative priors. The additional variance is reflected in $\underset{\sim}{\pi}$ being a little more variable than the frequentist methods in Chapter 4, $\text{SD}(\pi_1) = 0.0462$ for Delta method versus the posterior $\text{SD}(\pi_1) = 0.0604$ for Bayes implicit representation. Much of this additional variance may be reduced with more informative priors on the parameters and this information may be available. Also, this difference may be due to the delta method underestimating the variability.

For model comparison we use the Deviance Information Criterion (DIC), a Bayesian method for model comparison, where the model with the smallest DIC is estimated to be the model that would best predict a replicate dataset which has the same structure as that currently observed (Spiegelhalter et al., 2002). Additionally, the pD is the effective number of parameters in the model, and for non-hierarchical models with little prior information pD should be approximately the true number of parameters (our model is hierarchical). The true number of parameters in the implicit representation model is 11. The implicit representation with a noninformative diet prior has a DIC of 1422.42 with a pD of 9.75685, underestimating the number of parameters.

Table 5.2 provides selected summaries using the informative diet prior. The implicit representation with informative priors has a DIC of 1423.24 with a pD of 9.19569, similar as with the noninformative prior.

Table 5.3 and Figure 5.7 compare the Bayes implicit representation results with noninformative and informative prior on $\underset{\sim}{\pi}$. The noninformative prior gives similar results to the frequentist method, but with larger variance. The informative prior

draws the distribution slightly towards the prior value and inflates the variance even more. The posterior distribution for π_1 has larger variance with the informative prior than the noninformative prior, possibly reflecting that the informative prior was concentrated in a different location than what the data suggested.

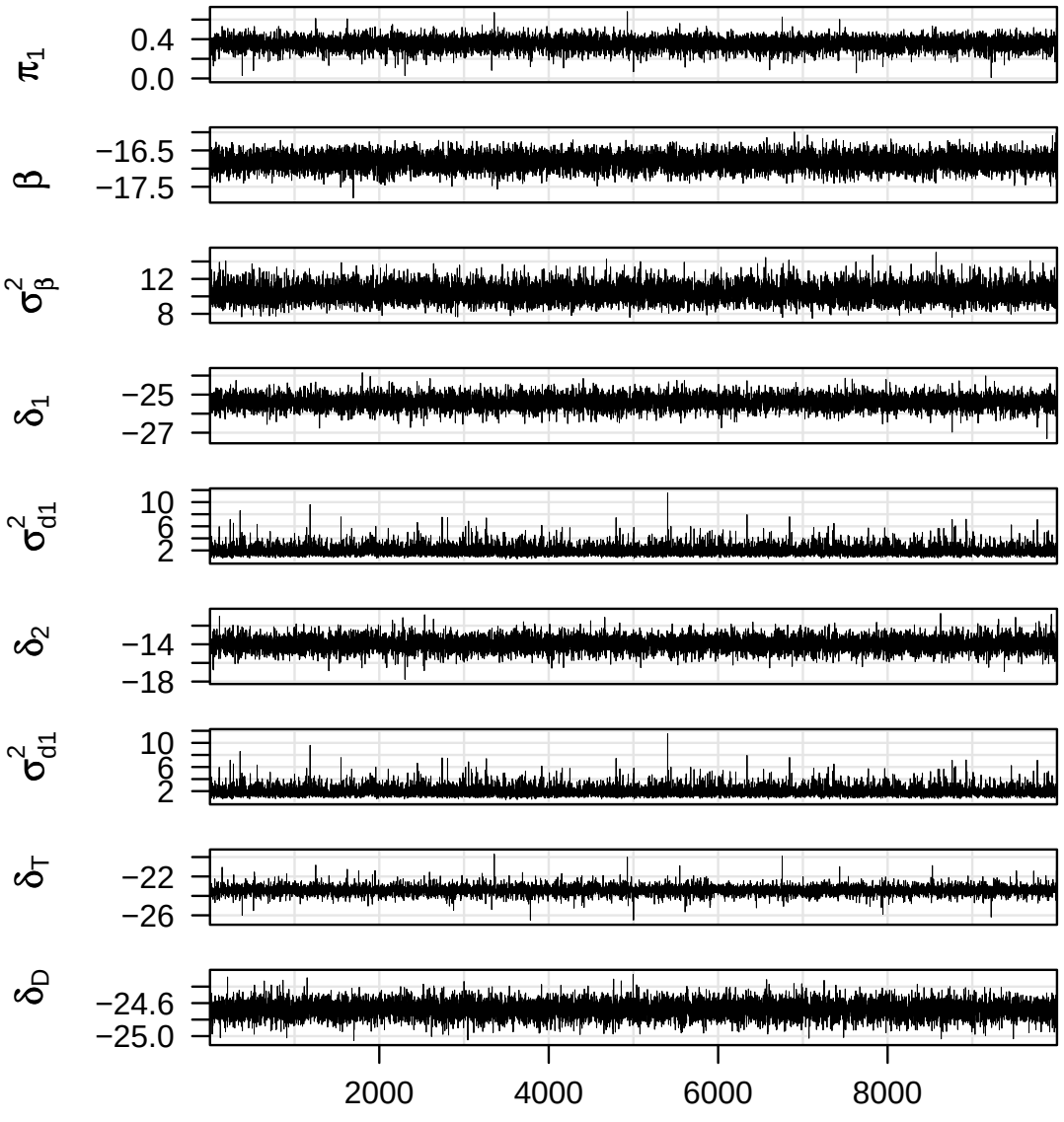

Figure 5.5: Dunlin Bayes implicit representation results for noninformative diet prior, trace plots for MCMC chain 1 of 5 using noninformative diet prior. Parameters shown are (1) π_1, (2) β, (3) σ^2_β, (4) δ_1, (5) $\sigma^2_{d_1}$, (6) δ_2, (7) $\sigma^2_{d_2}$, (8) δ_{T}, and (9) δ_{D}.

Table 5.1: Dunlin Bayes implicit representation posterior estimates with noninformative diet prior.

Parameter	mean	sd	MC error	val2.5pc	median	val97.5pc
π_1	0.356	0.0604	0.000272	0.23	0.358	0.468
π_2	0.644	0.0604	0.000272	0.532	0.642	0.77
β	−16.8	0.211	0.000902	−17.2	−16.8	−16.4
σ^2_β	10.4	0.964	0.00442	8.64	10.3	12.4
δ_1	−25.4	0.338	0.00147	−26.1	−25.4	−24.7
$\sigma^2_{d_1}$	1.95	0.789	0.00354	0.95	1.78	3.95
δ_2	−14.0	0.71	0.00304	−15.4	−14.0	−12.6
$\sigma^2_{d_2}$	11.1	3.77	0.0174	5.93	10.4	20.5
δ_{T}	−23.4	0.428	0.00205	−24.3	−23.4	−22.6
$\sigma^2_{d_{\mathrm{T}}}$	0.933	2.45	0.011	0.174	0.568	3.84
δ_{D}	−24.7	0.0961	0.000404	−24.9	−24.7	−24.5
$\sigma^2_{d_{\mathrm{D}}}$	0.266	0.0766	0.00033	0.156	0.254	0.451
δ'_1	−24.1	0.553	0.0025	−25.2	−24.1	−23.1
δ'_2	−12.7	0.838	0.0038	−14.4	−12.7	−11.1

Table 5.2: Dunlin Bayes implicit representation posterior estimates with informative diet prior. The other parameters are effectively the same as the noninformative case in Table 5.1.

Parameter	mean	sd	MC error	val2.5pc	median	val97.5pc
π_1	0.294	0.065	0.000298	0.149	0.299	0.407

Table 5.3: Dunlin Bayes implicit representation (Irep) posterior estimates for proportion contribution of terrestrial sources to dunlin diet. Freqentist result is from Table 4.5, and for Bayesian analyses, $\hat{\pi}_1$ and and $\mathrm{Var}[\hat{\pi}_1]$ are the posterior mean and variance.

Method	$\hat{\pi}_1$	$\mathrm{Var}[\hat{\pi}_1]$	$\mathrm{SD}[\hat{\pi}_1]$	95% Int
Freq mixture and source	0.367	0.00213	0.0462	(0.277, 0.458)
Irep noninformative	0.356	0.00365	0.0604	(0.230, 0.468)
Irep informative	0.294	0.00422	0.0650	(0.149, 0.407)

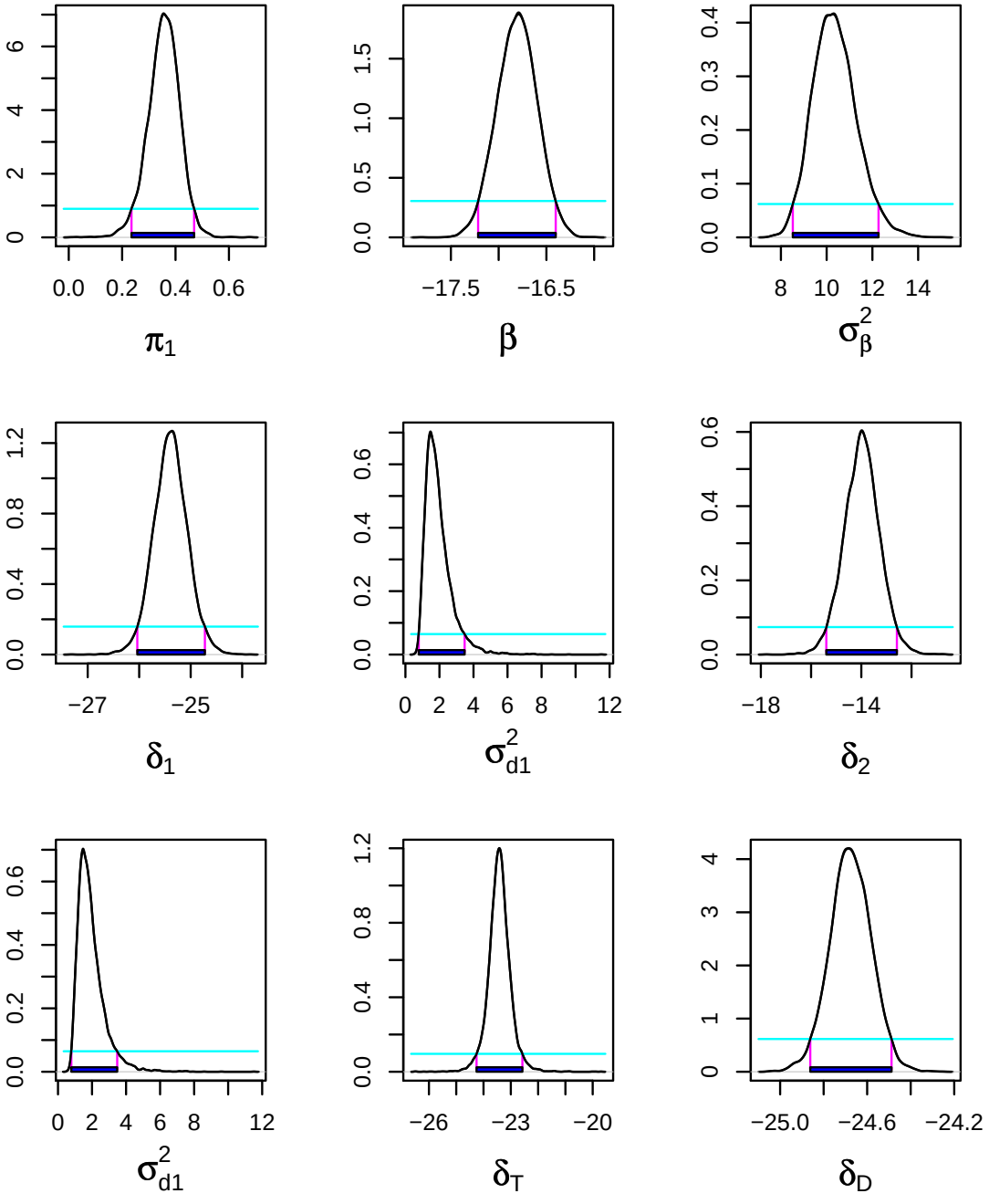

Figure 5.6: Dunlin Bayes implicit representation results for noninformative diet prior, posteriors for MCMC chain 1 of 5 with 95% HPD interval.

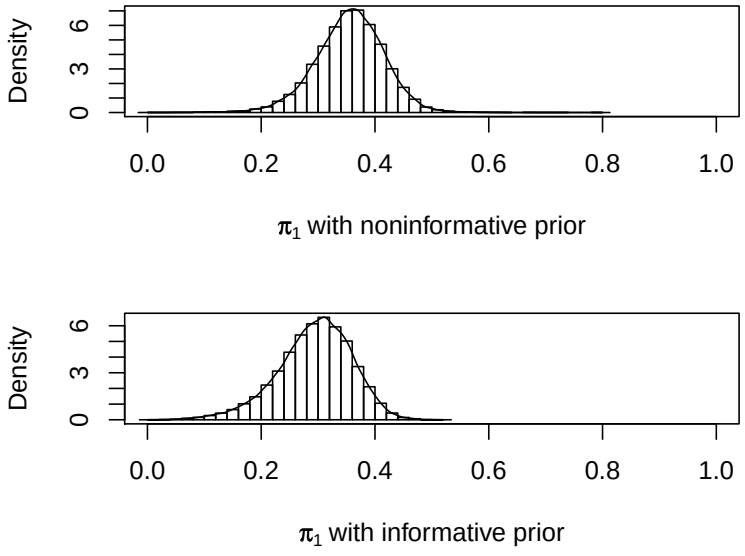

Figure 5.7: Dunlin Bayes implicit representation posterior estimates for proportion contribution of terrestrial sources to dunlin diet using noninformative and informative diet prior.

5.3 Explicit representation, subject-specific diets

In Section 5.2, the implicit representation modeled the observed mixture isotope ratio value centered at the mean isotope ratio conditional on the mean population diet. Here, the two primary contributions in the explicit representation are modeling subject-specific diets and modeling isotope ratio measurement error. This model is based on first principles and is the model we favor for analysis when the data conform to the more strict assumptions of this model. The explicit representation models each observed mixture isotope ratio value centered at the subject-specific mean isotope ratio subject to measurement error (Jardine and Cunjak, 2005). Because differences in diet lead to differences in the isotope ratios of individuals within a population (Angerbjörn et al., 1994; Matthews and Mazumder, 2004; Urton and Hobson, 2005; Araújo et al., 2007; Layman et al., 2007), the subject-specific mean isotope ratio is conditional on the subject-specific diet. These subject-specific diets, in turn, follow a distribution centered at the population mean diet with between-individual diet variation estimated with a precision parameter.

We will show that the explicit representation is sensitive to certain model assumptions, while the implicit representation is more robust to violations and may not indicate that problems exist. An important condition in this method is that the individual mixture diets, $\underset{\sim}{\pi}_j$, $j = 1, \ldots, J$, are probability vectors, which is equivalent to all individual mixture isotope ratios, $\underset{\sim}{\beta}_j$, $j = 1, \ldots, J$, being in the convex hull of the discrimination-corrected source values, $\underset{\sim}{\delta}'_s$, $s = 1, \ldots, S$. If there are mixture outliers, then the model may compensate by adjusting the $\underset{\sim}{\delta}'_s$ values to satisfy the convex hull condition, which is not desirable. We explore this as part of the dunlin example in Section 5.3.4. In the case of highly variable mixture measurements outside the convex hull of the source means, the implicit representation may be the desirable analysis because it only requires that the mixture mean be inside the convex hull, but the practitioner should consider explanations for the mixtures being beyond the

scope of the source means.

We model the subject-specific mean isotope ratio vector for mixture j, $\underset{\sim}{\beta}_j$, as

$$\underset{\sim}{\beta}_j \;=\; \mathbf{A}\underset{\sim}{\pi}_j, \tag{5.32}$$

assuming that $\mathbf{A}$ is common for all mixtures. Let the isotope ratio observation from mixture j, $\underset{\sim}{b}_j$, depend on its subject-specific mean isotope ratio, $\underset{\sim}{\beta}_j$, with

$$\underset{\sim}{b}_j|\underset{\sim}{\pi}_j, \Sigma_\varepsilon, \delta^* \overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{\beta}_j, \Sigma_\varepsilon), \quad j = 1, \ldots, J \tag{5.33}$$

$$\Sigma_\varepsilon|\Sigma_{0\varepsilon}, \nu_{0\varepsilon} \sim \text{Inv-Wishart}(\Sigma_{0\varepsilon}, \nu_{0\varepsilon}), \tag{5.34}$$

where $\Sigma_{0\varepsilon}$ and $\nu_{0\varepsilon}$ specify the prior covariance matrix of the isotope ratio measurement error (or alternatively, a within-individual variation, or both) and an effective prior sample size. This specification assumes that measurements for mixtures have a common measurement error covariance matrix, Σ_ε. Measurement error is an important parameter in this model which has not been previously estimated separate from other parameters.

As an aside, we note that in the overconstrained case ($S < I + 1$) that $\underset{\sim}{b}_j$ can have a density because of the measurement error (and within-individual) variability. In the absence of measurement error variability the issue of a density for $\underset{\sim}{\beta}_j$ would be more critical as the conditional distribution in (5.33) would be degenerate.

5.3.1 Measurement error

Often there is one measurement per individual, which does not allow for an estimate of Σ_ε. However, as introduced in Section 4.1, measurements using a continuous-flow isotope ratio mass spectrometer (IRMS) use known laboratory standards to calibrate the measurements. For the dunlin example, every five unknowns were

separated by both high and low laboratory standards, one on each end of the isotope ratio range of the measurements (Evans Ogden et al., 2005). Variation is observed in these known standards during the gas analysis and this information can be used to partially inform $\Sigma_{0\varepsilon}$ and $\nu_{0\varepsilon}$. Other sources of measurement error relating to sample collection, storage, and handling, and other conditions not applying to the standards, could be estimated with multiple measurements on individual mixtures. Thus, in every analysis, standards not only calibrate the isotope ratio measurements but are also an important component to assess the measurement error.

There are different ways of using standards to calibrate the measurements and each requires an appropriate method for modeling measurement error. A few ways include (1) using a linear regression fit to the standards over a single measurement run (possibly 10 to 100 measurements), (2) running pairs of identical standards at intervals and doing a "connect-the-means" linear interpolation between each pair, and (3) using splines to smoothly account for slow nonlinear measurement drift over the run. We restrict our attention to (1) because our example data come from labs that use this method, but others may be considered in future work. A linear regression is used by fitting the isotope ratio of the standards to the run sequence number to account for possible measurement drift over the run. The regression line is assigned the known isotope ratio value of the standards as the reference and the unknowns are assigned isotope ratio values relative to this line. Let the residual for the standard, $\underset{\sim}{z}_{rk}$, be the difference of the kth observed standard to the regression line for run r. By construction the residuals have expected value $\underset{\sim}{0}$ and an assumed common variance that may be instrument specific, but might be assumed to be the same over runs for a given instrument. The measurement error is typically much smaller than the between-individual variation.

Let the residuals of laboratory standards, $\underset{\sim}{z}_k$, have

$$\underset{\sim}{z}_k|\Sigma_\varepsilon \overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{0}, \Sigma_\varepsilon), \quad k = 1, \ldots, K_\varepsilon,$$

as in linear regression method (1) above. Let the error covariance have the prior in (5.34). Let $\mathbf{z}^* = [\underset{\sim}{z}_1, \ldots, \underset{\sim}{z}_{K_z}]$. This specification results in the posterior for the measurement covariance as in (5.8)

$$\Sigma_\varepsilon | \mathbf{z}^*, \Sigma_{0\varepsilon}, \nu_{0\varepsilon} \sim \text{Inv-Wishart}(\Sigma_{K_z\varepsilon}, \nu_{K_z\varepsilon}), \tag{5.35}$$

where $\nu_{K_z\varepsilon} = \nu_{0\varepsilon} + K_z$ and $\Sigma_{K_z\underset{\sim}{\delta}_\varepsilon} = \Sigma_{0\underset{\sim}{\delta}_\varepsilon} + (K_z - 1)\mathbf{S}_{\underset{\sim}{d}_z}$, where $(K_z - 1)\mathbf{S}_{\underset{\sim}{d}_z} = \sum_{k=1}^{K_z} \underset{\sim}{z}_k \underset{\sim}{z}_k^\top$. This gives an easy way to combine the measurement error of the current experiment with the (possibly instrument-specific) prior error assessment to inform the current analysis. Also, it gives the practitioner a choice of simply using (5.35) as the measurement error prior, or both the prior and likelihood in the modeling.

5.3.1.1 Univariate measurement error

Similar to as in Section 5.1.1.2, for the univariate case ($I = 1$) let residuals of laboratory standards, z_k, have

$$z_k | \sigma_\varepsilon^2 \overset{\text{ind}}{\sim} \text{Normal}(\underset{\sim}{0}, \sigma_\varepsilon^2), \quad k = 1, \ldots, K_z. \tag{5.36}$$

The conjugate prior distribution for the variance is

$$\sigma_\varepsilon^2 | \sigma_{0\varepsilon}^2, \nu_{0\varepsilon} \sim \text{Scaled-inv-}\chi^2(\sigma_{0\varepsilon}^2, \nu_{0\varepsilon}). \tag{5.37}$$

The posterior can be written,

$$\sigma_\varepsilon^2 | \mathbf{z}^*, \sigma_{0\varepsilon}^2, \nu_{0\varepsilon} \sim \text{Scaled-inv-}\chi^2(\sigma_{K_z}^2, \nu_{K_z}), \tag{5.38}$$

where

$$\begin{aligned} \nu_{K_z\varepsilon} &= \nu_{0\varepsilon} + K_z \\ \sigma_{K_z\varepsilon}^2 &= \frac{\nu_{0\varepsilon}}{\nu_{K_z\varepsilon}} \sigma_{0\varepsilon}^2 + \frac{K_z - 1}{\nu_{K_z\varepsilon}} s_z^2, \end{aligned}$$

where $s_z^2 = K_z^{-1} \sum_{k=1}^{K_z} z_k^2$.

5.3.2 Subject-specific diets

To model the mean population diet we adopt a hierarchical Dirichlet model, though other models may also work well, such as a hierarchical logit-Normal. Let the subject-specific diets, $\underset{\sim}{\pi}_j$, follow a Dirichlet distribution

$$\underset{\sim}{\pi}_j|\underset{\sim}{\pi}, \alpha_b, S \overset{\text{ind}}{\sim} \text{Dirichlet}(\alpha_b S\underset{\sim}{\pi}), \quad j = 1, \ldots, J \tag{5.39}$$

$$\alpha_b|\alpha_{01b}, \alpha_{02b} \sim \text{Uniform}(\alpha_{01b}, \alpha_{02b}), \tag{5.40}$$

with α_b representing the between-subject diet precision about the population mean diet, $\underset{\sim}{\pi}$. Larger values of α_b specify that the diets of the individuals are more similar. Hyperparameters $(\alpha_{01b}, \alpha_{02b})$ specify the range of α_b, and may need to be adjusted if the MCMC chain appears restricted by one of the boundaries. Let the population mean diet be as in (5.27) from the implicit representation, $\underset{\sim}{\pi}|\underset{\sim}{\pi}_0, \alpha_{0\pi}, S \sim \text{Dirichlet}(\alpha_{0\pi} S\underset{\sim}{\pi}_0)$.

The use of a shared parameter vector, $\underset{\sim}{\pi}$, introduces a sharing of information between the population and subject-specific distributions. This is appropriate in cases where the underlying distribution for each mixture is believed to be similar, but where some degree of variation is believed to exist. Observations made for one mixture can affect the posterior distribution of $\underset{\sim}{\pi}$, and therefore alter future predictions for observations for other mixtures, which makes sense in a biological system where animals are feeding in a common spatial-temporal domain.

5.3.3 Comparing implicit representation and explicit representation models

To study when the implicit representation or explicit representation model should be preferred we present the results of a small simulation study. In particular, we wish to understand the effects on the ability to estimate the diet proportion relative to

the true mean diet, measurement error magnitude, and whether information is available on the measurement error. We simulated data under the explicit representation model similar to the dunlin example varying the true diet, $\pi_1 = \{0.1, 0.25, 0.5\}$, the measurement error variation $\sigma^2_\varepsilon = \{2, 0.1\}$, and whether the prior for the measurement error, σ^2_ε, was informative or noninformative. We did all cases three times to see whether the results were consistent, and because they were we report the first of the three simulations.

Summarizing the results that follow, when the explicit representation model is the true model then the implicit representation model provides sensible answers about the average population diet, is an easier model to fit with fewer parameters, and the MCMC chains always mix well. The explicit representation model provides more detailed answers for the subject-specific diets and the average population diet at a cost of careful specification of the measurement error. Also, when the measurement error is large, then between-subject diet precision α_b is uninformed by the data and the MCMC chains mix poorly. Fortunately, there is always information on measurement error from the standards included as part of the calibration in the mass spectrometer, and this measurement error is typically very small. Therefore, the precision of the diet estimates can be improved by using the explicit representation.

5.3.3.1 Data simulation and modeling

For the simulation, values of the true mean population diet were taken as $\pi_1 = \{0.1, 0.25, 0.5\}$ and measurement errors were taken as $\sigma^2_\varepsilon = \{2, 0.1\}$. Let the between-subject diet precision be $\alpha_b = 8$. Following the explicit representation model, for the mixture we simulate $J = 100$ subject-specific diets following the distribution

$$\pi_{j1}|\pi_1, \alpha_b, S \sim \text{Beta}(16\pi_1, 16(1-\pi_1)).$$

The top two rows in Figure 5.8 show the distribution of π_{j1} and 100 sampled π_{j1}. For the two sources, we use "true" values given from the dunlin data as in Section 5.2.4.3. The mean defining equation for the mixture isotope ratio gives

$$\beta_{j1} \;=\; -24.14725\pi_{j1} + -12.68268\pi_{j2} \;=\; -11.46456\pi_{j1} - 12.68268.$$

We add a lot or a little measurement error to the mixture isotope ratio measurements given measurement errors $\sigma^2_\varepsilon = \{2, 0.1\}$ via

$$b_{jk}|\delta'_{11}, \delta'_{12}, \pi_{j1}, \sigma^2_\varepsilon \;\sim\; \text{Normal}(\beta_{j1}, \sigma^2_\varepsilon).$$

The bottom two rows in Figure 5.8 show the 100 sampled $\underset{\sim}{b}_{jk}$ with a little or a lot of measurement error. With a little measurement error, the isotope ratio measurements mirror the diets, but with a lot of error the isotope ratio distribution is much more disperse compared to the underlying diet distribution. We considered taking 1 or 5 measurements from each mixture, but taking multiple measurements did not substantially affect the results, so we use 1 measurement as is typically done in practice.

For modeling, we use a three-step prior on the sources from Section 5.2.4.3,

$$\begin{aligned}
\delta'_1|\sigma^2_{d_1}, \upsilon_{16,\delta_1}\ , \sigma^2_{d_\text{D}}, \upsilon_{28,\delta_\text{D}}, \sigma^2_{d_\text{T}}, \upsilon_{4,\delta_\text{T}} & \\
&\sim \text{Normal}(-24.14725, \sigma^2_{16,\delta_1}/17 + \sigma^2_{28,\delta_\text{D}}/29 + \sigma^2_{4,\delta_\text{T}}/5) \\
\delta'_2|\sigma^2_{d_2}, \upsilon_{21,\delta_2}\ , \sigma^2_{d_\text{D}}, \upsilon_{28,\delta_\text{D}}, \sigma^2_{d_\text{T}}, \upsilon_{4,\delta_\text{T}} & \\
&\sim \text{Normal}(-12.68268, \sigma^2_{21,\delta_2}/22 + \sigma^2_{28,\delta_\text{D}}/29 + \sigma^2_{4,\delta_\text{T}}/5) \\
\sigma^2_{d_1}|\sigma^2_{16,\delta_1}, \nu_{16,1} &\sim \text{Scaled-inv-}\chi^2(1.306338^2, 16.1) \\
\sigma^2_{d_2}|\sigma^2_{21,\delta_2}, \nu_{21,2} &\sim \text{Scaled-inv-}\chi^2(3.171903^2, 21.1) \\
\sigma^2_{d_\text{D}}|\sigma^2_{28,\delta_\text{D}}, \nu_{28,\text{D}} &\sim \text{Scaled-inv-}\chi^2(0.4973354^2, 28.1) \\
\sigma^2_{d_\text{T}}|\sigma^2_{4,\delta_\text{T}}, \nu_{4,\text{T}} &\sim \text{Scaled-inv-}\chi^2(0.6907896^2, 4.1).
\end{aligned}$$

We put flat priors are put on both π_1 and α_b

$$\begin{aligned} \pi_1|\pi_{01}, \alpha_{0\pi}, S &\sim \text{Beta}(1,1), \quad \text{and} \\ \alpha_b &\sim \text{Uniform}(0.2, 20). \end{aligned}$$

For measurement error, we consider either an informative prior on σ^2_ε

$$\sigma^2_\varepsilon|\sigma^2_{0\varepsilon}, \nu_{0\varepsilon} \sim \begin{cases} \text{Scaled-inv-}\chi^2(2, 25) & \text{if the known variance is 2} \\ \text{Scaled-inv-}\chi^2(0.1, 25) & \text{if the known variance is 0.1} \end{cases}$$

or a less informative prior

$$\sigma^2_\varepsilon|\sigma^2_{0\varepsilon}, \nu_{0\varepsilon} \sim \text{Scaled-inv-}\chi^2(0.5, 1).$$

For modeling using the implicit representation model for comparison, the prior on the precision around β_1 is defined as in the dunlin example as

$$(\sigma^2_b)^{-1} = \tau_b|\sigma^2_{0b}, \nu_{0b} \sim \text{Gamma}(0.05, 0.528). \tag{5.41}$$

Using WinBUGS we simulate 110,000 posterior samples prethinned by 10, excluding the first 10,000 as burn-in and keeping every 10th thereafter for 10,000 posterior samples.

Table 5.4 summarizes four comparisons between the implicit representation (irep) and explicit representation (erep) models. We are primarily interested in how the models compare in the inference of the population mean π_1. We'll see how the precision parameters in the models are influenced by the amount of measurement error in the data and whether we know *a priori* what that measurement error is, and how this influences our inference of π_1. The two models have different parameters, but there is an interpretive relationship between the parameters in the two models. For example, σ^2_b has similar information as σ^2_ε and α_b taken together. The true number of parameters is 2 in the irep model and 103 in the erep model.

In all four cases, both the irep and erep estimate π_1 accurately, all within 1 posterior standard deviation of the true value. When the measurement error is large ($\sigma_\varepsilon^2 = 2$), the erep model does not perform well, even when we have prior information on σ_ε^2. In the cases labeled "11i" and "31i" in the table (true π_1 is 0.1 and 0.5, respectively), the erep and irep are effectively the same in the sense that in the erep the large α_b indicates the subject-specific diets are concentrated at the mean diet $\underset{\sim}{\pi}$ and the posterior isotope ratio measurement error variance σ_ε^2 in the erep is similar to the posterior variation between isotope ratios σ_b^2 in the irep. The prior for α_b was modified to Uniform$(0.2, 2000)$ because the posterior was clearly restricted by the upper bound, but even at 2000 it was restricted. The between-diet precision α_b reflected the prior, being uniform with mean nearly 1000 (true α_b is 8), which makes the diets concentrate very tightly around π_1, so the mixture isotope ratio variation is not attributed to differences in individual diets but almost entirely to measurement error about a common mean, π_1. Thus, the posterior mean for measurement error parameter σ_ε^2 of the erep is close to the posterior mean for the mixture isotope ratio variation σ_b^2 of the irep.

When the measurement error is small ($\sigma_\varepsilon^2 = 0.1$), in the cases labeled "12n" and "22i" in the table (true π_1 is 0.1 and 0.25, respectively), the erep performs much better than the irep. The erep estimates the between-diet precision α_b well. When there is a noninformative prior on σ_ε^2 ("12n"), the posterior mean of the measurement error variance is close to the true error variance σ_ε^2, much better than when the error was large. And the measurement error is correctly estimated very precisely with an informative prior ("22i").

Table 5.4: Four illustrative examples comparing the implicit representation and explicit representation models for two measurement error scenarios. Values are posterior mean (sd).

	$\pi_1 = 0.1$, $\sigma^2_\varepsilon = 2$, info (11i)		$\pi_1 = 0.5$, $\sigma^2_\varepsilon = 2$, info (31i)	
param	irep	erep	irep	erep
π_1	0.090 (0.052)	0.090 (0.048)	0.480 (0.049)	0.479 (0.047)
σ^2_b	2.015 (0.294)		3.321 (0.476)	
σ^2_ε		1.986 (0.260)		2.982 (0.428)
α_b		991.1 (574.2)		939.1 (600.1)
	$\pi_1 = 0.1$, $\sigma^2_\varepsilon = 0.1$, noninfo (12n)		$\pi_1 = 0.25$, $\sigma^2_\varepsilon = 0.1$, info (22i)	
param	irep	erep	irep	erep
π_1	0.096 (0.053)	0.122 (0.036)	0.223 (0.059)	0.269 (0.036)
σ^2_b	1.052 (0.153)		1.505 (0.217)	
σ^2_ε		0.230 (0.157)		0.115 (0.037)
α_b		9.425 (3.218)		10.45 (2.584)

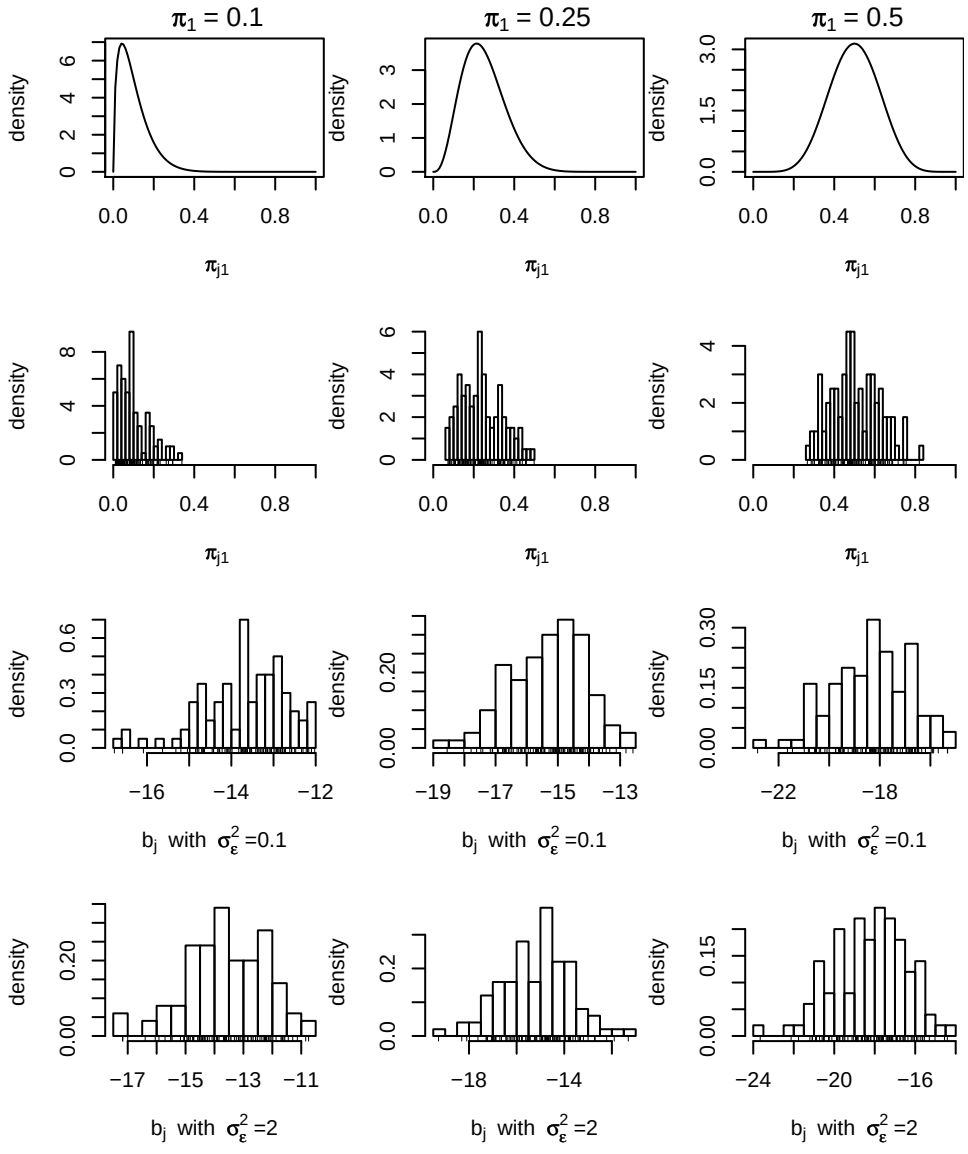

Figure 5.8: Comparing the implicit representation and explicit representation. For $\pi_1 = \{0.1, 0.25, 0.5\}$, (row 1) the distributions for π_{j1}, (row 2) $J = 100$ samples of π_{j1}, and (row 3) samples of b_{jk} with a little measurement error ($\sigma^2_\varepsilon = 0.1$) and (row 4) a lot ($\sigma^2_\varepsilon = 2$).

5.3.4 Dunlin explicit representation example, two sources and one isotope

We continue the analysis of the dunlin data from Section 5.2.4 using the explicit representation, using the three-step estimation for the source parameters.

5.3.4.1 Measurement error prior

To assess measurement error, we use the 59 low standards (reference-1) and 35 high standards (reference-2) from 8 runs, fit a linear regression to each reference in each run, and calculate the residuals associated with each regression assuming the error is common among all runs. The sample variation for the standards is $s_z^2 = 0.297^2$ on $K_z = 94$ measurements. Using a prior of $\sigma_{0\varepsilon}^2 = 1$ with a weight of $\nu_{0\varepsilon} = 1$ prior measurements, we calculate the parameters for (5.35),

$$\sigma_\varepsilon^2 | z_1, \ldots, z_{94}, \sigma_{0\varepsilon}^2, \nu_{0\varepsilon} \sim \text{Scaled-inv-}\chi^2(0.313^2, 95). \tag{5.42}$$

The data suggest there is very little measurement error.

5.3.4.2 Results using three-step estimation

We ran 5 chains for 110000 iterations prethinned every 10. Based on the convergence in traceplots we use iterates 10001–110000 further thinning every 10 to reduce the sample size to 50000. We checked that the Monte Carlo error for each parameter of interest was less than about 5% of the sample standard deviation (p. 55, Spiegelhalter et al., 2003). Figure 5.9 shows that the MCMC chains are mixing reasonably well for a selection of parameters.

Table 5.5 provides posterior summaries for most of the parameters in the model (excluding only the remaining π_{j1} and β_{j1}) and Figure 5.10 shows selected posterior

distributions with highest posterior density (HPD) interval. First, note that the posterior estimates for the source discrimination-corrected isotope ratio parameters are different from the implicit representation estimates in Table 5.1. In particular, the explicit representation posterior means for δ_1' and δ_2' are -30.5 and -12.4 compared to -24.1 and -12.7 for the implicit representation. In practice, if the source estimates are much different than what is given by the method of moment estimates in Section 4.1.5, then this should be investigated. The lower source endpoint has been shifted lower, and the upper has been moved higher. This is the key interesting feature in this example so we consider this phenomenon in detail in the following section. Because the source endpoints have shifted, this affects the estimate of the mean diet parameter, π_1, which is 0.249 instead of 0.356. In the trace plots in Figure 5.9, the dynamic relationship between δ_1' and the diet and between-diet precision parameters, π_1 and α_b, can be seen when δ_1' dips three times between iterations 4000 and 7000 and the diet parameters respond.

5.3.4.3 Mixtures outside sources

In the explicit representation, when the isotope ratios of individual mixtures are outside the range of the source means, then the estimation of at least one parameter may be compromised. An individual mixture's diet, $\underset{\sim}{\pi}_j$, is constrained to be a probability vector, and with two sources, as in the dunlin example, this is natural when $\delta_{11}' \leq \beta_{j1} \leq \delta_{12}'$. Looking at carbon in Figure 4.5, there are several dunlin isotope ratio observations less than the lower terrestrial source mean and several just greater than the upper marine source mean. In Table 5.5 we have already seen that the explicit representation posterior means for δ_1' and δ_2' were far from the method of moment estimates. Thus, while the distributions for δ_1' and δ_2' were centered at -24.1 and -12.7, these extreme mixtures outside this range forced these endpoints out so that all $\underset{\sim}{\pi}_j$ were between 0 and 1.

Table 5.5: Dunlin Bayes three-step explicit representation posterior estimates with noninformative prior.

Parameter	mean	sd	MC error	val2.5pc	median	val97.5pc
π_1	0.249	0.0255	0.000713	0.193	0.251	0.294
π_2	0.751	0.0255	0.000713	0.706	0.749	0.807
π_{11}	0.37	0.0366	0.00104	0.288	0.374	0.434
π_{21}	0.308	0.0319	0.000855	0.238	0.31	0.365
π_{31}	0.597	0.0552	0.00169	0.468	0.604	0.684
π_{41}	0.333	0.0337	0.000925	0.259	0.336	0.392
π_{51}	0.375	0.0369	0.00105	0.292	0.378	0.438
α_b	2.55	0.548	0.0148	1.76	2.46	3.9
β_{11}	−19	0.317	0.0015	−19.7	−19	−18.4
β_{21}	−17.9	0.316	0.0014	−18.5	−17.9	−17.3
β_{31}	−23.1	0.315	0.00145	−23.7	−23.1	−22.5
β_{41}	−18.4	0.316	0.00146	−19	−18.4	−17.7
β_{51}	−19.1	0.314	0.00138	−19.7	−19.1	−18.5
σ^2_ε	0.1	0.0147	0.00007	0.0753	0.0985	0.133
δ'_1	−30.5	1.84	0.0621	−35.3	−30.1	−28.3
δ'_2	−12.4	0.186	0.00168	−12.7	−12.4	−11.9
$\sigma^2_{d_1}$	1.96	0.8	0.00367	0.951	1.79	3.99
$\sigma^2_{d_2}$	11.1	3.74	0.017	5.95	10.4	20.3
$\sigma^2_{d_\mathrm{T}}$	53.5	71.6	1.33	8.56	34.3	213
$\sigma^2_{d_\mathrm{D}}$	0.266	0.0769	0.000349	0.157	0.253	0.453

Another indication that something may be amiss is MCMC nonconvergence when we use the two-step estimation method instead of the three-step. In Section 5.2.3 we showed that the two- and three-step estimation methods are equivalent with regards to the discrimination-corrected source isotope ratio means, δ'_s, $s = 1, \ldots, S$. This two-step estimation method illustrates the point made in Section 5.2.3 that in some situations the sum converges, $\underset{\sim}{\delta}'_s = \underset{\sim}{\delta}_s + (\underset{\sim}{\delta}_{\mathrm{T}s} - \underset{\sim}{\delta}_{\mathrm{D}s})$, but the individual components $(\underset{\sim}{\delta}_s, \underset{\sim}{\delta}_{\mathrm{D}s}, \underset{\sim}{\delta}_{\mathrm{T}s})$ may synchronously vacillate between multiple modes. In Figures 5.11 and 5.12 we see that the source and diet isotope ratio means appear to have two "basins of attraction" within the MCMC chains for two-step estimation, even though their sum is stable. The individual distributions for δ_1, δ_2, δ_T, and

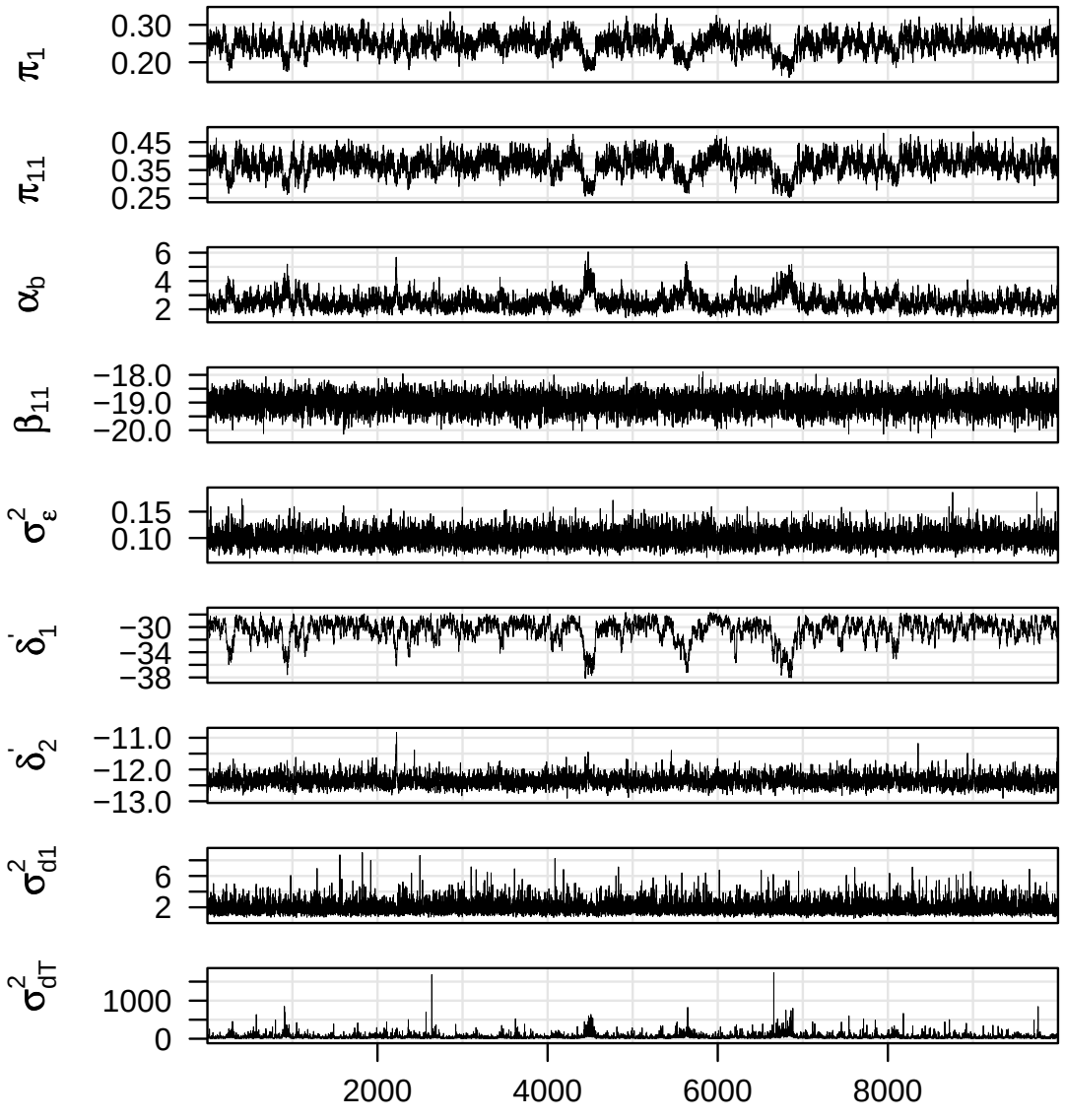

Figure 5.9: Dunlin Bayes three-step explicit representation results, trace plots for MCMC chain 1 of 5 using noninformative diet prior. Parameters shown are (1) π_1, (2) π_{11}, (3) α_b, (4) β_{11}, (5) σ^2_ε, (6) δ'_1, (7) δ'_2, (8) $\sigma^2_{d_1}$, and (9) $\sigma^2_{d_\mathrm{T}}$.

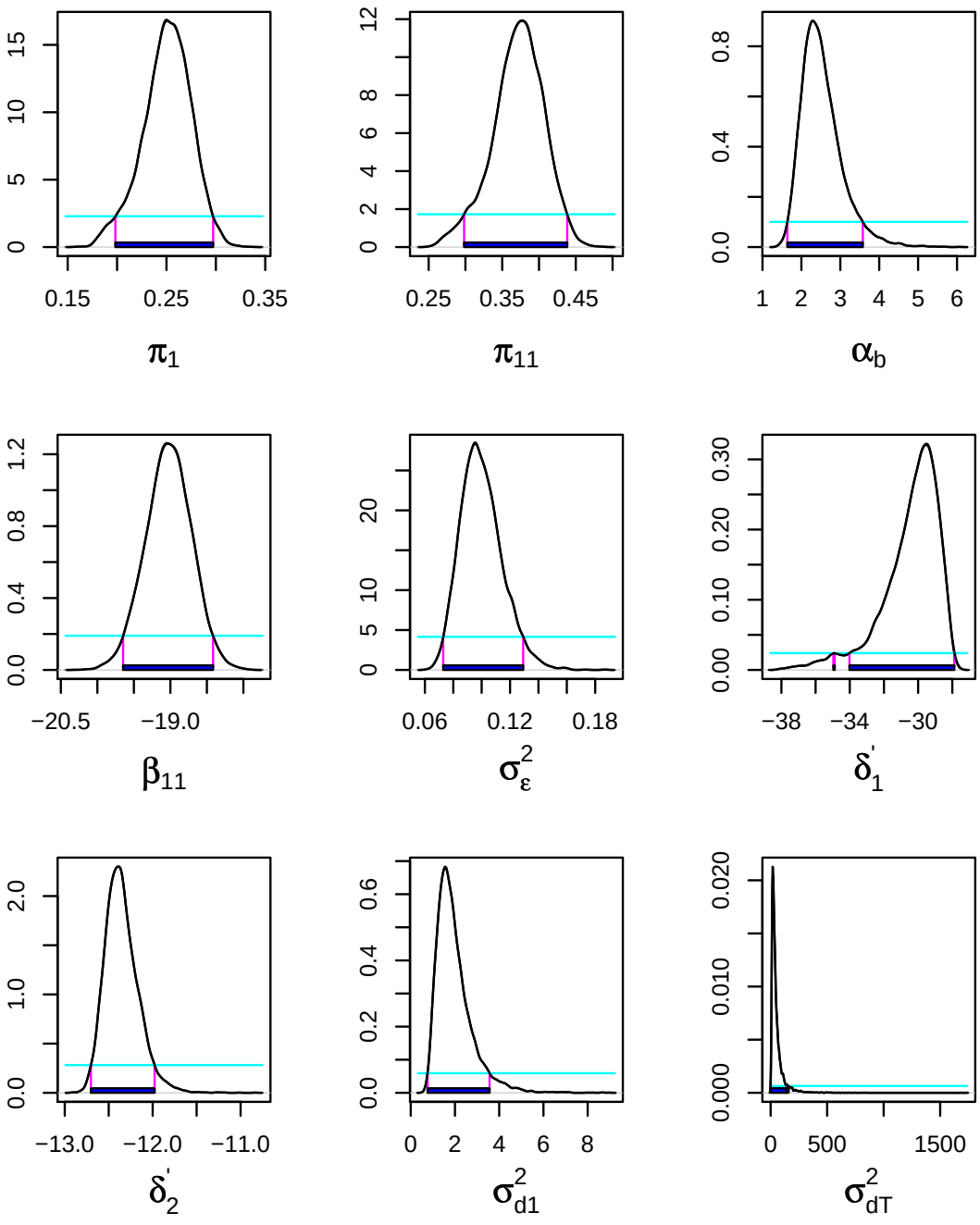

Figure 5.10: Dunlin Bayes three-step explicit representation results, posteriors for MCMC chain 1 of 5 with 95% HPD interval using noninformative diet prior.

δ_{D} each attempt to keep the posterior draws near their prior values, however the extreme mixtures and the probability constraint of their associated diets, π_j, force the endpoints, δ'_s, outward. This tug-of-war of the individual posteriors affects the individual chains. The three-step estimation method suggests that this may not be an MCMC convergence issue, but possibly an issue of model misspecification that those extreme mixtures can not be composed of the mean sources since they are beyond the endpoints. This behavior may also possibly be the result of weakly informed parameters or indicate other issues.

5.3.4.3.1 Trimmed mixtures to stay within the source endpoints To study parameter estimation when individual mixtures are beyond the scope of the source means we compare the fit of the implicit representation and explicit representation models using (1) the full mixture dataset ($J = 234$) in the previous sections with (2) a trimmed dataset ($J = 209$) including only those mixtures within a small margin of the source mean endpoints, $-23.5 \leq b_{j1} \leq -13$. This will show that it is the extreme mixtures that are most affecting the parameter estimation of the sources and diets in the explicit representation model.

Table 5.6 shows the results of three paired Bayesian model fits of the source parameters. For comparison, the method of moments estimates of the discrimination-corrected isotope ratio means (SD), $\hat{d}'_{1s}$, from Table 4.4 are -24.01 (0.373) and -12.63 (0.731) for terrestrial and marine. The first section of Table 5.6 gives the corresponding frequentist estimates of the source and discrimination values. The implicit representation model (irep) estimates the source means well for both the full data and trimmed data. The explicit representation model (erep) works well in the trimmed data case, but for the full data the terrestrial source endpoint shifts by -6‰. This endpoint shift serves to accommodate the mixture isotope ratios beyond the source means. If we also assume that the measurement error is not as small as

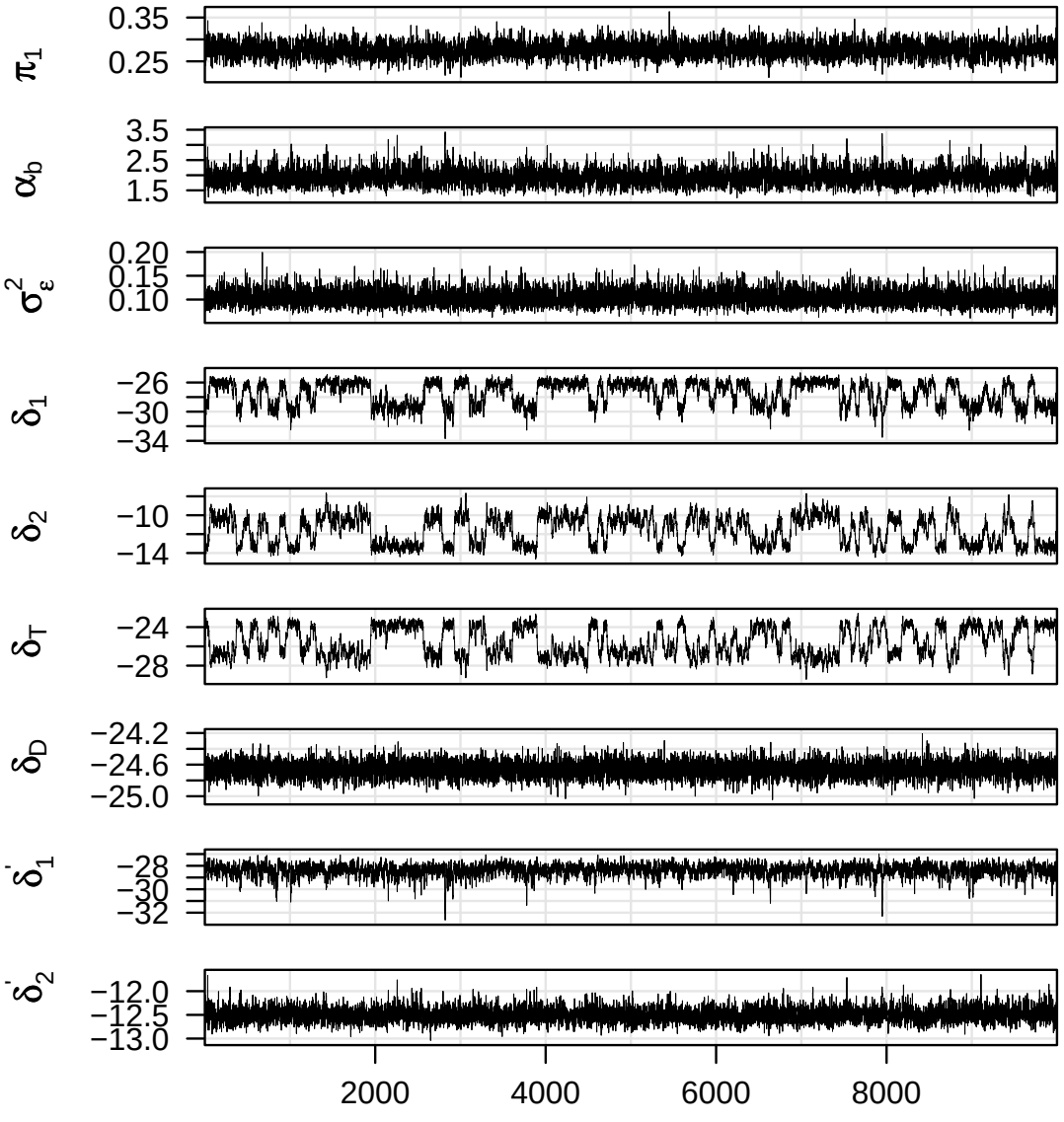

Figure 5.11: Dunlin Bayes two-step explicit representation results, trace plots for MCMC chain 1 of 5 using noninformative diet prior. Parameters shown are (1) π_1, (2) α_b, (3) σ^2_ε, (4) δ_1, (5) δ_2, (6) δ_{T}, (7) δ_{D}, (8) δ'_1, and (9) δ'_2.

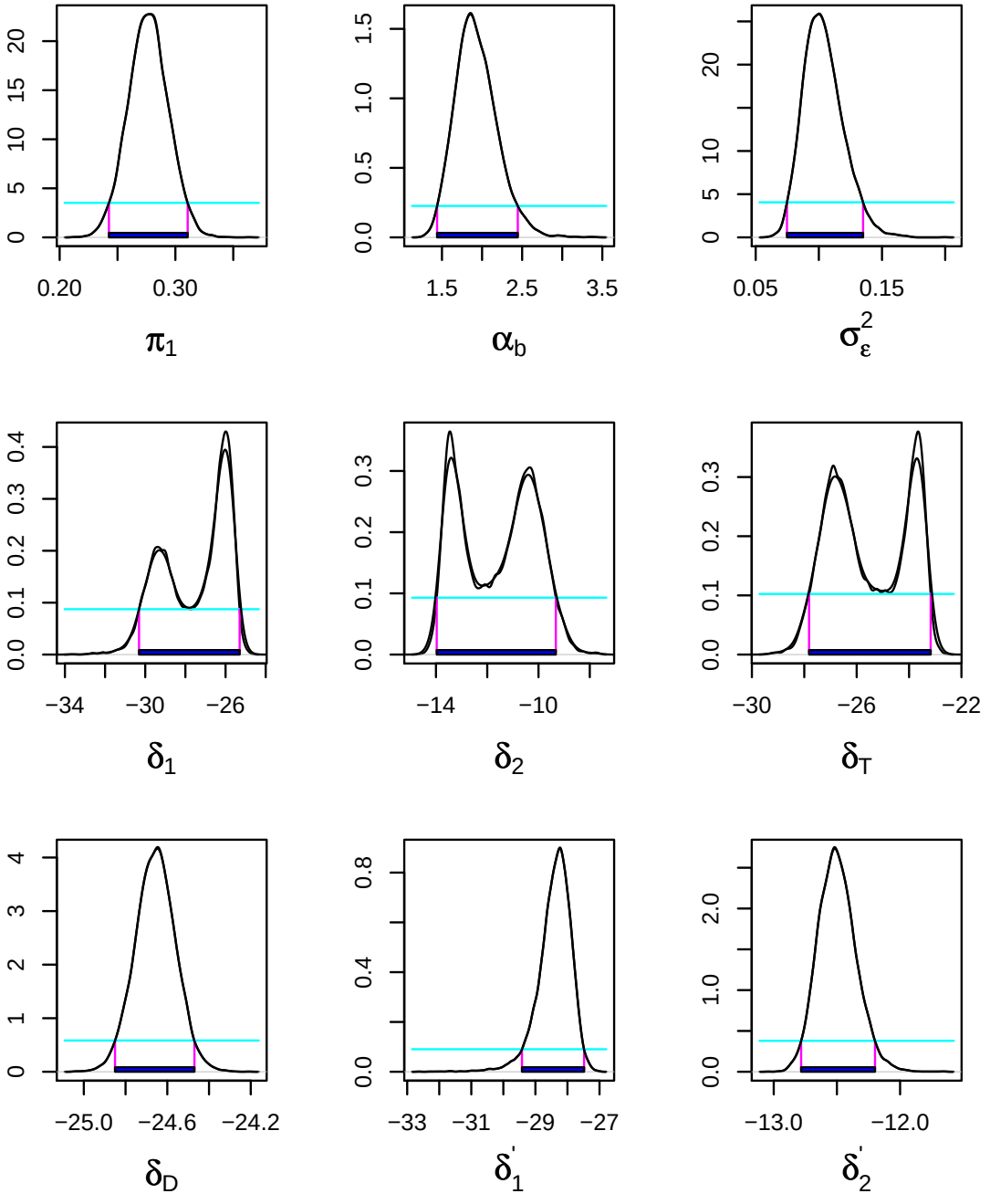

Figure 5.12: Dunlin Bayes two-step explicit representation results, posteriors for MCMC chain 1 of 5 with 95% HPD interval using noninformative diet prior.

we estimate it to be from the standards (erep big ME), by using

$$\sigma_\varepsilon^2|\sigma_{0\varepsilon}^2, \nu_{0\varepsilon} \sim \text{Scaled-inv-}\chi^2(4^2, 1),$$

then for the full data we see only a 1‰ shift to the left for the left endpoint and also about a 1.5‰ shift to the left for the right endpoint. The additional spread is accommodated for by the measurement error, σ_ε^2, now larger by over an order of magnitude, as well as the between precision of the mixture diets, α_b, decreasing. For the trimmed dataset, the measurement error stays smaller since a large error is not required to "reach" over the source endpoints to capture those extreme mixture isotope ratios. Therefore, the implicit representation is sensitive only to where the mean mixture isotope ratio is and not to where individual mixtures are, while the explicit representation is sensitive to individual mixture isotope ratios and will erroneously adjust parameter estimates in order to satisfy the model constraints. Thus, more care in the analysis is necessary when using the explicit representation model.

There are a number of reasons why mixture isotope ratios might fall outside the range of the source means, including that all necessary sources were not considered, the sample sources do not reflect the sources consumed by the sampled mixtures, source discriminations are in error, or many other possibilities. One reason highlighted by the last model in Table 5.6 is model misspecification, in particular, that individual mixtures have source preferences that can vary from the mean source isotope ratio. Consider the following replacement for (5.32) in the context of the dunlin example,

$$\beta_{j1} = \delta'_{j11}\pi_{j1} + \delta'_{j12}(1 - \pi_{j1}), \tag{5.43}$$

where $\delta'_{j11} = \delta_{j1s} + (\delta_\text{T} - \delta_\text{D})$, and there are now subject-specific source isotope ratios consisting of the mean source isotope ratio and a subject-specific offset, $\delta_{j1s} =$

Table 5.6: Dunlin Bayes illustration of a source endpoint shift under the explicit representation model when subject-specific mixture isotope ratios are outside the range of the discrimination-corrected isotope ratio means of the sources.

		Full data		Trimmed data	
Model	Param	mean	sd	mean	sd
Freq	$\bar{d}_{11}$	−25.36	1.313	same	
	$\bar{d}_{12}$	−13.98	3.248		
	$\bar{d}_{1\mathrm{T}}$	−23.28	0.323		
	$\bar{d}_{1\mathrm{D}}$	−24.63	0.380		
	$\hat{d}'_{11}$	−24.01			
	$\hat{d}'_{12}$	−12.63			
	π_1	0.367	0.046	0.367	0.045
irep	δ'_1	−24.1	0.553	−24.1	0.554
	δ'_2	−12.7	0.838	−12.7	0.841
	σ^2_β	10.4	0.964	6.96	0.690
	π_1	0.356	0.060	0.355	0.054
erep	δ'_1	**−30.5**	1.84	**−24.2**	0.406
	δ'_2	−12.4	0.186	−13.2	0.150
	σ^2_ε	**0.10**	0.015	0.098	0.014
	α_b	**2.55**	0.548	1.481	0.225
	π_1	0.249	0.026	0.332	0.023
Erep big ME	δ'_1	**−25.0**	0.934	−24.0	0.515
	δ'_2	−14.1	0.420	−13.8	0.350
	σ^2_ε	**2.54**	0.666	**1.15**	0.359
	α_b	**0.936**	0.284	1.39	0.330
	π_1	0.245	0.034	0.294	0.031

$\delta_{1s} + \xi_{j1s}$. We can rewrite this as (5.32) plus an additional term,

$$\beta_{j1} \;=\; \{\delta'_{11}\pi_{j1} + \delta'_{12}(1-\pi_{j1})\} + \{\xi_{j11}\pi_{j1} + \xi_{j12}(1-\pi_{j1})\}, \tag{5.44}$$

where the additional term will become absorbed in the error associated with $\underset{\sim}{b}_{jk}$ in (5.33). Such a model can explain the large estimate of σ^2_ε in the last model (erep big ME) in Table 5.6. Model comparison is a possible avenue for future research.

Having the condition that all the subject-specific mixture isotope ratio measure-

ments fall, in the language of Chapter 3, within the convex hull of the discrimination-corrected isotope ratio means of the sources is a considerably strict condition for the explicit representation to give accurate estimates of all model parameters. However, when individual mixtures fall outside the range of the source means, the practitioner should consider carefully whether the sampled sources and mixtures accurately reflect the system under investigation.

5.4 Underconstrained example, mink

We revisit the underconstrained mink example from Ben-David et al. (1997) to illustrate the implicit representation and explicit representation models when there are many sources and few isotopes. We reproduce the data in Table 5.7, where the mink summaries were calculated from 5 observations provided by Merav Ben-David (personal communication) after removing an outlier. We have no covariance estimates and no real prior information on any of the source isotope ratios so we use a univariate model for each isotope and Jeffreys priors to model the source parameters. From the convex hull plot in Figure 5.13 there is no clear covariance among the mixture observations so we treat the two isotopes independently. Thus, this example realistically illustrates how a researcher might use summarized source data from the literature to perform a diet analysis on a small sample of mixtures. Note that differences exist between the data in Table 5.7 and in Phillips (2001, Table 3) and Phillips and Gregg (2003, Figure 4), where there are errors in the nitrogen discrimination for the invertebrates and the carbon isotope ratio value for ducks, and the mink values are different since they are obtained from a plot (Ben-David et al., 1997, Figure 2).

For the sources we assume a univariate Jeffreys prior for isotopes independently as in (5.22) and (5.23). For example, for tidal fish carbon, we use the two-step

Table 5.7: Mink example, five observations of mink blood as a mixture of $S = 7$ sources, tidal fish, blue mussels, crabs, shrimps, rodents, amphipods, and ducks using $I = 2$ isotopes of carbon ($i = 1$) and nitrogen ($i = 2$) (Ben-David et al., 1997).

Observations		Sample	Isotope Ratios	
Coastal mink		j	carbon b_{j1}	nitrogen b_{j2}
	(b1)	1	−15.23	14.44
	(b2)	2	−14.37	14.18
	(b3)	3	−14.61	14.10
	(b4)	4	−15.76	12.50
	(b5)	5	−15.60	13.82
outlier	(b6x)	*6*	*−19.99*	*10.23*

Summaries	Sample	Isotope Ratios		Discrim	
Mixture	J	$\bar{b}_1$ (SE)	$\bar{b}_2$ (SE)		
Mink	5	−15.11 (0.27)	13.81 (0.34)		
Sources	K_s	$\bar{d}_{1s}$ (SE)	$\bar{d}_{2s}$ (SE)	Δ_{1s}	Δ_{2s}
Tidal fish ($s = 1$)	15	−15.23 (0.22)	12.68 (0.09)	1	2
Blue mussels ($s = 2$)	11	−19.51 (0.26)	7.74 (0.15)	1	3
Crabs ($s = 3$)	20	−16.28 (0.22)	9.20 (0.16)	1	3
Shrimps ($s = 4$)	6	−17.90 (0.39)	9.96 (0.21)	1	3
Rodents ($s = 5$)	18	−26.61 (0.50)	7.07 (0.28)	2	3
Amphipods ($s = 6$)	25	−19.69 (0.18)	12.00 (0.19)	1	3
Ducks ($s = 7$)	6	−23.38 (1.06)	11.92 (0.98)	2	3

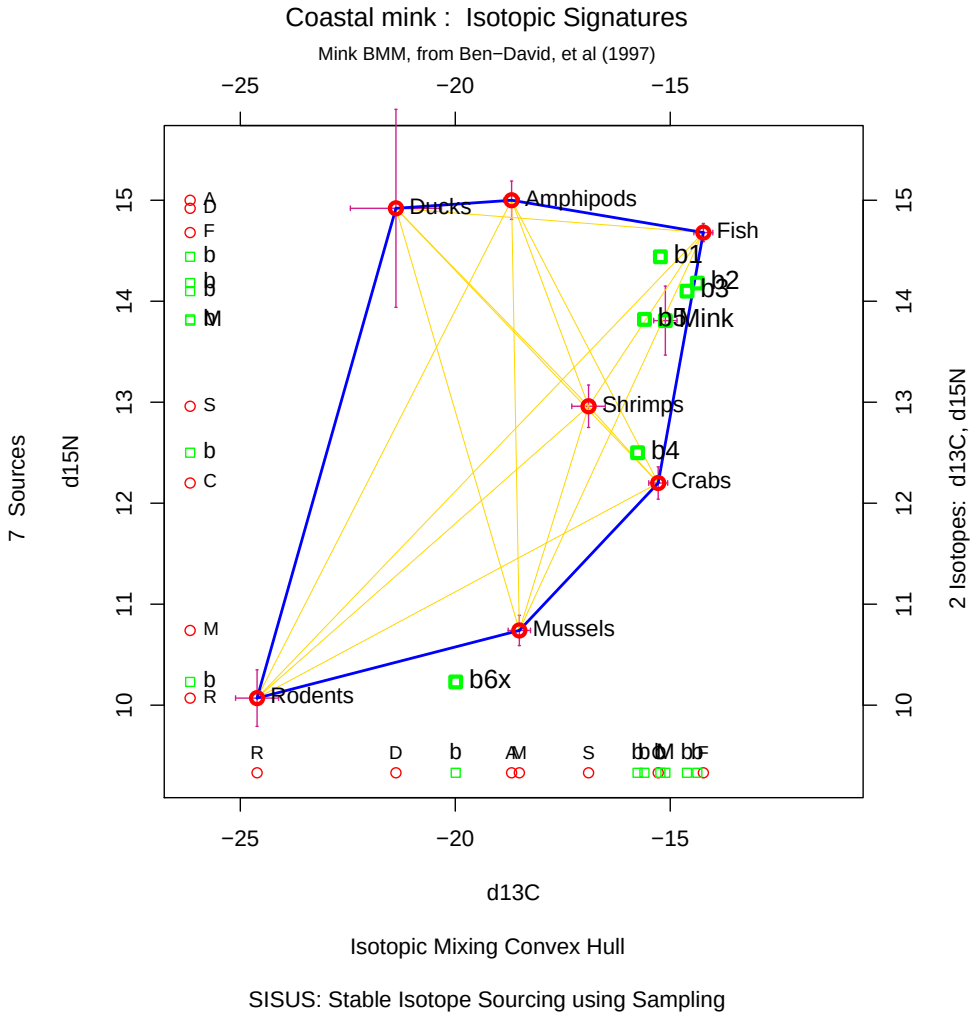

Figure 5.13: Convex hull plot of the mink example from Table 5.7, with means and standard error bars.

"updated" priors

$$\begin{aligned}\delta_{11}|d^*_{11},\sigma^2_{\delta_{11}} &\sim \text{Normal}(-15.23,\sigma^2_{\delta_{11}}/15)\\ \sigma^2_{\delta_{11}}|d^*_{11} &\sim \text{Scaled-inv-}\chi^2(15(0.22)^2,14),\end{aligned}$$

and for tidal fish nitrogen, we have

$$\begin{aligned}\delta_{21}|d^*_{21},\sigma^2_{\delta_{21}} &\sim \text{Normal}(12.68,\sigma^2_{\delta_{21}}/15)\\ \sigma^2_{\delta_{21}}|d^*_{21} &\sim \text{Scaled-inv-}\chi^2(15(0.09)^2,14).\end{aligned}$$

Note that it is possible to treat the covariance of carbon and nitrogen as missing values that are estimated in the Bayesian model, but we are not doing that. For simplicity, the discrimination values, $\underset{\sim}{\Delta}_s$, are assumed fixed and known, $s = 1,\ldots,S$. The mean diet proportion vector prior is uniform, with $\alpha_{0\pi} = 1$ and $\underset{\sim}{\pi}_0 = \underset{\sim}{1}/7$,

$$\underset{\sim}{\pi}|\underset{\sim}{\pi}_0,\alpha_{0\pi},S \sim \text{Dirichlet}(\underset{\sim}{1}).$$

5.4.1 Implicit representation

Because we have only five mink observations and there is not a strong indication of carbon and nitrogen correlation in Figure 5.13 we assume a simpler model for $\underset{\sim}{b}_j$ where the covariance is zero and the variance prior is noninformative. The sampling distribution for the mink carbon and nitrogen isotope ratios are

$$\begin{aligned}b_{ji}|\underset{\sim}{\pi},\sigma^2_{b_i},\delta^*_{\text{S}} &\overset{\text{ind}}{\sim} \text{Normal}(\beta_i,\sigma^2_{b_i}), \quad j = 1,\ldots,5\\ \sigma^2_{b_i}|\sigma^2_{0b_i},\nu_{0b_i} &\sim \text{Scaled-inv-}\chi^2(1,0.01).\end{aligned}$$

5.4.2 Results

We ran 5 chains for 110000 iterations prethinned every 10. Based on the convergence in traceplots we use iterates 10001–110000 further thinning every 10 to reduce the sample size to 50000. We checked that the Monte Carlo error for each parameter of interest was less than about 5% of the sample standard deviation (p. 55, Spiegelhalter et al., 2003). Figure 5.14 shows the MCMC chains have converged and are mixing well for a selection of parameters.

Table 5.8 provides posterior summaries for all the parameters and Figure 5.15 shows selected posterior distributions with highest posterior density (HPD) interval, and Figure 5.16 shows the same for $\underset{\sim}{\pi}$. Note that the posterior summaries include an equal-tailed interval which is slightly different from the HPD. First, note that the posterior estimates for the mixture and sources isotope ratio parameters are similar to the method of moment estimates in Table 5.7. The posterior means for $\underset{\sim}{\pi}$ are different from what is given in the no variation case in Table 3.2, and the posterior standard deviations are uniformly larger than in the no variation case. The scatterplot and marginal histograms in Figure 5.17 show more variability in the posterior distribution of $\underset{\sim}{\pi}$ than in the no variation case in Figure 3.4.

The diet estimates indicate that fish and crab contribute roughly 70% to an average mink's diet, which is not surprising since the mean isotope ratio for the mink is near the boundary of the convex hull with fish and crab as vertices. Furthermore, fish is estimated at just over half the diet.

5.4.3 Explicit representation

Let the isotope ratio observation from mixture j, $\underset{\sim}{b}_j$, depend on its subject-specific mean isotope ratio, $\underset{\sim}{\beta}_j$, with

$$b_{ji}|\underset{\sim}{\pi}_j, \sigma^2_{\varepsilon_i}, \delta^*_{\mathrm{S}} \overset{\text{ind}}{\sim} \text{Normal}(\beta_{ji}, \sigma^2_{\varepsilon_i}), \quad j = 1, \ldots, 5.$$

We assume measurement error for carbon and nitrogen are independent and we base our prior for measurement error variance on what we observed for the dunlin example in (5.42),

$$\sigma^2_{\varepsilon_i}|\sigma^2_{0\varepsilon_i}, \nu_{0\varepsilon_i} \sim \text{Scaled-inv-}\chi^2(0.3^2, 25).$$

Let the subject-specific diets, $\underset{\sim}{\pi}_j$, follow a Dirichlet distribution with a uniform distribution on the between-mixture precision parameter,

$$\begin{aligned} \underset{\sim}{\pi}_j|\underset{\sim}{\pi}, \alpha_b, S &\overset{\text{ind}}{\sim} \text{Dirichlet}(\alpha_b S\underset{\sim}{\pi}), \quad j = 1, \ldots, J \\ \alpha_b|\alpha_{01b}, \alpha_{02b} &\sim \text{Uniform}(0, 40). \end{aligned}$$

5.4.4 Results

We ran 5 chains for 17000 iterations prethinned every 10. Based on the convergence in traceplots we use iterates 2001–17000 without further thinning for a sample size of 75000. We checked that the Monte Carlo error for each parameter of interest was less than about 5% of the sample standard deviation (p. 55, Spiegelhalter et al., 2003). Figure 5.18 shows the MCMC chains have converged and are mixing well for a selection of parameters.

Table 5.9 provides posterior summaries for the mixture-related parameters and Figure 5.19 shows selected posterior distributions with HPD interval, and Figure 5.20

shows the same for $\underset{\sim}{\pi}$. Not shown are the posterior estimates for the source isotope ratio parameters, but they are similar to the method of moment estimates in Table 5.7. The posterior means for $\underset{\sim}{\pi}$ are more concentrated than in the implicit representation Table 5.8. The scatterplot and marginal histograms of $\underset{\sim}{\pi}$ in Figure 5.21 show this concentration about a single mode, which was not as clear in the implicit representation Figure 5.17.

To compare the posterior summaries for $\underset{\sim}{\pi}$ between the implicit representation and explicit representation, Figure 5.22 plots the mean and central 95% interval for the population mean $\underset{\sim}{\pi}$, as well as the mixture-specific $\underset{\sim}{\pi}_j$ from the explicit representation model. The explicit representation intervals for $\underset{\sim}{\pi}$ are uniformly less variable than the implicit representation. Many of the intervals for the elements of $\underset{\sim}{\pi}_j$ are also less variable than the implicit representation estimate for elements of $\underset{\sim}{\pi}$. This example illustrates that when the measurement error variance is small and the individual mixtures have isotope ratios within the scope of the sources, then the explicit representation model can provide more efficient estimation of the diet than the implicit representation model.

Finally, Figure 5.23 plots the posterior densities for the population mean $\underset{\sim}{\pi}$ from the implicit representation and explicit representation models, as well as the mixture-specific $\underset{\sim}{\pi}_j$ from the explicit representation model. The explicit representation densities for $\underset{\sim}{\pi}$ are uniformly less variable than the implicit representation.

In Chapter 3 we showed an improved algorithm to provide solutions to the underconstrained case without accounting for uncertainty. This example illustrates how our Bayesian models can provide diet estimates while also accounting for uncertainty. Also, the explicit representation model provided more efficient estimates than the implicit representation.

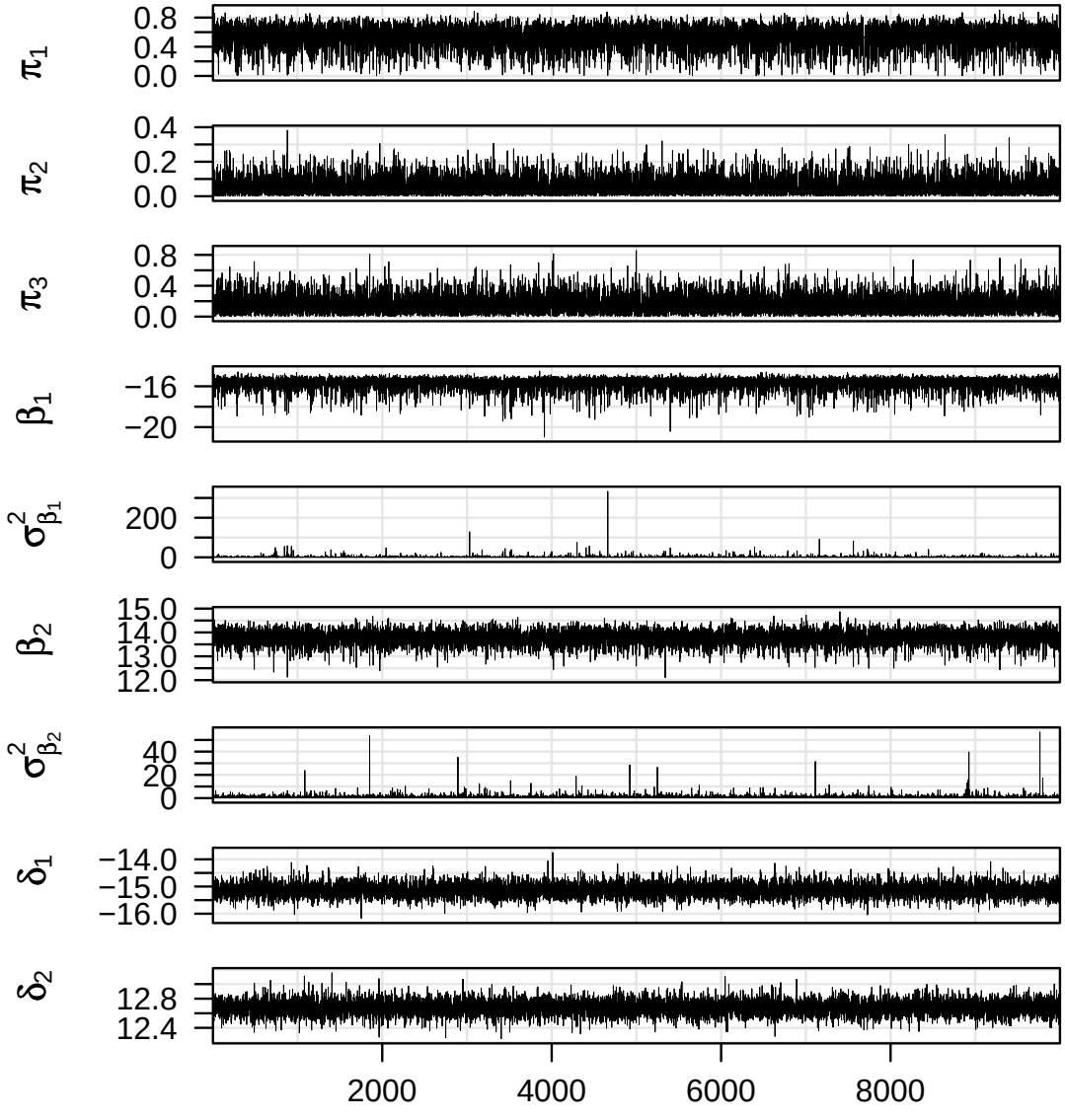

Figure 5.14: Mink Bayes implicit representation results, trace plots for MCMC chain 1 of 5. Parameters shown are (1) β_1, (2) $\sigma^2_{\beta_1}$, (3) β_2, (4) $\sigma^2_{\beta_2}$, (5) δ_{11}, and (6) δ_{21}.

Table 5.8: Mink Bayes implicit representation posterior estimates.

Parameter	mean	sd	MC error	val2.5pc	median	val97.5pc
π_1	0.534	0.156	0.000858	0.145	0.558	0.773
π_2	0.0591	0.0514	0.000237	0.00176	0.0453	0.19
π_3	0.161	0.12	0.000561	0.00669	0.138	0.447
π_4	0.104	0.0956	0.000464	0.00294	0.076	0.354
π_5	0.0277	0.0267	0.000125	0.000778	0.02	0.0978
π_6	0.0664	0.0679	0.000337	0.00177	0.0462	0.254
π_7	0.0483	0.0591	0.000412	0.00112	0.0307	0.21
β_1	−15.7	0.598	0.00284	−17.4	−15.6	−15
β_2	13.8	0.301	0.00132	13.1	13.8	14.3
$\sigma^2_{\beta_1}$	1.68	4.54	0.0203	0.156	0.688	9.38
$\sigma^2_{\beta_2}$	0.937	2.07	0.00917	0.202	0.622	3.48
δ_{11}	−15.1	0.237	0.000976	−15.6	−15.1	−14.6
δ_{21}	12.7	0.0969	0.000443	12.5	12.7	12.9
δ_{12}	−19.5	0.292	0.00133	−20.1	−19.5	−18.9
δ_{22}	7.74	0.17	0.000803	7.4	7.74	8.08
δ_{13}	−16.3	0.233	0.00107	−16.7	−16.3	−15.8
δ_{23}	9.21	0.17	0.000775	8.87	9.2	9.54
δ_{14}	−17.8	0.547	0.0025	−18.8	−17.8	−16.7
δ_{24}	9.96	0.267	0.00119	9.43	9.96	10.5
δ_{15}	−26.6	0.531	0.00247	−27.6	−26.6	−25.5
δ_{25}	7.07	0.298	0.00134	6.48	7.07	7.66
δ_{16}	−19.7	0.189	0.000835	−20.1	−19.7	−19.3
δ_{26}	12	0.197	0.000913	11.6	12	12.4
δ_{17}	−23	1.78	0.0125	−25.7	−23.2	−19
δ_{27}	11.9	1.25	0.00581	9.44	11.9	14.4
$\sigma^2_{\delta_{11}}$	0.856	0.388	0.00174	0.391	0.77	1.84
$\sigma^2_{\delta_{21}}$	0.141	0.0638	0.000285	0.0651	0.127	0.303
$\sigma^2_{\delta_{12}}$	0.929	0.537	0.00243	0.362	0.796	2.29
$\sigma^2_{\delta_{22}}$	0.31	0.178	0.000811	0.121	0.266	0.758
$\sigma^2_{\delta_{13}}$	1.08	0.393	0.00175	0.559	1	2.05
$\sigma^2_{\delta_{23}}$	0.571	0.208	0.000941	0.298	0.53	1.09
$\sigma^2_{\delta_{14}}$	1.59	2.37	0.0131	0.358	1.06	5.89
$\sigma^2_{\delta_{24}}$	0.442	0.592	0.00264	0.102	0.304	1.6
$\sigma^2_{\delta_{15}}$	5.1	2	0.00872	2.52	4.68	10.1
$\sigma^2_{\delta_{25}}$	1.59	0.62	0.00289	0.797	1.46	3.14
$\sigma^2_{\delta_{16}}$	0.883	0.28	0.00122	0.493	0.833	1.57
$\sigma^2_{\delta_{26}}$	0.985	0.313	0.00141	0.546	0.928	1.75
$\sigma^2_{\delta_{17}}$	13.5	31.8	0.223	2.64	8	54.8
$\sigma^2_{\delta_{27}}$	9.54	11.2	0.0515	2.24	6.59	34.5

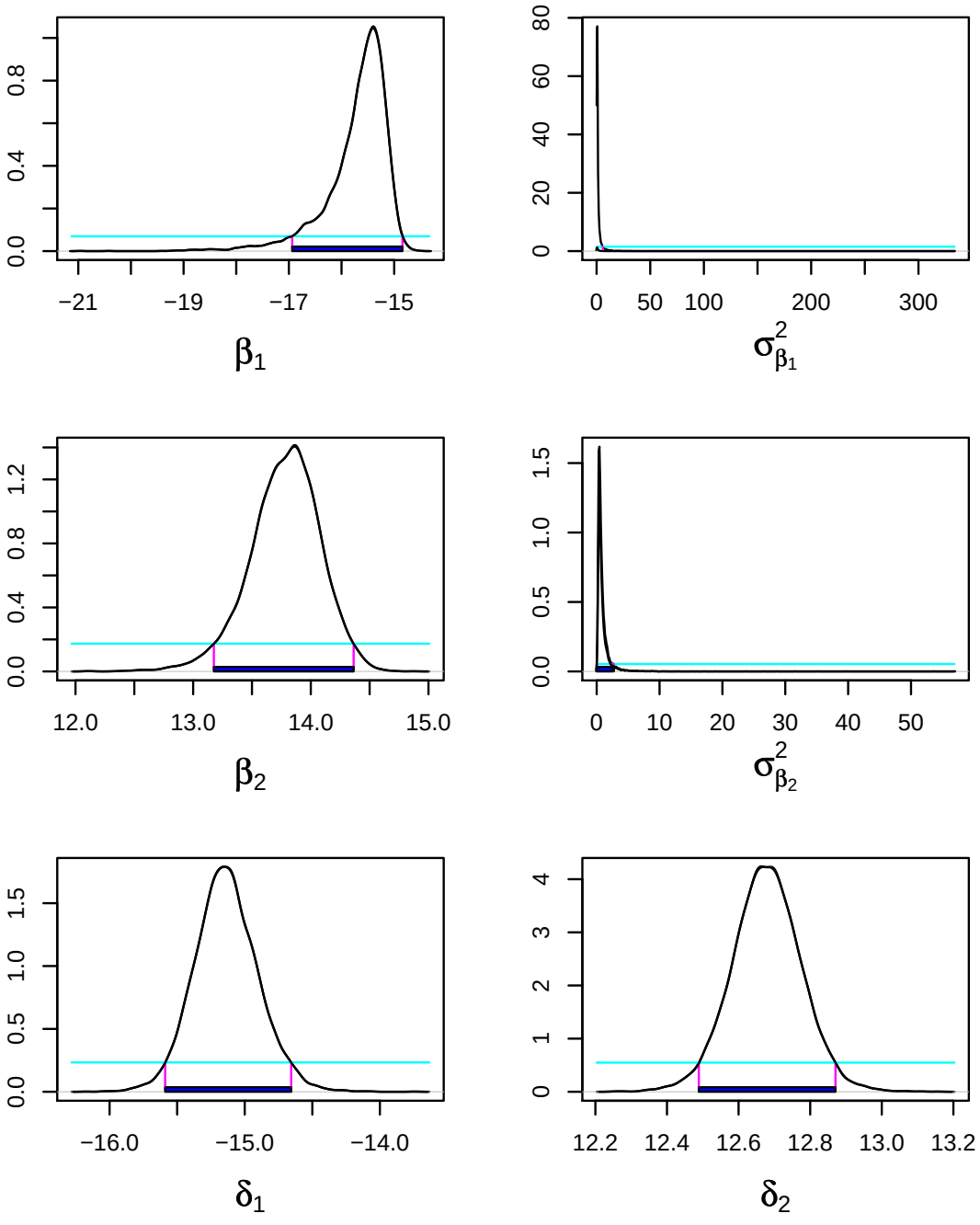

Figure 5.15: Mink Bayes implicit representation results, posteriors for MCMC chain 1 of 5 with 95% HPD interval.

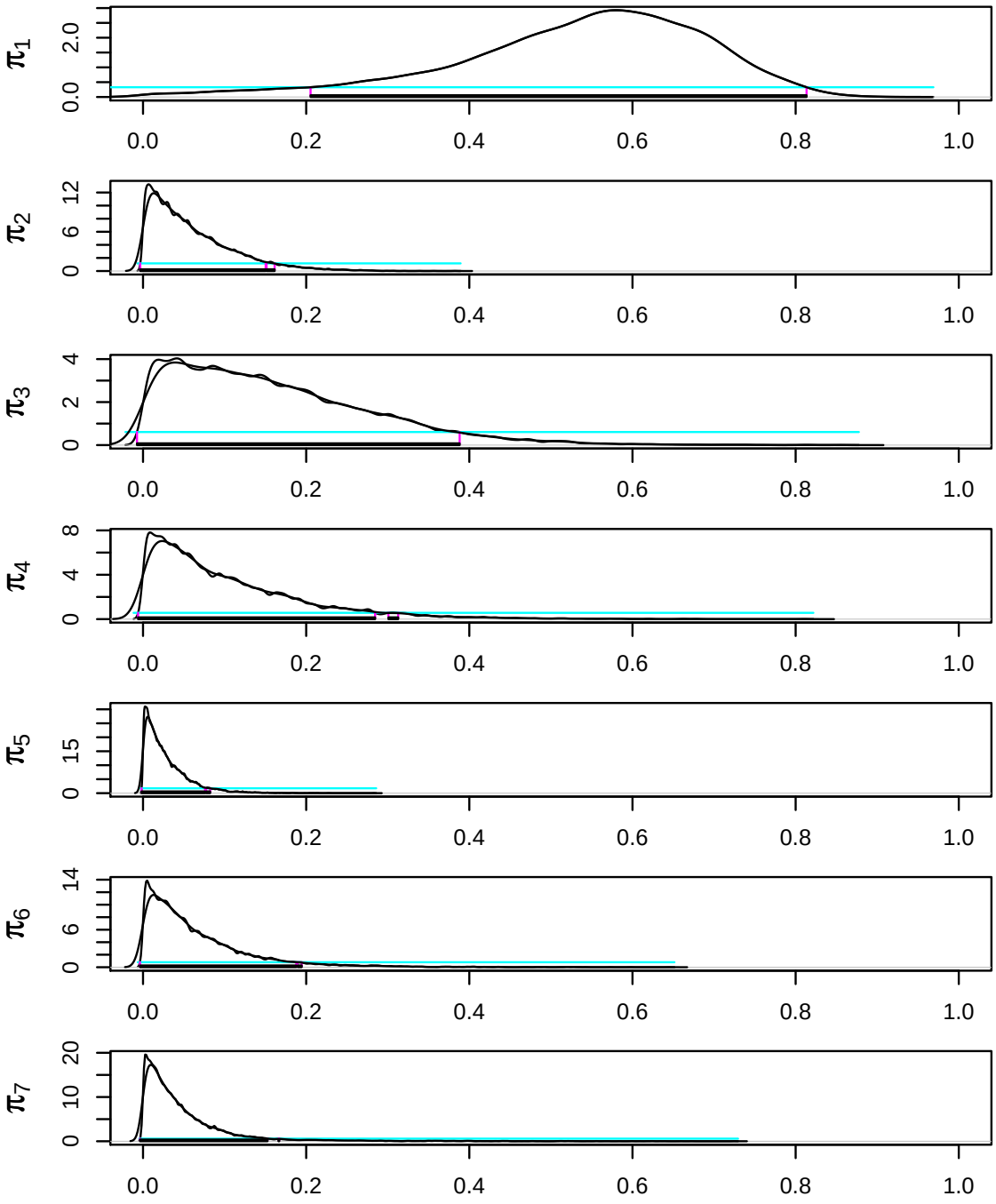

Figure 5.16: Mink Bayes implicit representation results, posteriors for $\pi_1, \ldots, \pi_7$ for MCMC chain 1 of 5 with 95% HPD interval.

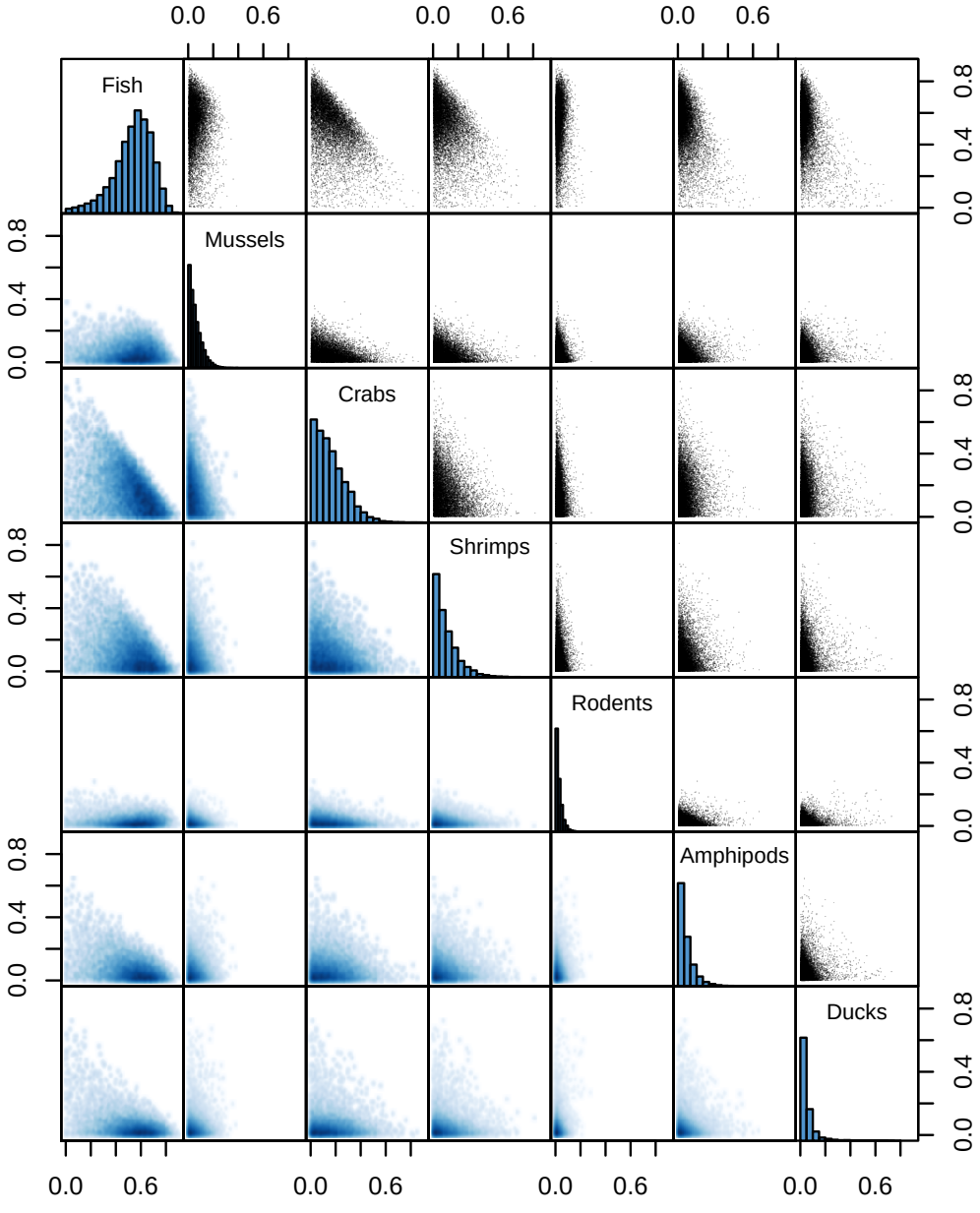

Figure 5.17: Mink Bayes implicit representation example, $\underset{\sim}{\pi}$ histograms along diagonal, scatterplot of paired source contributions on the upper diagonal, and corresponding two-dimensional density histograms on the lower diagonal.

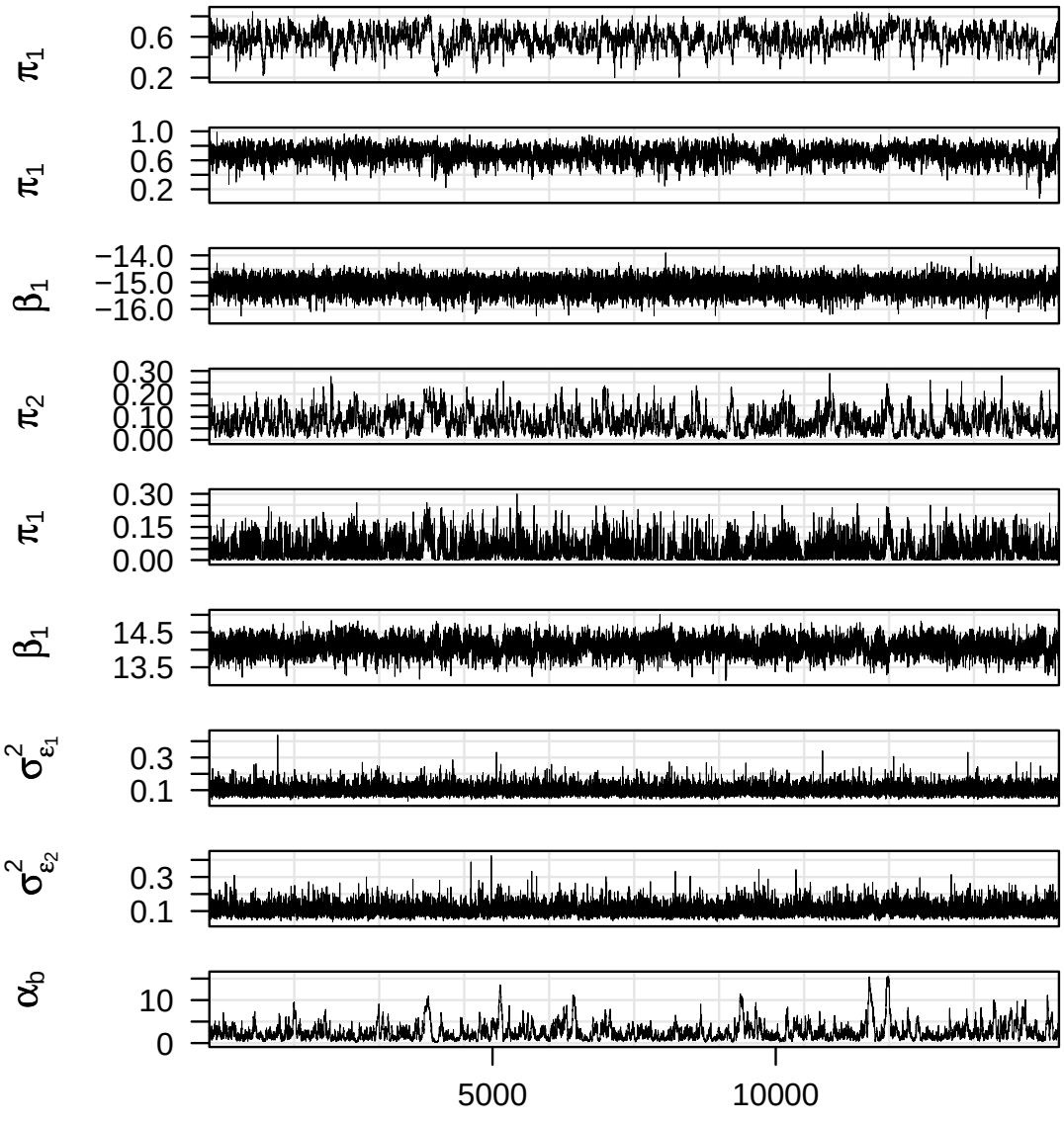

Figure 5.18: Mink Bayes explicit representation results, trace plots for MCMC chain 1 of 5. Parameters shown are (1) π_1, (2) π_{11}, (3) β_{11}, (4) π_2, (5) π_{12}, (6) β_{12}, (7) $\sigma^2_{\varepsilon_1}$, (8) $\sigma^2_{\varepsilon_2}$, and (9) α_b.

Table 5.9: Mink Bayes explicit representation posterior estimates for mixture parameters. Source parameters are similar as in implicit representation.

Parameter	mean	sd	MC error	val2.5pc	median	val97.5pc
π_1	0.572	0.107	0.00264	0.332	0.583	0.751
π_2	0.076	0.0471	0.00113	0.0112	0.0674	0.187
π_3	0.146	0.0845	0.00208	0.018	0.135	0.334
π_4	0.0852	0.0638	0.00159	0.0101	0.0684	0.25
π_5	0.0322	0.0252	0.000677	0.00448	0.0252	0.0987
π_6	0.0478	0.0347	0.000857	0.00651	0.0392	0.136
π_7	0.0409	0.0361	0.00106	0.00525	0.0317	0.129
π_{11}	0.686	0.104	0.00191	0.469	0.692	0.865
π_{12}	0.0425	0.0443	0.000783	0.0000139	0.0282	0.155
π_{21}	0.756	0.104	0.00213	0.538	0.764	0.928
π_{22}	0.0341	0.0366	0.000686	0.00000176	0.0218	0.129
π_{31}	0.723	0.105	0.00201	0.502	0.73	0.902
π_{32}	0.0407	0.0423	0.000759	0.00000167	0.0272	0.148
π_{41}	0.406	0.146	0.00327	0.0925	0.419	0.663
π_{42}	0.16	0.122	0.00235	0.0000394	0.147	0.415
π_{51}	0.586	0.115	0.00207	0.34	0.595	0.787
π_{52}	0.0719	0.0684	0.00118	0.0000109	0.0535	0.237
α_b	2.86	3.02	0.0979	0.42	1.89	11.8
β_{11}	−15.2	0.277	0.00167	−15.7	−15.2	−14.6
β_{12}	14.1	0.24	0.00337	13.6	14.1	14.6
β_{21}	−14.7	0.248	0.00318	−15.2	−14.7	−14.3
β_{22}	14.1	0.219	0.00295	13.7	14.2	14.5
β_{31}	−14.8	0.251	0.00266	−15.4	−14.8	−14.4
β_{32}	14.1	0.223	0.0026	13.6	14.1	14.5
β_{41}	−15.8	0.292	0.00291	−16.4	−15.8	−15.3
β_{42}	13.1	0.342	0.0067	12.4	13.1	13.8
β_{51}	−15.5	0.282	0.00193	−16.1	−15.5	−15
β_{52}	13.8	0.243	0.00205	13.3	13.8	14.3
$\sigma^2_{\varepsilon_1}$	0.101	0.0311	0.000169	0.0566	0.0952	0.177
$\sigma^2_{\varepsilon_2}$	0.108	0.0353	0.000377	0.0591	0.102	0.194

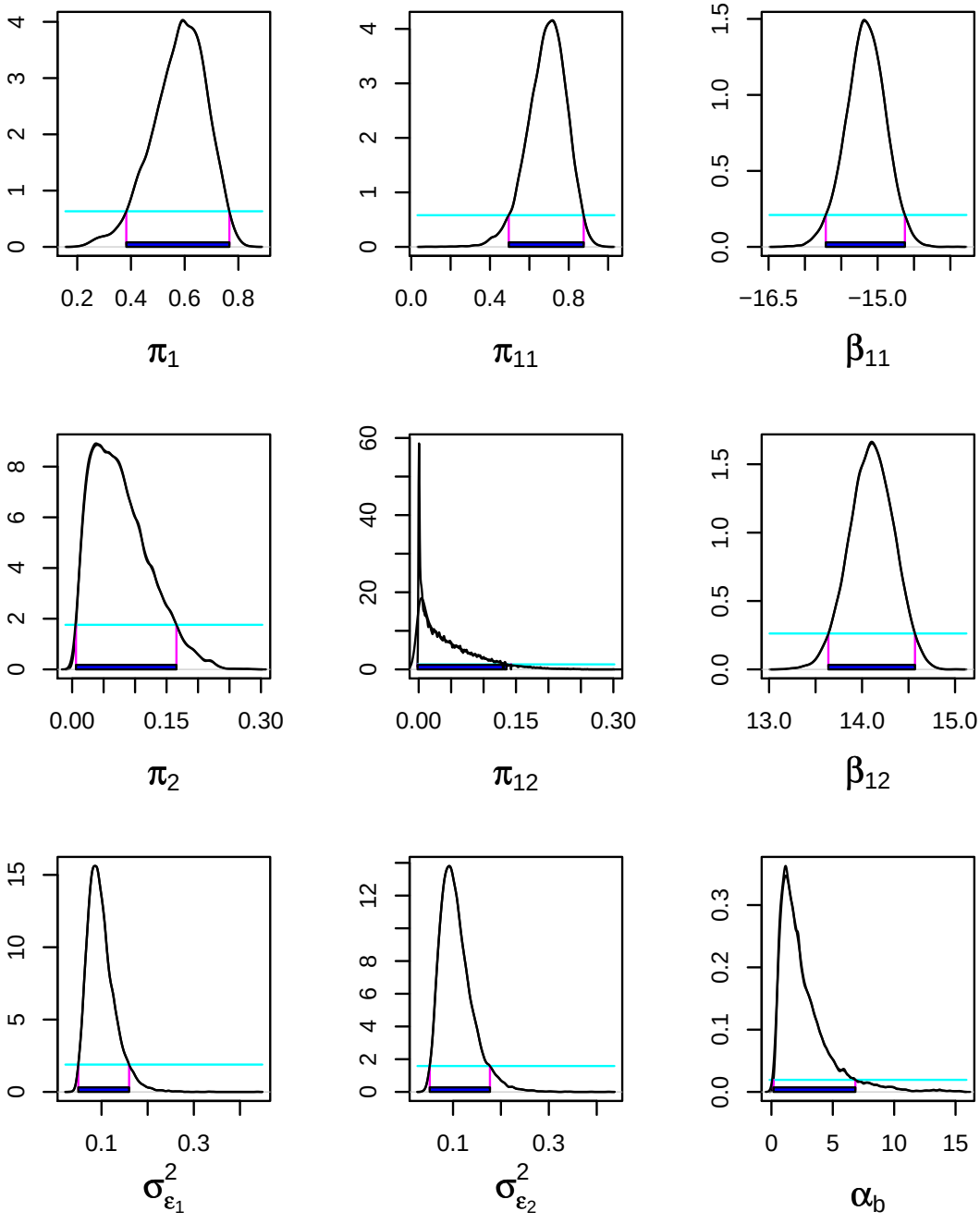

Figure 5.19: Mink Bayes explicit representation results, posteriors for MCMC chain 1 of 5 with 95% HPD interval.

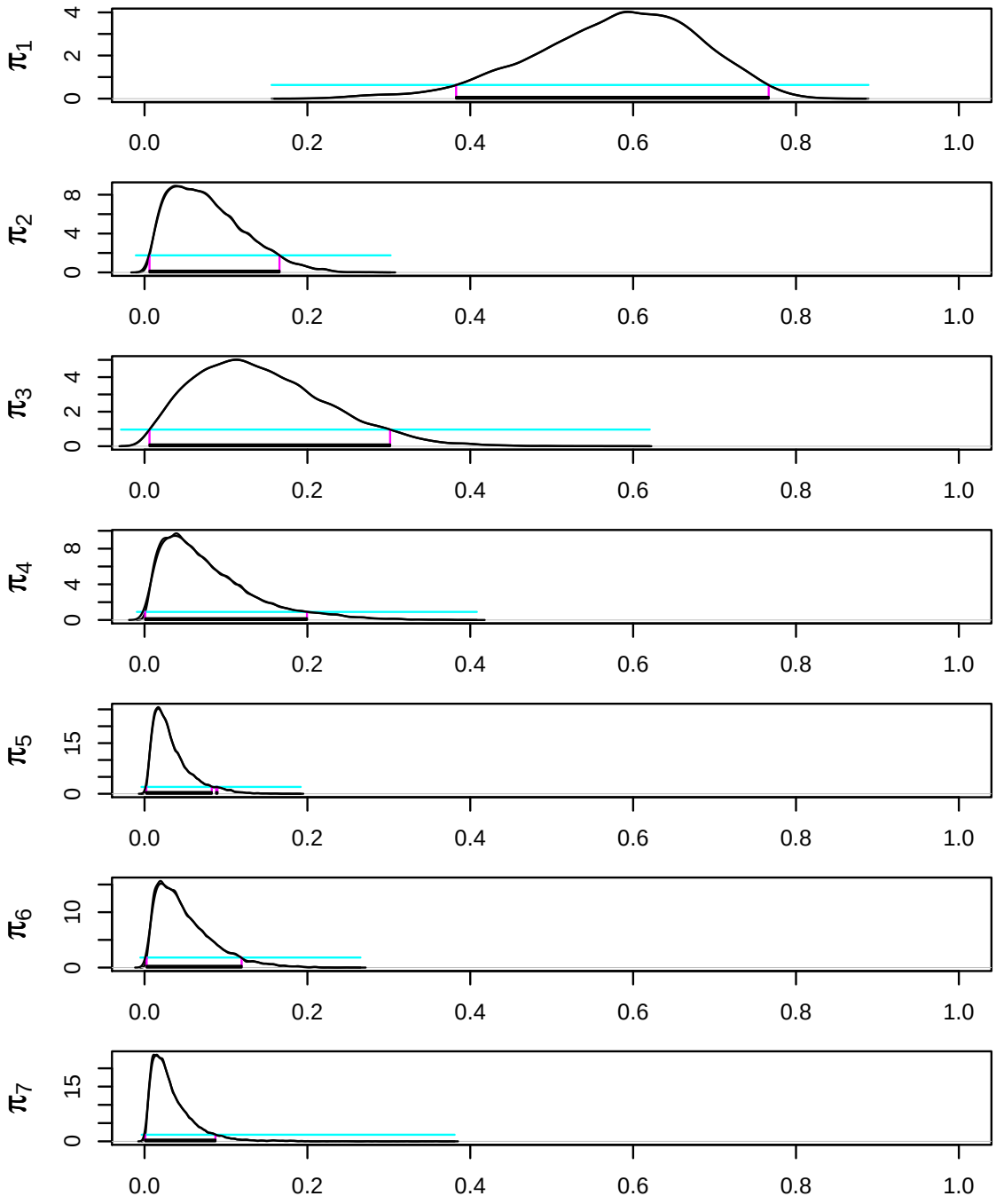

Figure 5.20: Mink Bayes explicit representation results, posteriors for $\pi_1, \ldots, \pi_7$ for MCMC chain 1 of 5 with 95% HPD interval.

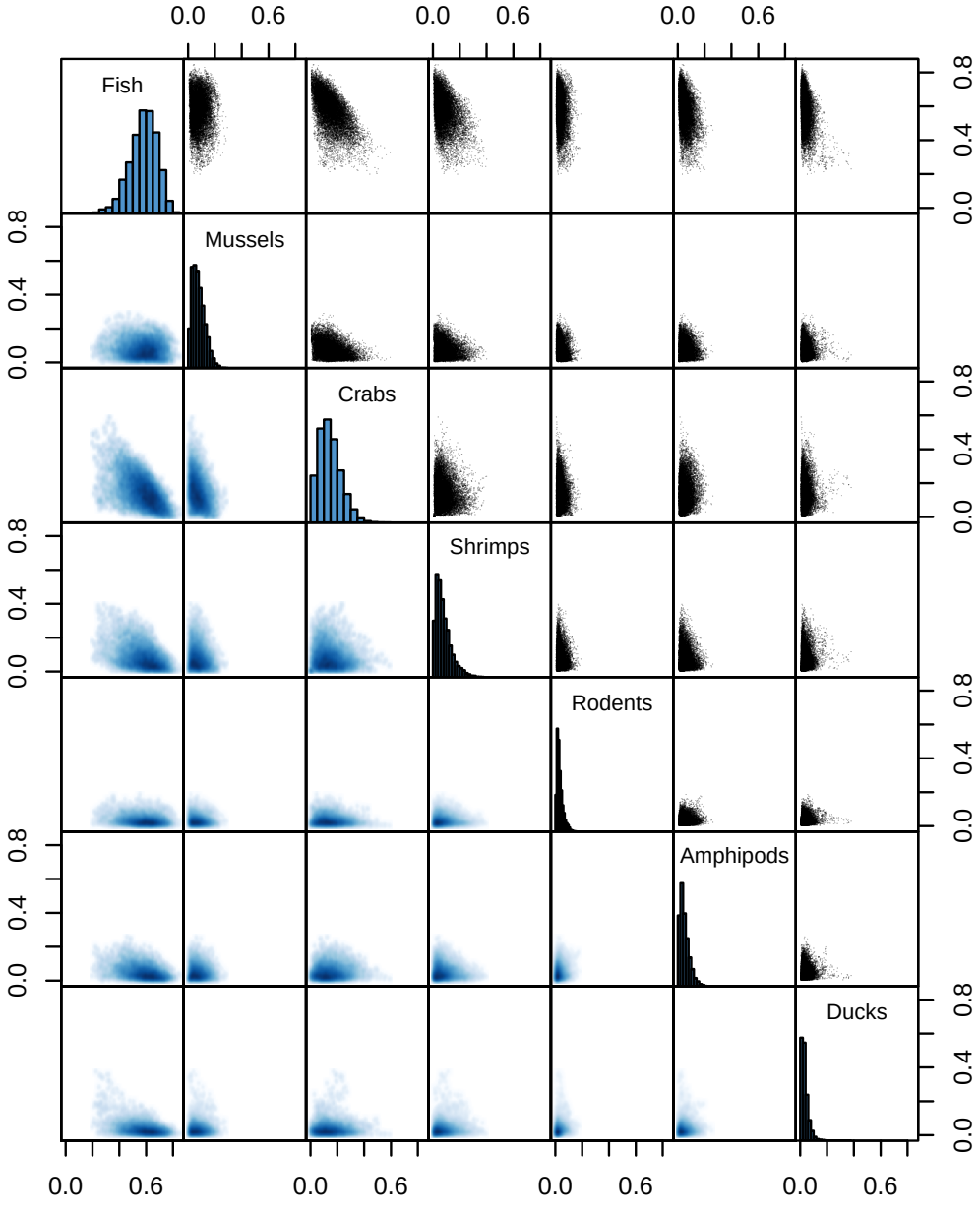

Figure 5.21: Mink Bayes explicit representation example, $\underset{\sim}{\pi}$ histograms along diagonal, scatterplot of paired source contributions on the upper diagonal, and corresponding two-dimensional density histograms on the lower diagonal.

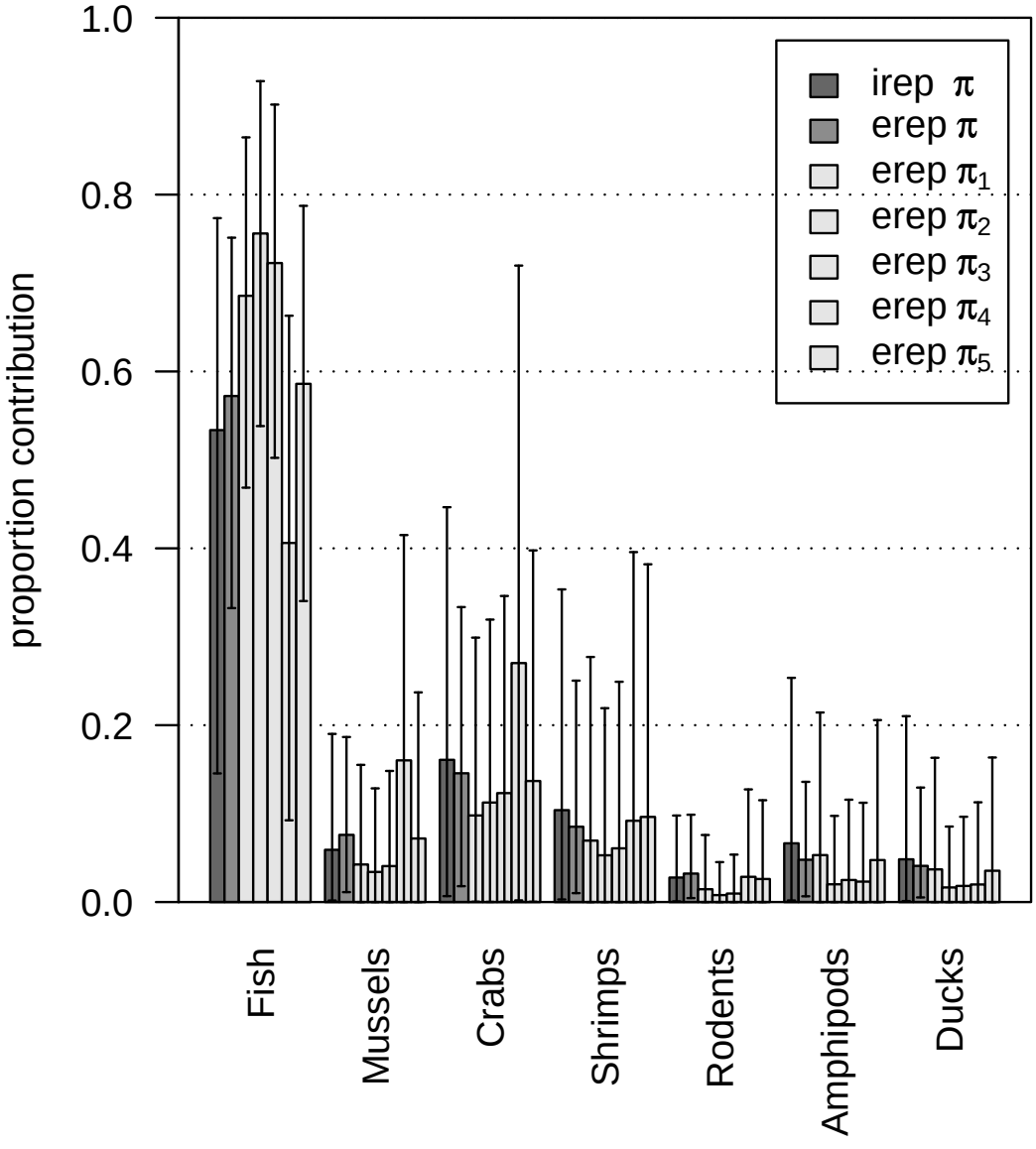

Figure 5.22: Mink Bayes implicit representation versus explicit representation results, posteriors intervals for $\pi_1, \ldots, \pi_7$ and $\pi_{j1}, \ldots, \pi_{j7}$, $j = 1, \ldots, 5$.

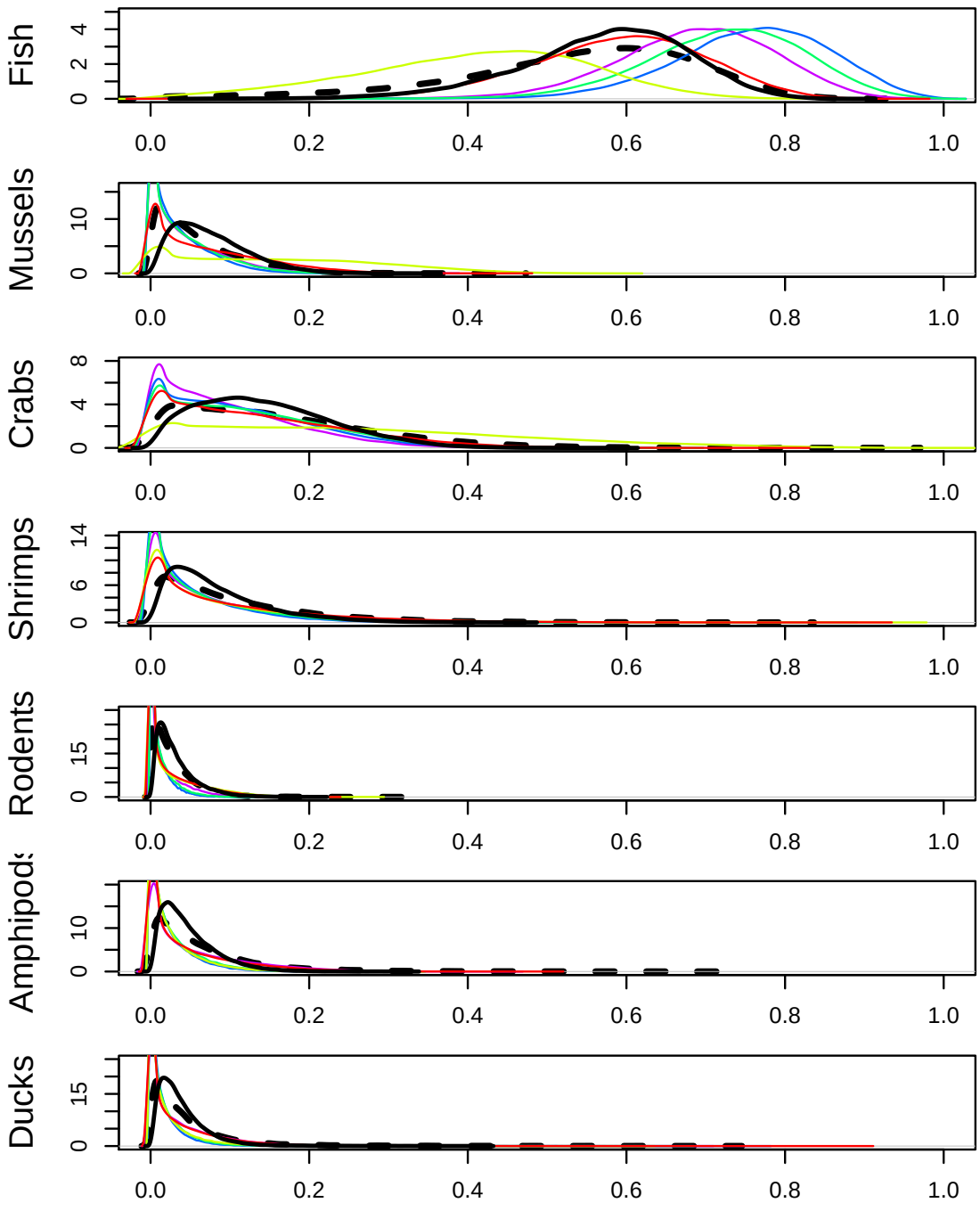

Figure 5.23: Mink Bayes implicit representation (dashed) versus explicit representation (solid) results, posteriors densities for $\pi_1, \ldots, \pi_7$ (black) and $\pi_{j1}, \ldots, \pi_{j7}$, $j = 1, \ldots, 5$ (colors).

Chapter 6

Extensions, future work, and conclusion

6.1 Extensions

6.1.1 Modeling an over-constrained population mean proportion

In ecology, carbon and nitrogen are often measured together. Also, two-source mixing model applications are very common, such as estimating diet proportions of C3 and C4 photosynthetic plants. In this situation ($S = 2, I = 2$) the diet proportions can be estimated using carbon or nitrogen. Practitioners often use their judgement to choose one of the isotopes over the other, but estimation can potentially be more precise by combining the solutions.

6.1.1.1 Frequentist implicit representation, weighted least squares

When we are in the over-constrained situation of having at least as many isotopes as sources, $I + 1 > S$, there are $u = I!/((S-1)!(I-S+1)!)$ combinations of isotopes to calculate univariate estimates $T_{s1}, \ldots, T_{su}$ of the proportion for source s, π_s. These solutions are correlated because the isotope ratios are measured jointly for each sample. Some of these estimates are better in mean squared error than others. Each estimate using the implicit-function theorem from Section 4.2.2.1 has an asymptotically normal distribution and these correlated estimates are optimally combined using weighted least squares to provide a single best estimate. Let $\underset{\sim}{T} = (T_{s1}, \ldots, T_{su})^\top$ have covariance matrix Σ_T, for which there is a consistent estimate $\mathbf{S}_T$. The weighted least squares estimate of π_s is

$$\tilde{\pi}_s = \frac{\underset{\sim}{1}^\top \mathbf{S}_T^{-1} \underset{\sim}{T}}{\underset{\sim}{1}^\top \mathbf{S}_T^{-1} \underset{\sim}{1}}, \tag{6.1}$$

with estimated variance $\mathrm{Var}[\tilde{\pi}_s] = (\underset{\sim}{1}^\top \mathbf{S}_T^{-1} \underset{\sim}{1})^{-1}$.

Example 6.1.1. Dunlin carbon and nitrogen

We use the dunlin example with two-sources and two-isotopes to illustrate how to incorporate individual estimates into a combined estimate. For this example, a modification in notation uses the additional first subscript on $\underset{\sim}{\pi}_i$ and π_{is} to indicate the isotope used for estimation, rather than a subject-specific diet as previously defined. Let $\underset{\sim}{\pi}_1 = [\pi_{11}, \pi_{12}]^\top$ and $\underset{\sim}{\pi}_2 = [\pi_{21}, \pi_{22}]^\top$ be the proportion vectors for carbon and nitrogen, respectively, with the understanding that $\underset{\sim}{\pi} = \underset{\sim}{\pi}_1 = \underset{\sim}{\pi}_2$. Using carbon alone the estimate of $\underset{\sim}{\pi}_1$ is

$$\hat{\underset{\sim}{\pi}}_1 = \begin{bmatrix} \hat{\pi}_{11} \\ \hat{\pi}_{12} \end{bmatrix} = \begin{bmatrix} 0.3673 \\ 0.6327 \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{V}}_1 = 0.002246 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

and using nitrogen alone the estimate of $\underset{\sim}{\pi}_2$ is

$$\hat{\underset{\sim}{\pi}}_2 = \begin{bmatrix} \hat{\pi}_{21} \\ \hat{\pi}_{22} \end{bmatrix} = \begin{bmatrix} 0.3596 \\ 0.6404 \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{V}}_2 = 0.004220 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Notice that the estimates and associated covariances are different for the carbon and nitrogen isotopes, with nitrogen being more variable. Let the parameter of interest be the proportion for the terrestrial diet, π_1. Above, we have the estimates of π_1 using carbon and nitrogen as $\underset{\sim}{T} = (T_1, T_2)^\top = (\hat{\pi}_{11}, \hat{\pi}_{21})^\top$, and we need to derive the covariance matrix, Σ_T. Using the implicit-function theorem from Section 4.2.2.1, we can estimate the variance for the two estimates $\hat{\underset{\sim}{\pi}}_1$ and $\hat{\underset{\sim}{\pi}}_2$ simultaneously, as well as the covariance between the estimates. Let

$$\underset{\sim}{\mu}_y = [\delta_{11}, \delta_{21}, \delta_{12}, \delta_{22}, \delta_{1\mathrm{T}}, \delta_{2\mathrm{T}}, \delta_{1\mathrm{D}}, \delta_{2\mathrm{D}}, \beta_1, \beta_2],$$

and

$$\underset{\sim}{y}_k = [\bar{d}_{11}, \bar{d}_{21}, \bar{d}_{12}, \bar{d}_{22}, \bar{d}_{1\mathrm{T}}, \bar{d}_{2\mathrm{T}}, \bar{d}_{1\mathrm{D}}, \bar{d}_{2\mathrm{D}}, \bar{b}_1, \bar{b}_2].$$

Let $\mathbf{G}_{CN}$ be block diagonal, diag$[\mathbf{G}_C, \mathbf{G}_N]$, giving the required equation for (4.19)

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{G}_C & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_N \end{bmatrix} \begin{bmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{21} \\ \pi_{22} \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \text{where}$$

$$\mathbf{G}_C = \begin{bmatrix} \delta_{11} + (\delta_{1\mathrm{T}} - \delta_{1\mathrm{D}}) - \beta_1 & \delta_{12} + (\delta_{1\mathrm{T}} - \delta_{1\mathrm{D}}) - \beta_1 \\ 1 & 1 \end{bmatrix} \quad \text{and}$$

$$\mathbf{G}_N = \begin{bmatrix} \delta_{21} + (\delta_{2\mathrm{T}} - \delta_{2\mathrm{D}}) - \beta_2 & \delta_{22} + (\delta_{2\mathrm{T}} - \delta_{2\mathrm{D}}) - \beta_2 \\ 1 & 1 \end{bmatrix}.$$

Then (4.23) is

$$\mathbf{H} \;=\; \begin{bmatrix} \pi_{11} & 0 & \pi_{12} & 0 & 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \pi_{21} & 0 & \pi_{22} & 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The covariance for $\underset{\sim}{y}_k$ is

$$\hat{\mathbf{\Omega}}_k \;=\; \begin{bmatrix} \mathrm{Cov}[\bar{d}_1] & & & & & \\ & \mathrm{Cov}[\bar{d}_2] & & & & \\ & & \mathrm{Cov}[\bar{d}_{\mathrm{T}}] & & & \\ & & & \mathrm{Cov}[\bar{d}_{\mathrm{D}}] & & \\ & & & & \mathrm{Cov}[\bar{b}] \end{bmatrix}.$$

Finally, the covariance for (4.22) is

$$\mathrm{Cov}[(\hat{\pi}_{11}, \hat{\pi}_{12}, \hat{\pi}_{21}, \hat{\pi}_{22})^\top] \;=\; \hat{\mathbf{V}} \;=\; \hat{\mathbf{G}}_{CN}^{-1} \hat{\mathbf{H}} \hat{\mathbf{\Omega}}_k (\hat{\mathbf{G}}_{CN}^{-1} \hat{\mathbf{H}})^\top,$$

which is estimated by substituting the natural estimates in the matrices above.

Now we apply weighted least squares to combine the correlated carbon and nitrogen estimates for an improved estimate for π_1. We use the Kronecker product to represent the covariance matrix,

$$\widehat{\mathrm{Cov}}[(\hat{\pi}_{11}, \hat{\pi}_{12}, \hat{\pi}_{21}, \hat{\pi}_{22})^\top] \;=\; \hat{\mathbf{V}} \;=\; \begin{bmatrix} 0.002246 & 0.000217 \\ 0.000217 & 0.004220 \end{bmatrix} \otimes \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

We can now form the covariance for the estimates for π_1 by extracting the four elements corresponding to the rows and columns associated with $\hat{\pi}_{11}$ and $\hat{\pi}_{21}$ in $\hat{\mathbf{V}}$, in this case this is the intersection of the first and third rows and columns,

$$\widehat{\mathrm{Cov}}[(\hat{\pi}_{11}, \hat{\pi}_{21})^\top] \;=\; \mathbf{S}_T \;=\; \begin{bmatrix} 0.002246 & 0.000217 \\ 0.000217 & 0.004220 \end{bmatrix}.$$

By (6.1), we have the estimate

$$\tilde{\pi}_1 \;=\; \frac{[1,1]\mathbf{S}_T^{-1}[0.3673, 0.3596]^\top}{[1,1]\mathbf{S}_T^{-1}[1,1]^\top} \;=\; 0.3647$$

with estimated variance

$$\mathrm{Var}[\tilde{\pi}_1] \;=\; ([1,1]\mathbf{S}_T^{-1}[1,1]^\top)^{-1} \;=\; 0.001564.$$

The estimate $\tilde{\pi}_1$ is a weighted average of the individual carbon and nitrogen estimates and has a smaller estimated variance than either of the individual estimates. Table 6.1 compares the three estimates. This can be applied in an analogous way to larger problems.

Table 6.1: Dunlin frequentist combined estimate has a smaller variance than the individual carbon and nitrogen estimates.

Isotope	Estimate		Var	SD
Carbon	$\hat{\pi}_{11}$	0.3673	0.002246	0.04739
Nitrogen	$\hat{\pi}_{21}$	0.3596	0.004220	0.06496
Combined	$\tilde{\pi}_1$	0.3647	0.001564	0.03954

△

6.1.1.2 Bayesian implicit representation overconstrained situatione

We estimate the population mean diet using the dunlin carbon and nitrogen data introduced in Section 4.1.5 with noninformative priors on all parameters and modeling carbon and nitrogen isotope ratio means for sources and the mixture as bivariate normal. Using the Bayesian paradigm the model and method is the same regardless of the relationship of S and I. Thus, treatment of the overconstrained situation ($I+1 > S$) is no different than the perfectly ($I+1 = S$) or underconstained ($I \leq S$)

cases. Furthermore, this works in the explicit representation model because of measurement error as discussed in Section 2.2.1.

Example 6.1.2. Dunlin carbon and nitrogen

The plot in Figure 6.1 and Table 6.2 shows the posterior distribution and summaries of π_1 using carbon alone, nitrogen alone, and carbon and nitrogen together. Note that the posterior summary using carbon alone is different from in Chapter 5 since different priors were used. As in the frequentist case in example 6.1.1, inference of π_1 is more precise using carbon than using nitrogen, and using both together provides even greater precision. △

Table 6.2: Dunlin Bayes implicit representation comparison showing that combined estimate has a smaller variance than individual carbon and nitrogen estimates.

Parameter		mean	sd	MC error	val2.5pc	median	val97.5pc
Carbon	π_1	0.361	0.054	0.0005	0.248	0.364	0.461
Nitrogen	π_1	0.351	0.076	0.0009	0.191	0.355	0.489
Combined	π_1	0.362	0.042	0.0004	0.275	0.363	0.442

6.1.2 Frequentist explicit representation

Because multiple measurements are rarely taken on mixtures, subject-specific modeling in the frequentist paradigm is not a valuable contribution at this time. But we can outline the steps to take. We can apply results from multivariate linear models for the data likelihood components of the mixing model. We can use general results for the multivariate linear model from Christensen (2001, sec. 1.1), applying these results for our four data components: mixtures, sources, discriminations, and assimilation efficiencies. For mixtures we can consider a mixed model with subject-specific random effects, for sources a fixed effects model, for discriminations pairs of fixed effects models or a regression model. For each model optimal parameter

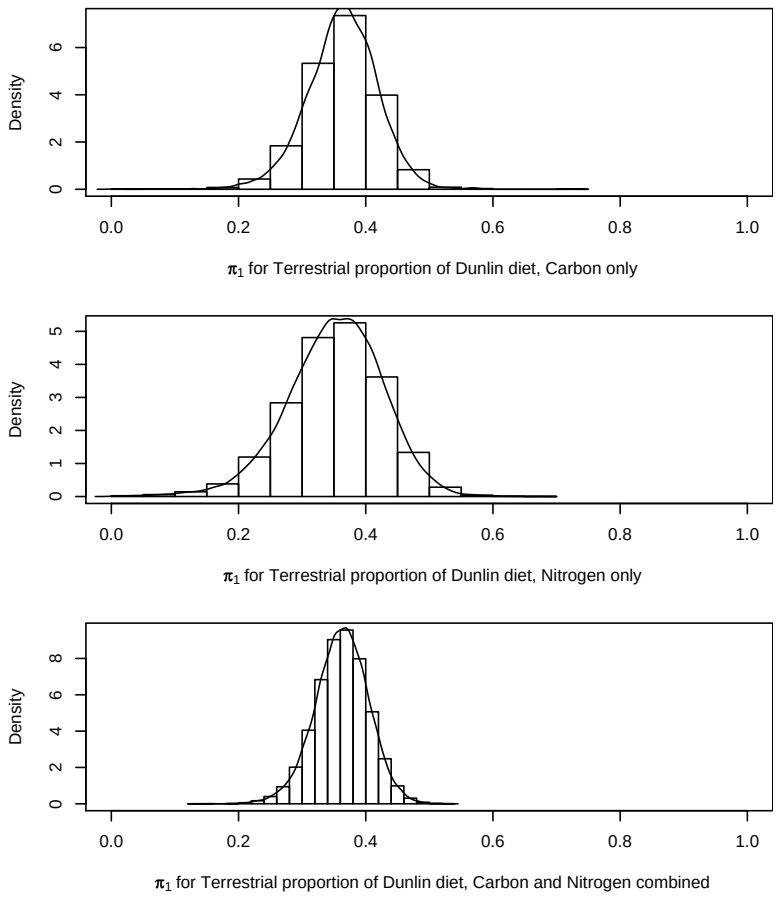

Figure 6.1: Dunlin Bayes implicit representation posterior distribution for π_1 for carbon alone, nitrogen alone, and using both isotopes together.

estimates are given by specifying a model in terms of the general multivariate linear model and applying the results to the vectorized univariate model. The individual mixture mean isotope ratio random effects are predicted with best linear unbiased predictors (BLUPs) (Robinson, 1991), and the mean diet vector and individual mixture diet vectors are estimated using the delta method. However, a nonparametric bootstrap should not be used with BLUPs because the variation will be consistently underestimated (Morris, 2002).

6.1.3 Extended mixing model

The extended mixing model introduced in Section 1.2.1.2 has two additional components for concentration and assimilation efficiency. In practice we would have samples on each of these and based upon their mean values, the implicit function theorem is modified to provide estimates of the mixture proportions and corresponding large sample distribution. Similarly, in the Bayesian approach one would simply append to the model the sampling distribution and the corresponding priors for each of the two additional components.

6.2 Future work

With this dissertation as a foundation, there are many more exciting statistical methods to develop for stable isotope sourcing.

6.2.1 Estimating diet-shift timing

The models in previous chapters are all cross-sectional in nature in that they concern a single point in time assuming the mixture is consuming from a set of sources

simultaneously. When animals undergo rapid dietary shifts due to migration, metamorphosis, seasonal food availability, or other reasons, the isotope ratio composition of their tissues begins changing to reflect their diet. After a shift in diet, animal tissues isotopically equilibriate with their new diet from both growth of new tissue and metabolic turnover of existing tissue (Fry and Arnold, 1982). Different tissues vary in their turnover rates with blood plasma and liver having high rates, muscle having a medium rate, and bone having a low rate (Tieszen et al., 1983). Additionally, during growth the increases in the mass of tissues has an additional dilution effect which results in faster equilibration to the new diet than would occur by metabolic turnover alone. Assuming a dietary step change has recently occurred, and given tissue measurements before and after the shift, the goal is to use the isotope ratio difference between tissues with different turnover rates to estimate the time of the diet shift.

Many studies support a modeling assumption that after a diet shift tissue isotope ratio composition changes follow an exponential model with a constant half-life period (Ayliffe et al., 2004; Bosley et al., 2002; Hobson and Clark, 1992; MacAvoy et al., 2001; Tieszen et al., 1983). Both growth and metabolic turnover contribute to the change in isotope ratio composition (Fry and Arnold, 1982; Hesslein et al., 1993). Hesslein et al. (1993) chose an exponential model for the isotope ratio change following a diet shift in their study of broad whitefish,

$$\beta_{iut} = \delta'_{iu\text{new}} + (\delta'_{iu\text{old}} - \delta'_{iu\text{new}}) \exp\{-(g_u + m_u)t\}, \tag{6.2}$$

where β_{iut} is the true isotope ratio in the mixture for isotope i in tissue type u at time t since the diet shift, $\delta'_{iu\text{old}}$ and $\delta'_{iu\text{new}}$ are the discrimination-corrected isotope ratios from the old and new diets, and g_u and m_u are the instantaneous rate constants for growth and metabolic turnover. Phillips and Eldridge (2006) introduce a method for estimating the time when the diet shift occurred, t, but variability in the estimation is not considered.

Given the work in Chapter 5 of this dissertation, a Bayesian approach to estimate the diet-shift time parameter is straightforward. Consider the model

$$\beta_{ut} = \begin{cases} \underset{\sim}{\delta}'_{u\text{new}} + (\underset{\sim}{\delta}'_{u\text{old}} - \underset{\sim}{\delta}'_{u\text{new}}) \exp\{-(\gamma_u + \mu_u)(t - t_0)\} & , \ t > t_0, u = 1, \ldots, U \\ \underset{\sim}{\delta}'_{u\text{old}} & , \ t \leq t_0 \end{cases} \tag{6.3}$$

Let t be time and let t_0 be the time of the diet shift, where t_0 is the primary parameter of interest. Assume the mixture's tissues, $u = 1, \ldots, U$, are in equilibrium with the "old" diet, $\underset{\sim}{\delta}'_{u\text{old}}$. Assume we know what the source of the new diet is, $\underset{\sim}{\delta}'_{u\text{new}}$. Let there be data to estimate $\underset{\sim}{\delta}'_{u\text{old}}$ and $\underset{\sim}{\delta}'_{u\text{new}}$ as in Section 5.1. Let there be a growth model to estimate γ_u and a metabolism model to estimate μ_u. Assume we have mixture data $\underset{\sim}{b}_{jut}$, $j = 1, \ldots, J$, $u = 1, \ldots, U$, before and after the time of the diet shift, t_0. For tissue type u, let $\pi_{u2} = \exp\{-(\gamma_u + \mu_u)(t - t_0)\}$, then it is immediate how to estimate t_0 as we have estimated π_{u2}, previously.

6.2.2 Differentiating between a diet shift or a mixed diet

Given a development of the diet shift model in Section 6.2.1 using (6.3), we can develop a test for whether a mixture has been consuming sources concurrently or in sequence. The ability to distinguish between the two diet scenaries will, in part, be a function of differences in tissue turnover rates, tissue-specific isotope discrimination differences, and isotope ratio differences in diet scenarios. For tissue type u, let $\pi_{u2} = \exp\{-(\gamma_u + \mu_u)(t - t_0)\}$. The idea is to test whether the estimated diets, π_{u2}, estimated from two tissues are different. If the diets differ then that suggests a diet shift, while if they are the same that suggests a constant diet. This is equivalent to testing whether t_0 differs between tissues.

6.2.3 Diet as a function of covariates

Researchers are interested in relating diet to covariates such as year, region, rainfall, altitude, species, body size, and more. The diet model we have developed here can be included in a larger model where these hypotheses can be tested.

6.2.4 Spatial models for migration

A migratory model will have components of both spatial and temporal models. Hobson and Wassenaar (2008) provide a detailed account of the use of stable isotope ratios for tracking animal migration. The plot in Figure 6.2 shows some information entering into the problem (from Hobson and Wassenaar (Figs. 2.2 and 2.3 2008)). On the left is an "isoscape" of $\delta^2 H$ of summer rainfall indicating a gradient from the equator to the poles. The water is incorporated into the tissues of plants and animals. On the right are $\delta^2 H$ values along the vane material for a feather grown at one location and a feather grown during a southward migration.

6.2.5 Modeling the diet vector with logit-Normal

The specification of the Dirichlet distribution for the population mean diet and subject-specific diets as in (5.27) and (5.39) is a convenient choice. The Dirichlet S-dimensional parameter vector at once defines the means and covariances of the elements of the random vector, $\underset{\sim}{\pi}$ in (5.27). A practitioner may prefer to define a more flexible covariance among at most $S-1$ elements of $\underset{\sim}{\pi}$. For example, rather than the always negative correlations from the Dirichlet distribution, one may wish to specify positive correlations between food sources that are typically consumed together. In this case it will be desirable to have a more flexible covariance structure for the diet proportion vector. One alternative is to use the logit-Normal distribution on an $S-1$-dimensional subset of $\underset{\sim}{\pi}$ with the constraint that $\underset{\sim}{\pi}$ is probability vector.

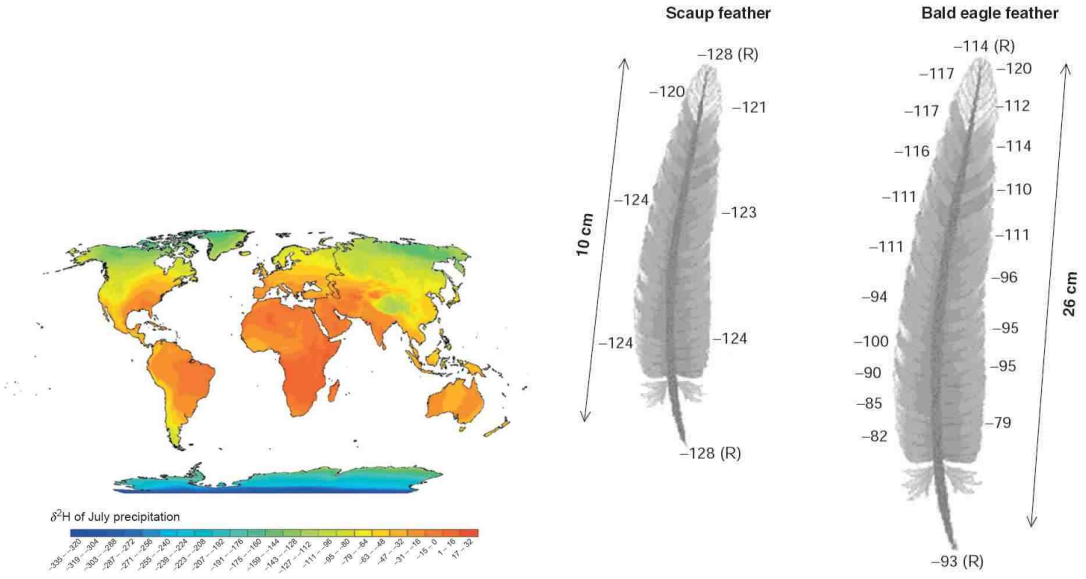

Figure 6.2: (Left) The average hydrogen isotopic ($\delta^2 H$) composition of rainfall across the globe in the summertime. Map available from Gabe Bowen at www.waterisotopes.org. (Right) Illustration of measured intrasample hydrogen ($\delta^2 H$) isotopic heterogeneity in migrating birds. Illustrated is a comparison of a flight feather from lesser scaup (*Aythya affinis*), left, that grew the entire feather at one location, and a bald eagle (*Haliaetus leucocephalus*), right, that grew the feather along its southward migration. Subsamples for hydrogen isotope analyses were taken from vane material (350 mg) up and down the left and right side of a single feather, and from the top and bottom portions of the rachis (R). Note that illustrated differences in feather length are not to scale and are in centimeters from the tip to base. Both plots from Hobson and Wassenaar (2008, Figs. 2.2 and 2.3) and $\delta^2 H$ numbers are in permil (VSMOW).

It is critical that the covariance matrix maintains an exchangable structure so that the results do not depend on the choice of the reference source. Below is an example using the explicit representation model.

Let the subject-specific diet vectors $\underset{\sim}{\pi}_j$ be defined in terms of $\underset{\sim}{\lambda}_j$, where

$$\pi_{js} = \frac{\exp\{\lambda_{js}\}}{1+\sum_{s=2}^{S}\exp\{\lambda_{js}\}}, \qquad \lambda_{j1} \equiv 0$$

with inverse

$$\lambda_{js} = \log\left(\frac{\pi_{js}}{1-\sum_{s=2}^{S}\pi_{js}}\right),$$

$s = 1, \ldots, S$ and $j = 1, \ldots, J$. In the explicit representation model, replace (5.39) by

$$\begin{bmatrix} \lambda_{j2} \\ \vdots \\ \lambda_{jS} \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} \lambda_2 \\ \vdots \\ \lambda_S \end{bmatrix}, \Sigma_{\lambda_b}\right) \tag{6.4}$$

$$\Sigma_{\lambda_b}|\Sigma_{0\lambda_b}, \nu_{0\lambda_b} \sim \text{Inv-Wishart}(\Sigma_{0\lambda_b}, \nu_{0\lambda_b}). \tag{6.5}$$

The mean diet $\underset{\sim}{\pi}$ is specified in terms of the mean $[\lambda_2, \ldots, \lambda_S]^\top$ from (6.4),

$$\pi_s = \frac{\exp\{\lambda_s\}}{1+\sum_{s=2}^{S}\exp\{\lambda_s\}}, \qquad \lambda_1 \equiv 0$$

$s = 1, \ldots, S$. Let $\underset{\sim}{\pi}_0$ be the prior mean for $\underset{\sim}{\pi}$, which induces a prior on $\underset{\sim}{\lambda}$, where

$$\lambda_{0s} = \log\left(\frac{\pi_{0s}}{1-\sum_{s=2}^{S}\pi_{0s}}\right),$$

$s = 1, \ldots, S$. Finally, (5.27) is replaced by

$$\begin{bmatrix} \lambda_2 \\ \vdots \\ \lambda_S \end{bmatrix} \sim \text{Normal}\left(\begin{bmatrix} \lambda_{02} \\ \vdots \\ \lambda_{0S} \end{bmatrix}, \Sigma_{\underset{\sim}{\lambda_\pi}} \right) \tag{6.6}$$

$$\Sigma_{\underset{\sim}{\lambda_\pi}} | \Sigma_{0\underset{\sim}{\lambda_\pi}}, \nu_{0\underset{\sim}{\lambda_\pi}} \sim \text{Inv-Wishart}(\Sigma_{0\underset{\sim}{\lambda_\pi}}, \nu_{0\underset{\sim}{\lambda_\pi}}). \tag{6.7}$$

The additional modeling flexibiliy requires additional parameters, $\Sigma_{0\lambda_b}$ and $\Sigma_{0\underset{\sim}{\lambda_\pi}}$, to be specified in this model such that the sources are exchangable.

Note that in this model specification it is not generally the case the $\text{E}[\underset{\sim}{\pi}_j] = \underset{\sim}{\pi}$. Instead, the hierarchy is defined in terms of $\underset{\sim}{\lambda}_j$, $j = 1, \ldots, J$, and $\underset{\sim}{\lambda}$ such that $\text{E}[\underset{\sim}{\lambda}_j] = \underset{\sim}{\lambda}$, that is, $\text{E}[\text{logit}(\underset{\sim}{\pi}_j)] = \text{logit}(\underset{\sim}{\pi})$.

6.2.6 Measurement error estimation

There are different ways of using standards to calibrate the measurements and each requires an appropriate method for modeling measurement error. A few ways include (1) using a linear regression fit to the standards over a single measurement run, (2) running pairs of identical standards and doing a "connect-the-means" interpolation between each pair, and (3) using splines to smoothly account for nonlinear measurement drift over the run. It would be valuable to have the estimation procedures detailed for each method in use, as well as a suggested method for calibration that maximizes the information relative to the number of standards run.

Additionally, we may wish to know whether the measurements over the range isotope ratio values is heteroscedastic. That is, does the error variation depend on the isotope ratio of the material, or the material itself, etceteras? If this is the case, then additional standards may be desired over others for certain isotope ratio ranges.

6.2.7 Subject-specific source model comparison

The explicit representation model performs poorly when there are mixture isotope ratios outside the solution polytope of the mean discrimination-corrected isotope ratios of the sources. Also, interpretation of the implicit representation model may be questionable when this is the case. A model with subject-specific sources can be conceived as in (5.43) and (5.44) and we can assess which of the two models is preferred.

6.2.8 Animal energy requirements

There are ranges of $\underset{\sim}{\pi}$ that are feasible to satisfy the energy or health requirements of the mixture (Martínez del Rio and Wolf, 2005, p. 155). For example, essential amino acids are required to be consumed as part of the diet for the synthesis of proteins in the mixture, thus a diet that excludes sources of these essential amino acids may not be feasible (at least it is not sustainable). In the frequentist paradigm, this can be done by restricting the range of inference for $\underset{\sim}{\pi}$. In the Bayesian paradigm, this can be done by restricting the range of the prior on $\underset{\sim}{\pi}$.

6.2.9 Dietary routing model

Dietary proteins and lipids may be preferentially "routed" to synthesis of body proteins or lipids, respectively (Krueger and Sullivan, 1984; Ambrose and Norr, 1993; Tieszen and Fagre, 1993). In such a situation, C isotopes in body proteins would be disproportionately labeled by dietary proteins, leading to an over-estimate of the fraction of protein-rich foods in the consumer's diet. Similarly, dietary lipids may be routed to synthesis of body fat (Stott et al., 1997). A model accounting for the preferential routing of specific source material to specific tissues is a simple extension accounting for this complexity.

6.2.10 Proportions of sources consumed mixing model

The proportions of sources consumed mixing model (PSCMM) is the same model as the EMM except that there is no simplex constraint in the defining equation, but instead the estimated vector is restricted to the unit cube. Given the biomass in the environment of each source before consumption, the PSCMM estimates the proportion of each source population consumed by the mixture population. Schuur et al. (2003) use this model to estimate proportions of fuel sources consumed during a boreal forest fire with isotope measurements of $\delta^{13}C$ and $\delta^{14}C$ in the carbon dioxide released from the fire. There is a limited biomass to be consumed by the fire consisting of needles, fine branches, cones, surface moss, etc., and it may be possible to completely consume some or all of these sources.

Let $\underset{\sim}{\mu}_s$ be the vector of elemental biomasses of source s in the population and let ϕ_s be the proportion of source s consumed (which takes the place of π_s in the BMM), $0 \leq \phi_s \leq 1,\ s = 1, \ldots, S$. The defining equation replacing (1.1) is

$$\beta_i \ = \ \frac{\sum_{s=1}^{S}(\delta_{is} + \Delta_{is})\mu_{is}\phi_s}{\sum_{s=1}^{S}\mu_{is}\phi_s}, \tag{6.8}$$

$i = 1, \ldots, I$ and $s = 1, \ldots, S$. The model has the same linear structure as (1.1) where

$$0 \ = \ \sum_{s=1}^{S}\left\{(\delta_{is} + \Delta_{is}) - \beta_i\right\}\mu_{is}\phi_s, \quad i = 1, \ldots, I. \tag{6.9}$$

This model is different from (1.1) in that it excludes the simplex constraint in the last row. The consumable biomass $\underset{\sim}{\mu}_s$ can be expressed as the product of the biomass per individual and the number of consumable individuals.

Some sources may have a tendency to be consumed faster than others. In the boreal fire example, most of the needles might burn before the moss begins to burn, and a severe fire could burn all the needles. An expert can put inequality constraints

(specified as prior information) on how much of each source will burn relative to other sources and relative to the severity of the fire. For example, the bole (tree trunk) bark proportion combusted is greater than the bole wood proportion combusted. One future challenge is to define priors over the unit cube that capture the propensity of some sources to be consumed at faster rates than others.

6.2.11 Trace element and isotope geochemistry

Many of the methods in this dissertation may naturally use trace elements in place of stable isotopes. For example, a two-source mixing model can be applied using barium as a proxy for salinity since the concentration of barium is high in freshwater (runoff from the continents) and low in the ocean. In addition to the use of stable isotopes, reconstructing past environments of organisms relies on ambient chemicals, primarily the elements strontium (Sr), calcium (Ca), barium (Ba), and to a lesser extent manganese (Mn), being incorporated from the environment into discrete layers of calcified material such as a shell of an organism. Environmental histories can be interpreted as either a change in elemental concentration, or where ambient chemicals correlate to environmental variables. An application is the reconstruction of the salinity experienced by fish (Elsdon and Gillanders, 2006). Ba/Ca in planktonic foraminiferal calcite (Ba/Ca_{foram}) provides an innovative tool to assess past variability in regional riverine runoff (Hall and Chan, 2004). Seawater Ba (Ba_{sw}) concentrations at oceanic sites influenced by riverine runoff have a notably high inverse correlation to salinity, with high Ba and low salinity at sites closest to the river mouth because dissolved Ba is high in riverine water and Ba desorbs from suspended sediments in estuaries (Edmond et al., 1978; Coffey et al., 1997). Laboratory experiments on living planktonic foraminifera demonstrate that Ba incorporation in foraminiferal calcite varies linearly with changes in Ba_{sw} concentration, independent of temperature changes within about 7 °C (Lea and Spero, 1994). Therefore, the variation of Ba/Ca_{foram} is controlled by the Ba_{sw} concentration, and the temporal

variation of Ba/Ca_{foram} provides valuable insights into changes in riverine freshwater input. Applications to the movement of barium in estuaries (Coffey et al., 1997; Shaw et al., 1998; Hanor and Chan, 1977) and the use of the Ba/Ca ratio in calcifying organisms to reconstruct salinity has a variety of applications (Elsdon and Gillanders, 2006; Weldeab et al., 2007; Hall and Chan, 2004). Trace elements can be coupled with stable isotope proxies that also can act as salinity proxies in certain cases. For example, in estuaries, the $\delta^{13}C$ value of dissolved inorganic carbon varies from near-zero down to -8 to -13 in freshwater because it is respired CO_2 from soil organic matter, which carries the same $\delta^{13}C$ value as the organic matter (with no fractionation).

6.3 Conclusions

In this dissertation we have considered the statistical issues associated with estimating the proportional contributions of sources to a mixture using stable isotopes, with particular attention to the issues in animal ecology. We have addressed the situation in three ways. First, by ignoring the issues of estimation, we applied a better algorithm implemented in our SISUS R package to provide representative samples of the solution polytope. Next, when the diet solution is perfectly-constrained, we developed general methods of frequentist estimation using the delta method and the bootstrap. Finally, for problems of any size, we develop two Bayesian models to account for uncertainty in estimating all model parameters. The implicit representation model estimates the population mean diet through the mean mixture isotope ratio, and SISUS can be viewed as this model where the variance of all model parameters goes to zero. The explicit representation model estimates the population mean diet through a distribution for subject-specific mixture diets centered at a population mean diet, and when measurement error is small and individual mixtures are within the scope of the sources, more efficient estimation is possible using this model. A number of examples have been used to illustrate the methods and compare the models. Model simplification and numerical stability has been addressed with the use of the three-step estimation approach. Other theoretical and practical details have been addressed throughout and a range of additional related research areas have been outlined.

References

SH Ambrose and L Norr. *Prehistoric human bone: archaeology at the molecular level*, chapter 1: Experimental evidence for the relationship of the carbon isotope ratios of whole diet and dietary protein to those of bone collagen and carbonate, pages 1–37. Springer, Berlin Heidelberg New York, 1993.

A Angerbjörn, P Hersteinsson, K Lidén, and E Nelson. Dietary variation in arctic foxes *(aiopex lagopus)* - an analysis of stable carbon isotopes. *Oecologia*, 99:226–32, 1994. Between individual variation.

MS Araújo, DI Bolnick, G Machado, AA Giaretta, and SF dos Reis. Using δ^{13}C stable isotopes to quantify individual-level diet variation. *Oecologia*, 152:643–54, 2007. between individual variation.

LK Ayliffe, TE Cerling, T Robinson, AG West, M Sponheimer, BH Passey, J Hammer, B Roeder, MD Dearing, and JR Ehleringer. Turnover of carbon isotopes in tail hair and breath CO_2 of horses fed an isotopically varied diet. *Oecologia*, 139 (1):11–22, 2004.

M Ben-David, TA Hanley, DR Klein, and DM Schell. Seasonal changes in diets of coastal and riverine mink: the role of spawning pacific salmon. *Canadian Journal of Zoology*, 75:803–811, 1997.

J Benichou and MH Gail. A delta method for implicitly defined random variables. *The American Statistician*, 43(1):410044, February 1989.

JP Benstead, JG March, B Fey, KC Ewel, and CM Pringle. Testing isosource: Stable isotope analysis of a tropical fishery with diverse organic matter sources. *Ecology*, 87:327–33, 2006.

A Boneh and A Golan. Constraints' redundancy and feasible region boundedness by random feasible point generator (rfpg). In *Third European Congress on Operations Research (EURO III)*, Amsterdam, April 9–11, 1979.

KL Bosley, DA Witting, RC Chambers, and SC Wainright. Estimating turnover rates of carbon and nitrogen in recently metamorphosed winter flounder Pseudopleuronectes americanus with stable isotopes. *Marine Ecology Progress Series*, 236:233–240, 2002.

CB Bouthitt and K Garnett. The evolution of the multicollector in isotope ratio mass spectromety. In *Proceedings of the 18th AMZSMS Conference: THO07*, 2006.

RW Butler and RW Campbell. The birds of the Fraser River delta: Populations, ecology and international significance. Canadian Wildlife Service Occasional Paper 65, Canadian Wildlife Service, Delta, British Columbia, Canada, 1987.

G Cabana and JB Rasmussen. Comparison of aquatic food chains using nitrogen isotopes. *Proceedings of the National Academy of Sciences*, 93:10844–7, 1996.

G Casella and EI George. Explaining the Gibbs sampler. *American Statistician*, pages 167–174, 1992.

S Caut, E Angulo, and F Courchamp. Variation in discrimination factors ($\Delta^{15}N$ and $\Delta^{13}C$): the effect of diet isotopic values and applications for diet reconstruction. *Journal of Applied Ecology*, 46(2):443–453, 2009.

TE Cerling, JM Harris, BJ MacFadden, MG Leakey, J Quade, V Eisenmann, and JR Ehleringer. Global vegetation change through the miocene/pliocene boundary. *Nature*, 389(6647):153–158, 1997.

S Chib and E Greenberg. Understanding the metropolis-hastings algorithm. *American Statistician*, pages 327–335, 1995.

R Christensen. *Advanced Linear Modeling*. Springer, 2 edition, 2001.

M Coffey, F Dehairs, O Collette, G Luther, T Church, and T Jickells. The behaviour of dissolved barium in estuaries. *Estuarine, Coastal and Shelf Science*, 45(1):113–121, 1997.

A Cree, GL Lyon, L Cartland-Shaw, and C Tyrrell. Stable carbon isotope ratios as indicators of marine versus terrestrial inputs to the diets of wild and captive tuatara (Sphenodon punctatus). *New Zealand Journal of Zoology*, 26(3):243–253, 1999.

AC Davison and DV Hinkley. *Bootstrap methods and their application*. Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, 1997.

M DeNiro and S Epstein. Influence of the diet on the distribution of carbon isotopes in animals. *Geochimica et Cosmochimica Acta*, 42:495–506, 1978.

M DeNiro and S Epstein. Influence of the diet on the distribution of nitrogen isotopes in animals. *Geochimica et Cosmochimica Acta*, 48:341–51, 1981.

JM Dickey, DV Lindley, and SJ Press. Bayesian estimation of the dispersion matrix of a multivariate normal distribution. *Communications in Statistics-Theory and Methods*, 14(5):1019–1034, 1985.

AP Dickin. *Radiogenic isotope geology*. Cambridge University Press, 2005.

JM Edmond, ED Boyle, D Drummond, B Grant, and T Mislick. Desorption of barium in the plume of the Zaire (Congo) River. *Netherlands Journal of Sea Research*, 12 (3-4):324–328, 1978.

B Efron. Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7: 1–26, 1979.

B Efron and RJ Tibshirani. *An Introduction to the Bootstrap*, volume 57 of *Monographs on Statistics & Applied Probability*. Chapmann & Hall, November 1993. ISBN 0412042312.

TS Elsdon and BM Gillanders. Temporal variability in strontium, calcium, barium, and manganese in estuaries: Implications for reconstructing environmental histories of fish from chemicals in calcified structures. *Estuarine, Coastal and Shelf Science*, 66(1-2):147–156, 2006.

LJ Evans Ogden. *Non-breeding shorebirds in a coastal agricultural landscape: winter habitat use and dietary sources*. PhD thesis, Dissertation, Simon Fraser University, Burnaby, British Columbia, Canada, 2002.

LJ Evans Ogden, KA Hobson, and DB Lank. Blood isotopic (δ^{13}C and δ^{15}N) turnover and diet-tissue fractionation factors in captive dunlin (Calidris alpina pacifica). *The Auk*, 121(1):170–177, 2004.

LJ Evans Ogden, KA Hobson, DB Lank, and S Bittman. Stable isotope analysis reveals that agricultural habitat provides an important dietary component for nonbreeding Dunlin. *Avian Conservation and EcologyÉcologie et conservation des oiseaux*, 1(3), 2005.

LA Felicetti, CC Schwartz, RO Rye, MA Haroldson, KA Gunther, DL Phillips, and CT Robbins. Use of sulfur and nitrogen stable isotopes to determine the importance of whitebark pine nuts to yellowstone grizzly bears. *Canadian Journal of Zoology*, 81:763–70, 2003.

B Fry and C Arnold. Rapid $^{13}C/^{12}C$ turnover during growth of brown shrimp (Penaeus aztecus). *Oecologia*, 54(2):200–204, 1982.

K Fukuda. *C-Library cddlib (version 0.94f)*, August 2005. URL `www.ifor.math.ethz.ch/~fukuda/cdd_home/cdd.html`.

K Fukuda and A Prodon. Double description method revisited. *Lecture Notes in Computer Science*, 1120:91–111, 1996. ISSN 0302-9743.

AE Gelfand and AFM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.

A Gelman, JB Carlin, HS Stern, and DB Rubin. *Bayesian Data Analysis*. Chapman and Hall, 1995.

J Geweke. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J M Bernado, J O Berger, A P Dawid, and A F M Smith, editors, *Bayesian Statistics 4*, Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics, pages 169–194. Oxford University Press, 1992.

JM Hall and LH Chan. Ba/Ca in Neogloboquadrina pachyderma as an indicator of deglacial meltwater discharge into the western Arctic Ocean. *Paleoceanography*, 19(1), 2004.

JS Hanor and LH Chan. Non-conservative behavior of barium during mixing of Mississippi River and Gulf of Mexico waters. *Earth Planet. Sci. Lett*, 37(2):242–250, 1977.

P Heidelberger and PD Welch. A spectral method for confidence interval generation and run length control in simulations. *Communications of the Association for Computing Machinery*, 24:233–45, 1981.

P Heidelberger and PD Welch. Simulation run length control in the presence of an initial transient. *Operations Research*, 31:1109–44, 1983.

LG Herrera M, KA Hobson, JC Martínez, and G Méndez C. Tracing the Origin of Dietary Protein in Tropical Dry Forest Birds. *Biotropica*, 38(6):735–742, 2006.

RH Hesslein, KA Hallard, and P Ramlal. Replacement of sulfur, carbon, and nitrogen in tissue of growing broad whitefish (Coregonus nasus) in response to a change in diet traced by $\delta^{34}S$, $\delta^{13}C$, and $\delta^{15}N$. *Canadian Journal of Fisheries and Aquatic Sciences*, 50(10):2071–2076, 1993.

GV Hildebrand, CC Schwartz, CT Robbins, ME Jacoby, TA Hanley, SM Arthur, and C Servheen. The importance of meat, particularly salmon, to the body size, population productivity, and conservation of north american brown bears. *Canadian Journal of Zoology*, 77:132–8, 1999.

GV Hilderbrand, SD Farley, CT Robbins, TA Hanley, K Titus, and C Servheen. Use of stable isotopes to determine diets of living and extinct bears. *Canadian journal of zoology*, 74(11):2080–2088, 1996.

KA Hobson. Tracing origins and migration of wildlife using stable isotopes: a review. *Oecologia*, 120(3):314–326, 1999.

KA Hobson and RG Clark. Assessing avian diets using stable isotopes I: turnover of ^{13}C in tissues. *Condor*, pages 181–188, 1992.

KA Hobson and RG Clark. Turnover of ^{13}C in cellular and plasma fractions of blood: Implications for nondestructive sampling in avian dietary studies. *Auk*, 110(3):638–641, 1993.

KA Hobson and LI Wassenaar. Stable isotope ecology: an introduction. *Oecologia*, 120(3):312–313, 1999.

KA Hobson and LI Wassenaar, editors. *Tracking Animal Migration with Stable Isotopes*. Academic Press, 2008.

AL Jackson, R Inger, S Bearhop, and A Parnell. Erroneous behaviour of MixSIR, a recently published bayesian isotope mixing model: a discussion of Moore & Semmens, Ecology Letters, 2008. *Ecology Letters*, 12(3):E1–E5, 2009.

ME Jacoby, GV Hilderbrand, C Servheen, CC Schwartz, SM Arthor, TA Hanley, CT Robbins, and R Michener. Trophic relations of brown and black bears in several western North American ecosystems. *Journal of Wildlife Management*, 63: 921–9, 1999.

TD Jardine and RA Cunjak. Analytical error in stable isotope ecology. *Oecologia*, 144:528–33, 2005.

WH Karasov. Digestion in birds: Chemical and physiological determinants and ecological implications. In M L Morrisona, C J Ralph, J Verner, and J R Jehl, editors, *Avian Foraging: Theory, Methodology, and Applications*, number 13 in Studies in Avian Biology, pages 391–415. Cooper Ornithological Society, Kansas, 1990.

PL Koch and DL Phillips. Incorporating concentration dependence in stable isotope mixing models: a reply to Robbins, Hilderbrand and Farley (2002). *Oecologia*, 133:14–18, 2002.

S Kotz, N Balakrishnan, and NL Johnson. *Continuous Multivariate Distributions, Models and Applications*. Wiley-Interscience, 2004.

HW Krueger and CH Sullivan. Models for carbon isotope fractionation between diet and bone. *Stable Isotopes in Nutrition*, 205:220, 1984.

CA Layman, JP Quattrochi, CM Peyer, and JE Allgeier. Niche width collapse in a resilient top predator following ecosystem fragmentation. *Ecology Letters*, 10: 937–44, 2007.

R Lazar and G Meeden. Isipta '03, proceedings of the third international symposium on imprecise probabilities and their applications, Lugano, Switzerland, July 14-17, 2003. In Jean-Marc Bernard, Teddy Seidenfeld, and Marco Zaffalon, editors, *ISIPTA*, volume 18 of *Proceedings in Informatics*, pages 361–371. Carleton Scientific, 2003. ISBN 1-894145-17-8.

DW Lea and HJ Spero. Assessing the reliability of paleochemical tracers: Barium uptake in the shells of planktonic foraminifera. *Paleoceanography*, 9(3):445–452, 1994.

DR Lide, editor. *CRC Handbook of Chemistry and Physics*. CRC, 89 edition, June 2008.

SC Lubetkin and CA Simenstad. Multi-source mixing models to quantify food web sources and pathways. *Journal of Applied Ecology*, 41:996–1008, 2004.

DJ Lunn, A Thomas, N Best, and D Spiegelhalter. WinBUGS-a bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing*, 10 (4):325–337, 2000.

SE MacAvoy, SA Macko, and GC Garman. Isotopic turnover in aquatic predators: quantifying the exploitation of migratory prey. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(5):923–932, 2001.

C Martínez del Rio and BO Wolf. Mass-balance models for animal isotopic ecology. In J Matthias Starck, Tobias Wang, and Tobias Wang, editors, *Physiological and Ecological Adaptations to Feeding in Vertebrates*, chapter 6, pages 141–174. Science Publishers, Inc., Enfield, NH, USA, February 2005.

B Matthews and A Mazumder. A critical evaluation of intrapopulation variation of δ^{13}C and isotopic evidence of individual specialization. *Oecologia*, 140:361–71, 2004. between individual variation.

JH McCutchan, Jr, WM Lewis, Jr, C Kendall, and CC McGrath. Variation in trophic shift for stable isotope ratios of carbon, nitrogen, and sulfur. *Oikos*, 102:378–90, 2003.

G Meeden and R Lazar. *polyapost: Simulating from the Polya posterior*, 2006. R package version 1.1.

W Meier-Augenstein. Applied gas chromatography coupled to isotope ratio mass spectrometry. *Journal of Chromatography A*, 842(1-2):351–371, 1999.

M Minagawa. Reconstruction of human diet from δ^{13}C and δ^{15}N in contemporary Japanese hair: a stochastic method for estimating multi-source contribution by double isotopic tracers. *Applied Geochemistry*, 7:145–158, 1992.

M Minagawa and E Wada. Stepwise enrichment of ^{15}N along food chains: further evidence and the relation between δ^{15}N and animal age. *Geochimica et Cosmochimica Acta*, 48:1135–40, 1984.

JW Moore and BX Semmens. Incorporating uncertainty and prior information into stable isotope mixing models. *Ecology Letters*, 11:470–80, 2008.

JS Morris. The BLUPs are not "best" when it comes to bootstrapping. *Statistics and Probability Letters*, 56:425–30, 2002.

CM O'Reilly, RE Hecky, AS Cohen, and PD Plisnier. Interpreting stable isotopes in food webs: recognizing the role of time averaging at different trophic levels. *Limnology and Oceanography*, 47:306–9, 2002.

A Parnell and A Jackson. *siar: Stable Isotope Analysis in R*, 2008. R package version 3.3.

D Paul, G Skrzypek, and I Forizs. Normalization of measured stable isotopic compositions to isotope reference scales–a review. *Rapid Communications in Mass Spectrometry*, 21(18):3006–14, 2007.

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

DL Phillips. *IsoSource (Version 1.3.1)*, August 2006. URL `www.epa.gov/wed/pages/models/stableIsotopes/isosource/isosource.htm`. See (Phillips and Gregg, 2003).

DL Phillips and PM Eldridge. Estimating the timing of diet shifts using stable isotopes. *Oecologia*, 147(2):195–203, Mar 2006. doi: 10.1007/s00442-005-0292-0. URL `dx.doi.org/10.1007/s00442-005-0292-0`.

DL Phillips and JW Gregg. Uncertainty in source partitioning using stable isotopes. *Oecologia*, 127:171–179, 2001.

DL Phillips and JW Gregg. Source partitioning using stable isotopes: coping with too many sources. *Oecologia*, 136(2):261–269, July 2003.

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

M Plummer, N Best, K Cowles, and K Vines. CODA: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1):7–11, March 2006.

DM Post. Using stable isotopes to estimate trophic position: models, methods, and assumption. *Ecology*, 83:703–18, 2002.

DM Post, CA Layman, DA Arrington, G Takimoto, J Quattrochi, and CG Monta na. Getting to the fat of the matter: models, methods and assumptions for dealing with lipids in stable isotope analyses. *Oecologia*, 152:179–89, 2007.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2006.

AE Raftery and SM Lewis. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, 7:493–7, 1992.

AE Raftery and SM Lewis. *Practical Markov Chain Monte Carlo*, chapter The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms. Chapman and Hall, 1995.

CT Robbins, GV Hilderbrand, and SD Farley. Incorporating concentration dependence in stable isotope mixing models: a response to Phillips and Koch (2002). *Oecologia*, 133:10–13, 2002.

GK Robinson. That BLUP is a good thing – the estimation of random effects. *Statistical Science*, 6(1):15–51, 1991.

DB Rubin. Using the SIR algorithm to simulate posterior distributions. *Bayesian statistics*, 3:395–402, 1988.

HC Sarakinos, ML Johnson, and MJ Vander Zanden. A synthesis of tissue-preservation effects on carbon and nitrogen stable isotope signatures. *Canadian Journal of Zoology*, 80:381–7, 2002.

FE Satterthwaite. An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2:110–4, 1946.

DA Schoeller, M Minagawa, R Slater, and IR Kaplan. Stable isotopes of carbon, nitrogen and hydrogen in the contemporary North American human food web. *Ecology of Food and Nutrition*, 18:159–170, 1984.

EAG Schuur, SE Trumbore, MC Mack, and JW Harden. Isotopic composition of carbon dioxide from a boreal forest fire: Inferring carbon loss from measurements and modeling. *Global Biogeochemical Cycles*, 17(1):1001–1010, 2003. doi: 10.1029/2001GB001840.

HP Schwarcz. Some theoretical aspects of isotope paleodiet studies. *Journal of Archaeological Science*, 18:261–275, 1991.

TJ Shaw, WS Moore, J Kloepfer, and MA Sochaski. The flux of barium to the coastal waters of the southeastern USA: The importance of submarine groundwater discharge. *Geochimica et Cosmochimica Acta*, 62(18):3047–3054, 1998.

PCF Shepherd. *Space Use, Habitat Preferences, and Time-activity Budgets on Non-breeding Dunlin (Calidris Alpina Pacifica) in the Fraser River Delta, BC.* PhD thesis, Simon Fraser University, Burnaby, British Columbia, Canada, 2001.

RL Smith. Monte Carlo procedures for generating random feasible solutions to mathematical programs. In *ORSA/TIMS Conference*, Washington, May 1980.

RL Smith. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32(6):1296–1308, 1984.

D Spiegelhalter, A Thomas, N Best, and D Lunn. *WinBUGS user manual*, 2003.

DJ Spiegelhalter, NG Best, BP Carlin, and A van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, pages 583–639, 2002.

RW Sterner and JJ Elser. *Ecological Stoichiometry.* Princeton University Press, 2002.

AW Stott, E Davies, and RP Evershed. Monitoring the routing of dietary and biosynthesised lipids through compound–specific stable isotope ($\delta^{13}C$) measurements at natural abundance. *Naturwissenschaften*, 84(2):82–6, 1997.

AE Taylor and WR Mann. *Advanced Calculus.* Wiley, 3rd edition, January 1983.

LL Tieszen and T Fagre. *Prehistoric human bone: archaeology at the molecular level*, chapter Effect of diet quality and composition on the isotopic composition of respiratory CO_2, bone collagen, bioapatite, and soft tissues, pages 121–155. Springer, Berlin Heidelberg New York, 1993.

LL Tieszen, TW Boutton, KG Tesdahl, and NA Slade. Fractionation and turnover of stable carbon isotopes in animal tissues: Implications for δ^{13}C analysis of diet. *Oecologia (Berlin)*, 57:32–7, 1983.

B Tummers. Datathief III. datathief.org, 2006.

EJM Urton and KA Hobson. Intrapopulation variation in gray wolf isotope (δ^{15}N and δ^{13}C) profiles: implications for the ecology of individuals. *Oecologia*, 145: 317–26, 2005. between individual variation.

MJ Vander Zanden and JB Rasmussen. Primary consumer δ^{13}C and δ^{15}N and the trophic position of aquatic consumers. *Ecology*, 80:1395–1404, 1999.

MJ Vander Zanden and JB Rasmussen. Variation in δ^{15}N and δ^{13}C trophic fractionation: implications for aquatic food web studies. *Limnology and Oceanography*, 46:2061–6, 2001.

MA Vanderklift and S Ponsard. Sources of variation in consumer-diet δ^{15}N enrichment: a meta-analysis. *Oecologia*, 136:169–182, 2003.

E Vulla, KA Hobson, M Korsten, M Leht, AJ Martin, A Lind, P Männil, H Valdmann, and U Saarma. Carnivory is positively correlated with latitude among omnivorous mammals: evidence from brown bears, badgers and pine martens. *Annales Zoologici Fennica*, 46:(preprint), 2009.

S Weldeab, DW Lea, RR Schneider, and N Andersen. 155,000 years of West African monsoon and ocean thermal evolution. *Science*, 316(5829):1303, 2007.

Appendices

Appendix A

Appendix to Chapter 1

A.1 Nature of the data

A detailed description of the nature of each data component will help inform modeling. This section can stand apart from the rest of the dissertation, used for reference as necessary. We start with nuclide definitions (modified from Lide (2008)), provide terminology for the model expressions we use, discuss relevant properties of common elements used in stable isotope sourcing, and mention how the data are collected.

A.1.1 Nuclide definitions

nuclide: a nuclear species which is characterized by the number of protons and neutrons that every atomic nucleus of this species contains. It can be used to distinguish isotopes among nuclei, as well as other properties, represented with ${}^{A}_{Z}E$, for example for standard carbon, ${}^{12}_{6}C$. The subscript Z is often dropped because the elemental symbol uniquely defines the atomic number.

atomic mass: the mass of a neutral atom of a nuclide

atomic mass unit (AMU): a measure of atomic masses, where 1 AMU = 1/12 the mass of ^{12}C. 1 AMU = 931.4812 MeV, 1.660565×10^{-27} kg.

atomic number (Z): integer that expresses the number of protons in a nucleus. Each element has a unique Z.

decay constant (λ): decay rate of a radioactive isotope. $\frac{\partial N}{\mathrm{d}t} = -\lambda N$, $N = N_0 e^{-\lambda t}$.

electron: subatomic particle having negative charge, with a mass of 1/1837 that of a hydrogen nucleus. rest mass: $0.9109534 \times 10^{-30}$ kg. All atoms consist of a nucleus and one or more electrons.

half-life: time required for one half of the atoms in a sample of radioactive isotope to decay. $t_{1/2} = \ln(2)/\lambda$.

isobar: same atomic mass but different atomic number (constant A, different Z).

isomer: one of two or more nuclides having the same mass number (A) and atomic number (Z) but existing for measureable times in different quantum states with different energies and radioactive half-lives.

isotone: family of isotopes having the same number of neutrons (N).

isotope: one of several nuclides having the same number of protons (Z) in their nuclei, but with different numbers of neutrons (N).

mass number (A): the sum of protons and neutrons in a nucleus, A=N+Z.

neutron: subatomic particle with zero charge and a rest mass of 1.67482×10^{-27} kg. It is unstable with respect to beta decay with a half-life of 12 minutes. The number of neutrons in a nucleus is represented by N.

proton: a subatomic particle with a positive charge and a rest mass of 1.67252×10^{-27} kg. The number of protons in an atom defines the element.

radioactive decay: decay of an isotope to a different isotope with a more stable energy configuration.

A.1.2 Isotope Ratio Mass Spectroscopy

The isotope ratio mass spectrometer (IRMS) allows the precise measurement of mixtures of stable isotopes (Paul et al., 2007). This technique has two different applications in the earth and environmental sciences. The analysis of stable isotopes is normally concerned with measuring isotope ratio variations arising from mass-dependent isotopic fractionation in natural systems.

Most instruments used for precise determination of isotope ratios are of the magnetic sector type. This type of analyzer is superior to the quadrupole type in this field of research for two reasons. Firstly, it can easily be set up for multiple-collector analysis, and secondly, it gives high-quality "peak shapes". Both of these considerations are important for isotope ratio analysis at very high precision and accuracy (Dickin, 2005).

The sector-type instrument designed by Alfred Nier was such an advance in mass spectrometer design that this type of instrument is often called the "Nier type". In the most general terms the instrument operates by ionizing the sample of interest, accelerating it over a potential in the kilo-volt range, and separating the resulting stream of ions according to their mass to charge ratio (m/z), see Figure A.1. The current of each ion beam is then measured using a "Faraday" detector or multiplier detector.

Many radiogenic isotope measurements are made by ionization of a solid source, whereas stable isotope measurements of light elements (e.g. H, C, O) are usually made in an instrument with a gas source. In the latter case, dual gas inlets enable reliable repetition of measurements by supplying continuous streams of the refer-

ence and sample gases, which are sequentially switched by a changeover valve. The IRMS's collector also features an array of Faraday cups (conductive, metal vessels which neutralise ions that hit them whilst themselves becoming charged), or "multicollector", which allows the simultaneous detection of multiple isotopes (Bouthitt and Garnett, 2006). Samples must be introduced as pure gases, achieved through combustion, gas chromatographic feeds (Meier-Augenstein, 1999), or chemical trapping. By comparing the detected isotopic ratios to a measured standard, an accurate determination of the isotopic make up of the sample is obtained, see Table A.2 below. For example, carbon isotope ratios are measured relative to the international standard for CO_2.

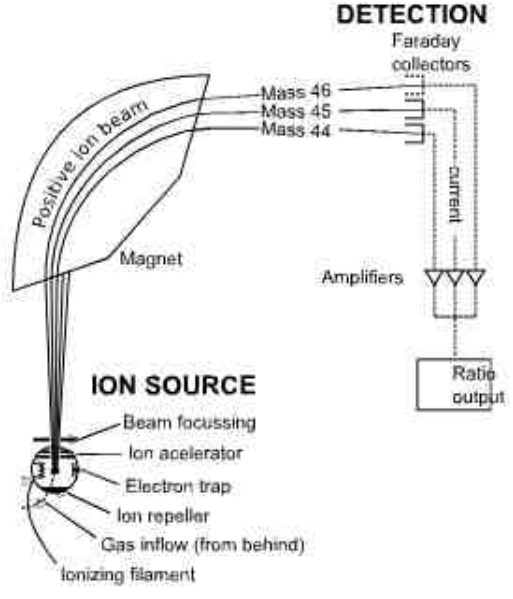

Figure A.1: Schematic of an isotope ratio mass spectrometer measuring CO_2 at the United States Geological Survey.

A.1.3 Terminology

The isotope ratio, ${}^aR$. The relative abundance of two isotopes a and b of element E can be expressed as an isotopic ratio, ${}^aR$,

$$ {}^aR \;=\; \frac{{}^aE}{{}^bE}, \qquad \text{e.g., for carbon} \quad {}^{13}R \;=\; \frac{{}^{13}C}{{}^{12}C}. $$

${}^aE$ and ${}^bE$ is the number (or proportion) of the rare and common isotopes in a sample, respectively. By convention, the more abundant isotope is placed in the denominator. This notation is not particularly helpful in itself, however, as changes in ${}^aR$ could result from changes in either ${}^aE$ or ${}^bE$. Typically two isotopes comprise nearly all the abundance of an element, and thus the ratio R can be thought of similar to an odds. It is however the basis for several expressions that are very useful.

Fractional abundance, ${}^aF$.

$$ {}^aF \;=\; \frac{{}^aE}{{}^aE + {}^bE}, \qquad \text{e.g., for carbon} \quad {}^{13}F \;=\; \frac{{}^{13}C}{{}^{13}C + {}^{12}C}. $$

This expression is particularly useful in artificially enriched systems where the ratio of ${}^aE$ to ${}^bE$ is increased by the intentional addition of a pure source of one of the two isotopes (usually the heavy isotope). The pure source is referred to as a spike, the process of adding the source as "spiking" a sample. Spikes allow us to monitor the movement of isotopes from one reservoir to another in response to one of more chemical reactions.

Relationship between ${}^aR$ and ${}^aF$. Note that

$$ {}^aF \;=\; \frac{{}^aR}{1 + {}^aR}, \quad \text{or alternatively,} \quad {}^aR \;=\; \frac{{}^aF}{1 - {}^aF}. $$

The delta notation, $\delta^a E$. Studies examining stable isotopes at or near natural abundance levels are usually reported as delta, δ. Lowercase delta values are not absolute isotope abundances but differences between sample readings and one or another of the widely used natural abundance standards which are considered delta = zero, $\delta^a E_{\text{standard}} \equiv 0$. The isotopic ratio ${}^aR$ exhibits variance in the range of the third to fifth decimal place. Numbers this small are better presented in terms of per mil, or parts per thousand (‰). This variance can be expressed relative to the isotopic ratio of a standard ${}^aR_{\text{standard}}$, which should be taken to be a constant. This difference relationship known as the delta notation ($\delta^a E$),

$$\delta^a E \;=\; \left(\frac{{}^aR_{\text{sample}} - {}^aR_{\text{standard}}}{{}^aR_{\text{standard}}}\right) \times 1000, \quad \text{or alternatively,} \quad \delta^a E \;=\; \left(\frac{{}^aR_{\text{sample}}}{{}^aR_{\text{standard}}} - 1\right) \times 1000.$$

This formulation of $\delta^a E$ scales the relative abundance values near zero and in magnitude of 10s. It is clear that $\delta^a E$ is on a nonlinear scale since the value is a function of the ratio of the abundance of one isotope to another. The plot in Figure A.2a shows that within the range of natural abundance for H, C, and N, the delta ratio is indistinguishable from linear. Figure A.2b shows what $\delta^2 H$ looks like when the range goes from 0 to 0.1, 0.5, and 0.99 to illustrate that $\delta^2 H$ is globally nonlinear. Note that a relative abundance of 0.1 for ${}^2H$ is 1000 times larger than its natural abunance. Because the heavy isotope is relatively rare in comparison to the light isotope the isotope ratio is locally approximately linear, but nonlinear thoughout the range of possible heavy to light proportions.

Change of isotopic reference scale. To go from scale 1 to scale 2, for example for oxygen that has two reference scales, [this may not be correct — check]

$$\delta^a E_{\text{sample}_2} \;=\; \left\{\left(\frac{\delta^a E_{\text{sample}_1}}{1000} + 1\right)\left(\frac{\delta^a E_{\text{standard}_2}}{1000} + 1\right) - 1\right\} \times 1000.$$

This relationship can be used to convert an isotope ratio value from one reference scale to another reference scale.

Measures of fractionation (discrimination) α, $1000\ln(\alpha)$, ε, and Δ. We are interested in measuring the isotopic offset between substances A and B. Such offsets arise from the expression of an isotope effect due to equilibrium or kinetic fractionation during a physical process or chemical reaction. The size of this isotope ratio fractionation can be expressed in several ways.

Fractionation factor α.

$$\alpha_{AB} \;=\; K^{1/n} \;=\; \frac{{}^{a}R_A}{{}^{a}R_B},$$

where K is the equilibrium constant for the associated reaction (such as going from vapor to liquid phase) and S is the number of atoms exchanged. It is assumed that A and B are in isotopic equilibrium. For simplicity, isotope exchange equations are usually written such that $S = 1$ so that $\alpha = K$. It is worth noting that most values of α are close to 1, with variability in the third through fifth decimal place. We can see that this is the case if we express a in terms of the delta notation

$$\alpha_{AB} \;=\; \frac{1+\frac{\delta^a E_A}{1000}}{1+\frac{\delta^a E_B}{1000}} \;=\; \frac{1000+\delta^a E_A}{1000+\delta^a E_B}.$$

Because α is close to unity, it is convenient to express fractionation in ways that accentuate the differences between $\delta^a E_A$ and $\delta^a E_B$. This can be done in one of three ways which yield approximately the same value for the per mil fractionation.

Natural log notation, $1000\ln(\alpha)$. One would think that this notation would have its own symbol, but it does not. We can approximate the fractionation in per mil from the fractionation factor α by taking the natural log of α and multiplying it by 1000. This is written simply as $1000\ln(\alpha)$. Mathematically, this works because we can think of α as being composed of $1+\varepsilon$, a small deviation. Now if we take the natural log of $1+\varepsilon$, the result can be expanded as a Maclaurin series in which

the first term, ε, is the largest and thus serves as a reasonable approximation of the natural log of α. If we retain more terms, we would obtain a more accurate result, but by convention, only the first term is retained,

$$1000\ln(\alpha) \;=\; 1000\ln(1+\varepsilon) \;\doteq\; \varepsilon.$$

Multiplying by 1000 yields the result in per mil. But since we are interested in ε anyway, there is an easier way to express the per mil fractionation.

The epsilon notation, ε. The isotope enrichment is

$$\varepsilon \;=\; (\alpha_{AB}-1)\times 1000 \;=\; \left(\frac{{}^{a}R_A}{{}^{a}R_B}-1\right)\times 1000.$$

The epsilon notation has the advantage over the $1000\ln(\alpha)$ notation in that it is an exact expression of the per mil fractionation. Note, in this body of this dissertation ε denotes measurement error rather than fractionation. There is a final way of determining the per mil fractionation. This last method is the least accurate, but most commonly applied because of its simplicity.

The capital delta notation, Δ. The isotope separation (or discrimination) is

$$\Delta_{AB} \;=\; \delta^{a}E_A - \delta^{a}E_B.$$

This method is least accurate because the errors in the two isotope ratio measurements do not cancel as is the case for random errors when calculating a ratio. In most cases, however, it is sufficiently accurate. Note that Δ_{AB} will depend on the reference material, for example, for oxygen since slope is different than 1 in the relationship $\delta^{18}O_{\mathrm{VSMOW}} = 1.030870 \times \delta^{18}O_{\mathrm{VPBD}} + 30.86$, where $1.030870 = R_{\mathrm{VPBD}}/R_{\mathrm{VSMOW}}$.

Concentration. The instrument software uses the following approach to calculate elemental concentration (Craig Cook SIRFER lab, Univeristy of Utah department of Biology, personal communication): 1) the user defines a known weight percentage reference material, 2) the user accurately weighs the reference material and analyzes it along with the unknowns, 3) the software integrates the peak area of the reference material for a given isotope, 4) the software calculates a kfactor relating peak area, weight, and elemental composition (weight percentage) of the reference material, 5) the software applies the kfactor to all unknowns. What the software is not capable of doing is checking for instrument linearity by allowing one to run a series of weighed reference materials to correct for any non-linearity.

the software simply divides by the weight so you can divide by the weight offline. When we do that, we always input a weight of 1 into the software so the software never has to divide by zero.

Assimilation Efficiency. The "true" assimilation efficiency used in the models is the fraction of ingested element absorbed, where e_{is} is the biomass of element i from source s assimilated divided by the biomass of source s ingested. This is different from the apparent assimilation efficiency (aae), which includes endogenous fecal losses, where aae_{is} is one minus the sum of the biomass of element i from source s not assimilated and the endogenous fecal losses of source s divided by the biomass of source s ingested (Karasov, 1990).

Notational Caveats. Different fields use these terms and equations in different ways. Because these ideas primarily came out of geology and oceanography, I have given their terms above. However, since biology has used stable isotopes, they have confused some of the notation. For example, $\alpha_{AB} = \frac{{}^{a}R_A}{{}^{a}R_B}$ is called the "isotope effect" where A is the source and B is the product and equilibrium is not assumed. Also, sometimes "isotope discrimination" is $\Delta_{AB} = (\alpha_{AB} - 1) \times 1000$, which earlier was

the equation for ε. So, it is important to define the equation for the notation being used and not to make assumptions.

A.1.4 Common Stable Isotopes

Table A.1 shows the relative abundance of common stable isotopes used for biological sourcing, note that an element may have multiple heavy isotopes. Table A.2 shows the (somewhat arbitrary) standards used as the constant in the denominator for δE. Table A.3 illustrates what the δE values are for the natural abundance of the common isotopes.

Table A.1: Commonly used stable isotopes for biological sourcing applications with natural abundance.

symbol	Z(p)	N(n)	mass	fraction	range	R
^{1}H	1	0	1.0078	.999885	.999816— .999974	
^{2}H	1	1	2.0141	.000115	.000026— .000184	.0001150132
^{12}C	6	6	12 by def	.9893	.98853— .99037	
^{13}C	6	7	13.003	.0107	.00963— .01147	.01081573
^{14}N	7	7	14.003	.99636	.99579— .99654	
^{15}N	7	8	15.000	.00364	.00346— .00421	.003653298
^{16}O	8	8	15.995	.99757	.99738— .99776	
^{17}O	8	9	16.999	.00038	.00037— .00040	.0003809256
^{18}O	8	10	17.999	.00205	.00188— .00222	.002054994
^{32}S	16	16	31.972	.9499	.94454— .95281	
^{33}S	16	17	32.971	.0075	.00730— .00793	.007895568
^{34}S	16	18	33.968	.0425	.03976— .04734	.04474155
^{36}S	16	20	35.967	.0001	.00013— .00019	.0001052742

Table A.2: Primary stable isotope standards. Standards are Vienna Standard Mean Ocean Water (VSMOW), Vienna Pee Dee Belemnite (VPDB), Atmosphere air (Air), and Canyon Diablo Troilite (Meteor Crater, AZ) (CDT).

Standard	isotopes	R
VSMOW	${}^{2}H/{}^{1}H$	0.00015576
VSMOW	${}^{18}O/{}^{16}O$	0.00200520
VPDB	${}^{13}C/{}^{12}C$	0.0112372
VPDB	${}^{18}O/{}^{16}O$	0.0020671
Air	${}^{2}He/{}^{4}He$	0.00000138
Air	${}^{15}N/{}^{14}N$	0.0036765
CDT	${}^{34}S/{}^{32}S$	0.04500451

Table A.3: Natural abundance δE for common isotopes.

Standard	isotopes	δE	natural range	sig dig
V-SMOW	${}^{2}H/{}^{1}H$	$-262‰$	600‰	1
V-PDB	${}^{13}C/{}^{12}C$	$-37.5‰$	100‰	0.1
Air	${}^{15}N/{}^{14}N$	$-6.3‰$	30‰	0.1
V-SMOW	${}^{18}O/{}^{16}O$	24.8‰	100‰	0.1
V-PDB	${}^{18}O/{}^{16}O$	$-5.9‰$	100‰	0.1
CDT	${}^{34}S/{}^{32}S$	$-5.8‰$	100‰	0.1

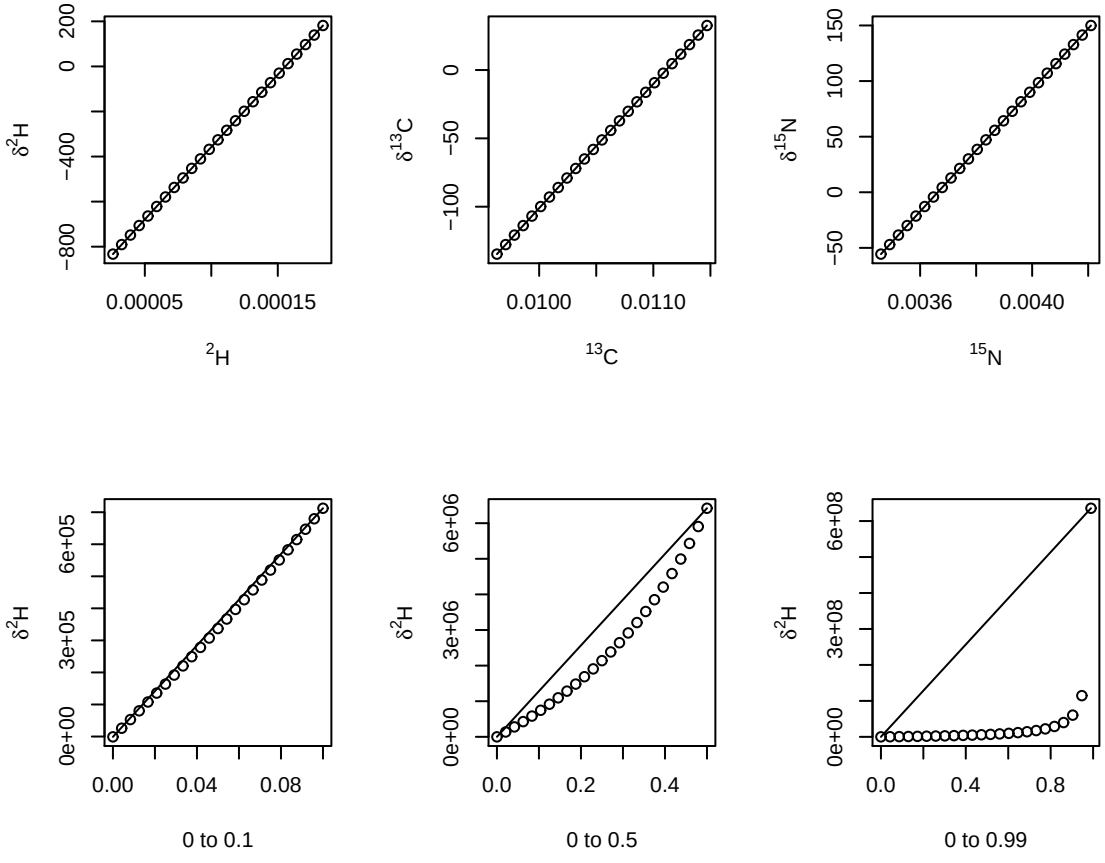

Figure A.2: (a, top) $\delta^a E$ is linear in the range of natural abundance for H, C, and N. The range of natural abundance was taken from Table A.1 and $\delta^a E$ was calculated using the standards given in Table A.2. (b, bottom) However, $\delta^a E$ is globally nonlinear, using H for this example. Note the change of scale in the horizontal axis.