Ecological Applications, 15(1), 2005, pp. 315–325

COMBINING GENETIC AND ECOLOGICAL DATA TO ESTIMATE SEA TURTLE ORIGINS

TOSHINORI OKUYAMA[1,3] AND BENJAMIN M. BOLKER[1,2]

[1]*Zoology Department, Box 118525, University of Florida, Gainesville, Florida 32611-8525 USA*
[2]*Archie Carr Center for Sea Turtle Research, University of Florida, Gainesville, Florida 32611-8525 USA*

Abstract. Many species of sea turtles spend part of their life cycle gathered in large feeding aggregations that combine individuals from widely separated rookery populations. Biologists have applied methods of mixed-stock analysis to mitochondrial DNA samples from rookeries and mixed populations to estimate the contributions of different rookeries to the mixed stock. These methods are limited by the amount of genetic overlap between rookeries and fail to incorporate ecological covariates such as rookery size and location within major ocean currents that are strongly suspected to affect rookery contributions. A new hierarchical Bayesian model for rookery contributions incorporates these covariates (and potentially others) to draw stronger conclusions from existing data. Applying the model to various simple scenarios shows that, in some cases, it can accurately estimate turtle origins even when turtles come from rookeries with high degrees of genetic overlap. Applying the model to more complex simulations shows that it performs well in a wide range of scenarios. Applying the model to existing data on green turtles (*Chelonia mydas*) narrows confidence intervals but does not change point estimates significantly. Applying it to loggerhead turtles (*Caretta caretta*) strengthens the dominance of the large rookery in south Florida, and brings estimates from a small data set on sea turtle strandings into line with those from rookery data. Used appropriately, hierarchical Bayesian methods offer great potential for introducing multiple levels of variation and ecological covariates into ecological models.

Key words: *Bayesian hierarchical model;* Caretta caretta; Chelonia mydas; *ecological covariate; genetic overlap; mixed-stock analysis; mtDNA haplotypes; rookery; sea turtles; spatial population structure.*

INTRODUCTION

Many species of sea turtles, including some species that nest and feed throughout the Atlantic, are threatened or endangered. These species spend part of their life cycle gathered in large feeding aggregations that combine individuals from widely separated rookery populations (Bowen et al. 1996, Bolten et al. 1998, Lahanas et al. 1998). Biologists have used mixed-stock analysis to estimate the contributions of different rookeries to the mixed stock, allowing them to assess either the importance of particular rookeries to the health of the mixed populations, or the impact of mortality in the mixed populations on the health of particular rookeries.

Mixed-stock analysis estimates the contributions of different source populations (rookeries) to a mixed population by comparing the distributions of genetic or phenotypic traits of individuals in the rookeries and in the mixed population (Fournier et al. 1984). For sea turtles, mtDNA haplotypes have established contributions to central mixed populations from rookeries throughout a surprisingly wide geographic catchment basin (Bowen et al. 1996, Bolten et al. 1998, Lahanas et al. 1998). These estimates have wide confidence intervals (Bolker et al. 2003). As sea turtle biologists and managers strive for more precise estimates of rookery contributions, they will need either to gather more data or to squeeze more information out of existing data. There are diminishing returns to gathering additional data of the same type; once one has enough data to characterize the genotype frequencies within rookeries and the mixed population accurately, the overlap of genetic information between rookeries becomes limiting. Current analytic methods are near their limit; computational Bayesian methods (Pella and Masuda 2001, Bolker et al. 2003) can incorporate prior information and improve the accuracy of confidence intervals, but do not improve point estimates (Bolker et al. 2003). One solution to this dilemma is to develop models that incorporate ecological covariates. Adding ecological covariates will extend the limits of mixed-stock analysis and will provide information on the ecological processes driving population dynamics.

This paper explores techniques to incorporate two simple ecological covariates, rookery size and location within ocean gyres, into mixed-stock analyses for sea turtles. A Bayesian hierarchical model (Gelman et al.

Manuscript received 27 February 2003; revised 31 March 2004; accepted 5 April 2004; final version received 24 May 2004. Corresponding Editor: A. B. Hollowed.
[3] E-mail: toshi@zoo.ufl.edu

1996, Sauer and Link 2002) is developed that accounts for variation at the level of sampling error, variation in the true contributions from each rookery, and variation in the relationship between size and expected contribution. This technique incorporates ecological covariates while maintaining the flexibility to accept that a strong genetic match really represents a larger-than-expected contribution from a small rookery. At one extreme, standard mixed-stock analysis methods, such as unconditional maximum likelihood, UML (Pella and Milner 1987), ignore ecological covariates. At the other extreme, one could construct a maximum likelihood regression model in which the parameters are the rookery genotype profiles (as in unconditional maximum likelihood) and the intercept and slope of the relationship between rookery size and rookery contribution. One could then estimate the parameters using the combined multinomial likelihood of the rookery samples (as in UML) and the mixed-stock sample based on the estimated size-based contributions. However, this procedure would classify any deviation from the best-fit line as error. To gain the benefits of ecological covariates without losing the ability to use the information given by apparent outliers, we use hierarchical models (Gelman et al. 1996). Hierarchical models fit into Bayesian frameworks and are also natural extensions of frequentist mixed-model approaches (Dennis 1996).

In this paper we construct and evaluate the performance of hierarchical models that address ecological processes. We apply the models first to ''data'' from simple simulations of genetic sampling, to assess their validity and power, and then to existing data from loggerhead (*Caretta caretta*) and green (*Chelonia mydas*) turtles. Finally, we discuss some of the lessons learned, questions opened, and broader issues raised by using more complex statistical models incorporating ecological covariates to understand patterns in ecological communities.

Methods

Hierarchical models allow us to incorporate relationships between rookery size and rookery contribution without assuming that any deviation from these relationships is solely due to sampling error. Rather than representing ecological relationships as fixed patterns with a single layer of error superimposed, hierarchical models represent variation at multiple scales. For example, a hierarchical regression model could allow variation of the slope and intercept parameters in different groups in addition to variation of individual points around the regression line. The roots of hierarchical modeling lie in Bayesian estimation: although frequentist approaches to hierarchical (or mixed) models exist (Clayton 1996, Pinheiro and Bates 2000), a broader and more powerful range of tools is available using a Bayesian framework.

Perhaps the most important aspect of the Bayesian perspective is that there are no fixed parameters: every parameter has a probability distribution. The goal is to use the data to estimate the probability distributions of the parameters, called the posterior distributions. We combine a prior distribution, which can be uninformative or weak (if we have little previous information or want to make few assumptions about the possible range of parameters), with a statistical model for the data that gives the likelihood of observing our data, given a particular set of parameters. In nonhierarchical (or non-covariate) Bayesian models, we specify only the prior distributions of the parameters themselves. Distributions are typically specified in terms of their parameters, called hyperparameters: for example, the means and variances of the slope and intercept would be the hyperparameters in a linear regression model. Hierarchical models redefine these constant hyperparameters as having their own distributions, which in turn have prior distributions that are defined by another level of hyperparameters. In the linear regression example just given, instead of specifying constant values for the prior means and variances of the slope and intercept, we would make these means and variances into distributions and specify their prior uncertainty. In principle, the top layer of the hierarchy can always be replaced by yet another set of hyperparameters with their own prior distributions.

The model

This section describes the general framework of the model; details are in Appendices A and B. The model extends the original model of Pella and Masuda (2001). The covariate model adds a hierarchical layer on top of the rookery contribution parameter; rookery contribution parameters are drawn from probability distributions based on rookery size. Fig. 1 shows the directed acyclic graph (DAG), which illustrates the structural dependencies of the model.

In a DAG, circles represent unknown parameters and squares represent sources of data: both kinds of elements are called nodes. A node with an outgoing arrow is called a ''parent'' and a node with an incoming arrow is called a ''descendant'' of that parent. The actual estimates of different unconnected parameters (such as the rookery haplotype frequencies and contributions) may be correlated once the model is estimated from data. Solid arrows in a DAG indicate probabilistic dependencies, where the value of a descendant node is drawn from a probability distribution determined by the parent node(s); dashed arrows indicate logical dependencies, where the descendant parameter has a fixed value that is simply calculated from the parent node(s). For example, ''Rookery Contribution Mean'' is determined by an algebraic function of three elements: ''Rookery Size'' and the two parameters ''Intercept'' and ''Slope'' (Fig. 1). When identifying parent–descendant relationships in the DAG, the deterministic

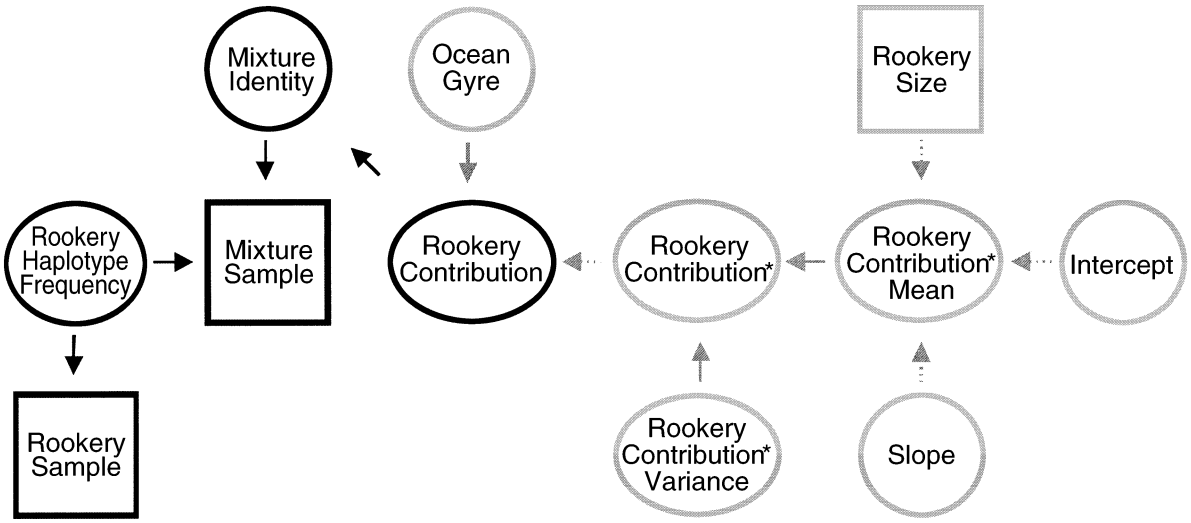

FIG. 1. Directed acylic graph (DAG) of a non-covariate mixed-stock analysis model. The non-covariate model (represented by the black nodes) is embedded within a covariate model. The covariate model adds the model components shown in gray and changes the distribution of rookery contributions from Dirichlet to additive logistic normal. ''Rookery Contribution'' in the gray nodes is additive log-ratio-transformed (indicated by ''⋆''; see *Methods: Additive log ratio transformation/logistic normal distribution*). Circles represent unknown parameters; squares represent data. Solid arrows indicate probabilistic dependencies; dashed arrows indicate logical dependencies.

dependencies (dashed arrows) are collapsed so that a node that depends on another through a series of calculations is counted as a descendant.

Previous non-covariate models (Pella and Masuda 2001) used a Dirichlet distribution as the prior distribution for the contribution parameters (black ''Rookery Contribution'' node in Fig. 1). In the covariate model, we use a more flexible logistic normal distribution (to be discussed) with mean parameters based on the rookery sizes (Fig. 1).

Additive log ratio transformation/logistic normal distribution

Analyzing compositional data or estimating compositional parameters such as rookery contributions presents particular technical challenges. Each rookery contribution must be between 0 and 1 (inclusive), and the contributions must sum to 1. The summation constraint is particularly challenging, as it induces a correlation between the parameters and limits one's choice of statistical models. Ordinary linear regression models neither limit the range of possible contributions nor constrain the sum to be 1; logistic regression models limit the range, but can still fail the summation constraint. One can fit constrained versions of these models, but brute-force attempts to fit constrained models often lead to artifacts.

The additive logistic normal distribution (Aitchison 1986, Billheimer et al. 2001) deals with these problems, at the cost of making our model slightly less intuitive. Suppose we pick $R - 1$ values $(y_1, \ldots, y_{R-1})$ from a multivariate normal distribution. We can transform these values to a set of R values that sum to 1 and that are each between 0 and 1 (i.e., that satisfy the constraints of compositional data) by saying:

$$p_i = \frac{\exp(y_i)}{1 + \sum_{j=1}^{R-1} \exp(y_j)} \qquad i = 1, \ldots, R - 1$$

$$p_R = 1 - \sum_{j=1}^{R-1} p_j. \tag{1}$$

These data are now distributed (by definition) according to the logistic normal distribution. The additive log ratio (alr) transformation of $\mathbf{p} = (p_1, \ldots, p_R)$ is

$$\text{alr}(\mathbf{p}) = \left[\log\left(\frac{p_1}{p_R}\right), \ldots, \log\left(\frac{p_{R-1}}{p_R}\right)\right]. \tag{2}$$

The alr is the inverse transformation that allows us to go from a set of compositional data to a (putatively) multivariate normal data set. If rookery size and rookery contribution data are logistic-normally distributed, a variety of sensible statistical models can be defined in terms of the multivariate normally distributed set of points that result from the additive log ratio transformation of the data. For example, we assume that the additive log-ratio-transformed contributions of different rookeries are independently normally distributed with means that depend linearly on the additive log-ratio-transformed rookery sizes. The additive log ratio transformation takes care of the necessarily negative correlation that occurs because all of the rookeries are sharing contributions to a single mixed population; beyond this correlation, we have assumed that the contributions are uncorrelated. Additional correlations could be incorporated by modeling the correlation

structure among the additive log-ratio-transformed rookery contributions (Billheimer et al. 2001).

Model evaluation

Performance criteria are needed to compare models. Mean squared error (MSE) is a common criterion for the performance of statistical methods, but for compositional data like ours, the Aitchison distance (Aitchison 1992) is superior. Like the MSE, the Aitchison distance accounts for both bias and variance; also like the MSE, a smaller Aitchison distance indicates superior model performance. Our results for both MSE and Aitchison distances were qualitatively similar. Models that incorporate added complexity can only perform better when they incorporate structure that is present in the data (and sometimes not even then; Ludwig and Walters [1985]). Comparing the covariate and the non-covariate model, we expect the non-covariate model to perform better (to have smaller Aitchison distance/MSE) when the covariate model incorporates poor data, e.g., if there is really no association between contribution and rookery size.

Another performance criterion of a model is the *coverage*, which indicates the accuracy of the estimated confidence limits. A region that includes the true value α% of the time in repeated samples is called the α% confidence region. Coverage can be estimated by running many simulations and evaluating the fraction of the time when the nominal α% confidence region includes the true value. If the coverage is less than α%, the confidence region is too narrow, and vice versa. We must distinguish between the univariate and multivariate coverage; the multivariate coverage describes the proportion of the time that the confidence intervals simultaneously include the true values for all n of the true parameters, whereas the univariate coverage only describes the proportion of time that confidence intervals for each individual parameter include the true values. Therefore, the multivariate coverage is always more conservative (lower) than the univariate coverage. For example, if the estimated rookery contributions in a five-rookery model are uncorrelated and the univariate coverage for each of the five rookeries is 95%, then the multivariate coverage will be only $(0.95)^5 \approx 77\%$.

When data are collected in the field, the true parameters are unknown. In this case, plausible fits to the data can be derived by testing summary statistics sampled from the posterior distribution (Gelman et al. 1996, Congdon 2001). Model selection is a more difficult problem, which we will discuss further.

Simulation procedures

Simulations were run to illustrate the strengths and weaknesses of the hierarchical covariate model. The relative performance of the covariate and non-covariate model was explored for a wide range of parameter values. Models were evaluated by comparing performance under different assumptions regarding the correlation strength between rookery size and rookery contribution. When the correlation strength is weak, the covariate model should not perform better than the simple non-covariate model. The other parameters will determine how strong a true correlation is necessary to justify using the covariate model, and how much better or worse the covariate model does for high and low correlations. Simulations were defined by specifying three suites of parameters: (1) true rookery contributions and rookery sizes, (2) true rookery haplotypes, and (3) the sizes of samples from the rookeries and from the mixed population.

Rookery sizes and rookery contributions.—To simulate rookery contributions and rookery sizes for R rookeries, we start by sampling $R - 1$ pairs of values from a bivariate normal distribution ($\{x_i,y_i\} \sim \mathcal{N}(\mathbf{0},\mathbf{M})$, $i = 1, \ldots, R - 1$), where $\mathbf{M}$ specifies the variances of each variable and the correlation between them (i.e., the variance–covariance matrix). Using the inverse of the additive log ratio transform (Eq. 1) leads to a bivariate distribution for the relative rookery sizes and rookery contributions. The maximum possible variance of a set of R non-negative values that add to 1 is $(1 - 1/R)/R$. Numerical simulations show that below this maximum there is a fairly tight one-to-one relationship between the variance of the original normally distributed variable and the variance of the transformed data. Two different normal variances were used in the simulations. For the ''high'' variance case, the bivariate normal variance σ^2 is set to 2.5, corresponding to realized variances (s_c^2 (in rookery contribution) or s_r^2 (in rookery size)) of 0.068 $\pm$ 0.004 (mean $\pm$ 1 SD for five rookeries) and 0.025 $\pm$ 0.001 (for 10 rookeries). For the ''low'' variance case, $\sigma^2 = 1.5$, corresponding to realized variances (s_c^2 or s_r^2) of 0.044 $\pm$ 0.002 (for five rookeries) and 0.016 $\pm$ 0.001 (for 10 rookeries). The correlation observed in simulations with different bivariate normal variances and correlations varies widely. Therefore simulations were conducted over a range of bivariate normal correlations (ρ) between 0 and 1. Subsequently, each simulated data set was assigned a ρ value on the basis of its observed correlation.

The statistical distribution underlying the simulations does not match the statistical model underlying the covariate model (logistic normality of contributions with a linear dependence on rookery size). One advantage of this slight mismatch is that it helps to ensure that the covariate models are robust to misspecification of the underlying relationships between rookery size and rookery contribution.

Haplotype frequency profiles.—An R (number of rookeries) by H (number of haplotypes) matrix represents the entire distribution of haplotype frequencies within rookeries (Table 1).

Each rookery has a characteristic haplotype (the diagonal elements in the matrix) that is present at high frequency, classified as common (c). In each rookery,

TABLE 1. Rookery haplotype frequency structure: *c*, common haplotype; *i*, intermediate haplotype; *r*, rare haplotype.

Rookery	Haplotype				
	H1	H2	H3	H4	H5
R1	*c*	*i*	*r*	*r*	*i*
R2	*i*	*c*	*i*	*r*	*r*
R3	*r*	*i*	*c*	*i*	*r*
R4	*r*	*r*	*i*	*c*	*i*
R5	*i*	*r*	*r*	*i*	*c*

the haplotypes on either side of the characteristic haplotype are present at intermediate frequencies (*i*). The remaining haplotypes are rare (*r*). The genetic divergence among the rookeries is parameterized by the ratio of the common and intermediate haplotype frequencies (c/i): the value $c/i = 2$ was used for overlapping genetic profiles, and $c/i = 6$ for distinct genetic profiles. Specifying the frequency of rare haplotypes, and applying the constraint that the sum of the haplotype frequencies equals 1, completes the parameterization. For example, for five rookeries ($R = 5$) with haplotype parameters $r = 0.01$ and $c/i = 2$, we can solve $c + 2(c/2) + 2(0.01) = 1$ to find the appropriate values of $c = 0.49$ and $i = 0.245$. In general, $c + 2c/(c/i) + r(R - 3) = 1$. For convenience, the number of rookeries was set equal to the number of haplotypes ($H = R$). In some of the specific scenarios that we will present, we impose a further block structure on the haplotype frequency matrix that mimics the case of two different gyres (large, circular surface ocean currents) with similar haplotype profiles within gyres and dissimilar profiles between gyres.

The combination of the haplotype frequencies and true contributions determines the haplotype composition of the mixed population. The frequency of the *j*th haplotype in the mixture population, q'_j, is

$$q'_j = \sum_{i=1}^{R} p_i q_{ij} \qquad j = 1, \ldots, H$$

where p_i is the contribution from the *i*th rookery and q_{ij} is the frequency of the *j*th haplotype in the *i*th rookery.

Sample sizes of the mixed population and rookery populations.—To complete the simulated samples, multinomial samples were drawn from the rookeries and from mixed populations. Baseline sample sizes were 25 for each rookery and 50 for the mixed population. The samples were distributed among haplotypes according to the multinomial distribution with the frequencies determined previously.

RESULTS

Scenario analysis

Three scenarios illustrate the properties and capabilities of covariate modeling. These scenarios are intentionally simplified. The first scenario shows the basic ability of the covariate model to detect structure when it is present. The second scenario distinguishes between size–contribution relationships among rookeries in different ocean gyres. The third scenario shows a slightly more realistic case involving a block of three different rookeries with overlapping mtDNA profiles.

Scenario 1: an extreme case.—This first scenario presents a classical problem for mixed-stock analysis: What happens when different combinations of source stocks can combine to produce precisely the same profile in the mixed stock (Table 2)? Rookeries R1 and R2 are each composed entirely of a distinctive haplotype; if they were the only source stocks known, mixed-stock analysis would be trivially easy. When a small rookery with haplotypes from both large rookeries is added (R3), separating contributions from the small and large rookeries becomes impossible. Without more information, even our covariate approach is powerless. Add a small rookery with a distinct haplotype (R4/H3), however, and we can estimate the relationship between rookery size and rookery contribution, which in turn gives us a better estimate for the contribution of R3. (We could also improve our estimate by assuming prior knowledge of the relationship between rookery size and rookery contribution, but we prefer to be as conservative as possible by using weak or uninformative priors.)

In this case, the true rookery contributions exactly equal the relative rookery sizes, and so a purely size-

TABLE 2. Scenario 1: confounding between large and small rookeries with three distinct haplotypes. R1 and R2 have population sizes 19 times larger than R3 and R4.

Rookery	Haplotype			True contribution	Noncovariate	Covariate
	H1	H2	H3			
R1	100	0	0	47.5	21.3 (†, 56.3)	39.6 (3.6, 59.1)
R2	0	100	0	47.5	20.3 (0.14, 55.5)	40.7 (3.4, 60.2)
R3	50	50	0	2.5	55.6 (†, 98.1)	16.3 (0.1, 86.7)
R4	0	0	100	2.5	2.8 (†, 8.4)	3.3 (0.3, 9.4)

Notes: All numbers are expressed as percentages. Rookery contribution estimates are shown as posterior mean (2.5 and 97.5 percentiles). Sample sizes: 25 (rookeries) and 50 (mixed population). Dagger symbols (†) indicate values smaller than 0.01%. Relative sizes of the rookeries are equal to the true contributions.

TABLE 3. Scenario 2: estimates for a scenario with eight rookeries in two gyres and eight haplotypes, where genetic overlap $c/i = 2$ and rare-haplotype frequency $r = 0.01$.

Gyre and rookery	Haplotype								True contribution	Non-covariate	Covariate
	H1	H2	H3	H4	H5	H6	H7	H8			
G1											
R1	*c*	*i*	*i*	*i*	*r*	*r*	*r*	*r*	60.4	38.96 (0.80, 96.92)	47.01 (4.77, 88.71)
R2	*i*	*c*	*i*	*i*	*r*	*r*	*r*	*r*	24.02	19.72 (†, 84.43)	23.85 (0.40, 68.51)
R3	*i*	*i*	*c*	*i*	*r*	*r*	*r*	*r*	9.61	18.99 (†, 82.13)	14.76 (0.10, 55.73)
R4	*i*	*i*	*i*	*c*	*r*	*r*	*r*	*r*	3.84	18.99 (†, 80.15)	9.58 (0.06, 41.51)
G2											
R5	*r*	*r*	*r*	*r*	*c*	*i*	*i*	*i*	1.54	0.92 (†, 7.15)	2.17 (0.01, 9.00)
R6	*r*	*r*	*r*	*r*	*i*	*c*	*i*	*i*	0.61	0.86 (†, 6.87)	1.29 (†, 6.46)
R7	*r*	*r*	*r*	*r*	*i*	*i*	*c*	*i*	0.25	0.84 (†, 6.89)	0.83 (†, 4.94)
R8	*r*	*r*	*r*	*r*	*i*	*i*	*i*	*c*	0.10	0.71 (†, 5.95)	0.51 (†, 3.31)

Notes: All values are expressed as percentages. Rookery contribution estimates are shown as posterior mean (2.5 and 97.5 percentiles). Dagger symbols (†) indicate values smaller than 0.01%. Relative sizes of the rookeries are equal to the true contributions within each gyre (e.g., 6004, 2402, 961, 384 for R1–R4 [gyre 1] and 15 400, 6100, 2500, 1000 for R5–R8 [gyre 2]; combined sizes of rookeries within a gyre do not enter the model).

based prediction would provide better point estimates than the covariate model. However, the covariate model improves the point estimates, and narrows the confidence intervals to the point where all four rookeries are contributing at least some turtles to the mixed population. The covariate model's lack of built-in assumptions about the strength of the size–contribution relationship allows it to fit well over a broad range of true scenarios ranging from perfect correlation between rookery size and rookery contribution to no correlation.

Scenario 2: incorporating ocean gyres.—Assuming contributions based on rookery size alone can fail if rookeries in different ocean gyres contribute very different numbers of turtles, or if the relationship between rookery size and rookery contribution differs in the different gyres. Adding a factor to the linear model that allows the slope and intercept to vary by ocean current mitigates this problem (as described in *Methods*).

Table 3 shows a test with rookeries in two separate ocean currents, where the rookeries in the first current contribute much more to the mixed population under analysis. A model incorporating rookery size alone would fail because contribution depends on both rookery size and currents. When ocean currents are included in the model, it successfully estimates the relative contributions from the rookeries in the first ocean current, and correctly estimates that the rookeries in the second ocean current contribute little to the mixed population. Use of a non-covariate model correctly estimates the low contribution from rookeries in the second current, but poorly estimates the relative contributions in the first current, at least in part, because this scenario is set up with a great deal of overlap in mtDNA profiles between rookeries ($c/i = 2$).

Scenario 3: an unresolved region.—Finally, consider a slightly more realistic situation with eight rookeries within a single ocean current, but where one block of three rookeries (R2–R4 in Table 4) is very poorly resolved by genetic information ($c/i = 1.1$). As in the previous scenarios, the covariate model can resolve the differences in contribution between these five rookeries. The necessary information exists because the mtDNA profiles are different enough between this block and the other rookeries (and among the other rookeries) to provide information on the relationship between rookery size and rookery contribution.

Exploring parameter space

Each of the scenarios just presented illustrates a particular feature of covariate models rather than giving a general overview of their performance. All three are also best case scenarios for covariate modeling: there are ambiguities in the genetic data, but there is always useful genetic information present in some part of the data set, and the true relationships between rookery size and rookery contribution are strong. Some features of our data sets are known (sample size and the distribution of rookery sizes), but others are unknown (e.g., the correlation between rookery size and rookery contributions). By exploring the relative performance of the non-covariate and covariate under a range of conditions, we can decide whether (and when) the covariate model is preferred.

For each combination of other parameters (variance, sample size, etc.), the model is tested across the entire range of positive correlations between rookery sizes and contributions. The non-covariate model is tested only once, because it does not incorporate rookery size and thus would give the same answer, regardless of the correlation. Estimates of Aitchison distance and coverage were derived from estimates of rookery contribution resulting from 100 different simulations for each set of parameters. The general pattern is the same across all parameters. For high correlations, the covariate model outperforms the non-covariate model (has lower Aitchison distance). The performance of the covariate model decreases approximately linearly with decreasing correlations, until at low correlations its

TABLE 4. Scenario 3: unresolved region, with rookery contribution estimates shown as posterior mean (2.5 and 97.5 percentiles), where $c/i = 1.1$ for rookeries R2–R4, $c/i = 6$ for rookeries R5–R8, and $r = 0.01$.

Rookery	Haplotype								True contribution	Noncovariate	Covariate
	H1	H2	H3	H4	H5	H6	H7	H8			
R1	*c*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	60.4	60.38 (44.44, 77.74)	60.62 (45.30, 75.94)
R2	*r*	*c*	*i*	*i*	*r*	*r*	*r*	*r*	24.02	11.92 (†, 43.99)	20.34 (1.85, 40.88)
R3	*r*	*i*	*c*	*i*	*r*	*r*	*r*	*r*	9.61	12.04 (†, 43.63)	10.35 (0.39, 31.48)
R4	*r*	*i*	*i*	*c*	*r*	*r*	*r*	*r*	3.84	11.68 (†, 44.15)	5.31 (0.10, 23.31)
R5	*r*	*r*	*r*	*r*	*c*	*i*	*i*	*i*	1.54	1.30 (†, 7.03)	1.65 (0.04, 6.17)
R6	*r*	*r*	*r*	*r*	*i*	*c*	*i*	*i*	0.61	0.86 (†, 5.88)	0.89 (0.01, 4.14)
R7	*r*	*r*	*r*	*r*	*i*	*i*	*c*	*i*	0.25	0.84 (†, 5.75)	0.52 (†, 2.88)
R8	*r*	*r*	*r*	*r*	*i*	*i*	*i*	*c*	0.10	0.97 (†, 6.23)	0.31 (†, 1.97)

Notes: Notation is as in Tables 2 and 3. Relative sizes of the rookeries are equal to the true contribution.

Aitchison distance rises above that of the non-covariate model (Fig. 2).

The maximum improvement in performance and the rate of performance loss with decreasing correlation vary with the other parameters. In general, when there is little information present in the genetic markers (e.g., small c/i, small sample sizes, etc.), the covariate model performs better than the non-covariate model, provided the ecological covariates give any true information at all. Overall, covariate models produce more accurate estimates of rookery contributions. They are also reasonably robust to variations in the underlying contribution structure: even in the limit of very low correlation, when the non-covariate model performs better, the covariate model is not much worse. Multivariate coverages of the non-covariate model are much lower than the nominal value of 0.95, whereas appropriate coverage is achieved with the covariate model with informative covariates (high correlation between rookery size and rookery contribution).

As illustrated in Table 5, the parameters of a stock mixture (the numbers of rookeries, variance in contributions and rookery sizes, and so forth) fall into two categories, those that can be known or guessed directly from ecological and genetic data (''known'') and those that can only be determined by mixed-stock analysis (''unknown''). Some of the parameters, such as sample size or the number of rookeries considered, are well known and are determined by the design of sampling programs. The variance in rookery size (calculated from the proportional or normalized rookery sizes, $n_i/\Sigma_j n_j$ where n_j indicates the jth rookery size) is known, albeit not very accurately, from observations of the number of nests at the rookeries. The genetic distinction (c/i) and frequency of rare haplotypes are not known. Bayesian and unconditional maximum likelihood methods are used to infer the genetic composition of a rookery using the genetic composition of the mixed population (Pella and Milner 1987). Nevertheless, in the spirit of a power analysis, we can estimate approximately where a particular metapopulation of turtles lies in parameter space, and whether covariate models are likely to be appropriate.

Estimating the ''unknown'' parameters is problematic. These quantities, which determine the expected performance of the covariate model, can themselves only be determined by mixed-stock analysis. There is a fundamental circularity involved in trying to estimate true values of these relationships; how do we know which assumptions we should make in trying to estimate the appropriateness of the assumptions themselves?

The solution to this problem is to develop methods of model selection that are informative for a given set of data, which is likely to produce better estimates, but these are not simple problems (Key et al. 1999, Spiegelhalter et al. 2002). We are working toward implementing and evaluating such methods in the context of stock analysis. A partial solution is to check whether the posterior distribution of the rookery size–contribution correlation is clearly bounded away from zero, which would suggest that there is a strong dependence of contribution and size and, thus, that the more complex model is appropriate. Even if rookery size and rookery contributions are correlated, the non-covariate model will tend to underestimate the correlation and overestimate its own relative performance. (Using the covariate model to estimate correlation would tend to overestimate correlations and would definitely risk circular logic.)

Results from data

What conclusions do the new methods reach about rookery contributions for Atlantic green (*Chelonia mydas*) and loggerhead (*Caretta caretta*) turtles from existing data (Bolten et al. 1998, Lahanas et al. 1998)? How different are the results, either in terms of point estimates or confidence intervals, from previous results from non-covariate models? (For details of the statistical models, see Appendix A.)

Posterior predictive checks, calculating the variance in mixed population sample for random draws from the posterior distribution, showed that this summary statistic consistently fell between the data and the posterior distribution. This result indicated that the model fit the data adequately. The results for green turtle or-

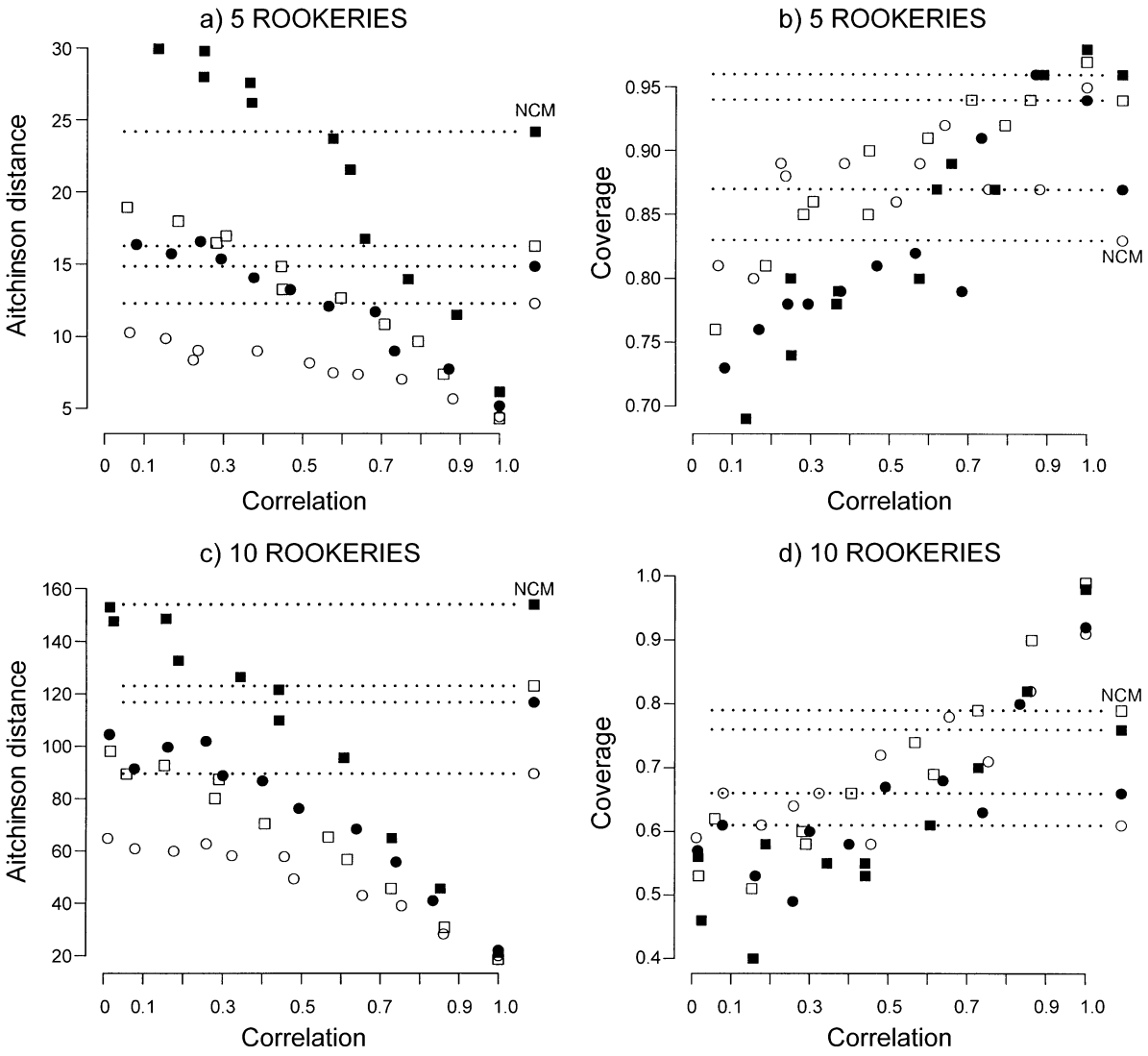

FIG. 2. Performance of covariate and non-covariate models across parameter space: open circles are low variance, low overlap ($c/i = 2$); open squares are low variance, high overlap ($c/i = 6$); solid circles are high variance, low overlap ($c/i = 2$); solid squares are high variance, high overlap ($c/i = 6$). Non-covariate model (NCM) results are shown at the right of each figure, along with dotted lines across all values of correlation. Plots show the Aitchison distance [panels (a) and (c)] or multivariate coverage [panels (b) and (d)] as a function of the simulated correlation between rookery size and rookery contribution; the covariate model performs better for high correlations.

TABLE 5. Simulation parameters and qualitative results.

Parameter	Symbol	Values	Covariate model preferred if . . .
Known parameters			
Rookery size variance	σ_r^2	high, low	low
Genetic distinction	c/i	2, 6	low
Number of rookeries	R	5, 10	high
Unknown parameters			
Rookery contribution variance	σ_c^2	high, low	low
contribution-size correlation	ρ	0.0–1.0	high

Notes: Known parameters are those that could be estimated at least approximately from basic genetic sampling and ecological knowledge; unknown parameters are those that can only be revealed by mixed-stock analysis.

TABLE 6. Contribution estimates (percentages) for green turtle data, with rookery contribution estimates shown as posterior mean (2.5 and 97.5 percentiles).

Rookery	Size (no. nests/yr)	Noncovariate	Covariate
FL	2300	3.86 (†, 15.73)	4.2 (0.3, 13.4)
MEXI	900	1.16 (†, 7.73)	1.5 (†, 6.1)
CR	93 500	78.95 (65.26, 89.44)	79.2 (66.2, 89.2)
AVES	2300	11.07 (†, 23.77)	6.4 (0.17, 18.6)
SURI	5300	3.44 (†, 17.98)	7.1 (0.1, 17.1)
BRAZ	900	0.35 (†, 2.94)	0.17 (†, 1.3)
ASCE	11 100	0.45 (†, 3.31)	0.67 (†, 3.4)
AFRI	7400	0.54 (†, 3.47)	0.56 (†, 2.9)
CYPR	400	0.17 (†, 1.64)	0.1 (†, 0.7)

Notes: Dagger symbols (†) indicate values smaller than 0.01%. Rookery abbreviations: FL, Hutchinson Island, Florida, USA; MEXI, Yucatán, Mexico; CR, Tortuguero, Costa Rica; AVES, Aves Island, Venezuela; SURI, Matapica, Suriname; BRAZ, Atol das Rocas, Brazil; ASCE, Ascension Island, UK; AFRI, Pailoa, Guinea Bissau; CYPR, Lara Bay, Cyprus. Gyre 1 contains the first five rookeries and the mixed stock, whereas gyre 2 contains the rest of the rookeries. The column ''Size'' reports average number of nests deposited each year.

igins are not radically different from those determined by non-covariate models (Table 6; Lahanas et al. 1998, Bolker et al. 2003). The major difference is that the covariate model predicts slightly higher contributions from Suriname than from Aves Island, whereas the non-covariate model predicts the opposite. (However, the confidence intervals are fairly wide in both cases, supporting either negligible contributions from either rookery or equal contributions from both rookeries.)

Table 7 shows the mixed-stock analysis results for loggerhead turtles. The main conclusion of the covariate model is to strengthen the dominance of the large (64 000 nests) south Florida rookery relative to the northwest Florida rookery (450 nests). The covariate model produced narrower confidence intervals. In particular, the covariate model allowed us to prove that at least 27% of the turtles in the mixed stock come from the south Florida rookery, whereas the non-covariate model was only able to bound the south Florida contribution between 0% and 95%.

How do these data relate to the previous simulations investigated? The variances of the rookery size within the main gyre (scaled to sum to one) are 0.179 for loggerhead turtles and 0.156 for green turtles. To compare genetics in real data sets with simple common/intermediate/rare structures from our simulations, we used the divergence statistics of Xu et al. (1994), which can be applied to more general haplotype frequency patterns as well as to our simulations. Divergence statistics based on observations in the main gyre were 0.78 (loggerhead turtles) and 1.52 (green turtles). Divergence statistics for simulated data were 0.802 (five rookeries, $c/i = 2$), 1.852 (five rookeries, $c/i = 6$), 2.250 (10 rookeries, $c/i = 2$), and 4.142 (10 rookeries, $c/i = 6$). The average rookery sample sizes are 41.5 ± 34.4 (mean $\pm$ 1 SD) for loggerhead turtles and 21.56 ± 9.34 for green turtles. The number of rookeries within the main gyre (which contains the mixed stock) is four for loggerhead turtles and five for green turtles. The data fall closest to the five-rookery case with $c/i = 2$ (loggerhead turtles) and $c/i = 6$ (green turtles). Fig. 2 (panels a and b; filled squares and circles) shows that in this case the covariate model often gives better and never gives much worse results than the non-covariate model. For example, when $c/i = 2$ with five rookeries, the covariate model has Aitchison distance ranging from 5 to 30 as correlation ranges from 1 to 0, whereas the non-covariate model results in a Aitchison distance of 24. The multivariate coverage of the covariate model ranges from 0.98 to 0.69, while the non-covariate model coverage is 0.96. Furthermore, the point estimates for the non-covariate model give correlations between

TABLE 7. Contribution estimates (percentages) for loggerhead turtle data, with rookery contribution estimates shown as posterior mean (2.5 and 97.5 percentiles).

Rookery	Size (no. nests/yr)	Noncovariate	Covariate
NWFL	450	17.14 (†, 64.0)	11.3 (1.1, 32.1)
SOFL	64 000	52.34 (†, 95.3)	64.2 (26.8, 91.8)
NEFL–NC	6200	10.53 (†, 48.85)	14.4 (0.8, 34.0)
MEXI	1800	16.09 (2.13, 48.17)	9.2 (0.2, 29.8)
GREECE	3000	3.76 (†, 29.24)	0.6 (†, 5.4)
BRAZ	4000	0.14 (†, 11.22)	0.2 (†, 1.4)

Notes: Dagger symbols (†) indicate values smaller than 0.01%. Rookery abbreviations: NWFL, northwest Florida; SOFL, south Florida; NEFL–NC, northeast Florida to North Carolina. MEXI, Quintana Roo, Mexico; GREECE, Kiparissia Bay, Peloponnesus Island, Greece; BRAZ, Bahia, Brazil. Gyre 1 contains the first four rookeries and the mixed stock; for details of gyre structure, see Appendix A.

contribution and size of 0.98 (for green turtles) and 0.92 (for loggerheads), and the posterior distribution of the slope parameter is bounded away from zero. Thus, our best guess is that the data for Atlantic green and loggerhead turtle populations are indeed appropriate for covariate modeling.

Conclusions

The covariate models presented here used information on ecological covariates to evaluate hypotheses regarding sea turtle origins. Results were generally consistent with previously published theories of the genetic origins of Atlantic sea turtle populations. For green turtles, the covariate models provide little additional information: their main effect is to lower the estimates and upper confidence limits of contributions from rookeries in Brazil, Africa, and Cyprus slightly, strengthening the conclusion that these rookeries contribute negligibly to the Atlantic mixed stock (Table 6). For loggerhead turtles, the covariate models make more substantive contributions. They strengthen the dominant contribution of the large south Florida rookery, relative to the contributions from northwest Florida and more northern rookeries, and establish that rookeries in Greece and Brazil are not contributing significantly (upper confidence limits of 5.4% and 1.4%, respectively, down from 29% and 11% in the non-covariate model).

The proposed modeling framework provides a technique for incorporating a wide range of ecological covariates in statistical models of sea turtle genetics. Rookery size and ocean currents are just the beginning; using the logistic-normal framework allows us to incorporate a wide range of ecological covariates in our statistical models. For example, more detailed geographic information could be incorporated in the model using conditional autoregressive models, which use geographic distances between rookeries to structure the variance–covariance matrix of transformed contributions (Billheimer et al. 2001). A range of options, from a simple neighborhood matrix (rookery contributions are correlated with those of their neighbors) to more sophisticated distance measures based on oceanographic data, are possible.

Before charging ahead with ever more complex models with ever larger sets of covariates, however, we must develop and test the methods to decide whether the added model complexity is appropriate. An inherent danger associated with complex models is that modelers may develop unrealistic or overly complex structural assumptions. Simulation models such as those presented here provide one tool to assess whether the added complexity is appropriate. However, even when we test performance across a broad range of simulation parameters, we can never be sure that our simulations are representative of reality; there can always be unrealistic structural assumptions built into the model, or axes of variation that we have not explored. In the frequentist estimation framework, model fit can be tested against data in the absence of a specific simulation model (although some null hypothesis is always necessary). For example, parameters can be tested for significant difference from null hypothesis values, and models of appropriate complexity can be selected, using either likelihood ratio tests or information criteria (Hilborn and Mangel 1997, Burnham and Anderson 2002). The Bayesian hierarchical framework does not provide a clear prescription for limiting model complexity. Methods for Bayesian model testing and selection do exist, including Bayes factors, Bayesian information criteria, posterior likelihoods, and posterior predictive densities (Carlin and Louis 1996), but these have rarely been applied to complex hierarchical models, and we have never seen them applied in ecological contexts.

In the long run, the way to build better ecological models is not just to include more plausible covariates, but rather to incorporate structures that reflect the mechanistic processes underlying observed patterns. Our goal is to build into the models any mechanistic processes that we are sure of, but to remain agnostic about other aspects of the ecological system. The models considered here all retain the basic structure that the mixed-stock haplotype profile is a composition of the haplotype profiles of the contributing rookeries, in contrast, for example, to cluster analyses that simply assess the similarity of different stocks without assuming specific underlying mechanisms (Phelps et al. 1994). Our models are semi-mechanistic (Ellner et al. 1998) because they incorporate ecological process but allow flexibility in the relationship between ecological covariates (size, location, etc.) and contribution. Semi-mechanistic models have two advantages. First, including well-established mechanisms in the basic definition of the model increases the overall amount of information in the model, leveraging the observational data for more accurate estimates. Second, when models are semi-mechanistic, the estimates of parameters themselves give information about the strengths of different mechanistic processes. For example, by including rookery sizes in the model, we both improve the estimates of individual rookery contributions and gain some insight into the relationship between rookery size and rookery contribution. Ultimately, we should think of building semi-mechanistic models at both short (proximate) and long (ultimate) time scales. Short-term processes that determine rookery contributions are the behavioral and energetic rules by which turtles determine their movements, and the ocean currents that form the context for movement. Long-term processes include rookery establishment, the behavioral and stochastic factors determining homing to and straying from natal beaches by females, and population genetic processes such as drift and mutation that change the genetic makeup of a rookery. Some frameworks already exist for estimating some of these processes (such as spatial

coalescent theory for the processes of migration, mutation, and drift: Beerli and Felsenstein [2001]), but much of it will have to be constructed and integrated as we proceed toward the goal of a fully integrated model that describes the ecology and evolution of sea turtle populations.

ACKNOWLEDGMENTS

We thank Karen Bjorndal and Alan Bolten (Archie Carr Center for Sea Turtle Research) for useful discussions and for providing the rookery size data for loggerhead turtles and green turtles, and George Casella for statistical advice, the authors of the R and BUGS packages for powerful tools, and three anonymous reviewers for useful comments. This project was funded by Cooperative Agreement NA17RJ1230 between the Joint Institute for Marine and Atmospheric Research (JIMAR) and the National Oceanic and Atmospheric Administration (NOAA). The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subdivisions.

LITERATURE CITED

Aitchison, J. 1986. The statistical analysis of compositional data. Chapman and Hall, New York, New York, USA.

Aitchison, J. 1992. On criteria of measures of compositional difference. Mathematical Geology **24**:365–379.

Beerli, P., and J. Felsenstein. 2001. Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. Proceedings of the National Academy of Sciences (USA) **98**:4563–4568.

Billheimer, D. P., P. Guttorp, and W. F. Fagan. 2001. Statistical interpretation of species composition. Journal of the American Statistical Association **96**:1205–1214.

Bolker, B. M., T. Okuyama, K. A. Bjorndal, and A. B. Bolten. 2003. Stock estimation for sea turtle populations using genetic markers: accounting for sampling error for rare genotypes. Ecological Applications **13**:763–775.

Bolten, A. B., K. A. Bjorndal, H. R. Martins, T. Dellinger, M. J. Biscotio, S. E. Encalada, and B. W. Bowen. 1998. Transatlantic developmental migrations of loggerhead sea turtles demonstrated by mtDNA sequence analysis. Ecological Applications **8**:1–7.

Bowen, B. W., A. L. Bass, A. Garcia-Rodriguez, C. E. Diez, R. van Dam, A. B. Bolten, K. A. Bjorndal, M. M. Miyamoto, and R. J. Ferl. 1996. Origin of hawksbill turtles in a Caribbean feeding area as indicated by genetic markers. Ecological Applications **6**:566–572.

Burnham, K. P., and D. Anderson. 2002. Model selection and multi-model inference. Springer-Verlag, New York, New York, USA.

Carlin, B. P., and T. A. Louis. 1996. Bayes and empirical Bayes methods for data analysis. Chapman and Hall, New York, New York, USA.

Clayton, D. G. 1996. Generalized linear mixed models. Pages 275–301 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov Chain Monte Carlo in practice. Chapman and Hall, London, UK.

Congdon, P. 2001. Bayesian statistical modeling. John Wiley, New York, New York, USA.

Dennis, B. 1996. Discussion: should ecologists become Bayesians? Ecological Applications **6**:1095–1103.

Ellner, S. P., B. A. Bailey, G. V. Bobashev, A. R. Gallant, B. T. Grenfell, and D. W. Nychka. 1998. Noise and nonlinearity in measles epidemics: combining mechanistic and statistical approaches to population modeling. American Naturalist **151**:425–440.

Fournier, D. A., T. D. Beacham, B. E. Riddell, and C. A. Busack. 1984. Estimating stock composition in mixed stock fisheries using morphometric, meristic, and electrophoretic characteristics. Canadian Journal of Fisheries and Aquatic Sciences **41**:400–408.

Gelman, A., J. Carlin, H. S. Stern, and D. B. Rubin. 1996. Bayesian data analysis. Chapman and Hall, New York, New York, USA.

Hilborn, R., and M. Mangel. 1997. The ecological detective: confronting models with data. Princeton University Press, Princeton, New Jersey, USA.

Key, J. T., L. R. Pericchi, and A. F. M. Smith. 1999. Bayesian model choice: what and why? Pages 343–370 *in* J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, editors. Bayesian statistics 6. Oxford University Press, Oxford, UK.

Lahanas, P. N., K. A. Bjorndal, A. B. Bolten, S. E. Encalada, M. M. Miyamoto, R. A. Valverde, and B. W. Bowen. 1998. Genetic composition of a green turtle (*Chelonia mydas*) feeding ground population: evidence for multiple origins. Marine Biology **130**:345–352.

Ludwig, D., and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? Canadian Journal of Fisheries and Aquatic Sciences **42**:1066–1072.

Pella, J., and M. Masuda. 2001. Bayesian methods for analysis of stock mixtures from genetic characters. Fisheries Bulletin **99**:151–167.

Pella, J. J., and G. B. Milner. 1987. Use of genetic markers in stock composition analysis. Pages 247–276 *in* N. Ryman and F. W. Utter, editors. Population genetics and fisheries management. University of Washington Press, Seattle, Washington, USA.

Phelps, S. R., L. L. Leclair, S. Young, and H. L. Blankenship. 1994. Genetic diversity of patterns of chum salmon in the Pacific-northwest. Canadian Journal of Fisheries and Aquatic Sciences **51**(Supplement 1):65–83.

Pinheiro, J. C., and D. M. Bates. 2000. Mixed effects models in S and S-Plus. Springer Verlag, New York, New York, USA.

Sauer, J. R., and W. A. Link. 2002. Hierarchical modeling of population stability and species group attributes from survey data. Ecology **83**:1743–1751.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society B **64**:583–639.

Xu, S., C. J. Kobak, and P. E. Smouse. 1994. Constrained least squares estimation of mixed population stock composition from mtDNA haplotype frequency data. Canadian Journal of Fisheries and Aquatic Sciences **51**:417–425.

APPENDIX A

The model specification and Markov Chain Monte Carlo are available in ESA's Electronic Data Archive: *Ecological Archives* A015-009-A1.

APPENDIX B

BUGS code is available in ESA's Electronic Data Archive: *Ecological Archives* A015-009-A2.