



Where were they from? Modelling the source stock of dolphins stranded after the *Deepwater Horizon* oil spill using genetic and stable isotope data

L. Thomas^{1,*}, C. G. Booth², P. E. Rosel³, A. Hohn⁴, J. Litz⁵, L. H. Schwacke⁶

¹Centre for Research into Ecological and Environmental Modelling (CREEM), University of St Andrews, The Observatory, Buchanan Gardens, St Andrews KY16 9LZ, UK

²SMRU Consulting Europe, New Technology Centre, North Haugh, St Andrews KY16 9SR, UK

³National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, 646 Cajundome Blvd., Lafayette, LA 70506, USA

⁴National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, 101 Pivers Island Rd, Beaufort, NC 28516, USA

⁵National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, 75 Virginia Beach Drive, Miami, FL 33149, USA

⁶National Oceanic and Atmospheric Administration, National Centers for Coastal Ocean Science, 331 Fort Johnson Road, Charleston, SC 29412, USA

ABSTRACT: Understanding the source stock of common bottlenose dolphins *Tursiops truncatus* that stranded in the northern Gulf of Mexico subsequent to the *Deepwater Horizon* oil spill was essential to accurately quantify injury and apportion individuals to the appropriate stock. The aim of this study, part of the Natural Resource Damage Assessment (NRDA), was to estimate the proportion of the 932 recorded strandings between May 2010 and June 2014 that came from coastal versus bay, sound and estuary (BSE) stocks. Four sources of relevant information were available on overlapping subsets totaling 336 (39%) of the strandings: genetic stock assignment, stable isotope ratios, photo-ID and individual genetic-ID. We developed a hierarchical Bayesian model for combining these sources that weighted each data source for each stranding according to a measure of estimated precision: the effective sample size (ESS). The photo- and genetic-ID data were limited and considered to potentially introduce biases, so these data sources were excluded from analyses used in the NRDA. Estimates were calculated separately in 3 regions: East (of the Mississippi outflow), West (of the Mississippi outflow through Vermilion Bay, Louisiana) and Western Louisiana (west of Vermilion Bay to the Texas–Louisiana border); the estimated proportions of coastal strandings were, respectively 0.215 (95% CI: 0.169–0.263), 0.016 (0.036–0.099) and 0.622 (0.487–0.803). This method represents a general approach for integrating multiple sources of information that have differing uncertainties.

KEY WORDS: Genetic assignment · Stable isotope ratios · Bayesian hierarchical model · Data integration · *Deepwater Horizon* · Stranding

INTRODUCTION

The explosion on the *Deepwater Horizon* (DWH) drilling platform on 20 April 2010 resulted in the release of millions of barrels of oil into the northern

Gulf of Mexico (nGoM). Following its release, the oil spread to coastal regions of the nGoM between Florida and Louisiana (Michel et al. 2013), areas inhabited by the western and northern coastal stocks, and a number of bay, sound and estuary (BSE) stocks

*Corresponding author: len.thomas@st-andrews.ac.uk

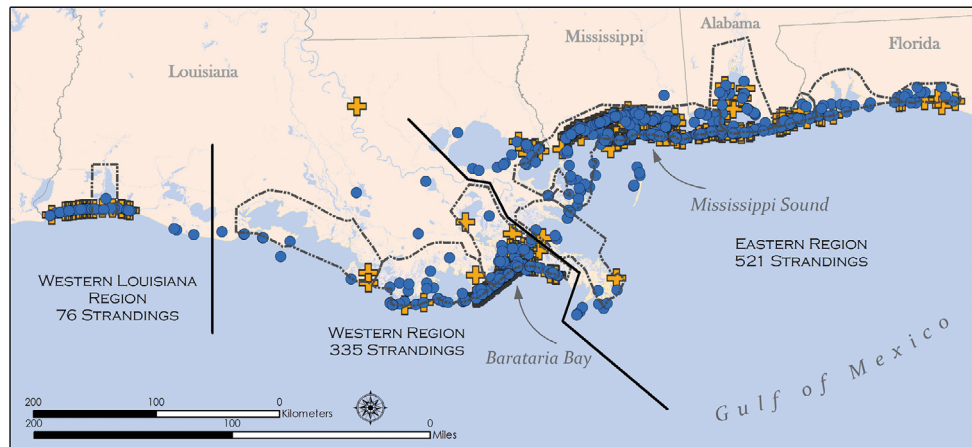


Fig. 1. Reported locations of strandings of common bottlenose dolphins *Tursiops truncatus* with (yellow crosses) and without (blue dots) additional information about stock of origin (i.e. either genetic assignment, stable isotopes, photo-ID or genetic ID). Analysis regions are shown, together with total strandings per region and approximate region boundaries (thick black lines). Approximate bay, sound and estuary (BSE) stock boundaries are shown as dot-dashed lines, with the location of the Barataria Bay (BB) and Mississippi Sound (MS) stocks indicated

of common bottlenose dolphins *Tursiops truncatus* (henceforth 'bottlenose dolphins') (Fig. 1) (Waring et al. 2015, Vollmer & Rosel 2013). Unusually large numbers of bottlenose dolphins were found stranded after the spill within the spill footprint, and a significant weight of evidence indicated that a large portion of these strandings were caused by exposure to petroleum-associated chemicals from the DWH spill (Venn-Watson et al. 2015a,b, DHNRDAT 2016). A Natural Resources Damage Assessment (NRDA) was initiated, one aim of which was to quantify injury to bottlenose dolphins from the BSE and coastal waters of the nGoM. Part of the assessment involved using the strandings to scale the number of mortalities by stock, for those stocks for which more direct estimates of mortality from mark–recapture studies were not available (DHNRDAT 2016, McDonald et al. 2017, this Theme Section). This requires accounting for the proportion of recorded strandings from each stock, the probability that an animal from each stock that died would strand at a given location (which differs between stocks) and the probability that a stranding at that location would be discovered and reported to the stranding network. Here, we focus only on the first question: What proportion of the recorded strandings came from each stock? The remaining questions are addressed, and an overview of the NRDA for bottlenose dolphins is given, in DHNRDAT (2016).

Dolphins stranded in a particular location are assumed to be either from that local BSE stock or the adjacent coastal stock, as movement between BSE stocks is negligible. Satellite telemetry demonstrated

that dolphins tagged within Barataria Bay (BB, Fig. 1) primarily remained within the Bay—and the few that entered coastal waters remained within 1.75 km of shore (Wells et al. 2017, this Theme Section). Strandings of offshore dolphins, including the offshore morphotype of bottlenose dolphins, are also very rare (Peltier et al. 2012, Carretta et al. 2016, DHNRDAT 2016). For example, in the nGoM, 620 stranded bottlenose dolphins were examined genetically during the time of the DWH oil spill and only 2 were of the offshore morphotype. In addition, a carcass drift model constructed specifically for the north central Gulf, using high resolution bathymetry and shoreline, wind and freshwater flow data revealed that the probability that a carcass originating in waters >20 m depth would beach was <1% (DWH MMIQT 2015). Therefore, our main question for each BSE was: What proportion of the stranded animals were from that BSE and what proportion were from the adjacent coastal stock? To help address this question, 4 sources of information were available, each on a different but overlapping subset of the stranded animals. First, multilocus genotypes (microsatellites) were obtained from samples collected via remote biopsy and live capture-release efforts to represent source stocks; specifically, the BB BSE (Fig. 1) and the adjacent western coastal stock. Genetic assignment tests were not attempted for strandings in the Western Louisiana region or the East region because sampling of potential baseline populations has not been done. The identified multilocus genotypes for both stocks were used as a baseline for comparison to perform a genetic stock assignment using available samples

from dolphins that stranded near BB (Rosel et al. 2017, this Theme Section). Second, stable isotope ratio (SIR) data were obtained from samples collected via remote biopsy from putative source stocks in both inshore and coastal waters near BB and Mississippi Sound (MS) (Fig. 1); SIR of carbon, nitrogen and sulfur were used to perform an isotope-based stock assignment for fresh dead or moderately decomposed stranded animals between western Louisiana and Choctawhatchee Bay, Florida, for which frozen skin had been collected (Hohn et al. 2017, this Theme Section). Samples used in the training datasets for the genetic assignment tests and stable isotope analysis were collected in geographically distinct locations that excluded areas of likely stock overlap. For example, samples representing the BB estuarine stock were collected within BB, and samples representing the coastal stock adjacent to BB were collected no closer than 2 km from shore. Third, photographic identification (photo-ID) matching was attempted by comparing photographs of dorsal fins from dead, stranded animals that were of sufficient quality (e.g. not too decomposed) to photo-ID catalogs available from the BB and MS estuarine stocks (Melancon et al. 2011). Last, genetic data including microsatellite, sex and mitochondrial DNA sequences were used to search for ‘genetic’ matches between the stranded and previously biopsied animals (as in Rosel et al. 2017). There was an attempt to collect all 4 sources of information wherever possible, but practical constraints such as carcass decomposition and sampling logistics meant that the majority of strandings were not suitable for these analyses; for those that were analyzed and did have additional information, more than half had information from 2 or more sources, but very few had all 4 (see Table 1).

Here, we present a Bayesian hierarchical model (Parent & Rivot 2012) for combining the sources of data. Our model represents the information from each data source for each animal in a flexible way (using a finite mixture of beta distributions); this information is then combined by weighting each data source using a measure of the precision of the estimates from that source (the effective sample size, Morita et al. 2008). The model was fitted to animals for which additional data were available, and then used to infer the stock of animals for which there was no additional information. The underlying assumption for the current study was that the probability of being from the coastal or BSE stock was the same for sampled versus unsampled dolphins. This approach represents a general method for the objective integration of multiple sources of information where

there is uncertainty associated with each source. We applied the model to dolphin stranding data collected between May 2010 and June 2014 to make joint inferences about the proportion of strandings from coastal versus BSE stocks in 3 regions affected by the oil spill.

MATERIALS AND METHODS

Hierarchical model for combining information

Basic model

Let N be the number of stranded animals recorded, of which N_a have additional information about the stock of the animal, and N_b do not, so $N = N_a + N_b$. Let $n = n_a + n_b$ be the unknown number of these that come from the coastal stock (where n_a is the number of coastal animals with additional information and n_b is the number without). Our inference goal is to estimate the proportion of coastal animals, which we denote p , where $p = n/N$, and hence the proportion of BSE animals is $1 - p$. The model is shown as a directed acyclic graph (DAG) in Fig. 2.

For each of the N_a individuals, there are up to M pieces of additional information. In the current study, $M = 4$, with potential information from genetic assignment, SIR assignment, photo-ID matching and genetic-ID matching. Each piece of information is in the form of a distribution on the probability that the animal is from the coastal stock, p_{ij} (where i indexes the individual and j indexes the data source). We use a distribution because there is uncertainty associated with each information source about the stock assignment probability, particularly for the genetic assignment and SIR sources. For the moment, we assume this uncertainty can be expressed in the form of a beta distribution (we will relax this assumption in the next subsection), so that:

$$p_{ij} \sim \text{beta}(\alpha_{ij}, \beta_{ij}) \quad (1)$$

where α_{ij} and β_{ij} are derived from the additional information and are assumed known.

We wish to combine the multiple sources of information for each animal (where multiple sources are available), to obtain an ensemble estimate of p_i : the probability that the i^{th} animal is coastal. To do this, we model p_i as a weighted sum of the p_{ij} components, so that p_i has a mixture distribution with density:

$$f(p_i; \underline{\alpha}_i, \underline{\beta}_i) = \sum_{j=1}^M \pi_{ij} f(p_{ij}; \alpha_{ij}, \beta_{ij}) \quad (2)$$

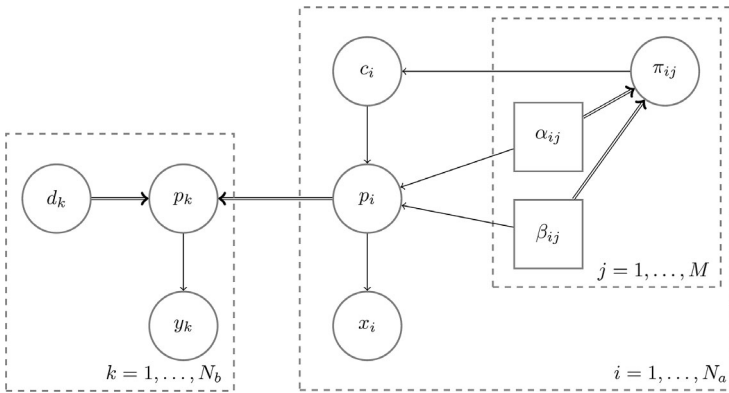


Fig. 2. Directed acyclic graph (DAG) indicating structure of the basic model for integrating multiple data sources. Square boxes denote constants, circles denote model parameters; single arrows denote stochastic dependency, double arrows denote deterministic dependency; dashed boxes denote groups of variables. Symbols are defined in 'Materials and methods: Hierarchical model for combining information'

where $f(p; \theta)$ denotes the probability density function (pdf) of the random variable p , with parameter vector θ , and π_{ij} denotes the mixing weights (which sum over j to 1); these weights should reflect how much information is contained in each data source about p_i . A natural weight would be the precision (i.e. the inverse variance); however, the precision of a beta distribution is not independent of the mean (precision is lowest at $p = 0.5$, and increases as p approaches 0 or 1). This could potentially cause issues with more extreme estimates of p getting larger weights, just because they are closer to 0 or 1. Instead, we use the effective sample size (ESS; Morita et al. 2008), which is a measure of precision that is independent of p . For a beta distribution, the ESS is $\alpha + \beta$, leading to weights:

$$\pi_{ij} = \frac{\alpha_{ij} + \beta_{ij}}{\sum_{l=1}^M (\alpha_{il} + \beta_{il})} \tag{3}$$

(the denominator ensures the weights for each animal sum to 1). For those animals for which a particular data type is not available, the corresponding α_{ij} and β_{ij} are set to zero. This has the effect of making the $\pi_{ij} = 0$ and hence information from that data type does not enter the mixture.

In practice, results are computed using posterior simulation (see 'Application of the hierarchical model' below). To facilitate sampling from the mixture distribution, the model is augmented with a latent allocation vector \underline{c} , such that each element of \underline{c} takes an integer value between 1 and M , indicating which of the data sources to use in the mixture. To create the mixture given in Eq. (2), we set:

$$c_i \sim \text{categorical}(\pi_{i1}, \dots, \pi_{iM}) \tag{4}$$

and then:

$$p_i | (c_i = j) = p_{ij}$$

This means that, on average, a proportion of the simulated values of π_{i1} will come from mixture component 1, π_{i2} from mixture component 2, etc. Hence, the relative values of the ESS for each data source govern the amount of influence that each data source has on the combined distribution of probability of being coastal for each stranded animal.

To estimate n_a (i.e. the number of stranded animals with additional information that come from the coastal stock), we assume the stock status of each of the N_a animals is a binary random variable, x_i with value 0 if BSE and 1 if coastal, and that:

$$x_i \sim \text{Bernoulli}(p_i) \tag{5}$$

so:

$$n_a = \sum_{i=1}^{N_a} x_i \tag{6}$$

Lastly, we need an estimate of n_b (the number of stranded animals without additional information that come from the coastal stock). To do this, we make the mild assumption that animals with and without additional information have the same distributions of probability of being coastal (see the Supplement at www.int-res.com/articles/suppl/n033p253_supp.pdf for more details, and alternative approaches that require stronger assumptions). For each of the n_b animals without additional information, we model the probability of being coastal as an unweighted mixture of the probability of being coastal for animals with additional information:

$$f(p_k; \underline{\alpha}, \underline{\beta}) = \frac{1}{n_a} \sum_{i=1}^{N_a} f(p_i; \underline{\alpha}_i, \underline{\beta}_i) \tag{7}$$

where k indexes individual ($k = 1, \dots, N_b$) (Analogous to the animals with additional information, we implement this mixture by defining a latent allocation vector, which we denote \underline{d} , where each element of \underline{d} takes a value between 1 and N_a , indicating which of the N_a animals with additional information distribution of p to use:

$$d_k \sim \text{categorical}(v_1, \dots, v_{N_a}) \tag{8}$$

where all v take the value $1/N_a$, and then:

$$p_k | (d_k = i) = p_i \tag{9}$$

Again analogous to the animals with additional information, we assume stock status of each of the N_b

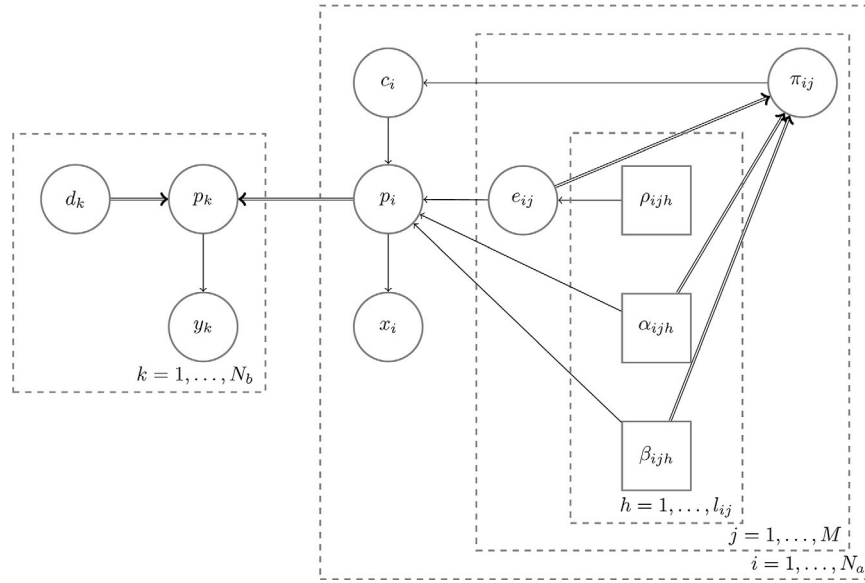


Fig. 3. Directed acyclic graph (DAG) indicating structure of the extended model for integrating multiple data sources. Notation is defined in Fig. 2 legend; symbols are defined in ‘Materials and methods: Hierarchical model for combining information’

animals without additional information is a binary random variable, y_j with value 0 if BSE and 1 if coastal, and that:

$$y_j \sim \text{Bernoulli}(p_j) \quad (10)$$

so:

$$n_b = \sum_{i=1}^{N_b} Y_j \quad (11)$$

Extended model

In practice, we found that a beta distribution did not fit the empirical distribution of probability of being coastal (i.e. p_{ij}) for some data sources and animals. This was apparent because the empirical distribution often had a lower variance than expected from a beta distribution with the same mean, and in some cases was even bi-modal. We therefore extended the basic model to allow p_{ij} to be a finite mixture of beta distributions:

$$p_{ij} \sim \sum_{h=1}^{I_{ij}} p_{ijh} \text{beta}(\alpha_{ijh}, \beta_{ijh}) \quad (12)$$

where I_{ij} is the number of mixture components, each component having parameters α_{ijh} and β_{ijh} , and mixture weighting p_{ijh} where $\sum_{h=1}^{I_{ij}} p_{ijh} = 1$. The DAG corresponding to this extended model is shown in Fig. 3. The ESS for the mixture is $\sum_{h=1}^{I_{ij}} p_{ijh}(\alpha_{ijh} + \beta_{ijh})$.

As with the other mixture distributions in the model, we implemented this extension via data augmentation, this time defining an N_a times M matrix \underline{E} , where each element, e_{ij} , takes a value from 1 to I_{ij} indicating which of the I_{ij} mixture components to use:

$$e_{ij} \sim \text{categorical}(p_{ij1}, \dots, p_{ijI_{ij}}) \quad (13)$$

Application to GoM strandings data

We fit the above model to strandings and associated data collected in the area inhabited by 10 BSE stocks and 2 coastal stocks (Fig. 1). For the purposes of analysis, these stocks were aggregated into 3 regions (shown on Fig. 1): East of the Mississippi outflow (E), West of the Mississippi outflow (W) and Western Louisiana (west of Vermilion Bay, Louisiana to the Texas-Louisiana border; WL). The WL region in particular was thought *a priori* to be likely to have a higher proportion of coastal strandings because there is only one estuarine population in the region, it is very small and there are no barrier islands. Data from common bottlenose dolphins stranded between the Texas–Louisiana border and Choctawhatchee Bay, Florida, from May 2010 to June 2014 were extracted from the NOAA Marine Mammal Health and Stranding Response Program (MMHSRP) database (<https://mmhsrp.nmfs.noaa.gov/mmhsrp/>) on 19 March 2015. Details on how marine mammal stranding data are collected can be found in Litz et al. (2014).

Here, we describe each of the additional datasets, and the analyses performed on them to yield distributions on p_{ij} (i.e. the probability of being coastal). Unless noted otherwise, analyses were performed in R version 3.2.0 (R Core Team 2015).

Genetic assignment data

Genetic assignment aims to identify the origin of individuals sampled from a mixture of genetically distinct populations (Manel et al. 2005). A complete description of the collection, processing and initial analysis of the genetic assignment data used here is given in Rosel et al. (2017). In summary, biopsy samples from putative source populations were collected from live animals in Barataria Bay (representing the BB stock) and adjacent coastal waters greater than 2 km from shore (representing the coastal stock); DNA was extracted from these samples and genotyped at 41 microsatellite loci. Microsatellite data were used as input to a genetic assignment analysis, using the software ONCOR (Kalinowski et al. 2007). Three genetically distinct groups were identified: Barataria Bay Estuary (BBE), Barataria Bay Island (BBI) and coastal. This assignment scheme was then applied to genetic samples taken from dolphins stranded along the barrier islands in the estuarine waters of Barataria Bay, yielding estimated probabilities of each stranded dolphin falling in the BBE, BBI and coastal groups (where the probabilities sum to 1). For our purposes, the BBE and BBI probabilities were summed to yield probability of being from a BSE stock (the 3 groups were initially identified to improve genetic assignment; see Rosel et al. 2017).

The ONCOR software is limited in that it does not produce estimates of uncertainty in the genetic assignment probabilities. Recall that a distribution of the probability of being from the coastal stock is required as an input to the hierarchical data-combining model. We therefore employed a stratified non-parametric bootstrap, resampling with replacement from individuals within the baseline dataset, but keeping the number of animals in each of the 3 distinct groups constant. This stratified scheme was similar to the original data collection, where separate sampling was performed on bay and coastal populations (although the existence of 2 genetically distinct populations within the bay was not known at that time, so no within-bay stratification was performed). Each of these resampled datasets was then used to construct a new assignment scheme in ONCOR, and a new set of assignment probabilities for stranded

animals was generated. This produced an empirical distribution of the probability of being coastal (and, conversely, of being BSE) for each stranded animal in the dataset. We used 1000 resamples. The empirical distribution was then fitted to a beta distribution, using maximum likelihood, to produce a parametric distribution of assignment probabilities that could be used in the subsequent hierarchical analysis. However, we found that the empirical distribution of assignment probabilities for some animals was bimodal with a mode at 0 or 1 (indicating certainty the animals were BSE or coastal, respectively), and another mode away from the boundary, a shape that cannot be matched by a simple beta distribution. Therefore, for each stranded animal, we fitted a simple beta, and also a 2-point mixture of beta distributions (Eq. 12, with $I_{ij} = 2$), and selected the model with lowest Akaike information criterion value (Akaike 2014). To increase robustness when fitting, we replaced bootstrap resample estimates of p_{ij} that were 0 with 1×10^{-10} , and 1 with $1 - 1 \times 10^{-10}$; we also constrained the fitting such that $\alpha_{ijh} + \beta_{ijh} \leq 10000$ —values this high represent a spike at 0 or 1. In cases where the 2 point mixture fitting did not converge, we used the 1-point model. Examination of the selected model for each animal showed that all fit the data well, so further extension to a 3-point mixture was not required.

SIR data

Habitat preferences and prey selection variations for common bottlenose dolphins in estuarine and coastal environments in the study area lead to different SIR signatures between these groups (Barros et al. 2010); hence, analysis of SIRs from carcasses can be used to identify recent habitat preferences for each dolphin, which may be used to infer stock (for more detail see Hohn et al. 2017); the data and analyses used here are described fully in Hohn et al. (2017). In summary, a training dataset of carbon, nitrogen and sulphur SIRs ($\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and $\delta^{34}\text{S}$ respectively) was created using biopsy samples from dolphins sampled within defined stock areas in Louisiana and Mississippi (including estuarine, barrier island and coastal) and assumed to represent those stocks. A classification tree was constructed using binary partition analysis ('rpart' library in R v.4.1.10; Therneau et al. 2015) with candidate discrete variable region (East or West—the latter being a combination of the W and WL regions used here) and continuous variables $\delta^{13}\text{C}$, $\delta^{15}\text{N}$ and $\delta^{34}\text{S}$. The

constructed classification tree was then applied to stranded animals for which SIR data were available, yielding an estimate of probability of being coastal for each sample. Estimates of the distribution of assignment probabilities were derived using a stratified nonparametric bootstrap (called 'bagging' in the binary partition literature; e.g. Sutton 2005), resampling 1000 times with replacement from individuals within the training dataset while keeping the same number of estuarine, barrier island and coastal animals in each region (East and West), and each time re-creating the regression tree and re-applying the tree to the stranded-animal data. The empirical distribution of probability of being coastal for each stranded animal was fitted to a beta mixture distribution using the same methods as applied to the genetic assignment data.

Photo- and genetic-ID data

Photographic catalogs exist for some BSE populations (McDonald et al. 2017) (but no coastal populations) in the study area, and these were searched in cases where adequate dorsal fin photographs existed for the stranded animals (many carcasses were too decomposed for accurate photo-identification). Where a match was made, the stranded animal was assigned a probability distribution of being coastal, $p_{ij} \sim \text{beta}(1, 9999)$, equivalent to assuming that it is almost certain the animal is not coastal (allowing for a small mis-identification error).

Likewise, the genotypic data from 19 microsatellite loci, the sex and the mitochondrial DNA control region sequence of stranded animals from an area slightly broader than the BB study area were compared to the same datasets collected from the biopsied dolphins sampled in BB and adjacent coastal waters. Where a stranded animal matched a live-sampled animal in all 3 genetic datasets, the stranded animal was assigned a probability distribution of being coastal, $p_{ij} \sim \text{beta}(1, 9999)$, similar to the photo-ID matches.

Note that no catalog exists for either data source for coastal populations (apart from the biopsied animals, which represent a very small fraction of the total population), and hence the chance of identifying matches to the coastal stock with these data sources is close to zero. Also, animals where a match was attempted but none found were treated identically to those where no match was attempted. These factors potentially bias the results in favor of BSE assignments. We therefore ran our analyses with and with-

out these data sources (see 'Discussion' for an alternative approach).

Application of the hierarchical model

The hierarchical model was fitted separately to each of the 3 regions (E, W, and WL) and was fitted twice in each region, once using all 4 sources of additional data and once using only the genetic assignment and SIR data.

Model fitting was performed using the software JAGS v.3.3.0 (Plummer 2012) (via the R library 'R2jags'). JAGS implements Markov chain Monte Carlo algorithms, although for this model all that is required is the ability to simulate from beta and categorical distributions. Arbitrary starting values were used, and a burn-in of 1000 iterations (very fast convergence was observed), after which 100 000 iterations were undertaken with thinning by 10, meaning that 10 000 iterations were used for inference. The reported estimates are the mean of the simulated values; Monte Carlo error was essentially negligible (results are accurate to at least 3 significant figures).

RESULTS

Strandings and additional data

A total of 932 strandings were recorded in the study region: 521, 335 and 76 respectively from the E, W and WL regions (Fig. 1). Of these, 366 (approx. 39%) had one or more sources of additional information (Table 1, Fig. S1 in the Supplement at www.int-res.com/articles/suppl/n033p253_supp.pdf): 129 with genetic assignment data (all from the W region), 217 with SIR data, 212 with attempted photo-ID matches and 175 with attempted genetic-ID matches (all from the W region). However, as stated previously, only successful photo-ID and genetic-ID matches were used in the analysis; after excluding unsuccessful matches there were 300 animals with one or more sources of additional information. For the analysis using only genetic assignment and SIR data there were 290 animals with one or both data types.

Genetic assignment data

A total of 140 biopsy and live capture-release samples from BB and adjacent coastal waters were used

Table 1. Number of common bottlenose dolphins *Tursiops truncatus* recovered after the *Deepwater Horizon* oil spill with or without additional data regarding their stock of origin (Y: yes; N: no), summarized by region: E: East; W: West; WL: western Louisiana. Photo- and genetic-ID columns indicate those animals that were checked for matches, not those for which positive matches were made

Genetic assignment	Stable isotope	Photo-ID	Genetic-ID	Region		
				E	W	WL
N	N	N	N	371	142	53
			Y	0	16	0
		Y	N	31	17	0
			Y	1	11	0
	Y	N	N	47	4	21
			Y	0	5	0
		Y	N	68	2	0
			Y	3	9	2
Y	N	N	N	0	1	0
			Y	0	50	0
		Y	N	0	0	0
			Y	0	22	0
	Y	N	N	0	0	0
			Y	0	10	0
		Y	N	0	0	0
			Y	0	46	0

to build the genetic assignment in ONCOR; this was then applied to the 129 stranded animals for which genetic data had been collected (see Rosel et al. 2017 for details). The mean estimated probability of being from the coastal stock, $\bar{p}_{.1}$, was 0.062 (the dot subscript is because we are averaging over stranded individuals and the 1 is because genetic assignment data are data source 1).

The bootstrap resampled estimates fell into 2 almost equal groups: those (64/129) where almost all bootstrap resamples had a very low p_{i1} (<0.1), indicating high confidence that the animal was not from the coastal stock, and those (65/129) where at least some bootstrap resample estimate of p_{i1} were higher (>0.1), indicating some uncertainty in assignment. This is illustrated in Fig. S2 in the Supplement, which shows bootstrap distributions of p_{i1} for selected individuals. The mean of the bootstrap estimates (i.e. mean over bootstraps and individuals) was 0.046 and the mean standard deviation (SD; i.e. mean over individuals of the SD of p_{i1} calculated within individuals using the bootstrap replicates) was 0.072.

Fig. S2 also shows the fit of the beta distributions to the bootstrap resamples: for 71 (55%) animals the 1-point mixture was chosen and for 58 (45%), the 2-point was chosen. Mean and SD of p_{i1} were 0.046 and 0.070, nearly identical to the bootstrap values, emphasizing the closeness of the fit. The mean ESS, i.e. mean over stranded animals of $\sum_{h=1}^{ij} \rho_{ijh}(\alpha_{ijh} + \beta_{ijh})$,

was 4302, with median 1183. There was a strong relationship between the estimated p_{i1} and the ESS (see Fig. S3 in the Supplement): very low values of p_{i1} had very spiked distributions and hence high ESS, while larger values of p_{i1} had flatter distributions and hence lower ESS. This means that for those animals for which there is strong evidence they are not from the coastal stock, the genetic assignment data will be weighted heavily if other data are available, while for those for which there is less certainty, the weighting will be less.

SIR data

A total of 205 remote biopsy and live capture-release samples from BB, MS, and adjacent coastal waters were used to build the binary partition, which was then applied to the 217 stranded animals for which SIR data had been collected (see Hohn et al. 2017 for details). Three of the 4 candidate variables (Region, $\delta^{13}\text{C}$ and $\delta^{34}\text{S}$) were selected, with the first split being on $\delta^{34}\text{S}$ (high values indicating coastal) and 4 subsequent partitions (6 terminal nodes, i.e. 6 distinct values that could be assigned to p_{i2}). The correct classification rate on the training dataset was 80%. For stranded animals, the mean estimated probability of being from the coastal stock, $\bar{p}_{.2}$, was 0.172. This is higher than that for the genetic assignment data, but many of the animals are different and the geographic area covered was larger.

The bootstrap distributions of p_{i2} for each animal were generally more dispersed than for the genetic assignment data, with few animals being assigned extreme values (either 0 or 1) in many resamples. A representative selection is shown in Fig. S4 in the Supplement. The mean of the bootstrap estimates was 0.247 and the mean SD 0.172. There was considerable regional variation: the bootstrap means for strandings in the E, W and WL regions were 0.216, 0.176 and 0.642 respectively.

As with the genetic assignment data, the beta distributions fit the bootstrap resamples very well (examples shown in Fig. S4): the mean and SD from the beta fits was 0.247 and 0.177 respectively—very close to those from the bootstrap. The 2-point mixture was chosen in the majority (191/217) of cases. The mean ESS was 49.9 and median 33.5; these values are approximately 2 orders of magnitude less than for the genetic assignment data. As with the genetic assignment data, the ESS is higher for values of p_{i2} closer to zero (Fig. S5 in the Supplement).

Photo- and genetic-ID data

Of the 212 strandings for which photographic matches were sought, matches were found for 38: 10 in the E and 28 in the W region. For the genetic-ID data, 175 stranded animals from the W region were compared to animals biopsied in the West region and 5 were positively matched.

Consistency among data sources

There were 73 animals with additional data from more than one data source (using just positive matches rather than all trials for photo- and genetic-ID data). Within animals, the probability of being from the coastal stock, p_{ij} , was generally consistent among sources, with only 12 (approx. 15%) having a difference of more than 0.2 between any pair of data sources (see Table S1 in the Supplement; the value 0.2 was chosen arbitrarily). In 10 cases, these were animals with a low probability of being coastal from genetic assignment (and, in 2 cases, photo-ID), but a higher probability from SIR assignment; these animals were all given high genetic assignment probability of being from BBI group rather than BBE group (Table S1). For the remaining 2 cases, estimated probabilities of being coastal from genetic assignment were 0.27 and 0.31, but the bootstrap 95% confidence intervals (percentile method) were 0 to 0.99, indicating very large uncertainty on the genetic assignment.

Combined inference from hierarchical model

Estimates of the proportion of coastal stock animals, p , by region are shown in Table 2. There is a clear difference between regions, with the West

region having the lowest p (0.06), East next lowest (0.21) and western Louisiana the highest (0.65). As expected, the estimates of p are higher when photo- and genetic-ID data are included than without these data sources, but the difference is less than 15% for all regions. Precision in all cases is good, with coefficients of variation between 11 and 26%.

DISCUSSION

Our hierarchical model for combining probability distributions from multiple sources, weighting by ESS, is very general: it can be thought of as a generalized precision-weighted average. In constructing the model and deriving the input distributions we used minimal assumptions. The model allows us to have a different distribution of the probability of being coastal for each animal, and to make inferences about dolphins that do not have additional information by sampling at random from those that do, i.e. we assume that animals are 'exchangeable'. Input distributions for the genetic assignment and SIR data sources were derived using a nonparametric bootstrap, which assumes that the animals in the remote biopsy datasets from putative source populations are representative of that population and are sampled independently, followed by a flexible semi-parametric method (finite mixtures) to obtain a form that can be used by the hierarchical model. We analyzed the data in 3 geographic regions, allowing for regional variation in stock structure.

Our method uses a 2-stage approach: first, separate analyses are performed on each dataset, and second, the results of these analyses are combined. This approach has the advantage that it allows specialist software and methods to be applied to each dataset, such as the use of ONCOR for the genetic assignment, and binary classification for stable isotope

Table 2. Estimated proportion of stranded animals (p) from the coastal stock of common bottlenose dolphins *Tursiops truncatus* by region, with associated standard error (SE), coefficient of variation (CV) and 95% credible interval (CI). Columns labelled '4 data sources' were derived by combining genetic assignment, stable isotope ratio, photo-ID and genetic-ID data; columns labelled '2 data sources' were derived from the genetic assignment and stable isotope ratio data only—these latter estimates are considered more reliable (see 'Discussion')

Region	4 data sources				2 data sources			
	p	SE(p)	CV(p)	95%CI (p)	p	SE(p)	CV(p)	95%CI (p)
East	0.208	0.024	0.11	0.163–0.255	0.215	0.025	0.12	0.169–0.263
West	0.058	0.015	0.26	0.030–0.090	0.065	0.016	0.25	0.036–0.099
Western Louisiana	0.653	0.081	0.12	0.487–0.816	0.652	0.081	0.12	0.487–0.803

analysis. These analyses can generally be led by the scientists with expertise in each dataset. On the other hand, the question of how to weight the different results arises at the second stage. An alternative would be a 1-stage approach, where the primary data analyses are integrated into a unified hierarchical model. The true stock of each animal would be a latent state, just as in our current model, but our beta-mixture observation model would be replaced by a much more complex model specific to each dataset, describing how that dataset contributes information about stock. Model construction and estimation would be more complex (and likely led by a statistician rather than subject matter expert), but the requirement for weighting may be alleviated. Note, however, that integrated models are common in other fields, such as fisheries, and questions of data weighting remain (e.g. Francis 2011). One reason to use weightings in this context is if there are thought to be other, un-modelled sources of error, or other subjective reasons to favor one source of data over another. Overall, we feel that given the application, our approach is a good compromise between overly simple (just using 1 dataset, or taking an unweighted average) and overly complex (an integrated 1-stage method).

We chose to use ESS as the basis for precision weighting. Our measure of ESS for the standard beta distribution is standard; on the other hand, there is no standard measure for beta mixtures. We took a weighted average of the mixture components, but this has the disadvantage of producing high ESS values when each beta component is spiked, even if the overall distribution is strongly multimodal and therefore quite uninformative about the parameter of interest. Alternative measures of ESS could be investigated, such as the approach of Morita et al. (2008, 2012), which is based on finding the number of observations required to obtain a posterior distribution close to the specified prior, starting from a maximally uninformative prior distribution.

We assume that all of the data sources and associated analyses give unbiased information on the animals' stock of origin. This requires that the putative source stock datasets contain animals that were correctly assigned to stock, that the data source is informative about stock origin and that the analysis methods are unbiased. These issues are discussed by Rosel et al. (2017) for genetic assignment, and Hohn et al. (2017) for SIR. One concern, in particular, is that SIR data tell us where animals have been foraging over the past few weeks to months (Browning et al. 2014, Giménez et al. 2016), not necessarily from

where they originated (Hohn et al. 2017). For example, in 10 cases, animals that were assigned via genetics to the BSE population with high probability had stable isotope signatures more indicative of a coastal stock animal (see Table S1 in the Supplement). These animals genetically were assigned to the BSE population in BB that prefers the inshore waters around the barrier islands separating BB from coastal waters. From satellite telemetry of dolphins in BB, Wells et al. (2017) found ranging patterns defining 3 patterns of habitat use. Overall, while dolphins occasionally entered Gulf coastal waters, most remained within BB. As with the genetic data, 1 group of dolphins was strongly associated with the barrier islands. The average maximum distance from shore for the BB island-associated dolphins was 1.75 km with a maximum distance of 4.2 km from shore. These movements, while not migration, would result in assimilation of the coastal isotopic signature in the skin sample if the dolphins had recently fed on fish of a more coastal origin during these trips into nearshore coastal waters. However, their genetic signature indicates that, evolutionarily, they originate from the estuarine stock of dolphins and not the coastal stock. We note that SIR data had, on average, a lower ESS than genetic assignment, indicating lower precision, perhaps partly as a result of this ambiguity.

Photo-ID and genetic-ID matches are more opportunistic sources of data. Matches between stranded carcasses and previously photographed or biopsied live dolphins are rare, but when a positive match is found it can offer some confirmation of the stock of origin for that individual carcass. As a result (and noted in the 'Materials and methods'), the photo- and genetic-ID data likely bias the result away from coastal and towards BSE assignments, which was problematic for this analysis. We therefore consider the analysis excluding these data sources to be more reliable, and only genetic assignment and SIR data were used in the final analysis. One issue with the photo- and genetic-ID data was that we did not use information about when a match was attempted but none was found—to use this requires additional information, as follows. If we use the notation C for coastal, C' for not coastal (i.e. BSE), M for a match and M' for no match then, using Bayes' rule:

$$\begin{aligned}
 P(C' \text{ given } M') &= \frac{P(M' \text{ given } C')P(C')}{P(M' \text{ given } C')P(C') + P(M' \text{ given } C)P(C)} \\
 &= \frac{P(M' \text{ given } C')P(C')}{P(M' \text{ given } C')P(C') + 1 \times P(C)} \quad (14)
 \end{aligned}$$

where $P(M' \text{ given } C')$ is the probability of no match given a BSE animal, which will be a function of the proportion of the BSE stock in the catalog and the false negative rate of the matching, and $P(C)$ and $P(C')$ are the *a priori* probability of the stranded animal being from the coastal and BSE stock, respectively. These could conceivably come from prior distributions on coastal and BSE stock sizes, mortality rates, stranding and reporting probabilities. However, since stock-specific mortality rate is the reason for attempting to divide strandings into stock in the first place, it only makes sense to attempt such an analysis in the context of a much larger, integrated analysis of mortality and stranding, something for which we did not have resources during the damage assessment process. Instead, we elected not to use the photo- and genetic-ID data, losing some precision in our estimates of stock structure of stranded animals, but without introducing any bias.

Our assumption that animals with additional data have the same probability of being from coastal versus BSE stocks as those without is broadly plausible. It is possible that animals from the coastal stock, which have farther to travel before stranding, would tend to be more decomposed and so less likely to be amenable to collection of additional data; however, decomposition mainly affected the ability to undertake photo-ID, and had less effect on the genetic assignment and stable isotope data used here. A possible bias in the other direction is that BSE animals may have a higher tendency to strand in remote locations and remain undetected.

We found that the proportion of stranded animals from the coastal stock was low (0.07) in the West region, moderate in the East (0.22) and high in western Louisiana (0.65). The higher value for western Louisiana is not surprising given that all strandings recovered were from the ocean side of the beaches, and there is very little estuarine habitat including only one small estuarine stock in Lake Calcasieu in far western Louisiana (Waring et al. 2015). In contrast, both the West and East regions sustain relatively large BSE stocks, with abundance estimates of >2000 dolphins for BB in the West region (McDonald et al. 2017), and >3000 dolphins for MS in the East region (Mullin et al. unpubl.). However, while the MS BSE stock is relatively large, the sound is a semi-open embayment, and prior observations of seasonal fluctuations in dolphin abundance have led to the hypothesis that members of the adjacent northern coastal stock may periodically enter the BSE waters, where they could potentially die and strand. This may account for the higher proportion of coastal

stock assignments for the East versus West region strandings.

Ultimately, the probabilities calculated by our model were used to apportion strandings to the appropriate coastal or BSE stock within the *Deepwater Horizon* oil spill footprint; the number of strandings that were in excess of expected, and presumed to be attributable to the oil spill, was estimated for each stock (DWH MMIQT 2015). The excess strandings were then scaled to estimate overall excess mortality using stock-specific probabilities for beaching and recovery estimated using a combination of modeling approaches (described in DWH MMIQT 2015). Of the 2 coastal stocks, the northern coastal stock experienced the greater number of mortalities, estimated at 3202 dolphins (DWH MMIQT 2015), partially due to the higher estimated probability of stranded animals being from the coastal stock for the East region. However, also contributing to the greater number of mortalities for the northern coastal stock was the high number of strandings tallied from across the East region (Fig. 1); all carcasses from this region classified as coastal were attributed to the northern coastal stock. The relatively large effect on this coastal stock is not surprising given that the majority of the range of the northern coastal stock intersects with the DWH oil spill footprint. Conversely, the estimated number of mortalities for the Western Coastal stock, which has relatively small overlap with the DWH oil spill footprint, was only 238 dolphins (DWH MMIQT 2015).

Acknowledgements. We thank Jill Bodnar for preparing Figs. 1 & S1; and Brian Quigley, John Venturella and Todd Speakman for conducting photo-analysis comparing images of stranded dolphin fins to existing photo-ID catalogs. This work was part of the *Deepwater Horizon* NRDA conducted cooperatively among NOAA, other US Federal and State Trustees, and BP. We thank the southeast US Marine Mammal stranding network for the collection of samples from stranded dolphins, particularly staff from Louisiana Department of Wildlife and Fisheries, Audubon Aquarium of the Americas, Institute for Marine Mammal Studies, Dauphin Island Sea Lab, Emerald Coast Wildlife Refuge and Gulf World Marine Park. We recognize Teri Rowles, Blair Mase-Guthrie, Erin Fougères, Elizabeth Stratton, Gina Rappucci, Sabrina Stevens, and Lauren Noble for their roles in stranding event and data management. We thank Eric Zolman, Brian Balmer, Todd Speakman (NOAA/NOS/NCCOS) and Carrie Sinclair (NOAA/NMFS/SEFSC) and biopsy and live-capture field crews for biopsy sample collection, Brian Quigley (NOAA/NOS/NCCOS) for photo-ID matching, and Lynsey Wilcox (NOAA/NMFS/SEFSC) for genetic data collection and laboratory technical support. We are grateful to Michail Papatomas for helpful discussions on the statistical methods. Lastly we thank Len Balthis and 3 anonymous reviewers for their very helpful comments. Response to cetacean stranding events is conducted by Marine Mammal

Stranding Networks (MMSN) authorized under Section 112c (Stranding Agreements from the NMFS regional offices) or Section 109h (Federal, State, or local government officials) of the Marine Mammal Protection Act. Remote biopsy sampling and capture-release sampling were conducted under NMFS Scientific Research Permit Nos. 779-1633-02, and 932-1905/MA-009526 respectively. Protocols for dolphin health assessments were reviewed and approved by the NOAA Institutional Animal Care and Use Committee.

LITERATURE CITED

- Akaike H (2014) Akaike's Information Criterion. In: Lovric M (ed) International encyclopedia of statistical science. Springer-Verlag, Berlin, p 25
- Barros NB, Ostrom PH, Stricker CA, Wells RH (2010) Stable isotopes differentiate bottlenose dolphins off west-central Florida. *Mar Mamm Sci* 26:324–336
- Browning NE, Dold C, Jack IF, Worthy GA (2014) Isotope turnover rates and diet-tissue discrimination in skin of *ex situ* bottlenose dolphins (*Tursiops truncatus*). *J Exp Biol* 217:214–221
- Carretta JV, Danil K, Chivers SJ, Weller DW and others (2016) Recovery rates of bottlenose dolphin (*Tursiops truncatus*) carcasses estimated from stranding and survival rate data. *Mar Mamm Sci* 32:349–362
- DHNRDAT (Deepwater Horizon Natural Resource Damage Assessment Trustees) (2016) *Deepwater Horizon* oil spill: final programmatic damage assessment and restoration plan and final programmatic environmental impact statement. www.gulfspillrestoration.noaa.gov/restoration-planning/gulf-plan (accessed 9 Jan 2015)
- DWH MMIQT (Deepwater Horizon Marine Mammal Injury Quantification Team) (2015) Models and analysis for the quantification of injury to Gulf of Mexico cetaceans from the *Deepwater Horizon* oil spill. DWH Marine Mammal NRDA Technical Working Group report. <https://www.doi.gov/deepwaterhorizon/adminrecord> (accessed 22 Mar 2016)
- Francis RICC (2011) Data weighting in statistical fisheries stock assessment models. *Can J Fish Aquat Sci* 68: 1124–1138
- Giménez J, Ramírez F, Almunia J, Forero MG, de Stephanis R (2016) From the pool to the sea: applicable isotope turnover rates and diet to skin discrimination factors for bottlenose dolphins (*Tursiops truncatus*). *J Exp Mar Biol Ecol* 475:54–61
- Hohn AA, Thomas L, Carmichael RH, Litz J and others (2017) Assigning stranded bottlenose dolphins to stock using stable isotope ratios following the *Deepwater Horizon* oil spill. *Endang Species Res* 33:235–252
- Kalinowski ST, Manlove KR, Taper MT (2007) ONCOR: a computer program for genetic stock identification. Montana State University, Bozeman, MT. www.montana.edu/kalinowski/Software/ONCOR.htm (accessed 29 June 2015)
- Litz JA, Baran MA, Bowen-Stevens SR, Carmichael R and others (2014) Review of historical unusual mortality events (UMEs) in the Gulf of Mexico (1990–2009): providing context for the multi-year northern Gulf of Mexico cetacean UME declared in 2010. *Dis Aquat Org* 112: 161–175
- Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *Trends Ecol Evol* 20:136–142
- McDonald TL, Hornsby FE, Speakman TR, Zolman ES and others (2017) Survival, density, and abundance of common bottlenose dolphins in Barataria Bay (USA) following the *Deepwater Horizon* oil spill. *Endang Species Res* 33:69–82
- Melancon RAS, Lane S, Speakman T, Hart LB and others (2011) Photo-identification field and laboratory protocols utilizing finbase version 2. NOAA Tech Rep Tech Memo NMFS-SEFSC-627
- Michel J, Owens EH, Zengel S, Graham A and others (2013) Extent and degree of shoreline oiling: *Deepwater Horizon* oil spill, Gulf of Mexico, USA. *PLOS ONE* 8:e65087
- Morita S, Thall PF, Müller P (2008) Determining the effective sample size of a parametric prior. *Biometrics* 64: 595–602
- Morita S, Thall PF, Müller P (2012) Prior effective sample size in conditionally independent hierarchical models. *Bayesian Anal* 7
- Parent E, Rivot E (2012) Introduction to hierarchical Bayesian modeling for ecological data. CRC Press, Boca Raton, FL
- Peltier H, Dabin W, Daniel P, Van Canneyt O, Dorémus G, Huona M, Ridoux V (2012) The significance of stranding data as indicators of cetacean populations at sea: modeling the drift of cetacean carcasses. *Ecol Indic* 18:278–290
- Plummer M (2012) JAGS version 3.3.0 user manual. <http://mcmc-jags.sourceforge.net/>
- R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Rosel PE, Wilcox LA, Sinclair C, Speakman TR, Tumlin MC, Zolman ES (2017) Genetic assignment to stock of stranded common bottlenose dolphins in southeastern Louisiana following the *Deepwater Horizon* oil spill. *Endang Species Res* 33:221–234
- Sutton CD (2005) Classification and regression trees, bagging and boosting. In: Rao CR, Wegman EJ, Solka JL (eds) *Handbook of statistics*, Vol 24. Elsevier, Amsterdam, p 303–329
- Therneau T, Atkinson B, Ripley B (2015) rpart: recursive partitioning and regression trees. R package version 4.1-9. <http://CRAN.R-project.org/package=rpart> (accessed 29 Oct 2015)
- Venn-Watson S, Colegrove KM, Litz J, Kinsel M and others (2015a) Adrenal gland and lung lesions in Gulf of Mexico common bottlenose dolphins (*Tursiops truncatus*) found dead following the *Deepwater Horizon* oil spill. *PLOS ONE* 10:e0126538
- Venn-Watson S, Garrison L, Litz J, Fougères E and others (2015b) Demographic clusters identified within the northern Gulf of Mexico common bottlenose dolphin (*Tursiops truncatus*) unusual mortality event: January 2010–June 2013. *PLOS ONE* 10:e0117248
- Vollmer NL, Rosel PE (2013) A review of common bottlenose dolphins (*Tursiops truncatus truncatus*) in the northern Gulf of Mexico: population biology, potential threats, and management. *Southeast Nat* 12:1–43
- Waring GT, Josephson E, Maze-Foley K, Rosel PE (eds) (2015) US Atlantic and Gulf of Mexico marine mammal stock assessments: 2014. NOAA Tech Memo NMFS-NE-231. US Department of Commerce, Woods Hole, MA
- Wells RS, Schwacke LH, Rowles TK, Balmer BC and others (2017) Ranging patterns of common bottlenose dolphins *Tursiops truncatus* in Barataria Bay, Louisiana, following the *Deepwater Horizon* oil spill. *Endang Species Res* 33: 159–180