NOAA Technical Memorandum NMFS

NOVEMBER 2011

EFFECTIVE STRIP WIDTHS FOR SHIP-BASED LINE-TRANSECT SURVEYS OF CETACEANS

Jay Barlow, Lisa T. Ballance, and Karin A. Forney

NOAA-TM-NMFS-SWFSC-484

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Marine Fisheries Service
Southwest Fisheries Science Center

The National Oceanic and Atmospheric Administration (NOAA), organized in 1970, has evolved into an agency that establishes national policies and manages and conserves our oceanic, coastal, and atmospheric resources. An organizational element within NOAA, the Office of Fisheries is responsible for fisheries policy and the direction of the National Marine Fisheries Service (NMFS).

In addition to its formal publications, the NMFS uses the NOAA Technical Memorandum series to issue informal scientific and technical publications when complete formal review and editorial processing are not appropriate or feasible. Documents within this series, however, reflect sound professional work and may be referenced in the formal scientific and technical literature.

NOAA Technical Memorandum NMFS

This TM series is used for documentation and timely communication of preliminary results, interim reports, or special purpose information. The TMs have not received complete formal review, editorial control, or detailed editing.

NOVEMBER 2011

EFFECTIVE STRIP WIDTHS FOR SHIP-BASED LINE-TRANSECT SURVEYS OF CETACEANS

Jay Barlow, Lisa T. Ballance, and Karin A. Forney

Protected Resources Division, Southwest Fisheries Science Center, National Marine Fisheries Service,
National Oceanic & Atmospheric Administration, 3333 N. Torrey Pines Ct., La Jolla, CA 92037

NOAA-TM-NMFS-SWFSC-484

U.S. DEPARTMENT OF COMMERCE
Gary F. Locke, Secretary
National Oceanic and Atmospheric Administration
Jane Lubchenco, Undersecretary for Oceans and Atmosphere
National Marine Fisheries Service
Eric C. Schwaab, Assistant Administrator for Fisheries

Effective Strip Widths for Ship-based Line-transect Surveys of Cetaceans

Jay Barlow[1], Lisa T. Ballance[1] and Karin A. Forney[2]

[1]NOAA Southwest Fisheries Science Center
3333 N. Torrey Pines Court, La Jolla CA 92037

[2]NOAA Southwest Fisheries Science Center
110 Shaffer Rd, Santa Cruz, CA 95060

Abstract

Effective strip width is a key parameter in estimating abundance and density from line-transect surveys. Here we estimate effective strip widths for 58 categories (genera, species, subspecies, stocks, or other groups) of cetaceans based on 13,500 sightings from 32 line-transect surveys conducted in the eastern Pacific Ocean by the Southwest Fisheries Science Center from 1991 to 2008. Generalized linear models (GLMs) are used to first identify factors that are important in determining the perpendicular sighting distances using stepwise model selection based on AIC. Six species groups of similar taxa are created and modeled separately. Important factors for most species groups include Beaufort sea state, swell height, visibility, group size, species, and a survey-specific categorical variable (*Cruise#*). Interactions between species and the other factors generally do not improve GLM fits, indicating that the effects of those factors are relatively consistent for species within a species group. Factors selected for the best-fit GLMs are included as potential covariates in a line-transect model fit to a subset of the same data, again using stepwise model selection based on AIC. The best-fit line-transect models do not include *Cruise#* and are generally simpler than the GLMs, likely because distant sightings were eliminated by truncation. Species-specific differences in ESW are seen within the species groups indicating that species of cetacean do differ in the distances at which they can be detected, even after accounting for the effects of group size and other covariates that affect sighting distances. Results from this analysis of multiple surveys can be used to improve estimates of effective strip widths for any survey using the same methods and similar vessels. This is especially true for seldom-seen species whose abundance is difficult to estimate from a single survey.

Introduction

Line-transect surveys have been used extensively to estimate the abundance of cetaceans worldwide. One of the key parameters in estimating abundance from line-transect surveys is the effective strip width, which provides a measure of how far animals are seen from the transect line and, therefore, how much area is effectively searched. The effective strip width is estimated as the inverse of a probability density function fit to the distribution of perpendicular sighting distances and evaluated at zero distance from the transect line (Buckland et al. 2001). Many factors have been found to affect estimates of effective strip width, including characteristics of the species being surveyed, search methods, the search platform, and sea and weather conditions at the time of the survey (Barlow et al. 2001; Gerrodette and Forcada 2005). Most estimates of cetacean abundance from line-transects have been based on simple, single-species analyses of a single survey and have not attempted to determine how effective strip width is affected by factors other than species and, in some cases, sea state. Because line-transect methods are robust to pooling data collected under a variety of conditions, this approach is reasonable and will typically give unbiased estimates of density if the assumptions of line-transect sampling are met (Buckland et al. 2001).

Recently, however, a multiple-covariate method of line-transect analysis has been developed that allows for explicit incorporation of factors that could affect effective strip width in fitting models to distributions of perpendicular sighting distances (Marques and Buckland 2003). This multiple-covariate approach can improve line-transect estimates by explaining some of the variation in effective strip widths. Prior to the development of this covariate approach, stratification was used to explain some of the variance associated with estimating effective strip widths and thereby improve those estimates (Barlow 1995). However, sample sizes from single surveys are often inadequate to apply either the multiple-covariate or stratification approaches to individual species (Barlow 2006; Zerbini et al. 2006). Ultimately, most estimation problems involve a trade-off between reducing bias (which favors more complicated models with more parameters) and increasing precision (which favors fewer parameters) (Burnham and Anderson 1998). Adding covariates and strata increases the number of estimated parameters, and the overall accuracy of estimates will depend upon having sufficient sample size to estimate those extra parameters.

Prior research has investigated the affect of covariates on detection distances and effective strip widths by pooling results from multiple surveys of multiple cetacean species. Barlow et al. (2001) used Generalized Additive Models (GAMs) to determine which factors are important in determining perpendicular sighting distances on cetacean line-transect surveys in the eastern Pacific. In that study, perpendicular distances were not truncated at a maximum distance (as is commonly done in line-transect analysis), and no attempt was made to determine whether the same factors are important in multiple-covariate line-transect models. They showed that a better fitting model could be obtained by pooling species into 13 categories rather than by estimating an independent effect for each of 34 species and showed that the specific ship used and survey year were important explanatory variables in some models. Zerbini et al. (2006) used multiple covariate line-transect models to estimate the abundance of whales in Alaska and also found that it was important to include the ship used for the survey as covariate in the best-fit model for two of three species. In that study, however, survey year was not included as a potential covariate. Barlow (2006) also used a multiple-covariate line-transect model but used stepwise model building to add covariates in estimating the density of cetacean species from a survey of Hawaiian waters. Sample sizes were small, so Barlow (2006) pooled this survey with many previous surveys in estimating effective strips widths. He found that the ship used was not an important predictor in the best-fit model for any species once the effects of group size and sightings conditions were included, but again the effect of different survey years was not investigated. Barlow and Forney (2007) used step-wise model building and model averaging in a multiple-covariate framework to estimate the abundance of cetacean species along the U.S. West Coast. They found that the ship used was an important covariate in some models for some species, but that survey year was not an important covariate (after the effects of ship, group size and survey conditions were included). All of these previous studies identified survey conditions (especially Beaufort sea state) as important in modeling variation in perpendicular sighting distances and effective strip widths in cetacean line-transect surveys.

Although these prior studies have been useful in identifying covariates that are important to include in multiple-covariate line-transect analyses, most of these have been focused on the specific objective of estimating the abundance of particular cetacean species in a particular area from a given set of surveys. Consequently, several generic questions have remained unanswered. Is it better to pool species with similar sighting characteristics or to model all species separately? Is it better to pool multiple surveys or to estimate abundance separately from each survey? Also, these prior studies have only considered linear (or log-linear) covariates and have not explicitly examined the form of the functional relationship between covariates and effective strip with in line-transect studies. Finally, these prior studies have not given the values for the coefficients estimated in multiple-covariate line-transect models, so others cannot estimate effective strip widths for novel survey data. A more detailed analysis of a large quantity of line-transect survey data is needed to address the shortcomings of prior analyses and to answer these more general questions.

In this paper we conduct detailed analyses of perpendicular sighting distance and effective strip widths from 32 surveys conducted by the Southwest Fisheries Science Center (SWFSC) in the eastern and central Pacific from 1991 to 2008. We use the power and flexibility of Generalized Linear Models (GLMs) to fit models of perpendicular sighting distance to determine which covariates are important, to determine the shape of the functional relationships, and to test for interaction effects. We then use stepwise model building to determine which covariates are most important in estimating effective strip widths in multiple-covariate line-transect models. We use polynomial functions to determine the function form of the relationships between covariates and measures of search distance in both approaches. We empirically explore whether line-transect estimates of cetacean abundance can be improved by pooling multiple species and surveys using the multiple covariate approach. Specifically, we investigate whether some of the factors that determine effective strip widths (such as sea state or group size) are common among species and surveys and can be more effectively estimated using this pooled approach. Akaike's Information Criterion (AIC) is used to select the best GLM models and to determine whether the line-transect model fits are improved by pooling species and surveys when estimating the covariate factors that affect sighting distances. Based on this much larger sample size, we show that survey conditions and group sizes are important covariates for most species and that species-specific effects are important within most species groups. However, we show that differences among surveys are often small (after controlling for the above factors), and pooling multiple surveys is often appropriate. We give the coefficients that were estimated in the multiple-covariate line-transect models so that others can estimate effective strip widths for their surveys conducted with similar methods.

Methods

Field Data Collection

The SWFSC has been using the same basic method for ship-based line-transect sampling since the early 1980s. Kinzey et al. (2000) describe these methods in detail. In brief, three experienced marine mammal observers search from the flying bridge deck of research ships. The ships follow pre-determined transect lines that are designed to systematically or representatively sample a defined study area. Two observers search using pedestal-mounted 25X Fujinon binoculars. The third observer records all data on a computer and also searches using unaided eyes and (occasionally) 7X binoculars. Survey conditions (Beaufort sea state, swell height and visibility) are recorded every 30-40 minutes or whenever conditions change. When cetaceans are seen within 3 nmi of the transect line, the ship is maneuvered to approach the animals so that the observers can better determine the species present and estimate the group size. Each observer makes independent estimates of group size and, when species occur in multi-species groups, the proportion of each species present.

SWFSC cetacean surveys from 1991-2008 have covered much of the eastern and central North Pacific Ocean and a small area of the eastern South Pacific Ocean (Fig. 1). The most frequently surveyed areas are the eastern tropical Pacific off Mexico and Central America and the California Current off the U. S. West Coast. Surveys have also included the Gulf of California, Alaska, the eastern temperate North Pacific, and the U.S. Exclusive Economic Zones around the Hawaiian Islands, Palmyra Atoll & Kingman Reef, and Johnston Atoll. Surveys often included more than one ship for multiple months. The largest-scale surveys were those in the eastern tropical Pacific, which included at least two ships for four months each. Most surveys were conducted on three National Oceanographic and Atmospheric Administration (NOAA) research ships: the *David Starr Jordan*, the *McArthur*, and the *McArthur II*. The University of Rhode Island ship *Endeavor* was used for one eastern tropical Pacific survey and the NOAA ship *Surveyor* was used for one Alaska survey. Data from each ship on each survey were assigned a unique cruise number.

Data Pre-processing

Line-transect data from all 1991-2008 SWFSC ship-based surveys were pooled and information on all sightings of marine mammal was extracted. The primary covariates used in this analysis included conditions at the time of the sightings: Beaufort sea state (*Beauf*, an ordinal integer variable taking values in our study from zero to six), swell height (*SwellHght*, estimated in feet from trough to crest), and visibility (*Vis*, the maximum distance in nmi at which observers estimated that they could see a dolphin, truncated at 6 nmi or roughly the distance to the horizon). *SwellHght* is correlated with *Beauf* (Pearson's R = 0.41), so swell height was expressed as the swell anomaly (*SwellAnom*, Barlow et al. 2001) or the deviation from the expected swell height for a given sea state:

$$SwellAnom = SwellHght - 2.19 - (0.5295 * Beauf) \quad .$$

Additional covariates included the identity of the most abundant species within a group (*Species*) and the natural logarithm of the total size of the group (*GroupSize*) which reflects the total number of individuals for all species present. Observers are known to, on average, underestimate the size of cetacean groups (Gerrodette and Forcada 2005), so calibration coefficients for individual observers were used to correct estimates of group size (Barlow and Forney 2007). Additional covariates included cruise number (*Cruise#*, a unique number assigned for a specific survey by a given ship in a given year), *Ship* (the ship from which the sighting was made), *Year* (the year of the survey as a categorical variable), and *Region* (one of seven geographic areas – Figure 1). A small number of sightings were excluded because group size was not estimated.

Five a-priori species[1] groupings were used for most analyses (Table 1): porpoises (including the genera *Phocoena* and *Phocoenoides*), delphinids (all species except killer whales), large odontocetes (including killer whales, sperm whales, and two large beaked whale species), small whales (including minke whales, dwarf and pygmy sperm whales, and small beaked whales in the genera *Ziphius* and *Mesoplodon*), and large mysticetes (all baleen whales except minke whales). The exceptions to purely taxonomic groupings were made to link species with similar sighting characteristics. Killer whales typically have a distinct blow and can be seen at greater distances than other delphinids (Barlow et al. 2001) and were grouped with other similar species. Similarly, Baird's and Longman's beaked whales typically have distinct blows and occur in larger groups than other beaked whales and were also grouped with large odontocetes. Minke whales surface cryptically, typically without a visible blow and were grouped with other taxonomically unrelated species with similar cryptic habits. Delphinids were further subdivided into large and small delphinids (Table 1) for some analyses, based on an adult length of approximately 2.4-2.6 m. Small delphinids typically occur in much larger groups than large delphinids. Species with adult lengths in the range of 2.4-2.6 m were grouped with either large or small delphinids depending on whether their mean group sizes were small (<40 individuals) or large (respectively).

Generalized Linear Model Selection

We used GLMs to identify the variables that affect perpendicular sighting distances within each of the five species groups, to determine the functional shape of relationships between ordinal variables and perpendicular distance, and to look for interaction effects between *Species* and the other variables. We modeled perpendicular distance as a function of categorical variables and polynomial functions fit to continuous and ordinal variables. We used the square of perpendicular distance as the dependent variable to approximate a maximum likelihood fit of an un-truncated half-normal distribution (the mean squared deviation from zero distance is the variance which defines the half-normal). We expect factors that affect sighting distance to be multiplicative so we used a logarithmic link function, and we used a gamma distribution of residuals (as is commonly used to describe the distribution of a variance). A constant (0.5 km) was added to perpendicular distances to avoid attempting to take the logarithm of a negative number (Barlow et al. 1999). Although we limited our analyses to linear models without smoothing functions, we used the *step.gam*[2] function in the R package *gam* (R Core

[1] Note that all species names are given in Table 1.
[2] Package 'gam' built under R version 2.12.2.

Development Team)to fit our GLM because of the function's efficient stepwise model-selection algorithm.

The forward-backward stepwise model selection algorithm in *step.gam* was used to select the covariates that provide the best fit using AIC. The model was initialized with all categorical variables (*Species* and *Cruise#*) and with linear terms for all continuous and ordinal variables (*GroupSize*, *SwellAnom*, *Visibility*, and *Beauf*). The use of *Species* as categorical variable within a model allows for differences in sighting characteristics for different species within a species group that are not captured by the other variables (e.g. species-specific behavior). At each step, the program evaluates whether AIC could be lowered by adding or deleting any of the categorical variables or by increasing or decreasing the polynomial order of the continuous or ordinal variables (a zero-order polynomial eliminated the variable from the prediction). The next step of the model was initialized with the single-variable change that lowered AIC by the greatest amount. The model selection continued until AIC was minimized. A maximum of four degrees of freedom was allowed for polynomial fits.

To determine whether the best-fit model might differ among species within a group, interaction effects were estimated between *Species* and each of the other factors that were included in the best-fit model. The potential for interaction effects between *Species* and other factors was also examined by exploring alternative subgroups within the delphinid species group. The conventional Cp statistic in R (which is closely related to AIC) was used to determine whether interaction effects improved GLM fits.

Multiple-covariate Line-Transect Modeling

Distributions of perpendicular sighting distances were fit with a truncated half-normal distribution with half-normal scale parameter determined as a function of the same covariates used in the GLM model selection (Marques and Buckland 2003). Models were fit in R using LT.fit [3], and a custom stepwise algorithm was used to select the best-fit model. Each model was initialized with all the factors that were included in the best-fit GLM models (again, with linear terms for *GroupSize*, *SwellAnom*, *Vis*, and *Beauf*), and additional covariates and higher-order polynomials were added until AIC could no longer be lowered. Separate line-transect models were fit to the subgroups identified in the GLM analysis of interaction effects. Truncation distances were 3 km for porpoises, 5.5 km for delphinids, large odontocetes, and large mysticetes, and 4 km for small whales; these values are similar to or slightly greater than truncation distances that have been used in previous published analyses of these data. Groups seen beyond these truncation distances were excluded from the analysis. Species or stocks (Table 1) with fewer than 4 observations were also excluded.

Results

GLM Model Selection

[3] LT.fit refers to R code written by Jeff Laake (NOAA National Marine Mammal Lab, Seattle Washington); a later version of this code is currently used in the fitting multiple-covariate line-transect models in the standard line-transect software, *Distance* (Thomas et al. 2010).

Using the GLM approach, the best-fit model for all species pooled included all covariates with second-order polynomials for *Beauf* and *Vis* and a third-order polynomial for *GroupSize* (Table 2). The best-fit GLM for the five main species groups did not always include all covariates or the same degrees of freedom in the polynomials (Table 2). *Species, Beauf,* and *Vis* were included in all best-fit models. When *Cruise#* was omitted and three related variables (*Ship*, *Year*, and *Region*) were substituted, *Ship* was included in all best-fit models except for porpoises and *Region* was included in all best-fit models except for large odontocetes. *Year* was included in the best-fit models only for large odontocetes and small whales. The best-fit models for porpoises, large odontocetes, and small whales did not include *GroupSize*, and the best-fit model for large odontocetes did not include *SwellAnom* or *Cruise#*. As is usually the case, the best-fit models for species groups with lower sample sizes were simpler and contained fewer covariates. AIC values were much higher when all cetaceans were pooled than when species were grouped and modeled separately (delta-AIC = 863); therefore, subsequent analyses used the species groupings. The subgroups of large and small delphinids gave a marginally better fit than that for all delphinids pooled (delta-AIC = 2) and were also used in subsequent GLM analyses.

The addition of interaction effects between *Species* and the other individual covariates improved the model fit in only three instances. For small delphinids, adding an interaction term between *Species* and *SwellAnom* improved the model very slightly (delta-Cp = 0.4), and for large delphinids and large odontocetes, adding an interaction term between *Species* and *Beauf* improved the models (delta-Cp = 7.2 and 10.2, respectively). For large delphinids, the interaction coefficient for *Species* and *Beauf* was significantly different from zero ($\alpha = 0.01$) only for short-finned pilot whales. For the other species groups whose best-fit model included an interaction term, none of the interaction coefficients for individual species were significantly different from zero ($\alpha = 0.01$).

Multiple-covariate Line-Transect Modeling

The best-fit line-transect model for most of the species groups included *Beauf*, *SwellAnom*, *Vis*, *GroupSize*, and *Species* covariates (Table 3). However, *Species* was not included in the best-fit model for large odontocetes or small whales, *GroupSize* was not included for porpoises, and *SwellAnom* and *Vis* were not included for large odontocetes. Contrary to the results of the GLM model fitting, *Cruise#* was not included in the best-fit models for any of the species groups. Some of the other covariates that were confounded with *Cruise#* were included in the best-fit models for small delphinids (*Region*), large delphinids (*Year*), porpoises (*Year*), and large mysticetes (*Ship*). The separate species groups of small and large delphinids gave best-fit models with a combined AIC that was much lower than for the pooled category of all delphinids (delta-AIC = 46).

The marginal, multiplicative effects of the ordinal covariates are illustrated in Figs 2a-d. The direction of most of the effects agrees with expectations. Sighting distances generally decreased in increasing sea state and swell height and increased with increasing visibility and group size. The sea state effect did not monotonically decrease for one species group (large mysticetes, Fig. 2a), but all values are near 1.0, indicating that sea state has little effect on perpendicular sighting distances for this group of species. All estimated coefficients are given in Appendix 2.

Discussion

GLM for Exploratory Line-transect Modeling

Generalized linear modeling is a powerful model-selection tool to evaluate which covariates are important when modeling line-transect data. The forward/backward model selection approach using AIC is built into most GLM packages. GLM also provides a convenient framework for evaluating interaction effects and for determining the shape of functional relationships for non-factor variables. This analysis showed that all the tested covariates (Beaufort sea state, visibility, swell anomaly, group size, species, and cruise number) were important in describing the observed perpendicular sighting distances for most species groups. However, interactions between these covariates and species were rare, indicating that the species groupings captured most of the differences between species. The variables *Ship, Year,* and *Region* were important when cruise number was excluded.

We used GLMs instead of GAMs (as was used by Barlow et al. 2001) because multivariate line-transect models are currently limited to linear models, and we used GLM to predict which covariates would be important in a multivariate line-transect model. Our approach also differs from that of Barlow et al. (2001) in modeling perpendicular distance squared with a log-link function instead of the logarithm of perpendicular distance with an identity link function and in using a gamma error distribution instead of a Gaussian distribution. Despite these differences, results were generally similar. Barlow et al. (2001) found that for 1991-96 data, Beaufort sea state, swell anomaly, group size, species, cruise number and region were important in determining perpendicular sighting distance.

Barlow et al. (2001) tested additional covariates and found that the search method (25X binoculars vs. naked eye & 7X binoculars), the sighting cue, and the observer making the sighting also affected estimates of perpendicular sighting distance. We deliberately did not use these covariates because we do not believe these to be useful covariates for line-transect analysis. On our surveys, observers search with 25X binoculars or 7X binoculars & unaided eyes, depending on their observation station. These methods are not independent because a group first seen by one method is not available to be first seen by another. Similarly, observers search in teams of three, and a group first seen by one observer is not available to be first seen by another. Finally, cetacean groups often display many types of behavior which can lead to ambiguity in designating a single sighting cue.

Multiple-covariate Line-Transect Modeling

The best-fit models using the multiple-covariate line-transect approach generally contained fewer covariates and lower-order polynomials than the GLMs. Most importantly, cruise number was not included in any of the line-transect models. This result is important because it means that data from multiple cruises can be pooled to estimate line-transect parameters, which can greatly increase the available sample size for estimating effective strip width and improve the precision of cetacean abundance estimates. However, all the ships used on these surveys were of similar size, and many of the same observers were used on multiple surveys, so this result may not apply generally when these conditions are not met.

The major difference between the data used for the GLM and the line-transect modeling is the exclusion of sightings beyond a truncation distance in the latter. This difference likely results in the simpler models selected for the line-transect approach and the exclusion of cruise number in those models. Differences in the likelihood of seeing distant groups of cetaceans makes very little difference in line-transect estimates of cetacean density because these groups are beyond the truncation distance. It is likely that the extra parameters selected in the GLM analysis are only needed to explain differences in the tails of the distributions of perpendicular sighting distances. However, the use of a gamma error distribution in the GLM models also could explain the GLM and line-transect modeling approaches.

The estimated coefficients for the selected covariates can be used to estimate perpendicular sighting distances given the appropriate covariates for any given sighting (see Appendices 1-3). The sum of the intercept and the products of coefficients times the covariates (or dummy variables) is equal to the natural logarithm of the standard deviation of the half-normal distribution that defines the sighting probability density.

Beaufort Sea State

Beaufort sea state is a subjective measure of wind speed as judged by the visible effects of the wind on the sea surface. This factor was included in all the best-fit GLM and line-transect models (Tables 2 & 3). Many previous analyses have shown that Beaufort sea state is an important factor affecting the ability to see cetaceans at sea. This effect is strongest for the species that are hardest to see, e.g. porpoises and small whales (Fig 2a). Although the factor was included for large mysticetes (a group that is easy to see in rough seas because of their tall blows), the effect is small and is not monotonically decreasing as it is for all other species groups.

Swell Height Anomaly

Swell height is a result of winds blowing over the surface, but swells can travel great distances with little loss of energy. The local swell height is therefore the result of both local winds and distant storms. The effect of local winds on sighting distances is already included in Beaufort sea state effects, but the effects of swells from distance storms is not. Our use of swell height anomaly effectively estimates the effect of swell height on sighting distances over-and-above the effect expected for a given Beaufort sea state. Swell height anomaly was included in the best-fit GLM and line-transect models for all species groups except for large odontocetes (Tables 2 & 3), and greater swell heights resulted in lower probabilities of detection with distance (Fig 2b). The effect of swell height is again greatest for the species that are hardest to see (small whales and porpoises) and for large delphinids. It is not clear why swell height anomaly was not important in models for large odontocetes, but this group includes sperm whales and killer whales, which are visible at great distances.

Visibility

Visibility is a subjective appraisal of the distance at which an observer could see a leaping dolphin when looking through 25X binoculars. Visibility is typically limited by haze, fog, and rain. Visibilities of 1-2 nmi generally indicate light fog. Search effort is usually discontinued in heavy fog. Visibility is included in the best-fit GLM and line-transect models for all species groups (Table 2 & 3). In all cases, sighting distances decreased as visibility decreased.

Group Size

Larger groups of animals are clearly easier to see at distance than smaller groups. Not surprisingly, group size (expressed as the natural log of group size) was included in the best-fit models for most species (Tables 2 & 3). Group size had the largest effects for delphinids, which occur in the largest range of group sizes (from solitary individuals to over 1,000 animals). Group size was not included in the best-fit GLM models for porpoises or for small whales (both of which characteristically occur in small groups) or for large odontocetes.

Species

Within a species group, the *Species* effect coefficients in Table 4 indicates whether each species is typically seen at greater or lesser distances than the average species within that group after controlling for the effect of different characteristic group sizes. Values greater than 1.0 indicate that species is typically seen at greater distances and values less than 1.0 indicate that they are typically seen at lesser distances. *Species* is an important factor in determining sighting distances within all species groups for the GLM models (Table 2) and within most species groups for the line-transect models (Table 3).

Within each species group, the greatest coefficients for the *Species* effect are for various categories of unidentified species (e.g. unidentified dolphin, unidentified small delphinid, unidentified large delphinid, etc.). This "unidentified" effect is because groups that are seen at greater distances are less likely to ever be identified. Typically a vessel needs to approach to within 1 nmi in order to reliably identify cetacean species and much closer than that for many species. If a group is seen at great distance, it is likely that its species composition will never be determined with certainty and will be recorded as one of the categories of unidentified species.

Similar and closely related species within a species group often have very different coefficients for the *Species* effect. For example, among small delphinids, the Central American subspecies of spinner dolphin (*Stenella longirostris centroamericana*) has a coefficient of 1.32 and the closely related stock of "Tres Marias" spinner dolphin has a coefficient of 0.68. This disparity is likely due to the uncertainty that results from small sample sizes (5 and 8, respectively); however the short-beaked and long-beaked species of common dolphins are both represented by large samples and have quite different coefficients (0.86 and 0.63, respectively). Differences that are due only to small sample size variation could be addressed by combining species or stocks with similar sighting characteristics, for example combining stocks of spinner dolphins. This pooling approach would also allow inclusion of rare stocks (with fewer than 4 sightings) which were excluded from this analysis. However, at least some of the differences in *Species* coefficients within a species group are likely related to real behavioral differences that make them more or less visible and hence more or less detectable at great distances. Careful consideration will need to be given to pooling species to avoid masking these real differences.

Cruise Number, Region, Year, and Ship Effects

In the GLM analysis, cruise number was an important explanatory variable for all species groups except large odontocetes, the group with the smallest sample size. *Cruise#* specifies a *Ship*, a *Year* and one or more *Regions*. Different ships can have observation decks at different heights and can have different tendencies to roll or vibrate, all of which can affect the distance at which cetaceans can be seen (Barlow et al. 2001; Gerrodette and Forcada 2005). *Year* specifies a particular group of observers who can have different abilities to detect distant groups of cetaceans. On surveys in the ETP, observers typically changed ships halfway through a cruise. Different *Regions* can have differences in sighting conditions that are not captured by the *Beauf* and *SwellAnom* variables. *Cruise#* can therefore act as a composite covariate that explains differences in sighting distances due to a wide variety of factors that were not explicitly included in the other covariates. When *Cruise#* was explicitly excluded, the best-fit GLMs typically included only *Ship* and *Region* (Table 2), indicating that these may be more important than *Year* as an explanatory variable. However, for most species groups, GLMs including *Cruise#* gave a better fit to the data (lower AIC) than models that included the optimal combination of *Region*, *Year*, and *Ship* (Table 2).

In contrast with the GLM analysis, the line-transect analysis did not include *Cruise#* in the best-fit model. Instead, only one or two of the related variables (*Region*, *Year*, and *Ship*) was usually selected (Table 3). This may indicate that the differences among different cruises, regions, years, and ships are subtle and are only apparent at sighting distances that are greater than the typical truncation distances used in line-transect analyses.

Other Potential Covariates

Barlow et al. (2001) examined the effects of additional covariates and found that some were included in the best-fit GAM for perpendicular sighting distance. These included the type of binocular that was used for making a sighting (*BinoCode* being either 25x binocular or naked eye/7x binocular), the actual cue that first was first detected (*Cue* being either bird flocks, splashes, blows and the body of the animal itself), the presence of glare on the waters directly ahead of the survey vessel (*Glare*), and the specific person who first detected the animals (*Observer*). *Glare* was also included as a potential covariate in the multiple-covariate analyses of Barlow (2006) and Barlow and Forney (2007), and *BinoCode* was also used by Barlow and Forney (2007). Although these covariates are sometimes found in best-fit models, we did not use them in our analysis for a variety of reasons. *Glare* was included in the best-fit model for only one of 20 models in Barlow (2006) and added almost imperceptibly to the best-fit model of Barlow and Forney (2007); *Glare* was excluded from the current analysis because it does not appear to add much other than additional complexity. *Cue* was excluded from the present analysis because, similarly, it was included in only one of nine models in Barlow and Forney (2007) and does not appear to be a very important covariate. *BinoCode* is always an important covariate whenever it is included (cetaceans can be seen at much greater distances with 25x binoculars), but, in hindsight, this covariate was found in increase rather than decrease the overall variance of line-transect density estimates. During SWFSC surveys, two people are always searching with 25x binoculars and one person (a data recorder) is always searching with naked eyes and (intermittently) with a 7x binocular. Most cetaceans are detected by 25x binoculars, but occasionally a group will escape detection until they are close enough to be seen by the data recorder. When that happens and data are analyzed using a Horvitz-Thompson estimator, the density for that species in that effort segment is estimated to be very high because the effective strip with for an observer searching with naked eyes and a 7x binocular is very narrow. However, it was likely just random chance that the observers using 25x binoculars missed the group, and the effective strip width for the entire search team would be much wider than the search width for the data recorder. When sample sizes are low, using *BinoCode* as a covariate can, by random chance, result in a greatly exaggerated estimate of density for some species. Finally, we did not use *Observer* as a covariate because each observer works as part of a team of three people. A cetacean group that is first detected by one person is not available to be first detected by another person. It would therefore be more appropriate to include the team of people as a covariate. However, the same team seldom worked together on different surveys, so this effect would be difficult to estimate. Including an effect for *Cruise#* effectively controls for differences in skill level among teams of observers on different surveys.

Correlations and Co-linearity

The interpretation of effects and coefficients is problematic whenever analyzing opportunistic data that were not collected with a factoral sampling design. It is often difficult to tease apart effects when covariates are highly correlated. In some cases, it is possible to express one variable as deviations from the expected value given another variable as we did with swell height and Beaufort sea state. However, this approach is not always possible. For example, swell heights are generally lower in the Gulf of California than in any of the other areas we surveyed, so the effect that should be seen in *SwellAnom* could be expressed in the coefficient for *Region = Gulf of California* instead. Similarly, different suites of species are found in different *Regions*, and *Species* effects can be confounded with *Region* effects. The best way to tease apart real factors from correlated factors is to base analyes on large quantities of data collected over a wide range of conditions, as we did. Negative *SwellAnom* values can be found in *Regions* with the roughest average seas if surveys are conducted over multiple years. Fortunately, most species are found in multiple areas, allowing separation of *Species* and *Region* effects. Although ours is arguably the largest data set for line-transect analysis, we still cannot be certain that estimated coefficients accurately reflect only the effect to which they are attributed.

Conclusion

Our analyses allow the estimation of effective strip width for a broad range of cetacean species in a wide range of survey conditions. Survey-specific information (*Region*, *Year* and *Ship*) was found to be of secondary importance relative to generic survey conditions (*Beauf*, *SwellAnom*, and *Vis*). The coefficients estimated here for the line-transect model can therefore be used to estimate effective strip widths for past surveys or even for a novel survey. Obviously the same survey methods would need to be used with a similar-sized survey vessel if these results were used to estimate effective strip width for a novel survey.

The more likely application of these results is in improving estimates of effective strip width for rarely seen species on past surveys. Some species are seldom seen and their density and abundance cannot be estimated because effective strip widths cannot be reliably estimated. In past analyses, similar species have been pooled in order to estimate effective strip width for the rarer species (Mullin and Fulling 2004), sometimes stratifying by group size to compensate for different characteristic group sizes among species (Barlow 1995). This previous approach fails to account for real differences that might exist due to species-specific behaviors. Multiple covariate approaches have been used in the past for single-species estimates of abundance (Gerrodette and Forcada 2005) which are effective for commonly seen species with large numbers of sightings. We have shown, however, that many of the covariates are not species-specific and that species groupings can be used to better estimate these covariates. Real differences between species can be accommodated by including a species-specific effect. Our multiple-covariate approach and estimated coefficients allow calculation of effective strip widths that are specific for observed survey conditions and for individual species. This approach will also allow line-transect-based habitat models to include area effectively searched as an offset (the product of effective strip width and transect length) instead of the previous approach of only using the transect length (Becker et al., in prep.).

Future Research

We believe that our multiple-covariate approach to estimating effective strip widths from pooled surveys will allow more accurate estimates of cetacean abundance than previous methods. However, additional research is needed in a few key areas. Most importantly, we were not able to develop models for a few very rare species (those with fewer than 4 sightings on all past SWFSC surveys) including Longman's, Hubb's, and Stejneger's beaked whales. By judiciously pooling species with similar sighting characteristics, models of effective strip width could be developed for more species. The species effect coefficients for several rarely seen species or stocks were widely disparate compared to other similar species, likely due to small sample size. These groups might also benefit from pooling with similar species. Finally, the GLM analysis identified three interaction effects between *Species* and sighting conditions for delphinids (*Beauf* and *SwellAnom*) and large odontocetes (*Beauf)*. The line-transect models might also be improved by including these interaction effects.

Acknowledgments

The data analyzed here were collected in often difficult conditions by a large team of dedicated observers, cruise leaders, survey coordinators, officers and crewmen, and all deserve a great deal of credit. The program LT.fit was graciously provided by Jeff Laake. This paper was improved by reviews by Amanda Bradford, Elizabeth Becker, Jeff Laake, and William Perrin.

Literature Cited

Barlow, J. 1995. The abundance of cetaceans in California waters. Part I: Ship surveys in summer and fall of 1991. Fishery Bulletin 93:1-14.

Barlow, J. 1999. Trackline detection probability for long-diving whales. pp. 209-221 In: G. W. Garner, et al. (eds.), Marine Mammal Survey and Assessment Methods. Balkema Press, Netherlands. 287pp.

Barlow, J. 2006. Cetacean abundance in Hawaiian waters estimated from a summer/fall survey in 2002. Marine Mammal Science 22(2):446-464.

Barlow, J. and K. A. Forney. 2007. Abundance and population density of cetaceans in the California Current ecosystem. Fishery Bulletin 105:509-526.

Barlow, J., T. Gerrodette, and J. Forcada. 2001. Factors affecting perpendicular sighting distances on shipboard line-transect surveys for cetaceans. J. Cetacean Res. and Manage. 3(2):201-212.

Becker, E. A., K. A. Forney, D. G. Foley, and J. Barlow. (in prep.). Density and spatial distribution patterns of cetaceans in the central North Pacific based on habitat models.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling: Estimating abundance of biological populations. Oxford University Press, Oxford. 432pp.

Burnham, K. P. and D. R. Anderson. 1998. Model selection and inference. A practical information-theoretic approach. Springer-Verlag, New York. 353pp.

Gerrodette, T., and J. Forcada. 2005. Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. Marine Ecology Progress Series 291:1-21.

Kinzey, D., P. Olson, and T. Gerrodette. 2000. Marine mammal data collection procedures on research ship line-transect surveys by the Southwest Fisheries Science Center. Southwest Fisheries Science Center Administrative Report. LJ-00-08. Available from the Southwest Fisheries Science Center, 8604 La Jolla Shores Dr., La Jolla CA, USA.

Marques, F. C. and S. T. Buckland. 2003. Incorporating covariates into standard line transect analysis. Biometrics 59:924-935.

McCullagh, P. and J.A. Nelder. 1989. Generalized Linear Models. Chapman and Hall: London. 511p.

Mullin, K. D. and G. L. Fulling. 2004. Abundance of cetaceans in the oceanic northern Gulf of Mexico, 1996-2001. Marine Mammal Science. 20:787-807.

Thomas, L., S.T. Buckland, E.A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R.B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology 47: 5-14.

Zerbini, A. N., J. M. Waite, J. L. Laake, and P. R. Wade. 2006. Abundance, trends and distribution of baleen whales off Western Alaska and the central Aleutian Islands. Deep Sea Research Part I: Oceanographic Research Papers. 53(11):1772-1790.

Table 1. Species groups used in this report with their scientific and common names and the species/subspecies/stock codes used by the SWFSC.

Species Group	Species Code	Scientific Name	Common Name
Small Delphinids	002	*Stenella attenuata attenuata*	Pantropical spotted dolphin
	006	*Stenella attenuata graffmani*	Coastal spotted dolphin
	090	*Stenella attenuata* (unid. subsp.)	Unidentified spotted dolphin
	003	*Stenella longirostris* (unid. subsp.)	Unidentified spinner dolphin
	010	*Stenella longirostris orientalis*	Eastern spinner dolphin
	011	*Stenella longirostris* (whitebelly)	Whitebelly spinner dolphin
	088	*Stenella longirostris centroamericana*	Costa Rican spinner dolphin
	100	*Stenella longirostris* (Tres Marias)	Tres Marias spinner dolphin
	101	*Stenella longirostris* (southwestern)	Southwestern spinner dolphin
	102	*Stenella longirostris* (Gray's)	Pantropical spinner dolphin
	103	*Stenella longirostris orientalis/centroamericana*	Undetermined eastern or Central American spinner dolphin
	013	*Stenella coeruleoalba*	Striped dolphin
	005	*Delphinus* sp.	Unidentified common dolphin
	016	*Delphinus capensis*	Long-beaked common dolphin
	017	*Delphinus delphis*	Short-beaked common dolphin
	022	*Lagenorhynchus obliquidens*	Pacific white-sided dolphin
	025	*Lagenorhynchus obscurus*	Dusky dolphin
	026	*Lagenodelphis hosei*	Fraser's dolphin
	027	*Lissodelphis borealis*	Northern right whale dolphin
	031	*Peponocephala electra*	Melon-headed whale
	077	Unidentified dolphin	Unidentified dolphin or porpoise
	177	Unidentified small delphinid	Unidentified small delphinid (*Delphinus, Lagenorhynchus, Lissodelphis or Stenella*)
Large Delphinids	015	*Steno bredanensis*	Rough-toothed dolphin
	018	*Tursiops truncatus*	Bottlenose dolphin
	021	*Grampus griseus*	Risso's dolphin
	032	*Feresa attenuata*	Pygmy killer whale
	033	*Pseudorca crassidens*	False killer whale
	034	*Globicephala* sp.	Unidentified pilot whale
	036	*Globicephala macrorhynchus*	Short-finned pilot whale
	277	Unidentified medium delphinid	Unidentified medium delphinid (*Feresa, Grampus, Steno or Tursiops*)
	377	Unidentified large delphinid	Unidentified large delphinid (*Pseudorca, Orca or Globicephala*)
Porpoises	040	*Phocoena phocoena*	Harbor porpoise
	044	*Phocoenoides dalli*	Dall's porpoise
	477	Unidentified porpoise	Unidentified porpoise (*Phocoena or Phocoenoides*)
Large Odontocetes	037	*Orcinus orca*	Killer whale
	110	*Orcinus orca* (transient)	Transient killer whale
	111	*Orcinus orca* (resident)	Resident killer whale
	112	*Orcinus orca* (offshore)	Offshore killer whale
	046	*Physeter macrocephalus*	Sperm whale
	063	*Berardius bairdii*	Baird's beaked whale
	065	*Indopacetus pacificus*	Longman's beaked whale
Large Mysticetes	066	*Eubalaena japonica*	North Pacific right whale
	069	*Eschrichtius robustus*	Gray whale
	070	*Balaenoptera sp.*	Unidentified rorqual
	072	*Balaenoptera edeni*	Bryde's whale
	073	*Balaenoptera borealis*	Sei whale
	074	*Balaenoptera physalus*	Fin whale
	075	*Balaenoptera musculus*	Blue whale
	076	*Megaptera novaeangliae*	Humpback whale
	079	Unidentified large whale	Unidentified large whale
	098	Unidentified whale	Unidentified whale
	099	*Balaenoptera borealis/edeni*	Rorqual identified as a Sei or Bryde's whale
Small Whales	047	*Kogia breviceps*	Pygmy sperm whale
	048	*Kogia sima*	Dwarf sperm whale
	080	*Kogia sp.*	Unidentified Kogia - dwarf or pygmy sperm whale
	049	Ziphiid whale	Unidentified small beaked whale
	001	*Mesoplodon peruvianus*	Pygmy beaked whale
	083	*Mesoplodon peruvianus*	Pygmy beaked whale originally identified as Species A
	051	*Mesoplodon sp.*	Unidentified *Mesoplodon*
	052	*Mesoplodon carlhubbsi*	Hubb's beaked whale
	059	*Mesoplodon densirostris*	Blaineville's beaked whale
	081	*Mesoplodon stejnegeri*	Stejneger's beaked whale
	109	*Mesoplodon perrini*	Perrin's beaked whale
	061	*Ziphius cavirostris*	Cuvier's beaked whale
	071	*Balaenoptera acutorostrata scammoni*	North Pacific minke whale
	078	Unidentified small whale	Unidentified small whale

Table 2. Sample size (number of sightings), AIC values and covariates selected for the best fit of a Generalized Linear Model to observed perpendicular sighting distances within the indicated groups of cetaceans. For ordinal covariates (*Beauf*- Beaufort sea state, *SwellAnom* – swell anomaly, *Vis* – Visibility, and *GroupSize*), numbers indicate the order of the polynomial that provided the best fit. For categorical covariates (*Species, Cruise#, Region, Year, and Ship*), a check mark indicates that this factor was included in the best-fit model. For all covariates, zero indicates that the covariate was not included in the best-fit model. Best-fit models are given both including *Cruise#* or alternatively including related covariates (*Ship, Year, and Region*), with delta-AIC giving the change in AIC from the model which included *Cruise#* as a potential covariate.

		Best-Fit Generalized Linear Model											
		Including *Cruise#*							**Excluding *Cruise#***				
Species Group	**Sample Size**	**AIC**	***Beauf***	***Swell Anom***	***Vis***	***Group Size***	***Species***	***Cruise#***	**delta-AIC**	***Ship***	***Year***	***Region***
All Cetaceans	13,497	84,306	2	1	2	3	✓	✓	45.8	✓	0	✓
Delphinids	7,751	47,777	2	1	1	3	✓	✓	37.4	✓	0	✓
Small Delphinids	5,764	36,700	2	1	1	3	✓	✓	11.2	✓	0	✓
Large Delphinids	1,987	11,075	1	1	1	1	✓	✓	15.2	✓	0	✓
Sub-total	7,751	47,775										
Large Odontocetes	596	4,379	3	0	2	0	✓	0	-1.2	✓	✓	0
Large Mysticetes	2,747	20,103	2	1	2	2	✓	✓	17.9	✓	0	✓
Porpoises	1,388	6,550	1	1	1	0	✓	✓	-8.8	0	0	✓
Small Whales	1,014	5,360	3	1	1	0	✓	✓	-5.7	✓	✓	✓

Table 3. Sample size (number of sightings), AIC values, truncation distances, and covariates selected for the best fit of a half-normal distribution to observed perpendicular sighting distances within the indicated groups of cetaceans. For ordinal covariates (*Beauf* – Beaufort sea state, *SwellAnom* – swell anomaly, *Vis* – Visibility, and *GroupSize*), numbers indicate the order of the polynomial that provided the best fit. For categorical covariates (*Species, Cruise#, Region, Year, and Ship*), a check mark indicates that this factor was included in the best-fit model. For all covariates, zero indicates that the covariate was not included in the best-fit model.

				Best-Fit Half-normal Line-transect Model								
Species Group	**Sample Size**	**AIC**	**Truncation Distance (km)**	***Beauf***	***Swell Anom***	***Vis***	***Group Size***	***Species***	***Cruise#***	***Region***	***Year***	***Ship***
Delphinids	7,242	21,372	5.5	2	1	1	1	✓	0	✓	0	0
Small Delphinids	5,312	16,083	5.5	1	1	1	1	✓	0	0	✓	✓
Large Delphinids	1,930	5,263	5.5	1	1	1	1	✓	0	0	✓	0
Sub-total	7,242	21,346										
Large Odontocetes	517	1,699	5.5	1	0	1	0	0	0	0	0	0
Large Mysticetes	2,342	7,531	5.5	2	1	1	1	✓	0	0	0	✓
Porpoises	1,261	2,057	3	2	1	1	0	✓	0	✓	✓	0
Small Whales	936	2,049	4	1	1	1	1	0	0	0	0	0

Table 4. Number of groups, mean group sizes, species effects, and mean effective strip widths (ESW) estimated for all species with adequate sample size and all groups within their truncation distances. Group sizes are the mean total size for all groups that had the indicated species as the majority member (including other species that may have been present). Species effects are the multiplicative factors by which the scale parameter of a half-normal distribution varies among species within a group after controlling for other significant covariates. ESWs are based on the best-fit model for each species group and include effects from all of the covariates (Table 3).

Species Group	Species Code	Scientific Name	# Groups	Mean Group Size	Species Effect	Mean ESW km
Small Delphinids	002	*Stenella attenuata*	891	139.5	0.98	3.77
	006	*Stenella attenuata graffmani*	231	91.0	0.81	2.74
	090	*Stenella attenuata* (unid. subsp.)	146	99.8	1.08	3.19
	003	*Stenella longirostris* (unid. subsp.)	15	393.0	1.20	4.15
	010	*Stenella longirostris orientalis*	309	218.6	1.02	4.11
	011	*Stenella longirostris* (whitebelly)	80	284.6	0.73	3.60
	088	*Stenella longirostris centroamericana*	5	887.2	1.32	5.19
	100	*Stenella longirostris* (Tres Marias)	8	473.0	0.68	4.21
	101	*Stenella longirostris* (southwestern)	44	334.9	0.69	3.53
	102	*Stenella longirostris* (Gray's)	8	91.8	0.73	2.56
	103	*Stenella longirostris orientalis/centroamericana*	9	256.6	1.01	4.09
	013	*Stenella coeruleoalba*	1080	57.2	1.10	3.42
	005	*Delphinus* sp.	74	140.0	1.42	3.60
	016	*Delphinus capensis*	122	432.2	0.63	2.88
	017	*Delphinus delphis*	1234	203.1	0.86	3.27
	022	*Lagenorhynchus obliquidens*	124	82.9	0.74	1.69
	025	*Lagenorhynchus obscurus*	26	142.8	0.76	2.61
	026	*Lagenodelphis hosei*	7	190.4	1.66	4.74
	027	*Lissodelphis borealis*	53	91.0	1.03	2.23
	031	*Peponocephala electra*	9	346.9	0.91	4.20
	077	Unidentified dolphin	724	23.5	1.90	3.10
	177	Unidentified small delphinid	113	37.7	2.04	4.05
Large Delphinids	015	*Steno bredanensis*	247	14.2	0.77	2.03
	018	*Tursiops truncatus*	818	36.3	0.91	2.72
	021	*Grampus griseus*	464	35.2	0.81	2.52
	032	*Feresa attenuata*	34	31.8	0.64	1.90
	033	*Pseudorca crassidens*	42	17.0	0.92	2.58
	034	*Globicephala* sp.	40	35.6	1.03	3.07
	036	*Globicephala macrorhynchus*	246	27.1	1.01	2.93
	277	Unidentified medium delphinid	34	11.9	0.98	2.36
	377	Unidentified large delphinid	5	3.2	2.96	4.05
Porpoises	040	*Phocoena phocoena*	106	3.6	0.74	1.35
	044	*Phocoenoides dalli*	1146	4.2	0.81	1.51
	477	Unidentified porpoise	9	2.5	1.67	2.51
Large Odontocetes	037	*Orcinus orca*	172	8.4	1.00	3.81
	046	*Physeter macrocephalus*	304	7.8	1.00	3.88
	063	*Berardius bairdii*	41	11.7	1.00	3.72
Large Mysticetes	066	*Eubalaena japonica*	4	1.7	0.56	2.03
	070	*Balaenoptera sp.*	270	1.7	1.25	3.85
	072	*Balaenoptera edeni*	203	1.7	1.00	3.44
	073	*Balaenoptera borealis*	17	2.9	1.24	3.87
	074	*Balaenoptera physalus*	519	2.8	1.02	3.54
	075	*Balaenoptera musculus*	361	1.9	0.90	3.16
	076	*Megaptera novaeangliae*	573	3.6	1.23	3.94
	079	Unidentified large whale	213	1.8	1.37	4.02
	098	Unidentified whale	52	1.3	0.88	2.95
	099	*Balaenoptera borealis/edeni*	130	1.4	0.85	2.98
Small Whales	047	*Kogia breviceps*	8	1.2	1.00	2.36
	048	*Kogia sima*	152	1.8	1.00	2.16
	080	*Kogia* sp.	18	1.7	1.00	2.13
	049	Ziphiid whale	203	1.8	1.00	1.76
	001	*Mesoplodon peruvianus*	21	2.3	1.00	2.03
	083	*Mesoplodon peruvianus*	14	2.8	1.00	2.02
	051	*Mesoplodon* sp.	59	2.2	1.00	1.83
	059	*Mesoplodon densirostris*	10	2.6	1.00	1.70
	061	*Ziphius cavirostris*	178	2.1	1.00	1.80
	071	*Balaenoptera acutorostrata scammoni*	37	1.1	1.00	1.79
	078	Unidentified small whale	137	3.6	1.00	1.77

Figure 1. Transect lines surveyed by the Southwest Fisheries Science Center (1991- 2008) and geographic strata used in this report.

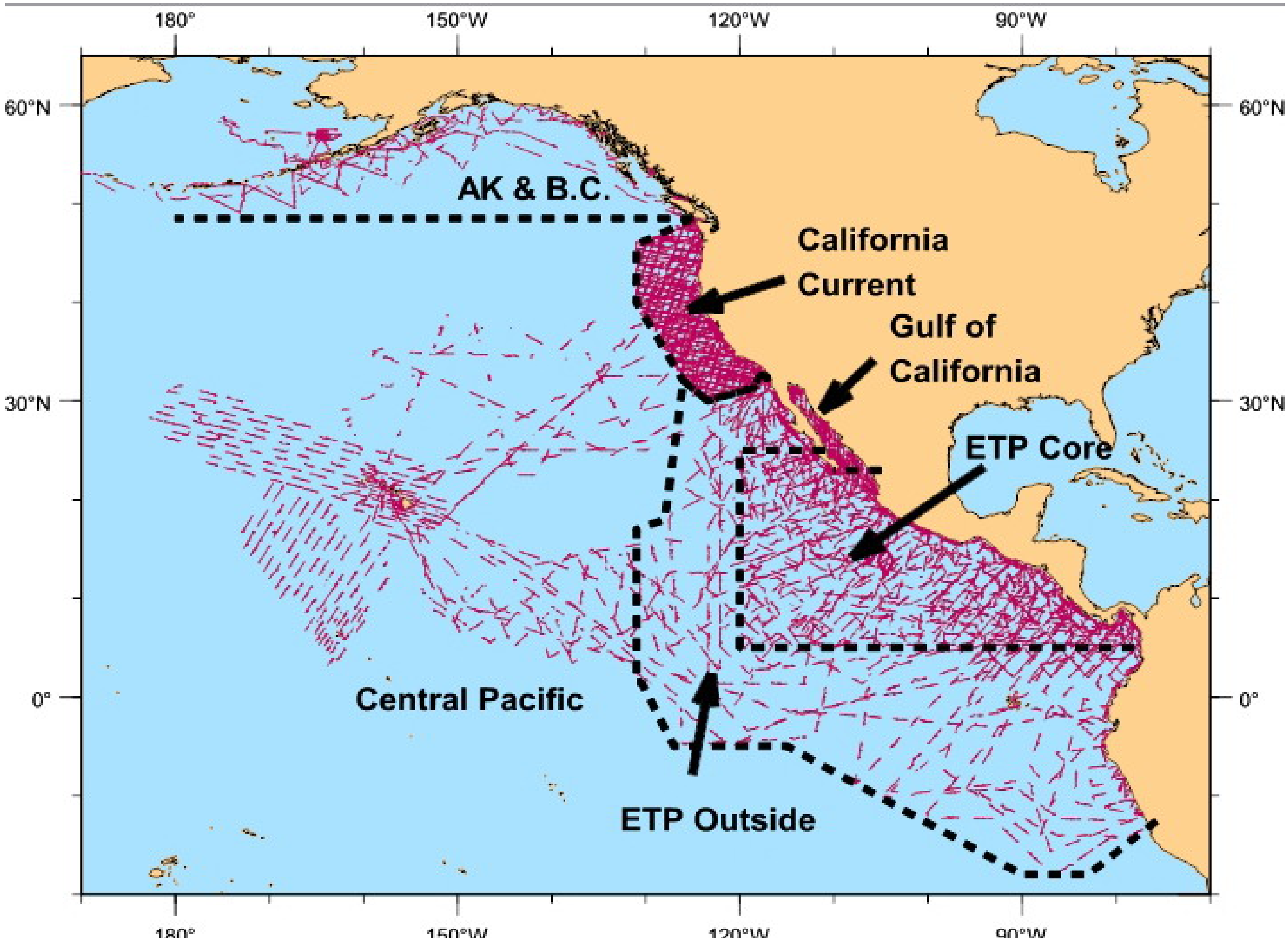
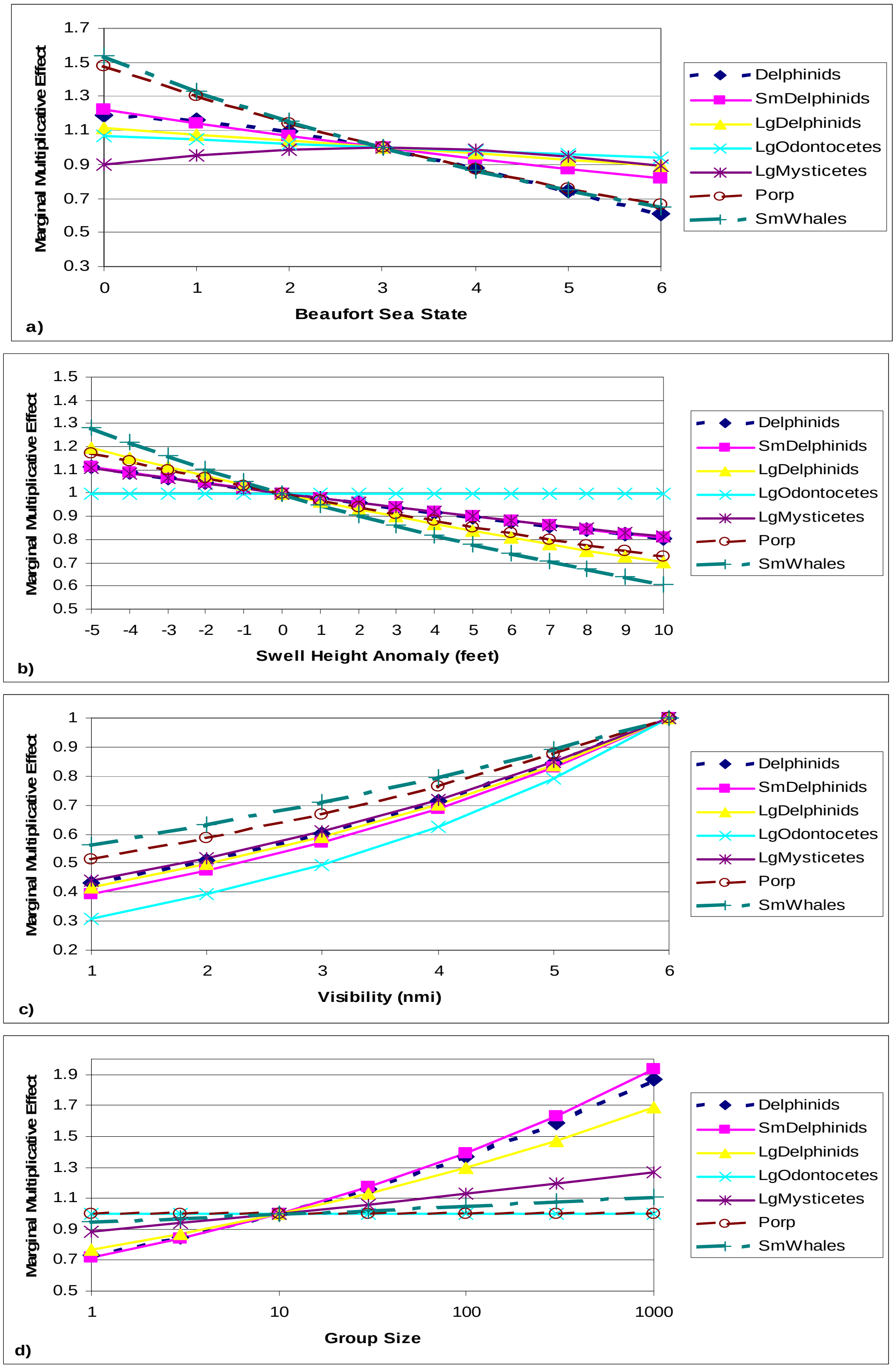


Figure 2. Multiplicative effects of a) Beaufort sea state (*Beauf*), b) swell height anomaly (*SwellAnom*), c) visibility (*Vis*) and d) group size on the scale parameter of a half-normal distribution fit to the observed distributions of truncated perpendicular sightings distances for seven species groups. Species groupings are given in Table 1. Beaufort sea state, swell anomaly, and visibility are normalized to their median values, and group size is normalized to a value of 10.



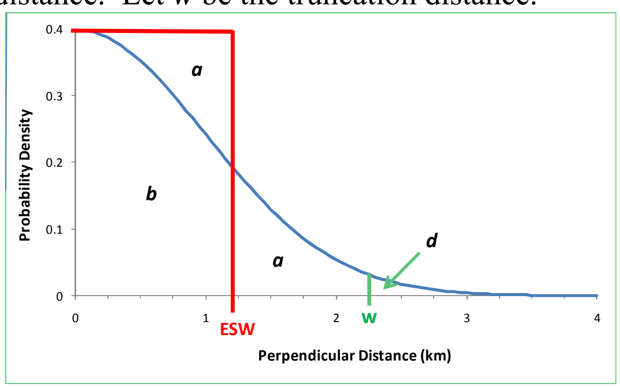
Appendix 1. Determining effective strip widths for sightings based on the coefficients estimated for covariates using a half-normal function.

Multiple-covariate line-transect models are based on the assumption that the natural logarithm of the scale parameter is an additive, linear function of covariates. The standard deviation (SD) is the scale parameter of the half-normal line-transect model. Therefore, the standard deviation of a half-normal line-transect model can be estimated as

$SD = \exp(\text{intercept} + \beta_1 \cdot \theta_1 + \beta_2 \cdot \theta_2 + \beta_3 \cdot \theta_3 + \beta_4 \cdot \theta_4 + \ldots)$, where
β_i = the estimated coefficient for the i-th covariate, and
θ_i = the value (or dummy value for categorical variables) of the i-th covariate.

Knowing the intercept, the coefficients, and the values (or dummy values) for the covariates, it is straightforward to estimate the standard deviation of the half-normal distribution.

Effective strip width (ESW) for a truncated half-normal distribution can be estimated from this standard deviation. The effective strip width is defined as the distance at which the probability of missing a group at lesser distance is equal to the probability of detecting a group at a greater distance. Let *w* be the truncation distance.



In the figure above, ESW is defined as the perpendicular distance at which the two areas labeled *a* are equal. The areas can be estimated as:

$$b = \int_0^{ESW} N_{pdf}\, dD$$

$$d = \int_w^{\infty} N_{pdf}\, dD$$

$$a = \int_{ESW}^{w} N_{pdf}\, dD = \int_0^{\infty} N_{pdf}\, dD - (b + d) = 0.5 - (b + d)$$

where N_{pdf} is the normal probability density function with a standard deviation of SD, and *D* is perpendicular distance. Area *a* can also be estimated as

$$a = \frac{0.3989}{SD} * ESW - b$$

where 0.3989/SD is the normal probability density evaluated at zero perpendicular distance. The latter two equations can be solved for *ESW*:

$$ESW = \frac{SD \cdot (0.5 - d)}{0.3989}$$

Effective strip width can therefore be calculated from the standard deviation (SD) that is estimated from the line-transect covariates and estimated intercept and coefficients (above).

An application of these formulae for estimating ESW is illustrated in an example below using the coefficients estimated in this paper. We estimate the ESW for a group of 53 striped dolphins seen on a 2003 survey aboard the McArthur II in survey conditions of *Beauf*= 5, *SwellAnom*= -0.5, and *Vis*= 5. Note that the *GroupSize* is expressed as the natural logarithm of the actual group size, the species coefficient is the natural logarithm of the value given in Table 4, and the net effect of the ordinal covariates is equal to the covariate value times the coefficient.

Covariate			
Name	Value	Coefficient	Effect
Intercept		-1.172	-1.172
Species	Striped Dolphin	0.092	0.092
Beauf	4	-0.067	-0.270
SwellAnom	-0.5	-0.021	0.011
Vis	5	0.187	0.935
GroupSize	3.970	0.330	1.311
Ship	McArthur II	0.159	0.159
Year	2003	0.116	0.116
Region	n/a	n/a	0.000
Cruise#	n/a	n/a	0.000
		Sum	1.183
SD=	3.265		
Trunc=	5.5		
ESW=	3.716		

Appendix 2. Data lists in R-language format giving the coefficient names and estimated coefficients for small delphinids, large delphinids, large odontocetes, large mysticetes, porpoises, and small whales. Coefficients are for a multiple-covariate, half-normal detection function fit to a truncated distributions of perpendicular sighting distances. Coefficients include an intercept and effects for Beaufort sea state (coeff.Beauf), swell height anomaly (coeff.SwellAnom), visibility (coeff.Vis), the natural logarithm of group size (coeff.LnTotSS), cruise number (coeff.CruzNo), the most abundant species in a group (coeff.SppMax), region (coeff.Region), ship (coeff.Ship), and year (coeff.Year). Numbers (i.e. 2, 3 or 4) after coefficient names refer to polynomial coefficients of the 2nd, 3rd, and 4th order. Cruise numbers refer to unique identification numbers used by the SWFSC. Regions names correspond to the regions shown in Figure 1. Ship names refer to the ships David Starr Jordan (DSJ), Endeavor (END), McArthur (MAC), and McArthur II (Mc2). Code numbers for SppMax are given in Table 1. R-language code to read these data structures and estimate effective strip widths is in Appendix 3.

Small Delphinids

```
#save this as text file 'SmDelphinid.coeff.txt'
structure(list(Intercept = -1.17154435244139, coeff.Beauf = -0.0674173577291749,
    coeff.Beauf2 = 0, coeff.Beauf3 = 0, coeff.SwellAnom = -0.0213534514082427,
    coeff.SwellAnom2 = 0, coeff.SwellAnom3 = 0, coeff.Vis = 0.187042612574128,
    coeff.Vis2 = 0, coeff.Vis3 = 0, coeff.LnTotSS = 0.330368047441034,
    coeff.LnTotSS2 = 0, coeff.LnTotSS3 = 0, coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426",
    "1467", "1468", "1508", "1509", "1601", "1604", "1605", "1607",
    "1610", "1611", "1612", "1613", "1614", "1615", "1616", "1617",
    "1621", "1622", "1623", "1624", "1627", "1628", "1629", "1630",
    "1631", "1634", "1635"), coeff.CruzNo = c(0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0), coeff.names.SppMax = c("002", "003", "005",
    "006", "010", "011", "013", "016", "017", "022", "025", "026",
    "027", "031", "077", "088", "090", "100", "101", "102", "103",
    "177"), coeff.SppMax = c(-0.0176986457157824, 0.186379204354464,
    0.347804740091564, -0.214502697255945, 0.0159944721383868,
    -0.320186812247046, 0.0918474763138507, -0.459268901763965,
    -0.148093350873714, -0.305523680996461, -0.269159167344642,
    0.508751245718006, 0.0297337875973701, -0.0904043657113494,
    0.642720445964805, 0.281220796307492, 0.0729986591844732,
    -0.379868815277969, -0.373632348470282, -0.319126315441983,
    0.00645453038911695, 0.71355974303961), coeff.names.Region = c("CentralNPac",
    "ETP-CORE", "ETP-OUTER", "GulfCal", "ORCAWA"), coeff.Region = c(0,
    0, 0, 0, 0), coeff.names.Ship = c("DSJ", "END", "MAC", "Mc2"
    ), coeff.Ship = c(-0.0561535087135978, -0.0699960307870748,
    -0.0329327664335339, 0.159082305934206), coeff.names.Year = c("1991",
    "1992", "1993", "1995", "1996", "1997", "1998", "1999", "2000",
    "2001", "2002", "2003", "2005", "2006", "2007", "2008"),
    coeff.Year = c(-0.11447696026202, 0.103720402056040, 0.00259826623933801,
    0.06454888539567, 0.00763850499694552, -0.115675763628974,
    -0.00682328150217962, 0.126609207096101, 0.0821027642521622,
    -0.205388266501776, 0.171766980799283, 0.116292959705944,
    -0.239377840583371, 0.236081421715201, 0.0351217252480271,
    -0.264739005026392)), .Names = c("Intercept", "coeff.Beauf",
    "coeff.Beauf2", "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2",
    "coeff.SwellAnom3", "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS",
    "coeff.LnTotSS2", "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo",
    "coeff.CruzNo", "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region",
    "coeff.Region", "coeff.names.Ship", "coeff.Ship", "coeff.names.Year",
    "coeff.Year"))
```

Large Delphinids

```
#save this as text file 'LgDelphinid.coeff.txt'
structure(list(Intercept = -0.830492414554335, coeff.Beauf = -0.0368793109360452,
    coeff.Beauf2 = 0, coeff.Beauf3 = 0, coeff.SwellAnom = -0.0355656661787991,
    coeff.SwellAnom2 = 0, coeff.SwellAnom3 = 0, coeff.Vis = 0.175413857778714,
    coeff.Vis2 = 0, coeff.Vis3 = 0, coeff.LnTotSS = 0.261081797250493,
    coeff.LnTotSS2 = 0, coeff.LnTotSS3 = 0, coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426",
    "1467", "1468", "1508", "1509", "1601", "1604", "1605", "1607",
    "1610", "1611", "1612", "1613", "1614", "1615", "1616", "1617",
    "1621", "1622", "1623", "1624", "1628", "1629", "1630", "1631",
    "1634", "1635"), coeff.CruzNo = c(0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
    coeff.names.SppMax = c("015", "018", "021", "032", "033",
    "034", "036", "277", "377"), coeff.SppMax = c(-0.263164674478506,
    -0.0963144614077351, -0.206327192934462, -0.446694561617205,
    -0.0886622810290301, 0.0253211966703576, 0.0131686147013674,
    -0.0209916948841358, 1.08366505497935), coeff.names.Region = c("CentralNPac",
    "ETP-CORE", "ETP-OUTER", "GulfCal", "ORCAWA"), coeff.Region = c(0,
    0, 0, 0, 0), coeff.names.Ship = c("DSJ", "END", "MAC", "Mc2"
    ), coeff.Ship = c(0, 0, 0, 0), coeff.names.Year = c("1991",
    "1992", "1993", "1995", "1996", "1997", "1998", "1999", "2000",
    "2001", "2002", "2003", "2005", "2006", "2007", "2008"),
    coeff.Year = c(0.354903797885561, -0.182631701443573, 0.142276185658412,
    -0.164499270062221, 0.188581523299023, -0.500885478251649,
    -0.0276532230881902, -0.0239365479175012, 0.205694573639803,
    0.276428845237926, -0.635858404609139, 0.145905178098018,
    0.285118876010395, 0.102711812247089, -0.042751188448352,
    -0.123404978255602)), .Names = c("Intercept", "coeff.Beauf",
    "coeff.Beauf2", "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2",
    "coeff.SwellAnom3", "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS",
    "coeff.LnTotSS2", "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo",
    "coeff.CruzNo", "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region",
    "coeff.Region", "coeff.names.Ship", "coeff.Ship", "coeff.names.Year",
    "coeff.Year"))
```

Large Odontocetes

```
#save this as text file 'LgToothedW.coeff.txt'
structure(list(Intercept = 0.0189843575539016, coeff.Beauf = -0.0219543263890091,
    coeff.Beauf2 = 0, coeff.Beauf3 = 0, coeff.SwellAnom = 0, coeff.SwellAnom2 = 0,
    coeff.SwellAnom3 = 0, coeff.Vis = 0.234231704633433, coeff.Vis2 = 0,
    coeff.Vis3 = 0, coeff.LnTotSS = 0, coeff.LnTotSS2 = 0, coeff.LnTotSS3 = 0,
    coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426", "1467", "1468",
    "1508", "1509", "1546", "1601", "1604", "1605", "1607", "1610",
    "1611", "1612", "1613", "1614", "1615", "1616", "1617", "1619",
    "1620", "1621", "1623", "1624", "1625", "1627", "1628", "1629",
    "1630", "1631", "1635"), coeff.CruzNo = c(0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0), coeff.names.SppMax = c("037", "046", "063"
    ), coeff.SppMax = c(0, 0, 0), coeff.names.Region = c("AK-BC",
    "CentralNPac", "ETP-CORE", "ETP-OUTER", "GulfCal", "ORCAWA"
    ), coeff.Region = c(0, 0, 0, 0, 0, 0), coeff.names.Ship = c("DSJ",
    "END", "MAC", "Mc2", "SUR"), coeff.Ship = c(0, 0, 0, 0, 0
    ), coeff.names.Year = c("1991", "1992", "1993", "1994", "1995",
    "1996", "1997", "1998", "1999", "2000", "2001", "2002", "2003",
    "2004", "2005", "2006", "2008"), coeff.Year = c(0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)), .Names = c("Intercept",
    "coeff.Beauf", "coeff.Beauf2", "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2",
    "coeff.SwellAnom3", "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS",
    "coeff.LnTotSS2", "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo",
    "coeff.CruzNo", "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region",
    "coeff.Region", "coeff.names.Ship", "coeff.Ship", "coeff.names.Year",
    "coeff.Year"))
```

Large Mysticetes

```
#save this as text file 'LgBaleenW.coeff.txt'
structure(list(Intercept = -0.0515548455976224, coeff.Beauf = 0.07199130834726,
    coeff.Beauf2 = -0.0123204003545985, coeff.Beauf3 = 0, coeff.SwellAnom = -0.0207923047254689,
    coeff.SwellAnom2 = 0, coeff.SwellAnom3 = 0, coeff.Vis = 0.164520399905488,
    coeff.Vis2 = 0, coeff.Vis3 = 0, coeff.LnTotSS = 0.119881602212297,
    coeff.LnTotSS2 = 0, coeff.LnTotSS3 = 0, coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426",
    "1467", "1468", "1508", "1509", "1546", "1601", "1604", "1605",
    "1607", "1610", "1611", "1612", "1613", "1614", "1615", "1616",
    "1617", "1619", "1620", "1621", "1623", "1624", "1625", "1627",
    "1628", "1629", "1630", "1631", "1634", "1635"), coeff.CruzNo = c(0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), coeff.names.SppMax = c("066",
    "070", "072", "073", "074", "075", "076", "079", "098", "099"
    ), coeff.SppMax = c(-0.581058546349015, 0.224018351087712,
    -0.00278365952424486, 0.215628475752496, 0.0159186322934409,
    -0.104879333272381, 0.210735155273783, 0.317281344164934,
    -0.129205014007626, -0.165655405419099), coeff.names.Region = c("AK-BC",
    "CentralNPac", "ETP-CORE", "ETP-OUTER", "GulfCal", "ORCAWA"
    ), coeff.Region = c(0, 0, 0, 0, 0, 0), coeff.names.Ship = c("DSJ",
    "END", "MAC", "Mc2", "SUR"), coeff.Ship = c(-0.0189278428418624,
    -0.111977205967702, 0.0720268616167258, 0.225332772065677,
    -0.166454584872838), coeff.names.Year = c("1991", "1992",
    "1993", "1994", "1995", "1996", "1997", "1998", "1999", "2000",
    "2001", "2002", "2003", "2004", "2005", "2006", "2007", "2008"
    ), coeff.Year = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0)), .Names = c("Intercept", "coeff.Beauf", "coeff.Beauf2",
    "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2", "coeff.SwellAnom3",
    "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS", "coeff.LnTotSS2",
    "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo", "coeff.CruzNo",
    "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region", "coeff.Region",
    "coeff.names.Ship", "coeff.Ship", "coeff.names.Year", "coeff.Year"))
```

Porpoises

```
#save this as text file 'Porp.coeff.txt'
structure(list(Intercept = 0.493111738311421, coeff.Beauf = -0.126334214797452,
    coeff.Beauf2 = -0.0013247430736434, coeff.Beauf3 = 0, coeff.SwellAnom = -0.0317421265216164,
    coeff.SwellAnom2 = 0, coeff.SwellAnom3 = 0, coeff.Vis = 0.134126120337354,
    coeff.Vis2 = 0, coeff.Vis3 = 0, coeff.LnTotSS = 0, coeff.LnTotSS2 = 0,
    coeff.LnTotSS3 = 0, coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426",
    "1508", "1546", "1604", "1605", "1607", "1617", "1619", "1620",
    "1625", "1627", "1628", "1635"), coeff.CruzNo = c(0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0), coeff.names.SppMax = c("040",
    "044", "477"), coeff.SppMax = c(-0.297844370503664, -0.212334623041841,
    0.510178993545506), coeff.names.Region = c("AK-BC", "ORCAWA"
    ), coeff.Region = c(1.26407043798311, -1.26407043798311),
    coeff.names.Ship = c("DSJ", "MAC", "Mc2", "SUR"), coeff.Ship = c(0,
    0, 0, 0), coeff.names.Year = c("1991", "1993", "1994", "1996",
    "1997", "2001", "2002", "2004", "2005", "2008"), coeff.Year = c(0.159852252188221,
    0.807582703197661, -1.66293806053158, 0.650484145408073,
    0.693422737856337, 0.723173655908542, -1.50340400717329,
    -1.71255718102924, 1.07932579491312, 0.765057959262156)), .Names = c("Intercept",
    "coeff.Beauf", "coeff.Beauf2", "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2",
    "coeff.SwellAnom3", "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS",
    "coeff.LnTotSS2", "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo",
    "coeff.CruzNo", "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region",
    "coeff.Region", "coeff.names.Ship", "coeff.Ship", "coeff.names.Year",
    "coeff.Year"))
```

Small Whales

```
#save this as text file 'SmWhale.coeff.txt'
structure(list(Intercept = 0.0690975342805672, coeff.Beauf = -0.142905590803002,
    coeff.Beauf2 = 0, coeff.Beauf3 = 0, coeff.SwellAnom = -0.0497583340456435,
    coeff.SwellAnom2 = 0, coeff.SwellAnom3 = 0, coeff.Vis = 0.114781645656011,
    coeff.Vis2 = 0, coeff.Vis3 = 0, coeff.LnTotSS = 0.0506425332818056,
    coeff.LnTotSS2 = 0, coeff.LnTotSS3 = 0, coeff.LnTotSS4 = 0, coeff.names.CruzNo = c("1426",
    "1467", "1468", "1508", "1509", "1601", "1604", "1605", "1607",
    "1610", "1611", "1612", "1613", "1614", "1615", "1616", "1617",
    "1619", "1620", "1621", "1623", "1624", "1625", "1628", "1629",
    "1630", "1631", "1634", "1635"), coeff.CruzNo = c(0, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0, 0, 0, 0), coeff.names.SppMax = c("001", "047",
    "048", "049", "051", "059", "061", "071", "078", "080", "083"
    ), coeff.SppMax = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), coeff.names.Region = c("AK-BC",
    "CentralNPac", "ETP-CORE", "ETP-OUTER", "GulfCal", "ORCAWA"
    ), coeff.Region = c(0, 0, 0, 0, 0, 0), coeff.names.Ship = c("DSJ",
    "END", "MAC", "Mc2"), coeff.Ship = c(0, 0, 0, 0), coeff.names.Year = c("1991",
    "1992", "1993", "1995", "1996", "1997", "1998", "1999", "2000",
    "2001", "2002", "2003", "2004", "2005", "2006", "2007", "2008"
    ), coeff.Year = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 0, 0)), .Names = c("Intercept", "coeff.Beauf", "coeff.Beauf2",
    "coeff.Beauf3", "coeff.SwellAnom", "coeff.SwellAnom2", "coeff.SwellAnom3",
    "coeff.Vis", "coeff.Vis2", "coeff.Vis3", "coeff.LnTotSS", "coeff.LnTotSS2",
    "coeff.LnTotSS3", "coeff.LnTotSS4", "coeff.names.CruzNo", "coeff.CruzNo",
    "coeff.names.SppMax", "coeff.SppMax", "coeff.names.Region", "coeff.Region",
    "coeff.names.Ship", "coeff.Ship", "coeff.names.Year", "coeff.Year"
    ))
```

Appendix 3. R-language program to read the data structures in Appendix 2 and estimates effective strip widths from survey data in an R dataframe (insert dataframe name in highlighted section below).

```
###########################################################################################
#  this routine estimates ESW from a half-normal distribution from the coefficients
#   that describe the standard deviation of that distribution and the associated sighting
#   data
#
#  this requires that you create a data frame of sightings with the following data for
#   each sighting:  SppMax (the 3-character species code for the most abundant species
#   in the group), Beauf (Beaufort sea state, as numeric), SwellAnom (the deviation of
#   swell height from that expected from a given Beauf, as numeric), Vis (the visibility,
#   as numeric, truncated at 6 nmi), LnTotSS (natural log of total group size, as numeric),
#   CruzNo (Cruise number as 4-character string), Region (region of survey as character
#   string: AK-BC, ETP-CORE, ETP-OUTER, GulfCal, ORCAWA, CentralNPac), Ship (ship as
#   3-character string: Mac, Mc2, DSJ, SUR, END), Year (survey year as 4-charater string)
#
#
###########################################################################################

###########################################################################################
# read in the coefficients used to estimate ESW as a data structure
###########################################################################################
 coeff.SmDelphinid= dget(file='SmDelphinid.coeff.txt')
 coeff.LgDelphinid= dget(file='LgDelphinid.coeff.txt')
 coeff.Porp= dget(file='Porp.coeff.txt')
 coeff.LgBaleenW= dget(file='LgBaleenW.coeff.txt')
 coeff.LgToothedW= dget(file='LgToothedW.coeff.txt')
 coeff.SmWhale= dget(file='SmWhale.coeff.txt')

###########################################################################################
# extract a list of species codes for each of 6 species groups
###########################################################################################
 listSmDelphinid= coeff.SmDelphinid$coeff.names.SppMax
 listLgDelphinid= coeff.LgDelphinid$coeff.names.SppMax
 listPorp= coeff.Porp$coeff.names.SppMax
 listLgBaleenW= coeff.LgBaleenW$coeff.names.SppMax
 listLgToothedW= coeff.LgToothedW$coeff.names.SppMax
 listSmWhale= coeff.SmWhale$coeff.names.SppMax


###########################################################################################
# nmax is the total number of sightings in your dataframe
###########################################################################################

#  attach([insert name of your dataframe here])
   nmax= length(SppMax)

# create null vectors to store the effects of each factor for each sighting
    c.Intercept= rep(0,nmax)
    c.SppMax= rep(0,nmax)
    c.CruzNo= rep(0,nmax)
    c.Region= rep(0,nmax)
    c.Ship= rep(0,nmax)
    c.Year= rep(0,nmax)
    c.Beauf= rep(0,nmax)
    c.SwellAnom= rep(0,nmax)
    c.Vis= rep(0,nmax)
    c.LnTotSS= rep(0,nmax)

###########################################################################################
# loop through all your sightings, estimating ESW for each
###########################################################################################

   for (i in 1:nmax) {

###########################################################################################
# determine which set of coefficients to use based on the SppMax for each sighting
###########################################################################################
```

```
      if (SppMax[i] %in% listSmDelphinid) {
         attach(coeff.SmDelphinid) } else {
      if (SppMax[i] %in% listLgDelphinid) {
         attach(coeff.LgDelphinid) } else {
      if (SppMax[i] %in% listPorp) {
         attach(coeff.Porp) } else {
      if (SppMax[i] %in% listLgBaleenW) {
         attach(coeff.LgBaleenW) } else {
      if (SppMax[i] %in% listLgToothedW) {
         attach(coeff.LgToothedW) } else {
      if (SppMax[i] %in% listSmWhale) {
         attach(coeff.SmWhale) } }}}}}

###########################################################################################
# estimate effect of categorical variable by matching names
###########################################################################################
      if (SppMax[i] %in% coeff.names.SppMax) { c.SppMax[i]=
coeff.SppMax[coeff.names.SppMax==SppMax[i]]
                                            } else { c.SppMax[i]= 0 }
      if (CruzNo[i] %in% coeff.names.CruzNo) { c.CruzNo[i]=
coeff.CruzNo[coeff.names.CruzNo==CruzNo[i]]
                                            } else { c.CruzNo[i]= 0 }
      if (Region[i] %in% coeff.names.Region) { c.Region[i]=
coeff.Region[coeff.names.Region==Region[i]]
                                            } else { c.Region[i]= 0 }
      if (Ship[i]   %in% coeff.names.Ship)     { c.Ship[i]= coeff.Ship[coeff.names.Ship==Ship[i]]
                                            } else { c.Ship[i]= 0 }
      if (Year[i]   %in% coeff.names.Year)     { c.Year[i]= coeff.Year[coeff.names.Year==Year[i]]
                                            } else { c.Year[i]= 0 }

###########################################################################################
# estimate effect of numeric variables as polynomials
###########################################################################################
      c.Intercept[i]= Intercept
      c.Beauf[i]= coeff.Beauf * Beauf[i] + coeff.Beauf2 * Beauf[i]^2 + coeff.Beauf3 * Beauf[i]^3
      c.SwellAnom[i]= coeff.SwellAnom * SwellAnom[i] + coeff.SwellAnom2 * SwellAnom[i]^2 +
                   coeff.SwellAnom3 * SwellAnom[i]^3
      c.Vis[i]= coeff.Vis * Vis[i] + coeff.Vis2 * Vis[i]^2 + coeff.Vis3 * Vis[i]^3
      c.LnTotSS[i]= coeff.LnTotSS * LnTotSS[i] + coeff.LnTotSS2 * LnTotSS[i]^2 +
                 coeff.LnTotSS3 * LnTotSS[i]^3 + coeff.LnTotSS4*LnTotSS[i]^4


      detach()

###########################################################################################
# end looping through sightings
###########################################################################################

   }

#  detach([insert name of your dataframe here])


###########################################################################################
# estimate overall effect as sum of intercept and individual effects
###########################################################################################

   coeff.sum= c.Intercept+c.Beauf+c.SwellAnom+c.Vis+c.LnTotSS+c.CruzNo+c.SppMax+c.Region+c.Ship+c.Year

###########################################################################################
# calculate the standard deviation of the half-normal from the sum of coefficients
###########################################################################################

   stndev= exp(coeff.sum)

###########################################################################################
# estimate ESW from a truncated half-normal (trunc= truncation distance)
###########################################################################################

   estESW=stndev*(0.5-pnorm(trunc,sd=stndev,lower.tail=FALSE))/0.3989
```

RECENT TECHNICAL MEMORANDUMS

SWFSC Technical Memorandums are accessible online at the SWFSC web site (http://swfsc.noaa.gov). Copies are also available form the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161 (http://www.ntis.gov). Recent issues of NOAA Technical Memorandums from the NMFS Southwest Fisheries Science Center are listed below:

NOAA-TM-NMFS-SWFSC-474 Global review of Humpback whale, *(Megaptera novaeangliae)*
A. FLEMING and J. JACKSON
(March 2011)

475 Status review update for Pacific salmon and steelhead listed under the Endangered Species Act: Central California Coast Coho Salmon ESU.
B.C. SPENCE and T.H. WILLIAMS
(March 2011)

476 U.S. Pacific marine mammal stock assessments: 2011.
J.V. CARRETTA, K.A. FORNEY, E. OLESON, K. MARTIEN, M.M. MUTO, M.S. LOWRY, J. BARLOW, J. BAKER, B. HANSON, D. LYNCH, L. CARSWELL, R.L. BROWNELL JR., J. ROBBINS, D.K. MATTILA, K. RALLS, and M.C. HILL
(June 2011)

477 Osteological specimens of tropical dolphins *(Delphinus, Grampus, Lagenodelphis, Stenella, Steno and Tursiops)* killed in the tuna fishery in the tuna fishery in the eastern tropical Pacific (1966-1992) and placed in museums by the Southwest Fisheries Science Center.
W.F. PERRIN and S.J. CHIVERS
(May 2011)

478 Ichthyoplankton and station data for surface (Manta) and oblique (Bongo) plankton tows for California Cooperative Oceanic Fisheries Investigations Survey cruises in 2007.
S.R. CHARTER, W. WATSON, and S.M. MANION
(May 2011)

479 Passive acoustic beaked whale monitoring survey of the Channel Islands, CA.
T.M. YACK, J. BARLOW, J. CALAMBOKIDIS, L. BALLANCE, R. PITMAN, and M. McKENNA
(May 2011)

480 Determining transmitter drag and best-practice attachment procedures for sea turtle biotelemetry studies.
T.T. JONES, B. BOSTROM, M. CAREY, B. IMLACH, J. MIKKELSEN, P. OSTAFICHUK, S. ECKERT, P. OPAY, Y. SWIMMER, J.A. SEMINOFF, and D.R. JONES
(November 2011)

481 Ichthyoplankton, paralarval cephalopod, and station data for surface (Manta) and oblique (Bongo) plankton tows for California Cooperative Oceanic Fisheries Investigations Survey and California Current Ecosystem Survey cruises in 2008.
W. WATSON and S.M. MANION
(May 2011)

482 Toward a national animal telemetry observing network (ATN) for our oceans, coasts and great lakes: Workshop synthesis report.
H. MOUSTAHFID, C GRIMES, J. KOCIK, B. BLOCK, K. HOLLAND, J. PAYNE, D. FOX, A. SEITZ, and C. ALEXANDER
(July 2011)

483 Fin whale acoustics as a tool to assess stock structure in the North Pacific.
B. JONES, S. RANKIN, and E. ARCHER
(November 2011)