Predicting trends in humpback whale (*Megaptera novaeangliae*) abundance using citizen science

NICOLÓ TONACHELLA[1,2], AURORA NASTASI[1], GREGORY KAUFMAN[1], DANIELA MALDINI[1] and ROBERT WILLIAM RANKIN[1]

The Great Whale Count (GWC) is an annual citizen science event that monitors changes in humpback whale (*Megaptera novaeangliae*) sightings in Maui County during the breeding season. The study includes 15 years of observations (1995–1996 and 1999–2011) with over 11 000 whale sightings. We provide a critical examination of the utility of the citizen science data given the challenges of observer-, site- and year-specific biases in counts, as well as an immeasurable and imperfect detection process. We estimate an annual increase of 5.16% per year (±2.76%), which closely resembles earlier trend estimates for Hawai'i. We demonstrate how uncertainty estimates in citizen science data can be strongly influenced by sampling processes, especially observer effects. Although such effects are now widely recognized in ecological studies, citizen science data often predate the mainstreaming of sampling protocols which measure and adjust for imperfect detectability. Here, we propose random effect models to minimize such effects in lieu of detectability techniques, and urge citizen science programs to adapt their protocols to handle observer processes at the planning and data collection stage.

Key words: *Megaptera novaeangliae*, Hawaii, citizen science, abundance, trends, shore-based survey, random effects, Generalized Linear Mixed Models.

INTRODUCTION

CITIZEN science is becoming a widespread tool for ecological and environmental monitoring (Silvertown 2009; Dickinson *et al.* 2010), especially in an era of fiscal restraint by governments and NGOs. Not only can citizen science engage members of the public and promote environmental stewardship (Trumbull *et al.* 2000; Cohn 2008), but the data can be useful in ecosystem monitoring and assessment. A classic example is the Christmas Bird Count sponsored by the National Audubon Society in the U.S.A., which has run every year since 1900 (Greenwood 1994; Cohn 2008; Silvertown 2009). Often, citizen science programs have a longevity and geographical breadth which surpasses other government or academic studies.

However, many citizen science programs have data collection protocols which complicate analyses and trouble ecological inferences. Protocols are commonly designed to facilitate volunteer coordination and training, rather than to address modelling assumptions such as independently and identically distributed (i.i.d.) errors. Even recognizing the need for updated and improved methodology can be stymied by fears of creating incompatibilities across years. Given the increasing popularity of citizen science for ecological inferences, there has been a growing call to scrutinize citizen science data and understand their sources of bias and variation (Dickinson *et al.* 2010).

This study presents a critical evaluation of the Maui Great Whale Count (GWC), a long-term humpback whale (*Megaptera novaeangliae*) dataset collected using citizen science. The GWC is one of the longest running cetacean projects involving citizen scientists, in operation since 1991. A consistent protocol was used over the years, consisting of one day of point-counts conducted simultaneously from multiple coastal sites along the southern shores of the island of Maui (Fig. 1). These sites are located within the Hawaiian Islands Humpback Whale National Marine Sanctuary. The Sanctuary waters between the islands of Maui, Moloka'i, Lana'i and Kaho'olawe are important for calving humpback whales (Aki *et al.* 1994) and host some of the highest densities of breeding humpback whales in Hawai'i (Mobley *et al.* 1999; Mobley *et al.* 2001). Monitoring the Hawaiian humpback whale population is important because of its status as a recovering species and its support of a thriving whalewatching industry (Gerber and DeMaster 1999; Utech 2000).

This study's primary objective is to provide an updated and accurate trend of humpback whale sightings for Maui coastal waters from 1995 to 2011 and to compare it to other humpback whale trend studies in Hawaiian waters. In order to meet this objective in the context of the available citizen science data, it is necessary to investigate some of the obvious sources of bias and variation, as well as assess the suitability of popular analytical techniques, such as distance sampling and the Generalized Linear Model (GLM). In particular, we investigate: i) imperfect detectability; ii) a catch-all notion of "overdispersion" to estimate extra variation; iii)

[1] Address: Pacific Whale Foundation, 300 Ma'alaea Road, Suite 211,Wailuku, HI 96793, U.S.A.
[2] Corresponding author: N.Tonachella@gmail.com

Table 1. Processes which may produce variation and correlation among counts.

Effect (*Notation*)	Sampling Process	Ecological Process
Observer (i)	Variation in ability to detect and count whales at distance; variation in sensitivity to local weather conditions.	
Site Location (s)	Local differences in weather and conditions which influence detectability, such as wind exposure and susceptibility to glare; differences in height-above-sea-level.	Whales aggregate in certain locations along the Maui Nui (i.e., Maui County).
Year (y)	Timing of Great Whale Count (January to March), and especially the mismatch between the peak of the migration and the date of sampling; annual differences in the weather during each year's survey date which influences detectability, such as wind, cloudiness or glare.	Population growth; timing of migration; variation in migration destination of the North Pacific stock.

observer effects; iv) point-count location (site) effects; and v) year effects.

The latter three effects are common in ecological studies (Kavanagh and Recher 1983; Gillies *et al.* 2006; Melbourne and Hastings 2008; Bolker *et al.* 2009). They manifest as correlations of the response variable within grouping variables, e.g., if whales prefer or avoid certain habitats (a site effect) or if some years have more favourable weather conditions for spotting whales (a year effect). Perhaps the most important effect for citizen science studies is the observer effect: the distribution of skills of amateurs in detecting objects and following protocols (Kavanagh and Recher 1983; Diefenbach *et al.* 2003; Koss *et al.* 2011). When such effects are not controlled for systematically through planning and sampling design, as is usually not the case in citizen science, then such inter-group correlations violate the assumptions of i.i.d. and can lead to improper inferences when relying upon GLMs. Table 1 lists a variety of ecological processes and artefacts of the sampling regime which may manifest as persistent site-, year-, and observer-specific biases in the GWC data.

Most of the processes listed in Table 1 influence whale sightings indirectly by affecting the ability of observers to detect whales at distance, such as localized weather phenomena or the different skills of observers. The explicit or tacit assumption when using point counts or cue counts for ecological inferences is that there is "constant proportionality" in detections (Norvell *et al.* 2003), whereby each observer, on average, detects objects with the same probability. If the proportion of detections changes by site, year, or observer, then trends in counts may reflect changes in detectability rather than changes in wildlife abundance (Simons *et al.* 2007). Therefore, count data usually require some means to either a) ensure that detection probability is constant, or b) measure detection probability as it varies and adjust the counts accordingly. A wealth of recent studies have demonstrated the pervasiveness of imperfect and variable detection probability in ecological studies (Nichols *et al.* 2000; Kéry and Schmid 2004; Alldredge *et al.* 2007; Conroy and Carroll 2009), and citizen science researchers should generally proceed from the assumption that detectability will be a major issue (MacKenzie and Kendall 2002).

The GWC was initially designed to facilitate detectability modelling via distance sampling (Buckland *et al.* 2004), by including bearing and distance estimates in the protocol. Distance sampling is perhaps the most popular method to correct for imperfect detectability. However, citizen science analysts should be aware that the method is difficult to apply and makes the strong assumption that objects are distributed uniformly across the scan-space (e.g., within 3 miles offshore in the GWC context), or that this distribution is easily estimable and has an integral function. This assumption appears as the $\pi(x) \sim x$ relationship in the distance sampling Likelihood (L_x):

$L_x(\theta)=\prod_i^n \frac{\pi(x)g(x)}{\int_0^{w}\pi(x)g(x)dx}=\prod_i^n \frac{xg(x)}{\int_0^{w}xg(x)dx}$, assuming uniform density,

where x is the distance from the observer; $g(x)$ is the detection function (e.g., the half-normal function); and $\pi(x)$ is the density of counts at distance x. Anecdotal evidence for this study suggested that whales were not uniformly distributed over the scan surface, thereby violating the assumptions of conventional distance sampling.

The challenge for the GWC, and for citizen science studies more generally, is to estimate an ecological process through the lens of an imperfect sampling regime. Without recourse to distance sampling or other direct measures of detection probability, it becomes necessary to strive for "constant proportionality" in detection probability. Here, we do this by addressing observer-, year- and site-specific biases in counts. Constant proportionality is central to meeting our primary objective of providing an updated and accurate trend of humpback whale abundance for Maui coastal waters.

An additional goal of this study is to provide a general framework to analyze the data and

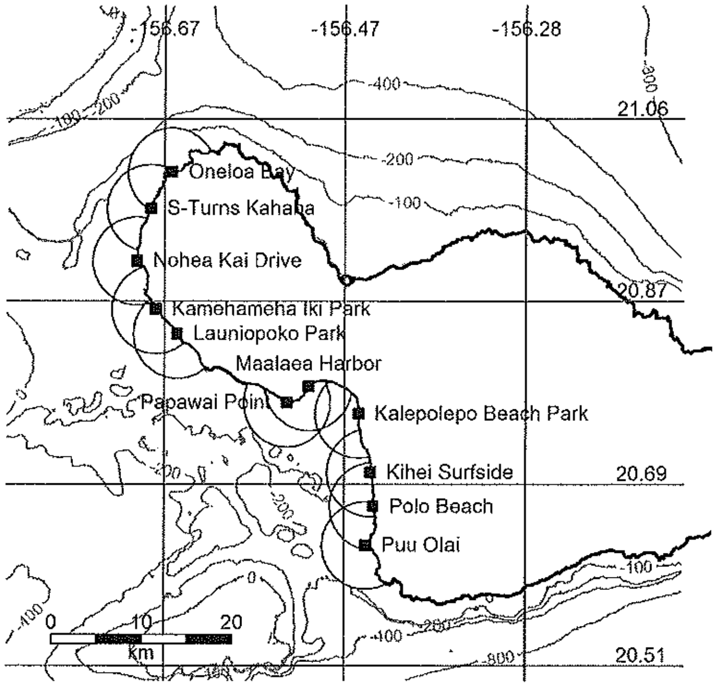

Fig. 1. Map of Maui, Hawai'i, including locations of the Great Whale Count sites considered in this analysis. Whale sightings were recorded within 4.8 km from shore, represented as the semi-circular buffers centred on each site (black squares).

facilitate similar efforts using GWC-like data on other Hawaiian islands, as well as citizen science more generally. We recognize that there may be better analytical techniques, such as Bayesian Hierarchical Models (Niemi and Fernández 2010; Moore and Barlow 2011) or tweaks to the distance sampling method which can estimate $\pi(x)$ (Marques *et al.* 2009). However, we feel that such methods are best considered at the study design phase, and are very complex and perhaps misleading to apply after-the-fact. Therefore, we suggest an analytical paradigm familiar to most science-trained professionals, namely the Generalized Linear Mixed Model (GLMM). GLMM are useful in reducing biases due to unmeasured processes (Halstead *et al.* 2012) as well as specific cases when detection probabilities cannot be measured directly but are suspected of being highly variable (Mac Nally *et al.* 2011).

METHODS

Study area

Humpback whales were counted at a number of shore-based sites on the island of Maui. We selected 11 of the most consistently surveyed sites for trend analysis (Fig. 1). The study area for each shore-based site's encompassed the coastal waters within approximately 4.8 km (3 statute miles) of each site's location. Survey areas varied from 23.44 to 40.03 km², due to the unique coastal morphology around each site (e.g., bays versus headlands).

Survey Methodology

Humpback whales were counted during one day between January 31st to March 11th each year, which generally coincides with the peak of the breeding and calving season (Herman and Antinoja 1977; Baker and Herman 1981; Tyack and Whitehead 1983; Baker and Herman 1984; Mobley and Herman 1985; Salden 1988, Au *et al.* 2000). The exact dates of the counts were generally decided based on convenience for volunteers and staff organizers. Data were available for 1995, 1996, and 1999–2011.

Sites were monitored by teams of observers consisting of a site leader (hereafter, referred to simply as 'observer') and volunteers. The observers were usually Pacific Whale Foundation staff and/or naturalists trained in the survey protocol. Volunteers did not receive formal training but received on-site instructions from the observers to assist in the counts.

The observers and volunteers scanned the waters within a 4.8 km radius of each site. Scans began at 8:30 and continued every 20 minutes until 11:30. Each scan lasted 10 minutes, followed by a 10 minute pause. During each scan, volunteers reported all sightings of humpback whale pods swimming within the survey area, estimated pod size (the number of whales in a pod), took bearing information, and estimated the distance from the pod to the shore-site location. Bearings were measured using a compass. Distances were estimated by comparing sighting location to a paper grid (with bearings) specific to each site. Whales were considered as being in a pod when they were seen within 10 body lengths (approximately 150 m) of one another and/or showed characteristic synchronization of surfacing patterns (Mobley and Herman 1985). Observers checked and confirmed all observations. The long scan duration (10 minutes) is assumed to facilitate accurate descriptions of pod size, composition, and distance information, as well as to increase the probability of detecting long-diving/singing males (Chu 1988).

During each scan, observers also recorded percent cloud cover, percent glare on the water, wind direction and Beaufort sea state (an ordinal variable for wind speed).

All site locations and whale sightings were subsequently entered into a Geographic Information System (GIS) in the 'R' programming language and environment (R Development Core Team 2010) using the 'rgeos' package (Bivand and Rundel 2012).

Analyses

When statistical analyses were performed, we used the term "whale sightings" to refer to counts of all individual whales during a scan. In these analyses, we are drawing conclusions from counts of whale sightings, but are extrapolating to the actual population of Maui County breeding humpback whales.

In the following sections, we briefly highlight a number of exploratory visualizations and analyses on topics including: imperfect detectability, non-uniform distribution of counts, bathymetric relationships, overdispersion and the presentation of two simple GLMs. These topics are reviewed only coarsely, inasmuch as they are a part of the decision process necessary to specify models 3 and 4. Models 3 and 4 represent the final tools to address this paper's ultimate objective, which is to estimate a trend of humpback whale population abundance in Maui County waters based on trends in sightings.

Detectability and Distance Sampling

There is much research on the phenomena of imperfect detectability and the problem it poses for making ecological inferences (Norvell *et al.* 2003; Simons *et al.* 2007; Conroy and Carroll 2009). Here, we do not document a full distance sampling analysis (which involves testing multiple key-functions, covariates, and random effects, including model selection), as we feel the technique is ultimately inappropriate for the GWC data. Nonetheless, we looked for the presence of imperfect detectability, by plotting density of counts versus distance-from-observer. Under imperfect detectability, one would expect higher densities close to the observer, and declining density away from the observer.

In a related exploration, we also plotted the density of counts versus distance-from-shore as (opposed to distance-from-observer). Distance-from-shore and distance-from-observer are related but distinct measures: distance-from-shore contours run parallel to the coast, and are assumed to be a dominant factor driving the distribution of whales, while distance-from-observer contours radiate in concentric circles from the point-count location and are assumed to be the dominant factor driving detectability. When observers are counting objects straight offshore, the two processes are confounded. In order to disentangle the confounding effect of imperfect detectability on whale density, our preliminary analysis only plotted densities in a single distance-from-observer bin: the bin with a distance interval of 4.02 to 4.82 km (2.5–3 miles). This semi-circular bin spans the entire range of distance-from-shore values in this study, and should have approximately the same detection probability along its contour. The goal of this visualization was to examine whether there were systematic changes in the counts of whales with distance-from-shore, which, if true, is a major violation of a fundamental assumption of the conventional distance sampling likelihood.

We also explored how whale counts changed according to bathymetry, supposing that depth, rather than distance-from-shore, was an alternative feature which might influence the nearshore distribution of whales. In the same distance-to-observer bin (4.02 to 4.82 km), we compared the observed depths of whale sightings versus the distribution of available depths. The empirical depth distribution also served to estimate a "minimum depth" for whales (i.e., the 97.5th depth percentile), under which we assumed that waters were too shallow to be considered whale habitat. This influenced our calculation of area offsets in the GLMs and GLMMs (described below).

Overdispersion

Overdispersion is a technical aspect of count modelling which can lead to incorrect inferences if present in count data (Vives *et al.* 2008; Fletcher 2012). We refer to counts being overdispersed when the variance in counts is much greater than the mean of counts (assumed to be equal in a Poisson distribution). Such extra variation is expected when a study is missing or cannot measure variables with high explanatory power. We expected citizen science sightings to be overdispersed because observers were likely to have a wide variety of competencies, thereby adding extra variation to the counts.

To check for overdispersion, we used a pre-analyses method similar to the simulations by Fletcher (2012). The simulations also serve as a data exploration exercise by testing which covariates (observer I.D., year, or site covariates) may explain extra variation in counts. The simulation proceeded as follows: i) calculate the mean of observed counts for each year; ii) randomly generate Poisson-distributed counts for each year, based on the observed means and the same number of observations (using the *rpois* function in R); iii) tally the number of instances when the randomly generated variance is as large or greater than the observed variance; iv) repeat for 1000 simulations. The percentage of simulations which result in variances greater than the observed variance is a measure of overdispersion: the percentage should be approximately 50% if the observed data is Poisson distributed, while a lower proportion is evidence of overdispersion.

Adding covariates with high explanatory power will always decrease overdispersion. Therefore, we repeated the above simulation, but calculated means and variances for each year-observer-site combination, supposing that observer and site effects were influencing detectability and counts.

This simulation also informed which family of count distributions was appropriate for the GLM/GLMM trend analyses (e.g., the Negative Binomial for overdispersed counts) and which covariates were likely sources of extra variation in the whale sightings.

GLM trend analyses (models 1 and 2)

We initially tested two GLMs. Model 1 was a simple regression of counts over years. The model is specified as:

$$C \sim NB(\lambda * A_s, \delta)$$

$$\log(\lambda) = \beta_0 + \beta_y * y$$

where C are the observed counts at scan t in year y and site s; λ is the mean for each year; A_s is the site (s) specific area for count sightings (i.e., waters within 4.8 km from the observer and deeper than 13 m, also known as an "offset"); δ is the dispersion parameter for the Negative Binomial distribution; $log(\lambda)$is the link function to relate the mean of counts to predictor variables (β); β_0 is the intercept; and β_y is the trend coefficient.

Model 2 was the same as model 1, but included a third explanatory variable for time-of-day (t) with coefficient β_t. This was motivated by anecdotal evidence that there were lower counts of humpback whales during the late-morning.

Both models 1 and 2 assume that there is a constant proportionality in the ratio of detected versus non-detected sightings, and that all counts are independently and identically distributed (i.i.d.). Rarely are such assumptions met in ecology, and probably least of all in citizen science. We examined the latter assumption by looking for patterns in the Pearson residuals (difference in expected versus observed counts, standardized by variance) across years, observers, sites, and other environmental variables. Models 1 and 2 allowed us to explore for patterns in the residuals, and whether there were persistent observer effects, site effects, and year effects after removing the variation due to the main trend over years and over time-of-day. The identification of such patterns was important for deciding how to specify the final models 3 and 4.

The effect sizes and standard errors of model parameters (δ, β_0, β_y and β_t) were estimated using Markov Chain Monte Carlo methods (MCMC; Gilks *et al.* 1996; MacKay 2003) using a component-wise Metropolis-within-Gibbs algorithm (Metropolis *et al.* 1953; Haario *et al.* 2005). We ran the MCMC chain until convergence was achieved, then ran 40 000 additional draws, thinning the results to about 2 000 samples from which inferences were made. MCMC techniques are useful for high-dimensional and/or random effects models (e.g., models 3 and 4), rather than simple GLMs, but were used for models 1 and 2 to ensure consistency across models.

GLMM trend analyses (models 3 and 4)

This section addresses our main goal: to provide an estimate of the trend in humpback whale abundance by regressing counts of whale sightings over time. As we cannot use conventional methods such as distance sampling to disentangle sampling processes from our trend of interest, we instead used random effects to parse the extra "nuisance" variation introduced by various grouping variables: observers, site, and year effects (Table 1).

GLMMs, also known as mixed effects model or random effects models, are very common in ecology and the reader is directed to excellent reviews on the subject (Gillies *et al.* 2006; Bolker *et al.* 2009). They are useful when there are unmeasured processes which can affect a parameter of interest (Halstead *et al.* 2012). In our case, the GLMMs treat the response variable as being correlated within each level of a grouping variable (such as years, observers, or sites) and adds an error term for each grouping variable (Schaub and Kéry 2012). In models 3 and 4, we consider there to be site-specific biases in counts, which altogether vary around the mean trend. We do the same for each individual observer and year. Technically, the addition of each random effect adds a variance parameter to our regression model: one for observers (σ_i^2), sites (σ_s^2), and year effects (σ_y^2). We can also use these variances to compare which processes are important in citizen science, i.e., how does variation in observer effects compare to the year-to-year variation?

We anticipated strong observer effects because only 33% of observers participated in the GWC for more than two years. Not only is this troubling from a data-reliability perspective, but it also presents technical challenges for solving the Maximum Likelihood for separate year and observer effects: there is no unique solution to parse the variation which is due to a year effect versus a common effect among all observers in the same year. For example, if we observe that one year has particularly high counts, we cannot conclude whether this is due to something about the whales or conditions unique to that year (e.g., a year effect) or if all observers in that year were coincidentally better at detecting whales (e.g., an observer effect). In order to parse the variation between years and observers, we added a separate observer effect likelihood for each year, but specified that each likelihood had a mean of zero and the same variance parameter for all 15 years. Philosophically, this assumes that the observers for each year were drawn from the same population, with the same mean and spread in abilities. Therefore, if counts were higher on average in one year versus another, the variation was absorbed by the year effects, rather than observer effects.

As in models 1 and 2, these GLMMs assume that there is constant proportionality in detections after accounting for correlations of counts within the same year, site and observer. In this manner, we do not make the assumption of i.d.d. The formal specification of model 3 is:

$$C \sim NB(\lambda^* A_s, \delta)$$

$$\log(\lambda) = \beta_0 + \beta_y{}^* y + \varepsilon_s + \varepsilon_y + \varepsilon_{i,y=1995}, + \varepsilon_{i,y=1996}, \ldots, \varepsilon_{i,y=2011}$$

$$\varepsilon_s \sim \mathrm{Norm}(0,\sigma_s^2)$$

$$\varepsilon_y \sim \mathrm{Norm}(0,\sigma_y^2)$$

$$\varepsilon_{i,y=1995} \sim Norm(0,\sigma_i^{\,2}),\ \varepsilon_{i,y=1996} \sim Norm(0,\sigma_i^{\,2}),\ldots,\ \varepsilon_{i,y=2011} \sim Norm(0,\sigma_i^{\,2})$$

where C, λ, A_s, δ, β_0, β_y are the same as in models 1 and 2 (representing variables for whale sightings, mean of counts, area offset, dispersion parameter, intercept, and trend coefficient, respectively); ε_s are site-specific random effects, normally distributed with a fixed mean of 0 and estimated variance of σ_s^2; ε_y are year-specific random effects, normally distributed with a fixed mean of 0 and estimated variance of σ_y^2; and $\varepsilon_{i,y=Y}$ are the per-year random effects of observers, each normally distributed with means fixed at 0, and sharing the same estimated variance parameter σ_i^2 across all years.

We also ran a complementary 4[th] model which included a third explanatory variable for time-of-day (t) with coefficient β_t.

We used MCMC techniques (the Metropolis-within-Gibbs algorithm; Haario *et al.* 2005) to estimate the distributions of all fixed effects, random effects, and variance parameters. Such high-dimensional problems require longer chains to ensure convergence and adequately sample the joint-distributions of parameters. Our

Table 2. Summary of the number of whales, pods and calves sighted during the 1995–1996 and 1999–2011 Great Whale Count surveys on Maui.

Year	Month	Day	Number of surveys	Number of stations	Total whale sightings	Total pods sightings	Total calf sightings
1995	Mar	11	36	4	173	102	13
1996	Mar	09	54	6	456	253	58
1997	-	-	-	-	-	-	-
1998	-	-	-	-	-	-	-
1999	Mar	06	63	7	293	189	30
2000	Feb	12	81	9	388	229	30
2001	Feb	24	54	6	615	425	42
2002	Feb	23	108	12	1020	498	91
2003	Mar	01	108	12	803	451	112
2004	Jan	31	108	12	685	421	61
2005	Feb	26	90	10	499	317	58
2006	Feb	25	108	12	1000	531	91
2007	Feb	24	108	12	1055	644	119
2008	Feb	23	90	10	1160	571	122
2009	Feb	28	108	12	711	439	71
2010	Mar	06	108	12	938	541	99
2011	Feb	26	99	11	1337	770	126
Total			1 323	147	11 133	6 381	1 123

routine required approximately 300 000 chains thinned to 2 000 samples. To our knowledge, there is no available software that can parse all four sources of variation (fixed effects and 3 variance parameters) and we wrote custom R functions for the joint-likelihood and MCMC sampler.

Model Comparison

We compared the four models by calculating each model's Akaike Information Criterion (AIC; Akaike 1974), checking for heterogeneity in Pearson residuals, and calculating a "posterior-predictive check" (Gelman 2003) which is a goodness-of-fit statistic for MCMC outputs. The AIC has a simple interpretation for GLMs, but there is not a standardized AIC for GLMMs, because the notion of what constitutes a "parameter" to be penalized is somewhat philosophical for random effects: at a minimum, there is an additional parameter for the variance of the likelihood function for each grouping variable; at a maximum, there is an additional parameter for every level of the random effect, which results in a larger AIC value (Bolker *et al.* 2009). The AIC values calculated for models 3 and 4 represent the more conservative approach, and consider every level of the random effect and its variance to be a parameter.

RESULTS

General Results

During the 1995–1996 and 1999–2011 Great Whale Counts, 1 323 shore-based scans were completed resulting in the 6 381 whale pod sightings and 11 133 humpback whales, 1 123 of which were calves (Table 2). The majority of pods sighted (49%) were composed of one whale, 34% of two, 12% of three and 5% of four or more individuals. The mean pod size was 1.4 (SE=±12.6) whales. Calves were mostly found in pods composed of two individuals (63%), which were assumed to be mother-calf pairs, as well as in pods of three individuals (28%), which we interpret as a mother-calf pair plus a male escort. Calf sightings (82%) occurred more often within 3.2 km from shore.

Imperfect Detectability

Figs. 2 and 3 show the change in count density with distance-to-observer and distance-to-shore. There appears to be a decrease in whale sightings with increasing distance-to-observer, which we interpret as evidence of imperfect detectability. Unfortunately, there also appears to be a strong increase in whale sightings with distance-to-shore, spiking around 3 km, which violates the central assumption of uniform density inherent to the conventional distance sampling method. Together, these imply that imperfect detectability may be an issue, but that conventional distance sampling is not appropriate for this situation. Our subsequent methods do not attempt to estimate or correct for this likely imperfect detection probability. Instead, we attempted to minimize systematic variation in the detection probability so we could assume "constant proportionality" of sightings versus true density.

The distribution of bathymetric values of observed sightings was significantly different

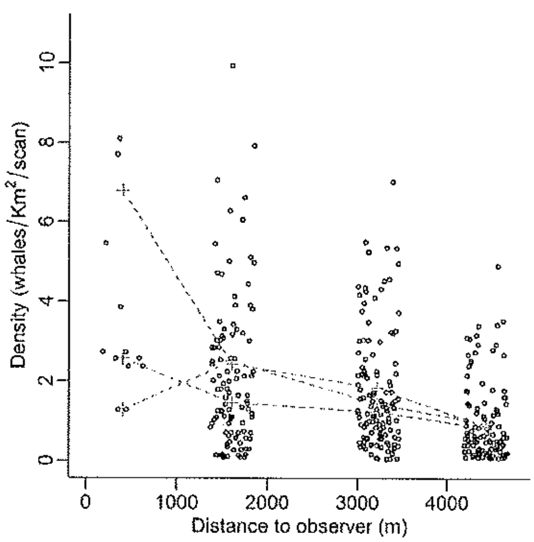

Fig. 2. Density of whale sightings versus distance to observer. Hatched lines show the moving-average for three separate years as an example of the variation in the way counts change with distance-to-observer.

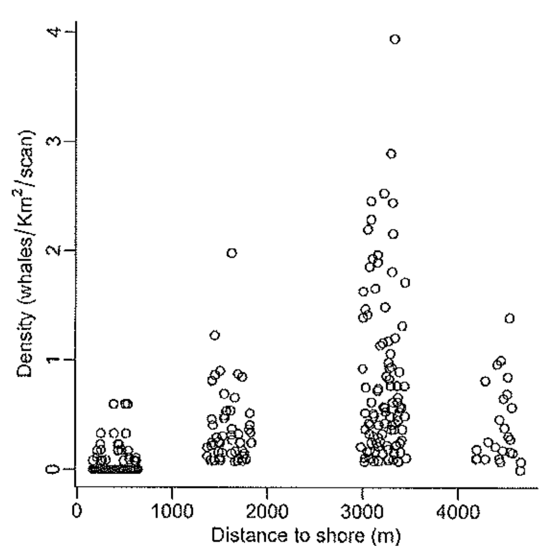

Fig. 3. Changes in density of whale sightings with distance-to-shore, pooled over all years and sites.

from the available depths (according to the Kolmogorov-Smirnov non-parametric test of distributions, with a p-value < 0.05). Both distributions peaked at around 50 m depth, whereas the distribution of whale sightings had a slightly longer tail at deeper values. 97.5% percent of sightings occurred at depths deeper than 13 m, and we used this value as an empirical "minimum depth" to calculate the offset areas in the GLMs and GLMMs (A_s, in the model specifications).

Overdispersion

For the overdispersion simulation, only 0.1% of the simulated counts yielded a variance greater than or equal to the observed variances. We considered this strong evidence of overdispersion, and proceeded with a Negative Binomial distribution to regress counts over time.

When we simulated the means and variances of counts at the level of each observer-year-site combination, the proportion of variances greater than or equal to the observed variances was 26.5%. This implies that a lot of the extra variation was due to observer and site effects. We also interpreted this as evidence for moderate overdispersion, and so we continued to use a Negative Binomial family for all four models.

Model Comparison

Table 3 shows the model fits and AIC values for the 4 models. All four models had adequate goodness-of-fit statistics, and similar trends estimates. However, visual inspection of Pearson residuals for models 1 and 2 revealed strong patterning across observers, sites and years. These patterns motivated the use of the more complex models 3 and 4, which included random effects for these grouping variables.

Based on AIC, there was strong support for model 4 (all other ΔAIC's where >> 3). The MLE trend estimate was 5.2%/year, with a confidence interval of 0.00 to 11.2%/year, and a p-value of 0.063. A one-tailed p-value (H_A: $\beta_y > 0$) was 0.031. There was strong evidence for a decline in whale abundance over the time-of-day, with a p-value of 0.002.

Fig. 4 compares the estimated variances of the three random effects. The variance parameters

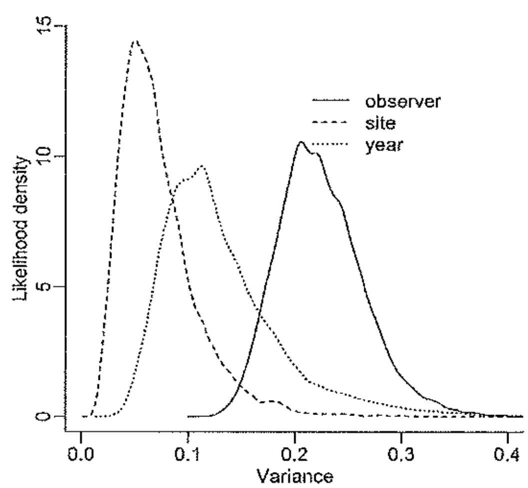

Fig. 4. Comparison of the estimated magnitude and distribution of variance parameters for the three random effects: observer, site and year effects.

Table 3. Comparison of models for trend-analyses.

Model	Fixed Effects	Random Effects	Parameters	AIC	ΔAIC	Goodness-of-fit	Trend ± S.E. (%/year)
1	$\beta_0 + \beta_y{*}y$	*n/a*	β_0, β_{year}, δ	7473	452	0.349	4.17 (±0.64)
2	$\beta_0 + \beta_y{*}y + \beta_t{*}t$	*n/a*	β_0, β_y, β_t, δ	7471	450	0.342	4.22 (±0.63)
3	$\beta_0 + \beta_y{*}y$	$\varepsilon_y \varepsilon_i$ $\varepsilon_{obs,y=0}, \varepsilon_{obs,y=1}, \ldots, \varepsilon_{obs,y=16}$	β_0, β_y, δ, σ_y^2, σ_i^2, σ_i^2	7038	17	0.458	4.34 (±2.76)
4	$\beta_0 + \beta_y{*}y + \beta_t{*}t$	$\varepsilon_y \varepsilon_s$ $\varepsilon_{i,y=0}, \varepsilon_{i,y=1}, \ldots, \varepsilon_{i,y=16}$	β_0, β_y, β_t, δ, σ_y^2, σ_i^2, σ_i^2	7021	0	0.452	5.16 (±2.76)

β coefficients for intercept *(0)*, year *(y)* and time of day *(t)* respectively.
ε random effects for each year, site *(s)* and observer *(i)*, respectively.
δ dispersion parameter for the Negative Binomial Distribution.
σ^2 variance parameters for the likelihood functions of year, site and observer random effects.

Table 4. Parameter estimates for final model (model 4).

Parameter	Notation	MLE	95% CI (lower)	95% CI (upper)	p value
Dispersion	δ	3.733	3.259	4.343	
Intercept	β_0	0.122	0.065	0.208	
Time of day	β_t	-0.059	-0.089	-0.019	0.002
Annual Trend	β_y	0.052	0.000	0.112	0.063
Site Variance	σ_s^2	0.050	0.018	0.144	
Observer Variance	σ_i^2	0.222	0.150	0.305	
Year Variance	σ_y^2	0.113	0.044	0.248	

suggest that a lot of the variation in the GWC is due to observer effects (variance of 0.22), followed by year effects and site effects. Figs. 5 and 6 show the temporal and spatial patterns of random effects for model 4.

All models showed patterning of Pearson residuals due to environmental variables, especially Beaufort sea state, whereby increasing Beaufort sea state led to lower counts (Fig. 7). However, this was less pronounced in models 3 and 4, where the decline was only prevalent at the highest (and least common) Beaufort categories of 6 (representing just 5 scans).

DISCUSSION

The contribution of citizen science to environmental and biological assessments is impressive, and we welcome its expanding importance. The vast scale of marine phenomena may depend on the distributed and inexpensive effort of many volunteers to effectively track large-scale changes, especially in

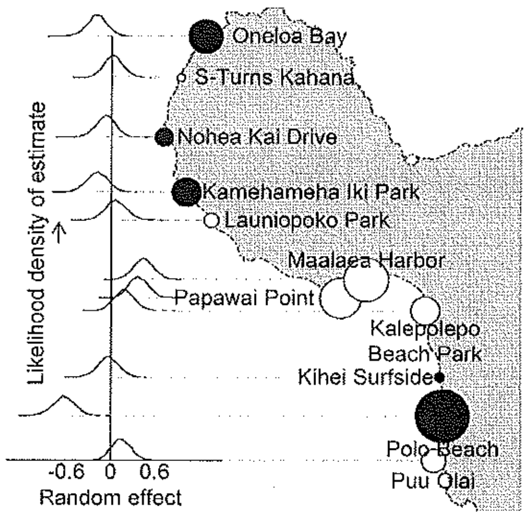

Fig. 6. Spatial distribution of site-specific random effects in whale sightings. The area of circles indicate the absolute magnitude of the effect, while colour indicates directionality: white for greater than 0 and black for less than 0.

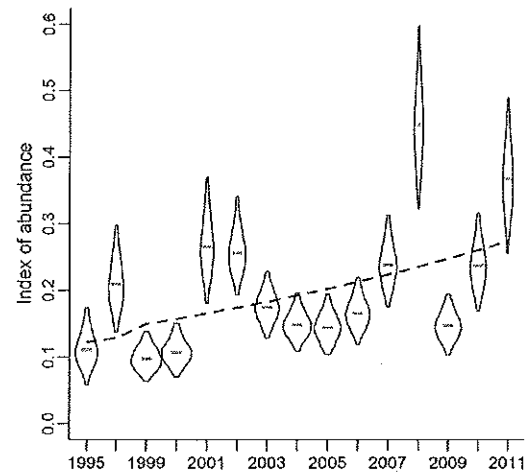

Fig. 5. Estimated overall trend (5.2%/year, hatched line) and distribution of year-specific random effects in the sightings of humpback whales in Maui coastal waters. Grey lines are means for year-specific random effects. Note: there is a break between 1996 and 1999.

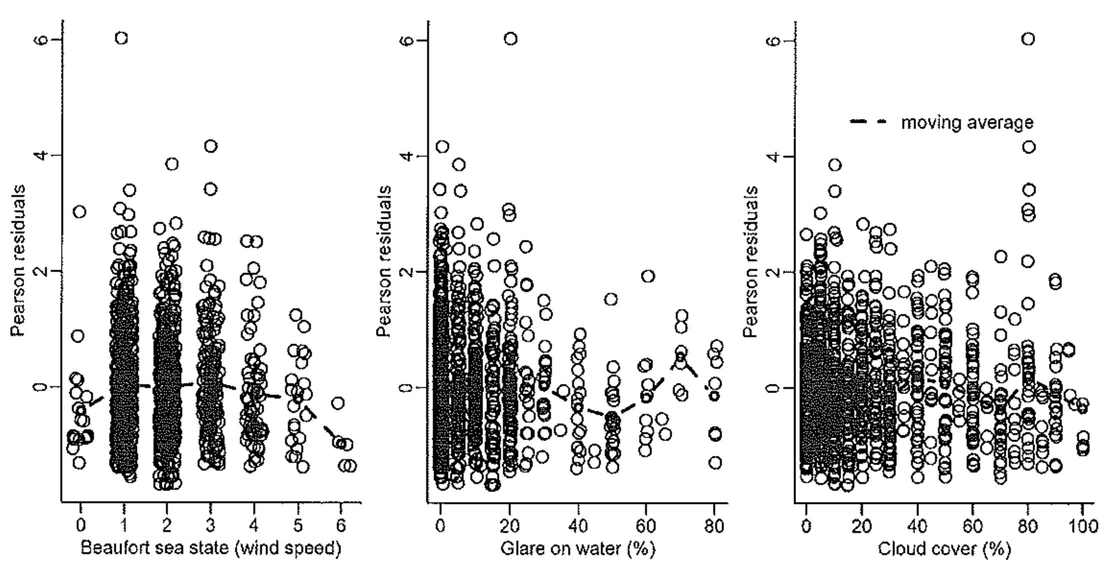

Fig. 7. Patterning of Pearson residuals for model 4, showing patterns in environmental covariates including: Beaufort sea-state, percent glare, and percent cloud cover. Hatched line is the moving average.

an era of budgetary constraint. However, it is increasingly important to understand the potential sources of variation and bias and to be aware that such long-term, multi-person studies can present problems for common analytical methods.

Our random effects model (model 4) estimated an increase of humpback whale sightings in the Maui coastal waters of 5.2% per year. Despite being limited to within 3 miles from the Maui southern shore, our trend is nonetheless similar to earlier Hawaiian trend estimates of a 5.5–7.0% per year proposed by Calambokidis (1999) and Mobley *et al.* (2001) based on photographic mark and recapture results and aerial survey data, respectively. The recent SPLASH report (Calambokidis *et al.* 2008) also estimated an increase of 5.5 to 6.0% per year for the entire Hawaiian Island chain. The confidence intervals of our estimate are very large, spanning 0 to 11.2%/year. The upper CI is close to the suggested physiological limit of population increase for humpback whales (Zerbini *et al.* 2010). Such large uncertainty means that the estimated trend is of marginal statistical significance (although the one-tailed p-value is <0.05). Strictly, this implies that the GWC and our GLMM do not have the statistical power to reject a null hypothesis of no trend. However, the 0.05 Type I error rate is perhaps not the most relevant statistic for wildlife managers for whom the mere identification and description of a trend is more important than hypothesis testing (Anderson *et al.* 2000; Gerrodette 2011).

Our GLMM facilitates the interpretation of those processes which may inflate uncertainty in the GWC. Most of the extra variation is due to sampling processes. For example, the variance for observer effects (strictly a sampling process) is much larger than the year or site effects (which are an unknown mix of both sampling and ecological processes; Fig. 4). This reaffirms earlier studies which document significant observer effects on wildlife abundance estimates (Kavanagh and Recher 1983; Sauer *et al.* 1994; Diefenbach *et al.* 2003). Such observer effects are perhaps especially important in citizen science programs, where differences between professionals and volunteers are likely to exist (Koss *et al.* 2011), and where the high turnover of participants, and the "first-year effect", has been well-documented and warned about (Kavanagh and Recher 1983; Dickinson *et al.* 2010). Given that GWC participants are only active for an average of 1.56 years, we would expect a large spread of abilities.

The mixture of both sampling and ecological processes within spatial (site location) and temporal (year of study) covariates makes it dangerous to test ecological hypotheses. For example, the GWC cannot definitively answer whether whales prefer certain habitats or locations, because the spatial distribution of counts is likely a function of detectability (e.g., different heights above the sea, exposure to the sun and wind) as much as habitat. This is evident in the lack of obvious spatial autocorrelation of counts among neighbouring sites (Fig. 6), despite overlapping fields-of-view (Fig. 1): if the spatial patterning of counts was truly grounded in ecological phenomena then we would expect close correlation among neighbouring and overlapping sites.

Likewise, the sharp breaks and large inter-annual variation in whale counts (Fig. 5, especially in 2007–2009) suggests a strong signal due to sampling processes instead of ecological processes. Such large fluctuations and cycles in abundance are well known for a variety of small mammals and marine taxa (Kendall *et al.* 1998), but allometric scaling suggests that cetaceans should have less pronounced, longer-period fluctuations (Krukonis and Schaffer 1991), unlike what is seen in the GWC data. Conversely, the periods of seeming stability in the GWC counts (2004–2006) are matched by estimated increases of approximately 146% from sight-resight models for the entire Hawaiian population during the same three year period (Calambokidis *et al.* 2008). We suggest that much of this inter-annual variation is not variation in the actual abundance of humpback whales, but is a result of sampling effects specific to each year's GWC, such as differences in weather and/or timing of the GWC versus timing of the whale migration.

In order for citizen science data to be useful for management decisions or ecological inferences, there needs to be a concerted effort to decrease the noise and bias during the data collection process. Protocols need to measure the variation due to the observer process, especially imperfect detectability. For the GWC, this is not possible with conventional distance sampling, and the existing protocol should be replaced with the double-observer method (Nichols *et al.* 2000) or the time-of-detection method (Farnsworth *et al.* 2002; Alldredge *et al.* 2007) which do not assume a uniform density of whales along the coast. Observer skill is also a major concern, which entails robust training of volunteers and validation of abilities. More attention should be given to retaining observers, especially given the documentation of "first-year effects." Having fewer, well-trained volunteers would also facilitate easier analysis of observer effects by having fewer parameters to model (e.g., this study involved 85 separate levels necessary in the GLMMs).

An important conclusion from this study is that simple Generalized Linear Model regressions are probably not suitable for citizen science data and may yield misleading results. Even in the case of the GWC, which avoids the common afflictions of other citizen science datasets, such as haphazard or opportunistic sampling (Dickinson *et al.* 2010), we demonstrate many violations to the central assumption of independently and identically distributed errors, as suggested by persistent patterns in observer, site and year effects (Figs. 4, 5 and 6). Despite continued criticism of GLMs for ecological analyses (Breiman 2001; Elith *et al.* 2006), they remain very popular. Managers should know that ignoring sampling processes can lead to underestimates of uncertainty, overconfidence in the reliability of the data, and perhaps controversies in drawing ecological inferences. For example, the standard errors of the GLM trend estimates (± 0.64) are much smaller than those of the GLMMs (± 2.72).

Our GLMM performs better than the GLMs in terms of satisfying the assumptions of constant proportionality of detections and heterogeneity of residuals. The lack of serious patterning in residuals across environmental variables, such as Beaufort sea state, suggests that much of the presumed variation in detectability is absorbed by the random effects. While random effects have been used previously to deal with imperfect detection probability (Mac Nally *et al.* 2011), we admit that this is not an ideal model specification, but a compromise in lieu of formal means to model detection probability. Only by specifying a mixture distribution that explicitly models the sampling process (e.g., detection) and the ecological process (e.g., count regression), and including appropriate covariates for either process, can we be confident that the population trend is not an artefact of changing sampling effects or environmental conditions.

We propose this method for an admittedly deficient protocol that predates the wealth of recent literature on detection probability and observer processes (Nichols *et al.* 2000; Farnsworth *et al.* 2002; Conroy and Carroll 2009). Increasingly, it is incumbent upon researchers to assume *a priori* that detection probability and observer effects will be an issue, and to either incorporate robust means of estimating their effects, or disprove their existence (MacKenzie and Kendall 2002). The utility of longterm and important datasets, such as citizen science efforts, will be increasingly contingent upon their ability to handle observer processes. We urge citizen science managers, of both new and historic programs, to incorporate protocols which explicitly handle observer effects and imperfect detection processes during the data collection procedure. We hope such changes in protocol can validate or invalidate *ad hoc* methods, such as the one proposed in this study, to deal with sampling processes in longterm datasets.

ACKNOWLEDGEMENTS

We would like to thank all the volunteers and Pacific Whale Foundation staff who have participated in the Great Whale Count over the many years of its operation. We would also like to thank the residents and businesses of the Island of Maui for their dedication to protecting and promoting the interests of the whales in Hawaiian waters. Finally, we thank two anonymous reviewers who provided comments which greatly improved this manuscript.

REFERENCES

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. on Automatic Control* **19:** 716–723.

Aki, K., Brock, R., Miller, J., Joseph, R., Mobley, P., Rappa J., Tarnas, D. and Yuen, M., 1994. A site characterization study for the Hawaiian Islands Humpback Whale National Marine Sanctuary, Technical Report 94–06, University of Hawai'i Sea Grant.

Alldredge, M. W., Pollock, K. H., Simons, T. R., Collazo, J. A., and Shriner, S. A., 2007. Time-of-detection method for estimating abundance from point-count surveys. *The Auk* **124:** 653–664.

Anderson, D. R., Burnham, K. P. and Thompson, W. L., 2000. Null hypothesis testing: problems, prevalence, and an alternative. *J. Wildl. Manage.* **64:** 912–923.

Au, W. W. L., Mobley, J., Burgess, W. C., Lammers, M. O. and Nachtigall, P. E., 2000. Seasonal and diurnal trends of chorusing humpback whales wintering in waters off western Maui. *Marine Mam. Sci.* **16:** 530–544.

Baker, C. S. and Herman, L. M., 1981. Migration and local movement of humpback whales (*Megaptera novaeangliae*) through Hawaiian waters. *Canadian J. Zool.* **59:** 460–469.

Baker, C. S. and Herman, L. M., 1984. Aggressive behavior between humpback whales, *Megaptera novaeangliae*, wintering in Hawaiian waters. *Canadian J. Zool.* **62:** 1922–1937.

Bivand, R. and Rundel, C., 2012. rgeos: Interface to Geometry Engine — Open Source (GEOS). Retrieved from http://CRAN.R-project.org/package=rgeos.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H. and White, J-S. S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecol. & Evol.* **24(3):** 127–135.

Breiman, L., 2001. Statistical Modeling: The Two Cultures. *Statistical Sci.* **16:** 199–215.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. and Thomas, L., 2004. Advanced Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, New York, NY.

Calambokidis, J., 1999. Status and trends of humpback whales in the North Pacific. Paper presented at the *Thirteenth Biennial Conference on the Biology of Marine Mammals, Hawaii, 1999.* p. 28.

Calambokidis, J., Falcone, E. A., Quinn, T. J., Burdin, A. M. J. P., Ford, J. K. B., Gabriele, C. M., Leduc, R., Mattila, D., Rojas, L., Straley, J. M., Taylor, B. L. R. J. U., Weller, D. H. B., Yamaguchi, M., Bendlin, A., Camacho, D., Flynn, K., Barlow, J. *et al.*, 2008. SPLASH : Structure of Populations, Levels of Abundance and Status of Humpback Whales in the North Pacific Final report for Contract AB133F-03-RP-00078.

Chu, K. C., 1988. Dive times and ventilation patterns of singing humpback whales (*Megaptera-novaeangliae*). *Canadian J. Zool.* **66:** 1322–1327.

Cohn, J. P., 2008. Citizen science: can volunteers do real research? *Biosci.* **58:** 192–197.

Conroy, M. J. and Carroll, J. P., 2009. Quantitative Conservation of Vertebrates. Wiley-Blackwell, Chichester, West Sussex, UK.

Dickinson, J. L., Zuckerberg, B. and Bonter, D. N., 2010. Citizen science as an ecological research tool: challenges and benefits. *Ann. Rev. Ecol. Evol. & Syst.* **41:** 149–172.

Diefenbach, D. R., Brauning, D. W. and Mattice, J. A., 2003. Variability in grassland bird counts related to observer differences and species detection rates. *The Auk* **120:** 1168–1179.

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A. T., Phillips, S. J. *et al.*, 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29:** 129–151.

Farnsworth, G. L., Pollock, K. H., Nichols, J. D., Simons, T. R., Hines, J. E., and Sauer, J. R., 2002. A removal model for estimating detection probabilities from point-count surveys. *The Auk* **119:** 414–425.

Fletcher, D. J., 2012. Estimating overdispersion when fitting a generalized linear model to sparse data. *Biometrika* **99:** 230–237.

Gelman, A., 2003. A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *Internat. Statistical Rev.* **71:** 369–382.

Gerber, L. R. and DeMaster, D. P., 1999. An approach to endangered species act classification of North Pacific humpback whales. *Cons. Biol.* **13:** 1203–1214.

Gerrodette, T., 2011. Inference without significance: measuring support for hypotheses rather than rejecting them. *Marine Ecol.* **32:** 404–418.

Gilks, W. R., Richardson, S. and Spiegelhalter, D. J., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

Gillies, C. S., Hebblewhite, M., Nielsen, S. E., Krawchuk, M. A., Aldridge, C. L., Frair, J. L., Saher, D. J., Stevens, C. E., and Jerde, C. L., 2006. Application of random effects to the study of resource selection by animals. *J. Anim. Ecol.* **75:** 887–898.

Greenwood, J. J. D., 1994. Trust the wildlife volunteers. *Nat.* **368:** 490.

Haario, H., Saksman, E., and Tamminen, J., 2005. Componentwise adaptation for high dimensional MCMC. *Computational Statistics* **20:** 265–273.

Halstead, B. J., Wylie, G. D., Coates, P. S., Valcarcel, P. and Casazza, M. L., 2012. "Exciting statistics": the rapid development and promising future of hierarchical models for population ecology. *Anim. Cons.* **15:** 133–135.

Herman, L. M. and Antinoja, R. C., 1977. Humpback whales in the Hawaiian breeding waters: Population and pod characteristics. *Scientific Reports Whales Res. Inst.* **29:** 59–85.

Kavanagh, R. and Recher, H. F., 1983. Effects of observer variability on the census of birds. *Corella* **7:** 93–100.

Kendall, B. E., Prendergast, J. and Bjørnstad, O. N., 1998. The macroecology of population dynamics: taxonomic and biogeographic patterns in population cycles. *Ecol. Letters* **1:** 160–164.

Kéry, M. and Schmid, H., 2004. Monitoring programs need to take into account imperfect species detectability. *Basic & App. Ecol.* **5:** 65–73.

Koss, R. S., Miller, K., Wescott, G., Bellgrove, A., Boxshall, A., McBurnie, J., Bunce, A., Gilmour, P. and Ierodiaconou, D., 2011. An evaluation of Sea Search as a citizen science programme in Marine Protected Areas. *Pac. Cons. Biol.* **15:** 116–127.

Krukonis, G. and Schaffer, W., 1991. Population cycles in mammals and birds: Does periodicity scale with body size? *J. Theo. Biol.* **148:** 469–493.

Mac Nally, R., Bowen, M., Howes, A., McAlpine, C. A. and Maron, M., 2011. Despotic, high-impact species and the subcontinental scale control of avian assemblage structure. *Ecol.* **93:** 668–678.

MacKay, D. J., 2003. Information Theory, Inference and Learning Algorithms. Cambridge University Press, Cambridge.

MacKenzie, D. I. and Kendall, W. L., 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecol.* **83:** 2387–2393.

Marques, T. A., Buckland, S. T., Borchers, D. L., Tosh, D. and Mcdonald, R. A, 2010. Point transect sampling along linear features. *Biomet.* **66:** 1247–1255.

Melbourne, B. A. and Hastings, A., 2008. Extinction risk depends strongly on factors contributing to stochasticity. *Nat.* **454:** 100–103.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E., 1953. Equation of state calculations by fast computing machines. *J. Chemical Physics* **21:** 1087–1092.

Mobley, J. R. Jr. and Herman, L. M., 1985. Transience of social affiliations among humpback whales (*Megaptera novaeangliae*) on the Hawaiian wintering grounds. *Canadian J. Zool.* **63:** 762–772.

Mobley, J. R. Jr., Bauer, G. B. and Herman, L. M., 1999. Changes over a ten-year interval in the distribution and relative abundance of humpback whales (*Megaptera novaeangliae*) wintering in Hawaiian waters. *Aquat. Mam.* **25:** 63–72.

Mobley, J. R. Jr., Spitz, S. S., Grotenfendt, R., Forestell, P., Frankel, A. and Bauer, G., 2001. Abundance of humpback whales in Hawaiian waters: Results of 1993–2000 aerial surveys. Report to the Hawaiian Islands Humpback Whale National Marine Sanctuary.

Moore, J. E. and Barlow, J., 2011. Bayesian state-space model of fin whale abundance trends from a 1991–2008 time series of line-transect surveys in the California Current. *J. App. Ecol.* **48:** 1195–1205.

Nichols, J. D., Hines, J. E., Sauer, J. R., Fallon, F. W., Fallon, J. E., and Heglund, P. J., 2000. A double-observer approach for estimating detection probability and abundance from point counts. *The Auk* **117:** 393–408.

Niemi, A. and Fernández, C., 2010. Bayesian spatial point process modeling of line transect data. *J. Agricul., Biolog. & Env. Statistics* **15:** 327–345.

Norvell, R. E., Howe, F. P. and Parrish, J. R., 2003. A seven-year comparison of relative-abundance and distance-sampling methods. *The Auk* **120:** 1013–1028.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Retrieved from http://www.r-project.org.

Salden, D., 1988. Humpback whale encounter rates offshore of Maui, Hawaii. *J. Wildl. Manage.* **52:** 301–304.

Sauer, J. R., Peterjohn, B. G. and Link, W. A., 1994. Observer differences in the North American Breeding Bird Survey. *The Auk* **111:** 50–62.

Schaub, M., and Kéry, M., 2012. Combining information in hierarchical models improves inferences in population ecology and demographic population analyses. *Anim. Cons.* **15:** 125–126.

Silvertown, J., 2009. A new dawn for citizen science. *Trends Ecol. & Evol.* **24(9):** 467–471.

Simons, T. R., Alldredge, M. W., Pollock, K. H., Wettroth, J. M. and Dufty, A. M., 2007. Experimental analysis of the auditory detection process on avian point counts. *The Auk* **124:** 986–999.

Trumbull, D. J., Bonney, R., Bascom, D. and Cabral A., 2000. Thinking scientifically during participation in a citizen-science project. *Sci. Education* **84:** 265–275.

Tyack, P. and Whitehead, H., 1983. Male competition in large groups of wintering humpback whales. *Behav.* **83:** 132–154.

Utech, D., 2000. Valuing Hawai'i's humpback whales: the economic impact of humpbacks on Hawai'i's ocean tour boat industry. Pp. 6–52 *in* The Economic Contribution of Whalewatching to Regional Economies: Perspectives from Two National Marine Sanctuaries. Marine Sanctuaries Conservation Series MSD-00-2. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Marine Sanctuaries Division, Silver Spring, MD.

Vives, J., Losilla, J. M., Rodrigo, M. F. and Portell, M., 2008. Overdispersion tests in count-data analysis. *Psychological Reports* **103:** 145–160.

Zerbini, A. N., Clapham, P. J. and Wade, P. R., 2010. Assessing plausible rates of population growth in humpback whales from life-history data. *Marine Biol.* **157:** 1225–1236.

LIST OF SYMBOLS

$L_x(\theta)$ equals the distance sampling likelihood function.

$\pi(x)$ *Greek letter pi* — equals the density of objects at distance x.

C equals total counts of whale sightings in a scan.

i equals the ith observer.

y equals the yth year of the study.

s equals the sth site in study.

t equals the time of day.

NB equals the Negative Binomial Distribution.

λ *Greek letter lambda* — equals the mean in the NB distribution.

A_s equals the site-specific area offset in NB regression.