SC/66a/SM/25

Report on Vaquita Rate of Change between 2011-2014 Using Passive Acoustic Data by the Expert Panel on Spatial Models

Armando Jaramillo-Legorreta, Lorenzo Rojas-Bracho, Jay VerHoef, Jeff Moore, Len Thomas, Jay Barlow, Tim Gerrodette, Barbara Taylor

INTERNATIONAL
WHALING COMMISSION

Report on Vaquita Rate of Change between 2011-2014 Using Passive Acoustic Data by the Expert Panel on Spatial Models

Armando Jaramillo-Legorreta, Lorenzo Rojas-Bracho, Jay VerHoef, Jeff Moore, Len Thomas, Jay Barlow, Tim Gerrodette, Barbara Taylor

Executive Summary

The Expert Panel on Vaquita Acoustic Monitoring had its second meeting on April 28, 2015 to analyze the first four seasons (2011-2014) of the acoustic monitoring program. Results indicate an estimated 67% decline in vaquita acoustic activity in the sampled area from 2011 to 2014. The average rate of decline of 31% per year (95% Bayesian Credible Interval -51% to -10% per year) is considerably worse than the previously estimated 18.5% per year (95% Bayesian Credible Interval -46% to +19% per year) for the 2011-2013 sampling period. These worsening results were caused by the very low number of detections in 2014, which resulted in an estimated rate of decline from 2013 to 2014 of 42%. The Panel found it is virtually certain that the acoustic activity has declined between 2011 and 2014 (prob. =

0.996) with very high probability of a rate of decline greater than 10% per year (prob. = 0.976).

The Panel considered the monitoring design to be sound but analyses were complicated by the loss of some monitoring devices (CPODs) in some years (2011, 2014) and low numbers of recording days for numerous CPOD devices in 2013. Several analyses were developed to account for the uneven sampling; all indicated substantial declines similar to the agreed estimate of 31% per year. The Panel agreed that year-to-year variation in the proportion of vaquitas present within the monitoring area could not be estimated both because the time series is short but also because the bycatch rate has likely changed from year-to-year. An earlier concern that the highest density acoustic detections were along the southeast boundary of the study area was addressed by adding CPODs further to the south in 2014. The new locations had approximately 10 times fewer detections than those to the northwest, which is consistent with the spatial pattern found in earlier visual survey data. The consistent spatial pattern of vaquita densities (both using acoustic and visual data) lends support to the conclusion that the decline is more likely a decline in vaquita population size rather than a shift in distribution.

Introduction

This paper updates analyses detailed in Annex 9 of the fifth meeting of the Comité Internacional para la Recuperación de la Vaquita (CIRVA V) by adding the 2014 acoustic monitoring data. In 2011, the passive acoustic monitoring program for vaquitas (*Phocoena sinus*) began the first full season of data collection. In April 2014, the Vaquita Acoustic Monitoring Steering Committee (SC) met to review data from the first 3 seasons of data (2011, 2012, 2013). Preliminary analysis suggested a dramatic decline in the vaquita population between 2011 and 2013 (Jaramillo-Legorreta et al. 2014). However, because the realized sampling effort was uneven across the sampling grid and over each sampling season, analysis of the data was not simple. Therefore, the SC recommended that a panel of experts with specific skills in spatial or trend modeling be convened to provide the best scientific analysis of trends in abundance of vaquita acoustic detections in a timeframe needed to manage this critically endangered species. This expert panel met on 24-26 June 2014 to develop statistical models based on the 2011-2013 results and held a web-based meeting on 28 April 2015 to present updated results that also included 2014 acoustic monitoring data. Here we present a summary of those results.

Background

The vaquita is a small species of porpoise found only in the northern Gulf of California, Mexico (Figure 1). It is subject to unsustainable bycatch in gillnet fisheries throughout its small range and, consequently, is classified as critically endangered by the International Conservation Union (IUCN). Although they are known to occur in waters 10-50 m deep, their distribution within the shallow water area is poorly characterized. The vaquita detections shown in Figure 1 are not fully representative of distribution in shallow water areas because most sightings are from a ship that cannot navigate shallow waters (see tracklines in Figure 1). The polygon within the figure is the Vaquita Refuge, which was agreed to in September 2005 (Protection Program published on December 2005) and within which no commercial fishing is allowed (no matter what fishing gear is used, even hooks). About half of vaquitas are estimated to be in the Refuge at any given time (Gerrodette and Rojas-Bracho 2011). Surveys in different years (1997 and 2008; Jaramillo-Legorreta *et al.*, 1999; Gerrodette *et al*., 2011) suggest that for the months of surveys (most from August through November) the distribution of vaquitas is remarkably constant. Within the Refuge, vaquitas are unevenly distributed.

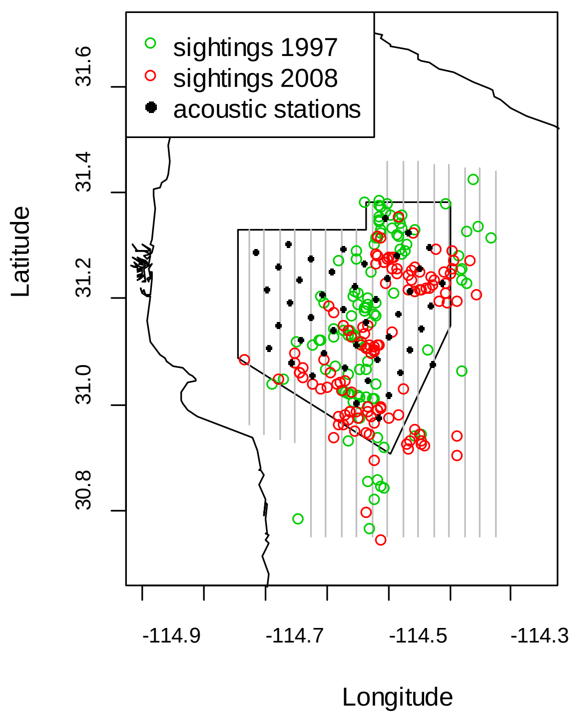


Figure 1. Visual detections (red and green circles) from two major ship surveys (in 1997 and 2008), with the survey track lines shown as light gray lines. The C-POD locations (deployed regularly since 2011) are shown as black dots and the Vaquita Refuge is outlined in black.

Because of the expense and imprecision of visual surveys (Jaramillo Legorreta, 2008; Rojas-Bracho *et al*., 2010), Jaramillo pioneered acoustic monitoring for vaquitas starting in 1997. Acoustic monitoring is possible because porpoises use echolocation to find their prey in the turbid waters of the northern Gulf of California. Jaramillo deployed boat-based acoustic detectors at fixed listening stations located throughout the range of vaquitas to examine the change in acoustic encounters over a period of 11 years (1997-2008) and showed a marked decline of 7.6%/year for a total decline of 58% (Jaramillo-Legorreta 2008). By the end of this study most stations recorded no vaquita acoustic activity and it became obvious that the level of acoustic monitoring effort achieved during the initial years of research were no longer sufficient to monitor vaquita activity accurately.

Thus, in 2008 several types of bottom-mounted passive acoustic devices, which are capable of recording autonomously for several months, were tested to increase the acoustic sampling effort for the dwindling numbers of vaquitas. A device called the CPOD had the best performance (Rojas-Bracho et al. 2010). The CPOD records characteristics of acoustic activity continuously over a period of several months. A Steering Committee (SC) was formed to design an acoustic monitoring project capable of detecting a ≥4%/year increase over a 5 year period (which would include 6 monitoring seasons). The SC created a grid design using 48 bottom-mounted CPODs deployed inside the Refuge for about 90 days each year. The original monitoring design also included CPODs located on Refuge perimeter buoys, but these CPODs were nearly all lost due to entanglement with fishing gear and likely active removal. A feasibility project was conducted using bottom-mounted CPODs just outside the southwestern boundaries of the Refuge but 6 of 8 were lost indicating that this area is still not possible to monitor with fixed CPODs (Jaramillo-Legorreta 2014).

After 2 years of initial testing and development, the acoustic monitoring program began its' first full season in 2011. The deployment and recovery of the bottom-mounted grid of CPODs was very successful over the first 4 seasons. However, the number of days recorded by individual CPODS differed because some CPODs were lost and never recovered, others shut off early within a season, and some filled their memory with background noise prior to retrieval. Figure 2 illustrates the achieved acoustic monitoring effort (i.e., days of acoustic monitoring per C-POD station) for the first 4 years.

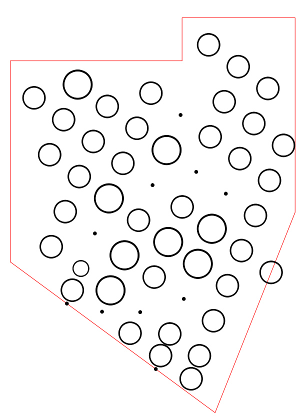


Figure 2. Locations of sampling sites, with number of days of monitoring effort indicated by circle size.

Effort also differed seasonally within year. CPODs were deployed later in 2012 and 2013 than in 2011 to avoid CPOD loss resulting from fishing activities (Figure 3), and deployment date now depends on information from aerial surveys that illegal fishing activities within the Refuge have largely ceased.

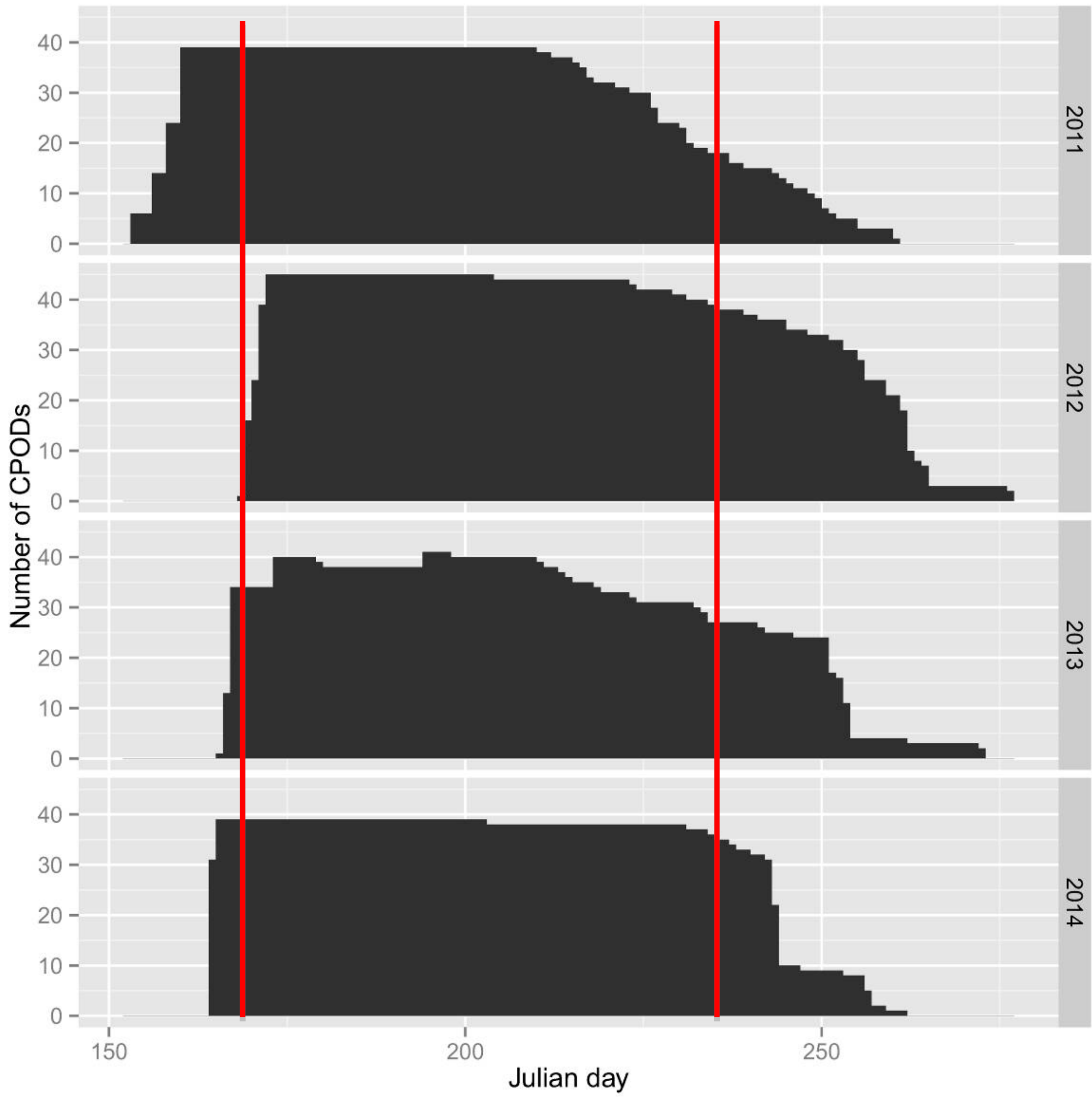


Figure 3. Effort by Julian day for each year. Julian dates shown run from May 30 (150) to October 2 (275). Vertical red lines enclose the core sampling period (from Julian day 170-231, June 19 to August 18, where ≥ 50% of the CPODs were operating in all years (discussed below). Julian dates actually vary slightly because of leap year.

Estimating the change in numbers of vaquita acoustic detections from 2011 to 2014 required an analytical treatment that accounts for the spatial and temporal differences in sampling within and between years, as shown in Figures 2 and 3. Conceptually, the analytical task is to best approximate the results that would have been obtained if all the circles in the grid shown in Figure 2 of were of equal size each year (same level of CPOD effort at all stations in all years). To do that, the Panel needed to consider all the factors that may make effort unequal and decide the best method of inference for stations that were un- or under-represented. In addition, the Panel needed to consider other factors besides differences in vaquita abundance or activity that may have caused differences in detections between years.

The simplest approach to measuring annual trends in vaquita clicks from C-POD data is to calculate the ratio of total clicks counted in consecutive pairs of years. However, this approach does not account for C-PODs that were lost or C-PODs that were not functional for the entire core sampling period. If C-PODs were lost predominately in high-density areas (which appears to be the case in 2011), this simple approach would produce biased estimates of trends. Likewise, if some sites received less effort, the total counts should be standardized to the number of days sampled, to avoid bias. To avoid both of these problems, analysis can be limited to data from only those sites that were sampled in consecutive years, and the mean number of clicks per day of sampling effort could be calculated for all these common sites. This direct-count method was used to produce estimates for comparison with other, better methods, which use more of the data (including data from sites that were only sampled in one or two years) and provide statistical estimates of uncertainty about the true trend given the data. The direct-count method does not make any estimate of certainty about the true trend but rather relies on an assumption that the data perfectly represent the true trend.

In contrast with the direct-count method, the Panel conducted statistical analyses that use spatial and temporal information within the dataset to estimate the probability that the acoustic data could have been observed by chance alone (noting that the data are a sample rather than perfect measurement of what we want to estimate) and to obtain a better estimate of trends that reflects uncertainty about the true trend for the population. The expert panel was directed to find the best method of statistical analysis to account for uncertainty and to make optimal use of all the available data.

Assumption and data selection analyses

The primary objective of the Panel was to estimate the annual mean rate of change in numbers of vaquita acoustic detections from 2011 to 2014 together with any uncertainties in that rate. A necessary assumption for analysis was that the annual rate of change in acoustic detections is a reasonable proxy for the rate of change in vaquita numbers. There are several important factors to keep in mind when interpreting the trend estimates from these first 4 years of acoustic detections. In June 2014, the Panel conducted analyses of those factors using the first 3 years of data. The Panel felt conclusions drawn were robust and elected not to repeat these assumption and data selection analyses with 2014 data.

First, if the monitoring grid covered the entire distribution of vaquitas, then inference about change in total vaquita population abundance would just depend on the assumption that click behavior remained the same through the time period (i.e., more recorded clicks would imply more vaquitas, not just more vocalizing, in the sampling area). Click behavior was investigated and there was no evidence of a change in clicks-per-vaquita in different years (see below). Additionally, there are data from past efforts covering the full range of vaquitas that support the assumption that acoustic detections and numbers of vaquitas decline at the same rate. For example, between 1997 and 2008 visual surveys and acoustic monitoring resulted in identical estimates of rate of change with a decline of 7.6%/year (Gerrodette et al. 2011, Jaramillo-Legoretta 2008). Therefore, the assumption that the number of recorded clicks is related to the level of use in the sampling area was judged to be reasonable.

Second, intense fishing outside the Refuge, even in the low summer fishing season, precludes using bottom-mounted CPODs outside the Refuge. Because the grid covers only a proportion of the vaquitas range, the other important assumption is that the proportion of vaquitas using the monitoring area over the summer period is the same each year. Over the 6-sampling seasons that the monitoring program was designed to cover, the changes in proportion in the Refuge would be expected to vary somewhat from year to year but not in any systematic way that would bias the rate-of-change estimate. However, with just four seasons of data (three periods of change), there is greater uncertainty about how much of the estimated annual change reflects change in overall population abundance vs. differences in the proportion of population using the sampling area each year. The length of the sampling period within a year mitigated this variability somewhat, but the Panel recognized these limitations to inference from the analysis. If future bycatch rates are constant (or reduced to zero), additional years of data may allow this issue to be addressed analytically.

Panelists agreed that the design of the monitoring program, which has systematic spatial coverage throughout the core of the Vaquita Refuge (and central to the distribution of the species) over a period of several months each year, was good, and that the analysis should rely primarily on this good design rather than on model-based spatial or temporal extrapolation to unsampled areas. The Panel carried out some basic descriptive analyses to consider factors other than a change in the number of vaquitas that might affect the number of acoustic detections observed.

Time of day: Because CPODs record data 24 hours per day and only whole days are used in the analysis, the sampling design is balanced with respect to time of day. **The Panel agreed that analysis could proceed without accounting for the influence of time of day on the data.**

Tide: The northern Gulf has a tidal range of over 10m (30 feet), which has potential to influence vaquita behavior and therefore acoustic detections. Therefore, the sampling of tidal states should be similar in different years if analyses are conducted without accounting for sampling of tidal states. Jaramillo stratified the 2011-2013 data into different tidal states. The tidal regime in the Upper Gulf of California is semidiurnal (two high and two low tides per day) and a cycle of spring-neap tides last approximately 15 days. Instead of using tide height as presented in tide tables, Jaramillo calculated the vertical speed of tide per hour as an index of tide current (using the tide height at the current hour minus the tide height at the previous hour). The absolute value was used, which does not distinguish between flood or ebb tides.Coverage of tidal states was similar between years (Table 1, 0.1 meters/hour intervals). A Kruskal-Wallis ANOVA by ranks indicated that the samples of every year originated from the same distribution, $H_{d.f\ 2,\ n=4464}=3.285$, p=0.1934. A median test gives similar non-significant results (Chi-squared=1.2, d.f.=2, p=0.5491). **The Panel agreed that analysis could proceed without accounting for the influence of tides on the data.**

Table 1. Number of hours sampled in eighteen vertical tide speed intervals for each sampling year period (2011-2013).

Tide speed interval		2011	2012	2013
≥ 0.0	≤ 0.1	151	150	130
> 0.1	≤ 0.2	153	156	144
> 0.2	≤ 0.3	159	160	145
> 0.3	≤ 0.4	151	133	156
> 0.4	≤ 0.5	125	134	131
> 0.5	≤ 0.6	139	126	138
> 0.6	≤ 0.7	121	115	128
> 0.7	≤ 0.8	106	117	111
> 0.8	≤ 0.9	99	90	95
> 0.9	≤ 1.0	73	75	73
> 1.0	≤ 1.1	76	77	76
> 1.1	≤ 1.2	62	57	57
> 1.2	≤ 1.3	36	42	44
> 1.3	≤ 1.4	27	34	24
> 1.4	≤ 1.5	8	15	20
> 1.5	≤ 1.6	2	7	12
> 1.6	≤ 1.7	0	0	3
> 1.7	≤ 1.8	0	0	1

Seasonal Effects: The Panel considered whether shifts in the amount of acoustic activity of vaquitas throughout the sampling season (generally from June through early September) could affect estimates of rate of change (see Appendix 3 for raw click data for each station and in each year). The distribution of sampling effort over the sampling season, as well as the pattern of apparent acoustic activity, differed somewhat among years (Figure 4). To avoid any potential biases caused by these differences, the Panel decided to analyze a seasonally reduced dataset that included dates chosen to be those within which at least 50% of the CPODs were operating across all 4 years, i.e., from Julian day 170-231[June 19 to August 19]. This core sampling period included 77.3% of the data, henceforth called the core dataset. The Panel used a Generalized Additive Model (details in Appendix 2) to assess whether the results from truncated dataset differed from using the full dataset (excluding data after September 14, the day prior to the earliest opening of shrimp season over the three years). This sensitivity test showed there were seasonal differences. This affirmed the choice to use the core dataset in order to avoid confounding inter-annual differences in seasonal sampling with potential seasonal differences in vaquita distribution. After discussion about whether it was necessary to model time within year (e.g., month), the Panel agreed that, for the purpose of estimating overall annual rate of change, using a common season across years and pooling data across that core period within a year would deal adequately with seasonal effects. **The Panel agreed that analysis could proceed using the core dataset and by averaging acoustic data within a year for each sampling point.**

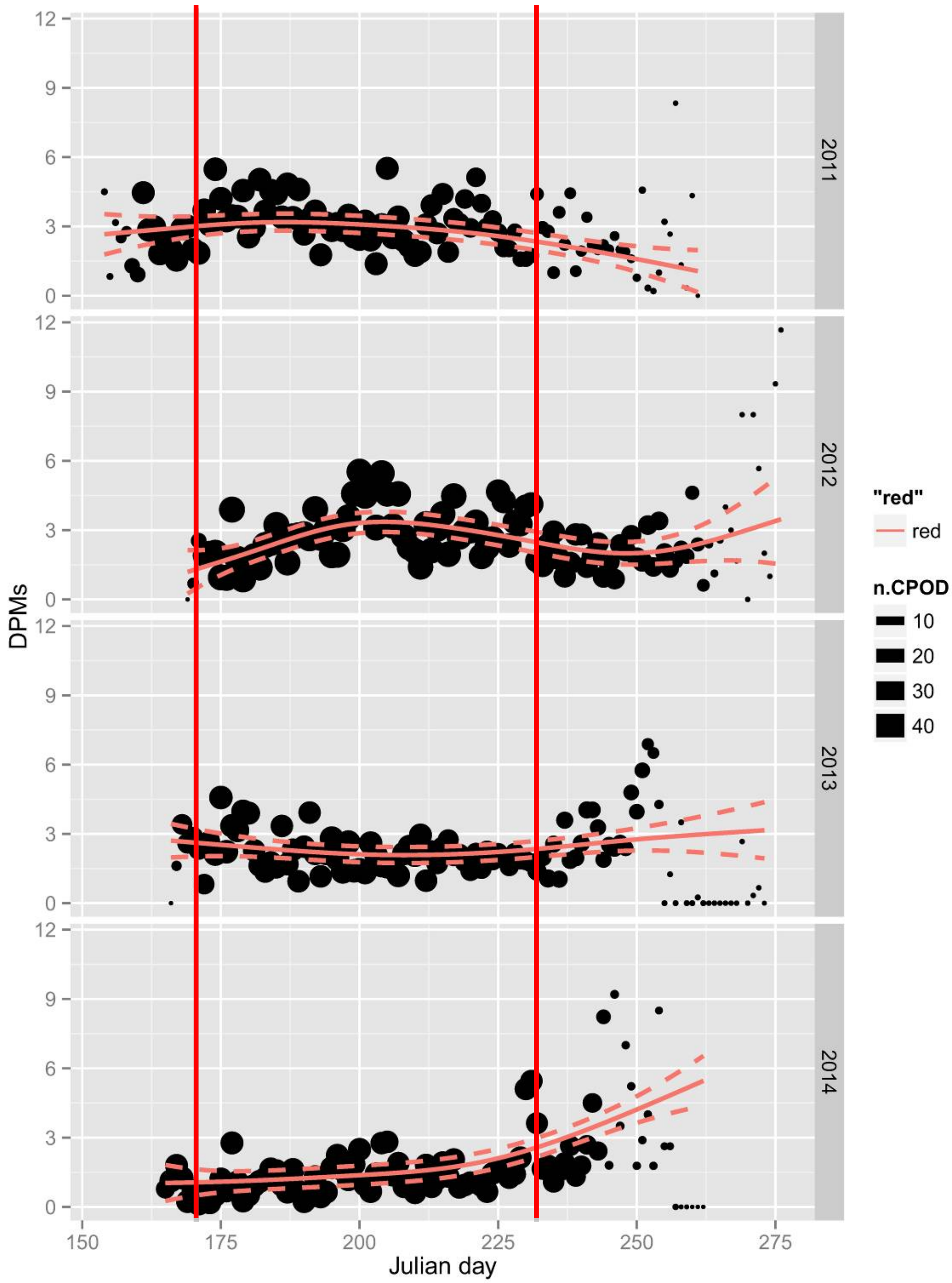


Figure 4. Mean acoustic detection positive minutes (see next section – Acoustic metric – for explanation), averaged across CPODs (y-axis) for each day of sampling (x-axis). Each dot represents a single day of sampling, with dot size proportional to the number of CPODs operating on that day. The red curves represent a smooth (a generalized additive mixed model with separate thin plate regression spline smooth per year, normal errors, identity link, weights that are number of CPODs and auto-regressive

error structure of order 1) with approximate 95% confidence interval shown as dashed lines. Vertical red lines indicate the core sampling period from Julian day 170-231.

Acoustic metric: The Panel focused its discussion on two types of measures of vaquita acoustics: clicks/day and detection positive time units (see below for discussion of appropriate time unit). Using acoustic events such as clicks/day to estimate trends in vaquita abundance assumes that acoustic events have a constant relationship with the number of vaquitas. Clicks are the most direct form of the acoustic data, and **Panelists agreed that clicks/day would be the preferred metric as long as the statistical properties were acceptable.** However, Panelists thought it useful to examine the data to see whether the amount of clicking per vaquita might have differed each year (e.g., due to annual differences in prey availability within the sampling area). The number of clicks per Detection Positive Minute (DPM, which is any minute that includes vaquita clicks) was variable, but with a similar pattern between years (Figure 5), which increased confidence in using clicks/day as a reliable acoustic index of vaquita abundance. Additionally, clicks/day was well characterized using a negative binomial distribution in generalized additive models (GAMs) and had no statistical issues in other models used (see details below and in Appendix 2). Nevertheless, the Panel thought analysis using a second metric that would be potentially less sensitive to changes in acoustic behavior would be useful as a sensitivity analysis. In addition to using DPMs, another metric explored was the number of times vaquitas were present ("positive") or not within a time unit that contained most vaquita encounters, where an encounter is determined as a period of detected activity (clicks) defined by silent gaps at each end of more than 30 minutes). The Panel considered different time units, and chose 30 minutes because just over 90% of vaquita encounters were less than 30 minutes in duration (Figure 6). These encounter units are called Detection Positive Half Hours (DPHH). The metric of vaquita positive 30-minute periods was thus used to examine the robustness of the results based on clicks/day.

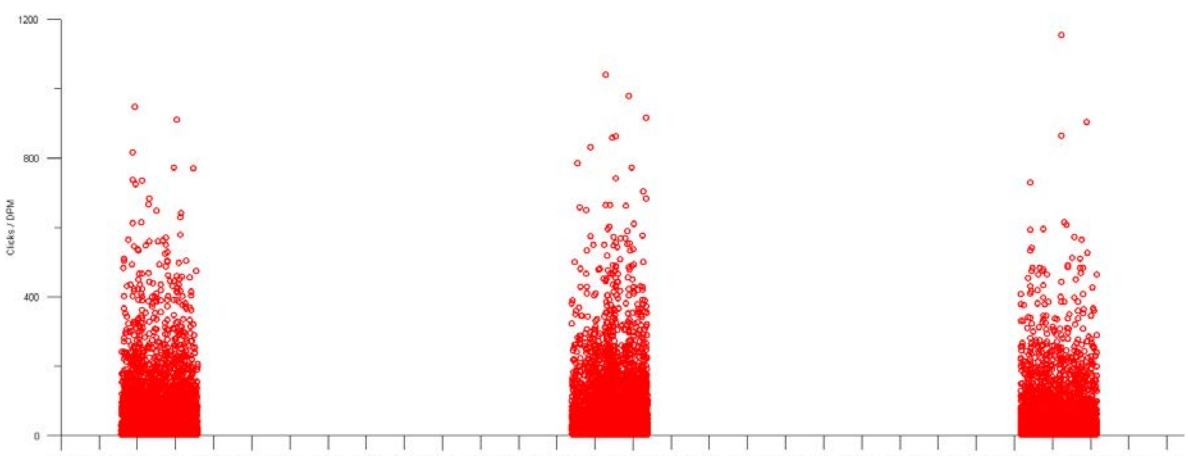


Figure 5. The number of clicks per Detection Positive Minute (DPM) over time.

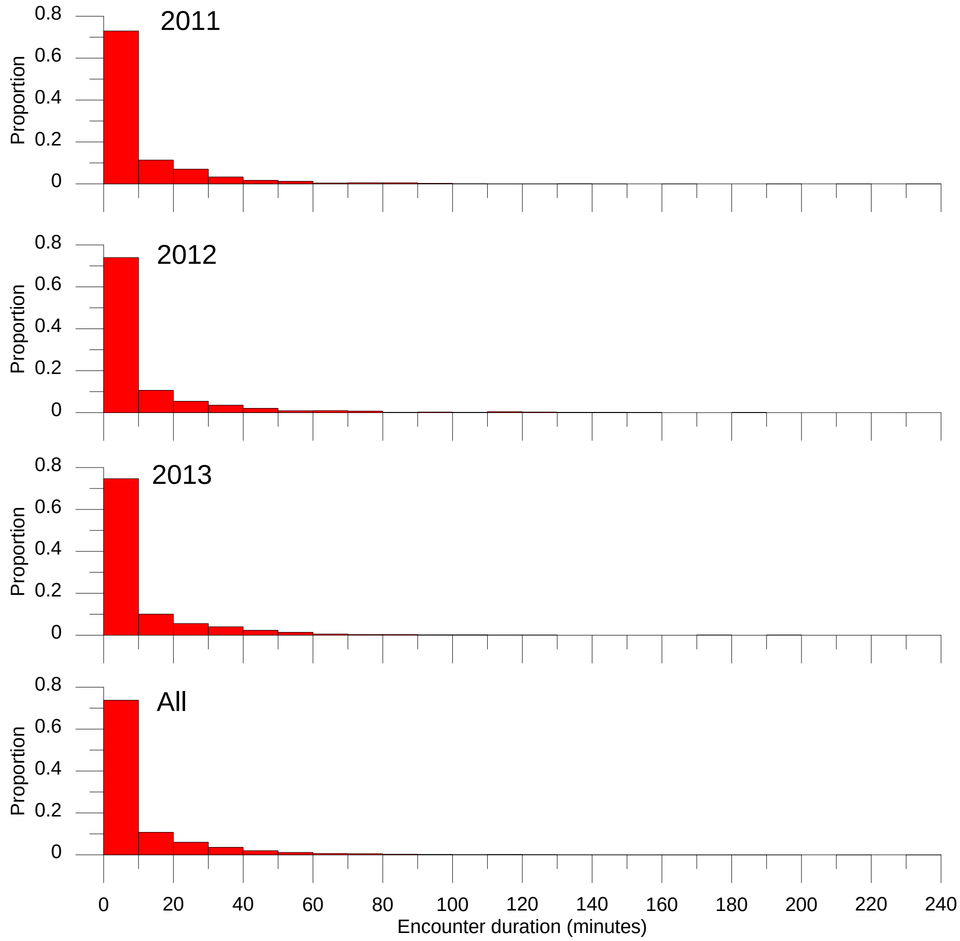


Figure 6. Proportion of vaquita encounters binned by encounter duration.

The relationship between number of DPMs per encounter and encounter duration appears to be linear, although with high variability (Figure 7). Thus, rates of echolocation (as indicated by slope) are nearly constant with increasing encounter duration. Different colors are shown for the three years (red, black and blue respectively from 2011-2013). No differences between years are apparent.

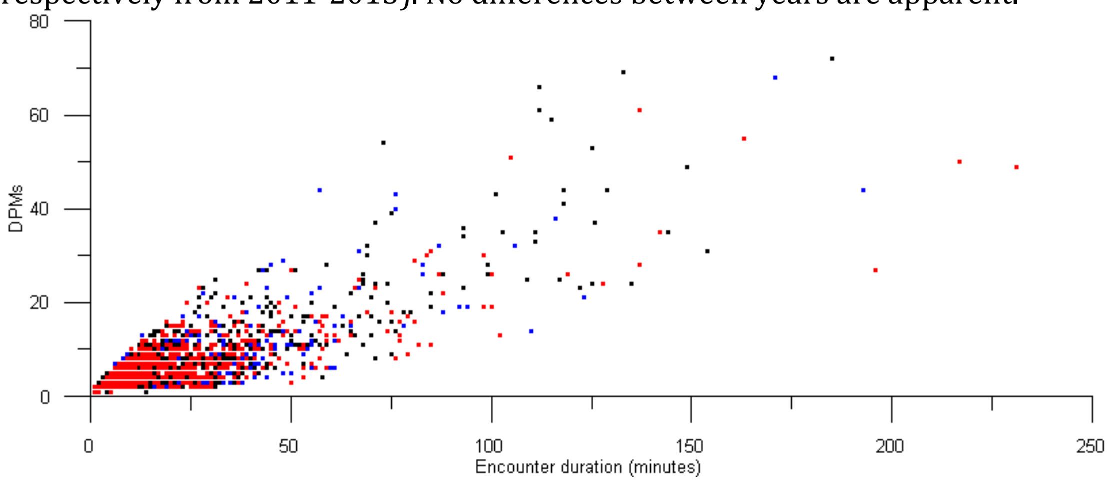


Figure 7. Scatterplot of DPMs for different encounters and for different years (2011-2013).

The GAM models using a negative binomial distribution had a poorer fit using either DPMs or Detection Positive Half Hours (DPHH) per day than using clicks/day (detailed below and in Appendix 2). The DPHH also tended to become saturated (Figure 8). An aggregation of 2 vaquitas could produce similar values of DPM and even more similar values of DPHH as an aggregation of 5 vaquitas, whereas total clicks would be expected to increase more linearly with average group size. This topic is further discussed below under the Spatial GAMs Model.

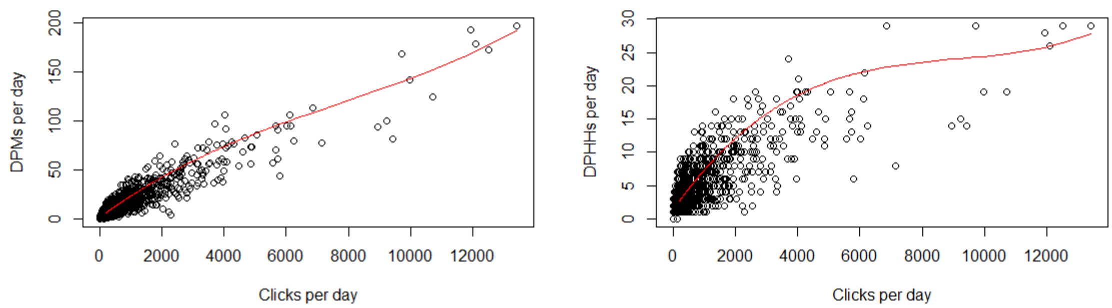


Figure 8. A loess smoothed fits of the number of detection-positive minutes (DPMs) per day (left) and the number of detection-positive half hours (DPHH) per day (right) as functions of the number of vaquita clicks per day for each site and year. Data are limited to the core sampling period 2011-2013.

The Panel agreed that the metric of choice was clicks/day because this metric uses the most raw form of the data and no statistical issues preventing its use.

Agreed scope of inference: The Panel discussed at length the types of analyses that could be performed on the data, and the inferences that could be drawn from the results.

1. **The Panel agreed that the spatial scope of inference should be limited to the CPOD sampling locations**. In other words, predictions from all models would be made only at the sample locations; no attempt would be made to extrapolate the predictions to some wider area such as the entire refuge. Such extrapolations cannot reliably be made from spatial models that omit biologically-relevant explanatory variables; in the present case constructing a detailed spatial habitat model would take far longer than the time available.
2. **Estimates would only be made covering the core sampling period, where at least 50% of the CPODs were operating in all years.** Any analysis would need to account for the fact that some locations did not have CPODs operating for the full time period in each year; data from each location and year should be weighted by the number of sample days.
3. **Inference from the analysis would be based on model-predicted click counts from the model at all sampled locations (n = 45)**. An alternative would have been to predict click counts only at locations with no sampling effort in a particular

year, and to use observed click counts at the other locations for making between year comparisons; this approach was rejected firstly because of the uneven number of sampling days across locations (higher sampling error and thus less confidence that the raw data accurately represent activity levels at less frequently sampled locations) and secondly because the observed click counts are extremely variable, likely reflecting variations in vaquita behavior in the vicinity of the CPODs (e.g., variation in animal speed, foraging behavior, etc.) – it was felt that using a model to "smooth out" this variability would result in more reliable inference about trend and provide a better assessment of the uncertainty associated with an estimates.

Description of Models

The Panel agreed to use Bayesian inference approaches for the main models used to estimate rate of change. There are many advantages of using Bayesian methods, but of particular value in the current context was the desire to obtain posterior probability distributions for annual rate-of-change, which in turn allow for straightforward estimation of the probability that the population declined between 2011 and 2014.

After consideration of numerous models, the Panel focused on two models with differing assumptions: the Spatial Model and the Non-Spatial Mixture Model. Here we describe the basis for these models with details in Appendix 2.

Spatial Model Description

The spatial model smoothed over the observed data, considering them to be a noisy version of an underlying smooth pattern of vaquita use. Vaquitas move throughout the study area, and the number of clicks encountered at a station are considered as an imperfect sampler due to stochastic movements of vaquitas. There is also unequal effort at locations, with some locations completely unsampled in some years. The model partitions variability into a spatially smooth surface plus independent random error, where the variance of the independent part decreases in proportion to increasing effort (number of sampling days). The estimated surface of vaquita use, then, is the predicted spatial surface. Each year is treated independently for predictions, but autocorrelation parameters are estimated by pooling across years.

The spatial model was a Gaussian log-linear mixed model (i.e., data assumed normal on log scale) with spatially autocorrelated error structure. Rationale for using this approach in favor of others is discussed below (see *Basis for model choice*). Details of this model are in Appendix 2. An overview is provided here.

The response variable data (W_{ti}) were the average number of clicks detected per day at each CPOD location i within a sampling year t. Thus the sample size for analysis was the sum of the number of CPODs functioning during the core sampling period in each year; this totaled 128 "CPOD-years". The data were transformed by adding 1 and taking the log of the values, i.e, Y_{ti}= log(W_{ti} + 1), because some functioning detectors recorded zero clicks during some years. The transformed data had reasonable variance:mean properties for using a Gaussian model (Appendix 3).

The transformed data were thus fit by the following model:

$Y_{ti} \sim \text{Normal}(\mu_t + Z_{ti}, \sigma_\varepsilon^2/n_{ti})$,

where μ_t is the expected mean number of clicks per day across locations in year t, Z_{ti} is a spatially autocorrelated random effect allowing the number of clicks per day at each location within a year to depart from the overall mean (with CPODs in closer proximity to each other expected to have more similar departures from the overall mean), and σ_ε^2 is the variance for spatially independent random error, weighted by variable sampling effort (number of CPOD-days, n_{ti}) across locations.

Details for estimating the spatial component of the model (Z_{ti}) are in Appendix 2. Worth noting here is that years were treated independently in the model, such that a different spatial surface was estimated from each year's data, but all years were assumed to have the same autocorrelation structure (same exponential decay in spatial random effect covariance as function of distance between locations). Also note that the spatial model is used to provide predictions for Y_{ti} at all K CPOD locations (K = 45), including those not sampled in some years, by drawing on information (through the spatial model parameters) from surrounding CPODs.

Inference was based on several summaries derived from the model parameter posterior distributions. Let S_{ti} be the predicted values for the average number of clicks per day (smoothed over the noisy process with variance σ_ε^2), back-transformed to the original scale of the data,

$S_{ti} = \exp(\mu_t + Z_{ti}) - 1$

An index of abundance (B_t) is taken to be the average of these values across all K CPOD locations for each year. Thus, given fitted estimates (predicted values) for S_{ti}:

$B_t = \frac{1}{K}\sum_{i=1}^{K} S_{ti}$.

An estimate of the geometric mean annual rate of population change between 2011 and 2014 is calculated as $\lambda = (B_{2014}/B_{2011})^{1/3}$. The proportion of the posterior distribution for this quantity that is less than 1 provides an estimate for the probability that the population in the sampled area has declined between 2011 and 2014.

Posterior summaries including means, medians, variances and credible intervals were obtained from MCMC samples. MCMC specifications (including priors) are detailed in Appendix 2.

Non-spatial Mixture Model Description

The non-spatial mixture model draws on the strength of the sampling design (repeated samples from a fixed semi-regular grid throughout the study area). Predicted click levels at individual CPOD locations were not based on a spatial model. Rather, within a generalized linear mixed model framework, individual CPOD locations were assigned probabilistically to one of $V = 3$ groups based on the level of detections they received across multiple years of sampling. Predictions for individual locations are given by estimated means and random effect variances for the groups to which CPOD locations are attributed.

The parameter of interest is $\theta_{v[k],t}$ the mean click rate (clicks per day) in year t for each of the V groups to which detector k is attributed. Because the data (total clicks per location per year, n_{kt}) were overly dispersed for a Poisson model, they were treated as negatively binomially distributed with the expectation given by the product of the estimated $\theta_{v[k],t}$ and effort (number of CPOD days, d_{kt}), i.e.,

$n_{kt} \sim$ Negative Binomial (p_{kt}, $r_{v[k],t}$),

where p and r are negative binomial parameters, and where $\mu_{kt} = \theta_{v[k],t} d_{kt} = r_{v[k],t}(1-p_{kt})/p_{kt}$ is the expectation for n_{kt}. Thus, variable sampling effort across CPOD locations is handled through its effect on the expectation and variance for n_{kt}.

Exploratory generalized additive model (GAM) analysis suggested that the click-rate data were well described by a negative binomial error distribution (see below).

Individual CPODs were probabilistically assigned to a use-intensity group v based on the data recorded at k across the years during which CPOD k was functioning. In OpenBUGS (Bayesian analysis software), this was done using the "categorical distribution" (multivariate generalization of the Bernoulli):

$v[k] \sim \text{cat}(\mathbf{s}_{vk})$,

where $\mathbf{s}_{vk}$ is the vector of estimated probabilities for k being in group v, which come from a Dirichlet prior distribution (see details in Appendix 2). The degree of certainty in assigning a CPOD location to a particular group depends on how correlated detections were through time; sites with consistently low or high levels of detections are assigned to a group with greater confidence, and all else being equal, CPODs with 4 years of data are assigned more confidently to a group that sites with one or two years of data. Uncertainty in group assignment is propagated through to estimates of other parameters.

In summary, the number of detections recorded across all CPODs are assumed to arise from a mixture of *V* negative binomial distributions. Information across years is shared for the purpose of assigning each CPOD location to a particular group *v*, but the means and variances for each *v, t* are independent. Predicted estimates for CPOD locations in years with missing data are based on the probability of belonging to group *v*, and the conditional mean and variance for group *v* in year *t*.

Inference is on the overall mean values for daily click rate (M_t), which are simply the means of the $\theta_{v[k],t}$ weighted by the number of CPODs belonging to each group *v*, for each *t*, i.e., $M_t = \frac{1}{K}\sum_{k=1}^{K} \theta_{v[k]t}$. The rate of change between 2011 and 2012 is M_2/M_1. The rate of change between 2012 and 2013 is M_3/M_2. The rate of change between 2013 and 2014 is M_4/M_3. The mean annual rate of change across years, $\bar{\lambda}$, is the geometric mean of these values. The probability that the population declined from 2011 to 2014 is the proportion of the Bayesian posterior distribution for $\bar{\lambda}$ that is less than 1 (or the probability that $\bar{\lambda}$ – 1 is less than zero). Inference about population change is based on posterior distribution summaries for these derived parameters.

Spatial GAM Models

In addition to the Bayesian models used to estimate the rate of change, the Panel agreed that a maximum likelihood approach (Generalized Additive Models) would be a useful comparison to the Bayesian methods. In preliminary analyses, Generalized Additive Models (GAMs) were developed to quickly evaluate and compare alternative models for estimating population change and to evaluate acoustic metrics before implementing those models in Bayesian spatial models. However, GAMs were not favored by the Panel as the approach for making inference because GAMs do not provide posterior probability estimates for key parameters of interest.

Five GAMs were fit to the 2011-2014 acoustic monitoring data with:

1) Site and year as independent categorical variables,
2) Site as a categorical variable and year as continuous (trend) variable,
3) A constant spatial smooth of location with year as a categorical variable,
4) A constant spatial smooth of location with year as a continuous (trend) variable, and
5) A year-varying spatial smooth of location.

All models were fit with the *mgcv* package in R. Spatial models used a two-dimensional thin-plate spline. The mean number of clicks per day for each site and year was used as the dependent variable with a negative binomial distribution and a log-link function. The best-fit model was selected with AIC.

Additional details on the GAMs are given in Appendix 2.

Basis for model choice

The Panel's charge was to give a best estimate of the current rate of change in vaquita detections. Although the spatial and mixture models gave similar results (see below), the Panel carefully considered the merits of each. Below we summarize the main differences between the two Bayesian approaches.

- The spatial model assumes that the spatial distribution of clicks is different each year but uses multiple years to estimate the spatial auto-correlation. The non-spatial mixture model assumes that each site falls (probabilistically) into categories of high, medium or low click density and that the probability of membership in these categories is shared between years for a given site.

- The spatial model uses information on site location to smooth over random spatial variations in click density. The non-spatial model uses no information on site location or proximity between sites.

- The spatial model assumes that the logarithm of mean clicks per day is normally distributed and the non-spatial model assumes that total click counts have a negative binomial distribution.

The Panel agreed that both approaches had merit and that averaging results of the two models would form the best basis for estimating rate of change.

Results and Discussion

Annual trends in vaquita clicks were first measured using the direct-count method, based only those sites that were sampled in both years of adjacent pairs of years. The direct-counts indicated a total change in the number of recorded clicks of 0% from 2011 to 2012, -33% from 2012 to 2013, and -53% from 2013 to 2014, which gives a geometric mean rate of -29.5% per year (negative changes are declines). However, as discussed previously, this method may be biased by non-random survey effort in space and time, and additionally does not provide any estimate of certainty in the true rate of change.

The exploratory GAM analysis showed that total clicks for each site and year could not be adequately modeled with common distribution functions (Poisson, negative binomial and Tweedie distributions). However, the negative binomial distribution provided a very good fit to mean clicks per day for each site and year (Appendix 2), and this distribution was used for subsequent analyses. An analysis with the entire summer dataset was compared to one based only the core sampling period (when at least 50% of CPODS were active in all years). Results showed that click rates trends differed for these two approaches. Of the two, the Panel decided to conduct remaining analyses and base inferences on the core sampling period data, to avoid potential biases caused by unbalanced spatial and temporal coverage in the full dataset (also see Seasonal Effects Section above).

GAM analyses were also used to explore two alternative acoustic measures of vaquita relative abundance: the mean number of minutes per day with vaquita clicks present (detection positive minutes – DPM) and the mean number of half-hour periods per day with vaquita clicks present (detection positive half-hours – DPHH). A negative binomial distribution function was used in a model that fit a common spatial pattern for all years. Results showed that the mean rates of decline for these two metrics were qualitatively similar to declines estimated using the Bayesian spatial model and non-spatial mixture model, but the model fit was not as good as with mean clicks per day (Appendix 2). DPM and DPHH only indicate the presence of vaquitas during a fixed time period and do not indicate the number of animals present. The vaquita distribution is very patchy, and these metrics tend to saturate at higher click count values (Figure 8) and are not thought to provide as much information on relative abundance as the number of clicks.

The Panel agreed to use the pooled posterior distributions from both the Bayesian spatial and non-spatial mixture models and to use posterior means as the central estimate. The average trend estimated from the spatial model is a change of -33% per year with a 95% posterior credibility interval from -56% to -11% per year, and the posterior probability of decline is 0.995. The estimated spatial density of vaquitas from the spatial model is illustrated in Figure 9, and the full posterior probability distribution is illustrated in Appendix 2. For the non-spatial mixture model, the average trend is a change of -28% per year. This non-spatial model gave a narrower 95% posterior credibility interval (from -43% to -10% per year, see Appendix 2 for the full posterior probability distribution) and a higher posterior probability of a decline (0.9998). Results of these two models are averaged by drawing equally from their respective Bayesian posterior samples for the growth rate parameter. **The model-averaged estimate for population change (Figure 10) has a mean rate of decline of 31%/year (95% Bayesian Credible Interval -51% – -10% per year). This decline is considerably worse than the previously estimated (CIRVA V) 18.5% per year (95% Bayesian Credible Interval -46% – +19% per year) because of the very low number of detections in 2014, which resulted in an estimated rate of decline from 2013 to 2014 of 42%. The Panel found that it is nearly certain the acoustic activity has declined (prob. = 0.996) with very high chance of a rate of decline greater than 10% per year (prob. = 0.976).**

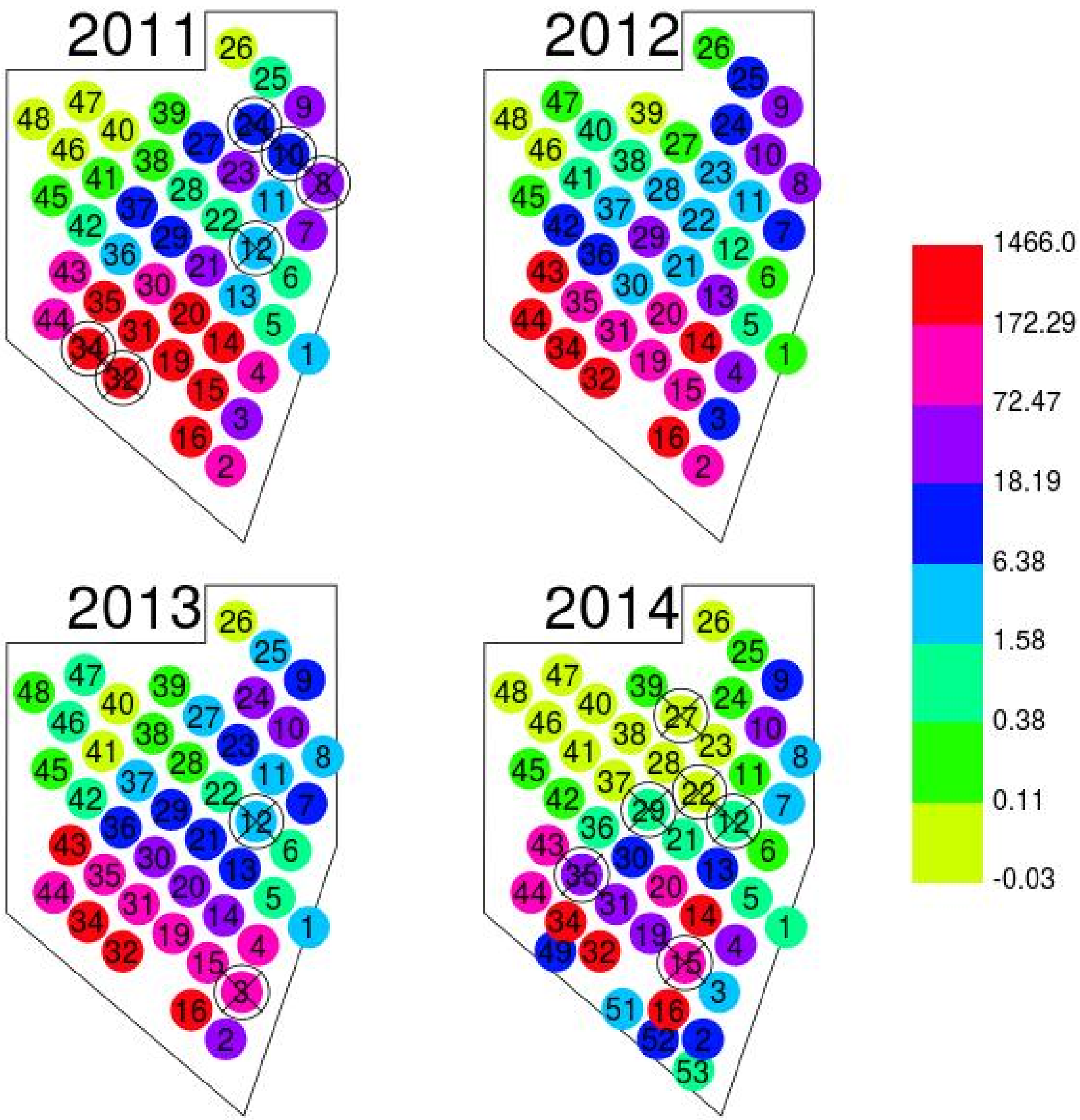


Figure 9. Estimated mean number of clicks per day predicted by the spatial model for the 45 C-POD sites with data for at least one year. Values are posterior medians. Sites with a circle/cross were missing in the indicated year. The analysis did not constrain the density surface to be the same each year. The sites 49—53 were added in 2014 (not included in the analysis) and show that clicks dropped off rapidly along the southwest border.

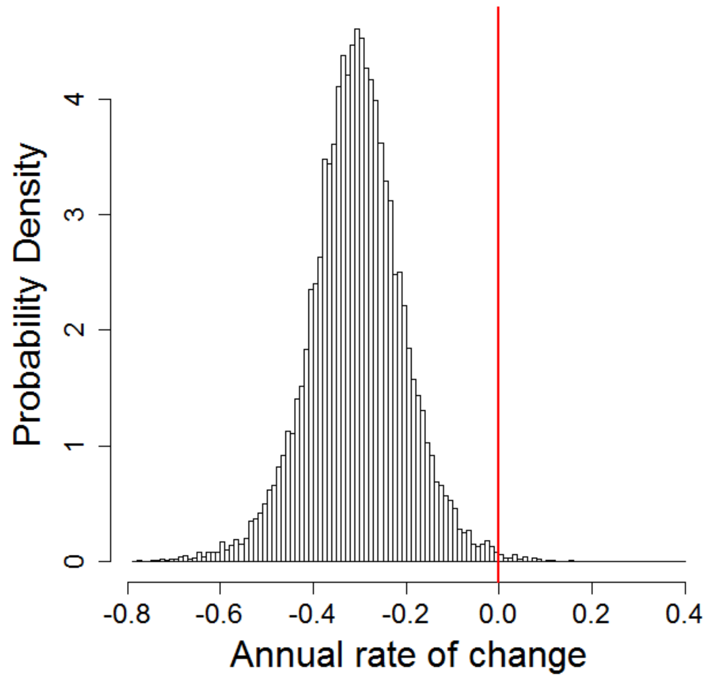


Figure 10. Posterior probability distribution from the pooled spatial and non-spatial mixture models. The mean is a -31% change (decline) per year.

The GAM analyses of the 2011-2014 acoustic monitoring generally gave higher rates of decline than the two Bayesian models (Table 2), especially those four models that assumed that the vaquita distribution did not change within the monitored area. The best-fit model was GAM #2 with site as a categorical variable and year as a continuous (trend) variable.

Table 2. Geometric mean annual rates of change from five GAM analyses using the 2011-2014 acoustic monitoring data.

GAM #	Variable Specification	Annual % Rate of Change	Delta-AIC
1	Site & year as categorical	-42	11.1
2	Site as categorical and year as continuous (trend)	-42	0.0
3	Spatial smooth with year as categorical	-43	40.0
4	Spatial smooth with year as continuous (trend)	-42	37.5
5	Year-specific spatial smooth	-15	126.3

Conclusions

The Panel agreed that the estimated rate of -31% should be considered as the best estimate of current rate of decline from the acoustic data alone. The greatest mean rate of decline for any annual increment was found between 2013 and 2014 by all three statistical approaches (spatial modeling –47%, non-spatial modeling –37%, GAMs model 3 -53%). This is consistent with reports of increased illegal fishing both for totoaba and within the Vaquita Refuge.

Acknowledgements

The workshop was funded by US Marine Mammal Commission. The research was funded between 2010 and 2013 by Mexico Minister of Environment, InstitutoNacional de Ecologia, The Ocean Foundation, Fonds de Dotation pour la Biodiversité, Cousteau Society, WWF México, WWF US, The Mohamed bin Zayed Species Conservation Fund and Barb Taylor. We thank the researchers dedicated to generating the data (in addition to AJL and LRB): G. Cardenas, E. Nieto, F.V. Esparza, M Sao, N. Tregenza. The participants expressed their gratitude to the Southwest Fisheries Science Center for hosting the workshop, to Dr. Lisa Balance for her welcoming remarks and Annette Henry for her support during the workshop. All participants were paid by their employers including the National Oceanographic and Atmospheric Administration, Instituto Nacional de Ecologia and the University of Saint Andrews. A special thanks to Francisco Valverde Esparza (monitoring program field team leader) for the effort to recover the C-POD deployed at site 32 during 2011 sampling period (returned by a fishermen June, 2014). Edwyna Nieto used her analysis skills to make data available for the final analyses provided in this report.

Literature Cited

CIRVA V reference

Diggle, Peter J., J. A. Tawn, and R. A. Moyeed. 1988. "Model-based geostatistics." Journal of the Royal Statistical Society: Series C (Applied Statistics) 47.3: 299-350.

Gerrodette, T, and L. Rojas-Bracho. 2011 Estimating the success of protected areas for vaquita (*Phocoena sinus*). Marine Mammal Science 4: E79-E100 DOI: 10.1111/j.1748-7692.2010.00438.x .

Gerrodette, T., B.L. Taylor, R. Swift, S. Rankin, A.M. Jaramillo-Legorreta and L. Rojas-Bracho. 2011. A combined visual and acoustic estimate of 2008 abundance, and change in abundance since 1997, for the vaquita, Phocoena sinus. *Marine Mammal Science* DOI: 10.1111/j.1748-7692.2010.00438.x.

Hohn, A.A., A.J. Read, S. Fernández, O. Vidal y L.T. Findley. 1996. Life history of the vaquita, *Phocoena sinus* (Phocoenidae, Cetacea). *Journal of Zoology (London)* 239:235-251.

Jaramillo-Legorreta, A.M., L. Rojas-Bracho y T. Gerrodette. 1999. A New Abundance Estimate for Vaquitas: First Step for Recovery. *Marine Mammal Science* 15:957-973.

Jaramillo-Legorreta, A.M. 2008.*Estatus actual de unaespecie en peligro de extinción, la vaquita (Phocoena sinus): Unaaproximaciónpoblacional con métodosacústicos y bayesianos.*Tesis Doctoral. Facultad de CienciasMarinas.UniversidadAutónoma de Baja California. Ensenada, Baja California, México. 115 pp.

Jaramillo-Legorreta, A.,L. Rojas-Bracho, G. Cardenas, E. Nieto, F.V. Esparza, M Sao, B.L. Taylor, J. Barlow, T. Gerrodette, A. Henry, N. Tregenza, T. Ragen. 2014. Report of the Second Meeting of the Steering Committee of the Vaquita Acoustic Monitoring Program April 24-25, 2014.

Rojas-Bracho, L., A. Jaramillo-Legorreta, G. Cardenas, E. Nieto, P. Ladron de Guevara, B.L. Taylor, J. Barlow, T. Gerrodette, A. Henry, N. Tregenza, R. Swift, and T. Akamatsu. 2010. *Assessing trends in abundance for vaquita using acoustic monitoring: within refuge plan and outside refuge research needs*. U.S. Department of Commerce, NOAA Technical Memorandum NMFS, NOAA-TM-NMFS-SWFSC-459. 39 pp.

Woodley, T.H. y A.J. Read. 1991. Potential rates of increase of a harbour porpoise (*Phocoena phocoena*) population subjected to incidental mortality in commercial fisheries. *Canadian Journal of Fisheries and Aquatic Sciences* 48:2429-2435.

Appendix 1: **Brief biographies of the Expert Panel**

Dr. Armando Jaramillo-Legorreta was raised in Mexico City and received his bachelor's degree at La Paz, Baja California Sur with a focus on marine biology. His main research interest since 1986 has been in the study of ecology and dynamics of marine mammal populations. He received his Masters and PhD degrees in Baja California with a research focuses on coastal oceanography, population ecology and population dynamics modelling. From 1996 to the present day, he is a researcher for the Marine Mammals Research and Conservation Group of the National Institute of Ecology in charge of the study of habitat use and acoustic monitoring of vaquitas. He was the lead author of the first estimate of abundance of vaquitas in 1997 and the first acoustic monitoring between 1997 and 2008 that informed Mexican Government of the decline of vaquita population. Since 2009 has led the current acoustic monitoring scheme. He has coauthored about 20 papers and chapters on different aspects of marine mammals as well as many technical reports. He is delegate for Mexico at the Scientific Committee of the International Whaling Commission and an advisor on the Mexican National Commission for Biodiversity. He is the current President of the Mexican Society of Marine Mammals.

Dr. Jay Barlow is a research scientist within the Marine Mammal and Turtle Division, SWFSC, La Jolla, where he has worked for 32 years. Jay received his PhD from Scripps Institution of Oceanography (SIO) in 1982. He is the leader of the EEZ Marine Mammals and Acoustics Program within PRD and is an Adjunct Professor at SIO. Dr. Barlow's research involves assessing human impacts on marine mammal populations, estimating their abundance and dynamics, the understanding role of mammals in marine ecosystems, and developing survey methods that use passive acoustics to detect and localize cetaceans. He currently is advisor of three PhD students and serves on dissertation committees of four others. At SIO, Jay teaches a 4-unit course called "Computer-intensive Statistics". Jay serves on many advisory committees both within NOAA (e.g., the NMFS Steering Committee on Assessing Acoustic Impacts on Marine Mammals and the Humpback Whale Biological Review Team) and internationally (e.g., the IUCN Cetacean Specialist Group). Jay has authored or co-authored 110 peer-reviewed journal articles and book chapters, 75 numbered government reports, and edited one book. He has been chief scientist on 12 NOAA and one Australian research surveys.

Dr. Jay VerHoef began his career as a statistician with the Alaska Department of Fish and Game, after receiving a co-major Ph.D. in statistics and ecology and evolutionary biology from Iowa State University. He now works as a statistician for a research lab, the National Marine Mammal Laboratory within the National Marine Fisheries Service of NOAA. For over 25 years, Dr. VerHoef has developed statistical methods and consulted on a wide variety of topics related to plant, animal, and environmental statistics. Dr. VerHoef is a fellow of the American Statistical Association (ASA) and past-Chair of the Section on Statistics and the Environment of ASA. He has over 100 publications, and he is a co-author of a book on spatial statistics. His CV can be found here https://sites.google.com/site/jayverhoef/Home/cv

Dr. Jeff Moore is a research scientist within the Marine Mammal and Sea Turtle Division of the NOAA Southwest Fisheries Science Center in La Jolla, California, where he has worked for five years. Jeff has B.Sc., M.Sc., and Ph.D. degrees in Wildlife Biology from UC Davis, Humboldt State, and Purdue University. Prior to coming to SWFSC, Jeff held post-doc and research faculty appointments at Duke University for four years, where he worked on a global fisheries bycatch assessment for marine 'megafauna' and developing methods for quantifying population impacts of bycatch on long-lived species, and for assessing interactions between developing country small-scale fisheries and coastal marine mammals and sea turtles. Jeff's current research involves assessing human impacts on marine vertebrate populations, estimating abundance and population dynamics parameters using Bayesian statistical methods, and developing quantitative tools to aid management and policy decisions. Jeff serves on advisory committees such as the Biological Review Team for reviewing the status of northeastern Pacific white sharks, and the IUCN Cetacean Specialist Group. He has authored or co-authored > 30 peer-reviewed scientific journal articles since 2004 (3/yr) in addition to numerous NOAA agency technical reports.

Dr. Len Thomas is a senior faculty member within the School of Mathematics and Statistics at the University of St. Andrews, Scotland. He is also director of the world-leading Centre for Research into Ecological and Environmental Modelling (CREEM), an inter-disciplinary research group at the interface between ecology and statistics. Two relevant major focuses of Len's research are statistical methods for population trend estimation (which has been working on since his PhD, at University of British Colombia, Canada, from 1993) and inferences from passive acoustic monitoring of cetaceans (which has been a major topic of research since 2007). One component of the latter has been his involvement in the design and analysis of the SAMBAH survey, a multi-national passive acoustic survey designed to estimate density of harbour porpoise in the Baltic by deploying CPODs at more than 300 sampling locations over a 2 year period, and performing associated calibration experiments. Over the past 21 years, Len has co-authored 73 peer-reviewed papers, 3 books, and a further 57 other publications or technical reports. He has been a keynote speaker at several major international conferences, most recently the International Statistical Ecology Conference (2012, topic: "The future of statistical ecology") and the European Cetacean Society Conference (2013, topic: "Interdisciplinary approaches in the study of marine mammals: ecology meets statistics").

Appendix 2: **Model details**

SPATIAL MODELING OF VAQUITA ACCOUSTIC DATA FROM 2011 – 2014.

Let $W_t(\mathbf{s}_i)$ denote a random variable for mean acoustic click counts at the ith spatial location in the tth year. Because some of the data were zero, we used $Y_t(\mathbf{s}_i) = \log(W_t(\mathbf{s}_i) + 1)$ for analysis.

To account for uneven effort per site, we divided the spatial model into a spatially structured component and an independent component (often called the nugget effect by geostatisticians). Then the set $\{Y_t(\mathbf{s}_i)\}$ were treated as spatially autocorrelated in a spatial linear mixed model,

$$[Y_t(\mathbf{s}_i) \mid \mu_t, Z_t(\mathbf{s}_i), \sigma_\varepsilon^2, n_{t,i}] = N(\mu_t + Z_t(\mathbf{s}_i), \sigma_\varepsilon^2 / n_{t,i}) \qquad \text{(A.1)}$$

where $n_{t,i}$ is the number of sampling days for each site for each year. The $n_{t,i}$ account for uneven sampling, and this can be also be viewed as measurement or sampling error in a hierarchical model. Let the vector $\mathbf{z}_t$ denotes all of the spatial random effects, $Z_t(\mathbf{s}_i)$, for the tth year,

$$[\mathbf{z}_t \mid \sigma_z^2, \rho] = N(\mathbf{0}, \sigma_z^2 \mathbf{R}_t(\rho)),$$

where we assumed that years were independent, but that the spatial stochastic process had the same autocorrelation model among years; that is,

$$\text{cov}\begin{pmatrix} \mathbf{z}_{2011} \\ \mathbf{z}_{2012} \\ \mathbf{z}_{2013} \\ \mathbf{z}_{2014} \end{pmatrix} = \sigma_z^2 \begin{pmatrix} \mathbf{R}_{2011}(\rho) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{2012}(\rho) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_{2013}(\rho) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R}_{2014}(\rho) \end{pmatrix}.$$

For spatial autocorrelation, we used the exponential model,

$$\text{corr}(Z_t(\mathbf{s}), Z_u(\mathbf{v})) = \begin{cases} \exp(-3h/\rho) & \text{for} \quad t = u, \\ 0 & \text{for} \quad t \neq u, \end{cases}$$

where h is Euclidean distance. That is, let $\mathbf{s} = (s_x, s_y)$ be the x- and y-coordinates of one point, and $\mathbf{v} = (v_x, v_y)$ be the x- and y-coordinates of another point, then

$$h = \sqrt{(s_x - v_x)^2 + (s_y - v_y)^2}\ .$$

For the spatial analysis, latitude and longitude coordinates were projected onto the plane using a Universal Transversal Mercator (UTM) projection with a user-defined central meridian. The central meridian was computed as the center of the vaquita refuge. This minimizes distortion from the projection, and UTM is a distance-preserving projection. After projection, the UTM coordinates were converted from meters to kilometers, and translated in space so that 0 on the x-coordinate corresponded with the western-most coordinate of the vaquita refuge, and 0 on the y-coordinate corresponded with the southern-most coordinate of the vaquita refuge.

To complete the model, we specified the following prior distributions,

$$\mu_{2011} \sim \text{UNIF}(-10,10)$$
$$\mu_{2012} \sim \text{UNIF}(-10,10)$$

$$\mu_{2013} \sim \text{UNIF}(-10,10)$$
$$\mu_{2014} \sim \text{UNIF}(-10,10)$$

$$\sigma_z \sim \text{UNIF}(0,10)$$
$$\sigma_\varepsilon \sim \text{UNIF}(0,10)$$
$$\rho \sim \text{UNIF}(0,500)$$

Because the data were modeled on the log-scale, these are flat and non-informative priors that encompassed any reasonable range of values for the parameters. The posterior distribution of the model is,

$$[\sigma_\varepsilon, \sigma_z, \rho, \mathbf{z}, \boldsymbol{\mu} \mid \mathbf{y}]. \tag{A.2}$$

We used Markov chain Monte Carlo (MCMC) methods, using the software package WinBUGS, to obtain a sample from the posterior distribution (A.2). We used a burn-in of 10,000 iterations, and then used 1,000,000 further iterations. For computer storage reasons, we kept a single iteration out of each 100, yielding a sample of 10,000 from the posterior distribution.

We were interested in several summaries derived from the posterior distribution. Let

$$\hat{S}_t^k(\mathbf{s}_i) = \exp[\hat{\mu}_t^k + \hat{Z}_t^k(\mathbf{s}_i)] - 1$$

be a spatially smoothed prediction for the *t*th year, at the *i*th site, and for the *k*th MCMC sample. Notice that these predictions smooth over the noisy process with variance σ_ε^2 contained in the model specification at the data level, and that we are putting the predictions back on the original scale of the data. Then, we take as an indicator of overall abundance, among all *n* sites for each year, as

$$\hat{B}_t^k = \frac{1}{n}\sum_{i=1}^{n} \hat{S}_t^k(\mathbf{s}_i).$$

Finally, we were interested in average rate of change, as a proportion, for the two time increments. We decided to use the geometric mean[1],

$$\hat{r}^k = \left(\frac{\hat{B}_{2012}^k}{\hat{B}_{2011}^k}\frac{\hat{B}_{2013}^k}{\hat{B}_{2012}^k}\frac{\hat{B}_{2014}^k}{\hat{B}_{2013}^k}\right)^{1/3} = \left(\frac{\hat{B}_{2014}^k}{\hat{B}_{2011}^k}\right)^{1/3},$$

and based on this, the posterior probability of a decreasing population can be computed from the mean of

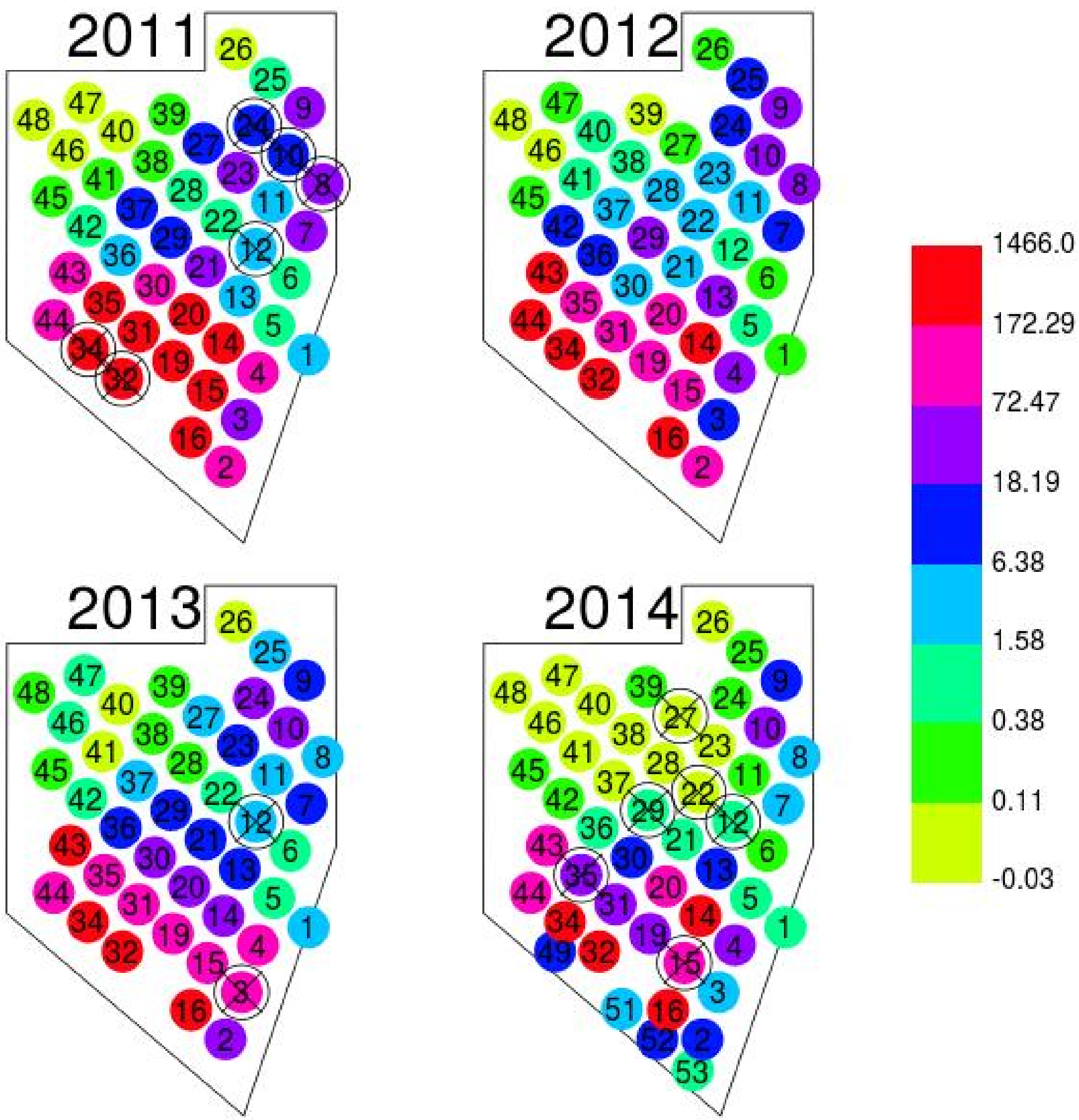
$$\hat{p}^k = I(\hat{r}^k < 1),$$

where *I*(.) is the indicator function. Posterior summaries including means, medians, and variances of $\hat{S}_t^k(\mathbf{s}_i)$, $\hat{B}_t^k$, $\hat{r}^k$, and $\hat{p}^k$, were obtained from the MCMC samples.

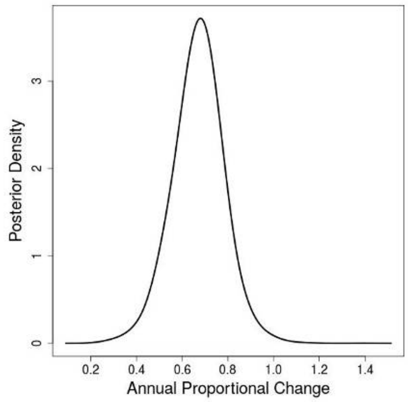
[1] Note that $\hat{r}^k$ is the parameter for proportional rate of change which is referred to using the symbol λ elsewhere in this Appendix and the body of the report.

RESULTS

Maps of $\hat{S}_t^k(\mathbf{s}_i)$ for each year and location are given below, where we used the median from the MCMC sample. The sites in 2011 - 2014 with circles around them and an 'x' through the circle indicate that data were missing for those years, so these are spatially interpolated values. Because modeling occurred on the log-scale, these missing values in particular had a wider variance, which had a large effect on the mean value when taking exponents to get back on the original scale of the data. So, for presentation purposes, we used the median. Also note that sites numbered 49 - 53 only occurred in 2014. There was some concern that the refuge was not capturing the majority of the vaquita population because, from 2011 - 2013, the highest abundance peaked along the southwest edge. The sites 49 - 53 were not included in the analysis, but their mean clicks per day are plotted, showing that clicks dropped off rapidly along the southwest border.



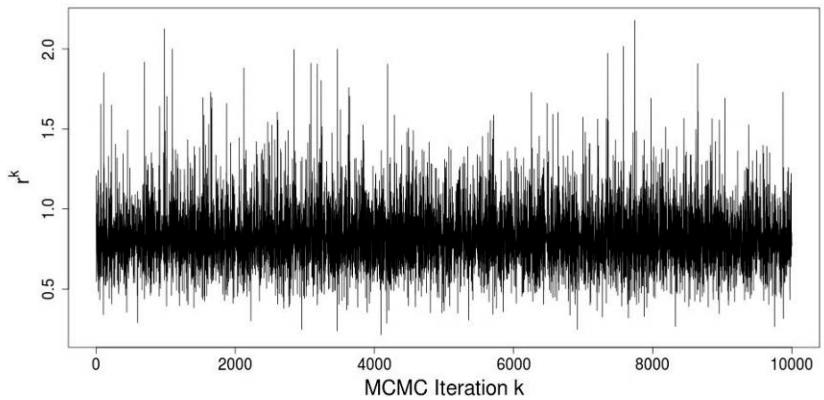
The posterior distribution of the annual proportional change, $\hat{r}^k$, is given below,



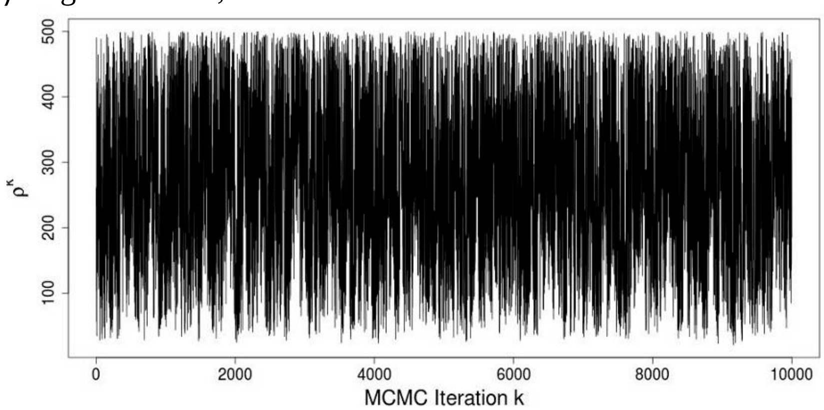
The mean of the posterior distribution for $\hat{r}^k$ was 0.670, and the median was 0.674, indicating about 33% per year decrease in clicks. The 95% credibility interval, based on the 2.5% and 97.5% quantiles, was 0.444 to 0.891. The probability $\hat{r}^k$ was less than one, i.e., $\hat{p}^k$, was 0.995. For annual increments, the mean of the posterior distribution for $\hat{B}^k_{2012} / \hat{B}^k_{2011}$ was 0.712 with a 95% credibility interval from 0.193 to 1.439; the mean of the posterior distribution for $\hat{B}^k_{2013} / \hat{B}^k_{2012}$ was 1.058 with a 95% credibility interval from 0.440 to 2.534; the mean of the posterior distribution for $\hat{B}^k_{2014} / \hat{B}^k_{2013}$ was 0.531 with a 95% credibility interval from 0.179 to 1.149.

ASSESSING THE MODEL AND MCMC

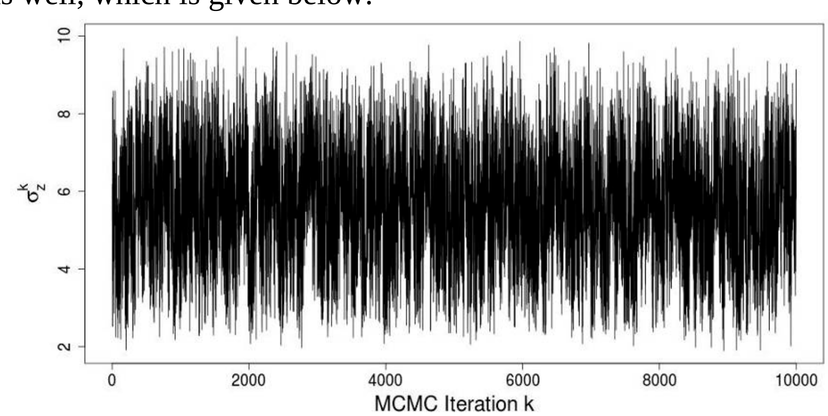
Our primary goal was to obtain a sample of $\hat{r}^k$ in order to project the current population estimate from 2010. To test for convergence in the MCMC chain, we used the Geweke test, found in the R coda package. The result was a *z*-value of 0.863, which is assumed to be a standard normal random variate under the assumption that the MCMC sample is from a stationary distribution. Our result indicates very little reason to be concerned that this particular MCMC chain had not converged. The MCMC trace is shown below.



The trace of ρ is given below,

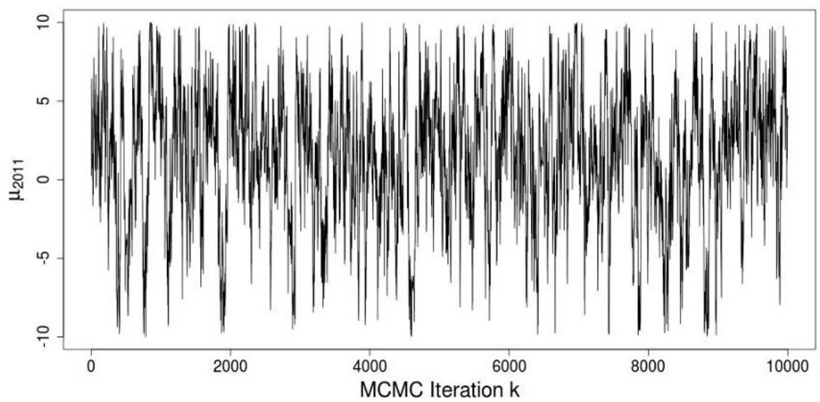


Note that values seem to be truncated by the prior, which has an upper bound of 500. We did a sensitivity analysis, and increased it to 1000. Part of the explanation requires the trace of σ_z as well, which is given below.



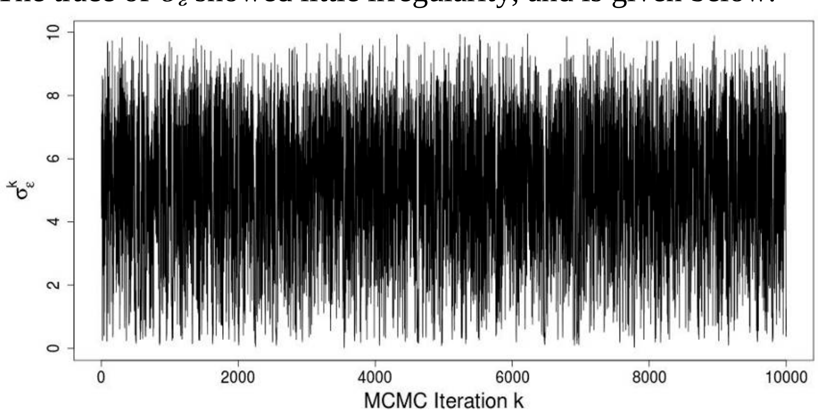
When ρ is increased to 1000, then σ_z becomes truncated by its upper bound of 10. This is a well-known phenomenon in spatial statistics, where the model explores a more linear form of the autocorrelation function by increasing both ρ and σ_z. In fact, the correlation between them, in the MCMC sample, is near 0.86. However, very large values of ρ and σ_z, when they occur together, have little effect on the autocorrelation within the spatial distances seen within the data set. We saw no change in our inferences by continuing to increase either ρ or σ_z, because eventually one of them would become truncated at their upper bound. We left the upper bound for the prior of ρ as 500 (km), as that allowed a lot of autocorrelation among sites, and was far beyond the maximum distance among plots in the vaquita refuge.

The trace of the mean parameters μ_{2011}, μ_{2012}, and μ_{2013} also wander throughout their whole prior distribution. This is shown as a trace of the MCMC sample for μ_{2011} in the following figure,



This may seem strange at first, especially since even the raw data (on the log scale) do not range from -10 to 10. The explanation lies in the fact that spatially autocorrelated random variables, such as the $Z_t(\mathbf{s}_i)$, can wander far from their mean of zero, so the whole set $\{Z_t(\mathbf{s}_i); i = 1, \ldots n\}$ may be positive or negative. To examine this effect, we just chose $Z_{2011}^k(\mathbf{s}_1)$ from the MCMC chain, and its correlation with μ_{2011}^k was -0.988. Thus, the MCMC sampler was behaving as expected.

The trace of σ_ε showed little irregularity, and is given below.



OTHER SPATIAL AND TEMPORAL MODELING CONSIDERATIONS

Table 3 shows the raw data used for spatial modeling. We tried several spatial models, including embedding the spatial linear model into a generalized linear model (called model-based geostatistics by Diggle et al.), where the untransformed data, conditional on the mean, followed a Poisson or negative binomial distribution. However, estimation of site mean values, and even means over sites, was very unstable resulting in average click rates per year, such as $\hat{B}_t^k$, that were often in the thousands.

We also considered a spatial model where the spatial random effects were constant across years, so that the conditional mean in (A.1) was $\mu_t + Z(\mathbf{s}_i)$ rather than $\mu_t + Z_t(\mathbf{s}_i)$. This resulted in much steeper rates of decline, with a mean $\hat{r}^k$ of nearer 0.7 rather than 0.8. The reason can be seen in Table 3, and in particular if we focus on site 34. If the random effects are held constant through the years, then the predicted values in 2011 will largely follow the pattern seen in 2012 and 2013. For 2012 and 2013, site 34 was one of the highest sites, so when that "surface" is shifted to 2011, the predicted values had average values that were nearer 900 to 1000, rather than around 300 seen in Table 3. We felt that it was a strong assumption to hold the spatial surface constant across years, so we rejected the use of that model. Although there are very few data to look at yearly trend (only 2 years for site 34) within site, the current model fits the general trend.

Table 3. Mean click rates per site for each year, along with sampling effort. Median values for $\hat{S}_t^k(\mathbf{s}_i)$ are shown in bold red for missing C-PODs in those years.

Site	Mean Clicks 2011	Mean Clicks 2012	Mean Clicks 2013	Sample Days 2011	Sample Days 2012	Sample Days 2013
1	6.05	0.00	2.00	62	60	58
2	157.12	75.13	52.05	41	60	58
3	56.60	9.27	**75.87**	62	60	
4	229.56	26.88	152.48	62	60	58
5	0.00	0.00	0.00	62	60	58
6	0.15	0.00	0.21	62	60	58
7	96.79	4.22	24.35	62	60	37
8	**21.36**	55.20	1.67		60	21
9	24.32	61.67	11.45	62	60	62
10	**13.82**	59.73	38.58		60	62
11	0.71	2.78	2.76	62	60	37
12	**1.91**	0.00	**2.86**		62	
13	0.40	81.65	6.66	58	60	62
14	1781.57	3800.00	83.48	58	60	62
15	158.75	83.57	83.98	57	60	62
16	808.15	287.40	218.06	62	60	62
19	694.00	81.68	23.40	62	59	10
20	365.56	116.37	14.00	62	59	29
21	48.36	1.47	13.78	58	59	37
22	0.00	1.53	0.64	62	59	55
23	37.47	3.73	19.55	62	59	62
24	**7.41**	13.31	37.56		59	62
25	0.00	17.47	1.76	49	62	62
26	0.00	0.00	0.00	52	60	62
27	12.65	0.00	4.66	62	54	62
28	0.00	2.84	0.00	62	62	62
29	10.33	53.81	15.82	57	62	62
30	84.79	3.02	34.58	62	62	62
31	548.44	136.71	115.95	62	62	42
32	527.70	695.37	2116.02	20	62	62
34	**311.95**	408.94	729.91		62	11
35	413.58	77.68	148.84	62	62	50
36	0.67	8.65	4.14	48	62	44
37	26.77	1.82	4.82	47	62	45
38	0.00	0.69	0.29	62	62	62
39	0.00	0.00	0.00	61	55	62
40	0.00	1.37	0.00	62	62	62
41	0.00	0.68	0.00	54	34	62
42	0.00	9.36	0.00	46	61	62
43	252.53	595.46	462.84	62	61	62
44	70.58	172.33	141.65	62	61	62
45	0.00	0.00	0.00	62	61	62
46	0.00	0.00	2.45	48	61	49
47	0.00	0.38	0.56	57	61	62
48	0.00	0.00	0.00	43	61	54

Non-spatial Mixture Model

Rationale: This approach attempts to draw on the strength of the sampling design; Spatial autocorrelation is not modeled.

Basic assumptions:

1. CPOD locations are representative of a sampled area that we wish to make inference about.

2. The mean number of clicks-per-effort-day for a CPOD is linearly related to the amount of use in the area considered to be sampled by that CPOD. Thus clicks-per-effort-day is taken as an index of use-days in the area.

If all CPOD locations had equivalent sampling effort, we could simply take the mean "clicks per effort-day" across CPODs in year *t* as a robust estimate of the use-index for that year. Inference would be based on comparing the means between years and assessing the probability that they are different (which would depend on the variances of the estimates).

However, data are missing for some CPOD locations in some years (call these missing "CPOD-years"), and precision of the overall mean detection rate could potentially be improved (thereby increasing the power to detect annual changes) by accounting for spatial heterogeneity in CPOD detection rates. Therefore, interpolating the value of the use-index for missing CPOD-years and improving precision in the annual estimates for the use-index are the analysis objectives.

Data

n_{kt} = number of clicks recorded at location k, year t

d_{kt} = number of effort-days at location k, year t

K = 45 = total number of CPOD locations with effort in at least one year

The data are truncated in time, i.e., only using recorded clicks and effort-days between Julian dates 170 and 231 (inclusive).

Model

The non-spatial mixture model draws on the strength of the sampling design (repeat samples from a fixed semi-regular grid throughout the study area), emphasizing a design-based rather than model-based approach to inference. Predicted click levels (mean number of clicks per season, n_{kt}) at individual CPOD locations are not based on a spatial model. Rather, within a generalized linear mixed model framework, individual CPOD locations are assigned probabilistically to one of V = 3 groups based on the level of detections they received across multiple years of sampling.

Predictions for individual locations are given by estimated means for the groups to which CPOD locations are attributed, i.e.,

$$n_{kt} \sim \text{Negative Binomial}\ (p_{kt}, r_{v[k],t}),$$

where p and r are negative binomial parameters. Exploratory generalized linear model (GAM) analysis suggested that the click-rate data were well described by a negative binomial error distribution (see GAM section below). The expectation for n_{kt} (which we denote μ_{kt}) is a function of the expected mean number of clicks per day ($\theta_{v[k],t}$) and sampling effort (d_{kt}). The former depends on the group membership for CPOD k and the year:

$$\mu_{kt} = \theta_{v[k],t} d_{kt}.$$

For the negative binomial, the expectation $\mu_{kt} = r_{v[k],t}(1-p_{kt})/p_{kt}$. We placed priors on $\theta_{v[k],t}$ and $r_{v[k],t}$ (see below), so that in each MCMC iteration, the value for $p_{kt} = r_{v[k],t}/(r_{v[k],t} + \mu_{kt})$.

CPOD location k is probabilistically assigned to a use-intensity group v based on the data recorded at k across the years during which CPOD k was functioning. In OpenBUGS, this was done using the "categorical distribution" (multivariate generalization of the Bernoulli):

$$v[k] \sim \text{cat}(\mathbf{s}_{vk}),$$

where $\mathbf{s}_{vk}$ is the vector of probabilities for k being in group v, which come from a Dirichlet prior distribution:

$$\mathbf{s}_{vk} \sim \text{Dirichlet}(\boldsymbol{\alpha}_v),$$

where $\boldsymbol{\alpha}_v$ are the Dirichlet intensity parameters. Setting $\alpha_1 = \alpha_2 = \alpha_3 = 1$ makes this distribution fairly uninformative, providing the flexibility for s_{vk} to take on any values that sum to 1 (across v for each k).

The degree of certainty in assigning a CPOD location to a particular group depends on how correlated detections are through time; sites with consistently low or high levels of detections (or with more years of information, since there were some missing CPOD-years) are assigned to a group with greater confidence. Uncertainty in group assignment is propagated through to estimates of other parameters.
In short, the number of detections recorded across all CPODs are assumed to arise from a mixture of V negative binomial distributions in each year. Information across years is shared for the purpose of assigning each CPOD location to a particular group v, but the means and variances for each v, t are independent. Predicted estimates for CPOD locations in years with missing data are based on the probability of belonging to group v, and the conditional expected mean and variance for group v in year t.

Inference is on the overall mean values for daily click rate (M_t), which are simply the means of the $\theta_{v[k],t}$ weighted by the number of CPODs belonging to each group v, for each t, i.e., $M_t = \frac{1}{K}\sum_{k=1}^{K}\theta_{v[k]t}$. The rate of change between 2011 and 2012 is M_2/M_1. The rate of change between 2012 and 2013 is M_3/M_2. The rate of change between 2013 and 2014 is M_4/M_3. The mean annual rate of change across years, $\bar{\lambda}$, is the geometric mean of these values. The probability that the population declined from 2011 to 2014 is the proportion of the Bayesian posterior distribution for $\bar{\lambda}$ that is less than 1 (or the probability that $\bar{\lambda} - 1$ is less than zero). Inference about population change is based on posterior distribution summaries for these derived parameters.

Additional assumptions

In addition to the basic assumptions above, we note the following:

1. We used V = 3 groups based on visual inspection of the data, which indicates locations for which the mean number of clicks per effort day is consistently extremely low (just a few clicks/day), very high (clicks/day = hundreds to low thousands), or in-between (clicks per day = tens). Using fewer groups, such as V = 1 (single group, no mixture), ignores this information, potentially biasing estimates of $\mu_{k,t}$ for missing CPOD-years (and hence for the M_t). On the other hand, assuming many groups (V > 3) may result in over-fitting the data, reducing precision in the estimates of $\mu_{k,t}$ and thus increasing uncertainty in M_t. In practice, data generated by a mixture of many processes tend to be well approximated by mixture models with just a few groups.

2. Justification for this general approach relies on the assumption that there are stable high-use and low-use areas through time, i.e., *on average*, locations with the highest click-rates in three years will also have the highest click rates in the fourth year. However, the assumption, as modeled, allows for some flexibility in how the implied spatial patterns of vaquitas vary through time. First, inconsistent relative use of individual detectors would result in uncertain group assignment and thus greater uncertainty in the expected click rates. Second, the mean click-rate differences between groupings are estimated independently for each year. Thus, for example, the mean click rate for "medium use" CPODs could theoretically be much higher than "low use" CPODS in one year but only slightly higher in another year. Simple Spearman correlations suggest that it is indeed reasonable to assume that relative use across individual CPODs was similar through time (e.g., $rs_{2011,2012} = 0.77$; $rs_{2012,2013} = 0.93$; $rs_{2011,2013} = 0.83$). Similarly, high certainty in the assignment of most CPODs to a particular one of the *V* groups (see below) provided additional support for this assumption.

3. In contrast with spatial models, we are not borrowing information from surrounding CPODs to estimate values for CPOD *k*. All CPOD locations are treated as independent sample locations. The expected value for CPOD *k,t* depends on which group *k* belongs to (which is informed by data in other years at *k*) and on the mean and variance parameters for the group in year *t* (which are informed by other CPODs in the same group, but irrespective of their proximity to *k*).

MCMC specifications

An MCMC chain of length 1,000,000 was run. The first 500,000 samples were discarded. Every 100th sample from the chain was retained, so that the posterior distributions were constructed from 10,000 samples.
The following prior distributions were used:

$\mathbf{s}_{vk}$ ~ Dirichlet(1, 1, 1) # Probability of CPOD k belonging to group v
$\log(\theta_{v[k],t})$ ~ Normal(−10, σ^2=1000), for v = 1;
$\theta_{v[k],t} = \theta_{v-1[k],t} + \exp[\Delta\log(\theta_{v[k],t})]$, for $v = 2, 3$
$\Delta\log(\theta_{v[k],t})$ ~ Normal(5, σ^2=1000) (left-truncated at zero to be positive)
$r_{v[k],t}$ ~ Categorical($\mathbf{z}$)[2], where $\mathbf{z}$ is a vector of probabilities for r = integers from 1 to 10;
z_r ~ Dirichlet(1) for all r
$p_{kt} = r_{v[k],t}/(r_{v[k],t} + \mu_{kt})$ # Negative binomial parameter

Results

Most CPODs were attributed to mixture group with high probability, though assignment was less clear (but still fairly confident) in a few cases (see examples in Figure 11).

[2] In WinBUGS and OpenBUGS, the negative binomial r parameter must be an integer ≥ 1.

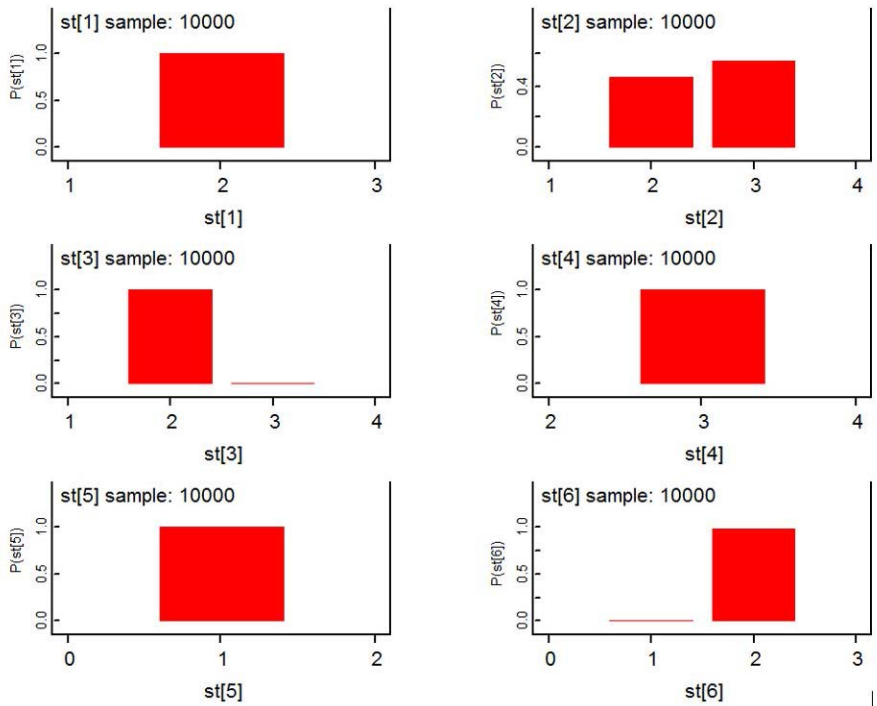


Figure 11. Sample OpenBugs output. Posterior densities for assignment of individual CPODs to one of three mixture groups. CPODs 1 through 6 shown here for example. CPODs 1, 3, and 6 were assigned to group 2 (medium-use) with high certainty. Detectors 4 and 5were assigned to groups 1 (low-use) and 3 (high-use), respectively, with high certainty. Assignment of CPOD2 was uncertain (similar probability of belonging to the medium- or high-use group).

Figure 12 shows annual predictions of mean click rate (average number of clicks per day) for the 45 CPODs that functioned in at least one year. Values depend on the mixture group to which the CPOD is most commonly assigned. Assignment of CPODs to mixture groups was generally clear. Detectors receiving almost no clicks in any year were assigned to one group; detectors receiving on the order of tens of clicks per day were assigned to a different group; and detectors receiving an average of hundreds of clicks per day in at least one year tended to be assigned to the third group. This third group was the most variable; hence the expected clicks/day for CPODs in this group had the highest variance, as indicated by broader credible interval bars, but overall the pattern of residuals indicated reasonable fit of this model to the data.

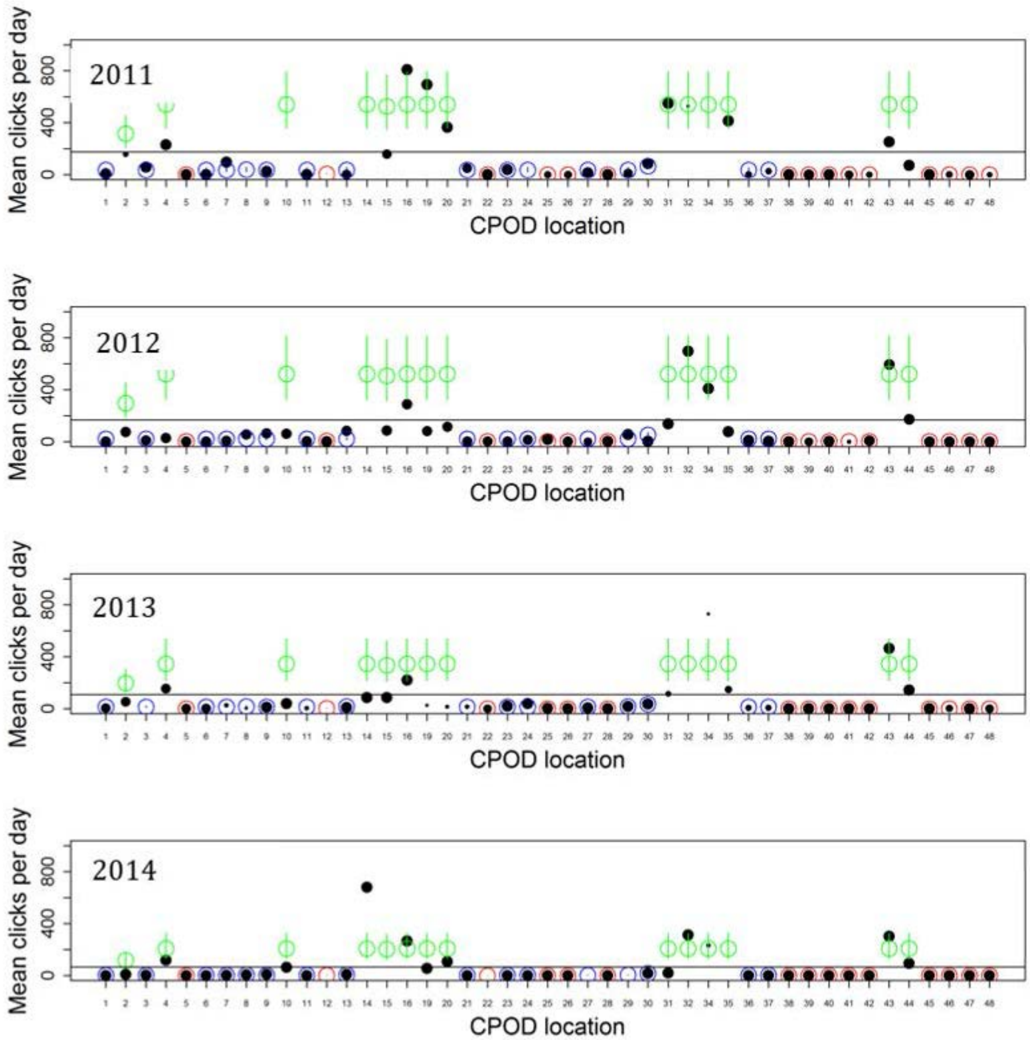


Figure 12. A) Observed and expected values for "mean clicks per day" at each CPOD location that functioned in at least one year, 2011 to 2014. Solid black points are the observed values, with point size indicating the relative level of effort (large circles = more days of sampling). Open circles are the model-expected values (with 90% credible intervals), $\theta_{v[k],t}$ *,for the three mixture groups (with most likely group indicated by different colors). Horizontal black line is the estimated overall mean for the year,* M_t*. Here, the y-axis only goes to 1000 (so that lower estimates may be visually resolved).*

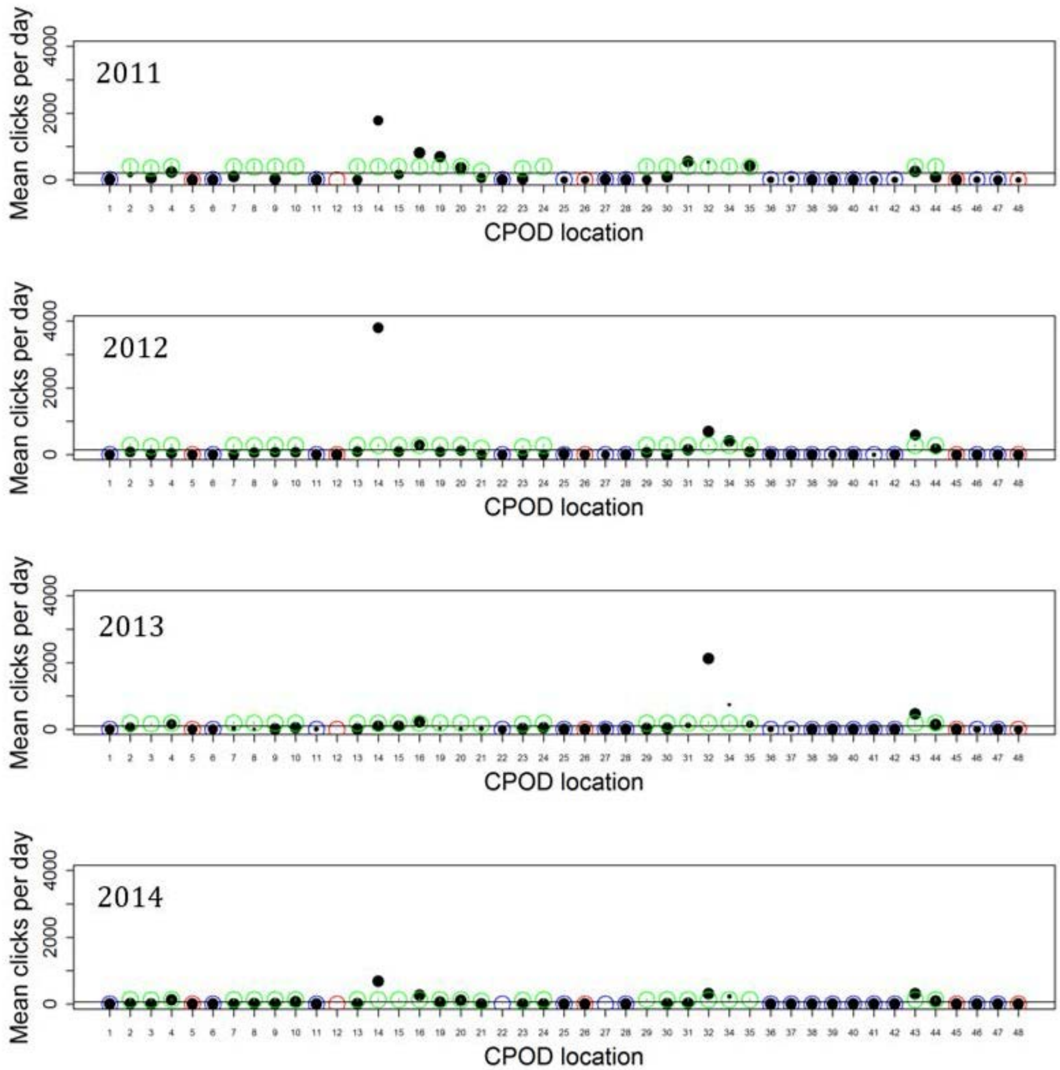


Figure 13. B) Same as in A but the y-axis goes to 4000 to show data extremes.

The posterior mean for $\bar{\lambda}$ was 0.72 with a 95% credible interval ranging from 0.57 to 0.90 (Figure 14). The probability that $\bar{\lambda}$ is less than 1 was 0.9998. Annual results are presented in the table below.

	Posterior Mean	2.50%	97.50%
$\bar{\lambda}$ (N.2012/N.2011)	0.99	0.48	1.87
$\bar{\lambda}$ (N.2013/N.2012)	0.71	0.32	1.37
$\bar{\lambda}$ (N.2014/N.2013)	0.63	0.27	1.26
Geo mean $\bar{\lambda}$	0.72	0.57	0.90
Geo mean r ($\bar{\lambda}$- 1)	-0.28	-0.43	-0.10

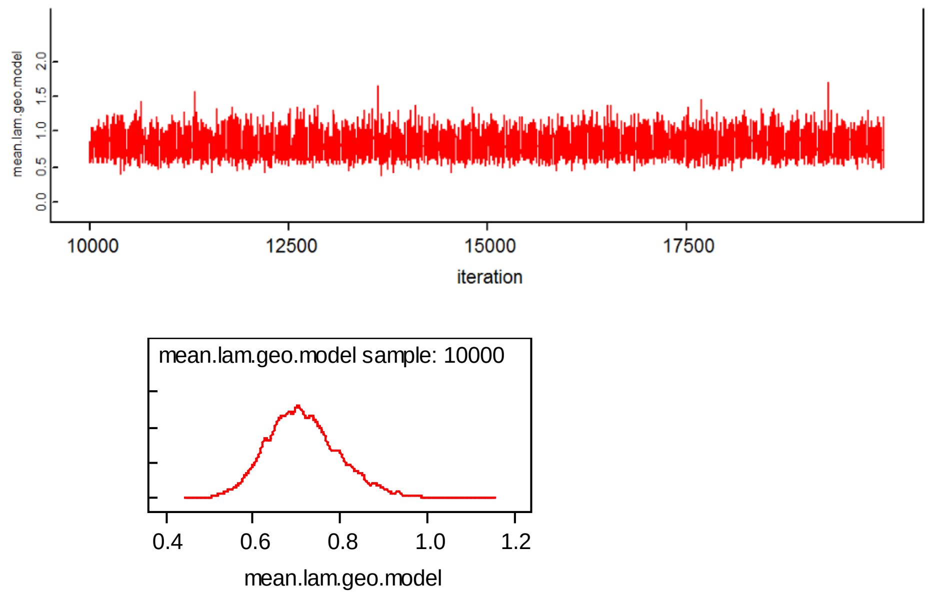


Figure 14. OpenBugs output for $\bar{\lambda}$, the geometric mean of M_4/M_3, M_3/M_2 and M_2/M_1. Top *panel shows the retained MCMC samples.*

Generalized Additive Models

Exploratory GAM Analyses: 2011-2013 DataGeneralized Additive Models (GAMs) were developed to quickly evaluate and compare alternative models for estimating population change before implementing those models in Bayesian models. In the GAMs, year was treated as a categorical explanatory variable (2011, 2012 and 2013) and spatial variation was modeled as a two-dimensional thin-plate spline using the *mgcv* package in R (v. 3.0.1). It was assumed that the spatial distribution of vaquitas was the same across years and that, between years, relative densities changed proportionately among all sites. GAMs that estimated different spatial patterns for each year were generally not stable and are not reported here. The spatial distribution was modeled using all years, but inference on the rate of change in population size was based on the ratio of mean of predicted values in 2013 to the mean predicted values in 2011. To maintain a balanced geographic coverage for this comparison, spatial predictions were made using *predict.gam* on the grid of 45 C-POD stations for which data were available in at least one year. Unless noted otherwise, the GAM analyses were based on the core sampling period (between Julian days 170 and 231, inclusive) when at least 50% of C-POD stations were deployed in each year.

Three common statistical distributions (Poisson, negative binomial and Tweedie distributions) were fit to each dependent variable used, and the best fit was evaluated by visual appraisal of the QQ plots. The negative binomial provided the best fit to all the dependent variables explored here. Within *mgcv*, the binomial parameter *theta* was specified as a range and that range was adjusted as necessary to ensure that best-fit value was not outside the range of potential values. When a mean of daily values was used as a dependent variable, the number of days was used as an offset to account for the unequal sample size.

Model Results

When total clicks per day were used as a dependent variable, none of the statistical models provided a good fit, but the negative binomial (Fig. 15) fit better than the Poisson or Tweedie distributions. When mean clicks per day (averaged over all days for a given site and year) was used as a dependent variable, a negative binomial distribution provided a very good fit to the data (Fig. 16).

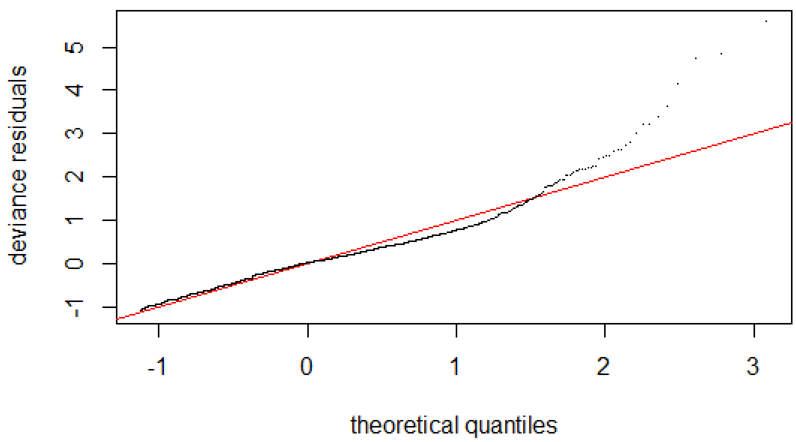


Figure 15. Quantile-quantile plot showing how well the best statistical distribution (negative binomial) fit the distribution of total clicks per day for 2011-2013. Ideally, all the points would fall on the line if the theoretical distribution fit the distribution of the data perfectly.

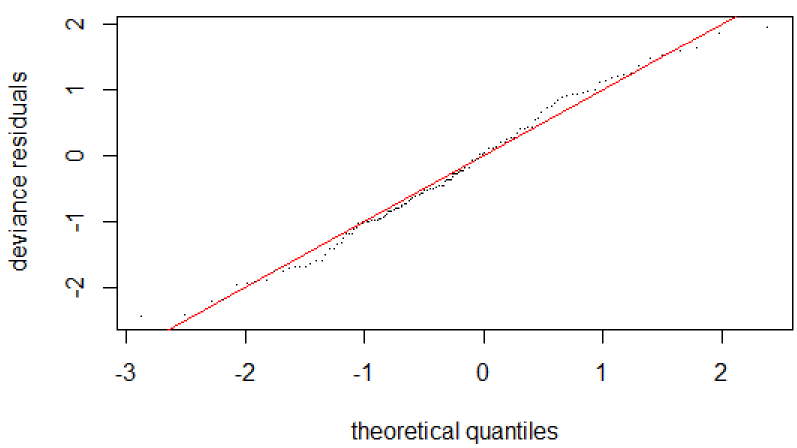


Figure 16. Quantile-quantile plot showing how well the best statistical distribution (negative binomial) fit the distribution of mean clicks per day (averaged over all days for each station and year). Note that the negative binomial distribution provided a much better fit for mean clicks per day than for total clicks per day (Figure 15).

Some previous studies of relative porpoise abundance using C-PODs have been based on detection positive minutes, that is the number of minutes per day with at least one porpoise click. When mean detection positive minutes (DPM) per day (averaged over all days for a given site and year) was used as a dependent variable, a negative binomial distribution provided a reasonable fit to the data (Figure 17).

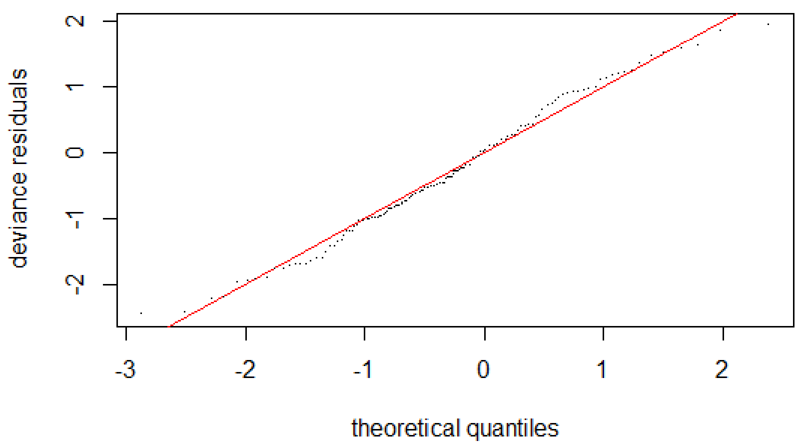


Figure 17. Quantile-quantile plot showing how well the best statistical distribution (negative binomial) fit the distribution of detection positive minutes (averaged over all days for each station and year).

We also explored the potential of using detection positive half-hour periods as a measure of relative vaquita density, that is the number of half-hour periods per day with at least one porpoise click. Preliminary analyses during the workshop showed that the vast majority of porpoise detections lasted less than half an hour, so half-hour periods should be relatively independent of each other. When mean detection positive half-hours (DPHH) per day (averaged over all days for a given site and year) was used as a dependent variable, a negative binomial distribution provided a marginally good fit to the data (Figure 18). A better fit was obtained using total DPHH per day instead of using the mean DPHH. The negative binomial distribution fit total DPHH (Figure 19) much better than total clicks (Figure 15). That total DPHH model result is given below.

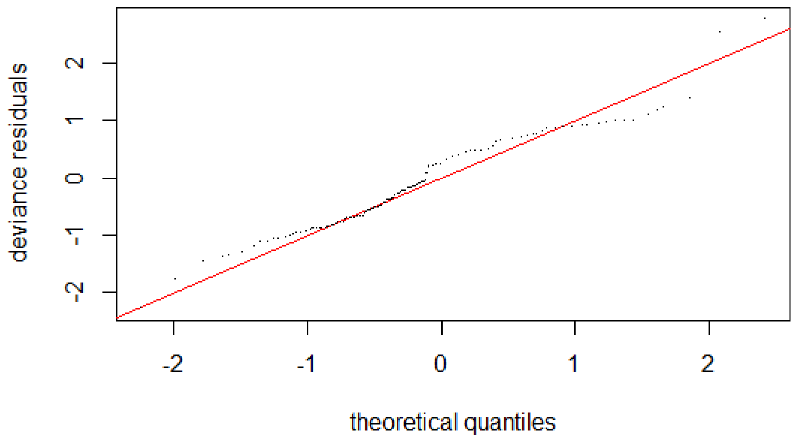


Figure 18. Quantile-quantile plot showing how well the best statistical distribution (negative binomial) fit the distribution of detection positive half-hours (averaged over all days for each station and year). Note that the negative binomial distribution did not fit detection positive half-hours as well as it fit detection positive minutes (Fig. 14).

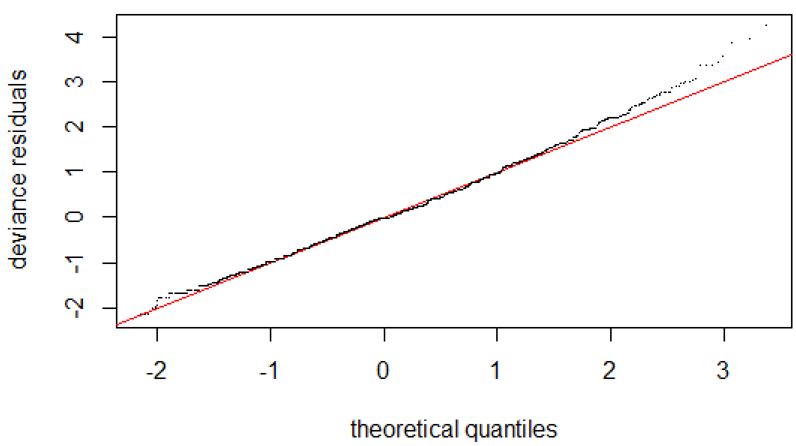


Figure 19. Quantile-quantile plot showing how well the best statistical distribution (negative binomial) fit the daily total of detection positive half-hours.

Appendix 3: **Further data description**

There were two efforts that were useful to Panelists in interpreting the data and considering how to choose an appropriate model. Station numbers are given in Figure 20. The first helpful effort was a representation of clicks through time for each station and for each year (Figure 21). The second analysis showed the relation between the mean and variance for mean clicks/CPOD day (Figure 22).

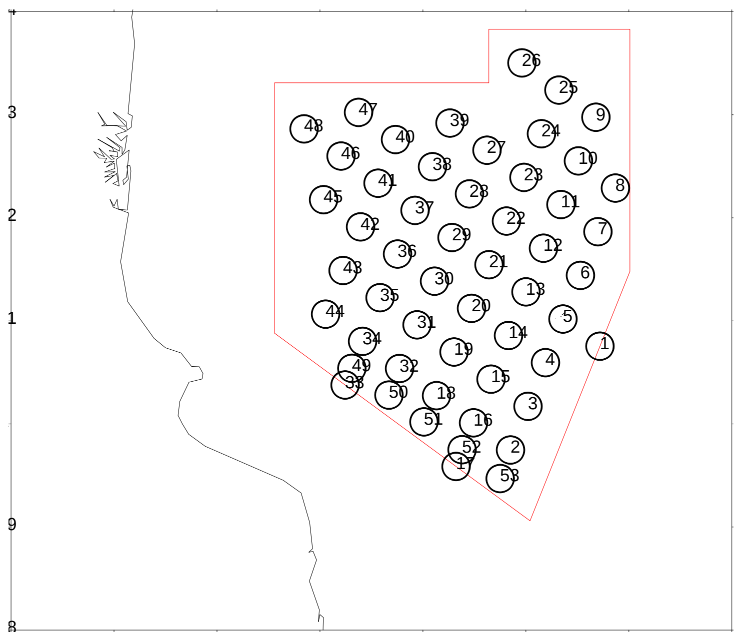
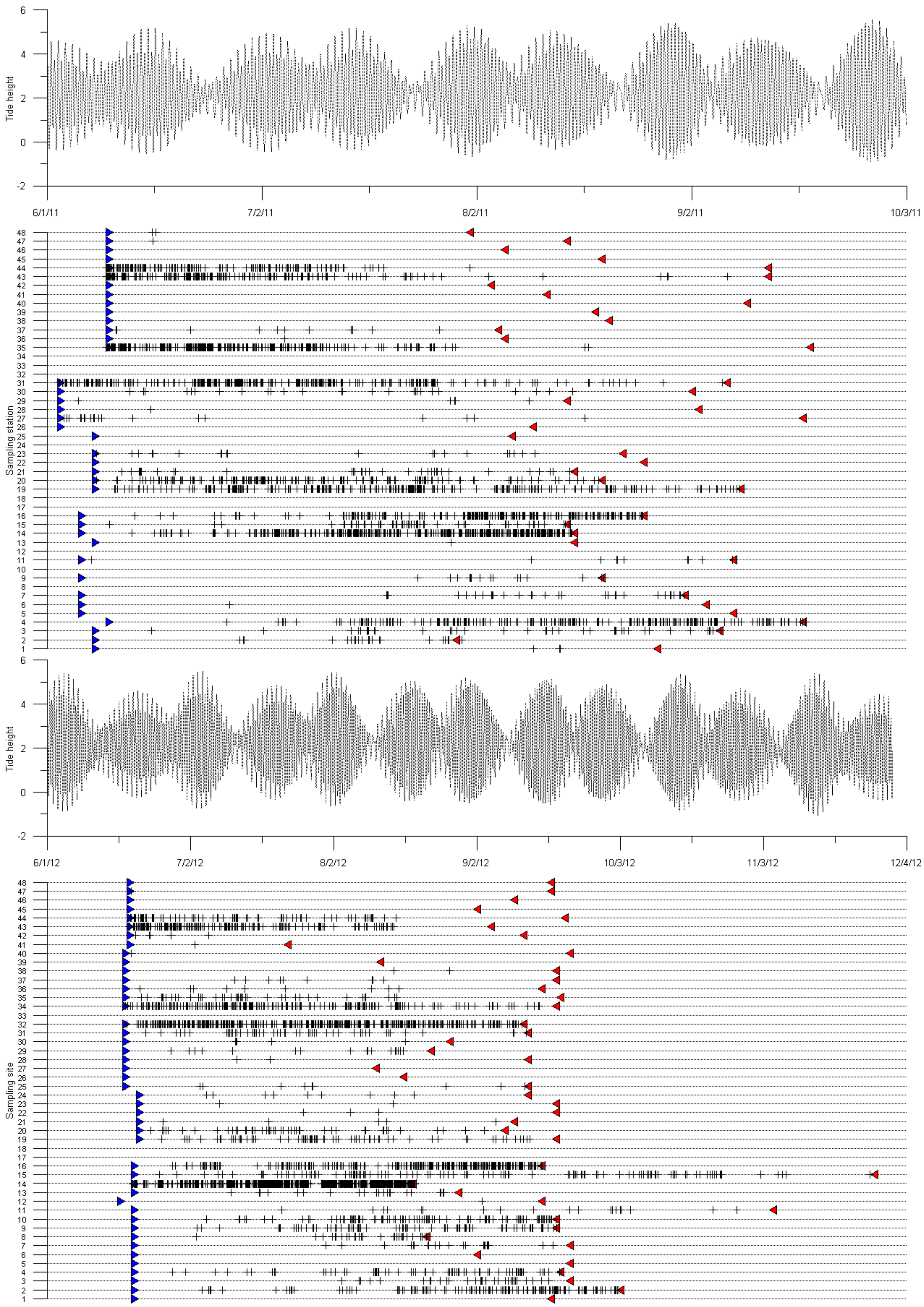


Figure 20. Position of the sampling sites inside the Vaquita Refuge (upper map, numbered circles). Below are the results of moorings and acoustic detectors deployed in 2011, 2012, and 2013. C-PODs were not deployed at sites 17, 18, and 33 in 2013 (X's). The CPOD at site 32 in 2011 was recovered June 25, 2013 and data were included in this analysis. Circles indicate sites where data are available, diamonds indicate all equipment lost at that site, and squares indicate sites where the mooring was recovered without the detector or the detector was recovered without any data.


Tide height
6/1/11
7/2/11
8/2/11
9/2/11
10/3/11
Sampling station
Tide height
6/1/12
7/2/12
8/2/12
9/2/12
10/3/12
11/3/12
12/4/12
Sampling site

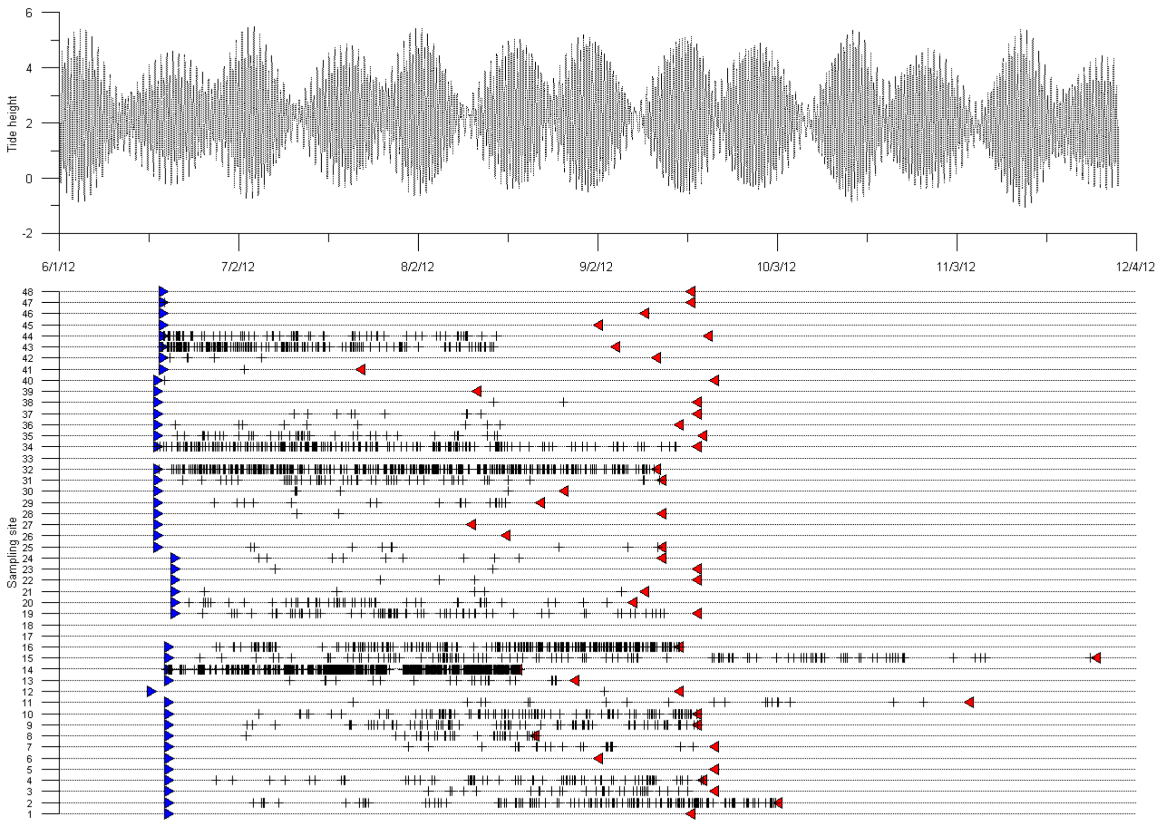


Figure 21. Detection Positive Minutes (DPM´s represented by crosses) (2011-2013) for every available sampling station. Tide heights for San Felipe (closest town to vaquita distribution area) are shown in the top panel for June – September, except for 2012 where period is extended because data were available through November for sites 11 and 15 (detectors recovered on 2013). In the lower panel blue triangles indicate the first sampling day and red triangles the last sampling day. C-PODs were turned throughout this period.

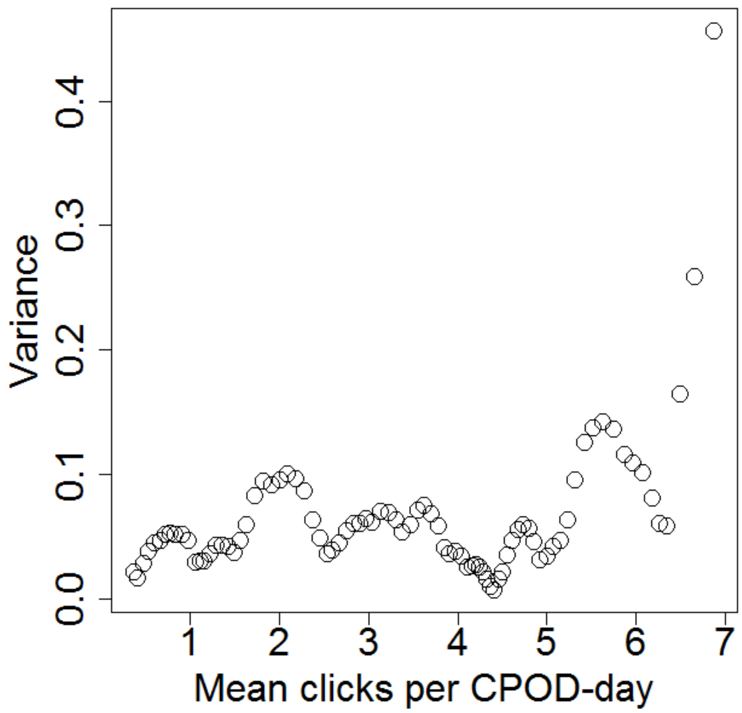


Figure 22. Variance and mean of log-transformed data, i.e., var[log(x_k+1)] and mean[log(x_k+1)], where x_k is the mean number of clicks per day for an individual CPOD location in a particular year (128 unique values) using data from the core sampling period. Each point represents the mean and variance of 10 ordered values (e.g., left-most point is mean and variance of the 10 lowest x_k values; next point is mean and variance of 2nd lowest to 11th lowest x_k,, etc.). Moving window approach results in serial autocorrelation in the variance values, but overall the variance is relatively constant with respect to the mean on the log scale (apart from a few outliers), justifying use of the Gaussian spatial model (constant variance assumption) of the log-transformed data.