



Noise levels received by endangered killer whales *Orcinus orca* before and after implementation of vessel regulations

Marla M. Holt^{1,*}, M. Bradley Hanson¹, Deborah A. Giles², Candice K. Emmons¹, Jeffery T. Hogan³

¹Conservation Biology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, 2725 Montlake Boulevard East, Seattle, WA 98112, USA

²Wildlife, Fish, & Conservation Biology, 1088 Academic Surge, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

³Cascadia Research Collective, 218½ 4th Avenue West, Olympia, WA 98501, USA

ABSTRACT: Whale watching is often conducted from motorized vessels, which contribute to underwater noise pollution and can disturb marine mammals. Protective measures can ameliorate some effects of disturbance, but it is crucial to empirically assess the effectiveness of such measures, particularly for endangered species. We quantitatively compared noise exposure to endangered southern resident killer whales before and after US federal vessel regulations were established to protect this population from disturbance by vessels and sound. We expected to see a reduction in noise exposure to this population from vessel sound propagation loss due to a doubling of the minimum viewing distance relative to a prior state law. Noise levels were empirically measured from digital acoustic recording tags (DTAGs) suction-cup attached to killer whales in transboundary critical habitat. We collected concurrent vessel data during DTAG deployments to relate to received noise levels at the animal. Results of a linear mixed model analysis that included 10 explanatory variables in candidate models revealed that noise was best predicted by animal ID, vessel count, vessel speed category, and year. Vessel count and speed category were positive predictors of noise levels. Vessel regulations (before vs. after implementation), country, and average vessel distance were not significant predictors of noise levels, although only 1 yr of baseline data limited assessment. These findings inform managers about the effectiveness of current regulations for viewing killer whales and are applicable to other cetacean species that are exposed to vessel noise from whale-watching activities.

KEY WORDS: Killer whale · *Orcinus orca* · Noise · Vessel regulations · Whale watching · Wildlife viewing · Archival tag

INTRODUCTION

Wildlife viewing is often a component of nature-based tourism activities, offering educational opportunities and generating a substantial amount of revenue (Leonard 2006). These benefits can provide incentives to promote wildlife populations in areas that might be otherwise prone to unsustainable harvesting or habitat destruction. However, wildlife

viewing can also contribute to disturbance of wildlife and/or important habitats when conducted irresponsibly (Geffroy et al. 2015, Senigaglia et al. 2016). In particular, there is considerable concern about anthropogenic disturbance of already vulnerable species that are struggling with population status and recovery.

Marine wildlife viewing often occurs from motorized vessels, which introduce incidental and sometimes in-

*Corresponding author: marla.holt@noaa.gov

tentional (e.g. navigational sonar) sound in the ocean. Marine wildlife viewing also often focuses on easily observable species that are coastally distributed and exposed to other sources of anthropogenic sound including other vessel traffic. Many animals use sound for key life processes in the ocean. Most notably, marine mammals use sound for activities involved in feeding, breeding, and navigation, and documented noise effects are numerous and reviewed by several sources (Richardson et al. 1995, NRC 2003, 2005). Protective measures can ameliorate negative impacts of anthropogenic disturbance. Marine mammal protective measures are usually established through a few key legislative authorities in the USA (e.g. the Marine Mammal Protection Act [MMPA 1972] and the Endangered Species Act [ESA 1973]). Follow-up investigations that provide empirical data to determine protective measure effectiveness, especially those relevant to vessel disturbance, are relatively rare but paramount for protected species conservation.

The southern resident killer whale population is currently listed as endangered in both the USA and Canada given a significant population decline in the mid- to late 1990s (Krahn et al. 2002). Their range extends from southeastern Alaska to central California coastal waters, but this fish-eating population is best known in the Salish Sea (inland waters of Washington State and British Columbia), where they feed extensively in the summer and fall (Hanson et al. 2010). Several threats challenge population recovery including prey scarcity, chemical contaminants, and disturbance from vessels and noise (NMFS 2008). This population is avidly observed by local whale-watching vessels, especially during summer daylight hours. In addition, commercial shipping, fishing, passenger, and recreational vessels frequently transit their summer/fall feeding grounds (Bassett et al. 2012, Veirs et al. 2016). Vessel counts collected within 0.5 mile (0.8 km) of these killer whales average around 15 to 20 vessels but can reach up to 80 vessels or more during peak summer periods (Seely 2015). Vessels and associated noise, which here refers to the broadband radiated sound of motorized vessels that contributes to underwater environmental noise levels, can affect behavioral and acoustic activities of killer whales, including their use of sound for echolocation, the predominant sensory modality for foraging, and communication (Barrett-Lennard et al. 1996, Williams et al. 2006, 2009, Holt et al. 2009, Lusseau et al. 2009, Noren et al. 2009).

Specifically, several studies have documented vessel and noise effects on southern resident and other killer whale populations of the northeast Pacific,

including the effects of increased noise as vessel numbers in close proximity increased (Holt et al. 2009). Behavioral responses to vessel sound and/or presence include changes in surface active behaviors, vocal behavior, and diving/movement parameters and a reduction in foraging activity (Williams et al. 2006, 2009, Holt et al. 2009, Lusseau et al. 2009, Noren et al. 2009). Given these documented effects and recognized risk factors, federal vessel regulations were established to protect the southern resident population from interference and noise. The regulations, effective in May 2011, make it unlawful for vessels to approach within 400 yards (yd) (366 m) of any killer whale's path or within 200 yd (183 m) from other directions in US waters corresponding to critical habitat (NOAA 2011). The regulation applies to all killer whales because most vessel operators cannot distinguish the ESA-listed southern resident types from other types of killer whales (NOAA 2011). Prior to federal regulations, a 2008 Washington state law established a minimum vessel approach distance of 100 yd (91 m) while viewing killer whales. Additionally, Be Whale Wise guidelines encourage all boaters to reduce speed to 7 knots within 400 yd (366 m), in addition to other recommendations which have changed periodically since 2002 (www.bewhalewise.org). Currently, only voluntary guidelines exist in Canadian waters.

A recent related study using digital acoustic recording tags (DTAGs) found that of 8 variables attributed to nearby vessels, the number of propellers and especially operating speed best predicted noise levels received by southern residents (Houghton et al. 2015). In the current investigation, we compared received noise levels before and after vessel regulations came into effect, also using DTAG data but involving additional field effort and a separate analysis from that of Houghton et al. (2015). We expected to find a reduction in received noise levels after regulations because of sound propagation loss. Specifically, if all vessels changed their minimum approach distance from 100 to 200 yd (91 to 183 m), on average, then the expected reduction would be in the range of 3 to 6 dB assuming cylindrical or spherical spreading loss (10 to $20 \times \log$ distance). The approach presented here allowed us to empirically and quantitatively assess any noise protective benefits to this population because of vessel regulations. The results can inform wildlife managers about the effectiveness of current regulations for viewing killer whales and are broadly applicable to a variety of other protected marine species, including other endangered cetacean species, on which vessel-based wildlife-viewing activities are focused.

MATERIALS AND METHODS

Data collection

Data were collected from a 6.7 m rigid-hull inflatable research vessel during daylight hours in September 2010, June 2011, September 2012, and September 2014 in transboundary waters surrounding the San Juan Islands, WA, USA (approximate range: 48.2° to 49.0° N, 122.7° to 123.6° W, Fig. 1). All field efforts were conducted under research permits (in the USA, NMFS No. 781–1824/16163 and in Canada, DFO SARA/Marine Mammal License No. MML 2010-01/SARA-106B) and approved by Northwest Fisheries Science Center’s Institutional Animal Care and Use Committee.

The DTAG is a suction-cup-attached multi-sensor archival tag. DTAGs used in this study included 2 hydrophones for audio recordings and temperature, pressure, accelerometer, and magnetometer sensors used to derive whale depth, pitch, roll, and heading that were synchronous with audio recordings (Johnson & Tyack 2003). In 2010, 2011, and 2014, larger version 2 tags (hereafter DTAGv2) were deployed, in which audio data were sampled at 192 kHz (16 bit)

with a 400 Hz 1-pole high-pass filter. In 2012, smaller version 3 tags (hereafter DTAGv3) were used, in which audio data were sampled at 240 kHz (16 bit), also with a 400 Hz 1-pole high-pass filter.

DTAGs were deployed using a pole from the bow pulpit of the research vessel. As conditions permitted, we collected geo-referenced data on the tagged whale and all vessels (including the research vessel) within 1.5 km (accuracy limit of the range finder) consistently across observations. Two integrated equipment packages designed to collect vessel data around killer whales (Giles 2014) were each operated by 2 experienced field observers consistently throughout the study. One collected all data at each focal whale surfacing, and the other collected all vessel data simultaneously. Each package consisted of a GPS/data collector (Trimble Geo HX or Geo XM), a laser range finder, and a compass that communicated as 1 unit to generate the geo-referenced (latitude/longitude) data.

The data collection protocol during focal follows is described in detail in Houghton et al. (2015). For the current investigation, whale data collection included a date/time stamp, latitude/longitude to determine country (either USA or Canada), and individual ID

from photo-ID records to determine age class (either adult or juvenile) and sex (either male or female) variables. Attribute data for each vessel included date/time stamp, latitude/longitude, vessel name, and speed estimated visually as follows: stationary, slow (1–2 knots), medium 3–4 knots, fast (5–6 knots), very fast (7+ knots). Speed categorization was the most reliable method to quantify operational speed of smaller vessels since they are not required to transmit automatic identification system (AIS) signals, repeated fast GPS fixes for individual vessels were logistically impossible, and radar guns can be inaccurate on the water (Erbe 2002). Additionally, it was not always possible to determine whether stationary vessels were in shutdown mode (engines turned off) or in idle. Despite this measurement limitation, vessel speed estimate was an important variable to include in candidate models to explain noise levels given results of previous investigations (McKenna et al. 2013, Houghton et al. 2015, Veirs

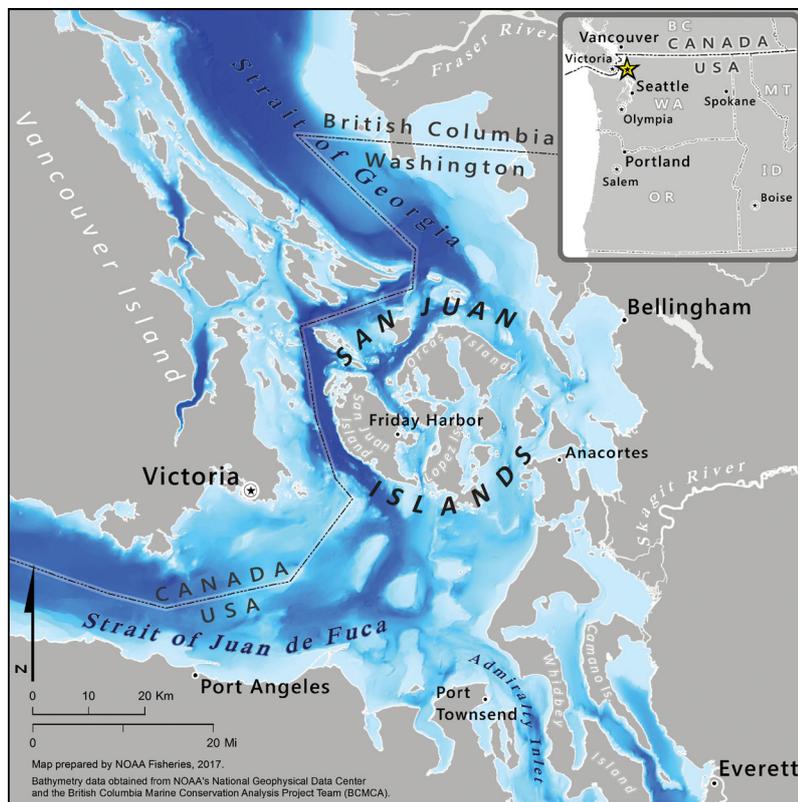


Fig. 1. Study area. Dashed line indicates the international border between US and Canadian waters

et al. 2016). Vessel data were collected in concentric rings starting with the closest distance to the focal whale at least every 5 min. Logistically, data could only be collected in conditions corresponding to Beaufort scales of 0 to 3. If vessel counts were too high to populate attribute data within this period, this was noted. It was often difficult to determine which vessels were actively whale watching relative to the focal whale, especially when vessel counts were high and whales were spread out. Regardless, all motorized vessels underway radiate underwater noise and are subject to US vessel regulations. During focal follows, the research vessel generally paralleled the whale at distance and speed consistent with other vessels, but the vessel operator frequently changed speed and distance relative to the focal whale to collect data. Once a tag fell off, it was retrieved, and raw data were offloaded using custom software (Johnson & Tyack 2003). Raw DTAG data were post-processed using the DTAG toolbox and custom-written routines in Matlab (version 7.10 or higher, Johnson & Tyack 2003). All non-audio data were calibrated and down-sampled to 5 Hz (Johnson & Tyack 2003).

Noise level and vessel data analysis

Flow noise often occurred in the acoustic record, particularly as swim speed increased (Goldbogen et al. 2006). Considerable effort was taken to ensure that noise levels were measured during periods unaffected by flow noise as well as other extraneous sounds (e.g. splashing from surfacings, killer whale sounds). First, an analyst determined periods absent of extraneous sounds and with minimal flow noise, which were typically easily identified via spectrographic analysis during manual screening of all audio data (Houghton et al. 2015). However, teasing out flow noise was occasionally difficult, particularly during high vessel noise periods given that both sources of noise overlap in (low) frequency. Therefore, periods unaffected by flow noise were objectively identified using a correlation-based method that separates the relative contributions of uncorrelated flow and correlated ambient noise between the 2 hydrophones in the tag (von Benda-Beckmann et al. 2016). The method is limited to mid-range frequencies because pressure fluctuations generated by flow are uncorrelated for frequencies greater than the speed through water divided by the hydrophone spacing, while ambient noise is correlated for frequencies less than one-tenth times the speed of sound divided by the hydrophone spacing (von Benda-Beckmann et al. 2016). Given the high

pass filter and hydrophone separation within the tag (2.5 and 4.5 cm for DTAGv2 and DTAGv3, respectively)¹, we assessed time periods unaffected by flow noise in which the RMS sound pressure level (SPL) of correlated noise exceeded the SPL of uncorrelated noise by 6 dB in each one-third octave band for the appropriate frequencies (0.5 to 6 kHz for DTAGv2, 0.5 to 3 kHz for DTAGv3, von Benda-Beckmann et al. 2016). During this assessment, we noted that many more time periods were affected by flow noise below 1 kHz compared to above 1 kHz given that flow noise is low-frequency-dominated noise. We assumed that flow noise would not affect noise levels above the upper frequency limit of which ambient noise remains correlated (3 or 6 kHz depending on the tag version) if flow noise was not an issue below these frequencies. Spectrographic assessment showed that flow noise spreads from lower to higher frequencies as swim speed increases and substantiated this assumption.

Noise levels were calculated using the receiving sensitivity of -177 dB re 1 V μPa^{-1} for DTAGv2 and -178 dB re 1 V μPa^{-1} for DTAGv3 across the bandwidth analysis, derived from reliable calibration tests conducted at tank facilities for this and other projects (with housing for both DTAGv3s used in 2012 and 2 of 6 DTAGv2s used in 2010, 2011, and 2014, Tyack et al. 2011, DeRuiter et al. 2013, Wensveen 2016). Received noise levels (in dB re 1 μPa) were calculated from RMS sound pressure averaged across both audio channels over a 1 s duration between 1 and 40 kHz. The time window was chosen to capture the dynamic nature of noise emitted by whale-watching vessels. Depth data were also averaged over the same 1 s period for each noise measurement. The frequency band was chosen based on the trade-offs between flow noise contamination (to maximize sample size) and the overlap of vessel noise and the hearing sensitivity in killer whales. Killer whales can hear below 1 kHz, but their hearing sensitivity is reduced by 50 dB or more below this frequency, relative to the lowest measured hearing threshold at approximately 20 to 40 kHz (Szymanski et al. 1999, Branstetter et al. 2017). Given that vessel noise is generally low-frequency noise, we also measured noise levels in the 0.5 to 40 kHz band during matched periods unaffected by

¹Typical swim speeds of killer whales average 1.6 m s^{-1} , with a maximum of 3.0 m s^{-1} (Williams & Noren 2009), well below the speed limit that flow noise remains uncorrelated for frequencies down to 0.5 kHz. Swim speeds would have to exceed 12.5 m s^{-1} for DTAGv2 (and 22.5 m s^{-1} for DTAGv3) to violate this assumption, and resulting flow noise would be obvious in the acoustic record during manual screening.

flow noise for the limited data available to determine the dB difference in noise levels when frequencies between 0.5 and 1 kHz were included.

Collecting vessel attribute data sometimes took up to 5 min (or more) during high vessel counts. Thus, vessel data were included in the analysis if collected within ± 5 min of the noise measurement. If multiple observations of the same vessel occurred within the interval, only the observation closest in time to the noise measurement was used. For each vessel within 1.5 km (limit of the range finder), horizontal distance from whale to vessel was calculated using latitude/longitude positions closest in time, and then true distance was calculated using whale depth at the noise measurement time. For vessel speed, the median was used as the central tendency of vessel speed, as this was coded as an ordinal variable for each vessel observation as follows: 0 = stationary; 1 = slow; 2 = medium; 3 = fast; 4 = very fast (speeds as given earlier in this Section) (similar to Houghton et al. 2015). Using these ordinal values, the median vessel speed was then calculated; rounded to the nearest integer, x ; and labeled as stationary ($x = 0$), slow ($x = 1$), medium ($x = 2$), fast ($x = 3$), or very fast ($x = 4$). Vessel variables were then populated for each noise level measurement as follows: vessel count (number of unique vessels), category of median vessel speed as a factor (hereafter vessel speed category), and mean vessel distance from the whale (linear average or log average, see next subsection). For this analysis, vessel speed category and vessel distance were aggregated across all vessels for each noise level measurement given that the focus of this investigation was to compare noise levels before and after vessel regulations came into effect. Vessel covariates were incorporated in the analysis in the simplest forms to ease in the interpretation of model results and inform the effectiveness of vessel regulations with respect to noise. A more thorough list of tested vessel variables involving a different analytical approach to relate to received noise levels is presented by Houghton et al. (2015). Any noise measurements that had missing vessel data were excluded.

Statistical analysis

Noise measurements separated by at least 30 s were initially considered independent for analysis upon inspecting data time series for each deployment; however, the vast majority of data points were more than 30 s apart given that large sections of the data had to be excluded due to flow noise contamina-

tion. Potential issues of temporal autocorrelation in the final model were tested and are described in the next paragraph. Using a linear (Gaussian) mixed model, several explanatory variables were then considered as potential predictors of noise level (response variable) using Akaike's information criterion (AIC) for model selection. Analysis was conducted using the lme4, lmerTest, and effects packages in R (Kuznetzova et al. 2012, Bates et al. 2014, Fox et al. 2015, R Core Team 2017). Explanatory variables included in candidate models were year (or before/after implementation of regulations); country (as regulations only apply in the USA); and vessel-, animal-, and tag-related variables. Explanatory variables were chosen based on our hypothesis with respect to vessel regulations (i.e. reduced noise after regulations), previous investigations of vessel noise (Holt et al. 2009, McKenna et al. 2013, Houghton et al. 2015, Veirs et al. 2016), potential effects related to animal- (animal ID, age class, sex) and tag-related variables (individual tag ID and deployment minutes) that might affect noise levels, and data exploration and availability (Table 1). We considered the interaction between year and country on noise levels as vessel regulations came into effect in 2011 but were only applicable in US waters. We did not consider other interaction terms among variables, given the lack of *a priori* reasons to assume multiplicative properties among the remaining variables, which can easily complicate the model structure and interpretation of results (Zuur et al. 2009). Additionally, other statistical comparisons (*t*-test, Chi-squared test) were conducted to compare vessel-related variables before and after vessel regulations came into effect to aid in interpretations of results.

Table 1. Explanatory variables considered during model selection. Number of levels for each factor are given in parentheses

Effect	Form (level)	Type
Year	Numeric, factor (4), before/after regulations (2)	Fixed
Country	Factor (2)	Fixed
Vessel count	Linear, log ₁₀	Fixed
Average vessel distance, closest vessel distance	Linear, log ₁₀	Fixed
Vessel speed category	Factor (4)	Fixed
Sex	Factor (2)	Fixed
Age class	Factor (2)	Fixed
Deployment min	Linear, log ₁₀	Fixed
Animal ID	Factor (24)	Random
Tag ID	Factor (8)	Random

Optimal data type (e.g. factor vs. numeric) or quantification (e.g. linear vs. log) for explanatory variables was first determined through data plotting, AIC selection, and interpretation of results (Table 1). Model selection occurred according to Zuur et al. (2009, p. 121–122) as follows: (1) optimal random effects structure was established starting with all fixed explanatory variables and interaction terms and adding random effects (with random intercept structure) if lower AIC values were obtained; (2) the top-ranked mixed model was identified starting with the full model and dropping explanatory variables until the lowest AIC value was found; and (3) significance of an explanatory variable was confirmed using the likelihood ratio test. We also explored potential issues of collinearity among covariates. Vessel count and average vessel distance were moderately correlated with one another ($r > 0.3$), but variance inflation factor values for these predictors as well as all others were < 2 . Thus, all were included in candidate models (Zuur et al. 2010, see the Supplement at www.int-res.com/articles/suppl/n034p015_supp.pdf for alternative approach). Model assumptions were checked visually by plotting residuals versus fitted values and with each of the fixed-effect variables (Zuur et al. 2009). Temporal autocorrelation was tested using the final model residuals within a deployment time series and the function *acf* in R. Any data points indicating significant time lags and, therefore, autocorrelation were removed from the data set, and then model selection was rerun according to the model selection protocol using the revised dataset. The *acf* test was then rerun on the revised dataset, again using the residuals of this final model, and no significant time lags were found. Given that our research vessel, which operated under a research permit, was always present and several noise level data points were associated with only this vessel (i.e. vessel count = 1), we also limited the analysis to data when vessel counts were > 1 and repeated the above model selection process.

RESULTS

Tag deployments and vessels

Data from 27 DTAG deployments on 24 individual animals were included in the analysis (Table 2). Note that 1 deployment was from a transient killer whale individual and included in the analysis, as vessel regulations apply to all killer whales (NOAA 2011). Noise levels in the 1 to 40 kHz band ranged from

Table 2. Summary of tag deployments by year, animal ID, age class (A: adult; J: juvenile), sex (M: male; F: female), tag ID (v2 denotes version 2 digital acoustic recording tag [DTAG], v3 denotes version 3 DTAG), and total tag deployment minutes for data included in the analysis of noise levels. Each row represents a separate tag deployment, listed in chronological order

Year	Animal ID	Age class	Sex	Tag ID	Total tag (min)
2010	L88	A	M	v2_244	270
2010	L72	A	F	v2_244	43
2010	L83	A	F	v2_236	163
2010	K33	J	M	v2_244	376
2010	J14	A	F	v2_236	240
2010	L86	A	F	v2_244	449
2010	L78	A	M	v2_236	68
2011	J30	J	M	v2_234	337
2011	J36	J	F	v2_244	107
2011	J35	A	F	v2_234	286
2011	J35	A	F	v2_232	206
2011	J8	A	F	v2_232	271
2012	L22	A	F	v3_105	417
2012	K33	J	M	v3_102	100
2012	L95	A	M	v3_105	422
2012	L109	J	M	v3_102	308
2012	L116	J	M	v3_105	166
2012	L84	A	M	v3_102	133
2012	L91	A	F	v3_105	159
2012	L47	A	F	v3_102	47
2012	J28	A	F	v3_102	156
2014	L113	J	F	v2_245	430
2014	L89	A	M	v2_242	533
2014	T65A2	J	M	v2_245	451
2014	L85	A	M	v2_242	383
2014	L91	A	F	v2_245	50
2014	K35	J	M	v2_245	286

approximately 96 to 127 dB re 1 μ Pa (mean = 108 dB \pm 5.4 SD, N = 384). Noise levels in the 0.5 to 40 kHz band were 1.4 dB higher, on average (difference range: 0.1 to 6.3 dB, N = 305), than those measured in the 1 to 40 kHz band for the same 1 s time period. Although noise levels hereafter only include frequencies ≥ 1 kHz (to maximize sample size), these are generally within a few decibels of broadband noise levels that include frequencies down to 0.5 kHz. Of the 384 noise level measurements (1 to 40 kHz) included in the analysis, 312 were measured in US waters, while 72 were measured in Canadian waters (by year, 2010: 73 in US, 18 in Canadian waters; 2011: 49 in US, 23 in Canadian waters; 2012: 124 in US, 15 in Canadian waters; 2014: 66 in US, 16 in Canadian waters). The country in which noise levels were measured was dictated by the swimming behavior of the whale and not necessarily where the tag was deployed (see Fig. 1 for the entire study area). Vessel counts ranged from 1 to 14 with a median of 3, and

average vessel distance ranged from 21 to 914 m with a median of 314 m. Individual vessel speeds ranged from stationary to very fast, but only 1 noise level observation had a vessel speed category, based on the median of all vessels present, of very fast. Thus, the data point was removed from the analysis given the limited sample size. Vessel types included commercial whale watching, private/pleasure, research, enforcement/monitoring, ferry, shipping, and other vessels, some of which appeared to be observing killer whales, while others appeared to be transiting the area.

Model selection results

The full model included year, country, the interaction between year and country, sex, age, and vessel speed category as factors, and vessel count, average vessel distance, and deployment minute in log form. Year as a factor was incorporated in the full model as opposed to a numeric data type or, alternatively, as before and after regulations, as it provided a better fit to the data (relative to the AIC of the model with year as a factor, year as a numeric: $\Delta AIC_i = 44.0$, before/after regulations: $\Delta AIC_i = 43.8$, where AIC_i is the difference in AIC of the i th model). When the before/after regulations variable was included in the full model, this variable was not significant, indicating that noise levels were not different after regulations were implemented. As expected, log vessel count also resulted in a better fit compared to when vessel count was kept in linear terms ($\Delta AIC_i = 7.7$). Log distance is consistent with simple approximations of spherical or cylindrical spreading, which is corroborated by empirical assessments of sound propagation at comparable vessel ranges conducted in the same study location as the current one (Holt et al. 2009, Veirs et al. 2016), but, unexpectedly, the distance parameter took on an apparent positive relationship

with noise when log transformed (i.e. noise was predicted to increase with increasing vessel distance). When distance was kept in linear terms, the parameter was slightly negative but close to zero, with a lower model AIC value ($\Delta AIC_i = 2.4$). This might be due to more complicated propagation effects in some locations and/or because distance was averaged across all vessels within a noise level observation, but was most likely because of its correlation with vessel counts (as count increased, the average distance increased due to spatial restrictions around whales). Thus, 2 competing models with distance in linear or log terms were used for model selection. Distance of the closest vessel (in linear and log form) was tested as an additional explanatory variable in candidate models, but it was not an important predictor of noise level. The best random effects structure included only animal ID. Tag ID was not an important predictor of noise level, indicating that each tag operated as expected in cases where calibration results of an individual tag were not available.

The top-ranked model included year, log vessel count, and vessel speed category along with animal ID as explanatory variables (Table 3). Results of the likelihood ratio test revealed that year ($\chi^2_3 = 24.16$, $p < 0.0001$), vessel speed category ($\chi^2_3 = 18.6$, $p < 0.001$), and log vessel count ($\chi^2_1 = 6.54$, $p = 0.011$) were significant explanatory variables. The year 2010 was considered the reference factor level in the model, as this was the baseline year before vessel regulations were implemented. The intercept estimate (see Table 5) represents the predicted noise level value for the year 2010 in the presence of 1 stationary vessel (SpeedCat0.stationary, LogCount = 0). Pairwise comparisons between years showed that 2012 noise levels were not significantly different from those of 2010 (lmerTest using type III sum of squares and Satterthwaite approximations, Fig. 2A). Although DTAGv3s were only used in 2012, noise levels were not different compared to 2010 when

Table 3. Model selection results for linear mixed model analysis. All models included (1|animalID) as a random effect. Note that year \times country is R notation and is equivalent to (year + country + year:country). ΔAIC_i : difference in Akaike's information criterion of i th model

Model	Fixed effects	ΔAIC_i
1	Year \times Country, LogCount, Distance, Speed, Age, Sex, LogDeployMin	13.4
2	Year, Country, LogCount, Distance, Speed, Age, Sex, LogDeployMin	8.8
3	Year, LogCount, Distance, Speed, Age, Sex, LogDeployMin	6.8
4	Year, LogCount, Distance, Speed, Sex, LogDeployMin	4.9
5	Year, LogCount, Speed, Sex, LogDeployMin	3.0
6	Year, LogCount, Speed, Sex	1.4
7	Year, LogCount, Speed	0.0

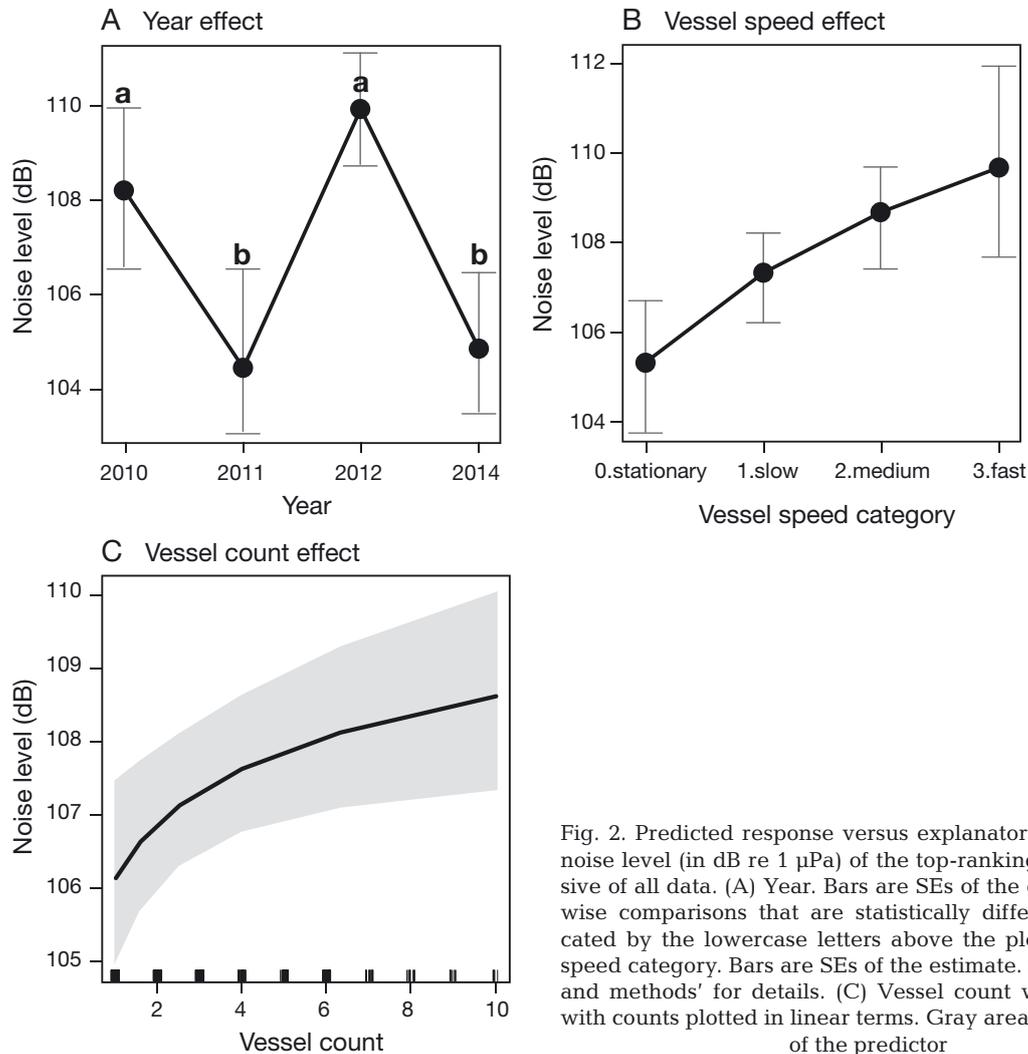


Fig. 2. Predicted response versus explanatory variables on noise level (in dB re 1 μ Pa) of the top-ranking model inclusive of all data. (A) Year. Bars are SEs of the estimate. Pairwise comparisons that are statistically different are indicated by the lowercase letters above the plots. (B) Vessel speed category. Bars are SEs of the estimate. See 'Materials and methods' for details. (C) Vessel count within 1.5 km, with counts plotted in linear terms. Gray area is the 95% CI of the predictor

DTAGv2s were deployed, indicating that tag version did not explain noise level differences among years. Both 2011 and 2014 noise levels were lower than those of 2010 ($p < 0.05$) but were not significantly different from each other (Fig. 2). Notably, after vessel regulations came into effect, there were fewer noise level observations (as a proportion of total observations) when vessel speed category was stationary (after regulations: 12.8% [35/273] vs. before regulations: 25.3% [23/91], $\chi^2_1 = 8.61$, $p < 0.01$). Average vessel distance, either in linear or log form, was not included in the top-ranking model, nor was country or the interaction between year and country included. Likewise, average vessel distance was not significantly different before compared to after regulations (t -test, $t = 1.042$, $p = 0.30$). Variance of animal ID and residuals, parameter estimates of fixed effects, and fixed-effect plots of the top-ranked model are shown in Tables 4 & 5 and Fig. 2, respectively.

The next-best model also included sex, but sex was not a significant variable in this case ($\chi^2_1 = 0.54$, $p = 0.4$, Table 3). Limiting the analysis to vessel counts above 1 resulted in a sample size of 295 and yielded similar results including the same random effect structure, but the top-ranking model in this case included year ($\chi^2_3 = 20.75$, $p < 0.001$), vessel speed category ($\chi^2_3 = 16.35$, $p < 0.001$), log vessel count ($\chi^2_1 = 2.77$, $p = 0.10$), and log deployment minute ($\chi^2_1 = 2.56$, $p = 0.11$, Table 6). However, log vessel count

Table 4. Variance and standard deviation of animal ID, included as a random effect in the top-ranking model (Model 6 in Table 3) to relate to received noise levels for all data

Group	Name	Variance	SD
AnimalID	(Intercept)	4.15	2.04
Residual		19.62	4.43

Table 5. Estimated regression coefficients of fixed effects included in top-ranking model (Model 7 in Table 3) related to received noise levels for all data. The year 2010 was considered the reference factor level in the model, and the intercept estimate represents the predicted noise level value for year 2010 in the presence of 1 stationary vessel (LogCount = 0, SpeedCat0.stationary)

	Estimate	SE
Intercept	105.1	1.30
2011	-3.9	1.52
2012	1.7	1.14
2014	-3.6	1.37
LogCount	2.5	0.99
SpeedCat1.slow	1.9	0.74
SpeedCat2.medium	3.2	0.83
SpeedCat3.fast	4.3	1.26

and log deployment minute were not significant parameters when the data set was limited to vessel counts >1. In this case, the vessel count variable was not significant, likely because of a more limited sample size (N = 295 vs. N = 384 with the full data set) and restricted range of values (vessel counts range 2 to 14) of this variable.

DISCUSSION

We compared noise levels received by killer whales before and after May 2011 to determine if there was a reduction in noise exposure because of vessel regulations. Noise levels across all years were highly consistent with those measured in the same frequency band in a previous investigation despite

Table 6. Estimated regression coefficients of fixed effects included in top-ranking model related to received noise levels for data when vessel counts >1 (excluding research vessel only observations). The year 2010 was considered the reference factor level in the model, and the intercept estimate represents the predicted noise level value for year 2010 in the presence of 1 stationary vessel (LogCount = 0, SpeedCat0.stationary)

	Estimate	SE
Intercept	103.5	1.81
2011	-5.3	1.86
2012	0.7	1.57
2014	-5.1	1.77
LogCount	2.5	1.51
SpeedCat1.slow	2.0	0.79
SpeedCat2.medium	3.4	0.94
SpeedCat3.fast	4.8	1.59
LogDeployMin	1.2	0.69

differences in equipment, proximity to killer whales, and study periods (Holt et al. 2009). We expected to see a reduction in noise exposure as a result of regulations. If all vessels changed their minimum approach distance from 100 to 200 yd (91 to 183 m), on average, then the expected noise reduction would be in the range of 3 to 6 dB, assuming cylindrical or spherical spreading loss (10 to 20 × log distance). This is likely an oversimplified assumption, as there are many other variables that contribute to the amount of noise radiated by vessels, such as vessel type, size, propulsion system, operational speed, and number of propellers (Holt et al. 2009, Bassett et al. 2012, McKenna et al. 2013, Houghton et al. 2015, Veirs et al. 2016). Indeed, differences in noise levels were not explained by the implementation of vessel regulations or vessel distance. Additionally, more complicated sound propagation effects (beyond spherical/cylindrical spreading loss assumption) were not considered, although such sources of error are likely minimal given the relatively deep well-mixed water, the short ranges of vessels to the focal whale, and previous empirical assessments of propagation loss in the study area (Holt et al. 2009, Veirs et al. 2016).

Results of the mixed model analysis that included all data revealed that received noise was best predicted by year, vessel count, vessel speed category, and animal ID. Animal ID addressed any variation in noise due to repeated measures effect. Noise levels varied depending on the year, with 1 yr before (2010) and 1 yr after regulations (2012) being higher than the other 2 post-regulation years (2011 and 2014). Variability in noise levels among years could be due to differences in oceanographic conditions among years, although data could only be collected in Beaufort scales of 0 to 3, which limits variability associated with wind/wave-generated noise. Of the vessel variables considered in this paper, count and speed category were significant positive predictors of received noise levels. That is, an increase in vessel counts or an increase in speed category predicts an increase in received noise level, consistent with other studies (Holt et al. 2009, McKenna et al. 2013, Houghton et al. 2015, Veirs et al. 2016). Data collection occurred from a research vessel as required for tagging, and whale and vessel observations occurred over dynamic spatial scales. Thus, effects of the research vessel could not be removed, but excluding periods in which it was the only vessel revealed very similar findings, with the exception of the added variable of log deployment minute. However, log deployment minute and vessel count were not significant terms in the final model, the latter likely because of a more

limited sample size and restricted range of count values. Range restriction on other vessel parameters was also unavoidable given constraints of this observational study. Noise levels could not be measured at night when counts were likely lower given the requirement to conduct visual observations. However, there were many observations when the research platform was the only vessel present within 1.5 km (i.e. count = 1). Fundamental spatial restriction during high counts is also evident from the positive correlation between vessel counts and distance. Local whale-watching vessels often follow voluntary guidelines including reduced speed suggestions (typically <7 knots), which affected the range of observed vessel speeds. Additionally, limitations of the vessel variables tested are due to the challenges of obtaining certain attribute information on smaller vessels that affect levels of radiated sound (e.g. propulsion system). Houghton et al. (2015), using a different analytical approach, tested an expanded list of vessel variables to relate to noise levels with a more limited sample size and found that of the 8 vessel variables tested, speed and number of propellers were the best predictors of received noise level. In the current study, the number of propellers was an untested vessel variable, as an increased sample size was desired to compare noise levels before and after the implementation of vessel regulations.

Parameter estimates of fixed effects in the top-ranking model (Tables 5 & 6) predict changes in noise levels received by killer whales when nearby vessel counts or vessel speed category change. Speed categorization was the most accurate method to quantify operational speed of the smaller whale-watching vessels, as many did not transmit AIS signals. Despite potential error associated with estimated vessel speed and speed category being averaged based on median value across all vessels, the parameter estimate for each speed category in our top-ranked model is plausible relative to other investigations that measured vessel speed more accurately (McKenna et al. 2013, Veirs et al. 2016; our Tables 5 & 6). For example, if vessels change their speed from stationary to slow, on average, then the noise level is predicted to increase by approximately 2 dB, while a change from slow to medium or medium to fast speed, on average, would result in a predicted increase in noise level of approximately 1 dB (Table 5). Likewise, noise levels are predicted to increase as vessel counts increase according to $2.5 \times \log_{10}$ count (Table 5). Given the log relationship, the effect of vessel count is predicted to be smaller as count linearly increases (Fig. 2C). For example, if vessel

counts increase from 1 to 10, noise levels are predicted to increase by approximately 2.5 dB, but if counts increase from 11 to 20 (also a 9 count increase), then noise would increase by about 0.6 dB. While these scenarios predict noise differences that seem small, it is important to interpret predictions in a biological context and across a broader range of vessel scenarios. For example, vessel counts can reach a maximum of over 80 vessels, and vessels do not always follow regulations and guidelines (Seely 2015). Furthermore, if killer whales are listening to faint sounds in a noise-limited environment, even a decrease of a few dB from reduced vessel counts or reduced speed could improve sound use for communication and echolocation and/or lower energetic costs associated with noise-induced vocal modification (Holt et al. 2009, 2015). These biological considerations are critical considering the risk factors for this population (e.g. prey limitation and vessel disturbance).

Disturbance from vessels and noise can lead to behavioral and acoustic effects in killer whales. According to US regulations, vessel operators are prohibited from approaching, in any manner, killer whales within 200 yd (183 m, NOAA 2011). In the current study, differences in noise levels were not explained by the implementation of vessel regulations nor by an effect of average vessel distance. Comparisons of noise levels by year revealed that 2 of the 3 yr after regulations had lower noise levels, but 2012 noise levels were not different from those of 2010 before regulations came into effect (Fig. 2A). The results of this investigation should be interpreted with caution given the observed inter-annual variability, the limitation of having only 1 yr of baseline data for comparison, and other aforementioned logistical constraints of the observational study. Vessel regulation adherence may also depend on the presence of enforcement, but this potential factor could not be included in the analysis due to small sample size. It is also possible that vessel behavior has changed after regulations came into effect including more likely operating under power at relatively close ranges (e.g. Seely 2015, p 35–36). Any potential noise reduction from an increased viewing distance would result in little overall change in noise if this was also associated with increased vessel speed. Indeed, the proportion of noise levels during which median vessel speed was categorized as stationary significantly decreased following the implementation of vessel regulations. Some vessel regulations in the USA do include speed restrictions, particularly those to protect right whales *Eubalaena glacialis* (NOAA 2008) and humpback whales *Megaptera novaeangliae* (NOAA 2001) from

vessel collision risk. A speed restriction was considered during the regulations process for killer whales, but this component was deemed too difficult to enforce given lack of vessel tracking technology with smaller vessels (NOAA 2009, 2011). Future technological advances may enable more feasible solutions to this specific enforcement challenge. Other measures to reduce nearby vessel counts around endangered killer whales, such as a protected area free of vessels (no-go zone), have also been proposed (NOAA 2009, 2011).

Received SPLs measured from the DTAGs provided spatially accurate representations of vessel noise exposure in killer whales, but flow noise and tagging logistics temporally limited data collection and availability. Potential shielding effects of the whale's body relative to the hydrophones on the tag, particularly at the higher frequencies relative to body size, might also affect measured noise levels. Nonetheless, received noise levels reported here were highly consistent with those measured by acoustic monitoring equipment in close proximity to individuals of the same killer whale population in the same study area (Holt et al. 2009). Future work that continues to monitor noise levels across a broader range of vessel scenarios and time periods is needed as vessel regulations continue to be implemented. While increased noise can have many effects on the hearing and use of sound by killer whales and other marine mammals, noise is not the only effect of nearby vessels. Disturbance from vessel presence itself is also a documented effect (Williams et al. 2011); thus, there is a complexity of factors that need to be considered by protected species managers. Future investigations that incorporate analyses of behavioral effects related to the use of sound and other mechanisms that are attributed to specific vessel variables and associated noise are needed to determine these additional effects.

Acknowledgements. This work was funded by the NOAA Ocean Acoustics Program and NOAA/NMFS Northwest Fisheries Science Center. The ECHO Program, Port of Vancouver, provided funds for DTAG hydrophone calibrations. We thank Juliana Houghton, Dave Haas, Robin Baird, and others for field assistance. We thank Eric Ward and Eli Holmes for statistical advice; Sander von Benda-Beckmann for providing advice and Matlab scripts to separate flow noise from ambient noise; Alessandro Bocconcelli, Ann Allen, Frants Jensen, Alison Stimpert, Stacy DeRuiter, Paul Wensveen, and Patrick Miller for providing valuable support and feedback on DTAG logistics, analysis, and calibration; Damon Holzer for the preparation of Fig. 1; and 4 anonymous reviewers for their constructive feedback on the manuscript.

LITERATURE CITED

- ✦ Barrett-Lennard LG, Ford JKB, Heise KA (1996) The mixed blessings of echolocation: differences in sonar use by fish-eating and mammal-eating killer whales. *Anim Behav* 51:553–565
- ✦ Bassett C, Polagye B, Holt MM, Thompson J (2012) A vessel noise budget for Admiralty Inlet, Puget Sound, Washington (USA). *J Acoust Soc Am* 132:3706–3719
- ✦ Bates D, Maechler M, Bolker B, Walker S (2014) lme4: linear mixed-effects models using Eigen and S4. R package version 1.1-7. <http://cran.r-project.org/package=lme4>
- ✦ Branstetter BK, St. Ledger J, Acton D, Stewart J, Houser D, Finneran J, Jenkins K (2017) Killer whale (*Orcinus orca*) behavioral audiograms. *J Acoust Soc Am* 141:2387–2398
- ✦ DeRuiter SL, Southall BL, Calambokidis J, Zimmer WMX and others (2013) First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biol Lett* 9:20130223
- ✦ Erbe C (2002) Underwater noise of whale-watching boats and potential effects on killer whales (*Orcinus orca*). *Mar Mamm Sci* 18:394–418
- ✦ ESA (Endangered Species Act) (1973) Endangered Species Act of 1973. 16 United States Code, Chap 35, Sections 1531–1544. www.nmfs.noaa.gov/pr/pdfs/laws/esa.pdf
- ✦ Fox J, Wesiberg S, Friendly M, Hong J, Andersen R, Firth D, Taylor S (2015) Package 'effects'. R package version 3.0-4. <http://cran.r-project.org/package=effects>
- ✦ Geffroy B, Samia DSM, Bessa E, Blumstein DT (2015) How nature-based tourism might increase prey vulnerability to predators. *Trends Ecol Evol* 30:755–765
- ✦ Giles DA (2014) Southern resident killer whales (*Orcinus orca*): a novel non-invasive method to study southern resident killer whales and vessel compliance with regulations. PhD dissertation, University of California, Davis, CA
- ✦ Goldbogen JA, Calambokidis J, Shadwick RE, Oleson EM, McDonald MA, Hildebrand JA (2006) Kinematics of foraging dives and lunge-feeding in fin whales. *J Exp Biol* 209:1231–1244
- ✦ Hanson MB, Baird RW, Ford JKB, Hempelmann-Halos J and others (2010) Species and stock identification of prey consumed by endangered southern resident killer whales in their summer range. *Endang Species Res* 11: 69–82
- ✦ Holt MM, Noren DP, Veirs V, Emmons CK, Veirs S (2009) Speaking up: Killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. *J Acoust Soc Am* 125:EL27–EL32
- ✦ Holt MM, Noren DP, Dunkin RC, Williams TM (2015) Vocal performance affects metabolic rate in dolphins: implications for animals communicating in noisy environments. *J Exp Biol* 218:1647–1654
- ✦ Houghton J, Holt MM, Giles DA, Hanson MB and others (2015) The relationship between vessel traffic and noise levels received by killer whales (*Orcinus orca*). *PLOS ONE* 10:e0140119
- ✦ Johnson MP, Tyack PL (2003) A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE J Oceanic Eng* 28:3–12
- ✦ Krahn MM, Wade PR, Kalinowski ST, Dahlheim ME and others (2002) Status review of southern resident killer whales (*Orcinus orca*) under the Endangered Species Act. NOAA Tech Memo NMFS-NWFSC-54. US Department of Commerce, Seattle, WA

- ✦ Kuznetzova A, Brockhoff PB, Christensen RHB (2012) lmerTest: tests in linear mixed effects models. R package version 2.0-30. <https://cran.r-project.org/package=lmerTest>
- Leonard J (2006) Wildlife watching in the U.S.: the economic impacts on national and state economies in 2006: addendum to the 2006 national survey of fishing, hunting, and wildlife-associated recreation report 2006. US Fish and Wildlife Service, Arlington, VA. [https://www.fws.gov/uploadedFiles/nat_survey2006_economics\(1\).pdf](https://www.fws.gov/uploadedFiles/nat_survey2006_economics(1).pdf)
- ✦ Lusseau D, Bain DE, Williams R, Smith JC (2009) Vessel traffic disrupts the foraging behavior of southern resident killer whales *Orcinus orca*. *Endang Species Res* 6: 211–221
- ✦ McKenna MF, Wiggins SM, Hildebrand JA (2013) Relationship between container ship underwater noise levels and ship design, operational and oceanographic conditions. *Sci Rep* 3:1–10
- ✦ MMPA (Marine Mammal Protection Act) (1972; amended through 2007) 16 United States Code, Sections 1361–1423h. www.nmfs.noaa.gov/pr/pdfs/laws/mmpa.pdf
- NMFS (National Marine Fisheries Service) (2008) Recovery plan for southern resident killer whales (*Orcinus orca*). National Marine Fisheries Service, Seattle, WA
- NOAA (National Oceanic and Atmospheric Administration) (2001) Regulations governing the approach to humpback whales in Alaska. *Fed Regist* 66:29502–29509
- NOAA (2008) Endangered fish and wildlife; final rule to implement speed restrictions to reduce the threat of ship collisions with north Atlantic right whales. *Fed Regist* 73: 60173–60191
- NOAA (2009) Protective regulations for killer whales in the northwest region under the Endangered Species Act and Marine Mammal Protection Act. *Fed Regist* 74: 37674–37686
- NOAA (2011) Protective regulations for killer whales in the northwest region under the Endangered Species Act and Marine Mammal Protection Act. *Fed Regist* 76: 20870–20890
- ✦ Noren DP, Johnson AH, Rehder D, Larson A (2009) Close approaches by vessels elicit surface active behaviors by southern resident killer whales. *Endang Species Res* 8: 179–192
- NRC (National Research Council) (2003) Ocean noise and marine mammals. The National Academies Press, Washington, DC
- NRC (2005) Marine mammal populations and ocean noise: determining when noise causes biologically significant effects. The National Academies Press, Washington, DC
- ✦ R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. www.r-project.org
- Richardson JW, Greene CR, Malme CI, Thomson DH (1995) Marine mammals and noise. Academic Press, San Diego, CA
- ✦ Seely E (2015) Final 2015 Soundwatch Program annual contract report. The Whale Museum, Friday Harbor, WA. <http://whalemuseum.org/pages/soundwatch-boater-education-program> (accessed on 15 July 2016)
- ✦ Senigaglia V, Christiansen F, Bejder L, Gendron D and others (2016) Meta-analyses of whale-watching impact studies: comparisons of cetacean responses to disturbance. *Mar Ecol Prog Ser* 542:251–263
- ✦ Szymanski MD, Bain DE, Kiehl K, Pennington S, Wong S, Henry KR (1999) Killer whale (*Orcinus orca*) hearing: auditory brainstem response and behavioral audiograms. *J Acoust Soc Am* 106:1134–1141
- ✦ Tyack PL, Zimmer WMX, Moretti D, Southall BL and others (2011) Beaked whales respond to simulated and actual navy sonar. *PLOS ONE* 6:e17009
- ✦ Veirs S, Veirs V, Wood JD (2016) Ship noise extends to frequencies used for echolocation by endangered killer whales. *PeerJ* 4:e1657
- ✦ von Benda-Beckmann AM, Wensveen PJ, Samara FIP, Beerens S, Miller PJO (2016) Separating underwater ambient noise from flow noise recorded on stereo acoustic tags attached to marine mammals. *J Exp Biol* 219:2271–2275
- Wensveen PJ (2016) Detecting, assessing, and mitigating the effects of naval sonar on cetaceans, Appendix III. PhD dissertation, University of St Andrews
- ✦ Williams R, Noren DP (2009) Swimming speed, respiration rate, and estimated cost of transport in adult killer whale. *Mar Mamm Sci* 25:327–350
- ✦ Williams R, Lusseau D, Hammond PS (2006) Estimating relative energetic costs of human disturbance to killer whales (*Orcinus orca*). *Biol Conserv* 133:301–311
- ✦ Williams R, Bain DE, Smith JC, Lusseau D (2009) Effects of vessels on behaviour patterns of individual southern resident killer whales *Orcinus orca*. *Endang Species Res* 6: 199–209
- Williams R, Ashe E, Sandilands D, Lusseau D (2011) Stimulus-dependent response to disturbance affecting the activity of killer whales. Report No. SC/63/WW5, Scientific Committee of the International Whaling Commission. IWC, Cambridge, p 1–27
- Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York, NY
- ✦ Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. *Methods Ecol Evol* 1:3–14

Editorial responsibility: Mike Noad,
 Gatton, Queensland, Australia

Submitted: September 9, 2016; Accepted: May 18, 2017
 Proofs received from author(s): July 7, 2017