

Quantifying spatial and temporal variation of North Pacific fin whale (*Balaenoptera physalus*) acoustic behavior

Frederick I. Archer¹  | Shannon Rankin¹ | Kathleen M. Stafford² | Manuel Castellote³ | Julien Delarue⁴

¹Marine Mammal and Turtle Division, Southwest Fisheries Science Center, La Jolla, California

²Applied Physics Laboratory, University of Washington, Seattle, Washington

³Joint Institute for the Study of the Atmosphere and Ocean (JISAO), University of Washington, and Marine Mammal Laboratory, Alaska Fisheries Science Center, Seattle, Washington

⁴JASCO Applied Sciences, Dartmouth, Nova Scotia, Canada

Correspondence

Frederick I. Archer, Southwest Fisheries Science Center, 8901 La Jolla Shores Drive, La Jolla, CA 92037.
Email: eric.archer@noaa.gov

Abstract

In order to help develop hypotheses of connectivity among North Pacific fin whales, we examine recordings from 10 regions collected in the spring and fall. We develop a Random Forest model to classify fin whale note types that avoids manual note classification errors. We also present a method that objectively quantifies the note and pattern composition of recordings. We find that fin whale recordings near Hawaii have distinctive patterns, similar to those found in other regions in the central North Pacific, suggesting potential migration pathways. Our results are consistent with previous studies that suggest there may be two different populations utilizing the Chukchi Sea and central Aleutians in the fall and mix to some degree in the southern Bering Sea. Conversely, we found little difference between spring and fall recordings in the eastern Gulf of Alaska, suggesting some residency of whales in this region. This is likely due to fine scale similarities of calls among the inshore regions of British Columbia, while offshore areas are being utilized by whales traveling from various distant areas. This study shows how our novel approach to characterize recordings is an objective and informative way to standardize spatial and temporal comparisons of fin whale recordings.

KEYWORDS

acoustics, *Balaenoptera physalus*, bout, classification, clustering, fin whale, INI, North Pacific

1 | INTRODUCTION

Describing patterns of population structure in large whales has always been a challenging endeavor. Historically, these studies have relied on morphological and genetic data, much of which was obtained during commercial whaling (Mizroch, Rice, Zwiefelhofer, Waite, & Perryman, 2009). Due to their large size and wide-ranging, pelagic distributions, it is often difficult to obtain a sufficient number of morphological and genetic samples to describe whale populations. Additionally, complex movement patterns between high latitude feeding areas and low latitude breeding areas are not always fully understood, making it difficult to analyze and interpret the results from data sets composed of samples collected from different portions of these ranges and in different times of the year.

Over the past several decades, acoustic data have proven to be a useful source of information for geographical differentiation in several species of baleen whales (Cerchio, Jacobsen, & Norris, 2001; McDonald, Mesnick, & Hildebrand, 2006; Oleson, Barlow, Gordon, Rankin, & Hildebrand, 2003). In contrast to the complex calls of humpback whales (*Megaptera novaeangliae*) (Payne & Payne, 1985; Winn & Winn, 1978), fin whales (*Balaenoptera physalus*) make stereotyped signals composed of short (<1 s), downswept notes (Watkins, 1981; Watkins, Tyack, Moore, & Bird, 1987). The most common of these easily recognizable notes occurs around 20 Hz, and has been named the 20-Hz pulse, 20-Hz note, or classic note (C). A variation of this note type is called a “backbeat” (B), that differs in center frequency and bandwidth (Clark, Borsani, & Notarbartolo di Sciarra, 2002). In some regions, including the North Atlantic and the Antarctic, there is also a higher frequency component associated with the 20-Hz pulses, which has been found to vary geographically (Simon, Stafford, Beedholm, Lee, & Madsen, 2010; Širović, Hildebrand, Wiggins, & Thiele, 2009). Fin whales produce other signals that have received considerably less research attention, including a ~45 Hz note that is not produced in a patterned sequence and is produced at different times of year than the 20-Hz note (Širović, Williams, Kerosky, Wiggins, & Hildebrand, 2013; Watkins, 1981).

These notes can be produced in stereotyped series that have often been referred to as “song,” and are thought to be a male reproductive display (Clark, 1990; Croll et al., 2002; Watkins, 1981). Fin whale song composition has been shown to vary in primarily two characteristics: the frequency (including bandwidth and center or peak frequency) and the time interval between successive 20-Hz notes (internote interval, INI; Hatch, 2004; Watkins et al., 1987). Most commonly, differences in the INI have been used to differentiate populations of fin whales in the Mediterranean Sea and the Atlantic, Pacific, and Southern Oceans (Castellote, Clark, & Lammers, 2012; Delarue, Todd, Van Parijs, & Di Iorio, 2009; Hatch, 2004; Morano et al., 2012; Širović et al., 2009; Thompson, Findley, & Vidal, 1992; Weirathmueller et al., 2017). These studies have characterized songs as “singlet” where 20-Hz notes that share the same time and bandwidth characteristics are repeated with a regular INI, “doublet” songs in which the INI alternates between relatively shorter or longer durations and sometimes different note types (e.g., Oleson, Širović, Bayless, & Hildebrand, 2014), and “triplet” songs composed of two note types and with fixed INIs (Delarue, Martin, Hannay, & Berchok, 2013). These studies suggest that fin whale song characteristics, specifically the temporal structure of note sequences, might be useful in differentiating among “acoustic populations” as has been done with blue whales (*Balaenoptera musculus*) (Buchan, Stafford, & Huckle-Gaete, 2015; McDonald et al., 2006; Mellinger & Clark, 2003; Samaran et al., 2013; Stafford, Chapp, Bohnenstiel, & Tolstoy, 2011; Stafford, Nieukirk, & Fox, 1999, 2001).

In the absence of a robust understanding of fin whale population structure and movement patterns within the North Pacific, passive acoustic data have been proposed as a useful line of evidence to generate hypotheses of connectivity. In conjunction with genetic and photo ID studies, acoustic data indicate that there is a resident population of fin whales in the Gulf of California, Mexico (Nigenda-Morales, Flores-Ramirez, Urban-R, & Vazquez-Juarez, 2008; Silber, Newcomer, Silber, Pérez-Cortés M., & Ellis, 1994; Širović, Oleson, Buccowich, Rice, & Bayless, 2017; Thompson et al., 1992); with possible seasonal influx or mixing of whales from the Pacific side of Baja California, Mexico. Acoustic monitoring has also shown that fin whales are present year-round in southern California, being more common in offshore waters in the winter (Širović et al., 2015). However, little is known about their population structure in the remainder of the North Pacific Ocean. Several studies have described, with a variety of methodologies, the

structure of songs and seasonal and interannual variations therein for areas of the North Pacific, including Hawaii, the Bering Sea, the southern Chukchi Sea, southern California, and the Sea of Cortez (Delarue et al., 2013; Oleson et al., 2014; Širović et al., 2017; Weirathmueller et al., 2017). The similarity in song structure within and between years for several pelagic areas of the North Pacific (Hawaii, Bering Sea, southern California) suggests that a single population might use this oceanic basin even though there is also evidence for multiple song types in the Bering Sea (Delarue et al., 2013) and the northeast Pacific (Koot, 2015).

In this study, we develop new tools for the objective description and analysis of sequences of fin whale calls. We use these descriptors to examine relationships of fin whale call structure from regions around the North Pacific. We also compare the variability within the North Pacific to a small selection of calls from fin whales around Antarctica. The approach we take characterizes calls based on note type proportion, INI, and stereotypic characteristics in their sequences. We then quantify differences and similarities of sequences of calls among regions using machine learning classification and clustering methods.

2 | MATERIALS AND METHODS

The goal of this study was to objectively compare fin whale acoustic sequences collected from several regions across the North Pacific at different times of the year. We wished to avoid preselecting specific characteristics of the sequences (e.g., only comparing internote intervals in doublets) as we did not want to inadvertently discard features that might be relevant to population structure. Given that our data consisted of recordings of different lengths and compositions of notes, we needed to quantify them in a comparable way that incorporates information about the types and patterns of notes, as well as their spacing in time. Because we do not know nor wish to assume the behavioral state of the whales producing the calls, we avoid characterizing our recordings as “songs.” For the purposes of this study, we are defining the elements of fin whale acoustic sequences as follows (Figure 1):

Note: a 20-Hz pulse represented by two primary types in fin whale (following Clark et al., 2002): B (backbeat) and C (classic).

Unit: a note plus its following internote interval (INI, the time from the centroid of the note to the centroid of the following note).

Bout: sequential groupings of distinct units of variable length, separated by “rests” (following Watkins et al., 1987).

The approach we developed first identified and categorized the fundamental components of bouts (i.e., notes and units). We then quantified how frequently these components, and extended patterns of them, occurred within bouts. Additionally, we quantified how often each pattern of notes and units were repeated in a given bout. With these measures, we then examined the similarity of bouts among regions and seasons in an analysis of spatial and temporal variation. The details of each step of this process follow.

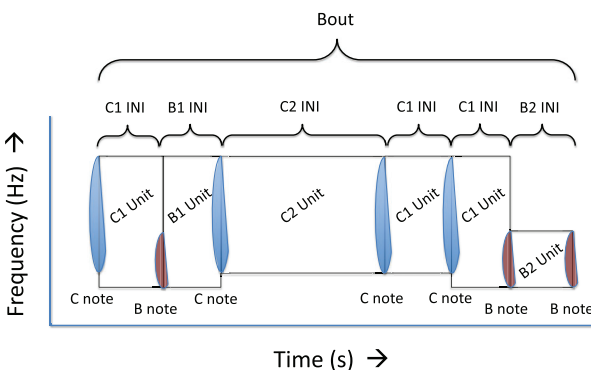


FIGURE 1 Example fin whale acoustic bout, demonstrating definitions of note types (B or C), INI (internote interval), and units (combination of note and INI).

Recordings were collected from 17 locations across the North Pacific, grouped into nine regions based on the proximity of several recording locations and large geographic gaps between regions (Figure 2). We also included recordings from the South Shetland Islands north of the Antarctic Peninsula ($\sim 62^{\circ}\text{S}$, 60°W) in order to compare variability within the North Pacific to variability between ocean basins. In order to minimize presence of overlapping bouts and address possible seasonal changes in bout structure, recordings were collected during the spring (February and March) and fall (September and October) between 2000 and 2014. All recordings were decimated to a sample rate of 512 Hz. In total, there were 15 unique combinations of regions and seasons, which are the strata used in all analyses in this study unless otherwise specified (Tables 1 and SM1).

Recordings were selected within the period sampled with the only restriction being that they contain recognizable fin whale bouts that did not appear to be created by more than a single animal. In each recording, an analyst used SpectroPlotter software (a custom software tool developed by JASCO Applied Sciences) to manually identify

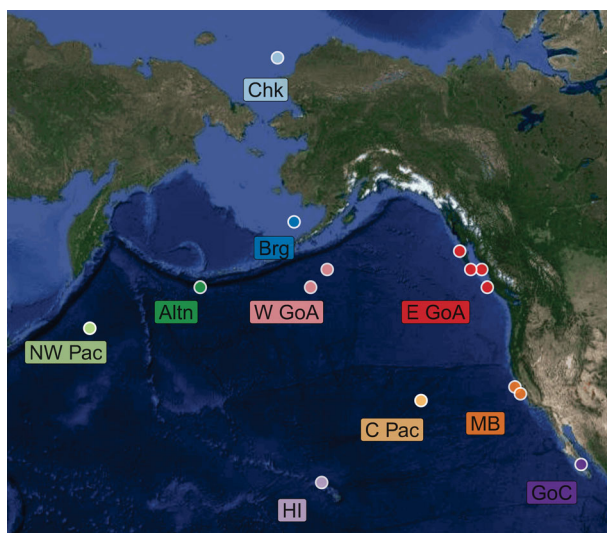


FIGURE 2 Location of recordings from each region. Each dot represents a separate location, with region abbreviations identified in the boxes.

TABLE 1 Summary of number of recordings, bouts, notes, and average note per bout in each stratum.

Region	Season	Recordings	Bouts	Notes	Notes/Bout
Chk	Fall	4	14	294	21.0
Brq	Fall	8	21	409	19.5
NW Pac	Spring	4	11	307	27.9
NW Pac	Fall	17	56	1665	29.7
Altn	Fall	1	6	113	18.8
W GoA	Spring	4	8	108	13.5
W GoA	Fall	8	30	718	23.9
E GoA	Spring	25	87	1,120	12.9
E GoA	Fall	36	112	1,573	14.0
C Pac	Fall	5	14	264	18.9
MB	Fall	5	16	409	25.6
HI	Spring	4	21	190	9.0
HI	Fall	1	20	231	11.6
GoC	Spring	9	20	399	19.9
Ant	Fall	15	37	790	21.4

discrete, characteristic fin whale notes. An analyst manually selected calls using detailed methods outlined in Rankin et al. 2018 (their appendix 1). A minimum of 500 notes were measured per season and region. The following spectral characteristics were measured: maximum frequency, minimum frequency, peak frequency, and center frequency. Frequency related measurements were based on annotation contents that contained a user-defined energy percentage selection, which varied for each bout. To facilitate consistency across analysts, user-defined energy selections followed detailed methods provided in appendix 2 of Rankin et al. 2018. The energy selection strategy was adopted to account for differences in background noise levels and spectra within the fin whale call frequency range among all deployment locations. INI was based on the time from the centroid of the note to the centroid of the following note to avoid the influence of variable signal-to-noise ratio within and between bouts among all deployment locations (see Rankin et al., 2018 for a detailed description of recorder hardware and processing methods).

2.1 | Note annotation and classification

As part of the note annotation process, each analyst identified notes as either (B)ackbeat or (C)lassic (Rankin et al., 2018). While high frequency notes were present in our Antarctic recordings, they were not found in the North Pacific recordings. Thus, we did not include them in further analysis.

Because note classification was done independently by multiple analysts with data from different regions and there appeared to be some interanalyst differences in classification, we developed a system to automatically classify notes based only on their measured spectral characteristics. The method we employed combined unsupervised clustering to identify note types with supervised classification to extend the clustering results to notes from all recordings.

For the unsupervised clustering phase, we used density clustering (Rodriguez & Laio, 2014), which is designed to identify points (peaks) in multidimensional data around which other points are clustered at a high density. Operationally, these peaks are defined as the k points surrounded by a large number of other points (ρ) within an a priori selected critical distance (dc), that are also a relatively large distance (δ) away from a point with a larger number of points within the same critical distance. The input for density clustering is a matrix of pairwise distances, which we calculated using the four measured frequency characteristics of the recorded notes and three computed values: bandwidth (maximum frequency minus minimum frequency), the difference between peak sound pressure level and centroid frequencies, and the difference between maximum and peak sound pressure level frequencies.

Rodriguez and Laio (2014) indicate that optimum values of dc are usually found between the 0.01 and 0.02 quantiles of the distribution of observed distances. In our optimization routine, we examined values of dc between the 0.01 to 0.03 quantiles of the observed distances, in increments of 0.001. For each value of dc and k clusters, ρ and δ were computed for each note, and threshold values of ρ and δ were selected as those halfway between the k and $k + 1$ points, with the largest values of the product of $\rho \times \delta$. We performed this clustering for $k = 2$ to 4.

We used Random Forest classification (Berk, 2006; Breiman, 2001) to determine the robustness and repeatability of our clustering methods. For every set of ρ and δ thresholds, if each cluster contained at least 100 notes, we then created a Random Forest classification model that predicted cluster membership based on the seven note spectral characteristics used above. In each tree in the model, n notes were randomly selected without replacement from each cluster, where n was half of the smallest cluster size. This produced a balanced model, ensuring that classification accuracy was not biased towards larger clusters (Archer, Martien, & Taylor, 2017; Berk, 2006). For each value of k , the value of dc producing a classification model with the lowest out-of-bag (OOB) error rate was selected as the optimum value, and the Random Forest model producing that error rate was chosen as the optimum classification model. If the OOB error rate at this optimum value was >0.05 , then that value of k was not considered.

2.2 | Unit identification

To form "units" (combinations of individual notes and the following INIs), we identified individual modes of INIs across all bouts using the Ckmeans.1d.dp method (Wang & Song, 2011), as implemented in the R package of the

same name. This method is a dynamic programming implementation of standard k -means clustering optimized for one-dimensional data, which is more appropriate for distributions of INIs than clustering methods designed for multivariate data sets. We clustered INIs following B-type notes separately from those following C-type notes. We evaluated $k = 2$ to 10 modes for INIs following each note. The optimum number of modes was chosen based on the BIC values of Gaussian mixture models fit to the data for each value of k .

Preliminary analyses identified a set of unusually short (<4.5 s) internote intervals (INIs), primarily in the Hawaii fall data. Given their short duration and regularity, they were suspected to be echoes resulting from multipath propagation. Additionally, throughout the data set there were longer INIs where it was possible that either the whale had periodically stopped calling (e.g., gaps or rests per Watkins et al., 1987) or notes within the recording were too faint for detection. Thus, to ensure that the distribution of INIs was appropriately representative of actual note production rates, we split individual recordings into separate bouts of sequential notes where all INIs were >4.5 s and <45 s.

Each unit was then labeled as the note followed by the note-specific INI mode cluster number (1 to k). For example, "B1" would describe a B followed by an INI that was part of B INI mode 1, while "C1" would be a fundamentally different unit, describing a C that was followed by an INI that was part of C INI mode 1 (Figure 1). The INI mode number was assigned sequentially, and does not represent interval duration.

2.3 | Bout composition

Previous studies have identified different kinds of singlet, doublet, and triplet patterns in fin whale bouts (Delarue et al., 2013; Koot, 2015; Oleson et al., 2014; Širović et al., 2017; Weirathmueller et al., 2017). For example, regularly spaced sequences of Cs are frequently observed, as are alternating BC or CC doublets, or even BCC triplets. In order to characterize bouts based on the composition of these patterns (p), we first identified all one- to three-note patterns. We did not explore four-note or larger patterns because triplets are the largest unit that has been described previously in the literature, longer patterns would start to significantly increase computational time, and most of the signal from these larger patterns should be captured by combinations of and correlations among the two- and three-note patterns that we are quantifying. The one note patterns are simply single instances of the two notes B and C. There are four unique two-note patterns (BB, BC, CB, and CC), and eight unique three-note patterns (BBB, CBB, BCB, CCB, BBC, CBC, BCC, and CCC) resulting in 14 unique one- to three-note patterns. Additionally, we identified all patterns of one- to three-units (combinations of note and INI). For each pattern (i) in a bout, we computed the following values (further described below):

N_b : Number of notes in bout b .

L_i : Number of notes in pattern i .

M_{ib} : Maximum number of nonoverlapping instances of pattern i possible in bout b ($= N_b/L_i$).

N_{ib} : Actual number of nonoverlapping instances of pattern i in bout b .

P_{ib} : Proportion of the bout b composed of pattern i ($= N_{ib}/M_{ib}$).

R_{ib} : Number of runs (sets of sequential occurrences) of pattern i in bout b .

$\bar{P}_{R(i)b}$: Average proportion of runs of pattern i in bout b ($= \sum_{j=1}^{R_i} \left(\frac{\text{length}(j)}{M_{ib}} \right) / R_{ib}$), where $\text{length}(j)$ is the length of the j -th run.

$\text{Max}[P_{R(i)b}]$: Maximum proportion of a run of pattern i in bout b ($= \frac{\max[\text{length}(j)]}{M_{ib}}$).

An example of these measures for a sample 20-note bout is given in Table 2. In each row, we show where the unique occurrences and runs of each pattern are in the bout and their summary metrics. Only sequential, non-overlapping patterns found starting from the beginning of the bout are counted as we wanted to quantify the presence and "clumpiness" of patterns as distinct components of a call. For example, the bout begins with BCBCB. If we are considering the BCB pattern, there is only one occurrence of this pattern in these first five notes and only three in the entire bout. We do not count the BCB that starts at the third note as this B is actually the last note of the BCB pattern that initiates the bout.

TABLE 2 Example of pattern proportions and runs for single, two-note, and three-note patterns in a sample 20-note bout. Runs for each pattern listed in the first column are underlined under the bout. Metrics used for the bout composition analyses are shown in the right-hand columns (see text for definitions).

Pattern (i)	Bout:																		M_i	N_i	P_i	R_i	$\bar{P}_{R(i)}$	Max($P_{R(i)}$)		
	B	C	B	C	B	B	C	C	B	C	B	C	B	C	B	B	C	C							B	C
B	<u>B</u>		<u>B</u>		<u>B</u>	<u>B</u>			<u>B</u>		<u>B</u>		<u>B</u>	<u>B</u>			<u>B</u>			20	10	0.5	8	0.06	0.1	
C		<u>C</u>		<u>C</u>			<u>C</u>	<u>C</u>		<u>C</u>		<u>C</u>				<u>C</u>	<u>C</u>		<u>C</u>	20	10	0.5	8	0.06	0.1	
BB					<u>B</u>	<u>B</u>								<u>B</u>	<u>B</u>					10	2	0.2	2	0.1	0.1	
BC	<u>B</u>	<u>C</u>	<u>B</u>	<u>C</u>		<u>B</u>	<u>C</u>		<u>B</u>	<u>C</u>	<u>B</u>	<u>C</u>	<u>B</u>	<u>C</u>		<u>B</u>	<u>C</u>		<u>B</u>	<u>C</u>	10	8	0.8	5	0.16	0.3
CB			<u>C</u>	<u>B</u>	<u>C</u>	<u>B</u>			<u>C</u>	<u>B</u>	<u>C</u>	<u>B</u>	<u>C</u>	<u>B</u>	<u>C</u>	<u>B</u>			<u>C</u>	<u>B</u>	10	7	0.7	3	0.23	0.4
CC							<u>C</u>	<u>C</u>									<u>C</u>	<u>C</u>			10	2	0.2	2	0.1	0.1
BBB																					6.67	0	0	0	0	0
CBB				<u>C</u>	<u>B</u>	<u>B</u>								<u>C</u>	<u>B</u>	<u>B</u>					6.67	2	0.3	2	0.15	0.15
BCB	<u>B</u>	<u>C</u>	<u>B</u>					<u>B</u>	<u>C</u>	<u>B</u>		<u>B</u>	<u>C</u>	<u>B</u>							6.67	3	0.45	3	0.15	0.15
CCB						<u>C</u>	<u>C</u>	<u>B</u>									<u>C</u>	<u>C</u>	<u>B</u>		6.67	2	0.3	2	0.15	0.15
BBC					<u>B</u>	<u>B</u>	<u>C</u>							<u>B</u>	<u>B</u>	<u>C</u>					6.67	2	0.3	2	0.15	0.15
CBC		<u>C</u>	<u>B</u>	<u>C</u>				<u>C</u>	<u>B</u>	<u>C</u>		<u>C</u>	<u>B</u>	<u>C</u>				<u>C</u>	<u>B</u>	<u>C</u>	6.67	4	0.6	4	0.15	0.15
BCC					<u>B</u>	<u>C</u>	<u>C</u>							<u>B</u>	<u>C</u>	<u>C</u>					6.67	2	0.3	2	0.15	0.15
CCC																					6.67	0	0	0	0	0

The value P_{ib} , the proportion of bout b that is composed of pattern i , is computed such that $P_{ib} = 1$ if the bout is composed entirely of the pattern. This is true regardless of the length of the pattern (e.g., a BBBBBB bout has $P_{ib} = 1$ for B, BB, and BBB patterns). In Table 2, the bout has equal proportions of Bs and Cs ($P_{ib} = 0.5$ for both), but higher proportions of BC and CB (0.8 and 0.7) than BB and CC (0.2) indicating that notes are not randomly arranged in this bout.

We define a “run” as a sequential occurrence of a pattern. In Table 2, there are 10 runs of the pattern B, but only two runs of the pattern BB, and none of pattern BBB. Focusing just on runs of the two-note patterns, we see that there are two runs of BB and CC each, three runs of CB, and five runs of BC. Since the number of runs of each pattern increases with the length of a bout, we summarize runs by computing the average proportion of the bout composed of runs of a given pattern, $\bar{P}_{R(i)}$, and the maximum proportion of the bout composed of a run of the pattern $\text{Max}(P_{R(i)})$. For example, a run of one occurrence of a two-note pattern is 0.1 of the length of a 20-note bout [1 occurrence/(20 notes in bout/2 notes in pattern) = 0.1]. Thus, in Table 2, the five runs of BC represent 0.2, 0.1, 0.3, 0.1, and 0.1 of the length of the bout respectively. Thus, $\bar{P}_{R(BC)} = 0.16$, and $\text{Max}(P_{R(BC)}) = 0.3$. Although there are more runs of BC in this bout, the CB runs are longer on average. So, even though the two patterns are inversions of one another, we would tend to characterize this bout as being more CB-like than BC-like. Finally, because $\bar{P}_{R(i)}$ and $\text{Max}(P_{R(i)})$ are the same (=0.15) for all of the three-note patterns except for BBB and CCC, which are both 0, we see that this bout is not significantly composed of any particular three-note pattern.

These three summary metrics, P_{ib} , $\bar{P}_{R(ib)}$, and $\text{Max}(P_{R(ib)})$ that quantify how patterns are distributed in bouts were computed for all one-, two-, and three-note patterns of the notes alone (B, C, BB, BC, etc.), as well as units (B1, B2, C1, C2, B1C1, B1C2, etc.) for each bout. These metrics were then used to cluster bouts within strata and build a model to classify bouts to strata as described below. We only considered bouts composed of nine notes or more to ensure that each bout contained a sufficient number of notes to characterize all summary metrics.

To identify groups of similar bouts within each of the 15 strata (10 regions and one or two seasons), we used Partitioning Around Medoids (PAM) clustering (Reynolds, Richards, de la Iglesia, & Rayward-Smith, 2006). Clustering

was based on pairwise Euclidean distances computed from the bout pattern proportion and run summaries described above. We selected the optimum number of PAM clusters in each stratum using the Gap method (Tibshirani, Walther, & Hastie, 2001) and the first local maximum criteria. We used the second power of the Euclidean distances to calculate the Gap statistic following Tibshirani et al. (2001) and standard errors of the gap statistic were computed using 500 bootstrap replicates. We examined one to three clusters per stratum with the restriction that all clusters had at least three bouts. We used the medoid bouts identified by PAM as representatives of their respective clusters in each stratum. Relationships among these representative bouts were then visualized with hierarchical clustering based on the pairwise Euclidean distances among them.

2.4 | Strata classification

We generated a set of Random Forest (Breiman, 2001) models that used the pattern proportion and run summaries as predictors to classify bouts to their strata (region and season) of origin. The Random Forest model was balanced with respect to number of bouts across strata in the trees as previously described. A total of 100,000 trees were run for each model to ensure convergence of the OOB error rate.

All analyses were run with R version 3.5.3 (R Core Team, 2019). Density clustering was conducted with the *densityClust* package version 0.3 (Pedersen, Hughes, & Qiu, 2017), PAM clustering and computation of the Gap statistic was performed with the *cluster* package version 2.0.9 (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2018), and the Random Forest models were run with the *randomForest* package version 4.6–14 (Liaw & Wiener, 2002).

3 | RESULTS

3.1 | Note classification

A total of 11,769 notes were recorded in all regions and seasons combined, of which 10,486 had data for all four spectral measures. The mean Pearson correlation coefficient among these spectral measures was 0.63. Based on these along with the three derived measures, Density Clustering identified two primary clusters (Figure 3). In the Random Forest model, 99.3% of the notes were correctly classified to their cluster (78 misclassifications). The clusters corresponded to notes with mean centroid frequencies of 18.8 Hz and 23.7 Hz (Table 3 and Figure 4), which correspond to the previously defined B and C note types, respectively. We then used these designations to create a classifier with which we predicted the note type of 1,069 of the notes with missing data based on centroid

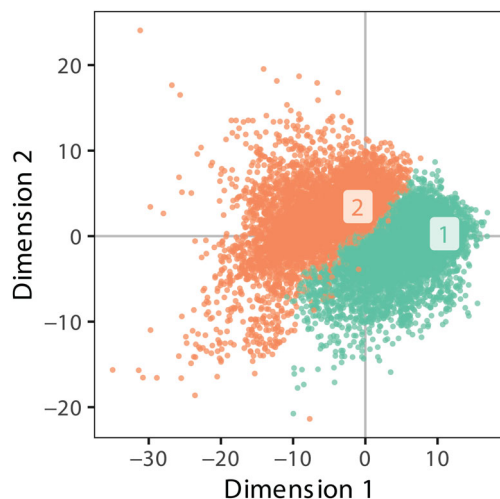


FIGURE 3 Multidimensional scaling of notes, colored according to clusters designated by Density Clustering. Cluster 1 corresponds to B notes and cluster 2 corresponds to C notes.

TABLE 3 Summary of spectral characteristics of notes based on Density Clustering. In each cell the top values are mean (standard deviation), and the bottom values are median (central 95th percentile).

Note	Centroid frequency	Peak SPL frequency	Maximum frequency	Minimum frequency
B	18.8 (1.5)	18.5 (1.5)	25.5 (3.1)	13.6 (2.4)
	18.6 (16.6–21.8)	18.2 (16.4–21.5)	25.1 (20.7–32.3)	4.2 (7.1–16.7)
C	23.7 (2.2)	22.9 (2.8)	33.9 (3.5)	17.2 (2.7)
	23.3 (20.5–29)	22.7 (19.1–28.6)	33.9 (27.9–40.8)	17.5 (10.7–21.3)

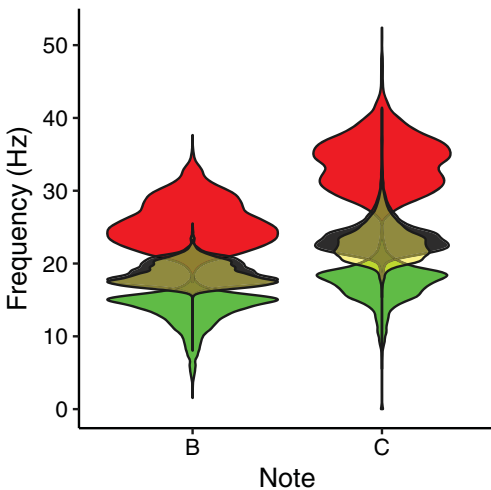


FIGURE 4 Overlay of distributions of spectral measures for B and C notes as designated by Density Clustering. Distributions are minimum (green), maximum (red), centroid (black), and peak sound pressure level frequency (yellow).

frequency, peak sound pressure level frequency, and the difference between them. The remaining 214 notes were then classified based on centroid frequency alone. Random Forest OOB error rates for these latter two models with reduced spectral measures were 6% and 7%, respectively.

We found considerable variability in the correspondence between note designations from the Density Cluster analysis and the original subjective designations from the analysts (Figures SM1 and SM2). In some regions, such as the Aleutians and Gulf of California, there was perfect correspondence between the two methods. However, in many others, the two methods disagreed. In some cases, such as Monterey Bay, Antarctica, and most notably Hawaii, the methods disagreed for over 90% of the notes. In the data from Hawaii, the analysts designated 92% of the notes as C, while Density Clustering designated all of these as Bs. Although not as pronounced, this was also the predominant pattern in other regions, where there was tendency for a large number of notes manually identified by readers as Cs to be designated as Bs by Density Clustering. Given that Density Clustering is more objective and repeatable, we use these note type designations in all analyses in this study.

The proportion of Bs and Cs varied across strata (Figure 5). All notes from Hawaii in the spring were Bs, as were 99% of the notes from Hawaii in the fall. Similarly, 98% of the notes from the western Gulf of Alaska in the spring were Bs. However, a majority of the western Gulf of Alaska notes in the fall were Cs (73%). Recordings from the Aleutians in the fall had the greatest proportion of Cs (99%). The B:C ratio in all strata was significantly different from 1:1 (binomial test $p \leq .05$).

3.2 | Bout composition

The Ckmeans.1d.dp cluster analysis identified five B INI modes and seven C INI modes (Table 4 and Figure 6). For the B INIs, the BIC value at $k = 5$ was 212 and 353 units greater than $k = 4$ and 6, respectively. For the C INIs, the

FIGURE 5 Proportion of B notes in each stratum for Spring (circles), and Fall (triangles).

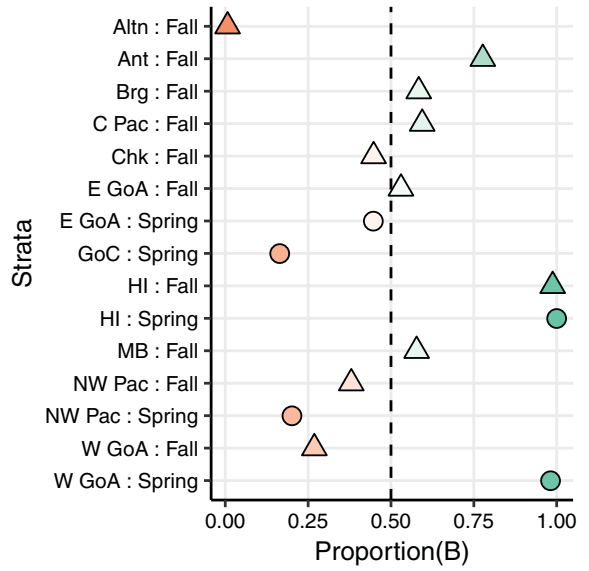


TABLE 4 Summary of INI distribution in each unit cluster identified by Ckmeans.1d.dp analysis.

Unit	n	Mean (s)	Median (s)	Range (s)
B1	1,271	7.5	8.2	1.1–110.8
B2	2,222	13.2	13.3	10.8–16.9
B3	980	20.7	20.8	16.9–24
B4	923	27.4	27.4	24.1–30.8
B5	330	34.1	33.8	30.8–44.9
C1	304	5.3	4.6	1.2–8.1
C2	862	10.8	10.7	8.2–12.6
C3	1,365	14.5	14.5	12.6–16
C4	1,298	17.6	17.5	16.1–19.9
C5	1,016	22.1	22.2	19.9–24.8
C6	269	27.5	27.2	24.8–32.1
C7	50	37.1	37	32.4–44.7

BIC value at $k = 7$ was 95 and 57 units greater than $k = 6$ and 8, respectively (Figure 6B). When the individual recordings were censored for INIs between 4.5 s and 45 s, a total of 473 bouts were defined that had nine notes or more (Table 1). There were 8,590 notes across all bouts, with bouts having an average of 18 notes, with a maximum of 94 notes (Figure SM3). Average bout duration was 294 s, ranging from 75 s to 1,449 s.

The variability of bouts across strata was enhanced when unit type is considered. As shown in Figure 7, although Hawaiian and Antarctic strata both have a high frequency of B notes, they are distinguished by having different proportions of B2 and B3 units, with INIs in Hawaii being longer than in the Antarctic. Fall and spring bouts in the eastern Gulf of Alaska show similar bout composition, being mainly B2 units and C notes, with a majority of C4 units. Likewise, bouts from Monterey Bay and Bering Sea (both from the fall) have similar distributions, characterized by relatively high proportions of B1, B3, C2, and C3 units.

The proportion and run summaries [pattern proportion, P_r , mean run proportion, $\bar{P}_{R(i)}$, and maximum run proportion, $\text{Max}(P_{R(i)})$] for one-, two-, and three-note patterns (notes alone and units) generated 1,869 summary measures

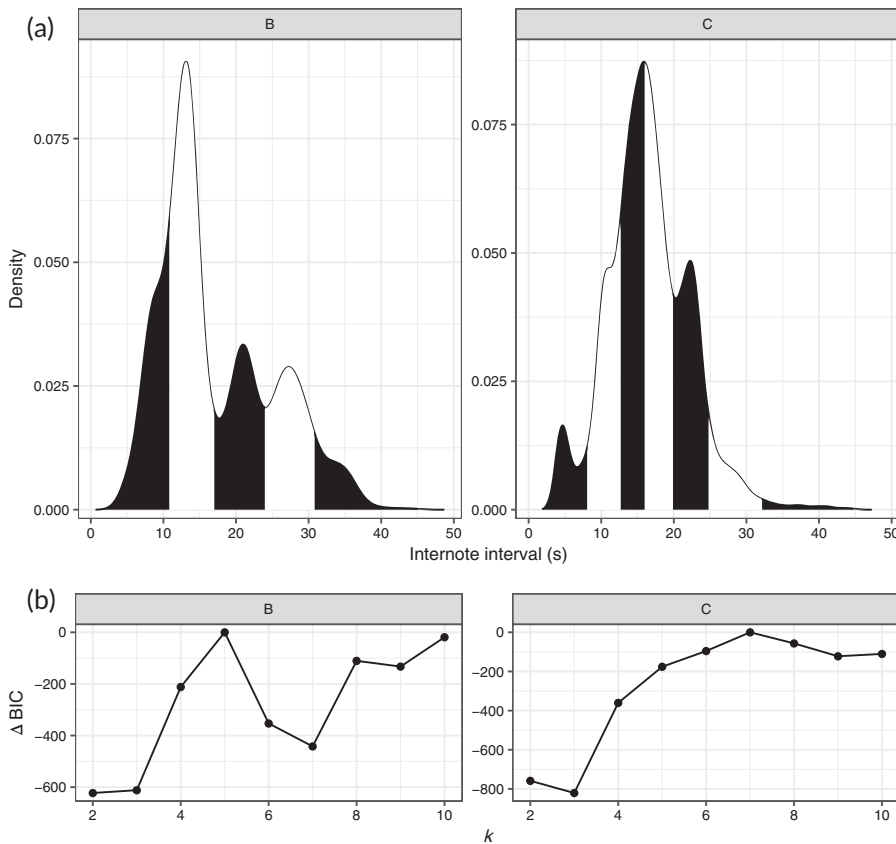


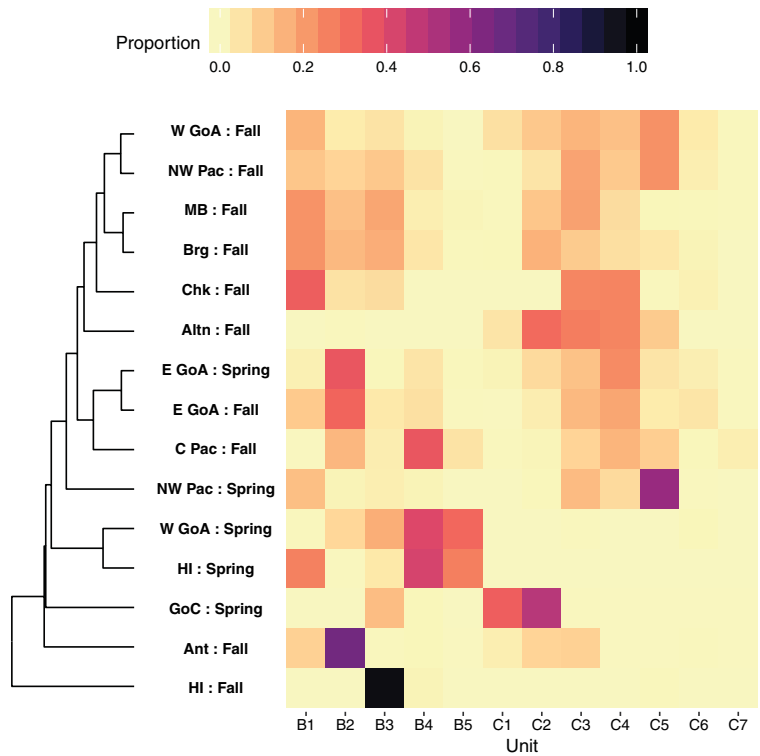
FIGURE 6 Result of Ckmeans.1d.dp clustering of internote intervals (INI) to create units. (a) Distribution of INI following B and C notes, with clusters identified in alternating black and white. (b) Delta-BIC values for $k = 2$ –10 clusters.

for the 473 bouts containing nine or more notes. Clustering based on pairwise distances of bouts within each stratum identified between one and three clusters of bouts per stratum (Figures 8, SM4, and SM5). These varied from clusters representing bouts made entirely of B or C notes and units to those representing combinations of both at different proportions. The Gulf of California spring stratum was unique in being the only one to have bouts that were either primarily B or primarily C. We also note that there were two types of primarily B bouts in the western Gulf of Alaska spring stratum, one having a single but spread INI mode around 15 s and the other with two INI modes at approximately 28 and 34 s. However, the sample sizes of these clusters were relatively small, so these may not be as anomalous as they first appear.

Several clusters had similar INI distributions within strata but were composed of different proportions and patterns of notes. For example, there were two INI modes in each of the three clusters in the Antarctic fall bouts, with means around 12 s and 14 s. In the Chukchi fall bouts, the three clusters had three peaks at approximately 8 s, 15 s, and 18 s, the first two of which could be characterized as BCC triplet, while the third is a BBB triplet. Although more diffuse, the pattern can also be seen among clusters in the Bering fall bouts, with three peaks at approximately 10 s, 15 s, and 20 s.

The dendrogram of representative bouts was composed of three distinct groups (Figure 9). The first comprised bouts composed primarily of Bs. Two of the Antarctic bout types clustered together in this group to the exclusion of the other North Pacific strata. The three Hawaii spring bout types clustered with one of the western Gulf of Alaska

FIGURE 7 Proportion of B and C units in each stratum with hierarchical clustering dendrogram on left indicating strata similarity.



bout types, while the two Hawaii fall bout types clustered with the Gulf of California spring “B” bout type. The Bering and Chuckchi “B” bout types also clustered together. The remainder was a collection of “B” bout types with pairings from the same or nearby regions, but different seasons (e.g., western Gulf of Alaska spring and fall), and very distant regions within the same season (e.g., Monterey Bay and northwest Pacific fall). The second major group was a cluster of bouts containing primarily Cs. Of these, bouts from the northwest Pacific fall and western Gulf of Alaska fall were the most similar. A bout type in the Gulf of California spring was notable in having very short INIs (C1 and C2 units from Figure 6). The third group was composed of bouts with combinations of the two notes. The two Chukchi Sea BCC triplet bout types were differentiated from the rest. A second clade in this group contained three strata with relatively regular B-C repeat units (eastern Gulf of Alaska spring and fall and western Gulf of Alaska fall). Bouts in the third cluster tended to have less clear repeat units, although there was a grouping of Monterey Bay fall and northwest Pacific spring, which seemed to have similar BC doublet bouts.

We also examined variation of sequence cluster assignment within the eastern Gulf of Alaska, off British Columbia, where we had recordings from four separate locations north of Vancouver Island and around Haida Gwaii (Figure 10). In both seasons, the distribution of clusters from Brooks Peninsula was significantly different (χ^2 p -value < .01) from that of locations to the north. In Brooks Peninsula, bouts tended to be regular BC runs. The two regions immediately north (Cape St. James and Caamano Sound) had similar proportions in the fall, with about half being BC repeats and the other half being BB (doublets). These two locations were both significantly different (χ^2 p -value < .01) from Langara Island to the northwest of Haida Gwaii where the four bouts recorded were BB. However, the INI distribution of fall bout cluster 3 from Cape St. James was greater than Caamano Sound and more similar to that of Langara Island (Figure SM6).

3.3 | Strata classification

The Random Forest model classifying bouts to strata had an overall correct classification rate of 53% compared to the 13% expected by chance alone (Table 5). The most distinctive strata were the Aleutians and both seasons in

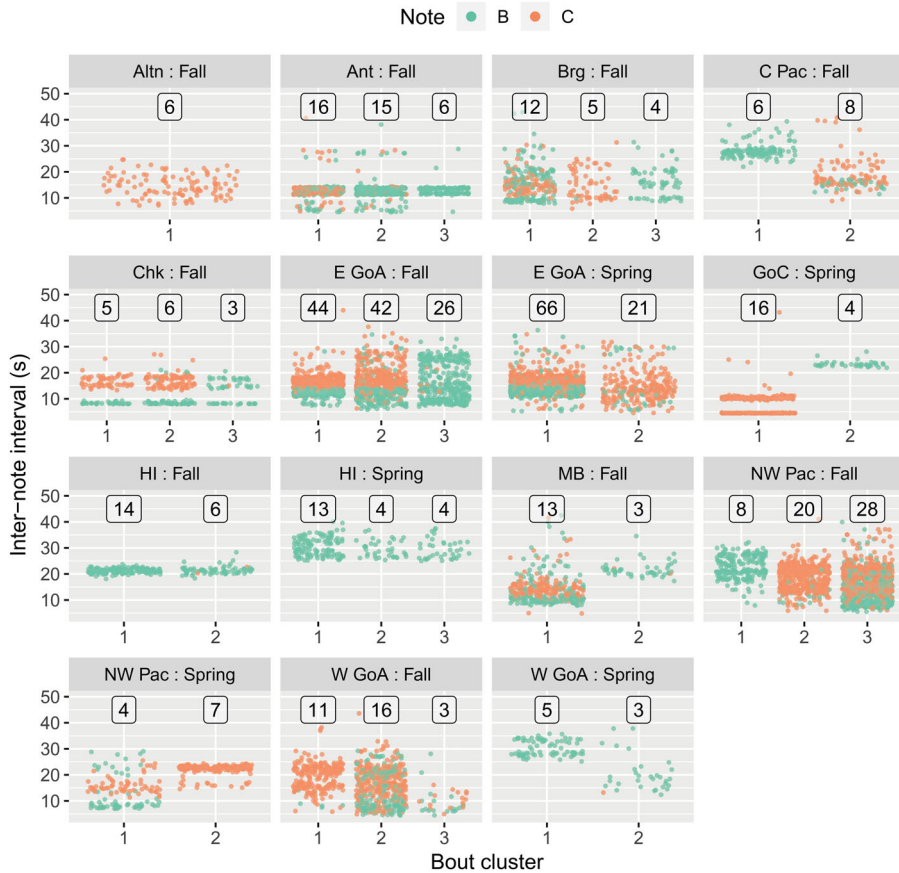


FIGURE 8 Distribution of INI in bout clusters identified in each stratum. Each point is a note from a bout colored by note type (green = B; orange = C). Notes are randomly placed on the x-axis around the vertical line designating their bout cluster to aid in visualization of the distributions. Numbers at the top of each vertical line denote the number of bouts present in that cluster.

Hawaii, both with 100% of the bouts correctly classified. The Antarctic bouts were the next most distinctive with a classification accuracy of 91% and three misclassifications: two to Monterey Bay and one to the Bering Sea. The stratum with the lowest classification accuracy was eastern Gulf of Alaska fall (8%), which had a majority of its bouts misclassified to eastern Gulf of Alaska spring.

4 | DISCUSSION

The impetus for this project was to examine the potential for acoustic monitoring to be used as a proxy for fin whale population or stock definition in the North Pacific. Typically, stock identification for large whales has been based on differences in morphology, genetics, or distribution, data which can be difficult to obtain for pelagic, wide-ranging species such as the fin whale. The use of acoustic differences, particularly song characteristics, has been proposed as one means of identifying stock boundaries and connectivity; however, identification of measure(s) robust enough to be considered can be problematic (e.g., Kershenbaum et al. 2016).

Previous studies of fin whale acoustic behavior have documented several kinds of 20-Hz singlet, doublet, and triplet songs (Koot, 2015; Oleson et al., 2014; Širović et al., 2017; Thompson et al., 1992; Weirathmueller et al., 2017).

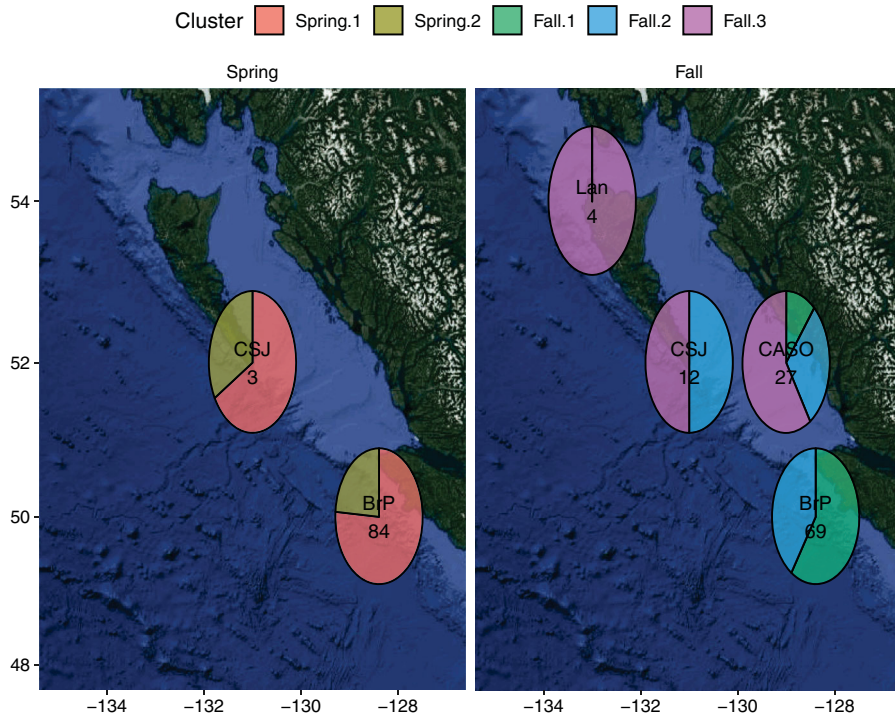


FIGURE 10 Proportion of bout clusters within Eastern Gulf of Alaska by season. BrP = Brooks Peninsula, CSJ = Cape St. James, CASO = Caamano Sound, Lan = Langara Island. Number in circle denotes total number of bouts. Colors denote bout cluster number within each season as shown in Figure 8.

If consistent across years, the presence of these distinctive B-unit patterns could be useful for population assignment. Fin whales are rarely sighted in Hawaiian waters and the true population size and boundaries are unknown, although estimates from summer and fall surveys within the EEZ suggest there are fewer than 200 individuals (Barlow, 2003; Bradford, Forney, Oleson, & Barlow, 2017; Carretta et al., 2018). It is also notable that similar B-unit bouts were observed in the western Gulf of Alaska spring, northwest Pacific fall, central Pacific fall, and Monterey Bay fall, and not observed in the Chukchi, Bering, or Aleutians. This suggests general connectivity among these greater central Pacific regions, also supported by the close similarity of BC patterns in the cluster including in Monterey Bay fall and NW Pacific spring (Figure 9).

Based on the results of genetic and photo ID analyses, there is strong evidence that there is a resident population of fin whales in the Gulf of California (Bérubé et al., 1998; Bérubé, Urban, Dizon, Brownell, & Palsbøll, 2002; Tershy, Urban-Ramirez, Breese, Rojas-Bracho, & Findley, 1993; Urban-Ramirez, Rojas-Bracho, Guerrero-Ruiz, Jaramillo-Legorreta, & Findley, 2005). Širović et al. (2017) describe long and short triplet songs recorded in the Gulf of California that were not found in the southern California Bight. The INI distribution of their short triplet (~5 s and 10 s) matches the INI distribution of the distinctive C sequences in our Gulf of California spring recordings. However, in our recordings this C sequence tended to more closely resemble a doublet pattern than a triplet. The INI distribution of the second cluster of B sequences in our Gulf of California spring recordings (~25 s) match that of the Širović et al. (2017) “long singlet” pattern found in the southern Gulf of California and southern California Bight. In our data, this B singlet was most closely related to sequences from Monterey Bay and northwest Pacific in the fall, suggesting that this pattern is either widespread across populations in the North Pacific, or that there is some connectivity between the central and southern U.S. West Coast and southern Gulf of California and the northwest Pacific.

TABLE 5 Confusion matrix from Random Forest classification model of segments. Rows are observed strata and columns are strata to which the segments were classified in the model.

Strata	Altn:		Ant:		Brg:		C Pac:		Chk:		E GoA:		E GoA:		GoC:		Hi:		MB:		NW Pac:		W GoA:		W GoA:		% Correct	CI	Prior
	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring			
Altn:Fall	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	54-100	1
Ant:Fall	1	34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	92	78-98	8	
Brg:Fall	4	0	9	0	0	3	0	0	2	0	0	2	0	0	0	0	0	0	2	0	0	0	1	0	0	43	22-66	4	
C Pac:Fall	0	0	0	10	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	71	42-92	3	
Chk:Fall	0	0	3	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	79	49-95	3	
E GoA:Fall	4	4	8	9	2	19	45	1	3	0	10	1	5	1	0	0	1	0	0	0	0	1	0	0	0	17	11-25	24	
E GoA:Spring	9	3	3	2	0	7	55	2	0	0	0	2	0	0	0	0	2	0	0	0	0	2	4	0	0	63	52-73	18	
GoC:Spring	0	0	0	0	0	0	0	16	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	80	56-94	4	
Hi:Fall	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	83-100	4	
Hi:Spring	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	84-100	4	
MB:Fall	1	0	0	0	1	0	0	0	3	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	69	41-89	3	
NW Pac:Fall	3	2	4	2	0	0	1	0	5	0	8	17	12	2	0	0	2	0	0	0	0	2	0	0	0	30	19-44	12	
NW Pac:Spring	0	0	0	0	0	0	0	0	0	0	3	0	8	0	0	0	0	0	0	0	0	0	0	0	0	73	39-94	2	
W GoA:Fall	1	0	2	0	0	0	0	1	0	0	4	3	9	10	0	0	0	0	0	0	4	3	9	10	0	33	17-53	6	
W GoA:Spring	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	3	38	9-76	2	
Overall																											53	48-57	13

In the northernmost region of our study, the Chukchi Sea, fin whales produced a BCC triplet pattern in the fall that was not observed elsewhere. There were also three bouts that were composed of BBB triplets with similar INIs to the BCC triplets. These bouts were misclassified in the Random Forest model to the Bering Sea fall, which had four similar but different BBB triplet bouts. Given the similarity of the INI distributions, the BBB and BCC triplets in these regions may be alternate forms of the same call type. It is also notable that the C bouts in the Aleutian fall recordings, with a relatively variable distribution of C INIs, are 100% diagnostic. They are very similar to the C bouts in Bering Sea fall Cluster 2, which were misclassified to the Aleutians. In combination with the distribution of the BBB and BCC triplets, this leads to a suggestion that in the fall there may be two different populations utilizing the Chukchi and central Aleutians that mix to some degree in the southern Bering Sea. This is consistent with observed movements of fin whales from the southern Bering Sea to the eastern Aleutians/Kamchatka as well as northward to the central Bering Sea as revealed by Discovery tag recoveries (Mizroch et al., 2009).

Recordings from the eastern Gulf of Alaska off British Columbia suggest a complex pattern of structure. The most common bout type in this region in both seasons was a BC doublet. The fact that BC doublets in spring and fall in this region were most similar to each other suggests that this song changes little over the year. However, there is also an indication of shifts in the overall composition of fin whales in this area. The C bout type present in the spring recordings (EGoA spring cluster 3) was only found in the Brooks Peninsula. In the fall, this C bout was not observed in Brooks Peninsula. However, a bout type composed primarily of Bs (EGoA fall cluster 3) was found in Langara Island, Cape St. James, and Caamano Sound, but absent from Brooks Peninsula. This bout type is most similar to the BBB triplets from the Chukchi and Bering fall recordings and has a multimodal INI distribution (Figures 8 and 9).

Our recordings from British Columbia are a subset of a larger data set previously analyzed by Koot (2015), which included recordings from other sites including offshore seamounts. In that analysis, two types of BC doublet songs were identified, a "Type 1" song with relatively long INIs (B-C interval > 20 s and C-B interval > 22 s), and a "Type 2" with shorter INIs (B-C interval 8–18 s and C-B interval 14–20 s). The more numerous Type 2 song corresponds well to bout cluster 1 in our EGoA spring and fall data. It is unclear if the Type 1 song is present in our data. In fig. 2–3 of Koot (2015), there is a cluster of three Type 1 songs from Langara Island. In our analysis, Langara Island is represented by four bouts of the primarily B type song with INIs that correspond to those of the Type 1 song presented by Koot (2015). A second cluster of Type 1 songs in Koot (2015) with larger INIs are found further offshore in recordings from Bowie Seamount to the north (which also has Type 2 songs) and Union Seamount to the south (which is only composed of Type 1 songs). The INIs of both groups of Type 1 songs correspond to BC doublets found in the Bering Sea, southern California Bight, and Hawaii by Oleson et al. (2014). Taken together, this suggests that while inshore regions of British Columbia may host a resident population, the offshore areas are being utilized by whales traveling from various distant areas as has also been shown with recoveries of Discovery tags (Mizroch et al., 2009). Further, it suggests that the method used here, which includes notes, units, and bouts, is capable of identifying more fine scale differences in acoustic behavior based on quantitative clustering of the data.

Our inclusion of a few recordings from the Southern Hemisphere for a comparison across ocean basins yielded inconclusive results. There were three bout types identified in the Antarctic recordings, however, the multimodal distribution of INIs was very similar across all three. Two of the bouts were primarily BB doublets, having relatively short intervals overall (10–15 s) and similar, but distinct intervals following each of the two notes. In the first combination, which contains both Bs and Cs, there does not seem to be a consistent pattern of which note precedes which INI class. Thus, given that the intervals were so similar among these three bout clusters, one possibility is that the bouts represent versions of the same song with different notes being substituted.

Koot (2015) found that the Type 2 songs from British Columbia were more similar to Southern Hemisphere songs than were the Type 1 songs. This led to a postulation that Type 2 singers were from a mitogenome clade (Clade C) that originated from a mitogenome haplotype in the Southern Hemisphere approximately 370 KYA (Archer et al., 2013). Although our study found no strong affiliation of Antarctic bouts with identified clusters in other regions in the North Pacific, we note that there is overlap with the lower portion of the INI distribution of bout cluster 3 in the eastern Gulf of Alaska fall recordings. This part of the distribution corresponds to recordings found in

Caamano Sound and are likely to be the same Type 2 recordings that Koot (2015) describe as being similar to Southern Hemisphere song.

Recent studies have shown that North Pacific fin whale song is not constant throughout the year and has also been changing over years. Oleson et al. (2014) describe an increase of BC doublet INIs of approximately 10 s from September to December that occurs synchronously in several regions in the North Pacific, resetting again the following spring. Additionally, Širović et al. (2017) show that this annual change is occurring on a backdrop of a slower but steady yearly increase in INI. Weirathmueller et al. (2017) found decadal changes in song type from singlet to doublet songs as well as intraannual changes similar to those described by Oleson et al. (2014). It is generally accepted that the regular sequences of 20-Hz singlets, doublets, or triplets, which are more frequently produced in the late fall and winter, are produced only by males and have a function in mating (Croll et al., 2002; Oleson et al., 2014; Širović et al., 2013; Watkins, 1981; Watkins et al., 1987). Given that the increase in doublet INIs also synchronously occurs across the North Pacific during this time, it is likely to also be a feature of the mating display (Oleson et al., 2014). Nonregular calls or sequences may be related to other functions such as social cohesion or feeding (Edds, 1988; McDonald, Hildebrand, & Webb, 1995; Moore et al., 1998; Širović et al., 2013).

Our fall (September/October) recordings were taken in months immediately preceding the annual INI increase reported by Oleson et al. (2014) and our spring (February/March) recordings were from months immediately after their peak. Overall, the data used in our study were collected randomly with respect to call type and content across a range of years where available from the regions examined. Because we do not have a single year sampled for all regions/seasons, nor the full range of years sampled for any single region, it is difficult to evaluate what effect these observed seasonal and annual changes have on interpretations of our results. It is possible that at the extreme of these INI shifts, bouts from the same area that are presumably being created by the same whales would be seen as different. In our study, we observe no difference in the INI distribution of one cluster of BC doublet bouts between spring and fall in the eastern Gulf of Alaska recordings, which suggests the possibility that the seasonally synchronous change observed in other parts of the North Pacific (Oleson et al., 2014) may not occur in some subset of these coastal British Columbia whales. Additionally, this shift has only been described in terms of a change in INI, not in pattern. Thus, our metrics that quantify bouts in terms of note pattern composition alone would continue to show similarity among calls throughout the INI shift. Although we did not do it here, it would be possible to compute similarities based on metrics from note patterns alone separate from those based on unit patterns alone separately to gauge the information content of INIs.

Our selection of calls within the same time frame from each year creates an unbiased snapshot of the acoustic repertoire of fin whales in each region. It is this unbiased nature of the data that allows us to identify diagnostic acoustic features that are potentially consistent over time, regardless of their behavioral context. We purposely chose to analyze data from the same months each year to limit potential biases caused by seasonal changes in the whales' behavioral state. Further, we restricted analyses to loud distinct signals to reduce the influence of ambient noise on the measurements. All data were decimated to the same sample rate and the same window size and overlap were used to ensure that all measurements had the same resolution. We did not, however, correct for different instrument type, deployment depth or water depth of the different locations. Instruments in shallow water or bottom-mounted have the potential to record multipath arrivals which can complicate selecting the direct-arrival pulse (McDonald & Fox, 1999; Weirathmueller et al., 2013, 2017), to reduce the likelihood of this, each measured pulse series was examined for the presence of multipath arrivals.

Another finding that this study highlights is the importance of developing an automated classification scheme for unique note types based on their spectral characteristics. This proved to be a critical component as there was considerable variation in the human analyst's determination of note types across regions with respect to the clustering of their spectral features. In some regions, like Hawaii, where Bs predominated, the analyst believed that a majority of notes were Cs. This may reflect preconceived expectations about the overall frequency of Bs in fin whale calls in the face of sequences of very similar repeating notes. This is supported by the observation that a majority of the

discrepancies in all regions were of analysts labelling notes as Cs that clustered as Bs. This can have an important effect on interpretation of patterns of variability among regions. For example, in the eastern Gulf of Alaska, where both note types occur at more equal frequencies, the discrepancies were more even. However, we still see a difference in the types of eastern Gulf of Alaska bouts present in each season depending on if the notes were manually designated or designated based on Density Clustering. The "Type 1" BC doublet reported by Koot (2015) may actually be a much more distinct BB doublet that would suggest greater degree of differentiation among whales found around British Columbia.

Although we have avoided characterizing the Density Clustering note determination as "correct," the evidence from the Random Forest model is that this process produces a more consistent determination than the analyst with little error (0.7%). Given that this model was based on a relatively large number of notes ($n = 10,486$), the results are expected to be robust. Thus, we strongly advocate that in future studies notes be classified in an objective and repeatable manner. Use of the model we developed in this study would make analyses of new data more consistent and comparable with those here.

The goal of this study was to produce a robust methodology for measuring and identifying the temporal structure of fin whale sequences to examine population identity in fin whales, but information on the timing of these acoustic units in sequences may eventually be used to provide insights including identity or fitness of individuals (Kershenbaum et al., 2016). This analytical approach may not be as appropriate for studies of species where calls and/or acoustic sequences are more complex, and biologically meaningful information is codified in the calls themselves rather than in the time intervals, where the drivers of variability are somewhat better understood, such as humpback whales (Garland et al., 2013; Rekdahl et al., 2018). In these species, other methods such as photo-identification of individual animals for mark-recapture population estimates and migratory patterns (Barlow et al., 2011; Smith et al., 1999; Stevick et al., 2003) and genetic data for migratory patterns and feeding ground philopatry (Baker et al., 1986, 2013) provide useful information on structure and patterns of connectivity.

In summary, we have shown that our full characterization approach is a useful and informative way to standardize comparisons of fin whale bouts across multiple regions and seasons. This analysis also suggests some hypotheses of population structure and connectivity across the north Pacific that can be further tested with genetic data, or extended acoustic data sets. Finally, the degree of variability in the calls in this study suggests that future biopsy efforts should strive to collect acoustic recordings at the same time. Building a paired set of the two data streams will greatly enhance interpretation of currently available acoustic and genetic data and likely lead to more informative new mixed analyses.

ACKNOWLEDGMENTS

The authors wish to thank Barbara Koot, Michael Richlen, Benjamin Jones, Jessica Thompson, John Ford, Rob Williams, and Janell Morano for providing data and/or analytical assistance. Bruce Martin and Briand Gaudet (JASCO Applied Sciences, Canada) provided us with the Spectroplotter software and were an incredible help with modifications to suit our measurements needs. We are very grateful to our external collaborators for providing data for this project: Catherine Berchok at Marine Mammal Lab at the Alaska Fisheries Science Center provided data from the Bering Sea; Sharon Nieu Kirk from Oregon State University provided data from the Gulf of Alaska; Chris Clark and Aaron Rice at the Bioacoustics Research Program at Cornell University provided recordings from Monterey Bay and Gulf of California; Bruce Howe from the University of Hawai'i provided data from the Aloha Cabled Observatory. The Chukchi Sea data were collected during a multiyear acoustic monitoring program funded by Shell Exploration and Production Company, ConocoPhillips, and Statoil. Data from the Pioneer Seamount were funded by NOAA's Ocean Exploration and VENTS programs and supported by the Pioneer Seamount Acoustic Observatory Program. Data from the northwestern Pacific are from the NOAA Pacific Marine Environmental Laboratory VENTS program at Newport, Oregon.

ORCID

Frederick I. Archer  <https://orcid.org/0000-0002-3179-4769>

REFERENCES

- Archer, F. I., Martien, K. K., & Taylor, B. L. (2017). Diagnosability of mtDNA with Random Forests: Using sequence data to delimit subspecies. *Marine Mammal Science*, 33(S1), 101–131.
- Archer, F. I., Morin, P. A., Hancock-Hanser, B. L., Robertson, K. M., Leslie, M. S., Bérubé, M., ... Taylor, B. L. (2013). Mitogenomic phylogenetics of fin whales (*Balaenoptera physalus* spp.): Genetic evidence for revision of subspecies. *PLoS ONE*, 8(5), e63396.
- Baker, C. S., Herman, L. M., Perry, A. L., Lawton, W. S., Straley, J. M., Wolman, A. A., ... Östman, J. (1986). Migratory movement and population structure of humpback whales (*Megaptera novaeangliae*) in the central and eastern North Pacific. *Marine Ecology Progress Series*, 31, 105–119.
- Baker, C. S., Steel, D., Calambokidis, J., Falcone, E., González-Peral, U., Barlow, J., ... Yamaguchi, M. (2013). Strong maternal fidelity and natal philopatry shape genetic structure in North Pacific humpback whales. *Marine Ecology Progress Series*, 494, 291–306.
- Barlow, J. (2003). *Cetacean abundance in Hawaiian waters during summer/fall of 2002* (Administrative Report LJ-03-13). La Jolla, CA: Southwest Fisheries Science Center.
- Barlow, J., Calambokidis, J., Falcone, E., Baker, C. S., Burdin, A., Clapham, P. J., ... Witteveen, B. (2011). Humpback whale abundance in the North Pacific estimated by photographic capture-recapture with bias correction from simulation studies. *Marine Mammal Science*, 27, 793–818.
- Berk, R. (2006). An introduction to ensemble methods for data analysis. *Sociological Methods and Research*, 34, 263–295.
- Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Di Sciara, G. N., Sears, R., ... Palsbøll, P. J. (1998). Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus 1758): Analysis of mitochondrial and nuclear loci. *Molecular Ecology*, 7, 585–599.
- Bérubé, M., Urban, J., Dizon, A. E., Brownell, R. L., & Palsbøll, P. J. (2002). Genetic identification of a small and highly isolated population of fin whales (*Balaenoptera physalus*) in the Sea of Cortez, Mexico. *Conservation Genetics*, 3, 183–190.
- Bradford, A. L., Forney, K. A., Oleson, E. M., & Barlow, J. (2017). Abundance estimates of cetaceans from a line-transect survey within the U.S. Hawaiian Islands Exclusive Economic Zone. *Fishery Bulletin*, 115, 129–142.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Buchan, S. J., Stafford, K. M., & Hucke-Gaete, R. (2015). Seasonal occurrence of southeast Pacific blue whale songs in southern Chile and the eastern tropical Pacific. *Marine Mammal Science*, 31, 440–458.
- Carretta, J. V., Forney, K. A., Oleson, E., Weller, D. W., Lang, A. R., Baker, J., ... Brownell, R. L. J., Jr. (2018). *U.S. Pacific marine mammal stock assessments: 2017* (NOAA Technical Memorandum NMFS-SWFSC-602). Washington, DC: U.S. Department of Commerce.
- Castellote, M., Clark, C. W., & Lammers, M. O. (2012). Acoustic compensation to shipping and air gun noise by Mediterranean fin whales (*Balaenoptera physalus*). *Advances in Experimental Medicine and Biology*, 730, 321.
- Cerchio, S., Jacobsen, J. K., & Norris, T. F. (2001). Temporal and geographical variation in songs of humpback whales, *Megaptera novaeangliae*: Synchronous change in Hawaiian and Mexican breeding assemblages. *Animal Behaviour*, 62, 313–329.
- Clark, C. W. (1990). Acoustic behavior of mysticete whales. In J. Thomas & R. A. Kastelein (Eds.), *Sensory Abilities of Cetaceans* (pp. 571–583). New York, NY: Plenum Press.
- Clark, C. W., Borsani, F., & Notarbartolo di Sciara, G. (2002). Vocal activity of fin whales, *Balaenoptera physalus*, in the Ligurian Sea. *Marine Mammal Science*, 18, 281–285.
- Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B. R., Flores, S., Gedamke, J., & Urban, J. (2002). Only male fin whales sing loud songs. *Nature*, 417, 809.
- Delarue, J., Martin, B., Hannay, D., & Berchok, C. L. (2013). Acoustic occurrence and affiliation of fin whales detected in the northeastern Chukchi Sea, July to October 2007–10. *Arctic*, 66, 159–172.
- Delarue, J., Todd, S. K., Van Parijs, S. M., & Di Iorio, L. (2009). Geographic variation in Northwest Atlantic fin whale (*Balaenoptera physalus*) song: Implications for stock structure assessment. *Journal of the Acoustical Society of America*, 125, 1774–1782.
- Edds, P. L. (1988). Characteristics of finback *Balaenoptera physalus* vocalizations in the St. Lawrence Estuary. *Bioacoustics*, 1, 131–149.
- Garland, E. C., Gedamke, J., Rekdahl, M. L., Noad, M. J., Garrigue, C., & Gales, N. (2013). Humpback whale song on the Southern Ocean feeding grounds: Implications for cultural transmission. *PLoS ONE*, 8(11), e79442.

- Hatch, L. T. (2004). *Male genes and male song: Integrating genetic and acoustic data in defining fin whale, Balaenoptera physalus, management units* (Doctoral dissertation). Cornell University, Ithaca, NY.
- Kershenbaum, A., Blumstein, D. T., Roch, M. A., Akçay, Ç., Backus, G., Bee, M. A., ... Zamora-Gutierrez, V. (2016). Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews*, 91, 13–52.
- Koot, B. (2015). *Winter behaviour and population structure of fin whales (Balaenoptera physalus) in British Columbia inferred from passive acoustic data* (Masters thesis). University of British Columbia, Vancouver, Canada.
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, 2(3), 18–22.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2018). cluster: Cluster analysis basics and extensions (R package version 2.0.7-1).
- McDonald, M. A., Hildebrand, J. A., & Webb, S. C. (1995). Blue and fin whales observed on a seafloor array in the Northeast Pacific. *Journal of the Acoustical Society of America*, 98, 712–721.
- McDonald, M. A., Mesnick, S. L., & Hildebrand, J. A. (2006). Biogeographic characterisation of blue whale song worldwide: Using song to identify populations. *Journal of Cetacean Research and Management*, 8(1), 55–65.
- Mellinger, D. K., & Clark, C. W. (2003). Blue whale (*Balaenoptera musculus*) sounds from the North Atlantic. *Journal of the Acoustical Society of America*, 114, 1108–1119.
- Mizroch, S. A., Rice, D. W., Zwiefelhofer, D., Waite, J., & Perryman, W. L. (2009). Distribution and movements of fin whales in the North Pacific Ocean. *Mammal Review*, 39, 193–227.
- Moore, S. E., Stafford, K. M., Dalheim, M., Fox, C., Braham, H. W., Polovina, J., & Bain, D. (1998). Seasonal variation in reception of fin whale calls at five geographic areas in the North Pacific. *Marine Mammal Science*, 14, 617–627.
- Morano, J. L., Salisbury, D. P., Rice, A. N., Conklin, K. L., Falk, K. L., & Clark, C. W. (2012). Seasonal and geographical patterns of fin whale song in the western North Atlantic Ocean. *Journal of the Acoustical Society of America*, 132, 1207–1212.
- Nigenda-Morales, S., Flores-Ramirez, S., Urban-R, J., & Vazquez-Juarez, R. (2008). MHC DQB-1 polymorphism in the Gulf of California fin whale (*Balaenoptera physalus*) population. *Journal of Heredity*, 99, 14–21.
- Oleson, E., Barlow, J., Gordon, J., Rankin, S., & Hildebrand, J. (2003). Low frequency calls of Bryde's whales. *Marine Mammal Science*, 19, 407–419.
- Oleson, E. M., Širović, A., Bayless, A. R., & Hildebrand, J. A. (2014). Synchronous seasonal change in fin whale song in the North Pacific. *PLoS ONE*, 9(12), e115678.
- Payne, K., & Payne, R. (1985). Large-scale changes over 19 years in songs of humpback whales in Bermuda. *Zeitschrift für Tierpsychologie*, 68, 89–114.
- Pedersen, T. L., Hughes, S., & Qiu, X. (2017). *densityClust: Clustering by fast search and find of density peaks* (R package version 0.3). Retrieved from <https://github.com/thomasp85/densityClust>
- R Core Team (2019). *R: A language and environment for statistical computing* (Version 3.5.3). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Rankin, S., Castellote, M., Delarue, J., Stafford, K., Archer, F., Koot, B., ... Morano, J. L. (2018). *Methods for characterizing fin whale song notes for comparative studies of geographic variation in song* (NOAA Technical Memorandum NMFS-SWFSC-592). Washington, DC: U.S. Department of Commerce.
- Rekdahl, M. L., Garland, E. C., Carvajal, G. A., King, C. D., Collins, T., Razafindrakoto, Y., & Rosenbaum, H. (2018). Culturally transmitted song exchange between humpback whales (*Megaptera novaeangliae*) in the southeast Atlantic and southwest Indian Ocean basins. *Royal Society Open Science*, 5(11). Retrieved from <https://doi.org/10.1098/rsos.172305>
- Reynolds, A. P., Richards, G., de la Iglesia, B., & Rayward-Smith, V. J. (2006). Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, 5, 475–504.
- Rodriguez, A., & Laio, A. (2014). Clustering by fast search and find of density peaks. *Science*, 344, 1492–1496.
- Samaran, F., Stafford, K. M., Branch, T. A., Gedamke, J., Royer, J.-Y., Dziak, R. P., & Guinet, C. (2013). Seasonal and geographic variation of southern blue whale subspecies in the Indian Ocean. *PLoS ONE*, 8(8), e71561.
- Silber, G. K., Newcomer, M. W., Silber, P. C., Pérez-Cortés, M. H., & Ellis, G. M. (1994). Cetaceans of the northern Gulf of California: Distribution, occurrence, and relative abundance. *Marine Mammal Science*, 10, 283–298.
- Simon, M., Stafford, K. M., Beedholm, K., Lee, C. M., & Madsen, P. T. (2010). Singing behavior of fin whales in the Davis Strait with implications for mating, migration and foraging. *Journal of the Acoustical Society of America*, 128, 3200–3210.
- Širović, A., Hildebrand, J. A., Wiggins, S. M., & Thiele, D. (2009). Blue and fin whale acoustic presence around Antarctica during 2003 and 2004. *Marine Mammal Science*, 25(1), 125–136.
- Širović, A., Oleson, E. M., Buccowich, J., Rice, A., & Bayless, A. R. (2017). Fin whale song variability in southern California and the Gulf of California. *Scientific Reports*, 7(1), 10126.
- Širović, A., Rice, A., Chou, E., Hildebrand, J. A., Wiggins, S. M., & Roch, M. A. (2015). Seven years of blue and fin whale call abundance in the Southern California Bight. *Endangered Species Research*, 28, 61–76.
- Širović, A., Williams, L. N., Kerosky, S. M., Wiggins, S. M., & Hildebrand, J. A. (2013). Temporal separation of two fin whale call types across the eastern North Pacific. *Marine Biology*, 160, 47–57.
- Smith, T. D., Allen, J., Clapham, P. J., Hammand, P. S., Katona, S., Larsen, F., ... Øien, N. (1999). An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Marine Mammal Science*, 15, 1–32.

- Stafford, K. M., Chapp, E., Bohnenstiel, D. R., & Tolstoy, M. (2011). Seasonal detection of three types of “pygmy” blue whale calls into the Indian Ocean. *Marine Mammal Science*, 27, 828–840.
- Stafford, K. M., Nieukirk, S. L., & Fox, C. (1999). Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific. *Journal of the Acoustical Society of America*, 106, 3687–3698.
- Stafford, K. M., Nieukirk, S. L., & Fox, C. (2001). Geographic and seasonal variation of blue whale calls in the North Pacific. *Journal of Cetacean Research and Management*, 3, 65–76.
- Stevick, P. T., Allen, J., Clapham, P. J., Friday, N., Katona, S. K., Larsen, F., ... Hammond, P. S. (2003). North Atlantic humpback whale abundance and rate of increase four decades after protection from whaling. *Marine Ecology Progress Series*, 258, 263–273.
- Tershy, B. R., Urban-Ramirez, J., Breese, D., Rojas-Bracho, L., & Findley, L. T. (1993). Are fin whales resident to the Gulf of California? *Revista de Investigacion Cientifica*, 1, 69–72.
- Thompson, P. O., Findley, L. T., & Vidal, O. (1992). 20-Hz pulses and other vocalizations of fin whales, *Balaenoptera physalus*, in the Gulf of California, Mexico. *Journal of the Acoustical Society of America*, 92, 3051–3057.
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2), 411–423.
- Urban-Ramirez, J., Rojas-Bracho, L., Guerrero-Ruiz, M., Jaramillo-Legorreta, A., & Findley, L. T. (2005). Cetacean diversity and conservation in the Gulf of California. In J. L. Cartron, G. Ceballos, & R. Felger (Eds.), *Biodiversity, ecosystems, and conservation in northern Mexico* (Vol. 1, pp. 276–297). New York, NY: Oxford University Press.
- Wang, H., & Song, M. (2011). Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. *R News*, 3(2), 29–33.
- Watkins, W. A. (1981). Activities and underwater sounds of fin whales. *Scientific Reports of the Whales Research Institute, Tokyo*, 33, 83–117.
- Watkins, W. A., Tyack, P., Moore, K. E., & Bird, J. E. (1987). The 20-Hz signals of finback whales (*Balaenoptera physalus*). *Journal of the Acoustical Society of America*, 82, 1901–1912.
- Weirathmueller, M. J., Stafford, K. M., Wilcock, W. S. D., Hilmo, R. S., Dziak, R. P., & Tréhu, A. M. (2017). Spatial and temporal trends in fin whale vocalizations recorded in the NE Pacific Ocean between 2003–2013. *PLoS ONE*, 12(10), e0186127.
- Winn, H. E., & Winn, L. (1978). The song of the humpback whale (*Megaptera novaeangliae*) in the West Indies. *Marine Biology*, 47, 97–114.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Archer FI, Rankin S, Stafford KM, Castellote M, Delarue J. Quantifying spatial and temporal variation of North Pacific fin whale (*Balaenoptera physalus*) acoustic behavior. *Mar Mam Sci*. 2020; 36:224–245. <https://doi.org/10.1111/mms.12640>