Investigating Group Behavioral Quantization of Oceanic Fish with Continental-shelf Scale Ocean-acoustic Sensing

by

Dong Hoon Yi

B.S., Pohang University of Science and Technology, Korea (2004)
M.S., Pohang University of Science and Technology, Korea (2006)

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2018

© Massachusetts Institute of Technology 2018. All rights reserved.

Signature redacted

Author .................

Department of Mechanical Engineering January 15, 2018

Signature redacted

Certified by ..............

Nicholas C. Makris
Professor of Mechanical Engineering Thesis Supervisor

Signature redacted

Accepted by ........................................ ........

Rohan Abeyaratne
Chairman, Department Committee on Graduate Theses
Investigating Group Behavioral Quantization of Oceanic Fish with Continental-shelf Scale Ocean-acoustic Sensing

by

Dong Hoon Yi

Submitted to the Department of Mechanical Engineering on January 15, 2018, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Abstract

The recent fish population decline due to increased human presence has led to calls for predictive methods to help reverse or stabilize the situation. It has been difficult, however, to establish such methods primarily due to the technical obstacles in observing fish populations in natural habitats. Here we use acoustics to observe the ocean environment and study fish behavior during the critical spawning period over continental-shelf scales. Fish are known to be one of the main sources of strong natural returns in the continental-shelf environment, and so identified as a major source of clutter for wide-area undersea surveillance. The first continental-shelf scale acoustic measurements of Atlantic cod over thousands of square kilometers using towed source and receiver arrays were made by an international, multi-disciplinary team led by MIT researchers including myself in the historic Lofoten cod spawning ground in Norway during the peak spawning period in Winter 2014, where extensive but spatially discrete groups of spawning cod were successfully imaged. These initial instantaneous wide-area observations of cod aggregations suggest that these observed spawning groups have quantifiable properties that are linked to essential collective behavioral functions. We find that the mean group population per annual spawning season of Northeast Arctic cod over the entire spawning ground in Lofoten Norway is remarkably invariant across the available 30 years of line-transect survey data. The marked stability of the annual mean spawning group size in contrast to the large variations in total spawning population across years supports the interpretation of the expected spawning group size over the 30-year data set as the group behavioral quantum empirically expected for reliable spawning. Time series of the total Atlantic cod spawning population for major spawning regions across the North Atlantic show that once the total spawning population declined below a quantum, recovery to pre-industrial levels did not occur in that region even after decades, which is an apparent consequence of large difference between the pre-industrial level and one quantum level. Quantized group behavior during spawning is also investigated for the Atlantic herring species. We find that the daily herring spawning group population is stable over the peak annual spawning period from wide-area acoustic measurements of spawning
herring in the Gulf of Maine in Fall 2006. This supports the quantum concept that the mean spawning group population has evolved to a stable optimal size to fulfill the essential behavioral function of reliable spawning for Atlantic herring. As with cod, time series of the Atlantic herring spawning population for major spawning grounds across the North Atlantic show that when total spawning population declined below the empirically determined quantum level, return to pre-industrial levels required decades. Our findings show that to be sustained at pre-industrial levels the total spawning population must greatly exceed the mean spawning group size found at pre-industrial levels for any oceanic fish population we investigated, and likely many others.

The migration of extensive social groups towards specific spawning grounds in vast and diverse ocean environments is an integral part of the regular spawning process of many oceanic fish species. Oceanic fish in such migrations typically seek locations with environmental parameters that maximize the probability of successful spawning and egg/larval survival. The 3D spatio-temporal dynamics of these behavioral processes are largely unknown due to technical difficulties in sensing the ocean environment over wide areas. Here we use ocean acoustic waveguide remote sensing (OAWRS) to instantaneously image immense herring groups over continental-shelf-scale areas at the Georges Bank spawning ground. Via multi-spectral OAWRS measurements, we capture a shift in swimbladder resonance peak correlated with the herring groups’ up-slope spawning migration, enabling 3D spatial behavioral dynamics to be instantaneously inferred over thousands of square kilometers. We show that herring groups maintain near-bottom vertical distributions with negative buoyancy throughout the migration. We find a spatial correlation greater than 0.9 between the average herring group depth and corresponding seafloor depth for migratory paths along the bathymetric gradient. This is consistent with herring groups maintaining near-seafloor paths to both search for optimal spawning conditions and reduce the risk of predator attacks during the migration to shallower waters where near-surface predators are more dangerous. This analysis shows that multi-spectral resonance sensing with OAWRS can be used as an effective tool to instantaneously image and continuously monitor the behavioral dynamics of swimbladder-bearing fish group behavior in 3 spatial dimensions over continental-shelf scales.

Recent research has found a high spatial and temporal correlation between certain baleen whale vocalizations and peak annual spawning processes of Atlantic herring in the Gulf of Maine. These vocalizations are apparently related to feeding activities of baleen whales with suggested functions that include communication, prey manipulation, and echolocation. Here the feasibility of the echolocation function is investigated. Physical limitations on the ability to detect large herring shoals and the seafloor by acoustic remote sensing are determined with ocean acoustic propagation, scattering, and statistical theories given baleen whale auditory parameters. Detection is found to be highly dependent on ambient noise conditions, herring shoal distributions, baleen whale time-frequency vocalization spectra, and geophysical parameters of the ocean waveguide. Detections of large herring shoals are found to be physically feasible in common Gulf of Maine herring spawning scenarios at up to $10\pm6$ km in range for humpback parameters and $1\pm1$ km for minke parameters but not for blue
and fin parameters even at zero horizontal range. Detections of the seafloor are found to be feasible up to 2±1 km for blue and humpback parameters and roughly 1 km for fin and minke parameters, suggesting that the whales share a common acoustic sensation of rudimentary features of the geophysical environment. No effect of anthropogenic sound on marine mammal vocalization behavior was found during our measurements. Some published statistical tests assessing the impact of anthropogenic sound on marine mammal behavior were found to have 98-100% false positive biases with no true positive confirmation, and so lack statistical significance.

Thesis Supervisor: Nicholas C. Makris
Title: Professor of Mechanical Engineering
3.5 Conclusions ....................................................... 71

4 Feasibility of Acoustic Remote Sensing of Large Herring Shoals and Seafloor by Baleen Whales 73
4.1 Introduction ...................................................... 73
4.2 Materials and Methods ........................................... 74
  4.2.1 Acoustic Parameters Relevant to Potential Active Acoustic Sensing in Baleen Whales and Corresponding Spatial Resolution .. 74
  4.2.2 Detection of Scattered Returns from Herring Shoals and the Seafloor ........................................... 75
  4.2.3 OAWRS Experiment during Peak Herring Spawning Processes in the Gulf of Maine in Fall 2006 .......................... 80
4.3 Results ......................................................... 80
4.4 Discussion ...................................................... 82
4.5 Conclusions ...................................................... 82

5 Conclusions ...................................................... 91

A Supplementary Information for Group Behavioral Quantization of Oceanic Fish 95
A.1 Acoustic sensing methods ........................................ 95
A.2 Cod areal population density obtained by OAWRS .............. 96
A.3 Critical boundary population density of cod spawning groups .... 97
A.4 Consistency between OAWRS and line-transect vertical echosounder measurement ........................................... 100
A.5 Cod areal population density obtained by line-transect vertical echosounding ........................................... 101
A.6 Mean spawning group size for Northeast Arctic cod ............. 102
A.7 Probability density of cod spawning-group size ................ 103
A.8 Mean spawning group size for Georges Bank herring ............. 104
A.9 OAWRS cod population density mosaic near Røst ................ 105
E.3 Fristrup, Hatch, and Clark statistical method to detect deviation from
null hypothesis ........................................ 169

E.4 False positive rate of Fristrup, Hatch, and Clark statistical method . 170

E.5 Conclusions ........................................ 172

F Preliminary analysis on the herring shoal population in Ålesund,
Norway ............................. 173
List of Figures

2-1 Spatial population distributions of entire gigantic cod groups of the Northeast Arctic obtained by instantaneous continental-shelf scale imaging (A) Primary Northeast Arctic spawning ground of Atlantic cod (*Gadus morhua*), Lofoten, Norway. Red circles indicate OAWRS imaging coverage in 50 s. Regions where zoomed OAWRS cod population density image appear in (B) and (C) are indicated by purple rectangles. (B) Vast 40-km-diameter temporally stable cod group containing roughly 40 million individuals (152,000 tons) south of Røst Island comprising 17% of the estimated Lofoten cod spawning population. (C) Temporally transient lineated 20-km-scale cod group containing roughly 1 million individuals (3,800 tons) near Andenes, Norway. Dashed white lines indicate water depth contours. Coordinate origins are at OAWRS Research Vessel Knorr locations: (B) 67.3 °N, 11.5 °E on February 23, 2014, 10:40:49 UTC; and (C) 69.0 °N, 14.3 °E on March 8, 2014, 00:16:49 UTC.
Time series of mean spawning group size for Northeast Arctic cod and regional cod spawning populations across the North Atlantic (A). Cod population density along conventional line-transect surveys in Lofoten Norway with downward directed echosounding at 1-nmi resolution (Supplementary Information). Discrete cod spawning groups are detected (black contours) and enumerated using group structural information obtained by OAWRS in decades of line-transect survey data (Supplementary Information). Example years of relatively high (2014) and relatively low (1999) cod abundance are shown. (B) Mean annual spawning group population over the Northeast Arctic spawning for 30 years of line-transect survey. Stability of the 30-year average of 10.6 million to within 36% (blue line with error bar showing 30-year standard deviation of the mean) suggests it serves as a mean behavioral unit or mean quantum for effective spawning. (C) Time series of total cod population at major spawning grounds in the North Atlantic during spawning show large variations over the last century. Time series for a total spawning population are shown in red from the time they decline below a cod quantum to the time they return to pre-industrial levels. The cod quantum is the mean spawning group size (10.6 million) found for Atlantic cod when the total population was at pre-industrial levels, where the only available data is from the Northeast Arctic. Gray and blue shaded areas indicate estimated spawning population levels in indicated regions before the onset of industrial fishing in the early 20th century (Supplementary Information).
Time series of mean spawning group size for Georges Bank herring and regional herring spawning populations across the North Atlantic (A) Location of the daily spawning herring group on the northern flank of Georges Bank during the 8-day peak spawning period in the Fall of 2006 determined by OAWRS imagery. Group boundaries occur at the group formation critical density of 0.2 fish/m² for herring [1]. (B) Daily spawning group population over the peak spawning period obtained from instantaneous OAWRS images of each group. Spawning group population has a mean of 204 million to within 35% (blue line and error bar indicate the standard deviation in group size) over the peak spawning period in 2006. (C) Time series of total herring population at major spawning grounds in the North Atlantic during spawning show large variations over the last century. Time series for a total spawning population are shown in red from the time they decline below a herring quantum to the time they return to pre-industrial levels. The herring quantum is the mean spawning group size (204 million) found for Atlantic herring when the total population was at pre-industrial levels, where the only available data is from Georges Bank. Gray and blue shaded areas are estimated spawning population levels in indicated regions before the onset of industrial fishing in the early 20th century (Supplementary Information). The OAWRS sensor location on October 4, 2006, 00:16:15 EDT, at 42.2°N, 67.7°W is the coordinate origin in (A).

Location of the OAWRS Gulf of Maine 2006 experiment. Yellow dashed circle shows 100-km diameter OAWRS areal coverage in 75 seconds. The white rectangular box represents the area investigated here. The gray scale shows the variation in the bathymetry in the region.
3-2 Spatial shoal segmentation pattern. The segmentation pattern is swept by a unit annular sector covering 4 x 4 grid area for analysis. A unit annular sector contains roughly 270 independent measurements.

3-3 (A) Highest measured scattering strength for each spatial position on the northern flank of Georges Bank during shoal formation and subsequent migration towards spawning grounds monitored from 17:25:00 to 21:37:30 EDT on September 29, 2006. Herring begin shoal formation process at locations with seafloor depth of roughly 140 m to 180 m at 17:25:00 EDT, further develop along the bathymetric contour and into the shallower waters, and then stay mostly at locations shallower than 140 m at 21:37:30 EDT. (B) Time evolution of high herring scattering area from 17:25:00 to 21:37:30 EDT on September 29, 2006. The OAWRS sensor location at 21:37:30 EDT on September 29, 2006 is the coordinate origin.

3-4 (A) Spatial distribution of frequency at which the strongest acoustic scattering occurred within vast herring groups imaged from 17:25:00 to 21:37:30 EDT on September 29, 2006. Gradual decrease in the strongest scattering frequency with decreasing seafloor depths indicates a consistent shift in the resonance frequency of herring scattering during the spawning migration towards the spawning grounds on Georges Bank. (B) The spatial distribution of the resonance peak frequency within vast herring groups imaged from 17:25:00 to 21:37:30 EDT on September 29, 2006.
3-5 The resonance peak of herring scattering consistently shifts from frequencies higher than 1125 Hz to lower than 735 Hz with decreasing seafloor depths. (A) A typical frequency roll-off of roughly 20 dB/octave with decreasing frequencies from the resonance peak frequency at roughly 1500 Hz is found for spatial positions with seafloor depths deeper than 115 m. (B-D) The resonance peaks are found within the measured OAWRS frequency range from 1125 Hz to 415 Hz for the positions with shallower seafloor depths.

3-6 The vertical distribution of herring during the spawning migration and neutral buoyancy as a function of relative distance from 140-m isobath contour along the migration path from 17:25:00 to 21:37:30 EDT on September 29, 2006. Herring are found to occupy roughly the bottom half of the water column with negative buoyancy throughout the migration. The black and red vertical lines represent the spatial variations in the mean shoal depth and neutral buoyancy depth, respectively. The blue color-filled boxes represent the vertical herring distribution for each shoal segment.

3-7 (A) 3D surface of mean herring group depth determined from multifrequency measurements taken from 17:25:00 to 21:37:30 EDT on September 29, 2006. (B) Seafloor depth of the sloping bottom on the northern flank of Georges Bank. The mean herring group depth and the seafloor depth are found to be highly correlated (0.9 correlation coefficient) for migratory paths along the bathymetric gradient. (C) The horizontal distribution of mean herring group depth. (D) Estimation uncertainty for mean herring group depth in percentage determined by Eq. (3.13). The rest of the estimated herring group parameters and corresponding estimation uncertainties are shown in Fig. B-2.

4-1 Beam patterns of two-element coherent array steered at broadside given baleen whale acoustic parameters given in Table 4.1.
4-2 Top view (A) and side view (B) of the geometry for possible acoustic remote sensing in a continental shelf environment by baleen whales. The resolution footprint given baleen whale acoustic parameters is an annular sector with range extent \( dp \) and cross-range extent \( \rho d\theta \).

4-3 (A) Migration of herring shoals from depths greater than 100 m to less than 50 m in the Gulf of Maine; (B) Corresponding frequency dependence of resonance scattering of herring swimbladders for deep and shallow shoals. Blue and red solid lines represent the herring target strength when herring layer is at depths \( z_{herring} \) of 130-180 m with a neutral buoyancy depth \( z_{nb} \) of 83 m on the northern flank of Georges Bank and at depths \( z_{herring} \) of 15-35 m with a neutral buoyancy depth \( z_{nb} \) of 20 m on Georges Bank, respectively.
Detections of herring shoals are feasible up to $10 \pm 6$ km in range for humpback parameters and $1 \pm 1$ km for minke parameters in common Gulf of Maine herring spawning scenarios of dense near-surface concentrations at depths of 15–35 m (red solid lines). Detections of the seafloor are found to be feasible up to $2 \pm 1$ km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over local ambient noise. Dark red dashed lines represent seafloor scattering for a seafloor depth of 50 m. Purple solid lines represent the beamformed ambient noise levels with mean wind speed of 5.6 m/s during peak herring spawning period in the Gulf of Maine with one detection threshold ($DT$) added. Light purple shaded areas range from the beamformed ambient noise levels at 2 m/s wind speed minus 5.6 dB to those at 10 m/s wind speed plus 5.6 dB, to illustrate the typical range of wind-speed-dependent ambient noise encountered. Gray solid lines represent the beamformed ambient noise level without significant shipping with wind speed of 5.1 m/s [89] with one $DT$ added. Light gray shaded areas range from beamformed ambient noise levels without significant shipping at 5.1 m/s wind speed plus/minus the instantaneous intensity standard deviation of 5.6 dB. The beamformed ambient noise levels are determined for frequency bands of baleen whale vocalizations given in Table 4.1 using Equations (4.9) and (4.10).
4-5 Detections of herring shoals at depths of 130–180 m (blue solid lines) are found to be unfeasible for the parameters of any baleen whale considered. Detections of the seafloor are found to be feasible up to 2 ± 1 km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over local ambient noise. Dark blue dashed lines represent seafloor scattering for a seafloor depth of 180 m. Purple solid lines represent the beamformed ambient noise levels with mean wind speed of 5.6 m/s during peak herring spawning period in the Gulf of Maine with one detection threshold (DT) added. Light purple shaded areas range from the beamformed ambient noise levels at 2 m/s wind speed minus 5.6 dB to those at 10 m/s wind speed plus 5.6 dB, to illustrate the typical range of wind-speed-dependent ambient noise encountered. Gray solid lines represent the beamformed ambient noise level without significant shipping with wind speed of 5.1 m/s [89] with one DT added. Light gray shaded areas range from beamformed ambient noise levels without significant shipping at 5.1 m/s wind speed plus/minus the instantaneous intensity standard deviation of 5.6 dB. The beamformed ambient noise levels are determined for frequency bands of baleen whale vocalizations given in Table 4.1 using Equations (4.9) and (4.10).

A-1 OAWRS system used for cod measurements during the Nordic Seas 2014 Experiment. Effectively monostatic OAWRS system used for cod measurement during the Nordic Seas 2014 Experiment, where source and receiver arrays were towed from the same research vessel RV Knorr. The OAWRS source was developed under the National Science Foundation and Sloan Foundation MRI program for wide-area sensing of marine life, and the ONR Five Octave Research Array (FORA) was used as the OAWRS receiver.
A-2 Measured vertical sound speed profiles. Measured XBT and CTD sound speed profiles during Nordic Seas 2014 Experiment near Andenes (A) and Røst (B). 113

A-3 Cod target strength calibration at OAWRS frequencies. Expected cod target strength \(\langle TS_{OAWRS} \rangle\) at OAWRS frequencies was determined (A) by matching the cod areal population density obtained from the 12 kHz RV Knorr vertical echosounder measurements at 01:51:00-02:20:24 UTC and that from the OAWRS image at 01:26:49 UTC on March 8, 2014 for corresponding locations along the RV Knorr’s line transect (B). 114

A-4 Northeast Arctic cod length measured over 1985-2014. Time series of mean (circles) and standard deviation (vertical lines) of cod length measured from annual trawl survey in the Lofoten area over 1985-2014. 115

A-5 Least squares estimate of cod group boundary critical density from decades of Northeast Arctic line-transect survey data and ICES spawning population time series. Cod group boundary critical density of roughly 0.016 fish/m² is found to provide the lowest least squares error between the annual spawning population time series and decades of line-transect survey data under the assumption that summing spawning groups approximates the total annual spawning population each year from Eq. A.3. Cod group boundary critical density range, in red, independently found from scaling herring critical density via Eq. A.3 is consistent with this least squares result. 116
A-6 Correspondence between OAWRS and RV Knorr line-transect vertical echosounder measurements of cod near Andenes. Correspondence between (A) an instantaneous OAWRS cod population density image and (B) 12 kHz RV Knorr vertical echosounder line-transect measurements near Andenes. The OAWRS cod population density image (A) was obtained near Andenes at 01:26:49 UTC on March 8, 2014. The 12 kHz RV Knorr vertical echosounder line-transect measurements (B) were made from 01:51:00 to 02:20:24 UTC on March 8, 2014. The white line transect in (A) passing through the group imaged by OAWRS is the RV Knorr track with corresponding vertical echosounder data shown in (B).

A-7 Correspondence between OAWRS and RV Knorr line-transect vertical echosounder measurements of cod near Røst. (A-D) Simultaneous correspondence between OAWRS imaging from RV Knorr and the 12 kHz RV Knorr vertical echosounder line-transect measurements near Røst where RV Knorr tracks through the OAWRS imaged groups are given in (A) and corresponding line vertical echo-sounder data is given in (B-D). (E-H) Correspondence between OAWRS imaging and the 12 kHz RV Knorr vertical echosounder line-transect measurements taken 5 hours after OAWRS imaging near Røst. (I-L) Correspondence between OAWRS imaging and the 12 kHz RV Knorr vertical echosounder line-transect measurements taken roughly 24 hours after OAWRS imaging near Røst.
A-8 Consistency between OAWRS and annual line-transect vertical echosounder measurements of cod near Røst. Consistency between cod population density distribution near Røst measured by OAWRS imaging on February 23, 2014 and by vertical echosounder measurements during the annual line-transect acoustic survey on March 17-31, 2014 (purple circles). High cod density areas near Røst and general tapering trends radially away from Røst are clear and consistent in both OAWRS and vertical echosounder measurements.

A-9 The probability density of cod spawning group size can be characterized by a single parameter, namely the mean spawning group size or quantum. (A) The histogram of the natural logarithm of the cod spawning group size obtained from the 30-year line-transect echosounder data [24] is found to be consistent with a Gaussian probability density distribution. (B) The standard deviation-to-mean ratio of cod spawning group size obtained from the 30-year line-transect echosounder data [24] is found to be stable with a mean of 2.0 to within 35% of the mean over 30 years such that \( q \approx 2\sigma \), where \( q \) is the mean and \( \sigma \) is the standard deviation of the cod spawning group size. Since the Gaussian density parametrically depends on only the mean and standard deviation, and the standard deviation can be expressed in terms of the mean, the probability density of cod group spawning size depends only on the mean group size, or the cod quantum, \( q \).

A-10 RV Knorr locations for cod population density mosaic RV Knorr locations for the Fig. 2-1B mosaic at 10:40:49, 11:29:09, 12:13:19, 13:09:59, 14:04:59, and 15:07:29 UTC on February 23, 2014.
A-11 Population growth rate as a function of total spawning population each year from 1920 to 2014 for each cod population investigated. The cod quantum is the mean spawning groups size of 10.6 million found for Atlantic cod from a total population at pre-industrial levels, where the only available data was from the Northeast Arctic. The figures do not consistently show higher variance in growth rate for smaller populations.

A-12 Population growth rate as a function of total spawning population each year from 1920 to 2014 for each herring population investigated. The herring quantum is the mean spawning groups size of 204 million found for Atlantic herring when the total population was at pre-industrial levels, where the only available data was from Georges Bank. The data show indications that for small populations below a quantum growth rate variance may be higher.

B-1 The OAWRS scattering strength images and corresponding exemplary coherence areas taken at (A-B) 17:25:00, (C-D) 18:42:30, (E-F) 19:52:30, (G-H) 21:37:30 EDT on September 29, 2006 (B,D,F,H). The e-folding lengths from the peak of the 2D autocorrelation function of herring scattering strength are plotted in black.

B-2 (A) The horizontal distribution of neutral buoyancy depth. (B) Estimation uncertainty for neutral buoyancy depth in percentage. (C) The horizontal distribution of herring population density. (D) Estimation uncertainty for herring population density in percentage. (E) The horizontal distribution of root-mean-square (RMS) shoal thickness $H^{RMS} = 2H/\sqrt{12}$, where $H = |\max z - \min z|$, and $z$ is the herring occupancy depth. (F) Estimation uncertainty for RMS shoal thickness in percentage.
C-1 Wind dependence of Gulf of Maine mean ambient noise levels for frequency bands of baleen whale vocalizations given in Table 4.1. Gray filled circles represent the mean ambient noise levels determined from measurements obtained during the OAWRS 2006 experiment [1, 14, 34] using Equation (4.9) for frequency bands of baleen whale vocalizations given in Table 4.1. Black solid lines represent the empirical best fit of the measured ambient noise levels to the model given in Equation (4.11). Empirically determined coefficients of the model are given in Table 4.1. Black vertical lines represent the variation of the measured mean ambient noise levels for given wind speed. Instantaneous measurements of ambient noise fluctuate around each mean ambient noise level with 5.6 dB standard deviation [41].

D-1 Humpback song occurrence rate is constant in the periods “before” and “during” OAWRS survey transmissions. The mean percentage of a diurnal cycle containing humpback whale song in the periods “before” and “during” OAWRS survey transmissions, as defined in Section D.1, remains constant, indicating the transmissions had no effect on humpback whale song over the entire passive 400-km diameter survey area of the Gulf of Maine including Stellwagen Bank.
Reported humpback whale Stellwagen Bank song occurrence [126] shows large natural variations within and across years. Large natural variations in humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise within and across years are common in the absence of sonar. Line plots of reported single sensor daily humpback whale song occurrence at Stellwagen Bank in hours/day (A) for the entire year and (B) from September 15 to October 17, in 2006 and 2008 [126]. Many periods lasting roughly weeks where high song occurrence episodes are found in one year but not in another when no sonars are present are indicated by black arrows in (A). The reported reducing change in humpback whale song occurrence, to zero [120, 126], occurred in the “before” period while the OAWRS vessels were inactive and docked on the other side of Cape Cod from Stellwagen Bank, at the Woods Hole Oceanographic Institution, due to severe winds for days before OAWRS transmissions for active surveying began on September 26, 2006, as marked by the black arrow in (B). This shows that Risch et al. [120] analysis violates temporal causality.
Quantifying large differences in the reported humpback whale song occurrence at Stellwagen Bank [126] across years. Difference in humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise across years exceeds that of the “during” period most of the time when no sonars are present. (A) Difference in mean humpback whale song occurrence at Stellwagen Bank over respective 11-day periods with 1-day increment in 2006 and 2008, (B) histogram of difference in mean humpback song occurrence over 11-day periods between 2006 and 2008 when no sonar is present, i.e. excluding the “during” period from September 26 to October 6. Periods when the difference in means of respective 11-day periods is greater than (red dots) and less than (blue dots) that of the “during” period are indicated in (A). The difference in means fluctuates randomly throughout the year, exceeding the “during” period 57.8% of the time (most of the time) when no sonars are present, indicating that there is nothing unusual about such differences, which are actually the norm.
D-4 Reported annual humpback song occurrence at Stellwagen Bank [126] are uncorrelated between years over 11-day periods. Annual humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise are uncorrelated over 11-day periods across years. (A) Correlation coefficient between 2006 and 2008 humpback whale song occurrence time series over 11-day period with 1-day increment (B) histogram of the correlation coefficient in (A). The correlation coefficient of the annual humpback whale song occurrence time series over 11-day periods across years obeys a random distribution peaking at zero correlation about which it is symmetric, showing that correlation in trend between years is random and quantitatively expected to be zero with roughly as many negative correlations as positive ones. The correlation coefficient between the humpback whale song occurrence across years smoothly transitions from negative values in the “before” period, showing no similarity or relation in trend between years just before the 2006 OAWRS survey transmission period, to some of the highest positive correlations obtained between years in the “during” period. This demonstrates high similarity and relation in trend between years during the 2006 OAWRS active survey transmission period, which contradicts the results of the Risch et al. [120] study.
Wind-dependence of mean detection range for single sensor at Stellwagen Bank [120], and OAWRS receiver array. The green shaded areas indicate the overall vocalizing humpback whale call rate densities (number of calls/[(min) (5 nmi)$^2$]) determined between September 22 and October 6, 2006 by our large aperture receiver array towed along several tracks (black lines). The mean detection ranges for the single sensor at Stellwagen Bank are in blue and for the OAWRS receiver array are in red, where Stellwagen Bank is marked by yellow shaded regions. These detection ranges are determined by the methods described in Section D.2.3 given a humpback whale song unit source level of approximately 180 dB re 1 $\mu$Pa and 1 m which is the median of all published humpback whale song source levels [188, 189, 191–194]. The error bars represent the spread in detection range due to typical humpback whale song source level variations (Section D.2.3). Under (A) low wind speed conditions vocalizing whales are within the mean detection area for a single Stellwagen Bank sensor but for (B) higher wind speeds most vocalizing whales are outside the mean detection area of the same sensor, which results in reduction of detectable whale song occurrence by the single sensor [120] at Stellwagen Bank. . . . . 156
D-6 Wind-speed increase causes reduction in humpback song occurrence at Stellwagen Bank. Average wind speed increase from the “before” to the “during” period at Stellwagen Bank causes reduction in the percentage of time humpback whale songs are within mean detection range of a single Stellwagen Bank sensor. (A) Averaged wind speed measured at the NDBC buoy [131] closest to Stellwagen Bank over the “before,” “during,” and “after” 11-day periods; and (B) percentage of the time vocalizing humpback whales localized by our large aperture array are within the mean detection range of the single sensor [120] at Stellwagen Bank in the “before” and “during” periods, using waveguide propagation methods and whale song parameters described in Section D.2.3. Since the OAWRS experiment was conducted only up to October 6, 2006, the humpback whale source distribution in the “after” period was not measured and we do not investigate the percentage of time that humpback whales are within the mean detection range of the single sensor at Stellwagen Bank [120] for the “after” period. The triangles represent the mean wind speed and the solid ticks represent the standard deviation of the wind speed over the respective 11-day periods.
Humpback song occurrence detectable by single sensor matches reported humpback song occurrence at Stellwagen Bank [120]. Average humpback whale song occurrence detectable by a single hydrophone at Stellwagen Bank in time-dependent ambient noise in the “before” and the “during” periods matches the reported humpback whale song occurrence by Risch et al. [120]. Using the measured wind speeds at Stellwagen Bank [131] (Figure D-6), the measured spatial distribution of vocalizing humpback whales (Figure 1 of Ref. [34]), and constant song production rates (Figure D-1) measured by our large-aperture array, the detectable song occurrence over the “before” and “during” period are found to be within ±18% of the reported means [120], much less than the standard deviations of reported song occurrence [120], using waveguide propagation methods and whale song parameters described in Section D.2.3. Before and during OAWRS survey transmissions, this figure shows that reported variations in song occurrence at Stellwagen Bank by Risch et al. [120] are actually due to detection range changes caused by wind-dependent ambient noise, through well-established physical processes [127, 132].

Histogram of the measured humpback whale song unit source levels. The humpback whale song unit source levels measured from more than 4,000 recorded song units during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location approximately follow a Gaussian distribution and are in the range 155 to 205 dB re 1 μPa and 1 m with a mean of 179.8 dB re 1 μPa and 1 m and a median of 179.4 dB re 1 μPa and 1 m, which are within 0.6 dB of the median of all published humpback whale song unit source levels of 180 dB re 1 μPa and 1 m [188, 189, 191–194]. The solid and dashed gray lines represent the mean and the median of the measured humpback song unit source levels, respectively.
Autocorrelation of Vu et al. [126] humpback whale song occurrence time series in 2006 and 2008. The e-folding time scale $\tau_e$ of the Vu et al. [126] annual humpback whale song occurrence time series is (A) 18 days for 2006 and (B) 21 days for 2008. The roughly 20-day coherence time scale shows that the humpback song occurrence gradually changes over periods longer than the 11-day periods analyzed in Risch et al. [120]. It is noteworthy that (1) the humpback song occurrence dropped to zero in the “before” period, and (2) only after a time period consistent with the measured coherence time scale of song occurrence, within which temporal processes are correlated, did song occurrence begin to increase in the “during” period (Figure D-2). The Risch et al. [120] analysis then violates temporal causality because the correlated processes that caused the reduction in humpback song occurrence started days before the OAWRS survey transmissions began, yet the analysis and conclusions of Risch et al. [120] offer no other explanation than these survey transmissions for the reduction. Both time series show high correlation at a time lag of roughly seven months due to increases in song occurrence during the spring and fall seasons (Figure D-2), separated by roughly seven months.
E-2 (Left) Uniform song length distribution incorrectly and implicitly assumed in the analysis of Fristrup, Hatch, and Clark [222]. The implicit assumption that song length distribution is uniform in the analysis in Fristrup, Hatch, and Clark [222] is derived because there is no other way to obtain the predictions of the null hypothesis of Fristrup, Hatch, and Clark [222] from the calculated probability density of overlap conditioned on song length without using the uniform song length distribution. (Right) Non-uniform and peaked distribution of finite support given as a control humpback song length distribution measured before any active transmissions on each day [222]. This non-uniform and peaked song length distribution is obtained from the measured song length histogram when no sonar is present with bin counts specified in FIG. 8 of Fristrup, Hatch, and Clark [222].

E-3 (Left) Joint probability density function of humpback song length and overlap with acoustic transmissions assuming a uniform song length distribution with no sonar present. White dashed line represents the predictions based on the null hypothesis of Fristrup, Hatch, and Clark [222]. (Right) Joint probability density function of humpback song length and overlap with active acoustic transmissions assuming a non-uniform and peaked song length distribution measured with no sonar present given in Fristrup, Hatch, and Clark [222] as a control song length distribution (Right panel of FIG. E-2). The assumption of uniform song length distribution with no sonar present leads to a drastically different results than the use of a measured song length distribution with no sonar present.
E-4  (Left) Measured relationship between humpback song length (black crosses) and minutes of overlap with active acoustic transmissions. The three black solid lines represent the relationship between the humpback song length and the ping overlap for 42-second, 84-second, and 126-second ping overlaps predicted by the “oblivious hypothesis” of Fristrup, Hatch, and Clark [222]. The solid lines are drawn through the median values for 42-second, 84-second, and 126-second ping overlaps with a slope equal to the inverse of the duty cycle of active acoustic transmissions. The solid lines are found to be consistent with the relationship between song length and minutes of overlap assuming a uniform song length distribution. (Right) The relationship between song length and minutes of overlap assuming a non-uniform and peaked song length distribution is consistent with the measured relationship given in left panel. The mean song length of the simulated samples for each overlap obtained from the joint probability density function using the non-uniform and peaked song length distribution (Right panel) is found to be within one standard deviation from the measured mean song length for each overlap. Red crosses represent the median song length for corresponding minutes of overlap with acoustic transmissions. No information is provided in Fristrup, Hatch, and Clark [222] to calculate the probability density for zero overlap.

F-1 Shoal population estimates of Ålesund herring on February 18 and 20, 2014

List of Tables

4.1 Parameters of baleen whale’s auditory reception and signal detection used for detection range estimation in the Gulf of Maine scenario 85

4.2 Parameters of herring shoals 85

D.1 Percentage of time the Risch et al. statistical test [120] incorrectly finds whales respond to sonar when no sonar is present using annual humpback whale song occurrence data reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise. 150

D.2 The Risch et al. statistical test is applied to the same humpback whale song occurrence data reported in Risch et al. [120] over the 33-day period from September 15 to October 17 for 2008 and 2009, with either of these two years as the test year and the other as the control year. It false-positively finds that whales react to sonar 100% of the time when no sonar is present, indicating self-contradictions in the Risch et al. [120] approach. 150

D.3 Typical anthropogenic noise sources at Stellwagen Bank. 151

D.4 Received mean intensity of typical anthropogenic noise sources at Stellwagen Bank. 161

D.5 Risch et al. statistical test statements [120]. 162

D.6 Possible outcomes of each pairwise comparison between the mean humpback whale song occurrence in the $j^{th}$ 11-day period of the $i^{th}$ 33-day period in the $k^{th}$ year and that in the $l^{th}$ 11-day period of the $i^{th}$ 33-day period in the $m^{th}$ year in the Risch et al. statistical test [120]. 162
Chapter 1

Introduction

Many oceanic fish species are known to form extensive social groups [1–5], typically referred to as fish shoals, and collectively behave in groups to accomplish essential tasks. For example, the annual group behavior of Atlantic cod to form extensive shoals for spawning is often associated with the regional cod population’s health given the reported observations that the recent population collapse of cod coincided with a disruption in shoaling behavior for spawning [6, 7]. It has been difficult, however, to quantitatively study such group behaviors primarily due to the technical obstacles [8] in underwater sensing over wide areas without spatial and temporal aliasing because the conventional, highly-localized line-transect methods from slow-moving vessels significantly undersample fish populations in space and time.

A wide-area ocean acoustic sensing method, known as Ocean Acoustic Waveguide Remote Sensing (OAWRS), has been demonstrated to be capable of instantaneously imaging and continuously monitoring underwater animal behaviors over continental-shelf scales [1–3, 9]. Over the last decade, the OAWRS measurements revealed the behavioral dynamics of fish shoal formation [1], horizontal migrations [1], shoal segmentation and dispersal [2], and neutral buoyancy regulation [9]. There still remain many intriguing fish group quantities to be determined including characteristic geometry of horizontal morphology, specific migration pattern in space and time, migration speed, expected fish population in each group, diel vertical migration, vertical distribution during migration, buoyancy regulation during migration, total fish pop-

35
ulation, and more. In this thesis, we experimentally determine biologically important but previously not accessible quantities and investigate the mechanisms for fish population collapse or recovery by further analyzing acoustic measurements from a recent OAWRS experiment in Winter 2014 in Lofoten, Norway and annual conventional line-transect surveys in the historic Lofoten cod spawning ground in Norway, as well as the past OAWRS experiments in the US Northeast coast.

We first investigate the mechanism for collective spawning of oceanic fish that form large spawning groups with ecosystem-scale ocean-acoustic sensing in combination with conventional acoustic survey methods. We do so by empirically determining the mean spawning group size for oceanic fish species. We find that the mean shoal population per annual spawning season of Northeast Arctic cod over the entire spawning ground in Lofoten Norway is stable across the available 30 years of line-transect survey data. Time series of the total Atlantic cod spawning population for major spawning regions across the North Atlantic show that once the total spawning population declined below a quantum, return to pre-industrial levels did not occur in that region even after decades. We find that the daily herring spawning group population is stable over the peak annual spawning period from wide-area acoustic measurements of spawning herring in the Gulf of Maine in Fall 2006. As with cod, time series of the Atlantic herring spawning population for major spawning grounds across the North Atlantic show that when total spawning population declined below the empirically determined quantum level, return to pre-industrial levels required decades. Our findings show that to be sustained at pre-industrial levels the total spawning population must greatly exceed the mean spawning group size found at pre-industrial levels for any oceanic fish population we investigated, and likely many others.

Via multi-spectral OAWRS measurements, we capture a shift in swimbladder resonance peak correlated with the herring groups’ up-slope spawning migration, enabling 3D spatial behavioral dynamics to be instantaneously inferred over thousands of square kilometers. We show that herring groups maintain near-bottom vertical distributions with negative buoyancy throughout the migration. We find a spatial correlation greater than 0.9 between the average herring group depth and corre-
sponding seafloor depth for migratory paths along the bathymetric gradient. This is consistent with herring groups maintaining near-seafloor paths to both search for optimal spawning conditions and reduce the risk of predator attacks during the migration to shallower waters where near-surface predators are more dangerous. This analysis shows that multi-spectral resonance sensing with OAWRS can be used as an effective tool to instantaneously image and continuously monitor the behavioral dynamics of swimbladder-bearing fish group behavior in 3 spatial dimensions over continental-shelf scales.

Finally, we investigate the feasibility of the echolocation function for certain baleen whale vocalizations that show a high spatial and temporal correlation with peak annual spawning processes of Atlantic herring in the Gulf of Maine. We determine physical limitations on the ability to detect large herring shoals and the seafloor by acoustic remote sensing with ocean acoustic propagation, scattering, and statistical theories given baleen whale auditory parameters. We find detection is highly dependent on ambient noise conditions, herring shoal distributions, baleen whale time-frequency vocalization spectra, and geophysical parameters of the ocean waveguide. No effect of anthropogenic sound on marine mammal vocalization behavior was found during our measurements. Some published statistical tests assessing the impact of anthropogenic sound on marine mammal behavior were found to have 98-100% false positive biases with no true positive confirmation, and so lack statistical significance.
Chapter 2

Group Behavioral Quantization of Oceanic Fish

2.1 Introduction

The decline of many oceanic fish species has been correlated with increased human presence in the industrial era [10–13]. An obstacle to help reverse or stabilize the situation is the difficulty in observing the behavior of large oceanic fish populations in the wild due to the vast undersea areas they inhabit [4, 8]. Here we investigate how the viability of oceanic fish may be related to their observed group behavior. We do so by studying group behavior during critical spawning periods in regions where populations are currently at pre-industrial levels. The essential observational tool enabling this analysis is ocean acoustic waveguide remote sensing (OAWRS), which monitors detailed spatio-temporal group behavioral processes over ecosystem scales in a manner that is not currently possible with other methods [1, 2, 14]. The most abundant oceanic fish species are known to gather in immense social groups during spawning [15]. Here we limit our study to species of oceanic fish for which the entire instantaneous horizontal morphologies of vast spawning groups have been measured. This focuses the study to Atlantic cod (Gadus morhua) and Atlantic herring (Clupea harengus), where the relevant measurements were made with OAWRS. Without such ground truth observations of group morphology for a given species, it is difficult
or impossible to unambiguously detect, quantify and enumerate spawning groups in available conventional data. This is because conventional data uses line-transects and other sparse sampling methods that have the potential to greatly under sample fish population distributions and groupings in space and time due to the vast regions that they inhabit and occupy [2] (Supplementary Information). A primary goal of this research is to investigate the relative magnitudes of spawning populations with respect to total pre-industrial levels and mean pre-industrial spawning group size. This leads to the finding that to be sustained at pre-industrial levels, the total spawning population must greatly exceed the mean spawning group size found at pre-industrial levels for the given species. As an apparent consequence of this large differential, in regions where total populations declined below this mean quantum, return to pre-industrial levels has not yet occurred or required decades for any oceanic fish population we investigated. While these results currently are only for the populations we investigated, they may apply to similar oceanic fish species that form vast spawning shoals. Byproducts of this research include to our knowledge the only available instantaneous images of the population density of entire cod spawning groups, stretching for tens of kilometers, the only available multi-decadal quantification of spawning group size distributions for a major oceanic fish species, and the accurate enumeration of the entire Georges bank herring stock with roughly 10 minutes of OAWRS sampling to within 7% percent of the NOAA estimate [16], which relied on weeks to months of conventional sampling.

Both Atlantic cod and Atlantic herring are major oceanic fish species that are extremely important to ocean ecosystems. Atlantic cod is of the fish-feeding Gadidae family of top level predators. In contrast, Atlantic herring is of the plankton feeding Clupeidae family of keystone species that are prey to many other fish, marine mammals and birds. Both families are found throughout the world’s oceans and are important in circulating biomass between open ocean and coastal waters via feeding to spawning cycles. Both families are extremely important economically and as a human food source [17, 18], as are many of the most abundant oceanic fish species. Atlantic cod and Atlantic herring in particular both have long economic and cultural
histories in coastal North Atlantic and Arctic regions, where collapses and near collapses of their populations in Canada, the United States, the North Sea, and Norway in the 20th to 21st centuries have led to economic turmoil for many coastal communities [12, 13], and calls for new survey methods with ecosystem-scale sensing [19, 20]. Various other species of gadids such as walleye pollock, haddock, whiting, and saithe, and clupeids such as sardines and sprats, also form large spawning groups that have traditionally been difficult or impossible to sense in their entirety.

2.2 Results and Discussion

We measured the instantaneous 2-D horizontal morphologies of entire vast cod groups extending for tens of kilometers in length, covering thousands of square kilometers, and containing millions of individuals, which to our knowledge have not been previously observed. These measurements were of Northeast Arctic cod (Gadus morhua) in their primary spawning ground off the coast of Lofoten Norway, during the 2014 spawning period using OAWRS (Fig. 2-1). The largest spawning group we observed spanned roughly 40 km in diameter, nearly 1/2 degree in latitude, and was populated by at least 40 million cod (Fig. 2-1B). It was found at the southern end of the Lofoten spawning ground where the densest concentrations surrounded the Island Røst. Here the cod’s preferred water depths of 100±20 m (Supplementary Information) approached rising seafloor about the island that curtailed further migrations inward. Such high cod concentrations surrounding Røst were observed to be spatially and temporally stable over days to weeks (Supplementary Information) apparently in relation to the island’s fixed bathymetric features. In contrast, at the northern end of the Northeast Arctic spawning ground, near Andenes, smaller less dense groups of roughly 20 by 2 km and 1 million cod were observed (Fig. 2-1C) to exhibit significant temporal and spatial variation over a period of roughly 1-2 days and were not associated with specific bathymetric features. All cod spawning groups were found to be bounded by a critical population density of roughly 0.016 fish/m² consistent with that measured for group formation of spawning Atlantic herring in the Gulf of Maine.
when adjusted for body length (Supplementary Information).

These initial instantaneous wide-area observations of cod aggregations suggest that the extensive but spatially discrete spawning groups observed have quantifiable properties that are linked to essential collective behavioral functions. During feeding, cod behave as individual predators [4, 21]. During spawning, a significant amount of energy is invested in group behavioral processes. Known advantages of cooperative group behavior of oceanic fish species during spawning include a highly increased mate-encounter rate which leads to improved mate choice and increased egg fertilization rates as well as time-space alignment to enable synchronized batch spawning [1, 22, 23]. Too large a spawning group, however, may potentially result in disadvantages including extra energy costs to maintain the group, nutrient depletion, environmental pollution, and increased risk contamination from diseases and parasites [22, 23]. Spawning group size likely evolves through a balance between such factors, and may have invariant properties if the influencing factors are stable.

We find that the mean group population per annual spawning season of Northeast Arctic cod over the entire spawning ground in Lofoten Norway is remarkably invariant across the available 30 years of line-transect survey data [24], averaging to 10.6 million cod to within 36% with no apparent temporal trend (Fig. 2-2A-B). In contrast, much larger variations in the total annual spawning population, including an increase of roughly 600% or more, were observed over this same 30 year period (Fig. 2-2C). The marked stability of the annual mean spawning group size in contrast to the large variations in total population across years supports the interpretation of the expected spawning group size over the 30-year data set as a mean unit or mean quantum empirically expected for reliable spawning. This stability is also seen by the fact that since the mid 2000s, when the Northeast Arctic cod population returned to pre-industrial levels (Fig. 2-2C), the mean spawning group size of 11.1 million cod to within 25% has insignificant difference from that of the 30-year mean. The annual group spawning size probability density is found to be consistent with a log-normal density (Supplementary Information, Supplementary Fig. A-9A), which is associated with growth processes similar to those of the spawning cod groups [25–
The log-normal density parametrically depends only on the mean and standard deviation, where it is important to note that the standard deviation of annual group size (Supplementary Fig. A-9) is not the same as the standard deviation of the mean group size across years (Fig. 2-2B). The annual spawning group size is found to have a standard-deviation-to-mean ratio of 2.0 that is constant to within 35% over the 30 years of data. As a result, the probability density governing annual spawning group size and the process leading to it can be roughly characterized by a single parameter, namely the mean spawning group size or cod quantum (Supplementary Fig. A-9). This is likely related to the single-parameter power law that has been observed previously in the wide-area spatial spectra of fish aggregations, which indicates a fractal process of structural self-similarity at all scales [2, 28].

To perform this analysis, discrete cod spawning groups were detected and counted in the 30-year line-transect survey data set. This was done by spatially interpolating between line transects for each annual survey and segmenting regions at and above the empirically determined group boundary threshold density of 0.016 fish/m² (±0.006 from fish length scaling) to identify discrete spawning groups, where this threshold was independently determined by a number of approaches using both line-transect and OAWRS measurements. Regional group morphologies estimated by analyzing line-transect data in this manner were found to be consistent with those directly measured in 2-D horizontal OAWRS imagery of dense spatial sampling (Supplementary Information).

Pre-industrial hindcasts for the 19th century (Supplementary Information) combined with available data over roughly the last century indicate that the total spawning population of Northeast Arctic cod have not fallen below the mean spawning group level in these time periods (Fig. 2-2C). Relevant environmental conditions in the Northeast Arctic spawning grounds are similar to those found in other major Atlantic cod spawning grounds [29]. Given this and the fact no other data on the mean size of entire Atlantic cod spawning groups at pre-industrial levels is available, the mean spawning group size derived for Northeast Arctic cod found here is used as the only available metric for the mean spawning group size for cod at pre-industrial
levels in other major Atlantic spawning grounds.

Time series of the total Atlantic cod spawning population for major spawning regions across the North Atlantic (Fig. 2-2C) show that once the total spawning population declined below the mean cod spawning group size for a population at pre-industrial levels, return of the total spawning population to pre-industrial levels did not occur in that region even after decades, as quantified in more detail in the Supplementary Information. This is apparently a consequence of the large difference between the total pre-industrial spawning population for each population and the mean spawning group size for the population at pre-industrial levels. Diversification to more than one spawning group, however, increases the probability of robust reproduction by taking greater advantage of spatial and temporal variations or fluctuations in the environment. The presence of more than one spawning group, for example, reduces uncertainties and mitigates associated risks for the overall population given the possibility that some groups may randomly encounter unfavorable local conditions for reproductive fitness. Conversely, the presence of more than one spawning group increases the chances that some groups will encounter highly favorable local conditions. The time series for cod do not consistently show higher variance in growth rate for smaller populations (Supplementary Information).

Quantized group behavior during spawning is also investigated for the Atlantic herring species, which are total breeders, releasing one batch of eggs per season, as opposed to cod which are batch spawners, whose individuals each strive to release multiple batches following multiple courtships across an entire annual spawning season. These two species then have fundamental differences in their spawning strategies that are representative of the two categories of capital spawning found in oceanic fish. For capital spawners, energy is stored in advance so that reproductive success is less dependent on the availability of food during spawning [30, 31]. In the fall of 2006 large Atlantic herring groups (Fig. 2-3A) were observed to follow a consistent daily process over the peak Georges Bank spawning period whereby a group would form on the northern flank of the Georges Bank near sunset and migrate to shallower regions on the bank to spawn under cover of darkness [1]. If each herring goes through the
spawning grouping process only once, as has been previously established [32], the sum of the daily group populations over the spawning period should equal the total Georges Bank spawning population. This is indeed demonstrated by consistency to within 7% of NOAA’s estimate of the total Georges Bank spawning population of 1.74 billion for 2006 over the course of the year [16] and the 1.63 billion estimate independently obtained by summing the daily spawning group populations over the peak spawning period from a single instantaneous OAWRS image each day (Fig. 2-3B).

We find the daily spawning group population of 204 million Atlantic herring on average to be stable with a standard deviation within 35% of the mean over the peak spawning period (Fig. 2-3B). This suggests the mean spawning group population may have evolved to a stable optimal size to fulfill the essential behavioral function of reliable spawning for Atlantic herring, in which case the mean spawning group size could be considered to be a mean behavioral quantum. Since the 2006 measurements were made in a period when the total spawning population was at or above estimated pre-industrial levels (Fig. 2-3C), it is unlikely that industrial-age pressures have significantly affected this empirically derived mean spawning group or quantum size for Atlantic herring. Given the similarity in relevant environmental conditions in the other major historic Atlantic herring spawning grounds, and the lack of data on entire Atlantic herring shoals in other regions, the mean quantum size derived for Georges Bank is used as the only available metric for the the mean spawning group size for herring at pre-industrial levels in other major Atlantic spawning grounds. As with cod, time series of the Atlantic herring spawning population for major spawning grounds across the North Atlantic (Fig. 2-3C) show that when total spawning population declined below the mean herring spawning group size for a population at pre-industrial levels, return of the total spawning population to pre-industrial levels required decades, as quantified in more detail in the Supplementary Information. Again, this is apparently a consequence of the large difference between the total pre-industrial spawning populations and the mean spawning group size of a population at pre-industrial levels.
2.3 Conclusions

We have empirically shown that two distinct species of oceanic fish at pre-industrial population levels quantize into vast groups of stable species-dependent mean population to cooperatively reproduce. These two species, Atlantic cod and Atlantic herring, bound the spectrum from large upper-trophic predators to mid-trophic keystone species, and from batch to total capital spawning. To be sustained at pre-industrial levels, we find the total spawning population must greatly exceed the mean spawning group size found at pre-industrial levels for any oceanic fish population we investigated. An apparent consequence of this large differential is that in regions where total populations declined below this mean species-dependent quantum, return to pre-industrial levels has not yet occurred or required decades. We have demonstrated this empirically for eight distinct major populations across the North-Atlantic over roughly the last century. While these results currently are only for the populations we investigated, they may apply to similar oceanic fish species that form vast spawning shoals. This is significant because many of the abundant oceanic fish species that form vast spawning groups are of critical importance to ocean ecosystems.
Figure 2-1: Spatial population distributions of entire gigantic cod groups of the Northeast Arctic obtained by instantaneous continental-shelf scale imaging (A) Primary Northeast Arctic spawning ground of Atlantic cod (*Gadus morhua*), Lofoten, Norway. Red circles indicate OAWRS imaging coverage in 50 s. Regions where zoomed OAWRS cod population density image appear in (B) and (C) are indicated by purple rectangles. (B) Vast 40-km-diameter temporally stable cod group containing roughly 40 million individuals (152,000 tons) south of Røst Island comprising 17% of the estimated Lofoten cod spawning population. (C) Temporally transient lineated 20-km-scale cod group containing roughly 1 million individuals (3,800 tons) near Andenes, Norway. Dashed white lines indicate water depth contours. Coordinate origins are at OAWRS Research Vessel Knorr locations: (B) 67.3 °N, 11.5 °E on February 23, 2014, 10:40:49 UTC; and (C) 69.0 °N, 14.3 °E on March 8, 2014, 00:16:49 UTC.
Figure 2-2 (facing page): Time series of mean spawning group size for Northeast Arctic cod and regional cod spawning populations across the North Atlantic (A) Cod population density along conventional line-transect surveys in Lofoten Norway with downward directed echosounding at 1-nmi resolution (Supplementary Information). Discrete cod spawning groups are detected (black contours) and enumerated using group structural information obtained by OAWRS in decades of line-transect survey data (Supplementary Information). Example years of relatively high (2014) and relatively low (1999) cod abundance are shown. (B) Mean annual spawning group population over the Northeast Arctic spawning for 30 years of line-transect survey. Stability of the 30-year average of 10.6 million to within 36% (blue line with error bar showing 30-year standard deviation of the mean) suggests it serves as a mean behavioral unit or mean quantum for effective spawning. (C) Time series of total cod population at major spawning grounds in the North Atlantic during spawning show large variations over the last century. Time series for a total spawning population are shown in red from the time they decline below a cod quantum to the time they return to pre-industrial levels. The cod quantum is the mean spawning group size (10.6 million) found for Atlantic cod when the total population was at pre-industrial levels, where the only available data is from the Northeast Arctic. Gray and blue shaded areas indicate estimated spawning population levels in indicated regions before the onset of industrial fishing in the early 20th century (Supplementary Information).
Figure 2-3 (facing page): Time series of mean spawning group size for Georges Bank herring and regional herring spawning populations across the North Atlantic (A) Location of the daily spawning herring group on the northern flank of Georges bank during the 8-day peak spawning period in the Fall of 2006 determined by OAWRS imagery. Group boundaries occur at the group formation critical density of 0.2 fish/m² for herring [1]. (B) Daily spawning group population over the peak spawning period obtained from instantaneous OAWRS images of each group. Spawning group population has a mean of 204 million to within 35% (blue line and error bar indicate the standard deviation in group size) over the peak spawning period in 2006. (C) Time series of total herring population at major spawning grounds in the North Atlantic during spawning show large variations over the last century. Time series for a total spawning population are shown in red from the time they decline below a herring quantum to the time they return to pre-industrial levels. The herring quantum is the mean spawning group size (204 million) found for Atlantic herring when the total population was at pre-industrial levels, where the only available data is from Georges Bank. Gray and blue shaded areas are estimated spawning population levels in indicated regions before the onset of industrial fishing in the early 20th century (Supplementary Information). The OAWRS sensor location on October 4, 2006, 00:16:15 EDT, at 42.2 °N, 67.7 °W is the coordinate origin in (A).
A

B

C

One quantum equals the mean size of a spawning group from a population at pre-industrial levels.
Chapter 3

Instantaneous 3D Continental-shelf Scale Imaging of Oceanic Fish by Multi-spectral Resonance Sensing Reveals Group Behavior During Spawning Migration

3.1 Introduction

During their reproductive period, oceanic fish of many species migrate to specific grounds to spawn where they form extensive social groups for cooperative reproduction that often span hundreds to thousands of square kilometers [1, 4, 33]. The 3D spatial behavioral dynamics of these immense social groups are largely unknown due to difficulties in sensing over such wide areas, but hold important clues about the ability of their respective species to survive.

A wide-area underwater sensing method known as ocean acoustic waveguide remote sensing (OAWRS) has been shown to be capable of instantaneously imaging and continuously monitoring the morphologies of such large social groups of oceanic fish
over continental-shelf scales [1, 2]. As typically used, however, OAWRS integrates scattered contributions from fish over depth, so yielding only the 2D horizontal distributions of areal fish population density at any instant without more specific depth information [1, 2].

Here the 3D spatial behavioral dynamics of large Atlantic herring groups is inferred via multi-spectral resonance sensing with OAWRS. We study herring groups during their migration to the Georges Bank spawning ground in Fall 2006. We capture a consistent spatial variation in the frequency response of scattered returns from the herring. The resonance peak is found to decrease as the herring migrate to shallower depths. Herring group parameters are then estimated via an exhaustive search for an empirical best-fit between the measured and modeled frequency responses at each spatial position. Assuming an approximately uniform vertical layer of herring [1, 2, 9], the group parameters estimated here include mean shoal depth, shoal thickness, neutral buoyancy depth, and areal population density.

Resonant scattering from herring has been previously studied during the shoal formation process in deep waters (~180 m depth) on the northern flank of Georges Bank. In that study typical damped-harmonic-oscillator sub-resonance scattering was found with a 20 dB/octave roll-off as frequency decreased [9]. The resonance peak was inferred to be at roughly 1500 Hz, outside the OAWRS measurement range between 415 to 1125 Hz. As the herring groups migrate from the deeper water to much shallower spawning grounds on Georges Bank, the resonance peak is found here to move into the frequency band of the current OAWRS measurements. The measured decrease in the resonance peak frequency during the migration is expected because a decrease in fish depth during migration leads to ambient pressure reduction and volume expansion in the swimbladder, assuming Boyle’s law. This then leads to increased swimbladder compressibility and reduced resonance frequency.

Our analysis indicates that herring groups maintain near-bottom vertical distributions with negative buoyancy throughout the upslope migration towards the spawning grounds on Georges Bank. We find a high spatial correlation of greater than 0.9 between the average herring group depth and corresponding seafloor depth for
migratory paths along the bathymetric gradient. From a behavioral perspective, the near-bottom migratory route is sensible because seafloor suitable for spawning can be better sought in the vicinity of the seafloor. Predation risk is also reduced by staying near the seafloor where hunting efficiency is reduced.

More generally, our results show that multi-spectral resonance sensing with OAWRS can be an effective tool to instantaneously image and continuously monitor the behavioral dynamics of swimbladder-bearing fish group in 3 spatial dimensions over continental-shelf scales.

3.2 Materials and Methods

3.2.1 OAWRS Gulf of Maine 2006 Experiment

Data presented here is from the OAWRS Gulf of Maine 2006 Experiment which was conducted to study the behavior of Atlantic herring during their peak spawning period on Georges Bank via continuous monitoring with instantaneous wide-area sensing...
The experiment took place from September 19 to October 6, 2006. The OAWRS data used here involves active transmissions of 1 s duration linear-frequency-modulated waveforms from a vertical source array. The array transmits sound omni-directionally in the horizontal in the frequency range of 390 to 1150 Hz. Scattered returns from environmental features are received by a horizontal towed line array with multiple nested sub-apertures containing a total of 160 hydrophones spanning a frequency range from below 50 to 3750 Hz for spatially unaliased sensing.

Three linear apertures of the receiver array, i.e. the low-frequency (LF) aperture, the mid-frequency (MF) aperture, and the high-frequency (HF) aperture, consist of 64 equally spaced hydrophones with respective inter-element spacing of 1.5 m, 0.75 m, and 0.375 m [35]. Images are generated by beamforming, matched filtering, and charting scattered returns. This leads to a fixed range resolution of roughly 15 m, and an azimuthal resolution varying as \( \lambda / (L \cos \theta) \) in radians away from array endfire, where \( \theta \) is the scan angle from array broadside, \( \lambda \) is the acoustic wavelength, and \( L \) is the array aperture length. Here a non-uniformly-spaced aperture [36] combining the LF and MF apertures is used for beamforming to obtain OAWRS images at frequencies centered at 735 and 950 Hz with enhanced angular resolution compared to that of the MF aperture alone. Similarly OAWRS images are generated at frequencies centered at 1125 Hz with a sub-aperture combining the LF, MF, and HF apertures leading to better angular resolution compared to that of the HF aperture alone by a factor of 2-3 [36]. A maximum-likelihood method [37] is used to deconvolve the array beam patterns of the non-uniformly-spaced array apertures used here from the beamformed OAWRS intensities, such that the azimuth-dependent artifacts resulting from the limited angular imaging resolution are further suppressed. Instantaneous snapshots of the ocean environment over the two-way travel times of the scattered returns are then obtained with sufficient imaging resolution to investigate the behavioral dynamics of oceanic fish group behavior over vast regions.
3.2.2 Instantaneous multi-spectral OAWRS scattering strength images of large herring groups

The OAWRS intensity measurements at frequencies centered at 415, 735, 950, and 1125 Hz from 17:25:00 EDT to 21:37:30 EDT on September 29, 2006 are compensated for two-way acoustic propagation in range-dependent continental-shelf waveguides through parabolic equation modeling, spatially varying OAWRS imaging resolution, and source power [1–3, 9, 14, 34] to investigate the frequency response of scattering from herring. This yields the OAWRS images of scattering strength $SS_{data}$, a quantity that depends on scattering from herring only, which can be expressed as:

$$SS_{data} = 10 \log_{10} \left( \frac{\Phi}{\Phi_{ref}} \right)^2 - SL - TLA,$$

where $\Phi$ is the measured acoustic pressure, $\Phi_{ref}$ is the reference acoustic pressure in water, $SL$ is the transmitted level, and $TLA$ is the depth-averaged two-way transmission loss to individual herring integrated over OAWRS imaging resolution. The $TLA$ is given [3] by

$$TLA = 10 \log_{10} \left( \int_{A_{R}(\rho_{C})} \frac{1}{H} \int_{z_{0} - H/2}^{z_{0} + H/2} \chi(r, r_{S}; r_{T}) dz_{T} d\rho_{T}^{2} \right),$$

where

$$\chi(r, r_{S}; r_{T}) = (4\pi)^{4} |G(r|r_{T}; f, c(r_{w}), d(r_{w}))|^{2} |G(r_{T}|r_{S}; f, c(r_{w}), d(r_{w}))|^{2},$$

and $G(r|r_{T}; f, c(r_{w}), d(r_{w}))$ is the Green function between the target location $r_{T} = (\rho_{T}, z_{T})$ and the receiver location $r$, $G(r_{T}|r_{S}; f, c(r_{w}), d(r_{w}))$ is the Green function between the source location $r_{S} = (\rho_{S}, z_{S})$ and the target location $r_{T}$, and $A_{R}(\rho_{C})$ is the OAWRS imaging resolution footprint [1–3, 9] centered at horizontal location $\rho_{C}$, and $c(r_{w})$ and $d(r_{w})$ are the sound speed and the density in the water column at any point $r_{w}$ in the propagation path, respectively. The transmitted level at each
OAWRS frequency is calibrated [38] from one-way propagated signals received by a desensitized hydrophone on the OAWRS receiver array on September 29 in 2006. A parabolic equation model [39] is used to calculate the Green functions in Eq. (3.2) in a range-dependent Gulf of Maine environment. The conditional expectation over the sound speed and the density in Eq. (3.2) is determined by averaging 5 Monte-Carlo realizations, where the Green functions are calculated along the entire propagation path in range and depth for each realization. Each Monte-Carlo realization employs sound speed profiles measured during the OAWRS 2006 experiment [9] every 500 m [38, 40] along the propagation path. The measured frequency response of herring scattering from each spatial position is determined from these scattering strength images of herring shoals at frequencies centered at 415, 735, 950, and 1125 Hz. Here two consecutive, instantaneous OAWRS intensity images and two consecutive 15-m range cells are averaged to reduce the measured OAWRS intensity standard deviation to roughly 1.3-1.8 dB per pixel [1–3, 41, 42]. The stochastic acoustic propagation modeling is employed to obtain the expected transmission loss with standard deviation of roughly 1-1.5 dB over the herring shoal thickness [1, 9, 34], leading to an overall error of roughly 1-2 dB per pixel [1, 9] in the obtained scattering strength. The changes in herring occupancy depth have a negligible effect on the relative level differences of the expected two-way transmission loss between frequencies leading to variations less than 0.5 dB.

### 3.2.3 Instantaneous 3D imaging via multi-spectral resonance sensing

Here herring groups’ depth information is inferred from the instantaneous OAWRS horizontal 2D images taken at multiple frequencies, yielding instantaneous 3D spatial distributions of large herring groups. This is done by determining the herring group vertical distribution parameters via an exhaustive search for an empirical best-fit between the measured and modeled frequency responses of resonant scattering from herring swimbladders (Eqs. (3.10)-(3.11)). The herring group parameters include
Figure 3-2: Spatial shoal segmentation pattern. The segmentation pattern is swept by a unit annular sector covering 4 X 4 grid area for analysis. A unit annular sector contains roughly 270 independent measurements.

mean shoal depth, shoal thickness, neutral buoyancy depth, and areal population density, assuming that herring are in an approximately uniform vertical layer [1, 2, 9]. The analysis is done for shoal segments that sub-divide the herring high-density area into a number of annular sectors (Fig. 3-2) such that the spatial variation in the herring groups' spatial structure is investigated in an efficient manner (Sec. 3.2.4).

The measured frequency response of herring scattering at each horizontal position is obtained from the instantaneous multi-spectral OAWRS scattering strength images (Sec. 3.2.2), whereas the modeled frequency response is obtained from a swimbladder resonance model [43, 44] as described in Sec. 3.2.5.

3.2.4 Spatial shoal segmentation

The horizontal area occupied by dense herring groups is spatially segmented into a number of annular sectors with spatial overlaps to investigate the spatial variation in the frequency response of herring scattering (Fig. 3-2). Each shoal segment has a 2-km extent in the radial direction from the OAWRS sensor location (approximately across-bank direction), and 10°-extent in azimuthal angle leading to roughly 3.1 km at 18 km range in the along-bank direction. This is comparable to the spatial coherence area of
the herring group formation process with an extent of $1.3 \pm 0.8$ km in the across-bank direction and $3.2 \pm 2.4$ km in the along-bank direction obtained from the OAWRS scattering strength images investigated (Fig. B-1). The spatial coherence area of the herring shoal formation process describes the extent of local population centers within which scattering strength should not vary significantly, and so the herring group formation process is assumed to be stationary. The spatial coherence area is determined from one e-folding length of the 2D scattering strength autocorrelation [45] function of local herring group area.

The shoal segmentation pattern (Fig. 3-2) exploits the fact that the OAWRS resolution footprint is much smaller than the spatial coherence area of the herring group formation process. The OAWRS imaging resolution footprint resolves independent and identically distributed samples of the herring shoal formation process. This is because the range resolution of OAWRS corresponds to the coherence time scale of the fluctuation in the fully randomized received acoustic field, i.e. $1/B = 0.02$ sec, where $B = 50$ Hz is the signal bandwidth [41]. Then each OAWRS resolution cell samples independent acoustic field fluctuations that are identically distributed as long as the biological process that is being observed is stationary. Each beamformed azimuthal direction corresponds to distinct propagation paths leading to independent random fluctuations in the received field. Each shoal segment contains up to roughly $N \approx 270$ independent measurements when fully occupied by fish, which are used to infer herring depth information in a least-squares sense (Sec. 3.2.6). This effectively leads to a reduction in the standard deviation of mean parameter estimates in each shoal segment by a factor of $\sqrt{N}$ compared to the case with $N = 1$. The analysis efficiency is also enhanced by reducing the number of parameter inversions that require exhaustive search. In addition, the temporal coherence scale of the herring group formation process is found here to be roughly 2.5-5 minutes from 18:31:35 to 19:56:15 EDT on September 29, 2006. The temporal coherence scale is determined as the e-folding time of the autocorrelation [45] function of the imaged herring population time series. The temporal coherence scale characterizes major temporal fluctuations in fish population.
3.2.5 Modeled resonant scattering from herring in a uniform vertical layer

The expected scattering strength \( SS^{\text{model}} \) of herring in a uniformly-distributed vertical layer with mean shoal depth \( z_0 \), shoal thickness \( H \), neutral buoyancy depth \( z_{\text{nb}} \), and areal population density \( n_A \) at frequency \( f \) is determined \([43, 44]\) as

\[
SS^{\text{model}}(z_0, H, z_{\text{nb}}, n_A, f) = 10 \log_{10} \left( \frac{1}{H} \int_{z_0-H/2}^{z_0+H/2} \int_0^l \frac{|S(z, z_{\text{nb}}, l, f)|^2}{k} g(l)dl dz \cdot n_A \right),
\]

where \( S \) is the far-field scatter function \([46]\) of a single herring, \( k \) is the acoustic wavenumber, \( l \) is the fork length of herring, \( g(l) \) is the truncated Gaussian probability density function \([44]\) to ensure positiveness of the parameter \( l \), and \( z \) is the herring depth. Here the backscattering cross section of herring is determined as

\[
|S(z, z_{\text{nb}}, l, f)|^2 = \frac{r^2(z, z_{\text{nb}}, l)}{f_0^2(z, z_{\text{nb}}, l) \eta^{-2}(z, z_{\text{nb}}, l, f) + \left( \frac{f_0^2(z, z_{\text{nb}}, l)}{f^2} - 1 \right)^2},
\]

where \( r(z, z_{\text{nb}}, l) \) is the equivalent radius of swimbladder, \( f_0(z, z_{\text{nb}}, l) \) is the resonant frequency of swimbladder, \( \eta(z, z_{\text{nb}}, l, f) \) is the damping factor. In Eq. (3.5), the equivalent radius of swimbladder \( r(z, z_{\text{nb}}, l) \) is determined by:

\[
r(z, z_{\text{nb}}, l) = \left[ \frac{3 c_{\text{nb}} m_{\text{flesh}}(l) (1 + z_{\text{nb}}/10)}{4\pi \rho_{\text{flesh}} (1 + z/10)} \right]^{1/3},
\]

assuming that the swimbladder volume varies with pressure according to Boyle's law \([47]\), where \( c_{\text{nb}} \) is the ratio of the swimbladder volume at neutral buoyancy to the volume of herring's flesh \( V_{\text{flesh}} \) assumed to be 0.05 \([48]\), \( V_{\text{flesh}} = m_{\text{flesh}}(l)/\rho_{\text{flesh}} \), \( m_{\text{flesh}}(l) \) is the mass of a single herring empirically determined by the fork length of herring \( l \) as \( 3.35 \times 10^{-6}l^{3.35} \) in kg where \( l \) is given in cm \([9]\), and \( \rho_{\text{flesh}} \) is the density of herring's flesh of 1071 kg/m\(^3\) \([49]\). The resonant frequency of herring swimbladder \( f_0(z, z_{\text{nb}}, l) \) in Eq. (3.5) is determined by:

61
\[ f_0(z, z_{nb}, l) = \frac{\kappa(\epsilon(z, z_{nb}, l))}{2\pi r(z, z_{nb})} \sqrt{\frac{3\gamma P_{atm}(1 + z/10)}{\rho_{flesh}}}, \]  

(3.7)

where \( \gamma = 1.4 \) is the ratio of the specific heats of air, and \( P_{atm} = 1.013 \times 10^5 \) Pa is the atmospheric pressure, \( \kappa(\epsilon(z, z_{nb}, l)) \) is the swimbladder correction term, and \( \epsilon(z, z_{nb}, l) \) is the swimbladder’s eccentricity. The correction term \( \kappa(\epsilon(z, z_{nb}, l)) \) for a prolate spheroidal swimbladder \([50]\) is given by:

\[ \kappa(\epsilon(z, z_{nb}, l)) = \frac{\sqrt{2}(1 - \epsilon^2(z, z_{nb}, l))^{1/4}}{\epsilon^{1/3}(z, z_{nb}, l)} \left[ \ln \left( \frac{1 + \sqrt{1 + \epsilon^2(z, z_{nb}, l)}}{1 - \sqrt{1 - \epsilon^2(z, z_{nb}, l)}} \right) \right]^{-1/2}. \]  

(3.8)

In Eq. (C.7), the ratio of the minor to major axis of a prolate spheroidal swimbladder \( \epsilon(z, z_{nb}, l) = \left( \frac{c_{sl}/l}{r(z, z_{nb}, l)} \right)^{-3/2} \), and the ratio of the major axis of the swimbladder to the herring’s fork length \( l \) is assumed to be \( c_{sb} \approx 0.364 \) for herring \([9]\). The damping factor \( \eta(z, z_{nb}, l, f) \) in Eq. (3.5) is obtained from:

\[ \frac{1}{\eta(z, z_{nb}, l, f)} = \frac{2\pi r(z, z_{nb}, l)f^2}{c f_0(z, z_{nb}, l)} + \frac{\xi}{\pi r^2(z, z_{nb}, l)f_0(z, z_{nb}, l)\rho_{flesh}}, \]  

(3.9)

where \( f \) is the frequency, \( c \) is the sound speed, and \( \xi \) is the viscosity of herring’s flesh \([49]\).

### 3.2.6 Least squares estimation of herring group parameters

The least-square estimates of the herring group parameters for each shoal segment are determined by minimizing the magnitude of the weighted sum of square differences between the measured and the modeled scattering strengths at frequencies centered at 415, 735, 950, and 1125 Hz. The negative weighted sum of square differences or cost function \( \Delta \) is defined as:

\[ \Delta(z_0, H, z_{nb}, n_A) = -\sum_{k=1}^N \sum_{j=1}^{N_f} \frac{1}{2\sigma_j^2} (SS_{data}^{k,j} - SS_{model}^{k,j}(z_0, H, z_{nb}, n_A, f_j))^2 \]  

(3.10)
where $\sigma_j^2$ is the variance of scattering strength in dB at frequency $f_j$ within the shoal segment, $SS_{jk}^{data}$ is the measured scattering strength level from the $k^{th}$ pixel in the shoal segment at frequency $f_j$, $SS_{model}(z_0, H, z_{nb}, n_A, f_j)$ is the modeled scattering strength level calculated by Eq. (3.4) given herring shoal parameter values, $N$ is the number of pixels containing independent measurements in the shoal segment, $N_f$ is the number of frequencies at which scattering strength of fish is at least one standard deviation greater than the background scattering strength level at frequencies $f_j$ ($2 \leq N_f \leq 4$). The estimators of parameters for the shoal segment $z_0, \hat{H}, \hat{z}_{nb}, \hat{n}_A$ are determined by an exhaustive search over ranges of the parameters such that the magnitude of the weighted sum of square differences $\Delta$ is minimized, which can be expressed as:

$$
\min_{z_0, H, z_{nb}, n_A} |\Delta(z_0, H, z_{nb}, n_A)|. \tag{3.11}
$$

The ranges of parameters $z_0, H, z_{nb}, n_A$ are determined such that herring groups physically stay within the water column with a feasible range of neutral buoyancy depth $z_{nb}$ for a given herring occupancy depth $z$, i.e.:

$$(1 - \mu/100)P_{z_{nb}} \leq P_z \leq (1 + \nu/100)P_{z_{nb}}, \tag{3.12}
$$

where $P_{z_{nb}} = P_{atm}(1 + z_{nb}/10)$ is the hydrostatic pressure at neutral buoyancy depth, $P_{atm}$ is the atmospheric pressure, $P_z = P_{atm}(1 + z/10)$ is the hydrostatic pressure at herring occupancy depth, $\mu = 32$ is the upper bound for percentage difference between $P_z$ and $P_{z_{nb}}$ for positive buoyancy, and $\nu = 100$ is the upper bound for percentage difference between $P_z$ and $P_{z_{nb}}$ for negative buoyancy [51, 52]. The least-squares estimation in log-transformed scattering strength domain in decibels (Eqs. (3.10)-(3.11)) leads to a negligible bias in the mean scattering strength estimates with a large sample size [41], which is the case for this analysis. We restrict our analysis to the acoustic measurements with no data contamination from other ocean acoustic sources including ambient noise, shipping, marine animal vocalizations, and seafloor scattering. To address this issue, only frequency responses exceeding the
mean background level by at least 10 dB were included in the analysis.

### 3.2.7 Estimation uncertainty measure from cost function

Here the estimation uncertainty for herring group parameters $\theta = [z_0, H, z_{nb}, n_A]^T$ is determined by treating the exponential of the cost function $\Delta(\theta)$ as a joint probability distribution. The one-dimensional cost function for each parameter $\theta_k$ is determined here as $\delta_k(\theta_k) = \max \Delta(\theta | \theta_k)$ such that the correlation between the parameters is included in $\delta_k$, where $k = 1, ..., 4$. An estimation uncertainty measure $\eta_k$ for a given parameter $\theta_k$ is then derived as the standard deviation of the parameter $\theta_k$ given the cost function $\delta_k$ for the parameter $\theta_k$, which can be expressed as:

$$\eta_k = \left( \frac{\int \theta_k^2 \exp[\delta_k(\theta_k)] d\theta_k}{\int \exp[\delta_k(\theta_k)] d\theta_k} - \left( \frac{\int \theta_k \exp[\delta_k(\theta_k)] d\theta_k}{\int \exp[\delta_k(\theta_k)] d\theta_k} \right)^2 \right)^{1/2}.$$  

\hspace{1cm} (3.13)

### 3.3 Results

We measure the spatial variation in the frequency response of scattering from large herring groups during an upslope migration towards the Georges Bank spawning ground via multi-spectral resonance sensing with OAWRS (Fig. 3-3). It is found that the frequency at which maximum acoustic scattering occurs consistently decreases with decreasing seafloor depth (Fig. 3-4A). We find this spatial variation in the frequency of the peak scattering is consistent with the gradual decrease in the resonance peak of herring swimbladders during the spawning migration (Fig. 3-4B). We do so by determining herring group parameters via an exhaustive search for an empirical best-fit between the measured and modeled frequency responses of herring scattering, as described in Sec. 3.2.6.

We show the variation in the frequency response of herring scattering as a function of seafloor depth in Fig. 3-5, where the resonance peak frequency decreases with decreasing seafloor depth. The resonance peak frequency of herring is found to be roughly at 1500 Hz outside the frequency band of the current OAWRS measurements.
when herring began their shoal formation process in deeper waters. As the herring groups fully develop and migrate towards the shallower spawning grounds (Figs. 3-4B and 3-5), the resonance peak is fully captured and studied because it moves into the frequency range of OAWRS measurements. Our findings are consistent with the results of a previous study [9], where the resonant scattering from herring was investigated during the initial shoal formation process in the basin area with the seafloor depth of 180-200 m on the northern flank of Georges Bank. In this study, a characteristic 20 dB/octave roll-off of the damped-harmonic-oscillator sub-resonance behavior as the frequency decreased was confirmed by the measured frequency response. The resonance peak frequency was consistently inferred at roughly 1500 Hz outside the

Figure 3-3: (A) Highest measured scattering strength for each spatial position on the northern flank of Georges Bank during shoal formation and subsequent migration towards spawning grounds monitored from 17:25:00 to 21:37:30 EDT on September 29, 2006. Herring begin shoal formation process at locations with seafloor depth of roughly 140 m to 180 m at 17:25:00 EDT, further develop along the bathymetric contour and into the shallower waters, and then stay mostly at locations shallower than 140 m at 21:37:30 EDT. (B) Time evolution of high herring scattering area from 17:25:00 to 21:37:30 EDT on September 29, 2006. The OAWRS sensor location at 21:37:30 EDT on September 29, 2006 is the coordinate origin.
Figure 3-4: (A) Spatial distribution of frequency at which the strongest acoustic scattering occurred within vast herring groups imaged from 17:25:00 to 21:37:30 EDT on September 29, 2006. Gradual decrease in the strongest scattering frequency with decreasing seafloor depths indicates a consistent shift in the resonance frequency of herring scattering during the spawning migration towards the spawning grounds on Georges Bank. (B) The spatial distribution of the resonance peak frequency within vast herring groups imaged from 17:25:00 to 21:37:30 EDT on September 29, 2006.

OAWRS measurement frequency range for the herring groups in this initial stage of the herring spawning process.

The measured spatial variation in acoustic scattering from herring groups is physically expected [44] because the herring group migration involves a significant decrease in occupancy depth by roughly a factor of 3. During this migration, an ambient pressure reduction and a swimbladder volume expansion assuming Boyle’s law result in the reduction in the effective swimbladder stiffness and an increase in the effective mass displaced by the swimbladder vibration. Consequently, the resonance peak
Figure 3-5: The resonance peak of herring scattering consistently shifts from frequencies higher than 1125 Hz to lower than 735 Hz with decreasing seafloor depths. (A) A typical frequency roll-off of roughly 20 dB/octave with decreasing frequencies from the resonance peak frequency at roughly 1500 Hz is found for spatial positions with seafloor depths deeper than 115 m. (B-D) The resonance peaks are found within the measured OAWRS frequency range from 1125 Hz to 415 Hz for the positions with shallower seafloor depths.

frequency $f_0$ should decrease following $f_0 \sim \frac{(1+z/10)^{5/6}}{(1+z_{nb}/10)^{1/5}}$, where $f_0$ is the resonance frequency, $z$ is the herring depth, and $z_{nb}$ is the neutral buoyancy depth. For example, the resonance peak frequency is calculated to shift roughly from 1.5 kHz to 600 Hz when herring migrate from $z = 155 \pm 20$ m with $z_{nb} = 83$ m to $z = 25 \pm 10$ m with $z_{nb} = 20$ m [44].

We find herring groups maintain near-bottom vertical distributions with negative buoyancy throughout the migration (Fig. 3-6). The mean herring group depths are found to follow the seafloor depths during the migration as shown in Figs. 3-6-3-7. For the entire migration path, herring are found to stay close to the seafloor and follow the sloping bottom towards Georges Bank, which is characterized by a high normalized correlation coefficient of greater than 0.9 between the average herring group depth and the corresponding seafloor depth for migration paths along the bathymetric gradient. This is reasonable given that herring would not have a priori knowledge of their exact final spawning positions and the detailed bathymetry along the entire migration path. The vertical distribution of herring obtained here for the deeper water region
Herring distribution (this study)

- Mean herring depth (this study)
- Neutral buoyancy depth (this study)

Herring distribution (Gong et al. [5])

- Neutral buoyancy depth (Gong et al. [5])

0.06

20

40

60

80

100

120

140

160

180

Distance (km)

10 8 6 4 2 0 2 4 6 8 10 12

Figure 3-6: The vertical distribution of herring during the spawning migration and neutral buoyancy as a function of relative distance from 140-m isobath contour along the migration path from 17:25:00 to 21:37:30 EDT on September 29, 2006. Herring are found to occupy roughly the bottom half of the water column with negative buoyancy throughout the migration. The black and red vertical lines represent the spatial variations in the mean shoal depth and neutral buoyancy depth, respectively. The blue color-filled boxes represent the vertical herring distribution for each shoal segment.

is consistent with that measured by downward-directed echosounders in the initial stage of the spawning process in the deeper basin area during the same OAWRS experiment [1], where dense herring groups were found to occupy below the water depths of roughly 100 m at roughly 180-200 m seafloor depth [1, 9].

The near-bottom vertical distribution of herring and negative buoyancy throughout the spawning migration are consistent with herring searching for locations with the environmental conditions optimal for effective spawning such as gravel seafloor, shallow seafloor depth, and bank areas [53, 54]. These environmental parameters for spawning can be better monitored in the vicinity of the seafloor where spawning
takes place. Predation risk may be reduced by staying close to the seafloor when
the structurally complex seafloor with benthic plants and animals provides shelter
availability and decreases predator's efficiency for hunting. In addition, the down-
ward movements towards the seafloor can be made with less energy consumption by
maintaining negative buoyancy during the migration.

We find the depth at which herring are neutrally buoyant is consistently shallower
than the depth herring are occupying such that majority of herring are negatively
buoyant throughout the migration (Fig. 3-6). The decrease in neutral buoyancy
depth during the migration (Fig. 3-6) is consistent with the previously observed gas
release from herring in upward vertical migrations [55, 56], whereas the mechanism
for gas generation at such depths is unclear [9]. The gas bubbles released from herring
will have a negligible impact on OAWRS imaging because the released bubbles with
an estimated mean radius of roughly 1 mm [57] have the resonance peak at frequencies
on the order of a few kHz [57, 58] much higher than the frequency band of the current
OAWRS measurements. The 20 dB/octave rapid roll-off in the sub-resonance scat-
ering from bubbles with decreasing frequency makes the scattering from the released
gas bubbles negligible in the frequency band of the current OAWRS measurements.

3.4 Discussion

The spatial variation in swimbladder resonance investigated here by OAWRS have
enabled fish group depths to be remotely estimated over wide areas spanning thou-
sands of square kilometers. This could not be achieved by conventional fisheries
acoustic line-transect methods for a number of reasons. Conventional line transect
methods, for example, lead to significant temporal and spatial aliasing due to the
Lagrangian particle nature of the sampling, whereas OAWRS provides a full Eulerian
population density field over large areas at any instant. The frequency range of op-
eration is also a key distinguishing factor, since OAWRS systems employ much lower
frequencies than those in conventional fisheries acoustics. Resonant scattering from
many species of swimbladder bearing fish can be investigated in the current OAWRS
Figure 3-7: (A) 3D surface of mean herring group depth determined from multifrequency measurements taken from 17:25:00 to 21:37:30 EDT on September 29, 2006. (B) Seafloor depth of the sloping bottom on the northern flank of Georges Bank. The mean herring group depth and the seafloor depth are found to be highly correlated (0.9 correlation coefficient) for migratory paths along the bathymetric gradient. (C) The horizontal distribution of mean herring group depth. (D) Estimation uncertainty for mean herring group depth in percentage determined by Eq. (3.13). The rest of the estimated herring group parameters and corresponding estimation uncertainties are shown in Fig. B-2.

frequency band of roughly 400 Hz to 2 kHz where the resonant peak can often be found [3, 19], as has been noted in Ref. [3]. In contrast, conventional fisheries acoustic systems operate at much higher frequencies (> 10 kHz) that are typically well above the swimbladder resonance peak of many oceanic fish species. The greatest differences in target strength between and within fish species typically occur at and below swimbladder resonance. Far less variation between and within species occur as frequency increases above swimbladder resonance and geometric scattering issues begin to dominate [59-62].
3.5 Conclusions

Multi-spectral resonance sensing with ocean acoustic remote sensing (OAWRS) is an effective tool to instantaneously image and continuously monitor the 3D group behavioral dynamics of swimbladder-bearing fish in natural habitats over continental-shelf scales. This was demonstrated for vast Atlantic herring groups by instantaneously determining their 3D spatial spawning migration distributions over thousands of square kilometers on the northern flank of Georges Bank. The peak resonance scattering frequency of herring was measured and shown to consistently decrease as herring migrate to shallower regions on the spawning ground. Herring groups were found in the lower half of the water column with mean depth highly correlated (0.9 correlation coefficient) with seafloor depth for migratory paths along the bathymetric gradient. This is consistent with both predator avoidance and the search for optimal seafloor spawning locations, as are our findings that herring maintain negative buoyancy during this final migration to the shallow regions of Georges Bank. Various other oceanic fish species that are known to form dense and extensive behavioral groups can be studied with multi-spectral resonance sensing with OAWRS. The spectral signatures obtained by OAWRS provide key information for the quantification of 3D group behavioral dynamics as well as fish species identification over wide areas. It is further expected that the fish group behavior across species such as predator-prey interaction in 3 spatial dimensions can be revealed over continental-shelf scales by multi-spectral sensing with OAWRS.
Chapter 4

Feasibility of Acoustic Remote Sensing of Large Herring Shoals and Seafloor by Baleen Whales

4.1 Introduction

Recent research has found a high spatial and temporal correlation between certain baleen whale vocalizations and peak annual spawning processes of Atlantic herring (Clupea harengus) in the Gulf of Maine [14, 34], indicating an apparent relationship between these vocalizations and feeding activities of baleen whales. Suggested functions of the vocalizations include communication [63–66], prey manipulation [67], and echolocation [68–71], given the fact that Atlantic herring is a keystone prey species common in the diets of many marine animals including large whales in the Gulf of Maine region [72–74]. Here, feasibility of the possible echolocation function is investigated for large and dense herring aggregations. This differs substantially from the possible ability of baleen whales to detect other whales by active acoustics, which has been previously discussed [75–80].

The approach is to first determine parameters relevant to possible active acoustic sensing in baleen whales, including source and receiver characteristics such as
time-frequency vocalization spectra and auditory system aperture. Detection of large herring shoals and the seafloor is then investigated using theoretical and numerical methods developed for analyzing Ocean Acoustic Waveguide Remote Sensing (OAWRS) of fish shoals [1–3, 9, 19] given these active sensing parameters. Sensing resolution in cross-range is determined by spatial array theory and in range by incoherent energy analysis [1–3, 9, 14, 19, 34, 41], since evidence suggests temporally coherent auditory processing is unlikely [81, 82].

Typical well documented scenarios during the herring spawning season on Georges Bank [1, 9, 14, 19, 34] are investigated to determine ranges up to which detection of large herring shoals and the seafloor is feasible for given whale acoustic parameters. Detection of large herring shoals and the seafloor is found to be highly dependent on local ambient noise conditions at whale locations, herring shoal distribution, baleen whale vocalization parameters such as time duration and source spectra, as well as oceanographic conditions including sound speed structure, bathymetry, and seafloor scattering amplitudes along the acoustic propagation path.

4.2 Materials and Methods

4.2.1 Acoustic Parameters Relevant to Potential Active Acoustic Sensing in Baleen Whales and Corresponding Spatial Resolution

Baleen whale vocalization for potential active sensing is parameterized by call source level $SL$, time duration $T$, center frequency $\bar{f}$, and one-third octave frequency bandwidth $BW_{1/3}$ with center frequency $\bar{f}$ (Table 4.1) following conventions for measuring these parameters in References [14, 34, 63, 64].

For potential active sensing, a baleen whale’s acoustic receiver is parameterized as a two-element spatially coherent array of omni-directional receivers of aperture $L$ corresponding to the separation between ears, which is consistent with the small ear-dimensions to wavelength ratio of vocalizations and the low impedance contrast
between tissue and sea water. The corresponding far-field beam pattern is obtained by spatial array theory (Figure 4-1) as:

\[ B(\sin \theta) = \cos \left( \frac{\pi f L}{c} \sin \theta \right) \] (4.1)

where \( \theta \) is the angle from array broadside, \( f \) is the frequency, and \( c \) is the sound speed. The aperture and frequency are taken from Table 4.1 for the baleen whales considered. The angular resolution of this two element array is best for broadside steering and is given by equivalent beamwidth:

\[ d\theta = 2 \sin^{-1}(\psi/2) \] (4.2)

where

\[ \psi = \int_{-1}^{1} |B(\sin \theta)|^2 d(\sin \theta) = 1 + \frac{c}{2\pi f L} \sin \left( \frac{2\pi f L}{c} \right) \] (4.3)

Cross-range resolution is then determined as \( \rho d\theta \) at array broadside, at horizontal range \( \rho \), for an effectively monostatic active sensing system. No ambiguity in directional sensing due to grating lobes [84] is expected because the grating lobes are located outside real space for given whale acoustic parameters (Table 4.1). Assuming incoherent energy analysis, range resolution \( d\rho \) is \( \frac{cT}{2} \), where the active signal duration \( T \) is given in Table 4.1 for the various baleen whales considered. The areal resolution of the parameterized active sensing system or the smallest area distinguishable by the sensing system, typically referred to as the resolution footprint, is then given by \( \rho d\rho d\theta \) for monostatic geometries (Figure 4-2).

### 4.2.2 Detection of Scattered Returns from Herring Shoals and the Seafloor

The expected magnitude square of the scattered field, \( \langle |\Phi_F|^2 \rangle \), received at location \( r \) from \( N \) omni-directional targets randomly distributed with uniform probability in a vertical layer of thickness \( H \) centered at \( z = z_0 \) within a resolution footprint \( A_R(\rho C) \)
centered at horizontal location $\rho_C$ can be determined from [3, 85]:

$$10 \log_{10} \left( \frac{\Phi_F}{\Phi_{ref}} \right)^2 \approx SL + TLA + TS + 10 \log_{10} \frac{\langle n_A \rangle}{n_{A,ref}}$$  \hspace{1cm} (4.4)

In Equation (4.4), $\Phi_{ref} = 1$ $\mu$Pa is the reference acoustic pressure in water, $SL$ is the whale call source level with mean and standard deviation [14, 34] given in Table 4.1 consistent with circular complex Gaussian random fields fluctuations [41], $TLA$ is the depth-averaged two-way transmission loss to individual targets integrated over one resolution footprint area [3] given in Equation (4.5), $TS = 10 \log_{10} \langle \frac{S}{r_{ref}^2} \rangle^2$ is the expected target strength of a single herring, $S$ is the plane wave scatter function of a single herring, $k$ is the acoustic wavenumber, $r_{ref} = 1$ m is the reference length, $n_{A,ref} = 1$ fish/m$^2$ is the reference areal fish population density, and $\langle n_A \rangle = N/A_R(\rho_C)$ is the expected areal density of the targets within a spatially varying resolution cell centered at horizontal location $\rho_C$. When the instantaneous bandwidth $BW$ of baleen whale vocalizations [14, 34] is greater than the one-third octave bandwidth $BW_{1/3}$ given in Table 4.1, an adjusted whale call source level $SL_{adj} = SL + 10 \log_{10} \frac{BW_{1/3}}{BW}$ is used. The expected target strength of a single herring in a vertical layer is determined by parameters such as mean layer depth, shoal thickness, neutral buoyancy depth, and herring length. With mean and standard deviation of the parameters empirically determined in Reference [9] (Table 4.2), $TS$ is determined using a swimbladder resonance model [43] for deep shoals as described in Appendix C.1. Similar analysis is performed for shallow shoals given measured constraints. The $TLA$ term in Equation (4.4) can be written [3] for monostatic sensing as:

$$TLA = 10 \log_{10} \left( \int_{A_R(\rho_C)} \frac{1}{H} \left( \int_{z_0-H/2}^{z_0+H/2} \right) (4\pi)^4 \langle |G(r|\rho_S,z_S;f,c(r_w),d(r_w))|^4 |\rho_S,z_S \rangle dz_Sd\rho_S^2/r_{ref}^2 \right)$$  \hspace{1cm} (4.5)

where $G(r|\rho_S,z_S;f,c(r_w),d(r_w))$ is the Green function between the target location $r_S = (\rho_S,z_S)$ and the receiver location $r$, and $c(r_w)$ and $d(r_w)$ are the sound speed and the density in the water column at any point $r_w$ in the propagation path, respectively.
The areal integration in Equation (4.5) should be taken over the intersection between resolution footprint and the area occupied by herring when the herring distribution does not fully overlap with the resolution footprint. A parabolic equation model [39] is used to calculate the Green function from the whale location to the herring shoal in a range-dependent Gulf of Maine environment. The conditional expectation over the sound speed and the density in Equation (4.5) is determined by averaging 10 Monte-Carlo realizations, where the Green functions are calculated along the entire propagation path in range and depth for each realization. Each Monte-Carlo realization employs a different sound speed profile measured during the OAWRS 2006 experiment [9] every 500 m [38, 40] along the propagation path.

The expected magnitude square of the scattered field \( \langle |\Phi_S|^2 \rangle \) from the seafloor is determined from [19]:

\[
10 \log_{10} \left( \frac{|\Phi_S|}{\Phi_{ref}} \right)^2 \approx SL + 10 \log_{10} \left( \iiint_{V_R(r_C)} (4\pi)^2 \langle |G(r|r_S; f, c(r_w), d(r_w))| dr_s dr_w \rangle \frac{r_s^2}{r_{ref}^3} \right) + 10 \log_{10} \left( \langle |A_S(f, \Gamma, \Gamma_d, V_C)|^2 \rangle / r_{ref}^{-1} \right)
\]

where \( SL \) is the whale call source level, \( V_R(r_C) \) is the volume of the resolution cell centered at horizontal location \( r_C \), \( G(r|r_S; f, c(r_w), d(r_w)) \) is the Green function between the target location \( r_S = (\rho_S, z_S) \) and the receiver location \( r, c(r_w) \) and \( d(r_w) \) are the sound speed and the density at any point \( r_w \) in the propagation path, respectively. The Green function and the conditional expectation over the water column sound speed and the density in Equation (4.6) are determined by following the same procedures given for those in Equation (4.5). The expected magnitude square of Rayleigh-Born seafloor scattering amplitude per coherence volume \( \langle |A_S(f, \Gamma, \Gamma_d, V_C)|^2 \rangle \) is defined as [19]:

\[
\langle |A_S(f, \Gamma, \Gamma_d, V_C)|^2 \rangle = k^4 V_C \left[ Var(\Gamma) + F_d Var(\Gamma_d) + F_c Cov(\Gamma, \Gamma_d) \right]
\]
where $\Gamma_{\kappa}$ is the fractional change in seafloor compressibility, $\Gamma_d$ is the fractional change in seafloor density, $V_C$ is the coherence volume of inhomogeneities, and $F_{\kappa}$ and $F_d$ are the proportionality constants of scattering contributions from dipole and cross terms to monopole scattering contributions, respectively. The frequency dependence of the expected magnitude square of Rayleigh-Born seafloor scattering amplitude per coherence volume $\langle |A_S(f, \Gamma_{\kappa}, \Gamma_d, V_C)|^2 \rangle$ is measured during the OAWRS 2006 experiment [19] as:

$$\langle |A_S(f, \Gamma_{\kappa}, \Gamma_d, V_C)|^2 \rangle = C \times f^q$$

(4.8)

where $C \approx 1.78 \times 10^{-8} \text{ m}^{-1} \text{Hz}^{-q}$ is the proportionality constant, $f$ is the frequency, and $q \approx 2.12$ is the power law coefficient of frequency dependence of the expected magnitude square of the seafloor scattering amplitude. The level of magnitude square of Rayleigh-Born seafloor scattering amplitude per coherence volume is assumed to vary with standard deviation up to 2 dB as empirically determined in Reference [19].

The expected magnitude square of ambient noise field is determined as [14, 34]:

$$\langle |\Phi_N|^2 \rangle \approx \frac{1}{\tau} \int_{BW_{1/3}} |\Phi_{N, data}(f)|^2 df \approx \frac{1}{\tau} |\Phi_{N, data}(\bar{f})|^2 \times BW_{1/3}$$

(4.9)

where $\tau$ is the measurement time, $f$ is the frequency, $\Phi_{N, data}(f)$ is the Fourier transform of the measured ambient noise field time series, $\Phi_{N, data}(\bar{f})$ is the Fourier transform of the measured ambient noise field time series evaluated at the center frequency $\bar{f}$ of baleen whale vocalizations given in Table 4.1, and $BW_{1/3}$ is the one-third octave frequency bandwidth of various baleen whale vocalizations given in Table 4.1. The expected magnitude square of the beamformed ambient noise field $\Phi_{N,BF}$ is determined from:

$$10 \log_{10} \left( \frac{\langle |\Phi_N|^2 \rangle}{\langle |\Phi_{N,BF}|^2 \rangle} \right) = AG$$

(4.10)

where $AG$ is the array gain of the receiver array given in Table 4.1. The wind-speed dependence of ambient noise is modeled as [34]:
\[
\left\langle \left| \frac{\Phi_N}{\Phi_{ref}} \right|^2 \right\rangle = \alpha v^n + \beta
\] (4.11)

where \( \Phi_{ref} = 1 \mu Pa \) is the reference acoustic pressure in water, \( v \) is the wind speed, \( n \) is the power law coefficient of wind-speed-dependent ambient noise, \( \alpha \) is the waveguide propagation factor [83], and \( \beta \) is the constant baseline ambient noise intensity. The coefficients \( n, \alpha, \) and \( \beta \) given in Table 4.1 are empirically determined [34] by least-square fitting between the measured and the modeled ambient noise levels as a function of measured wind speed during the OAWRS 2006 experiment [1] (Figure C-1). The instantaneous measurements of ambient noise vary with 5.6 dB standard deviation [41].

For target detection under seafloor scattering limited detection conditions, the expected magnitude square of the scattered field from herring shoals should be greater than that of the scattered field from the seafloor by a factor that corresponds to detection threshold \( DT \) given in Table 4.1:

\[
10 \log_{10} \frac{\left\langle |\Phi_F|^2 \right\rangle}{\left\langle |\Phi_S|^2 \right\rangle} \geq DT
\] (4.12)

where the maximum detection range is found when the equality holds. Similarly for target detection under noise limited conditions, the expected magnitude square of the scattered field from herring shoals should be greater than that of the beamformed ambient noise field \( \Phi_{N,BF} \) by a factor that corresponds to \( DT \) given in Table 4.1:

\[
10 \log_{10} \frac{\left\langle |\Phi_F|^2 \right\rangle}{\left\langle |\Phi_{N,BF}|^2 \right\rangle} \geq DT
\] (4.13)

where the maximum detection range is found when the equality holds. The detection threshold \( DT \) is determined by requiring that the scattered returns from fish targets stand at least one standard deviation [41, 42] above background levels from seafloor scattering or ambient noise.
4.2.3 OAWRS Experiment during Peak Herring Spawning Processes in the Gulf of Maine in Fall 2006

An OAWRS experiment was conducted on the northern flank of Georges Bank during the peak annual Atlantic herring spawning period in the Fall of 2006 [1, 9, 14]. The spatial distribution of herring was instantaneously imaged and continuously monitored [1]. The spawning process was found to follow a regular diurnal pattern, where herring were diffusely scattered near the seafloor during daylight hours, formed dense and large shoals near the 180-200 m depth contours near sunset, and then migrated synchronously towards spawning grounds at depths less than 50 m in the evening as shown in Figure 4-3A [1].

Parabolic equation transmission loss modeling [39] was calibrated and verified with roughly one hundred two-way transmission loss measurements made from calibrated targets with known scattering properties [86] and thousands of one-way transmission loss measurements [9, 42] during the same experiment. Herring target strength was determined by empirical fit to a resonance model [43] over multi-frequencies in ranges relevant to this study [9]. Thousands of baleen whale vocalizations were passively recorded and localized in the vicinity of large herring shoals and call parameters, as shown in Table 4.1, were determined [14, 34]. Ambient noise levels that include wind-driven and shipping noise contributions were measured and fit to empirical models at baleen whale vocalization frequencies [14, 34].

4.3 Results

We find acoustic detections of herring shoals are physically feasible up to $10 \pm 6$ km in range for humpback (Megaptera novaeangliae) parameters and $1 \pm 1$ km for minke (Balaenoptera acutorostrata) parameters in common herring spawning scenarios of dense near-surface concentrations at 15–35 m depth, but not for herring concentrations at 130–180 m depth even at zero horizontal range (Figures 4-4 and 4-5). This assumes detections are possible when scattered returns from herring shoals stand at
least one standard deviation above mean ambient noise and seafloor scattering levels, and the measured ambient noise conditions, herring shoal distributions, baleen whale time-frequency vocalization spectra, and geophysical parameters of the ocean waveguide in the Gulf of Maine during the 2006 herring spawning period. The detection range of 1 ± 1 km, where spherical spreading dominates, suggests that the corresponding minke parameters are better suited to direct-path waterborne rather than waveguide acoustic sensing. For fin (Balaenoptera physalus) and blue (Balaenoptera musculus) parameters, we find detection of herring shoals is not feasible even at zero horizontal range (Figures 4-4 and 4-5) for any of the shoal depths considered because the scattered intensity from the shoals is at least two orders of magnitude lower than those from the seafloor.

The observed variations in detection range are most strongly affected by the frequency dependence of resonant scattering from herring swimbladders. For shallow shoals, whale frequency ranges closer to the resonant scattering peak lead to greater detection ranges (Figures 4-3B). For deep shoals, all whale frequency ranges considered are too far below the resonant scattering peak to lead to detections. As herring shoals migrate to the south from the deep locations on the northern flank of Georges Bank to shallower than 50 m on Georges Bank (Figure 4-3A), it is found that the resonant frequency of herring swimbladders shifts to lower frequencies from roughly 1.5 kHz to roughly 600 Hz (Figure 4-3B). Herring target strength at baleen whale vocalization frequencies then increases by roughly two orders of magnitude (Figure 4-3B), which is the key factor enabling active acoustic detection of herring shoals for minke and humpback parameters.

We find seafloor detection is physically feasible up to 2 ± 1 km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters. This assumes scattered returns from the seafloor stand at least one standard deviation above ambient noise levels (Figures 4-4 and 4-5). This suggests that the baleen whales share a common acoustic sensation of rudimentary features of the geophysical environment.
4.4 Discussion

Little information exists on the mechanisms by which baleen whales detect prey [82, 87, 88]. Since the current analysis does not examine behavioral data, no conclusions can be drawn about whether the baleen whales considered actually employ active sensing to detect herring shoal prey. The analysis merely considers the physical feasibility of such sensing from an acoustic detection perspective. It is interesting, however, that the whale vocalization spectra do not appear to be optimized to take advantage of the peak resonant target strength of the deeper herring shoals, and only the humpback and perhaps the minke spectra are situated at or near the resonant peak for the shallower shoals.

Our analysis shows that local ambient noise is the primary limiting factor in detection of herring shoals for humpback and minke parameters as well as seafloor detections for the parameters of any whale considered (Figures 4-4 and 4-5). It is possible that detections of herring shoals would then be feasible at much longer ranges under ocean ambient noise conditions without significant shipping traffic [89, 90]. For example, under ambient noise conditions reported for Arafura and Timor Seas in Australia [89] with insignificant shipping, detections of herring shoals would be feasible up to 10-30 km in range for humpback parameters and 2-5 km for minke parameters (Figures 4-4). Similarly, detections of the seafloor would be feasible up to 5-8 km in range for blue parameters, 4-9 km for humpback parameters, 2-4 km for fin parameters, and 1-3 km for minke parameters (Figures 4-4 and 4-5) under ocean ambient noise conditions without significant shipping. These estimates could represent possible detection ranges in ancient ocean environments that lacked shipping noise contributions.

4.5 Conclusions

The feasibility of acoustic remote sensing of large herring shoals and the seafloor by baleen whales has been investigated. We have found that it is physically feasible to
detect dense near-surface herring shoals up to $10 \pm 6$ km in range with humpback acoustic parameters and $1 \pm 1$ km with minke acoustic parameters over ambient noise and seafloor scattering levels in the Gulf of Maine continental shelf environment. In contrast, detections of herring shoals are found to be unfeasible for fin and blue whale parameters even at zero horizontal range because scattered intensity from the seafloor are expected to be greater than the scattered intensity from herring shoals. These variations in detection range are primarily due to the strong frequency dependence of resonant scattering from herring swimbladders. We have found that detections of the seafloor are feasible up to $2 \pm 1$ km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over ambient noise, suggesting that the whales share a common acoustic sensation of rudimentary features of the geophysical environment.
Figure 4-1: Beam patterns of two-element coherent array steered at broadside given baleen whale acoustic parameters given in Table 4.1.
Table 4.1: Parameters of baleen whale’s auditory reception and signal detection used for detection range estimation in the Gulf of Maine scenario

<table>
<thead>
<tr>
<th>Baleen whale species</th>
<th>SL (dB re 1 μPa 1 m)</th>
<th>T (s)</th>
<th>f (Hz)</th>
<th>BW(_{1/3}) (Hz)</th>
<th>DT (dB)</th>
<th>AG (dB)</th>
<th>L (m)</th>
<th>α</th>
<th>β</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fin</td>
<td>189±5.6</td>
<td>0.8</td>
<td>20</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>3.0</td>
<td>1.88(\times 10^8)</td>
<td>6.68(\times 10^5)</td>
<td>0.6</td>
</tr>
<tr>
<td>Blue</td>
<td>189±5.6</td>
<td>2</td>
<td>40</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>3.0</td>
<td>1.19(\times 10^8)</td>
<td>2.51(\times 10^6)</td>
<td>0.4</td>
</tr>
<tr>
<td>Minke</td>
<td>179±5.6</td>
<td>0.1</td>
<td>315</td>
<td>72.5</td>
<td>3</td>
<td>1.4</td>
<td>1.5</td>
<td>5.96(\times 10^6)</td>
<td>3.76(\times 10^4)</td>
<td>0.6</td>
</tr>
<tr>
<td>Humpback</td>
<td>180±5.6</td>
<td>1.44</td>
<td>450</td>
<td>100</td>
<td>3</td>
<td>3.7</td>
<td>2.0</td>
<td>1.26(\times 10^6)</td>
<td>5.01(\times 10^6)</td>
<td>1.1</td>
</tr>
</tbody>
</table>

SL is the whale call source level [14], T is the time duration, \(\bar{f}\) is the center frequency of baleen whale vocalizations relevant for target detection, BW\(_{1/3}\) is the one-third octave bandwidth centered at frequency \(\bar{f}\), DT is the detection threshold, AG is the array gain given whale acoustic parameters, L is the spatial distance between the two omni-directional receiver with spatial coherence, α is the waveguide propagation factor [83], β is the constant baseline ambient noise intensity, and n is the power law coefficient of wind-speed-dependent ambient noise. The parameter values for SL are given in the form of mean ± standard deviation.

Table 4.2: Parameters of herring shoals

<table>
<thead>
<tr>
<th>Herring shoals</th>
<th>(z_0) (m)</th>
<th>(H) (m)</th>
<th>(z_{nb}) (m)</th>
<th>(l) (cm)</th>
<th>(\langle n_A \rangle) (fish/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep</td>
<td>155±6.5</td>
<td>50±15.6</td>
<td>83±31.3</td>
<td>24.2±1.7</td>
<td>5.0</td>
</tr>
<tr>
<td>Shallow</td>
<td>25±1.0</td>
<td>20±6.2</td>
<td>20±6.5</td>
<td>24.2±1.7</td>
<td>1.0</td>
</tr>
</tbody>
</table>

\(z_0\) is the mean depth of herring layer, \(H\) is the shoal thickness, \(z_{nb}\) is the neutral buoyancy depth of herring, \(l\) is the fork length of herring, and \(n_A\) is the areal population density of herring shoals. The parameter values for \(z_0\), \(H\), \(z_{nb}\), and \(l\) are given in the form of mean ± standard deviation.
Figure 4-2: Top view (A) and side view (B) of the geometry for possible acoustic remote sensing in a continental shelf environment by baleen whales. The resolution footprint given baleen whale acoustic parameters is an annular sector with range extent $d\rho$ and cross-range extent $\rho d\theta$. 
Figure 4-3: (A) Migration of herring shoals from depths greater than 100 m to less than 50 m in the Gulf of Maine; (B) Corresponding frequency dependence of resonance scattering of herring swimbladders for deep and shallow shoals. Blue and red solid lines represent the herring target strength when herring layer is at depths $z_{herring}$ of 130–180 m with a neutral buoyancy depth $z_{nb}$ of 83 m on the northern flank of Georges Bank and at depths $z_{herring}$ of 15–35 m with a neutral buoyancy depth $z_{nb}$ of 20 m on Georges Bank, respectively.
Figure 4-4: Detections of herring shoals are feasible up to 10 ± 6 km in range for humpback parameters and 1 ± 1 km for minke parameters in common Gulf of Maine herring spawning scenarios of dense near-surface concentrations at depths of 15–35 m (red solid lines). Detections of the seafloor are found to be feasible up to 2 ± 1 km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over local ambient noise. Dark red dashed lines represent seafloor scattering for a seafloor depth of 50 m. Purple solid lines represent the beamformed ambient noise levels with mean wind speed of 5.6 m/s during peak herring spawning period in the Gulf of Maine with one detection threshold (DT) added. Light purple shaded areas range from the beamformed ambient noise levels at 2 m/s wind speed minus 5.6 dB to those at 10 m/s wind speed plus 5.6 dB, to illustrate the typical range of wind-speed-dependent ambient noise encountered. Gray solid lines represent the beamformed ambient noise level without significant shipping with wind speed of 5.1 m/s [89] with one DT added. Light gray shaded areas range from beamformed ambient noise levels without significant shipping at 5.1 m/s wind speed plus/minus the instantaneous intensity standard deviation of 5.6 dB. The beamformed ambient noise levels are determined for frequency bands of baleen whale vocalizations given in Table 4.1 using Equations (4.9) and (4.10).
Figure 4-5: Detections of herring shoals at depths of 130–180 m (blue solid lines) are found to be unfeasible for the parameters of any baleen whale considered. Detections of the seafloor are found to be feasible up to 2 ± 1 km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over local ambient noise. Dark blue dashed lines represent seafloor scattering for a seafloor depth of 180 m. Purple solid lines represent the beamformed ambient noise levels with mean wind speed of 5.6 m/s during peak herring spawning period in the Gulf of Maine with one detection threshold (DT) added. Light purple shaded areas range from the beamformed ambient noise levels at 2 m/s wind speed minus 5.6 dB to those at 10 m/s wind speed plus 5.6 dB, to illustrate the typical range of wind-speed-dependent ambient noise encountered. Gray solid lines represent the beamformed ambient noise level without significant shipping with wind speed of 5.1 m/s [89] with one DT added. Light gray shaded areas range from beamformed ambient noise levels without significant shipping at 5.1 m/s wind speed plus/minus the instantaneous intensity standard deviation of 5.6 dB. The beamformed ambient noise levels are determined for frequency bands of baleen whale vocalizations given in Table 4.1 using Equations (4.9) and (4.10).
Chapter 5

Conclusions

In this thesis, we have experimentally determined biologically important but previously not accessible quantities such as the mean spawning shoal population for oceanic fish species, the depth information for large fish shoals, and fish swimbladder physiology, which enabled us to study the mechanisms for fish population collapse or recovery and fish behavior during the annual spawning migration. The essential observational tool enabling these analyses is the continental-shelf-scale ocean-acoustic sensing, known as OAWRS, capable of instantaneously imaging and continuously monitoring the ocean environment over thousands of square kilometers.

In Chap. 2, we have empirically shown that two distinct species of oceanic fish at pre-industrial population levels quantize into vast groups of stable species-dependent mean population to cooperatively reproduce. These two species, Atlantic cod and Atlantic herring, bound the spectrum from large upper-trophic predators to mid-trophic keystone species, and from batch to total capital spawning. To be sustained at pre-industrial levels, we find the total spawning population must greatly exceed the mean spawning group size found at pre-industrial levels for any oceanic fish population we investigated. An apparent consequence of this large differential is that in regions where total populations declined below this mean species-dependent quantum, return to pre-industrial levels has not yet occurred or required decades. We have demonstrated this empirically for eight distinct major populations across the North-Atlantic over roughly the last century. While these results currently are only
for the populations we investigated, they may apply to similar oceanic fish species that form vast spawning shoals. This is significant because many of the abundant oceanic fish species that form vast spawning groups are of critical importance to ocean ecosystems.

In Chap. 3, we have demonstrated that the multi-spectral resonance sensing with OAWRS can be used as an effective tool to instantaneously image and continuously monitor the 3D behavioral dynamics of swimbladder-bearing fish group behavior in natural habitats over continental-shelf scales. We have done so for vast Atlantic herring groups during annual spawning migration on the northern flank of Georges Bank by inferring the 3D spatial distributions from multi-spectral OAWRS measurements. This was possible because the spatial variation in the frequency response of scattering from herring was captured instantaneously over continental-shelf scales via multi-spectral resonance sensing with OAWRS. The resonance peak frequency of herring swimbladders is found to consistently decrease as herring migrate towards the shallower spawning ground. We have found that herring maintain near-bottom vertical distributions with negative buoyancy throughout the spawning migration. The herring groups' near-bottom vertical distribution is characterized by a high correlation greater than 0.9 between the mean herring group depth and corresponding seafloor depth for migratory paths along the bathymetric gradient. Our findings are consistent with the seasonal primary activity of herring seeking locations with optimal environmental parameters for spawning.

In Chap. 4, the feasibility of acoustic remote sensing of large herring shoals and the seafloor by baleen whales has been investigated. We have found that it is physically feasible to detect dense near-surface herring shoals up to 10 ± 6 km in range with humpback acoustic parameters and 1 ± 1 km with minke acoustic parameters over ambient noise and seafloor scattering levels in the Gulf of Maine continental shelf environment. In contrast, detections of herring shoals are found to be unfeasible for fin and blue whale parameters even at zero horizontal range because scattered intensity from the seafloor are expected to be greater than the scattered intensity from herring shoals. These variations in detection range are primarily due to the
strong frequency dependence of resonant scattering from herring swimbladders. We have found that detections of the seafloor are feasible up to $2 \pm 1$ km in range for blue and humpback parameters and roughly 1 km for fin and minke parameters over ambient noise, suggesting that the whales share a common acoustic sensation of rudimentary features of the geophysical environment.
Appendix A

Supplementary Information for
Group Behavioral Quantization of
Oceanic Fish

A.1 Acoustic sensing methods

Two research vessels were employed to measure cod in the Nordic Seas 2014 Experiment. Both the RV Knorr and RV Johan Hjort collected vertical echosounding data along line transects. The beamwidth of the vertical echosounder deployed from RV Knorr was 35.0° at 12 kHz, which yields a 60-m diameter resolution footprint at 100-m depth where many of the fish groups we imaged were concentrated, while that for RV Johan Hjort was 6.8° at 38 kHz, which yields a 12-m diameter resolution footprint at 100-m depth.

The RV Knorr towed the OAWRS source and receiver arrays (Fig. A-1) and collected OAWRS data. OAWRS images were produced by beamforming, matched filtering, and charting the received OAWRS returns and then correcting for transmission loss [1–3, 91, 92]. In the Nordic Seas 2014 Experiment, an effectively monostatic arrangement for acoustic measurements was used where the source and receiver arrays were towed by the same research vessel (Fig. A-1). Linear frequency modulated
(LFM) source waveform transmissions of 50 Hz bandwidth and 1 second duration centered at 955 Hz were sent every 50 seconds for the OAWRS cod measurements presented. Similar information about the Gulf of Maine 2006 herring measurements presented appear in Ref. [1].

A.2 Cod areal population density obtained by OAWRS

OAWRS cod population density distributions \( n_{A,OAWRS} \) were determined from measured OAWRS intensities converted to scattering strength \( SS_{OAWRS}(\rho_C) \) at horizontal location \( \rho_C \) following the standard procedures described in Refs. [3, 9, 91, 92]. This included averaging over 5 consecutive, instantaneous OAWRS images and two consecutive 15-m range cells to reduce the measured OAWRS intensity standard deviation to roughly 1 dB or 25 percent [1, 3, 9, 41, 42, 91, 92]. Stochastic propagation modeling [3, 9, 19, 34] was performed to obtain expected transmission loss [3, 9, 19, 34] as a function of \( \rho_C \). This was done by averaging 100 parabolic equation [39] realizations of the waveguide intensity field each employing a unique sound speed profile measured during the Nordic Seas 2014 Experiment (Fig. A-2) at 500 m increments [40] along the propagation path. It led to a standard deviation of roughly 1 dB in transmission loss over the group thickness. Mean cod target strength \( T_{SOAWRS} \) was determined at OAWRS frequencies in the Lofoten region during the 2014 spawning season. In particular, \( T_{SOAWRS} \) was found to be -25.9 dB re 1 m (Fig. A-3) by equating areal population density for nearly simultaneous OAWRS \( n_{A,OAWRS}(\rho_C) \) and RV Knorr 12 kHz echosounder \( n_{A,Knorr}(\rho_C) \) measurements in the Andenes region via \( T_{SOAWRS}(\rho_C) = SS_{OAWRS}(\rho_C) - 10 \log_{10}(n_{A,Knorr}(\rho_C)) \) and averaging the result over 116 independent samples or resolution cells, following Refs. [3, 9]. This led to a \( T_{SOAWRS} \) error of 0.12 dB or 3 percent. The resulting OAWRS cod population density estimate was found to have a standard deviation of roughly 1-2 dB per pixel, which is within 1 dB of that obtained by resonance swimbladder modeling [19, 43] for the mean cod spawning length measured for the Lofoten region in 2014 (Fig. A-4). This \( T_{SOAWRS} \) value of -25.9 dB re 1 m was used to estimate OAWRS cod areal population density.
throughout the Lofoten region for the 2014 spawning season.

A.3 Critical boundary population density of cod spawning groups

The critical population density defining the boundary of a gigantic cod shoal was consistently determined to be 0.016 fish/m² by a number of independent methods including: (1) scaling the critical density consistently obtained experimentally with OAWRS for gigantic group formation in another oceanic fish species [1] to account for body length ratios; (2) segmenting regions of high population density and gradient in OAWRS images; (3) segmenting regions of high population density and gradient in corresponding high-resolution line-transect vertical echosounder measurements; (4) least squares estimation with decades of line-transect survey data of cod in the Northeast Arctic spawning ground.

The cod and herring spawning groups we study are horizontally extensive, stretching for tens of kilometers, and vertically thin, roughly one thousand times less than their horizontal extent. We find that the mean distance between nearest neighbors in the cod spawning groups we studied is 5 times greater in the horizontal than in the vertical, from high-resolution echo-sounding data. This makes the vertical separation component negligible when determining the magnitude of the separation vector between any two cod on average. The herring spawning group formation process and attainment of critical density has been shown to occur as herring leave the effectively 2-D horizontal surface of the seafloor. As with the cod, this leads to a mean horizontal separation much greater than the mean vertical separation for the herring groups [1]. Additionally, a number of studies have consistently found that the distance between fish in groups is proportional to fish body length [93–95]. Given these facts, critical density for the horizontally extensive and vertically thin spawning groups of this study is expected to scale in proportion to fish body length squared. The separations between cod were determined from the positions of more than 150 cod individuals
in the Andenes spawning group (Fig. 2-1C). They were detected in a roughly 3 km transect through the spawning group at 02:40 UTC to 03:00 UTC on March 8, 2014 where the volume scattering coefficient ($s_v$) was greater than -42 dB re 1 m$^{-1}$. All detections were within the RV Johan Hjort 38 kHz echosounder’s 6.8° main beam.

The boundary population density of a cod spawning group of 0.016 fish/m$^2$ is found to be the expected value of the critical density for cod spawning group formation based on thousands of independent and consistent OAWRS measurements of group formation in Atlantic herring [1, 3, 9, 14, 34], which follow basic group behavioral principles [96]. Since the separation between fish in groups is found to scale with body length [93–95] and the mean vertical separation between fish in the cod and herring groups studied is negligible compared to the mean horizontal separation, the critical density for cod group formation is found to be

$$c_{\text{cod}} = c_{\text{herring}} \left( \frac{L_{\text{herring}}}{L_{\text{cod}}} \right)^2,$$

(A.1)

by scaling the critical density for herring group formation [1] by the squared ratio of herring to cod lengths where $L_{\text{herring}}$ is the herring body length in the Fall 2006 Georges Bank spawning ground [9], and $L_{\text{cod}}$ is the cod body length measured in the Winter-Spring 2012-2014 Northeast Arctic cod spawning area (Fig. A-4).

A simple contouring of population density at the critical density $c_{\text{cod}} = 0.016$ fish/m$^2$ in each wide-area OAWRS image, Figs. 2-1B and 2-1C, led to segmentation of the high population density region that defined the two gigantic shoals shown. More than 90% of the detected cod fell within the closed boundary contour determined by the critical density $c_{\text{cod}} = 0.016$ fish/m$^2$ in these images (Figs. 2-1B and 2-1C). The group boundary can also be detected from the spatial population density gradient, which is 5 times larger on the critical density contour than in the surrounding low density regions on average. The minimum cod population density detectable by OAWRS was roughly 0.005 fish/m$^2$, where seafloor scattering mechanisms become a contaminant [3, 19]. The population density standard deviation is roughly 1-2 dB [91]. Raw OAWRS images are each comprised of hundreds of thousands of indepen-
dent pixels since the OAWRS system has 15-m range resolution, and forms at least 64 independent horizontal beams, spanning ranges of many tens of kilometers. The angular resolution of any beam away from endfire varies as $\lambda/L \cos \theta$, where $\lambda$ is the acoustic wavelength, $L$ the receiver array aperture length and $\theta$ the horizontal angle from array broadside [91, 92]. In the raw OAWRS images corresponding to Figs. 2-1B and 2-1C respectively, within the shoal boundary, for example, there are roughly 10,000 and 300,000 independent measurements or resolution cells, respectively and on the contour defining the shoal boundary, roughly 3,000 and 6,000 independent measurements respectively. So the OAWRS imagery alone provided thousands of independent samples to obtain the cod critical density. To reduce the standard deviation of Gaussian field fluctuations, it is conventional for 5 consecutive instantaneous raw OAWRS images to be averaged at each pixel with the additional averaging of two pixels in range to further reduce the standard deviation of Gaussian field fluctuations by $1/\sqrt{10}$ [41, 42]. This was done for the OAWRS images presented Figs. 2-1B and 2-1C. Some additional processing was done as described later for the Fig 2-1B mosaic.

Similar results were found with high-resolution vertical echosounder measurements by the RV Johan Hjort which was directed to cross the cod group found by OAWRS in Fig. 2-1C. Again more than 90% of the cod detected with this echo-sounding data were within the region exceeding the 0.016 fish/m² critical density which corresponded to the shoal boundary identified by OAWRS. The population density gradient at this critical density was two orders of magnitude larger than that found on average in the bordering low density regions by vertical echo-sounding. The RV Johan Hjort made more than 18,000 independent measurements of cod at and within the spawning group boundary shown in Fig. 2-1C. It employed a 38 kHz high-resolution vertical echosounder system which has roughly 0.2-m range resolution, and forms a 6.8° main beam in the vertical direction.

The empirically derived critical density for cod group formation (Eq. A.3) is also independently found to have the best least squares consistency with the annual spawning population time series and decades of line-transect survey data as shown in Fig. S5, under the assumption that summing the populations of spawning groups will well
approximate the total annual spawning population each year. An unrealistically high value for group critical density will incorrectly exclude too much of the population. An unrealistically low value will incorrectly include background levels that have no cod information only noise. This consistency was found by searching for a threshold density $n_{A,T}$ at which the $i$-th year’s cod spawning population estimate $S_{i,Q}(n_{A,T})$ best-matches $i$-th year’s cod spawning population inferred by International Council for the Exploration of the Sea (ICES), $S_{i,ICES}$, over a 30-year period from 1984 to 2014 (Fig. A-5), i.e.

$$\min_{n_{A,T}} \sum_{i=1}^{M} (S_{i,ICES} - S_{i,Q}(n_{A,T}))^2,$$

(A.2)

where $M$ is the number of years. In Eq. A.3, the $i$-th year’s cod spawning population estimate $S_{i,Q}(n_{A,T})$ is determined by multiplying the number of mean spawning groups in $i$-th year $N_{i}(n_{A,T})$ by the mean spawning group size $Q(n_{A,T})$ for given threshold density $n_{A,T}$ as

$$S_{i,Q}(n_{A,T}) = N_{i}(n_{A,T})Q(n_{A,T}) = N_{i}(n_{A,T})\frac{1}{M} \sum_{k=1}^{M} \frac{S_{k,ICES}}{N_{k}(n_{A,T})}.$$  

(A.3)

The mean spawning group size $\langle Q \rangle$ is then determined at the critical density for cod group formation $c_{cod}$ by

$$\langle Q \rangle = Q(c_{cod}) = \frac{1}{M} \sum_{i=1}^{M} \left. \frac{S_{i,ICES}}{N_{i}(n_{A,T})} \right|_{n_{A,T}=c_{cod}}.$$  

(A.4)

### A.4 Consistency between OAWRS and line-transect vertical echosounder measurement

Near Andenes Norway (Fig. A-6), 12 kHz RV Knorr local vertical echosounder line-transect measurements were made simultaneously with OAWRS and showed consistent trends in cod population (Fig. A-6). This consistency was also found in the
Røst region where RV Knorr vertical echosounder measurements were made simulta-
neously with OAWRS measurements (Figs. A-7A-D), and then roughly 5 hours (Figs. 
A-7E-H) and roughly 24 hours after OAWRS imaging (Figs. A-7I-L). Cod layers were 
consistently found at their preferred spawning depths near the 100 m contour where 
the transition layer between warm Atlantic waters and cold Arctic waters is typically 
found in the Lofoten region [33, 97] (Figs. A-7A-L). Cod density distributions near 
Røst were temporarily stable over a roughly one-month period (Fig. A-8), as seen from 
the consistency between cod density distributions obtained by OAWRS on February 
23, 2014 and by the annual line-transect vertical echosounder survey of March 17-31, 
2014. The high cod density areas near Røst and general tapering trends radially 
away from Røst are clear and consistent in both OAWRS and vertical echosounder 
measurements.

**A.5 Cod areal population density obtained by line-transect vertical echosounding**

Cod areal population density was determined from RV Johan Hjort 38 kHz echosounder 
measurements near Andenes by:

\[
n_{A,JH} = \frac{1}{\sigma_{bs}} \int_{z_1}^{z_2} s_v dz,
\]

(A.5)

where \(n_{A,JH}\) is the cod areal population density along RV Johan Hjort line transects, 
\(s_v\) is the calibrated volume backscattering coefficient in \(m^{-1}\) [98] at 38 kHz from RV 
Johan Hjort, \(z_1\) and \(z_2\) are the measured depth bounds of the group’s vertical extent, 
which ranged roughly between 30 m and 100 m near Andenes, \(\sigma_{bs} = 10^{TS_{38kHz}/10}\) is 
the mean backscattering cross section of an individual fish at 38 kHz in units of \(m^2\), 
and \(TS_{38kHz}\) is the empirical cod target strength at 38 kHz [48] corresponding to the 
mean cod length \(L\) measured in the region in 2014 (Fig. A-4):

\[
TS_{38kHz} = 20 \log_{10} L - 67.4.
\]

(A.6)
Then, RV Knorr 12 kHz vertical echosounder line-transect measurements near An-
denes were calibrated with the cod population density obtained from the RV Johan
Hjort 38 kHz echosounder measurements in the same region. This was done by com-
paring the RV Knorr 12 kHz area backscattering strength \( (S_a) \) measured at 03:22:45
UTC on March 8, 2014 with the cod areal population density measured from RV
Johan Hjort at 03:15:00 UTC on March 8, 2014, such that peak areal population den-
sities from the two research vessels match within 8 minutes in time and within 250
m in space. Negligible clipping occurred in the RV Knorr echosounder measurements
used for this calibration.

A.6 Mean spawning group size for Northeast Arctic
cod

The mean spawning group population for Northeast Arctic cod was determined
by averaging the annual mean spawning group populations, where each year’s mean
spawning group population was obtained by dividing the ICES inferred cod spawn-
ing population for each year in the Lofoten region [99] by the number of discrete
spawning groups we counted each year from the 30-year Lofoten line-transect survey
data from 1984 to 2014 [24]. This was done by spatially interpolating between line
transects for each annual survey and segmenting regions at and above the empirically
determined group boundary threshold density of 0.016 fish/m² to identify discrete
spawning groups. The vertical echosounder measurements of cod archived in “Nau-
tical Area Scattering Coefficients (NASC, \( s_A \))” [98] were used to determine cod areal
population densities \( n_A \) along the survey’s line transects. This was done using the em-
pirical length-dependent cod target strength \( T_{S_{38kHz}} \) at 38 kHz (Eq. A.5) determined
from each year’s mean cod length (Fig. A-4) via \( n_A = s_A/(4\pi(1852)^2\sigma_{bs}) \), where \( s_A \)
is the nautical area scattering coefficients in m²nmi⁻², and \( \sigma_{bs} = 10^{T_{S_{38kHz}}/10} \) is the
mean backscattering cross section of cod at 38 kHz in m². The mean cod length
over 1985-2014 was used for the year 1984 where trawl catch data were not available.
The annual Lofoten line-transect survey data [24] have been used to estimate annual Northeast Arctic cod spawning populations by ICES since 1985 [99]. Each year's mean spawning group size estimate has 30% error (Fig. 2-2B) which is dominated by uncertainties of each year's ICES inferred cod population [99].

A.7 Probability density of cod spawning-group size

The histogram of the natural logarithm of the cod spawning group size is found to be roughly consistent with a Gaussian probability density for the limited data available, as shown in Figure A-9A. The probability density of cod spawning group size is then consistent with a log-normal density. Cod spawning group size was obtained from the 30-year line-transect vertical echosounder data [24] by (a) spatially interpolating between line transects for each annual survey, (b) segmenting regions at and above the empirically determined group boundary threshold density of 0.016 fish/m² to identify discrete spawning groups, and (c) integrating the cod population density over each segmented region. Each year's spawning group size distribution is normalized such that it integrates to the total ICES cod spawning population for that year [99].

It is found that each year's standard deviation-to-mean ratio for the individual cod spawning group size obtained here from the 30-year line-transect vertical echosounder data [24] is stable at a mean of 2.0 to within 35% over 30 years (Fig. A-9B). This can be expressed as $\sigma \approx 2q$, where $q$ is the mean and $\sigma$ is the standard deviation of the random variable $s$ for annual group spawning size.

A log-normal random variable has a probability density that parametrically depends only on the random variable's mean and variance. Given the effectively constant relationship found between the mean and variance of $s$, annual group spawning size for cod approximately follows a single-parameter log-normal probability density,

$$P(s) = \frac{1}{s\mu_2(q)\sqrt{2\pi}} e^{-\frac{(\ln s - \mu_1(q))^2}{2\mu_2(q)^2}}, \quad (A.7)$$

where $\mu_1(q) \approx \ln[q(1+2/q)^{-1/2}]$, and $\mu_2(q) \approx (\ln[1+2/q])^{1/2}$. The single parameter
governing this distribution is the mean annual spawning group size for cod, or the cod quantum, $q$.

### A.8 Mean spawning group size for Georges Bank herring

Mean spawning group size for Georges Bank herring was determined by averaging daily spawning group populations imaged by OAWRS over the 8-day peak herring spawning period from September 26 to October 3, 2006 on the northern flank of Georges Bank. The OAWRS herring population density images of fully developed large groups on each day were obtained from measured OAWRS intensities following the standard procedures described in Refs. [9, 91, 92]. The resulting OAWRS herring density images have an error of roughly 1-2 dB in regions where group populations follow a stationary random process in space and time [91]. The daily herring spawning populations were determined by integrating the OAWRS herring areal population densities at and above the measured group formation critical density of $c_{herring} = 0.2$ fish/m$^2$ for herring [1]. Analysis of historical herring line transect data [16] to estimate spawning group size was found to be impossible without unacceptable ambiguity due to severe spatial and temporal undersampling of spawning groups by the line transect method. This was due to the fact that the line-transect survey was primarily designed to survey all diffuse herring concentrations near the seafloor over the entire northern flank of Georges Bank with widely spaced transects over a two week period. The spawning groups, however, were found by OAWRS to form at specific times (evening) and locations (primarily central bank) over this same period [1]. As a result it was found that the predetermined line transects overwhelmingly missed spawning groups when both OAWRS and predetermined line-transect surveys were conducted simultaneously in the 2006 spawning season. Even when a predetermined line-transect passed through a spawning group in an annual survey, it did so only once in the group’s small across-bank dimension, leaving the group’s large along-bank dimension
ambiguous. To our knowledge, all simultaneous measurements of spawning groups between OAWRS and conventional echo-sounding during 2006 occurred after the OAWRS vessel directed the conventional echo-sounding vessel to a given spawning group. The spawning groups observed by OAWRS did not coincide with the space-time trajectory of the planned line-transect survey, whose research vessel behaved as a Lagrangian particle in a spatially vast and temporally dynamic field of fish.

A.9 OAWRS cod population density mosaic near Røst

Shadowing by very shallow bathymetry or islands prevented the entire area around the underwater peninsula surrounding the Røst Island from being imaged instantaneously by OAWRS from a single ship location (Fig. A-10). Consequently, six instantaneous OAWRS images taken at locations around the Røst bathymetric feature at 10:40:49, 11:29:09, 12:13:19, 13:09:59, 14:04:59, and 15:07:29 UTC on February 23, 2014 were combined to generate the cod population density mosaic of Fig. 2-1B. A running spatial averaging 1 km square window was applied to the raw OAWRS data of the Fig 1B mosaic to eliminate boundary discontinuities due to lack of temporal simultaneity of the six instantaneous images. Such smoothing was not necessary in Fig. 2-1C since that image is not a mosaic.

A.10 Historic total spawning population time series in the North Atlantic

Historic spawning stock abundance time series for Atlantic cod and Atlantic herring in four spawning grounds of the North Atlantic - US Northeast, Canada, North Sea, Norway - were reconstructed and investigated from before the industrial fishing age (earlier than the late 19th century) to present (Figs. 2-2C and 2-3C).

Northeast Arctic cod: Spawning stock abundance time series of Northeast Arctic cod for 1946-2014 were obtained from reported stock numbers-at-age and pro-
portion of mature-at-age data [99], as

\[ S_i = \sum_{j=0}^{a_{max}} s_j^i \gamma_j^i, \]  

(A.8)

where \( S_i \) is the spawning stock abundance in \( i \)-th year, \( a_{max} \) is the maximum fish age, \( s_j^i \) is the stock numbers-at-age at age \( j \) in \( i \)-th year, and \( \gamma_j^i \) is the proportion of mature-at-age for age \( j \) in \( i \)-th year. For 1876-1956, cod spawning stock biomass time series were derived from Catch-Per-Unit-Effort (CPUE) estimates obtained using official Norwegian annual cod catch and fishing effort data in Ref. [100]. Spawning stock abundance of Northeast Arctic cod in 1876-1945 was then estimated by dividing biomass time series by a representative weight per cod value of 3.8 kg/fish. This 3.8 kg/fish weight per cod value was obtained in a least squares sense between biomass time series divided by a weight per cod and abundance time series in 1946-2013 from

\[ \min_w \sum_{i=1}^{M} \left( S_i - \frac{SSB_i}{w} \right)^2, \]  

(A.9)

where \( w \) is the weight per fish, \( M \) is the number of years, \( S_i \) is the spawning stock abundance in \( i \)-th year, and \( SSB_i \) is the spawning stock biomass in \( i \)-th year. The spawning stock abundance in Lofoten spawning grounds where OAWRS imaging and annual line-transect acoustic survey took place was derived by multiplying the spawning stock abundance of entire Northeast Arctic cod population by a factor of 2/3 [29].

**US Gulf of Maine cod:** Spawning stock abundance time series of US Gulf of Maine cod for 1982-2013 were obtained from reported stock numbers-at-age and proportion of mature-at-age data [101] via Eq. A.10. For 1963-1994, cod spawning stock biomass time series of US Gulf of Maine cod was available from Ref. [102]. Spawning stock abundance of US Gulf of Maine cod in 1963-1981 was then estimated by dividing biomass time series by a representative weight per cod value of 1.75 kg/fish. This 1.75 kg/fish weight per cod value was obtained by minimizing the least square difference between biomass time series divided by a weight per cod and abundance time series in 1982-1994 (Eq. A.10). Historic spawning stock abundance
time series of US Gulf of Maine cod in 1891-1914 and 1931-1965 were estimated by determining a scale factor for CPUE time series for US Georges Bank cod in 1891-1914 and 1931-1965 [103] to match US Gulf of Maine spawning stock abundance in 1963-1965, where both time series were available, from

$$\min_{\lambda} \sum_{i=1}^{M} (I_i \cdot \lambda - S_i)^2$$  \hspace{1cm} (A.10)

where $\lambda$ is the scale factor, $M$ is the number of years, $I_i$ is the quantity to be scaled such as CPUE to match $S_i$ in $i$-th year, and $S_i$ is the spawning stock abundance in $i$-th year. The CPUE estimates for Gulf of Maine cod in 1891-1914 and 1931-1965 were assumed to have the same temporal trends as those for Georges Bank cod reported in Ref. [103].

**Canada Northern cod (Atlantic cod in the statistical area 2J3KL):** Spawning stock biomass time series of Canadian Northern cod for 1983-2012 and 1962-1982 were obtained from Ref. [104] and Ref. [105], respectively. Spawning stock abundance of Canadian Northern cod in 1962-1993 and 2010-2012 was then estimated by dividing spawning stock biomass by mean weight per cod value of 2.4 kg/fish obtained for 2009 [106]. Collapsed spawning stock abundance of Canadian Northern cod in 1994-2009 was estimated by dividing spawning stock biomass by each year's mean weight per cod values varying roughly between 0.5 and 3.1 kg/fish [106]. Each years' mean weight per cod was found from number of samples caught-at-age, proportion of mature-at-age, and mean sample weight-at-age [106]. The averaged weight per cod of 1995 and 1997 was used for the year 1996 where catch data were not available. Historic spawning stock abundance time series of Canadian Northern cod in 1550-1800 were estimated by determining a scale factor for cod stock biomass time series of the entire Eastern Canadian waters in 1550-2004 [107] to match Northern cod spawning stock abundance in 1962 [105] via Eq. A.10 assuming a steady-state in the regional distribution of historic cod populations in the Eastern Canadian waters.

**North Sea cod:** Spawning stock abundance time series of North Sea cod for 1963-2014 were obtained from reported stock numbers-at-age and proportion of mature-at-
age data [108] via Eq. A.10. For 1910-1960, cod spawning stock biomass time series was available from Ref. [109]. Spawning stock abundance of North Sea cod in 1910-1960 was then estimated by dividing biomass time series by a representative weight per cod value of 3.1 kg/fish. This 3.1 kg/fish weight per cod value was obtained by minimizing the least square difference between biomass time series divided by a weight per cod and abundance time series in 1963-2014 (Eq. A.10). Historic cod spawning stock abundance in 1889-1893 was obtained by determining a scale factor for the mean spawning stock abundance in 2003-2007 as the ratio of Landing Per Unit fishing Power (LPUP) in 1889-1893 to those in 2003-2007 [110].

**Norwegian Spring-spawning herring:** Spawning stock abundance times series of Norwegian Spring-spawning herring in 1988-2014 were obtained from reported stock numbers-at-age and proportion of mature-at-age data [111] via Eq. A.10. For 1907-1998, herring spawning stock biomass time series was available from Ref. [112]. Spawning stock abundance of Norwegian Spring-spawning herring in 1907-1987 was then estimated by dividing biomass time series by a representative weight per herring value of 0.28 kg/fish. This 0.28 kg/fish weight per herring value was obtained by minimizing the least square difference between biomass time series divided by a weight per herring and abundance time series in 1988-2014 (Eq. A.10).

**North Sea herring:** Spawning stock abundance time series of North Sea herring in 1947-2014 were obtained from reported stock numbers-at-age and proportion of mature-at-age data [113] via Eq. A.10. Historic herring spawning abundance time series in 1600-1800 was obtained by determining a scale factor for CPUE time series of North Sea herring in 1600-1966 [114] to match North Sea herring spawning stock abundance in 1947-1966 via Eq. A.10, where both time series were available.

**US Georges Bank herring:** Spawning stock biomass time series of Gulf of Maine-GB (GoM-GB) herring population complex in 1965-2014 were obtained from Ref. [115]. For 1880-1964, spawning stock biomass of GoM-GB herring complex was estimated by employing a surplus production model with fisheries data in 1880-2007 in the Gulf of Maine [116] via Eq. A.11. The herring fisheries data in 1880-2007 obtained from Ref. [116] was scaled to represent entire GoM-GB fisheries catch by an
The inverse of (1 - mean ratio of Georges Bank herring stock to entire GoM-GB herring complex [16]), and then used as an input to the surplus production model. Since the surplus production model outputs the total stock biomass, the mean ratio of the spawning stock biomass to total stock biomass as well as the mean ratio of Georges Bank herring stock to entire GoM-GB herring complex are multiplied to the output of the model to obtain spawning stock biomass of Georges Bank herring. The 3-year running averages of ratio of Georges Bank herring stock to entire GoM-GB herring complex was used for Figure 3C [16]. Herring spawning stock abundance was estimated by dividing spawning stock biomass by a representative weight per herring of 0.1 kg/fish [16].

**Canada Southwest (SW) Nova Scotia herring (SW Nova Scotia/Bay of Fundy Atlantic herring in the statistical area 4VWX):** Spawning stock biomass of SW Nova Scotia and Bay of Fundy herring in 4VWX statistical area in 1965-2006 was obtained from Ref. [117]. For 1880-2007, historic herring spawning stock biomass of SW Nova Scotia and Bay of Fundy herring was estimated by employing a surplus production model with fisheries data in 1880-2007 in Bay of Fundy [116] via Eq. A.11. The herring fisheries data in 1880-2007 obtained from Ref. [116] was used as an input to the surplus production model. Since the surplus production model outputs the total stock biomass, the ratio of spawning stock biomass to total stock biomass was multiplied to the output of the model to obtain spawning stock biomass of SW Nova Scotia/Bay of Fundy herring. Herring spawning stock abundance was estimated by dividing spawning stock biomass by a representative weight per herring of 0.1 kg/fish.

### A.11 Surplus production model for Atlantic herring populations

A surplus production model was employed to estimate historic herring spawning stock abundance in Georges Bank and Southwest Nova Scotia/Bay of Fundy from the late
19th century to present. A similar approach appears in Ref. [107] for Atlantic cod biomass in the entire Eastern Canadian waters. The employed surplus production model is based on an equation [107]:

\[ B_y = B_{y-1} + r_b \frac{B_{y-1}}{aK + B_{y-1}} B_{y-1} \left(1 - \frac{B_{y-1}}{K}\right) - C_y, \tag{A.11} \]

where \( B_y \) is the total stock biomass in year \( y \), \( r_b \) is the intrinsic rate of population growth expected without the depensation effect, \( K \) is the carrying capacity of the ecosystem or the initial biomass, \( a \) is the depensation parameter, and \( C_y \) is the fisheries catch data in year \( y \). The depensation parameter \( a \) is set to be zero for \( B_y > K/2 \) and given a value of 0.06 [107] for \( B_y \leq K/2 \). Historic stock biomass \( K \) before industrial fishing began and the parameter \( r_b \) were determined by searching for the values at which available stock biomass data time series best-match the surplus production model estimate time series such that the percentage error in surplus production model fit is minimized:

\[
\min_{r_b, K} \sqrt{\frac{\sum_{i=1}^{M} (B_{i,\text{data}} - B_{i,\text{SP}}(r_b, K))^2}{\sum_{i=1}^{M} B_{i,\text{data}}^2}}, \tag{A.12}
\]

where \( r_b \) is the intrinsic rate of population growth expected without the depensation effect, \( K \) is the carrying capacity of the ecosystem or the stock biomass before industrial fishing began, \( M \) is the number of years, \( B_{i,\text{data}} \) is the available stock biomass data in \( i \)-th year, and \( B_{i,\text{SP}} \) is the surplus production model stock biomass estimate in \( i \)-th year. The ranges of historic spawning stock abundance for Georges Bank herring and SW Nova Scotia/Bay of Fundy herring in Fig. 2-3C were determined at 30% error in surplus production model fit to data.

A.12 Time to return to pre-industrial levels after a decline below one quantum level

The time \( (T_{\text{return}}) \) required for a total spawning population time series \( (S) \) that has declined below the quantum level \( (A) \) to return to its pre-industrial level \( (B) \) is deter-
mined via the following procedure, where \( S_i \) is the total spawning population in the \( i \)-th year, \( i = 1, \ldots, N \), \( N \) is the total number of years, and \( \Delta T \) is the time interval between \( S_i \) and \( S_{i+1} \):

1. Define an index \( D = i \mid (S_i-A)(S_{i+1}-A) < 0 \) such that the total spawning population time series crosses the quantum level \( A \) at the years in \( S \).
2. Find the earliest year at which \( S \) declines below \( A \), \( i_D = \min(D) \).
3. Define an index \( P = i \mid (S_i-B)(S_{i+1}-B) < 0 \) and \( i > i_D \) such that \( S \) crosses pre-industrial level \( B \) at the years in \( S \).
4. Find the earliest year at which \( S \) returns to \( B \), \( i_P = \min(P) \) when \( n(P) > 0 \). If the total spawning population does not return to pre-industrial level, the largest year index \( N \) is used for \( i_P \), i.e. \( i_P = N \) when \( n(P) = 0 \).
5. \( T_{\text{return}} \) is determined to be \((i_P - i_D)\Delta T \) when \( n(P) > 0 \) and \( T_{\text{return}} \) is greater than \((i_P - i_D)\Delta T \) when \( n(P) = 0 \).

By this analytic algorithm, it is found that some spawning populations have returned to pre-industrial levels after declining below the size of a mean spawning group measured from a population at pre-industrial levels, or quantum. In particular, it took Norwegian herring and Georges Bank herring 19 and 13 years respectively to return to pre-industrial levels after declining below a herring quantum. It is also found, however, that some spawning populations have not yet returned to pre-industrial levels after going below the size of a quantum. In particular, Gulf of Maine cod, Canadian Northern cod and North Sea cod have not returned to pre-industrial levels after going below a cod quantum 22, 21 and 10 years ago respectively.

### A.13 Population growth rate

The population growth rate \( R_i \) in the \( i \)-th year is defined through a logarithmic function as \([118, 119]\):
\[ R_i = \frac{10 \log_{10} \frac{S_{i+1}}{S_i}}{\Delta T}, \]  

where \( S_i \) is the total spawning population in \( i \)-th year, \( \Delta T \) is the time between \( S_i \) and \( S_{i+1} \), \( i = 1, ..., N-1 \), and \( N \) is the total number of years. The population growth rates obtained each year from 1920 to 2014 are shown as a function of total spawning population for each population we investigated (Supplementary Figs. A-11 and A-12).

Figure A-1: OAWRS system used for cod measurements during the Nordic Seas 2014 Experiment. Effectively monostatic OAWRS system used for cod measurement during the Nordic Seas 2014 Experiment, where source and receiver arrays were towed from the same research vessel RV Knorr. The OAWRS source was developed under the National Science Foundation and Sloan Foundation MRI program for wide-area sensing of marine life, and the ONR Five Octave Research Array (FORA) was used as the OAWRS receiver.
Figure A-2: Measured vertical sound speed profiles. Measured XBT and CTD sound speed profiles during Nordic Seas 2014 Experiment near Andenes (A) and Røst (B).
Figure A-3: Cod target strength calibration at OAWRS frequencies. Expected cod target strength ($T_{S_{OAWRS}}$) at OAWRS frequencies was determined (A) by matching the cod areal population density obtained from the 12 kHz RV Knorr vertical echosounder measurements at 01:51:00-02:20:24 UTC and that from the OAWRS image at 01:26:49 UTC on March 8, 2014 for corresponding locations along the RV Knorr’s line transect (B).
Figure A-4: Northeast Arctic cod length measured over 1985-2014. Time series of mean (circles) and standard deviation (vertical lines) of cod length measured from annual trawl survey in the Lofoten area over 1985-2014.
Empirical best-fit boundary population density

Critical population density obtained for cod by scaling law

Figure A-5: Least squares estimate of cod group boundary critical density from decades of Northeast Arctic line-transect survey data and ICES spawning population time series. Cod group boundary critical density of roughly 0.016 fish/m² is found to provide the lowest least squares error between the annual spawning population time series and decades of line-transect survey data under the assumption that summing spawning groups approximates the total annual spawning population each year from Eq. A.3. Cod group boundary critical density range, in red, independently found from scaling herring critical density via Eq. A.3 is consistent with this least squares result.
Figure A-6: Correspondence between OAWRS and RV Knorr line-transect vertical echosounder measurements of cod near Andenes. Correspondence between (A) an instantaneous OAWRS cod population density image and (B) 12 kHz RV Knorr vertical echosounder line-transect measurements near Andenes. The OAWRS cod population density image (A) was obtained near Andenes at 01:26:49 UTC on March 8, 2014. The 12 kHz RV Knorr vertical echosounder line-transect measurements (B) were made from 01:51:00 to 02:20:24 UTC on March 8, 2014. The white line transect in (A) passing through the group imaged by OAWRS is the RV Knorr track with corresponding vertical echosounder data shown in (B).
Figure A-7 (facing page): Correspondence between OAWRS and RV Knorr line-transect vertical echosounder measurements of cod near Røst. (A-D) Simultaneous correspondence between OAWRS imaging from RV Knorr and the 12 kHz RV Knorr vertical echosounder line-transect measurements near Røst where RV Knorr tracks through the OAWRS imaged groups are given in (A) and corresponding line vertical echo-sounder data is given in (B-D). (E-H) Correspondence between OAWRS imaging and the 12 kHz RV Knorr vertical echosounder line-transect measurements taken 5 hours after OAWRS imaging near Røst. (I-L) Correspondence between OAWRS imaging and the 12 kHz RV Knorr vertical echosounder line-transect measurements taken roughly 24 hours after OAWRS imaging near Røst.
Figure A-8: Consistency between OAWRS and annual line-transect vertical echosounder measurements of cod near Røst. Consistency between cod population density distribution near Røst measured by OAWRS imaging on February 23, 2014 and by vertical echosounder measurements during the annual line-transect acoustic survey on March 17-31, 2014 (purple circles). High cod density areas near Røst and general tapering trends radially away from Røst are clear and consistent in both OAWRS and vertical echosounder measurements.
Figure A-9: The probability density of cod spawning group size can be characterized by a single parameter, namely the mean spawning group size or quantum. (A) The histogram of the natural logarithm of the cod spawning group size obtained from the 30-year line-transect echosounder data [24] is found to be consistent with a Gaussian probability density distribution. (B) The standard deviation-to-mean ratio of cod spawning group size obtained from the 30-year line-transect echosounder data [24] is found to be stable with a mean of 2.0 to within 35% of the mean over 30 years such that $q \approx 2\sigma$, where $q$ is the mean and $\sigma$ is the standard deviation of the cod spawning group size. Since the Gaussian density parametrically depends on only the mean and standard deviation, and the standard deviation can be expressed in terms of the mean, the probability density of cod group spawning size depends only on the mean group size, or the cod quantum, $q$. 
Figure A-10: RV Knorr locations for cod population density mosaic RV Knorr locations for the Fig. 2-1B mosaic at 10:40:49, 11:29:09, 12:13:19, 13:09:59, 14:04:59, and 15:07:29 UTC on February 23, 2014.
Figure A-11: Population growth rate as a function of total spawning population each year from 1920 to 2014 for each cod population investigated. The cod quantum is the mean spawning groups size of 10.6 million found for Atlantic cod from a total population at pre-industrial levels, where the only available data was from the Northeast Arctic. The figures do not consistently show higher variance in growth rate for smaller populations.
Figure A-12: Population growth rate as a function of total spawning population each year from 1920 to 2014 for each herring population investigated. The herring quantum is the mean spawning groups size of 204 million found for Atlantic herring when the total population was at pre-industrial levels, where the only available data was from Georges Bank. The data show indications that for small populations below a quantum growth rate variance may be higher.
Appendix B

Appendix for Instantaneous 3D Continental-shelf Scale Imaging of Oceanic Fish by Multi-spectral Resonance Sensing Reveals Group Behavior During Spawning Migration
Figure B-1: The OAWRS scattering strength images and corresponding exemplary coherence areas taken at (A-B) 17:25:00, (C-D) 18:42:30, (E-F) 19:52:30, (G-H) 21:37:30 EDT on September 29, 2006 (B,D,F,H). The e-folding lengths from the peak of the 2D autocorrelation function of herring scattering strength are plotted in black.
Figure B-2: (A) The horizontal distribution of neutral buoyancy depth. (B) Estimation uncertainty for neutral buoyancy depth in percentage. (C) The horizontal distribution of herring population density. (D) Estimation uncertainty for herring population density in percentage. (E) The horizontal distribution of root-mean-square (RMS) shoal thickness $H_{RMS} = 2H/\sqrt{12}$, where $H = |\max z - \min z|$, and $z$ is the herring occupancy depth. (F) Estimation uncertainty for RMS shoal thickness in percentage.
Appendix C

Appendix for Feasibility of Acoustic Remote Sensing of Large Herring Shoals and Seafloor by Baleen Whales

C.1 Estimation of Expected Target Strength of a Single Herring in a Vertical Layer

The probability density function of a Gaussian variable $x$ with mean $\mu$ and standard deviation $\sigma_x$ is:

$$f(x|\mu, \sigma_x^2) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-\mu)^2}{2\sigma_x^2}}$$  \hspace{1cm} (C.1)

with a cumulative distribution function:

$$F(x) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x-\mu}{\sigma_x \sqrt{2}} \right) \right]$$  \hspace{1cm} (C.2)

where $\text{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt$.

When the variable $x$ is truncated within $x_{low} < x < x_{high}$ where $x_{low}$ and $x_{high}$ are determined to ensure positiveness of the parameters for target strength estimation of a single herring in a vertical layer, the probability density function $g(x)$ for the
truncated Gaussian distribution is expressed as:

\[
g(x|\mu, \sigma^2, x_{\text{low}}, x_{\text{high}}) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{1}{F(x_{\text{high}}) - F(x_{\text{low}})}
\]  

(C.3)

This distribution is assumed for mean depth of herring shoal \(z_0\), shoal thickness \(H\), neutral buoyancy depth \(z_{nb}\), and herring length with mean \(\mu\) and standard deviation \(\sigma_x\) given in Table 4.2.

Using a fish swimbladder resonance model [3, 43], the expected target strength \(TS\) of a single herring in a uniform vertical layer of herring is determined from:

\[
\left\langle \left| \frac{S}{k} \right|^2 \right\rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{H} \int_{z_0-H/2}^{z_0+H/2} \frac{S^2}{k} d\zeta g(z_0)g(H)dg(l)d\eta(z_{\text{nb}})dz_{\text{nb}}
\]

\[
= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{H} \int_{z_0-H/2}^{z_0+H/2} \frac{r^2(z, z_{\text{nb}}, l)}{f^2(z, z_{\text{nb}}, l) - 1} \eta^{-2}(z, z_{\text{nb}}, l, f) + \left( \frac{f^2(z, z_{\text{nb}}, l)}{f^2} - 1 \right) dz_{\text{nb}} dz(l) d\eta(z_{\text{nb}})g(H)dg(z_{\text{nb}})
\]  

(C.4)

where \(S\) is the plane wave scatter function of a single herring, \(k\) is the acoustic wavenumber, \(f\) is the frequency, \(z\) is the herring depth, \(l\) is the fork length of herring, \(r(z, z_{\text{nb}}, l)\) is the equivalent radius of swimbladder, \(f_0(z, z_{\text{nb}}, l)\) is the resonant frequency of swimbladder, \(\eta(z, z_{\text{nb}}, l, f)\) is the damping factor, \(TS = 10 \log_{10} \left\langle \left| \frac{S}{r_{\text{ref}}k} \right|^2 \right\rangle\), and \(r_{\text{ref}} = 1\) m is the reference length. In Equation (C.4), the equivalent radius of swimbladder \(r(z, z_{\text{nb}}, l)\) is determined by:

\[
r(z, z_{\text{nb}}, l) = \left[ \frac{3}{4\pi} \frac{c_{\text{nb}} m_{\text{flesh}}(l)}{\rho_{\text{flesh}}} \right]^{1/3}
\]

(C.5)

assuming that the swimbladder volume varies with pressure according to Boyle’s law [47], where \(c_{\text{nb}}\) is the ratio of the swimbladder volume at neutral buoyancy to the volume of herring’s flesh \(V_{\text{flesh}}\) assumed to be 0.05 [48], \(V_{\text{flesh}} = m_{\text{flesh}}(l)/\rho_{\text{flesh}}\), \(m_{\text{flesh}}(l)\) is the mass of a single herring empirically determined by the fork length of herring \(l\) as \(3.35 \times 10^{-6}\) \(l^{3.35}\) in kg when \(l\) is given in cm [9], and \(\rho_{\text{flesh}}\) is the density of herring’s flesh of 1071 kg/m\(^3\) [49]. The resonant frequency of herring swimbladder \(f_0(z, z_{\text{nb}}, l)\) in Equation (C.4) is determined by:
\[ f_0(z, z_{nb}, l) = \frac{\kappa(\epsilon(z, z_{nb}, l))}{2\pi r(z, z_{nb})} \sqrt{\frac{3\gamma P_{atm}(1 + z/10)}{\rho_{flesh}}} \]  

(C.6)

where \( \gamma = 1.4 \) is the ratio of the specific heats of air, and \( P_{atm} = 1.013 \times 10^5 \) Pa is the atmospheric pressure, \( \kappa(\epsilon(z, z_{nb}, l)) \) is the swimbladder correction term, and \( \epsilon(z, z_{nb}, l) \) is the swimbladder's eccentricity. The correction term \( \kappa(\epsilon(z, z_{nb}, l)) \) for a prolate spheroidal swimbladder [50] is given by:

\[ \kappa(\epsilon(z, z_{nb}, l)) = \frac{\sqrt{2}(1 - \epsilon^2(z, z_{nb}, l))^{1/4}}{\epsilon^{1/3}(z, z_{nb}, l)} \left[ \ln \left( \frac{1 + \sqrt{1 + \epsilon^2(z, z_{nb}, l)}}{1 - \sqrt{1 - \epsilon^2(z, z_{nb}, l)}} \right) \right]^{-1/2} \]  

(C.7)

In Equation (C.7), the ratio of the minor to major axis of a prolate spheroidal swimbladder \( \epsilon(z, z_{nb}, l) \) is \( \left( \frac{csb/2}{r(z, z_{nb}, l)} \right)^{-3/2} \), and the ratio between the herring's fork length \( l \) and the major axis of the swimbladder is assumed to be \( csb \approx 0.364 \) for herring [9]. The damping factor \( \eta(z, z_{nb}, l, f) \) in Equation (C.4) is obtained from:

\[ \frac{1}{\eta(z, z_{nb}, l, f)} = \frac{2\pi r(z, z_{nb}, l)f^2}{\kappa f_0(z, z_{nb}, l)} + \frac{\xi}{\pi r^2(z, z_{nb}, l)f_0(z, z_{nb}, l)\rho_{flesh}} \]  

(C.8)

where \( f \) is the frequency, \( c \) is the sound speed, and \( \xi \) is the viscosity of herring's flesh given to be 50 Pa s [49].
Figure C-1: Wind dependence of Gulf of Maine mean ambient noise levels for frequency bands of baleen whale vocalizations given in Table 4.1. Gray filled circles represent the mean ambient noise levels determined from measurements obtained during the OAWRS 2006 experiment [1, 14, 34] using Equation (4.9) for frequency bands of baleen whale vocalizations given in Table 4.1. Black solid lines represent the empirical best fit of the measured ambient noise levels to the model given in Equation (4.11). Empirically determined coefficients of the model are given in Table 4.1. Black vertical lines represent the variation of the measured mean ambient noise levels for given wind speed. Instantaneous measurements of ambient noise fluctuate around each mean ambient noise level with 5.6 dB standard deviation [41].
Appendix D

Environmental Impact Considerations for OAWRS

D.1 Re-evaluation finds no effect of sonar on humpback whale song occurrence

Before and during OAWRS survey transmissions [1, 9], we measured a constant humpback whale song occurrence rate, as shown in Figure D-1, indicating no change of humpback song related to these transmissions over the entire survey area in the Gulf of Maine, a roughly 400-km diameter area, including Georges and Stellwagen Banks. Additionally, we find that the humpback whale song occurrence rate from Stellwagen Bank was constant before and during OAWRS survey transmissions, indicating no change of humpback song at Stellwagen Bank related to these transmissions. These direct measurements contradict the conclusions of Risch et al. [120].

To investigate this contradiction, we first follow the standard practice of checking for the bias [121, 122] of a statistical test by applying the test to control data where no stimulus is present to determine the false positive outcome rate [123–125]. Since the bias of Risch et al. statistical test [120] was not checked in Risch et al. [120], we do so here (Section D.2.1) with the available annual humpback whale song occurrence
data [126] from the same set of single sensors Risch et al. [120] used at Stellwagen Bank. We show that their statistical test false-positively finds whales react to sonar 98-100% of the time over a yearly period when no sonars are present. For example, when their statistical test is applied to annual humpback whale song occurrence data published in Ref. [126], with 2006 as the test year and 2008 as the control year, it false-positively finds whales react to sonar: (1) 100% of the time over the year before the “during” period; and (2) 98% of the time over the year when the “during” period is excluded from the test, as described in Section D.2.1 and Table D.1. Here the “during” period is defined as the 11-day period from September 26 to October 6 with active OAWRS survey transmissions, the “before” period is the 11-day period before the “during” period, and the “after” period is the 11-day period after the “during” period following the usage in Risch et al. [120]. When applied to the same humpback whale song occurrence data reported in Risch et al. [120] over the 33-day period from September 15 to October 17 for 2008 and 2009, with either of these two years as the test year and the other as the control year, the statistical test false-positively finds humpback whales respond to sonar 100% of the time when no sonar is present, as described in Section D.2.1 and Table D.2, indicating a self-contradiction in the Risch et al. [120] approach. No meaningful conclusions can be drawn from a statistical test with such high bias.

An explanation for the severe bias in the statistical test of Risch et al. [120] becomes evident upon inspection of the annual humpback whale song occurrence time series published in Ref. [126]. Very large natural variations within and across years are common in the humpback whale song occurrence time series when no sonars are present, as can be seen in Figure D-2. There are many periods lasting roughly weeks where high song occurrence episodes are found in one year but not in another, when no sonars are present (Figure D-2). For the majority of the time, greater than 57%, the difference in the song occurrence across years when no sonars are present exceeds that of the “during” period (Figure D-3), indicating that there is nothing unusual about such differences, which rather than “alterations” [120] are actually the norm.
The statistical test used by Risch et al. [120] is overwhelmingly biased because it mistakes natural variations in humpback whale song occurrence 98-100% of the time for changes caused by sonar when no sonar is present, lacks any true positive confirmation and so lacks the statistical significance to draw the conclusions of Risch et al. [120].

Since the reported reducing change in humpback whale song occurrence, to zero [120, 126], occurred in the “before” period (Figure D-2) while the OAWRS vessels were inactive and docked on the other side of Cape Cod from Stellwagen Bank at the Woods Hole Oceanographic Institution due to severe winds days before OAWRS transmissions for active surveying began on September 26, 2006, the Risch et al. analysis [120] severely violates temporal causality. Moreover, the annual humpback whale song occurrence time series are uncorrelated over 11-day periods across years, and the correlation coefficient obeys a random distribution peaking at zero correlation about which it is symmetric (Figure D-4), showing that correlation in trend between years is random and quantitatively expected to be zero with roughly as many negative correlations as positive ones. In fact, the correlation coefficient between the humpback whale song occurrence across years smoothly transitions from negative values in the “before” period, showing no similarity or relation in trend between years just before the 2006 OAWRS survey transmission period, to some of the highest positive correlations obtained between years in the “during” period (Figure D-4). This demonstrates high similarity and relation in trend between years during the 2006 OAWRS active survey transmission period, which contradicts the results of the Risch et al. [120] study. These causality violations are also discussed in the context of the measured temporal coherence of humpback whale song occurrence in Section D.2.4.

It is well known that wind speed variation can lead to severe detection range limitations in passive sensors, especially a single sensor that has zero array gain [127–130]. Risch et al. [120] did not investigate the effect of wind dependent ambient noise on the detection range of their single hydrophones located in the Stellwagen Bank
(Figure D-5). They did report that “Ambient noise levels over the whole analysis bandwidth (10-1000 Hz) and in the frequency band with most humpback whale song energy (70-300 Hz) did not vary dramatically within or between years.” Wind speeds varied, however, from calm to near-gale conditions within a period of a few hours or days, many times over the 33-day period examined by Risch et al. [120], as is common for Fall in Stellwagen Bank [131]. These natural wind speed variations must have significantly changed the local wind-dependent noise level according to known physics [127, 132]. Since noise “can have a tremendous, if not a dominating, influence on the detection range of any sonar system” [133], the dramatic changes in wind speed at Stellwagen Bank must have led to dramatic changes in the detection range of single sensors deployed there. The range at which signals, in this case humpback whale songs, can no longer be detected because they become indistinguishable from ambient noise is the detection range from the sensor. Since ambient noise is wind speed dependent, so is the detection range (Figure D-5), and so is humpback whale song occurrence measured at that sensor if variations in wind speed cause the detection range to pass through the range of the singing humpback whales (Figure D-6). In this case even if a whale sang at a constant rate, song occurrence measured at the sensor (Figure D-7) would vary with local wind noise (Figure D-6). Moreover, the annual humpback whale song occurrence reported in Ref. [126] had a standard deviation of 3.54 dB in the 33-day period examined by Risch et al. [120], which was less than the 3.8 dB standard deviation in ambient noise level reported by Risch et al. [120], and so local ambient noise variation could have caused all the variations in humpback whale song occurrence reported over that period.

Using the measured wind speeds at Stellwagen Bank [131], and the measured spatial distribution and constant rates of singing humpback whales determined by our large aperture array, we determine the song occurrence detectable by a single hydrophone at Stellwagen Bank, as shown in Figure D-7. We find it to match the song occurrence reported by Risch et al. [120] in the “before” and “during” periods with high accuracy, within ±18% of the reported means, which is much less than the standard deviation
of the humpback whale song occurrence reported by Risch et al. [120]. This match shows that the variation in reported song occurrence from the “before” to “during” period is due to detection range limitations of the single sensor at Stellwagen Bank from wind-dependent ambient noise, and is not due to the song production rate, which we show to be constant. The constant song production and occurrence rates in the “before” and “during” periods measured by our large aperture array are unaffected by wind noise because the array gain was sufficiently high to make the detection range well beyond the range of the vocalizing whales for all wind conditions (Figure D-5). Our data shows no humpback whale vocal activity originating from Stellwagen Bank in either the “before” or “during” periods, but vocalizing humpback whales located near Georges Bank could be heard at Stellwagen Bank during low wind noise conditions (Figure D-5). In high wind noise, the single sensor mean detection range at Stellwagen Bank is too short to include the regions with measured singing humpback whales, but in low wind noise, it is large enough to include the regions with measured singing humpback whales as shown in Figure D-5, making the mean song detection rate at Stellwagen Bank higher in lower wind noise. Noise from near gale force winds in the last 3 days of the “before” period, for example, caused a significant drop in the detection range of the single sensor and the corresponding significant drop in the song occurrence rate at Stellwagen Bank [126] while the OAWRS vessels were inactive and docked at the Woods Hole Oceanographic Institution. Since the OAWRS experiment was conducted only up to October 6, 2006, the vocalizing humpback whale distribution in the “after” period was not measured and we do not investigate the song occurrence for that period.

It has been previously shown that due to collapse of the herring stock at Stellwagen Bank, humpback whale populations drastically decline at Stellwagen Bank during the herring spawning period and correspondingly increase at other locations where spawning populations are large [134]. Moreover, in the Fall of 2006, herring populations were negligible in the Massachusetts Bay and Cape Cod area, including Stellwagen Bank [135], but in contrast were decadally high in the Georges Bank region [136],
consistent with the theory that humpback whales migrate to locations with large spawning herring aggregations [134]. This phenomenon was not mentioned or investigated in Risch et al. [120], but it is highly relevant because the time period Risch et al. [120] focused on is centered exactly on the peak annual herring spawning period of the Gulf of Maine for 2006. Indeed, it has been previously shown by OAWRS in Ref. [1] and by annual NEFSC acoustic echosounding and trawl surveys in Refs. [137] and [138] that this peak annual herring spawning period occurred from the last week of September to the first week of October 2006 on Georges Bank. Based on the results of Ref. [134], it should then be expected that the Stellwagen Bank humpback whale population would be low at this time and the population at Georges Bank would be high, as has been confirmed in Section 2.1 of Ref. [34] for vocalizing humpback whales.

The levels of the various anthropogenic noises at Stellwagen Bank were not discussed in Risch et al. [120], but only OAWRS levels were selected for analysis and discussion without this context. It is recommended by the National Academy of Sciences (NAS), however, that “A comprehensive noise impact assessment would include additional specific data regarding both sound levels and sources throughout the area for which impacts are being assessed [139].” Such an impact assessment should include “all aspects of the acoustic environment” [140] to avoid the problem another impact assessment had of being evaluated as “misrepresentative of the existing soundscape [139].” Here the soundscape of anthropogenic noise sources at Stellwagen Bank, from highest to lowest intensity or loudest to most quiet is delineated in Tables D.3 and D.4, following these NAS recommendations, where it is seen that the reported OAWRS transmissions fell at the quietest end of the noise spectrum when audible. Shipping traffic, on the other hand, contributes most to the anthropogenic component of mean acoustic intensity at Stellwagen Bank by many orders of magnitude. Most anthropogenic sources of underwater noise listed in Tables D.3 and D.4 continuously operate [141, 142] over a wide range of frequencies audible to whales, i.e. tens to hundreds of Hertz [127, 133, 142, 143], and result in received levels that may exceed the currently recommended NOAA guideline of 120 dB re 1 μPa received level [144–148] in water.
for continuous noise [149] for a range of whale distances (Table D.3). Even the maximum OAWRS received sound pressure level reported by Risch et al. [120] is orders of magnitude lower than the current 160 dB NOAA guideline for short duration signals such as the OAWRS 1-2 seconds duration pulse, and significantly lower than the 120 dB guideline for even continuous sources [149] which OAWRS is not. The maximum received acoustic intensities of OAWRS signals at Stellwagen Bank reported by Risch et al. [120] are the same as those of a quiet wooded forest or a quiet room with no conversation [150], whereas the acoustic intensities received at Stellwagen Bank from shipping traffic are often the same as those of a busy roadway or a busy airport runway [150, 151]. Risch et al. [120] reported that visual inspections of humpback whales in Stellwagen Bank were made during the OAWRS experiment, suggesting that humpback whales were within visible range of research vessels. Research vessels close enough to whales to sight them can easily have engine noise levels at the whales greatly exceeding the reported OAWRS levels over broader frequency bands and much greater time duration (Table D.3).

Before and during OAWRS survey transmissions, we measured constant humpback whale song occurrence and production rates over our entire survey area roughly 400-km in diameter covering most of the Gulf of Maine, including Stellwagen Bank, indicating the transmissions had no effect on humpback whale song production rate. Using annual humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time dependent ambient noise, we show the statistical test used by Risch et al. [120] for assessing the response of humpback whales to sonar transmission false positively finds humpback whales respond to sonar 98-100% of the time when no sonars are present. With this and the lack of any true positive confirmation for the Risch et al. [120] statistical approach, the analysis of Risch et al. [120] lacks the statistical significance to draw the conclusions of Risch et al. [120]. The fact that the Risch et al. [120] analysis only allows sonar causes for the reducing change reported in Risch et al. [120], yet the change occurred days before the sonar survey began, is consistent with a violation of temporal causality in the
Risch et al. [120] study. The Risch et al. statistical test [120] mistakes natural variations in whale song reception, from such factors as natural variations in whale distributions [152], singing behavior [153, 154], and ambient noise, for changes caused by sonar 98-100% of the time when no sonar is present. Before and during OAWRS survey transmissions, we find that the variations in song occurrence at Stellwagen Bank reported by Risch et al. [120] are consistent with the natural phenomena of detection range fluctuations caused by wind-dependent ambient noise, through well established physical processes [127, 132]. Misinterpretation of natural phenomena from flawed analytic methods such as biased testing and neglect of physical laws can have seriously negative consequences [155–159].

D.2 Materials and Methods

D.2.1 Risch et al. statistical test

To evaluate its bias and quantify the impact of this bias, the Risch et al. statistical test of Ref. [120] is applied to Stellwagen Bank humpback whale song occurrence data reported in Refs. [120, 126], since the bias of this test has not been previously investigated, and the implications of a bias have not been previously analyzed or discussed for this test.

The Risch et al. statistical test [120] applies the Tukey method [160] for simultaneous pairwise multiple comparison with the quasi-Poisson generalized linear model (GLM) and log link in the statistical programming language ‘R’ [120, 161, 162] to humpback whale song occurrence over non-overlapping 11-day periods within a 33-day period across years, and tests the resulting pairwise comparisons following the statements of Table D.5. The input to the statistical test of Ref. [120] is daily humpback whale song occurrence time series data over each 11-day period. Each pairwise comparison between the mean song occurrence in the \( j^{th} \) 11-day period of the \( i^{th} \) 33-day period in the \( k^{th} \) year and that in the \( l^{th} \) 11-day period of the \( i^{th} \) 33-day period in the \( m^{th} \)
year is assigned a value of $P_{(j,k),(l,m)}$. The value of $P_{(j,k),(l,m)}$ is the probability that the absolute value of the Tukey test statistic [160] is greater than the observed value of the test statistic, conditioned on the null hypothesis, i.e., all mean humpback whale song occurrences over 11-day periods are the same, and is denoted by the variable $P$ in Risch et al. [120]. If $P_{(j,k),(l,m)}$ is less than a threshold $P_T$ set by the user, then the means are classified by the user to be significantly different, otherwise they are classified by the user to be not significantly different.

Suppose there are daily humpback whale song occurrence time series over $M$ years, and for each year there are $N$ 33-day periods. Let $\mu_{(j,k)}^i$ be the mean humpback whale song occurrence over the $j^{th}$ 11-day period of the $i^{th}$ 33-day period in the $k^{th}$ year, where $i = 1, \ldots, N$, $j = 1, 2, 3$, and $k = 1, \ldots, M$. Let $k = 1$ be the test year and let $k = 2, \ldots, M$ be the control years.

For a given 33-day period over $M$ years, there are $3MC_2 = \frac{(3M)!}{2!(3M-2)!}$ pairs of 11-day periods. Comparing the $P_{(j,k),(l,m)}$ with $P_T$ for each of the $3MC_2$ pairs, outcome $T_{(j,k),(l,m)}$ is assigned for the comparison between the mean song occurrence pair $\mu_{(j,k)}^i$ and $\mu_{(l,m)}^i$. The possible outcomes $T_{(j,k),(l,m)}$ are: (1) $X(\mu_{(j,k)}^i < \mu_{(l,m)}^i)$, which is defined as: $\mu_{(j,k)}^i$ and $\mu_{(l,m)}^i$ are not significantly different and $\mu_{(j,k)}^i < \mu_{(l,m)}^i$; (2) $X(\mu_{(j,k)}^i < \mu_{(l,m)}^i)$, which is defined as: $\mu_{(j,k)}^i$ and $\mu_{(l,m)}^i$ are not significantly different and $\mu_{(j,k)}^i < \mu_{(l,m)}^i$; (3) $X(\mu_{(j,k)}^i < \mu_{(l,m)}^i)$, which is defined as: $\mu_{(j,k)}^i$ and $\mu_{(l,m)}^i$ are significantly different and $\mu_{(j,k)}^i < \mu_{(l,m)}^i$; and (4) $X(\mu_{(j,k)}^i < \mu_{(l,m)}^i)$, which is defined as: $\mu_{(j,k)}^i$ and $\mu_{(l,m)}^i$ are significantly different and $\mu_{(j,k)}^i < \mu_{(l,m)}^i$, as given in Table D.6.

The rate of false positive findings that whales respond to sonar when no sonar is present is

$$P_{FP} = \frac{\sum_{i=1}^{N_S} b_i}{N_S},$$

(D.1)
where
\[
b_i = \begin{cases} 
1 & \text{when } \sum_{n=1}^{4} a_{i,n} \text{ is non-zero}, \\
0 & \text{otherwise}, 
\end{cases} \tag{D.2}
\]

\(N_S\) is the number of 33-day periods when no sonars are present, the \(a_{i,n}\) are defined in Table D.5, and each \(i^{th}\) 33-day period, for \(i = 1, 2, ..., N_S\), has no sonar present.

### D.2.2 False positive rate and statistical bias of the Risch et al. statistical test

When the Risch et al. statistical test [120], as described mathematically in Section D.2.1 and Table D.5, is applied to the three 33-day humpback whale song occurrence time series data reported in Risch et al. [120], with 11-day time series indices \(j = 1\) for the “before” period from September 15 to September 25, \(j = 2\) for the “during” period from September 26 to October 6, and \(j = 3\) for the “after” period from October 7 to October 17, and indices \(k = 1\) for year 2006, \(k = 2\) for year 2008 and \(k = 3\) for year 2009, we obtain the same \(P\) values and results reported in the ‘Risch et al. Statement’ column of Table D.5. Specifically, daily humpback whale song occurrence time series denoted by \(\Psi_{2006}\) for year 2006, \(\Psi_{2008}\) for year 2008, and \(\Psi_{2009}\) for year 2009, from song occurrence data reported in Risch et al. [120] over the 33-day period from September 15 to October 17, are input to the Tukey tests of the statistical programming language ‘R’, as described in Section D.2.1. Since there is only one 33-day period from September 15 to October 17, \(i = 1\). This 33-day period consists of the three consecutive non-overlapping 11-day periods with indices \(j\) or \(l = 1, 2, 3\) and year indices \(k\) or \(m = 1, 2, 3\) for pairwise comparisons between periods within and across years. A value of \(p_{(j,k),(l,m)}^i\), the \(P\) value, and a corresponding \(T_{(j,k),(l,m)}^i\) outcome are determined for each pairwise comparison between the mean song occurrence in the \(j^{th}\) 11-day period of the \(k^{th}\) year and that in the \(l^{th}\) 11-day period of the \(m^{th}\) year from the Tukey tests, as described in Section D.2.1.

We apply the Risch et al. statistical test [120] to the two-year humpback whale
song occurrence daily time series data reported in Vu et al. [126] with the same statistical test settings used to obtain the \( P \) values and results reported in the ‘Risch et al. Statement’ column of Table D.5. The Vu et al. [126] daily humpback whale song occurrence time series (Figure 3 of Ref. [126]) over the \( i^{th} \) 33-day period, denoted by \( \Xi_{i,2006} \) for year 2006 and \( \Xi_{i,2008} \) for year 2008, are input to the Tukey tests of the statistical programming language ‘R’, as described in Section D.2.1. For the \( i^{th} \) 33-day period, consisting of three consecutive non-overlapping 11-day periods with indices \( j \) or \( l = 1, 2, 3 \), and year indices \( k \) or \( m = 1 \) for the test year 2006 and \( k \) or \( m = 2 \) for the control year 2008, a value of \( p_{(j,k),((l,m))}^{i} \), the \( P \) value, and a corresponding \( T_{(j,k),((l,m))}^{i} \) outcome are determined for each pairwise comparison between the mean song occurrence in the \( j^{th} \) 11-day period of the \( k^{th} \) year and that in the \( l^{th} \) 11-day period of the \( m^{th} \) year from the Tukey tests, as described in Section D.2.1. From the outcomes \( T_{(j,k),((l,m))}^{i} \), the corresponding \( a_{i,n} \) are determined based on Table D.5. This is repeated for all continuous 33-day periods, where the \( i + 1^{th} \) 33-day period begins 1-day after the \( i^{th} \) 33-day period. Only 33-day periods that have 11-day periods with reported whale song occurrence are included. If data is missing in any day from a 33-day period, then that 33-day period is excluded from both years. False positive rates are then determined from \( a_{i,n} \) via Equations (D.1) and (D.2). The Risch et al. statistical test [120] false-positively finds whales react to sonar in (a) 100\% of the 35 continuous 33-day periods before the “during” period (Table D.1) when no sonar is present; and (b) 98\% of the 50 continuous 33-day periods excluding the “during” period (Table D.1) when no sonar is present. No valid or meaningful conclusions can be drawn from such an overwhelmingly biased statistical test. This specific application of the Risch et al. statistical test [120] has not been previously reported.

When the Risch et al. statistical test [120] is applied to the same humpback whale song occurrence data, \( \Psi_{2008} \) and \( \Psi_{2009} \), reported in Risch et al. [120] over the 33-day period between September 15 and October 17, with 11-day time series indices \( j = 1 \) for the “before” period, \( j = 2 \) for the “during” period, and \( j = 3 \) for the “after” period, and year indices \( k = 1 \) for the test year 2008 and \( k = 2 \) for the control year 2009, as
well as with year indices $k = 2$ for the control year 2008 and $k = 1$ for the test year 2009, the test false-positively finds that whales react to sonar 100% of the time when no sonar is present, indicating self-contradictions in the Risch et al. [120] approach, as shown in Table D.2, which make their analysis and conclusions invalid. This specific application of the Risch et al. statistical test [120] has also not been previously reported.

D.2.3 Model for detectable humpback whale song occurrence

Detectable humpback whale song occurrence for a coherent sensor array can be quantified in terms of local wind-speed-dependent ambient noise for a given spatial distribution of vocalizing humpback whales. The humpback whale song occurrence depends on the presence of at least one singing humpback whale inside the mean wind-dependent detection range of the sensor array. The percentage of time in a day over which a humpback whale is within the mean detection area and is singing corresponds to the measured daily humpback whale song occurrence rate.

The detection range [127–130, 133], $r_d$, is defined as the range from the center of the array at which signals, in this case humpback whale songs, can no longer be detected above the ambient noise, and is the solution of the sonar equation [127, 128, 163–165],

$$NL(v) + DT - AG = RL(r_d(v)) = SL - TL(r_d(v)),$$  \hspace{1cm} (D.3)

where $NL(v)$ is the wind-speed-dependent ambient noise level, $v$ is the wind speed, $DT$ is the detection threshold, $RL$ is the received sound pressure level due to a humpback whale song source level $SL$ undergoing a transmission loss of $TL(r_d(v))$ at range $r_d(v)$ for some given source and receiver depths, and $AG$ is the array gain equal to $10 \log_{10} N_0$ for a horizontal array, where $N_0$ is the number of coherent sensors spaced at half wavelength [127, 128, 163–165]. The capability of sensor arrays with high array gain such as ours to detect sources orders of magnitude more distant in range than a single sensor is standard, well established and well documented in many
textbooks [127, 128, 163–166]. The array gain of our coherent horizontal OAWRS receiver array is 18 dB, which enables detection of whale vocalizations in an ocean acoustic waveguide [127, 164–166] up to either two orders of magnitude lower in SNR or two orders of magnitude more distant in range than a single hydrophone [127, 128, 163–166], which has zero array gain [127, 128, 163–166], by direct inspection of Equation (D.3). We set the detection threshold, DT, such that the sum of signal and noise is detectable at least 5.6 dB [167–170] above the noise. The ambient noise and the received signal are filtered to the frequency band of the source. Further, the wind-speed-dependent ambient noise level is modeled as

\[
NL(v) = 10 \log_{10} \left( \frac{\alpha v^n + \beta}{1 \mu Pa^2} \right)
\]  

where \( n \) is the power law coefficient of wind-speed-dependent ambient noise, \( \alpha \) is the waveguide propagation factor [171] and \( \beta \) corresponds to the constant baseline sound pressure squared in the frequency band of the source. The coefficients \( n, \alpha \) and \( \beta \) are empirically obtained by minimizing the root mean square error between the measured and the modeled ambient noise level as a function of measured wind speed during the OAWRS experiment in the Gulf of Maine [172]. We find \( n \approx 1.2 \) in the frequency range of the observed humpback song units, which is consistent with past ambient noise measurements in high shipping traffic regions [173–176]. (A value of \( n \approx 3 \) would have been consistent with wind-dependent ambient noise with no significant shipping component [89, 177, 178] but a value of \( n \approx 3 \) was not obtained.) The noise levels obtained from Equation (D.4) in Stellwagen Bank are consistent with those reported in Risch et al. [120].

A standard parabolic equation model of the US Navy and the scientific community, Range-dependent Acoustic Model (RAM) [164, 179–182], that takes into account range-dependent environmental parameters is used to calculate the transmission loss \( TL(r_d(v)) \) from the whale location to the sensor in a highly range-dependent continental-shelf environment in the Gulf of Maine including Stellwagen Bank. The
model uses experimentally measured sound speed profiles acquired during the OAWRS 2006 experiment [9] and standard bathymetry data for the Gulf of Maine [183]. Expected transmission loss [184] is determined along any given propagation path from source to receiver by Monte-Carlo simulation over range-dependent bathymetry [183] and range-dependent sound speed structures measured from oceanographic data [3, 9, 41, 185]. An estimate of detection range $\hat{r}_d(v)$ for a given humpback whale song unit source level can be obtained from Equation (D.3) by a minimum mean squared error method. Higher transmission loss occurs in shallower waters due to more intense and pervasive bottom interaction [127, 128, 163–165]. Transmission loss in deeper waters is typically significantly lower due to upward refraction [127, 164] which leads to far less intense and pervasive bottom interaction, as is the case in the deeper waters surrounding Georges Bank [127, 128, 163–165]. Highly directional transmission loss may then occur when there are large depth variations about a receiver. Indeed, this effect makes the detection range of whales in directions to the North of our receiver and Georges bank much greater than in directions to its South where the relatively shallow waters of Georges Bank are found (Figure D-5). The fact that we localized the sources of many whale calls at great distances along shallow water propagation paths on Georges Bank in directions where transmission loss was greater and found negligibly small vocalization rates much closer to the receiver in the deeper waters north of Georges Bank where transmission loss was much less greatly emphasizes the finding that the vocalization rates originating from north of Georges Bank were negligibly small. This indeed is expected based on general behavioral principles [134] since the whales’ dominant prey was on Georges Bank, where the majority of whale vocalizations originated (Figures 1-3 of Ref. [34]), and not in the deeper waters to the North, as we note in Section 2.1 of Ref. [34]. This is also consistent with the historical distribution of humpback whales in the Gulf of Maine during the fall season [186]. The ranges and propagation paths from deep to shallow waters between our receiver array and Stellwagen Bank are very similar to those between our receiver array and the distant whale call sources localized along Georges Bank (Figure D-5). The corresponding transmission losses have negligible differences. The fact that we
localized the sources of many whale calls on Georges Bank but found negligibly small vocalization rates originating from Stellwagen Bank in the “before” or “during” periods, then emphasizes the fact that vocalization rates originating from Stellwagen Bank were negligibly small in these periods. As noted in Section 2.1 of Ref. [34] and Section D.1, this is consistent with the well documented findings that humpback whales migrate away from Stellwagen Bank where herring stocks have collapsed to feed at other locations that support large herring aggregations such as Georges Bank [134]. Our transmission loss calculations with the standard RAM parabolic equation model have been extensively and successfully calibrated and verified with (1) thousands of one-way transmission loss measurements made during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location [9, 42]; (2) thousands of two-way transmission loss measurements made from herring shoal returns and verified by conventional fish finding sonar and ground truth trawl surveys during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location [1, 9, 187]; (3) roughly one hundred two-way transmission loss measurements made from calibrated targets with known scattering properties during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location [86]; and (4) thousands of one-way transmission loss measurements made during a past OAWRS experiment conducted in a similar continental shelf environment [41].

We find that the humpback whale song source levels measured from more than 4,000 song units recorded during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location approximately follow a Gaussian distribution and are in the range 155 to 205 dB re 1 μPa and 1 m (Figure D-8) with a mean of 179.8 dB re 1 μPa and 1 m and a median of 179.4 dB re 1 μPa and 1 m. The high array gain [127–129, 163–165] of our densely sampled, large aperture coherent OAWRS horizontal receiver array used here enables detection of whale songs two orders of magnitude lower in SNR than a single hydrophone, which has no array gain. Our measurements of humpback whale song source levels then have a high dynamic
range and span the wide range of published source levels [188–193], except for those in Ref. [194], which appear to be anomalously low compared to the rest of the literature as has been previously noted in Ref. [190]. The mean and median of our measured source levels match very well (within 0.6 dB) with the median of all published humpback whale song unit source levels of 180 dB re 1 μPa and 1 m [188, 189, 191–194].

Our song unit source levels are determined given our estimated whale positions and waveguide propagation modeling. Results in Figures D-6 and D-7 are computed using our measured whale positions and the median of all published humpback song source levels of 180 dB re 1 μPa and 1 m [188, 189, 191–194], which has negligible difference from our measured median and mean song source levels, for the range of measured humpback singing depths of 2 m to 25 m [191, 195]. Results in Figures D-6 and D-7 are insensitive to variations in whale position variations within the errors we report for our measured whale positions in Section 3.3 of Ref. [34], and so are insensitive to the whale position errors of our measurement system. Insensitivity here means the measured to modeled song occurrence match is within ± 18% as in Figure D-7.

The total humpback whale song occurrence in a day detectable by a sensor in varying wind speeds is

\[
T_{\text{song}} = \int_0^{T_{\text{day}}} S(t) dt, \quad (D.5)
\]

where \( S(t) = 1 \) when \( \hat{r}_d(v(t)) \) is greater than or equal to the minimum of \( r_i(t) \) over all \( i \), and \( S(t) = 0 \) when \( \hat{r}_d(v(t)) \) is less than the minimum of \( r_i(t) \) over all \( i \), where \( i = 1, 2, ..., N_w \), \( N_w \) is the total number of singing whales, \( v(t) \) is the measured wind speed, \( r_i(t) \) is the range of the \( i^{th} \) singing humpback whale from the sensor at time \( t \), and \( T_{\text{day}} \) is the full diurnal time period of 24 hours. The detectable humpback whale song occurrence rate is then \( \frac{T_{\text{song}}}{T_{\text{day}}} \).
D.2.4 Autocorrelation of annual humpback whale song occurrence time series in 2006 and 2008

We calculated the normalized autocorrelation function \( [196] \) of the Vu et al. \([126]\) 2006 and 2008 annual humpback whale song occurrence time series. The autocorrelation function at zero time lag, where perfect temporal correlation exists, is one. The time lag at which the autocorrelation function falls to \( 1/e \) is the e-folding time scale defining the width of the correlation peak, or coherence time scale, within which processes are conventionally taken to be correlated \([196, 197]\). The e-folding time scale of the Vu et al. \([126]\) annual humpback whale song occurrence time series is 18 days for 2006 and 21 days for 2008 (Figure D-9). The roughly 20-day coherence time scale shows that the humpback song occurrence gradually changes over periods longer than the 11-day periods analyzed in Risch et al. \([120]\). This time is consistent with the smooth and gradual transition in Figure D-4 of the correlation coefficient of 11-day periods across years from negative values in the “before” period to some of the highest positive correlations obtained between years in the “during” period, which contradicts the results of the Risch et al.\([120]\) study and is consistent with a violation of temporal causality in the Risch et al. \([120]\) study. It is noteworthy that (1) the humpback song occurrence dropped to zero in the “before” period, and (2) only after a time period consistent with the measured coherence time scale of song occurrence, within which temporal processes are correlated, did song occurrence begin to increase in the “during” period (Figure D-2). The Risch et al. \([120]\) analysis then also violates temporal causality because the correlated processes that caused the reduction in humpback song occurrence started days before the OAWRS survey transmissions began, yet the analysis and conclusions of Risch et al. \([120]\) offer no other explanation than these OAWRS survey transmissions for the reduction, when only other causes are causally possible. Indeed as we have shown in Section D.1 non-sonar causes regularly lead to such changes in song occurrence, and as we have shown in Section D.2.3 standard detection range variations from measured wind speed dependent noise variations at Stellwagen and measured humpback whale song sources
near Georges Bank completely account for the changes reported in Risch et al. [120].

Table D.1: Percentage of time the Risch et al. statistical test [120] incorrectly finds whales respond to sonar when no sonar is present using annual humpback whale song occurrence data reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise.

<table>
<thead>
<tr>
<th>Analysis period</th>
<th>Excluding “during” period</th>
<th>Before “during” period</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of time with false-positive response</td>
<td>98.0%(49/50)</td>
<td>100%(35/35)</td>
</tr>
</tbody>
</table>

Risch et al. statistical test [120] is applied to all continuous 33-day periods, as described in Section D.2.2, in the annual humpback whale song occurrence reported from single sensor detections at Stellwagen Bank in 2006 and 2008 [126], with 2006 as the test year and 2008 as the control year. The test false-positively finds humpback whales react to sonar 98-100% of the time over a yearly period when no sonars are present. The fraction of time when the Risch et al. statistical test [120] false-positively finds whales react to sonar is given in the parenthesis. The parenthetical numbers in the denominator represent the total number of 33-day periods with no sonar present within the analysis period and the parenthetical numbers in the numerator represent the number of 33-day periods when the Risch et al. statistical test [120] false-positively finds whales react to sonar when no sonar is present. aThe “during” period is defined in Section D.1.

Table D.2: The Risch et al. statistical test is applied to the same humpback whale song occurrence data reported in Risch et al. [120] over the 33-day period from September 15 to October 17 for 2008 and 2009, with either of these two years as the test year and the other as the control year. It false-positively finds that whales react to sonar 100% of the time when no sonar is present, indicating self-contradictions in the Risch et al. [120] approach.

<table>
<thead>
<tr>
<th>Risch et al. statistical test</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>With 2008 as the test year and 2009 as the control year</td>
<td>False positive response</td>
</tr>
<tr>
<td>With 2009 as the test year and 2008 as the control year</td>
<td>False positive response</td>
</tr>
</tbody>
</table>
Table D.3: Typical anthropogenic noise sources at Stellwagen Bank.

<table>
<thead>
<tr>
<th>Continuous anthropogenic noise source</th>
<th>Source level in dB re 1 μPa and 1 m</th>
<th>Frequency in Hz</th>
<th>Source range in km for received level above 120 (^a) dB re 1 μPa</th>
<th>Source range in km for received level between 88-110 (^b) dB re 1 μPa</th>
<th>Acoustic intensity in Watts/m(^2) 1 m away from anthropogenic noise source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cruise ship</td>
<td>219 [198]</td>
<td>10 to &gt;1,000 [199]</td>
<td>&lt;100</td>
<td>160 to &gt;200</td>
<td>5,000</td>
</tr>
<tr>
<td>Cargo vessel</td>
<td>192 [127, 200]</td>
<td>10 to &gt;1,000 [127, 200]</td>
<td>&lt;10</td>
<td>30-200</td>
<td>10</td>
</tr>
<tr>
<td>Research vessel</td>
<td>166-195 [198]</td>
<td>40 to &gt;1,000 [142, 198]</td>
<td>&lt;6</td>
<td>2-130</td>
<td>0.025-20</td>
</tr>
<tr>
<td>Outboard motor boat</td>
<td>176 [143, 201]</td>
<td>100 to &gt;1,000 [202, 203]</td>
<td>&lt;2</td>
<td>3-20</td>
<td>0.25</td>
</tr>
<tr>
<td>Whale watching boat</td>
<td>169 [204]</td>
<td>100 to &gt;1,000 [204]</td>
<td>&lt;1</td>
<td>3-25</td>
<td>0.05</td>
</tr>
</tbody>
</table>

\(^a\)Recommended received pressure level in the NOAA guideline for continuous-type sources [149]. \(^b\)Range of received pressure level at Stellwagen Bank single sensor reported by Risch et al. of OAWRS impulsive signal [120], of roughly 1-2 seconds duration and at least 75 seconds spacing between impulses. Source ranges are determined at the frequencies with maximum humpback whale vocalization energy, using the waveguide propagation methods described in Section D.2.3. Humpback whale vocalizations are known to have source levels in the range of 175 to 188 dB re 1 μPa and 1 m [188-190, 193], and have been reported to go up to 203 dB re 1 μPa and 1 m [205]. All data shown in the table is for sources and measurements in water where \(L_{s,\text{water}} = L_w + 171\) based on the sound speed and density of water, \(L_w\) is the power level in dB re 1 Watt, and \(L_{s,\text{water}}\) is the source level in dB re 1μPa and 1 m. Underwater noise from a typical low flying jet airplane [151] can lead to underwater sound pressure levels exceeding 120 dB re 1μPa in water at ranges less than 5 kilometers.
Figure D-1: Humpback song occurrence rate is constant in the periods “before” and “during” OAWRS survey transmissions. The mean percentage of a diurnal cycle containing humpback whale song in the periods “before” and “during” OAWRS survey transmissions, as defined in Section D.1, remains constant, indicating the transmissions had no effect on humpback whale song over the entire passive 400-km diameter survey area of the Gulf of Maine including Stellwagen Bank.
Figure D-2: Reported humpback whale Stellwagen Bank song occurrence [126] shows large natural variations within and across years. Large natural variations in humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise within and across years are common in the absence of sonar. Line plots of reported single sensor daily humpback whale song occurrence at Stellwagen Bank in hours/day (A) for the entire year and (B) from September 15 to October 17, in 2006 and 2008 [126]. Many periods lasting roughly weeks where high song occurrence episodes are found in one year but not in another when no sonars are present are indicated by black arrows in (A). The reported reducing change in humpback whale song occurrence, to zero [120, 126], occurred in the “before” period while the OAWRS vessels were inactive and docked on the other side of Cape Cod from Stellwagen Bank, at the Woods Hole Oceanographic Institution, due to severe winds for days before OAWRS transmissions for active surveying began on September 26, 2006, as marked by the black arrow in (B). This shows that Risch et al. [120] analysis violates temporal causality.
Figure D-3: Quantifying large differences in the reported humpback whale song occurrence at Stellwagen Bank [126] across years. Difference in humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise across years exceeds that of the “during” period most of the time when no sonars are present. (A) Difference in mean humpback whale song occurrence at Stellwagen Bank over respective 11-day periods with 1-day increment in 2006 and 2008, (B) histogram of difference in mean humpback song occurrence over 11-day periods between 2006 and 2008 when no sonar is present, i.e. excluding the “during” period from September 26 to October 6. Periods when the difference in means of respective 11-day periods is greater than (red dots) and less than (blue dots) that of the “during” period are indicated in (A). The difference in means fluctuates randomly throughout the year, exceeding the “during” period 57.8% of the time (most of the time) when no sonars are present, indicating that there is nothing unusual about such differences, which are actually the norm.
Figure D-4: Reported annual humpback song occurrence at Stellwagen Bank [126] are uncorrelated between years over 11-day periods. Annual humpback whale song occurrence reported from single sensor detections at Stellwagen Bank [126] in time-dependent ambient noise are uncorrelated over 11-day periods across years. (A) Correlation coefficient between 2006 and 2008 humpback whale song occurrence time series over 11-day period with 1-day increment (B) histogram of the correlation coefficient in (A). The correlation coefficient of the annual humpback whale song occurrence time series over 11-day periods across years obeys a random distribution peaking at zero correlation about which it is symmetric, showing that correlation in trend between years is random and quantitatively expected to be zero with roughly as many negative correlations as positive ones. The correlation coefficient between the humpback whale song occurrence across years smoothly transitions from negative values in the "before" period, showing no similarity or relation in trend between years just before the 2006 OAWRS survey transmission period, to some of the highest positive correlations obtained between years in the "during" period. This demonstrates high similarity and relation in trend between years during the 2006 OAWRS active survey transmission period, which contradicts the results of the Risch et al. [120] study.
Figure D-5: Wind-dependence of mean detection range for single sensor at Stellwagen Bank [120], and OAWRS receiver array. The green shaded areas indicate the overall vocalizing humpback whale call rate densities (number of calls/[(min) (5 nmi)$^2$]) determined between September 22 and October 6, 2006 by our large aperture receiver array towed along several tracks (black lines). The mean detection ranges for the single sensor at Stellwagen Bank are in blue and for the OAWRS receiver array are in red, where Stellwagen Bank is marked by yellow shaded regions. These detection ranges are determined by the methods described in Section D.2.3 given a humpback whale song unit source level of approximately 180 dB re 1 $\mu$Pa and 1 m which is the median of all published humpback whale song source levels [188, 189, 191–194]. The error bars represent the spread in detection range due to typical humpback whale song source level variations (Section D.2.3). Under (A) low wind speed conditions vocalizing whales are within the mean detection area for a single Stellwagen Bank sensor but for (B) higher wind speeds most vocalizing whales are outside the mean detection area of the same sensor, which results in reduction of detectable whale song occurrence by the single sensor [120] at Stellwagen Bank.
Figure D-6: Wind-speed increase causes reduction in humpback song occurrence at Stellwagen Bank. Average wind speed increase from the “before” to the “during” period at Stellwagen Bank causes reduction in the percentage of time humpback whale songs are within mean detection range of a single Stellwagen Bank sensor. (A) Averaged wind speed measured at the NDBC buoy [131] closest to Stellwagen Bank over the “before,” “during,” and “after” 11-day periods; and (B) percentage of the time vocalizing humpback whales localized by our large aperture array are within the mean detection range of the single sensor [120] at Stellwagen Bank in the “before” and “during” periods, using waveguide propagation methods and whale song parameters described in Section D.2.3. Since the OAWRS experiment was conducted only up to October 6, 2006, the humpback whale source distribution in the “after” period was not measured and we do not investigate the percentage of time that humpback whales are within the mean detection range of the single sensor at Stellwagen Bank [120] for the “after” period. The triangles represent the mean wind speed and the solid ticks represent the standard deviation of the wind speed over the respective 11-day periods.
Figure D-7: Humpback song occurrence detectable by single sensor matches reported humpback song occurrence at Stellwagen Bank [120]. Average humpback whale song occurrence detectable by a single hydrophone at Stellwagen Bank in time-dependent ambient noise in the “before” and the “during” periods matches the reported humpback whale song occurrence by Risch et al. [120]. Using the measured wind speeds at Stellwagen Bank [131] (Figure D-6), the measured spatial distribution of vocalizing humpback whales (Figure 1 of Ref. [34]), and constant song production rates (Figure D-1) measured by our large-aperture array, the detectable song occurrence over the “before” and “during” period are found to be within ±18% of the reported means [120], much less than the standard deviations of reported song occurrence [120], using waveguide propagation methods and whale song parameters described in Section D.2.3. Before and during OAWRS survey transmissions, this figure shows that reported variations in song occurrence at Stellwagen Bank by Risch et al. [120] are actually due to detection range changes caused by wind-dependent ambient noise, through well established physical processes [127, 132].
Figure D-8: Histogram of the measured humpback whale song unit source levels. The humpback whale song unit source levels measured from more than 4,000 recorded song units during the same 2006 Gulf of Maine experiment discussed here at the same time and at the same location approximately follow a Gaussian distribution and are in the range 155 to 205 dB re 1 μPa and 1 m with a mean of 179.8 dB re 1 μPa and 1 m and a median of 179.4 dB re 1 μPa and 1 m, which are within 0.6 dB of the median of all published humpback whale song unit source levels of 180 dB re 1 μPa and 1 m [188, 189, 191–194]. The solid and dashed gray lines represent the mean and the median of the measured humpback song unit source levels, respectively.
Figure D-9: Autocorrelation of Vu et al. [126] humpback whale song occurrence time series in 2006 and 2008. The e-folding time scale $\tau_e$ of the Vu et al. [126] annual humpback whale song occurrence time series is (A) 18 days for 2006 and (B) 21 days for 2008. The roughly 20-day coherence time scale shows that the humpback song occurrence gradually changes over periods longer than the 11-day periods analyzed in Risch et al. [120]. It is noteworthy that (1) the humpback song occurrence dropped to zero in the “before” period, and (2) only after a time period consistent with the measured coherence time scale of song occurrence, within which temporal processes are correlated, did song occurrence begin to increase in the “during” period (Figure D-2). The Risch et al. [120] analysis then violates temporal causality because the correlated processes that caused the reduction in humpback song occurrence started days before the OAWRS survey transmissions began, yet the analysis and conclusions of Risch et al. [120] offer no other explanation than these survey transmissions for the reduction. Both time series show high correlation at a time lag of roughly seven months due to increases in song occurrence during the spring and fall seasons (Figure D-2), separated by roughly seven months.
Table D.4: Received mean intensity of typical anthropogenic noise sources at Stellwagen Bank.

<table>
<thead>
<tr>
<th>Continuous anthropogenic noise source</th>
<th>Received level in water in dB re 1 μPa (or corresponding mean intensity in Watts/m²) 500 m away from an anthropogenic noise source over a minute or longer</th>
<th>How many decibels higher (or times greater) the mean intensity of the given anthropogenic noise source over a minute or longer at 500 m is than that reported for OAWRS at Stellwagen Bank [120]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cruise ship</td>
<td>177 (0.33)</td>
<td>85 (300,000,000)</td>
</tr>
<tr>
<td>Cargo vessel</td>
<td>147 (0.00033)</td>
<td>55 (300,000)</td>
</tr>
<tr>
<td>Research vessel</td>
<td>121-144 (0.00000083-0.00017)</td>
<td>29-52 (750-150,000)</td>
</tr>
<tr>
<td>Outboard motor boat</td>
<td>131 (0.0000083)</td>
<td>39 (7,500)</td>
</tr>
<tr>
<td>Whale watching boat</td>
<td>124 (0.0000017)</td>
<td>32 (1,500)</td>
</tr>
</tbody>
</table>

*Whale watching vessels [206] are allowed to approach humpback whales at ranges much less than 500 m according to NOAA Whalewatching Guidelines [207].
Table D.5: Risch et al. statistical test statements [120].

<table>
<thead>
<tr>
<th>Risch et al. Statement</th>
<th>Algorithmic representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 “While ‘before’ and ‘after’ periods differed significantly within the years 2008 and 2009 (P &lt; 0.001), with more song recorded in the later period in both years, this increase was not significant in 2006 (P = 0.2147).”</td>
<td>If $T_{(1,1),(3,1)}^i = T_{(1,k),(3,k)}^i$ for all $k \neq 1$, then $a_{i,1} = 0$, otherwise $a_{i,1} = 1$.</td>
</tr>
<tr>
<td>2 “In 2006, the ‘during’ period was significantly different from the period ‘after’ (P = 0.0093), with more song recorded later. The 2006 ‘during’ period was not detectably different from the period ‘before’ (P = 0.5226).”</td>
<td>If $T_{(2,1),(3,1)}^i = X(\mu_{(2,1)}^i &lt; \mu_{(3,1)}^i)$ or $T_{(2,1),(3,1)}^i = X(\mu_{(2,1)}^i \not\approx \mu_{(3,1)}^i)$ AND $T_{(1,1),(2,1)}^i = Y(\mu_{(1,1)}^i &lt; \mu_{(2,1)}^i)$ or $T_{(1,1),(2,1)}^i = Y(\mu_{(1,1)}^i \not\approx \mu_{(2,1)}^i)$, then $a_{i,2} = 0$, otherwise $a_{i,2} = 1$.</td>
</tr>
<tr>
<td>3 “When comparing the ‘during’ period across years, 2006 differed significantly from 2009 (P = 0.0057). The same time period did not differ significantly between 2006 and 2008 (P = 0.1842), or between 2008 and 2009 (P = 0.4819).”</td>
<td>If $T_{(2,1),(2,k)}^i = Y(\mu_{(2,1)}^i \not\approx \mu_{(2,k)}^i)$ for all $k &gt; 1$, then $a_{i,3} = 1$, otherwise $a_{i,3} = 0$.</td>
</tr>
<tr>
<td>4 “Yet, overall there was considerably less song recorded in the 11 ‘during’ days in 2006 compared to both 2008 and 2009.”</td>
<td>If $\mu_{(2,1)}^i \not\approx \mu_{(2,k)}^i$ for all $k &gt; 1$, then $a_{i,4} = 0$, otherwise $a_{i,4} = 1$.</td>
</tr>
</tbody>
</table>

Table D.6: Possible outcomes of each pairwise comparison between the mean humpback whale song occurrence in the $j^{th}$ 11-day period of the $i^{th}$ 33-day period in the $k^{th}$ year and that in the $l^{th}$ 11-day period of the $i^{th}$ 33-day period in the $m^{th}$ year in the Risch et al. statistical test [120].

<table>
<thead>
<tr>
<th>Outcome $T_{(j,k),(l,m)}^i$</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X(\mu_{(j,k)}^i &lt; \mu_{(l,m)}^i)$</td>
<td>Means are not significantly different and $\mu_{(j,k)}^i &lt; \mu_{(l,m)}^i$</td>
</tr>
<tr>
<td>$X(\mu_{(j,k)}^i \not\approx \mu_{(l,m)}^i)$</td>
<td>Means are not significantly different and $\mu_{(j,k)}^i \not\approx \mu_{(l,m)}^i$</td>
</tr>
<tr>
<td>$Y(\mu_{(j,k)}^i &lt; \mu_{(l,m)}^i)$</td>
<td>Means are significantly different and $\mu_{(j,k)}^i &lt; \mu_{(l,m)}^i$</td>
</tr>
<tr>
<td>$Y(\mu_{(j,k)}^i \not\approx \mu_{(l,m)}^i)$</td>
<td>Means are significantly different and $\mu_{(j,k)}^i \not\approx \mu_{(l,m)}^i$</td>
</tr>
</tbody>
</table>
Appendix E

Comment on "Variation in humpback whale (Megaptera novaeangliae) song length in relation to low-frequency sound broadcasts [J. Acoust. Soc. Am. 113, 3411-3424 (2003)]"

E.1 Introduction

Concerns about the effect of anthropogenic sound on marine life [208] and recent advances in passive listening capability and monitoring technology enabled many behavioral response studies [209-212], which are primarily responsible for a large increase in the number of publications in the related field. However, little attention has been given to the quality of the studies [213] compared with that given in other fields of study [214-221].

We find the null hypothesis of Fristrup, Hatch, and Clark [222] ("oblivious hypothesis") is based on an incorrect and hidden assumption that humpback song length
follows a uniform probability distribution. The uniform song length distribution assumption is inconsistent with and contradicts the measured humpback song length when no sonar is present given in Fristrup, Hatch, and Clark [222]. When this measured humpback song length distribution when no sonar is present is used instead of the assumed uniform song length distribution, the null hypothesis of no effect is found to be statistically indistinguishable from humpback song length distributions measured with sonar present given in Fristrup, Hatch, and Clark [222]. The conclusion of Fristrup, Hatch, and Clark [222] that the length of humpback song is affected by sonar is then apparently inconsistent with and not supported by the data presented in Fristrup, Hatch, and Clark [222].

In particular, we calculate the joint probability density function of humpback song length and overlap with active acoustic transmissions given an assumption that humpback songs and active acoustic transmissions occur independently in time. In doing so, we use two humpback song length distributions for comparison: (1) a uniform
distribution of finite support and (2) a non-uniform and peaked distribution of finite support given as a control humpback song length distribution measured before any active transmissions on each day by Fristrup, Hatch, and Clark [222]. The measured relationship between humpback song length and overlap with acoustic transmissions is then reproduced using the calculated joint probability density functions and the number of song length samples reported in Fristrup, Hatch, and Clark [222]. We then determine the false positive rate of the Fristrup, Hatch, and Clark statistical method [222], which is used to detect humpback whale response, or “deviation from prediction” of the null hypothesis in Fristrup, Hatch, and Clark [222].

We find that the joint probability density function of humpback song length and overlap with acoustic transmissions determined using a uniform song length distribution is consistent with the null hypothesis of Fristrup, Hatch, and Clark [222]. This indicates that the null hypothesis of Fristrup, Hatch, and Clark [222] implicitly assumes a uniform humpback song length distribution. In Fristrup, Hatch, and Clark [222], a whale response to sonar or “deviation from prediction” [222] is found by comparing the measured relationship between humpback song length and overlap with acoustic transmissions [222] with the prediction of the null hypothesis of Fristrup, Hatch, and Clark [222] that assumes a uniform song length distribution with no justification.

The uniform song length distribution is apparently unlikely to occur given that the control song length distribution measured before any active transmissions on each day reported in Fristrup, Hatch, and Clark [222] is non-uniform and peaked [222]. We find that the joint probability density function of humpback song length and overlap with transmissions determined using this non-uniform and peaked song length distribution measured without any active transmissions [222] is consistent with the measured relationship between humpback song length and overlap with acoustic transmissions reported in Fristrup, Hatch, and Clark [222]. We then show the Fristrup, Hatch, and Clark statistical method [222] false positively finds whale response 99-100% of the time when the non-uniform and peaked song length distribution is used, suggesting
Figure E-2: (Left) Uniform song length distribution incorrectly and implicitly assumed in the analysis of Fristrup, Hatch, and Clark [222]. The implicit assumption that song length distribution is uniform in the analysis in Fristrup, Hatch, and Clark [222] is derived because there is no other way to obtain the predictions of the null hypothesis of Fristrup, Hatch, and Clark [222] from the calculated probability density of overlap conditioned on song length without using the uniform song length distribution. (Right) Non-uniform and peaked distribution of finite support given as a control humpback song length distribution measured before any active transmissions on each day [222]. This non-uniform and peaked song length distribution is obtained from the measured song length histogram when no sonar is present with bin counts specified in FIG. 8 of Fristrup, Hatch, and Clark [222].

E.2 Joint probability density function of humpback song length and overlap with active acoustic transmissions

Here the joint probability density function of humpback song length and overlap with acoustic transmissions is calculated with an assumption that humpback songs and active transmissions occur independently. For the given humpback song length $Y = y$, we increase relative time between the beginnings of song and active transmissions with a 1-second increment and determine overlap $X$ between humpback song
and active transmission to calculate conditional probability density $f_{X|Y=y}(x|y)$. We use humpback song lengths ranging from 3 to 30 minutes and the active transmission sequence lasting for roughly 54 minutes with 30-minute time periods without transmissions added to the transmission sequence at its beginning and at the end for numerical calculation. Within one transmission sequence, ten 42-second signals are transmitted with 6-minute intervals [222]. The determined conditional probability density function $f_{X|Y=y}(x|y)$ is then multiplied by the marginal probability density function of humpback song length $f_Y(y)$ such that joint probability density function of humpback song length and overlap with acoustic transmissions $f_{XY}(x,y)$ is obtained. Here two marginal probability density functions of humpback song length are used for comparison: (1) a uniform distribution of finite support from 3 to 30 minutes (Left panel of FIG. E-2) and (2) a non-uniform and peaked distribution of finite support (Right panel of FIG. E-2) reported as a control humpback song length distribution before active transmissions in Fristrup, Hatch, and Clark [222].
Figure E-4: (Left) Measured relationship between humpback song length (black crosses) and minutes of overlap with active acoustic transmissions. The three black solid lines represent the relationship between the humpback song length and the ping overlap for 42-second, 84-second, and 126-second ping overlaps predicted by the “oblivious hypothesis” of Fristrup, Hatch, and Clark [222]. The solid lines are drawn through the median values for 42-second, 84-second, and 126-second ping overlaps with a slope equal to the inverse of the duty cycle of active acoustic transmissions. The solid lines are found to be consistent with the relationship between song length and minutes of overlap assuming a uniform song length distribution. (Right) The relationship between song length and minutes of overlap assuming a non-uniform and peaked song length distribution is consistent with the measured relationship given in left panel. The mean song length of the simulated samples for each overlap obtained from the joint probability density function using the non-uniform and peaked song length distribution (Right panel) is found to be within one standard deviation from the measured mean song length for each overlap. Red crosses represent the median song length for corresponding minutes of overlap with acoustic transmissions. No information is provided in Fristrup, Hatch, and Clark [222] to calculate the probability density for zero overlap.

We find that the relationship between humpback song length and overlap with acoustic transmissions (FIG. E-3) significantly varies depending on the humpback song length distribution (FIG. E-2). We find that the joint probability density function determined using a uniform song length distribution (Left panel of FIG. E-3) is consistent with the null hypothesis of Fristrup, Hatch, and Clark [222] (White dashed line in left panel of FIG. E-3). The joint probability density function determined using a non-uniform and peaked song length distribution (Right panel of FIG. E-3) is found to be consistent with the measured relationship between song length and overlap with acoustic transmissions (Three right-most columns of crosses in left panel of
E.3  Fristrup, Hatch, and Clark statistical method to detect deviation from null hypothesis

The Fristrup, Hatch, and Clark statistical method [222] applies the student t-test for pairwise comparison of song length measurements that have distinct minutes of overlaps with active acoustic transmissions, after adjusting for the slope expected by the null hypothesis. The input to the method is $M$ sets of humpback song lengths with distinct overlaps. Each pairwise comparison between song length with the $T \times i$-second overlap and that with the $T \times (i + 1)$-second overlap is assigned a value of $p_{i,i+1}$, where the time duration of active transmission unit $T = 42$. The value of $p_{i,i+1}$ is the probability that the value of the student’s t-test statistic is greater than the observed value of the test statistic, conditioned on the null hypothesis, i.e. two mean humpback song lengths are the same after adjusting for the slope expected by the null hypothesis. This $p_{i,i+1}$ is referred to as the “p-value” in Fristrup, Hatch, and
Clark [222]. If \( p_{i,i+1} \) is less than a convention for test of statistical significance, or 0.05 used in Fristrup, Hatch, and Clark [222], then the difference in means is classified by the user to be statistically significant, otherwise it is classified by the user to be not statistically significant.

Suppose there are song lengths with overlaps with active transmissions over \( T \times j \) seconds, where \( j = 1, ..., M \). Let \( \mu_j \) be the mean humpback song length with \( T \times j \)-second overlap, where \( j = 1, ..., M \). For a given maximum number of overlaps \( M \), there are \( M - 1 \) pairs for comparison. Comparing the \( p_{i,i+1} \) with a convention for test of statistical significance, or 0.05 used in Fristrup, Hatch, and Clark [222], outcome \( Q_{i,i+1} \) is assigned for the comparison between the mean song length pair \( \mu_i \) and \( \mu_{i+1} \), where \( i = 1, ..., M - 1 \). The possible outcomes \( Q_{i,i+1} \) are (1) \( SD \), which is defined as: the difference between \( \mu_i \) and \( \mu_{i+1} \) is statistically significant; and (2) \( ND \), which is defined as: the difference between \( \mu_i \) and \( \mu_{i+1} \) is not statistically significant.

The rate of false positive findings for statistically-significant differences between song lengths with overlaps of length \( T \times i \) seconds and \( T \times (i + 1) \) seconds is

\[
FPR_{i,i+1} = \frac{NSD_{i,i+1}}{N_{i,i+1}},
\]

(E.1)

where \( NSD_{i,i+1} \) is the number of outcomes \( Q_{i,i+1} = SD \) from Monte Carlo simulation sets using a peaked song length distribution given in Fristrup, Hatch, and Clark [222], \( N_{i,i+1} \) is the total number of Monte Carlo simulations using a peaked song length distribution given in Fristrup, Hatch, and Clark [222], and \( i = 1, ..., M - 1 \).

### E.4 False positive rate of Fristrup, Hatch, and Clark statistical method

To quantify statistical bias of the method used to detect whale response, or deviation from prediction of the null hypothesis, described in Fristrup, Hatch, and Clark [222]
and Section E.3, humpback song length samples are generated (Black crosses in right panel of FIG. E-4) using (1) the number of song length samples reported in Frístrup, Hatch, and Clark [222] for each value of overlap (Numbers on top of FIG. E-4 left panel) and (2) conditional probability density function $f_{Y\mid X=x}(y\mid x)$ of humpback song length $Y$ given an overlap value $X = x$. The conditional probability density function $f_{Y\mid X=x}(y\mid x)$ of humpback song length $Y$ given a value for overlap $X = x$ is calculated from the joint probability density function $f_{XY}(x, y)$ obtained using a non-uniform and peaked song length distribution before any active transmissions shown in the right panel of FIG. E-2 [222]. The generated set of song length samples simulates the expected song length measurements when humpback song length follows a non-uniform and peaked probability distribution. The Frístrup, Hatch, and Clark statistical method mathematically described in Section E.3 is then applied to the simulated set of song lengths to compare the mean humpback song lengths with 42 $\times$ 2-second and 42 $\times$ 3-second overlaps against the previous categories with 42 $\times$ 1-second and 42 $\times$ 2-second overlaps, respectively, after adjusting for the slope expected by the null hypothesis. Here we generate $N = 500$ Monte Carlo simulated sets of humpback song lengths.

We find the measured relationship between song length and acoustic transmissions (Left panel of FIG. E-4) is consistent with the relationship between the simulated song length and overlap with transmissions obtained using a peaked song length distribution (Right panel of FIG. E-4). We find that the Frístrup, Hatch, and Clark method [222] false positively finds the differences between simulated song lengths with 42 $\times$ 2-second overlap with active transmissions and those with 42 $\times$ 1-second overlap 100 % (500/500) of the time, after adjusting for the slope expected by the null hypothesis, suggesting false positive detection of deviation from prediction of the null hypothesis, or whale response, 100 % of the time. Similarly, we find that the Frístrup, Hatch, and Clark statistical method [222] false positively finds the differences between simulated song lengths with 42 $\times$ 3-second overlap with transmissions and those with 42 $\times$ 2-second overlap 99 % (497/500) of the time, after adjusting for the
slope expected by the null hypothesis. No meaningful conclusions can be drawn from a statistical method with such high statistical bias.

E.5 Conclusions

In summary, we have shown the null hypothesis of Fristrup, Hatch, and Clark [222] ("oblivious hypothesis") is based on a hidden assumption that humpback song length in the absence of sonar follows a uniform probability distribution which is apparently inconsistent with measured data presented in Fristrup, Hatch, and Clark [222] and so incorrect. The uniform song length distribution is inconsistent with and contradicts the measured humpback song length when no sonar is present given in Fristrup, Hatch, and Clark [222]. When this humpback song length distribution when no sonar is present is used instead of the assumed uniform song length distribution, the obtained relationship between song length and minutes of overlap with active acoustic transmissions is found to have statistically insignificant difference from that given in Fristrup, Hatch, and Clark [222] with sonar present. This significantly differs from the conclusion of Fristrup, Hatch, and Clark [222] that the relationship between song length and minutes of overlap is affected by sonar. The statistical method used to detect whale response, described in Fristrup, Hatch, and Clark [222], is found to have a 99-100 % false positive rate, and so no meaningful conclusions can be made from a method with such high statistical bias. Similar problems with the use of incorrect null hypothesis have been previously discussed in the environmental literature [217]. The results of this Appendix E then have significant implications for improving the assessment of environmental impacts in the ocean.
Appendix F

Preliminary analysis on the herring shoal population in Ålesund, Norway

Herring shoal populations detected by OAWRS in the Ålesund area are found to be roughly between 7 and 66 million from the OAWRS measurements taken at 23:38:59-01:23:59 UTC on February 18-19, 2014 and between 4 and 38 million from the OAWRS measurements taken at 22:54:19-00:22:39 UTC on February 20-21, 2014 (Fig. F-1). The herring shoal population estimates are subject to large variations (Fig. F-1) due to changes in herring vertical distributions and neutral buoyancy. It is found that the target strength of Ålesund herring can easily vary by more than ±3 dB depending on occupancy depth and neutral buoyancy depth, leading to a large variation in the herring population estimate by roughly a factor of 3. Roughly up to 10 dB variation in the target strength of herring is possible for extreme cases. This is because the resonance peak of herring scattering occurs near 955 Hz. The herring shoal population of Ålesund herring is found to be relatively lower than that of Georges Bank herring. This can be attributed to (1) range limitations due to a number of fish attenuation issues during the Nordic Seas Experiment in 2014; and (2) the fact that the peak spawning of Ålesund herring had ended and most of the herring had migrated south at the time of the experiment according to the captain of the Fishing Vessel Artus and others. There is also thought that the herring shoals that we observed in the Ålesund area had spawned already, and so the behavioral purpose and forces leading
The herring shoal populations in the Ålesund area were obtained by integrating herring density within the shoal boundary defined at a boundary density of 0.11 fish/m² from the OAWRS measurements made at 955 Hz on February 18-21, 2014 (Fig. F-2). The target strength of Ålesund herring at 955 Hz $TS$ is calibrated to be -29.5 dB re 1 m from simultaneous measurement of herring by OAWRS and the vertical echosounder from 01:51:50 to 01:55:59 UTC on February 21, 2014 via $TS \approx SS - 10 \log_{10} n_A$, where $SS$ is the scattering strength from OAWRS images taken at 955 Hz, and $n_A$ is the herring areal population density obtained from vertical echosounder measurements. The shoal boundary density of 0.11 fish/m² for Ålesund herring $c_{Ålesund}$ was obtained from scaling boundary density for Georges Bank herring $c_{GB}$ by fish body length ratio squared, i.e. $c_{Ålesund} = \left( \frac{L_{GB}}{L_{Ålesund}} \right)^2 c_{GB}$, where $L_{GB}$ is the body length of Georges Bank herring, and $L_{Ålesund}$ is the body length of Ålesund herring.

Figure F-1: Shoal population estimates of Ålesund herring on February 18 and 20, 2014
Figure F-2: Horizontal distribution of herring population density observed with OAWRS (Left) at 23:38:59-01:23:59 UTC on February 18-19, 2014, (Right) at 22:54:19-00:22:39 UTC on February 20-21, 2014
Bibliography


179


[59] Multifrequency backscattering properties of atlantic herring (clupea harengus) and norway pout (trisopterus esmarkii), author=Fässler, Sascha Mm and Santos, Rita and García-núñez, Norma and Fernandes, Paul G, journal=Canadian Journal of Fisheries and Aquatic Sciences, volume=64, number=2, pages=362–374, year=2007.


189


190


