Passive acoustic monitoring of the deep ocean using ambient noise

A Thesis
Presented to
The Academic Faculty

By

Katherine F. Woolfe

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in the
George W. Woodruff School of Mechanical Engineering

August 2015

Passive acoustic monitoring of the deep ocean using ambient noise

Approved by:

Dr. Karim Sabra
Committee Chair
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Levent Degertekin
School of Mechanical engineering
Georgia Institute of Technology

Dr. Julien Meaud
School of Mechanical Engineering
Georgia Institute of Technology

Dr. William Kuperman
Scripps Institution of Oceanography
University of California

Dr. Emanuele Di Lorenzo
School of Earth & Atmospheric Sciences
Georgia Institute of Technology

Date Approved: June 9, 2015

Acknowledgments

There are many people to whom and am indebted for assistance in various aspects of this thesis formulation and my graduate school experience in general. First I would like to thank my advisor, Karim Sabra, for his inspiration and guidance as I pursued this work. I would also like to thank my fellow graduate student friends- Shima Shahab, Jason Kulpe, Ellen Skow, Shane Lani, Brendan Nichols, Bernie Shieh, Katie Matlack, and Adrienne Little- for their friendship, support, and encouragement. Thank you to my faculty mentors-Aldo Ferri, Christine Valle, Wayne Whiteman, Peter Rogers- who provided valuable advice and insights through the highs and lows of graduate school. A very big thanks to my parents, Jim and Barbara Faist, for encouraging a love of knowledge from a young age and instilling in me the determination that is necessary for the completion of so many years of school. Thank you to my husband, Chris Woolfe, for providing love, patience, perspective, and delicious food for me to eat. Lastly, I give my heartfelt thanks and appreciation to my savior, Jesus Christ, for providing this fascinating world for me to study.

Table of Contents

List of Figures

Summary

In the ocean, changes in the speed of sound can be related to changes in water temperature. By leveraging this relationship, acoustic methods – namely acoustic tomography- have been used to monitor temperature changes in the deep ocean for the purposes of providing inputs to climate change models. Traditionally, these acoustic methods involve loud, active sound sources which can be logistically challenging to operate and have been criticized for potentially disturbing marine animals. Therefore, this work demonstrates a passive acoustic method - previously only used in shallow water for short monitoring durations- that uses only recordings of low-frequency (1-40 Hz) ambient noise to continuously monitor variations in deep ocean temperature with an unprecedented degree of precision and temporal resolution. Numerical simulations were conducted to show the portions of the ocean that are monitored with this passive method. This work also provides recommendations (regarding sensor placement around the world) for future development of a global passive acoustic sensor network that makes use of distant noise sources (sea-ice or seismic sources) to extract meaningful information (whether temperature, currents, etc.) about the ocean. Finally, an optimization method is proposed to overcome one of the fundamental limitations of previous applications of this passive monitoring method: tracking oceanic fluctuations that occur over short time scales. Hence, the results of this study may assist in the development of more reliable climate models that include an enhanced understanding of the ocean's role as a global heat sink. Finally, an optimization method was proposed to enhance the emergence rate

of coherent arrivals from ambient noise correlations, thus allowing this passive monitoring method to track acoustic medium fluctuations on a shorter time scale. This optimization could also be used in other applications of noise-based passive monitoring in a rapidly fluctuating medium (such as seismic monitoring, structural health monitoring, biomedical imaging, etc.).

.

Chapter 1: Introduction

1.1 Motivation

Since deep oceans play a major role in absorbing atmospheric heat, measuring their temperature variations is necessary to quantify air-sea heat exchanges in order – among others- to assess global warming trends [1] and calibrate climate change models [2], [3]. However, in contrast to ocean surface temperatures, deep ocean temperatures cannot be readily inferred from satellite-based remote sensing methods [4]; instead, they are most commonly measured with only limited spatial and temporal resolution using globally sparse, free-drifting profiling oceanographic floats [5]. Acoustic thermometry provides another option for detecting fine variations in deep ocean temperatures over large distances, based on measuring acoustic propagation travel times between sources and receivers inserted in the ocean [6], [7]. However, these active acoustic monitoring methods typically involve sounds sources with large power requirements and have been criticized for potentially causing disturbance to marine animals. An emerging alternative to measurements with active acoustic sources is the use of ambient noise correlation processing.

The goal of this work is to use the coherent portions of the ambient noise to monitor the ocean, developing passive modality of the active counterpart of acoustic thermometry (i.e. using acoustic travel times to track deep ocean temperatures for the purposes of climate monitoring). This dissertation focuses on obtaining a signal from a random underwater noise field, and then using that signal to monitor the ocean. A signal

is desired information, while noise is typically viewed as an annoying, unwanted distortion of that signal. Work within the past decade has focused on extracting useful information from ambient, diffuse noise fields [8], [9], [10]. Specifically, an estimate of the Green's function between two points can be obtained by extracting the coherent noise that passes through both sensors. By averaging over time, the coherent portions of the noise emerge from the mostly incoherent noise [11], [12]. These coherent portions can be used to track ocean fluctuations- such as temperature fluctuations- that occur on a time scale greater than the averaging time used to extract the coherent wave fronts.

1.2 Ocean Ambient Noise

When obtaining a coherent signal from ambient noise, it is important to understand the properties of the ambient noise field (e.g., frequency content, noise sources spatial origin and mechanisms, etc.). The ocean is filled with many types of noise, including shipping noise, surface noise, biological noise, and seismic noise. Each noise source has a particular frequency band. When recording underwater noise, it is of paramount importance to understand which noise sources are likely to show up in the data. Infrasound (sound below the range of human hearing) is considered to be less than 20 Hz and is predominately due to seismic noise and internal waves, although this dissertation also discusses the possible effects of noise generated by sea ice moving and cracking in the Polar Regions. Man-made sounds tend to dominate the spectrum from 20 Hz – 500 Hz. Sound caused by marine animals (whales, dolphins, etc.) can also contribute to noise within this frequency band. Attenuation is so low within this frequency band that sound can propagate 1000+ km with the help of the Sound Fixing

and Ranging (SOFAR) waveguide [13]. Noise from 500 Hz to 50 kHz is primarily due to ocean surface noise, although biological noise is also present. The surface noise is generated by movement of the wind, waves, and bubbles. Noise above 50 kHz is caused by the thermal-mechanical random motion of individual molecules, called Brownian noise [14]. A diagram illustrating the sources of ambient ocean noise from 1 Hz – 100 kHz is shown in Figure 1. This diagram is a representation of multiple researchers' work and was compiled by Wenz [15].

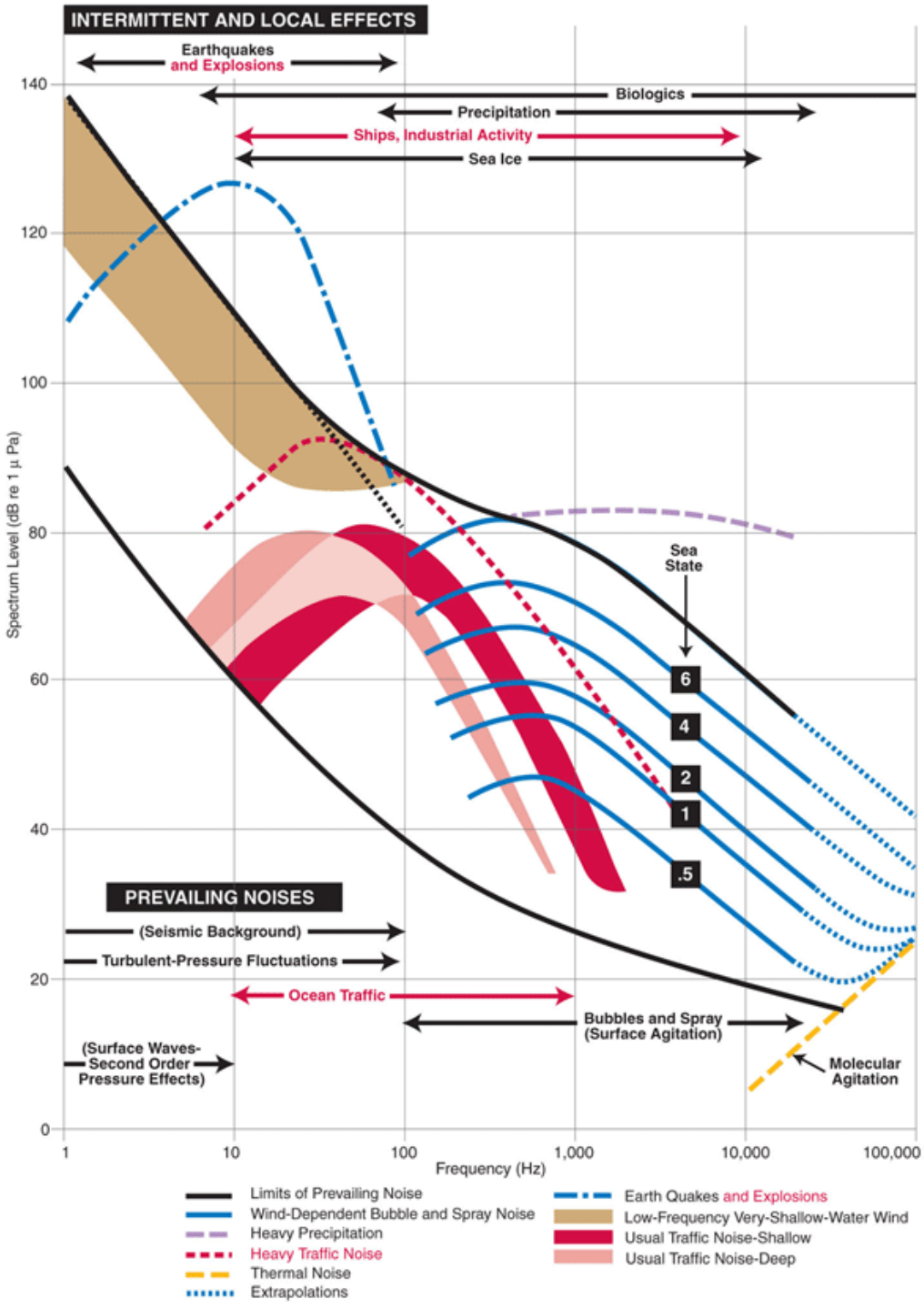

Figure 1: Ambient noises in the ocean from different sources and levels [15].

1.3 Overview of Acoustic Tomography

Global climate change research has generated a need for precise temperature measurements of the deep ocean. Over the past 50 years, the oceans have absorbed about 90% of the total heat added to the climate system, and the deep oceans (depth > 700 m) continue to increase in temperature although the surface layers appear to have stabilized [16], [17]. The data which are largely responsible for these assessments come from the Argo float program, which consists of thousands of autonomous floats distributed throughout the oceans. These floats sample the upper 2000 m of ocean, recording hydrostatic pressure and temperature in the water column. These data can be spatially interpolated using a least-squares algorithm that averages data collected from the nearest 100 float profiles, creating global estimates for ocean temperature [5]. While these estimates are likely precise enough when describing upper ocean temperature changes, where seasonal variations are on the order of 3°C near Ascension Island, uncertainty grows when estimating temperature fluctuations in the deep ocean, where seasonal variations are on the order of 0.1°C.

An alternative to point measurements is acoustic tomography, which is the process of inferring the state of the ocean from precise measurements of travel time of sound. The tomographic method was introduced by Munk and Wunsch in 1979 [6]. Traditional tomography over the past 40 years has made use of large acoustic sources to emit long coded signals lasting 30 seconds or more that travel across ocean basins or smaller sections of ocean. These pulses are picked up by receivers which use the measured arrival time of the signal to infer ocean states such as temperature or current

[18]. Acoustic tomography has the benefit of being able to sample and average the large-scale ocean thermal structure. The temperature of the water at each point on the acoustic ray paths between sensors is spatially integrated, effectively suppressing unwanted small scales that contaminate conventional direct measurements and lead to aliasing [6]. These rays sample a range of water depths, discussed in Appendix E. Acoustic tomography has been shown to provide more precise temperature measurements than thermometer measurements at point locations in the ocean [18].

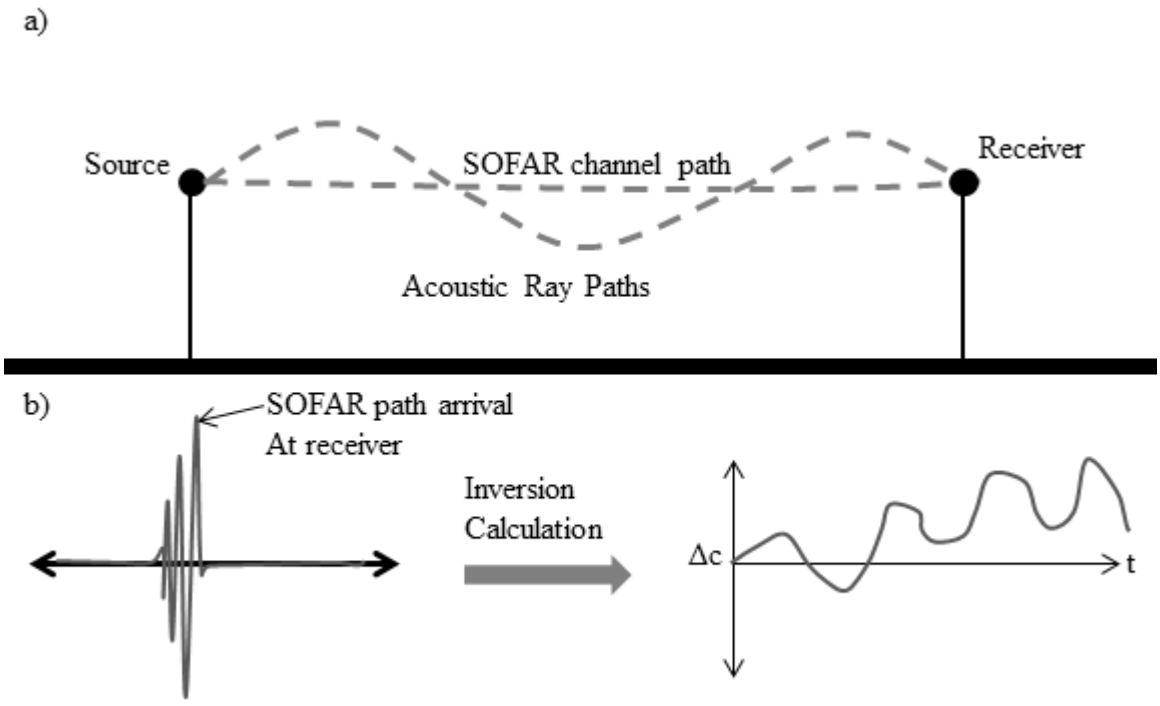

Figure 2: a) A simplified schematic of a deep water active acoustic tomography experiment, showing an acoustic source, receiver, and two typical ray paths. The SOFAR channel path is the stable feature that is useful for acoustic tomography. b) Time fluctuations of the last arrival (i.e. the SOFAR arrival) are inverted [6] to track changes in sound speed in the water along the SOFAR channel path.

One early example of acoustic tomography is the Heard Island feasibility experiment, where a signal of 221 dB (re 1μPa) was projected from Heard Island in the southern Indian Ocean to 16 receiver sites in the Indian, Atlantic, and Pacific oceans [13]. The ray paths for this experiment are shown in Figure 3. The major goal of this experiment was to determine the transmission range of acoustic signals and to see if the signals could be used to track climate change. Signals were detected and travel times were estimated up to a distance 18 Mm, but the experiment was shut down because of concerns about the effects of noise on marine mammals. A more successful tomography experiment was the Acoustic Thermometry of Ocean Climate (ATOC) experiment. The ATOC acoustic array was located in the North Pacific Ocean, and was active from 1996 until 2006. Results from this study show more precise temperature tracking than point measurements and good agreement with satellite data [4]. However, the ATOC project was dogged by environmental concerns as well, ultimately leading to its shut down. Although active-source acoustic tomography has been demonstrated to be the most reliable way to track small changes in temperature averaged over large distances, concerns over the impact of noise on marine animals have largely stopped progress in this field of study.

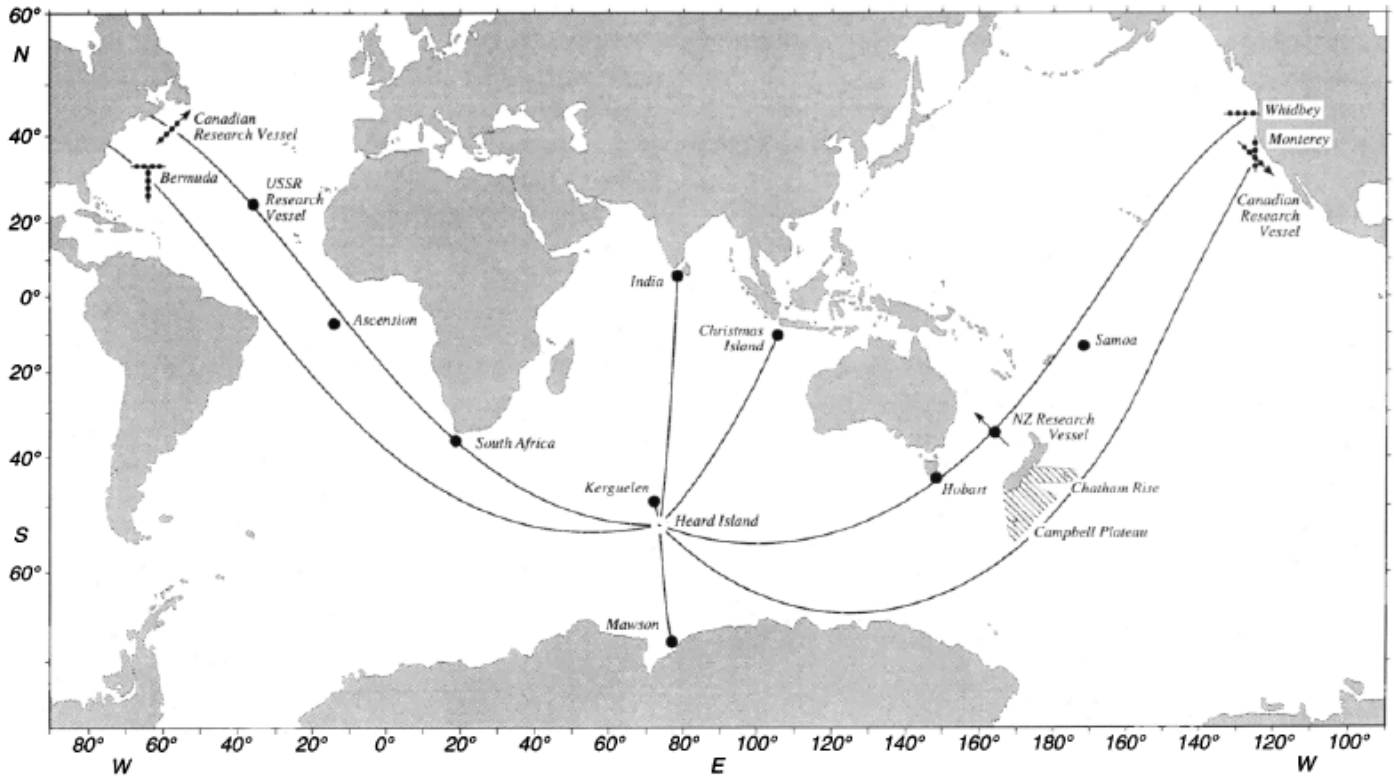

Figure 3: Ray paths for the Heard Island test [13].

1.3 International Monitoring Stations Background

The hydroacoustic data used in this dissertation are publically available from the Comprehensive Nuclear-Test-Ban Treaty Organization (CTBTO). The CTBTO was formed to enforce the Comprehensive Nuclear-Test-Ban Treaty, a treaty that outlaws nuclear test explosions. The CTBTO established the International Monitoring System (IMS), a worldwide network of sensors that are used to detect explosions. The IMS includes 5 hydrophone stations spread throughout the world's oceans that are used in this study. Each IMS hydroacoustic station is composed of one or two triangular-shaped horizontal array of three hydrophones (referred to hereafter as triad). The locations of the hydrophones for each station are listed in Appendix D. The sides of each triad are approximately 2 km long and the hydrophones are located within the ocean deep sound channel (or SOFAR channel) at each site (see Figure 5). At stations with two triangular arrays, the two arrays are typically ~130 km apart. For the purpose of consistency, this dissertation uses the CTBTO naming convention for each monitoring station (e.g. H01, H10N, etc). Each station is located near an island or landmass for data collection purposes (see Figure 5). H01 is a single triad located near Cape Leeuwin in Australia, H08 consists of a north and south triad near Diego Garcia Island, H10 consists of a north and south triad near Ascension Island, and H11 consists of a north and south triad near Wake Island. There is a direct line-of-sight sound path between north and south triads at H11 and H10, but not at H08 where Diego Garcia Island lies between the north and south triads. H03 has both a north and south triad near Juan Fernandez Island, but only the north triad was operational for the year 2009. All hydrophones are physically cabled to shore to transmit data for archival purposes to CTBTO

headquarters in Vienna in near-real-time via satellite link. The layout for a single triad is shown in Figure 4.

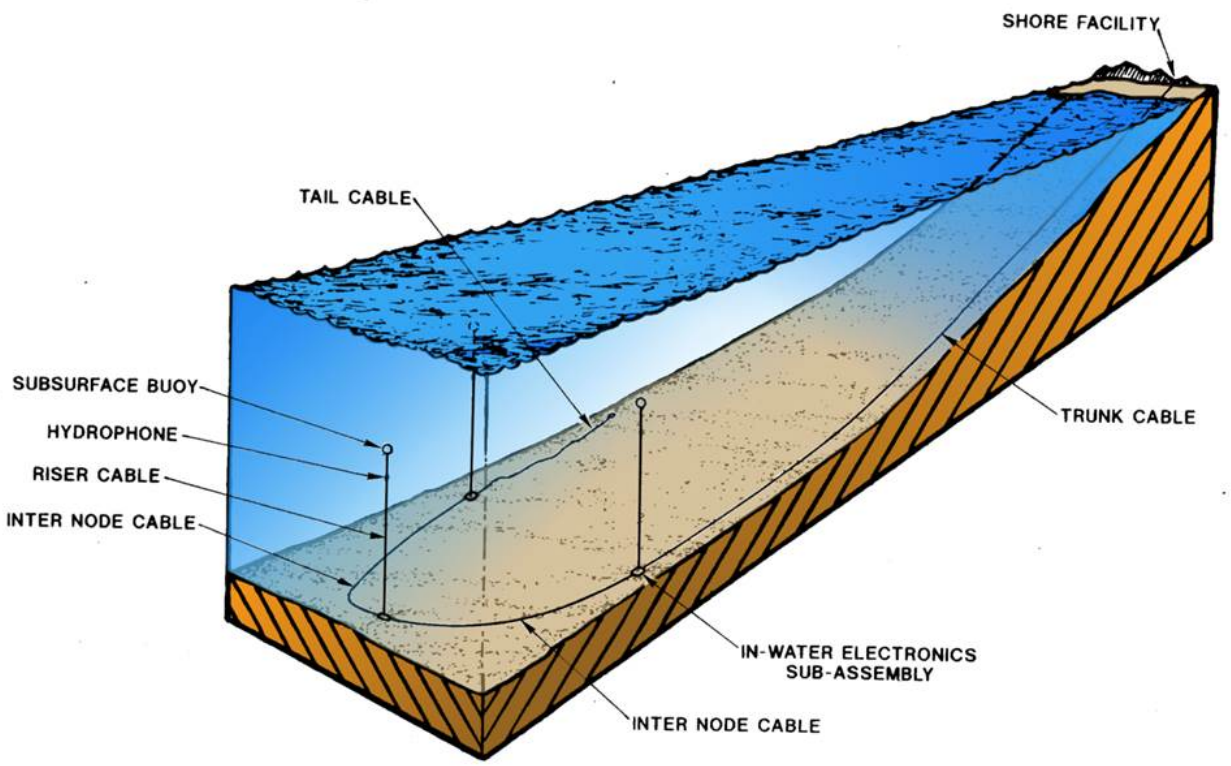

Figure 4: Schematic for a single IMS hydrophone triad (image provided by L-3 Maripro Company).

The hydrophone stations constantly sample sound pressure data at a rate of 250 Hz. The hydrophones themselves are positioned on the Sound Fixing and Ranging (SOFAR) channel: a water depth that corresponds to the minimum sound speed in the water column. Because the sound speed profile is at a minimum at the SOFAR depth, sound is refracted around this depth. The SOFAR channel acts like a waveguide, enabling underwater sound to travel great distances with little attenuation. Because of the properties of the SOFAR channel, the IMS hydrophone stations can detect acoustic perturbations that originate thousands of kilometers from the station.

The positioning of the hydrophones also allows the CTBTO to triangulate the location of the sound source.

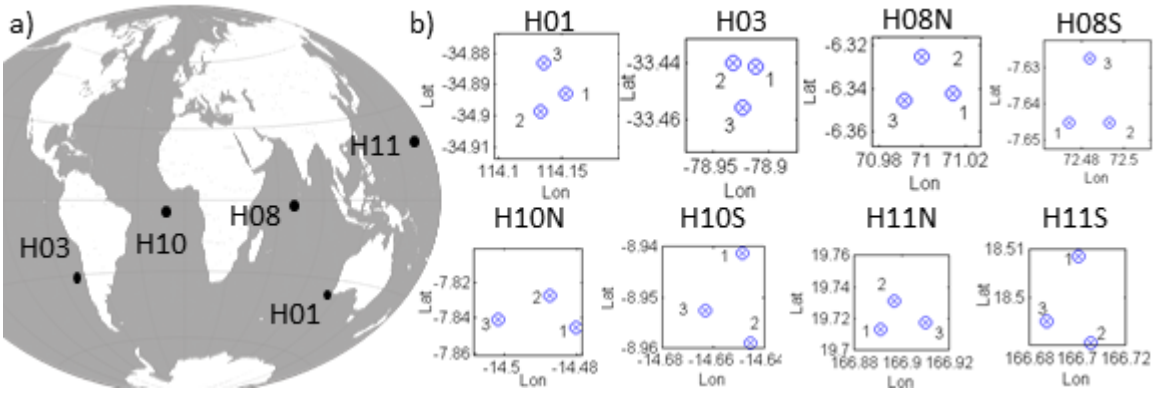

Figure 5: a) The locations of all the IMS hydroacoustic stations used in this study. b) A zoomed-in view of the hydrophone triads at each station. H08, H10, and H11 stations each have a north and south triad, and H01 and H03 stations only consist of one triad each.

1.4 Contributions to the Literature

This work makes the following contributions to the nascent fields of passive acoustic tomography and passive monitoring using ambient noise correlations:

1. This is the first deep ocean passive acoustic thermometry study. Previous work has been limited to shallow water, where the acoustic propagation physics and frequency range are very different. Acoustic tomography is traditionally conducted in the deep ocean to make use of the SOFAR waveguide for long propagation paths. Extending passive monitoring to the deep ocean will enable sensing of the ocean over longer ranges, similar to traditional acoustic tomography.

2. Previous studies in shallow water have collected coherent arrivals over time periods of up to a month. This work examines up to 9 continuous years of data, showing that the ocean environment contains enough stable sound sources to monitor passively over long time periods.

3. Previous work in passive acoustic thermometry has stopped at extracting the coherent wavefronts from ambient noise correlations to estimate Green's function arrivals. This work completes the inverse problem, converting from acoustic travel times to average temperature shifts along the acoustic travel path. The resultant temperature shifts are compared with independent direct temperature measurements.

4. Although there have been many studies on noise coherence in the ocean, this research presents the first comparison of the noise spatial coherence of multiple sites around the world, with an emphasis on the effects of noise coherence on passive acoustic tomography.

5. Previous studies have discussed one of the major limitations of using noise correlations for passive monitoring: that it is difficult or impossible to monitor events that occur on a time scale that is smaller than the emergence time of the coherent components of the ambient noise. This work presents an optimization method to overcome this limitation,

enabling the potential to monitor events that occur on a shorter time scale than the averaging time.

1.5 Outline of the Thesis

The organization of the thesis is as follows. Chapter 2 describes the signal processing techniques that are used to obtain the components of the ambient noise field that are useful for passive monitoring. Chapter 3 presents the deep ocean passive acoustic thermometry analysis and results, along with insights gleaned from multiple propagation models and temperature measurements. Chapter 4 shows the low-frequency (1-40 Hz) ambient noise horizontal directionality and coherence measured at each IMS hydroacoustic station (for stations functioning in the years 2009 and 2010) around the world. This chapter also shows how results from hydrophones separated by short ranges can be potentially used to extrapolate how well the passive monitoring technique would work at larger scales (such as an ocean basin). Chapter 5 shows how a genetic algorithm can be used to optimize a matched filter output, enabling the tracking of acoustic medium fluctuations that occur on shorter time scales than the averaging time. This reduces the impact of one of the major drawbacks of the noise correlation technique.

Chapter 2: Signal Processing Techniques

2.1 Extracting Green's function estimates from Cross-Correlations of Ambient Noise

Research on obtaining coherent wave fronts from random noise was first done in a paper by Weaver and Lobkis in 2001 [19]. The authors showed that the cross correlation of two sensors recording a diffuse field would yield an estimate of the Green's function (transfer function) between the sensors. Weaver and Lobkis foresaw that this method could be applied to other fields of study beyond ultrasonics. Campillo and Paul were the first to apply this method to seismology in 2003 [20]. This concept was also applied to underwater acoustics by Roux *et al*. in 2004 [11]. The field of passive imaging has grown tremendously over the past ten years, and the fundamental principle behind this field is the cross-correlation.

The cross-correlation is a statistical measure of the coherence, or similarity, between two signals [21]. The energy-normalized cross-correlation $C_{12}(t)$ of two signals $S_1(t)$ and $S_2(t)$ is given by the following formula.

$$C_{1,2}(t) = \frac{\int_{-\infty}^{\infty} S_1(\tau) S_2(\tau+t) d\tau}{\sqrt{\int_{-\infty}^{\infty} (S_1(\tau))^2 d\tau} \sqrt{\int_{-\infty}^{\infty} (S_2(\tau))^2 d\tau}} \quad (1)$$

If the two signals are similar at a given time shift, *t*, then the normalized cross-correlation will return a value very close to 1 for that time shift. If the signals are very different for a given time shift, then the cross-correlation will return a value close to zero.

In 2005, Roux *et al.* showed that the derivative of the average cross-correlation function for signals recorded at two sensors would result in the Green's function in free space from sensor 1 to sensor 2, and from sensor 2 to sensor 1 [22]. Roux's result is shown in Equation 2.

$$\frac{d}{dt}\langle C_{1,2}(t)\rangle \propto G(r_2, 0; r_1, -t) - G(r_1, t; r_2, 0) \quad (2)$$

In this equation, $G(r_2, 0; r_1, -t)$ is the time reversed Green's function of the second sensor to the first, $G(r_1, t; r_2, 0)$ is the Green's function from sensor 1 to sensor 2, and the brackets denote the ensemble average of the cross-correlation. Additionally, r_1 and r_2 are the positions of the sensors. If the acoustic power and location of each ambient noise source are known, then Equation 2 can be modified for an exact (rather than proportional) equality. It should be noted that this formula for the Green's function assumes an infinite medium (or free space) with no attenuation.

To illustrate this process, an example is shown in the figure below. Part A shows the process of active sensing, where a pulse is sent from sensor 1 and received by sensor 2. The time delay depends on the sound speed of the medium, and the signal travels in a direct line from sensor 1 to sensor 2. Part B illustrates how both sensor 1 and sensor 2 record the ambient noise. The cross-correlation serves to filter out wave arrivals that are not common to sensor 1 and sensor 2. The acoustic waves that are common to sensors 1 and 2 are detected by the cross-correlation.

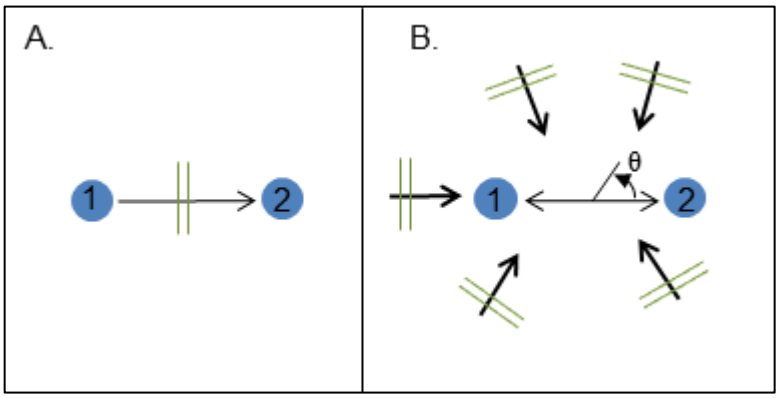

Figure 6: Comparison of active and passive sensing. (A) In active sensing, a signal is sent by Sensor 1 and travels through the medium to Sensor 2, which receives a time-delayed replica of the original signal. (B) In passive sensing, Sensor 1 and Sensor 2 both record ambient noise, ideally consisting of wave fronts propagating from all directions. Performing a time average on the cross-correlation of the noise recorded at both sensors yields the coherent components of that noise (i.e. the noise that is common to both sensors). In a medium with a constant sound speed, the coherent waves travel in a straight line between sensors.

We will consider a scenario involving a medium with a constant sound speed of $c=1500$ *m/s*, a sensor separation of $x=300$ *m*, and plane waves impinging on the sensors from all directions. For any single plane wave, the signals recorded at each sensor are identical except for a time delay of $t_d = x*cos(\theta)/c$ (θ is defined in Figure 6). When the signals are cross-correlated, the signals will align at t_d and the cross-correlation will have a value of 1. Depending on the incoming angle of the wave, the time delays can have a maximum value of $\pm x/c$ if the wave is on-axis with the receivers ($\theta=0$ or $\theta=180$) or a minimum value of 0 if the wave is perpendicular to the receiver axis ($\theta=90$ or $\theta=270$). If plane waves impinge equally from all angles, the cross-correlation of the recorded signals would result in a square pulse. The pulse starts at $-x/c$ and ends at $+x/c$. All arrivals between these extremes are resultant from plane waves arriving off of the receiver axis. The derivative of this cross-correlation results in two dirac-delta functions at

$t=-x/c$ and $t=x/c$, the step discontinuities. These arrivals are equivalent to the Green's function in free space from sensor 1 to sensor 2 and vice versa. The results are shown in Figure 7.

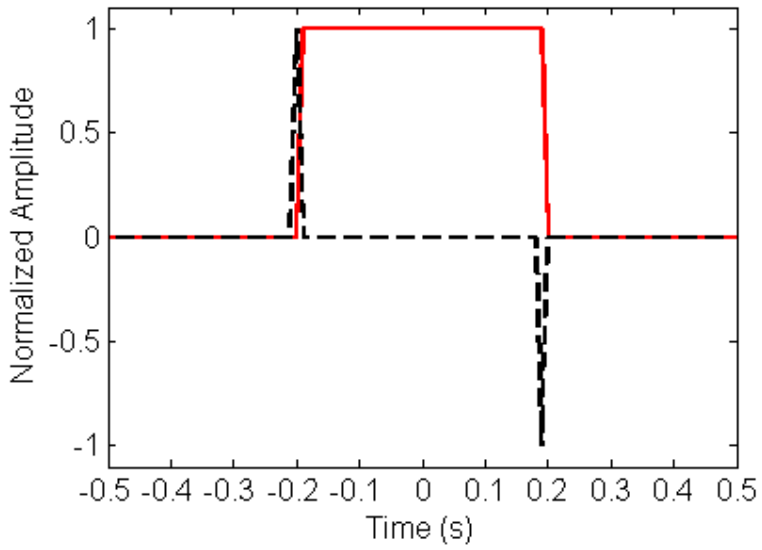

Figure 7: Cross-correlation (red) and time derivative of the cross-correlation (black dashed) between two sensors with infinite bandwidth. The time derivative is directly proportional to the Green's function between the two sensors.

In experiments, it is unrealistic to have infinite bandwidth. All real signals are band-limited, and therefore will not have the zero-frequency component that is present in the square wave. The cross-correlation of the finite bandwidth case actually gives a good estimate of the Green's function without taking the derivative. Using the same setup as before, the incoming plane waves are now band limited (20-30 Hz). The resulting cross-correlation waveform along with its derivative is shown in Figure 8. Note that the two waveforms are very similar. Roux *et al.* showed that for the finite bandwidth case, the difference between the two waveforms is a $\pi/2$ phase shift, which means that a band-limited cross-correlation is a decent approximation of the

Green's function. In fact, the cross-correlation may be the preferred method for approximating the Green's function because a time derivative can introduce additional noise.

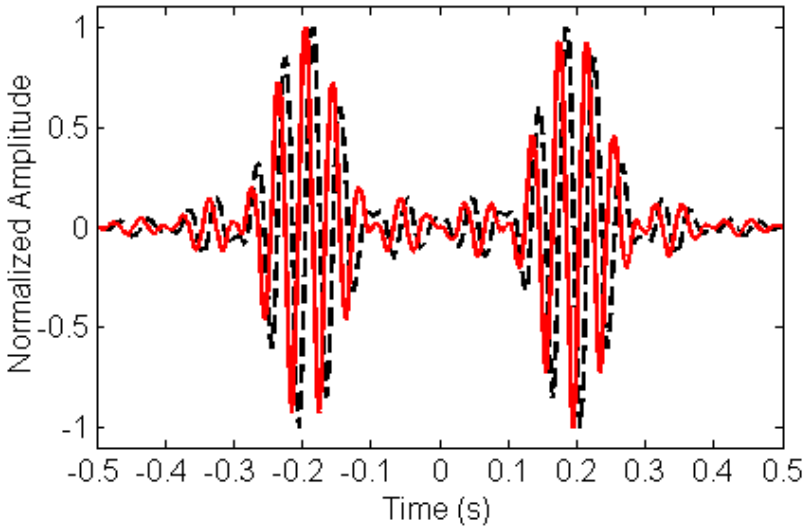

Figure 8: Cross-correlation (red) and time derivative of the cross-correlation (black dashed) for two sensors with a finite bandwidth. Both waveforms can be used as an approximation of the Green's function.

There are three limitations to the cross-correlations of ambient noise technique. These limitations can make it difficult or impossible to find a Green's function in certain cases. The limitations are as follows.

1. **A Large Dominant Source**: When there are one or more loud sources that are above the ambient noise level, the loud sources can drastically interfere with the measured arrivals [11]. If the loud sources happen to be in line with the receiver axis, the sources can drastically improve the emergence of the Green's function. However, if the sources are not in the end fire direction, the loud sources will "drown out" the coherent components of the ambient noise.

2. **Low SNR:** In this case, "signal" in SNR corresponds to coherent waves that are common to both sensors and travel in a straight line between them, and "noise" refers to any incoherent noise. This incoherent noise may consist of a combination of electrical noise and waves that do not pass through both sensors. Only a small fraction of the overall ambient noise ends up being coherent. The longer the averaging time the more coherent noise can add up to obtain a better estimate of the Green's function. If the measurements are dominated by electrical noise (which would be unique to each sensor and therefore lacking in coherence), then an extremely long averaging time is necessary to obtain the Green's function. It would also be necessary for the measurement environment to remain stable over the averaging time duration.
3. **Limited Bandwidth:** While the advantage in limiting the bandwidth of the signal is that the cross-correlation (rather than its derivative) is a good enough estimate of the Green's function, there are drawbacks to narrowband signals. The main disadvantage in having a narrow noise spectrum is that the time-domain signal will be spread out compared to a signal with a wider frequency band. For imaging and monitoring purposes, it is useful to have a broader band signal to ensure a clear arrival in time. If the frequency content of the recorded signals is not evenly weighted across the limited band, the bandwidth will be limited even more.

The effects of bandwidth, loud interferers, and low SNR must be mitigated to obtain an accurate Green's function estimate. The next section on amplitude clipping and frequency whitening will discuss the signal processing techniques that are used to reduce the effects of these sources of error.

2.2 Preprocessing of the ambient noise data: Amplitude Clipping and Frequency Whitening

Amplitude clipping and frequency whitening are nonlinear signal processing techniques that are used to mitigate the undesirable effects caused by weighted frequency spectrum and a large dominant noise source. These techniques do not have well-set parameters, and each data set requires a trial-and-error process to get the best Green's function estimate.

Amplitude clipping is used to decrease the influence of signal outliers in the time domain. Amplitude clipping is done by clipping peaks from the time signal that are larger than three standard deviations of the noise. This threshold can be changed depending on the environment and the noise.

Frequency whitening is similar to amplitude clipping, but instead of removing the influence of peaks in the time domain, frequency whitening removes peaks in the frequency domain. Frequency whitening sets all the amplitude information in the frequency domain to unity, while keeping the phase information intact. This process enhances the effects of the weaker bands that may also have useful information to contribute to the cross-correlation. The drawback to frequency whitening is that it might amplify undesired noise, such as self-noise, making it more difficult to extract coherent waves between the sensors. The best method to find a frequency band for passive monitoring is to select the band that delivers the highest SNR. It is important to note that both frequency whitening and amplitude clipping are done before calculating the cross-correlation.

2.3 Array Beamforming

A single sensor or sensor pair is limited in the information it can provide about the propagating wave. To determine the direction of the wave source, it is necessary to use an array of sensors. The array acts as a spatial filter, attenuating all signals save those propagating from certain directions. This general process is called beamforming. Beamforming is used to focus the array's signal-capturing abilities in a particular direction. When discussing beamforming, this paper will follow the formulation presented in Johnson and Dudgeon [23].

Delay-and-sum beamforming is the oldest and simplest array signal processing algorithm, and remains a powerful approach today. The underlying idea is very simple: If a propagating signal is present in an array, the sensor outputs, delayed by appropriate amounts and added together, reinforce the signal with respect to noise. The delays that reinforce the signal are directly related to the length of time it takes for the signal to propagate between sensors. Therefore, the spatial orientation of the sensors as well as the sound speed of the medium must both be known.

To be specific, consider *s(t)* as a signal emanating from a source located at the point $\vec{x}_0$. Several sources may be present, and their emanations sum to constitute the wave field $f(\vec{x}, t)$ measured by the sensors. Consider an array of M sensors located at $\{\vec{x}_m\}$, $m = 0,\ldots,M-1$. The origin of the coordinate system is selected to coincide with the center of the array.

$$\sum_{m=0}^{M-1} \vec{x}_m = \vec{0} \tag{3}$$

The coordinate axes and array orientation are shown in Figure 9. The waveform measured by the *m*th sensor is $y_m(t) = f(\vec{x}_m, t)$. The sensor samples the wave field spatially at the sensor's location as well as temporally. The delay-and-sum beamformer consists of applying a delay Δ_m and an amplitude weight w_m to the output of each sensor, then summing the resulting signals. The beamformer output signal is defined as:

$$B(t) = \sum_{m=0}^{M} w_m y_m(t - \Delta_m) \tag{4}$$

If all the signals y_m are merely delayed replicas of each other, then the SNR of the beamformer output *B(t)* will increase by a factor of *M* over y_m.

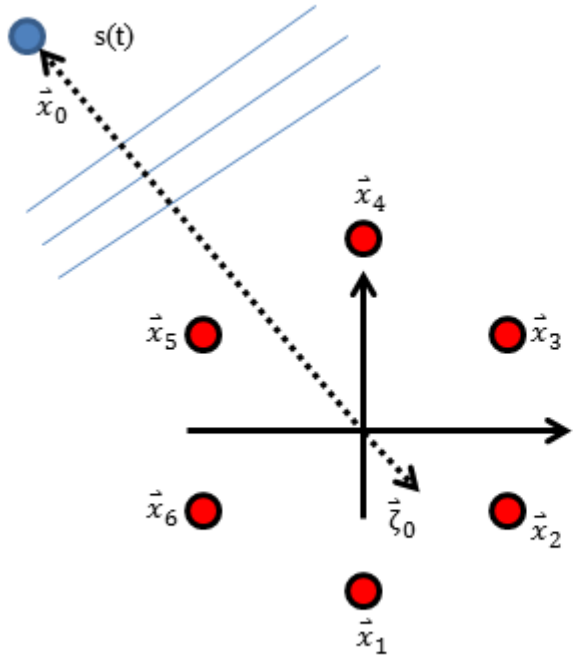

Figure 9: A source (blue) emits a signal that is detected by the array (red). The direction of propagation is the unit vector $\vec{\zeta}_0$.

The delay Δ_m is determined by the position of the sensors, speed of the wave in the medium, and either the search angle or the angle of propagation.

$$\Delta_m = \frac{-\vec{\zeta}_0 \cdot \vec{x}_m}{c} \tag{5}$$

Beamforming can also be done in the frequency domain. The frequency-domain beamformer output is given as:

$$B(e,f) = e(f)^H AY(f)Y(f)^H A^H e(f) \tag{6}$$

where e(f) is the steering vector given by:

$$e(f) = \begin{bmatrix} \exp\{-j\vec{k}_0 \cdot \vec{x}_0\} \\ \vdots \\ \exp\{-j\vec{k}_{M-1} \cdot \vec{x}_{M-1}\} \end{bmatrix} \quad (7)$$

A is an optional *MxM* diagonal matrix knows as the weight matrix, *Y(f)* is the Fourier transform matrix of the signals, and $\vec{k}_m$is the wave number vector equivalent to $2\pi f \cdot \vec{\zeta}_0/c$. The superscript *H* indicates the Hermitian (conjugate transpose operation) of the matrix. It is common practice to combine the steering vector and weight matrix into a single weighted steering vector, *W*.

In common practice, each signal detected will have noise and other parameters that will decrease the cross-coherence of the signals, meaning that the signals are not just shifted replicas but have other changes in addition to the time shift. In these noisy environments, it is necessary to use adaptive beamforming to increase the performance of the beamformer. However, simple delay-and-sum beamforming is included in this dissertation because it provides the fundamental understanding of the mechanics of beamforming that also are applicable to the adaptive beamforming algorithms used in this research.

2.4 Adaptive Split-Beam Beamforming

Split-beam beamforming coherently combines cross-correlations between signals collected at arrays or array elements rather than beamforming the individual signals [24], [25]. Split-beam beamforming (also known as product array processing) is known to have a higher spatial resolution than the simpler beamforming methods discussed in the previous section [25]. In this case, the beamformer output $B(f,T_r)$ for a given frequency f and recording duration T_r is given by [26]:

$$B(f,T_r) = W_1^H(f)\hat{C}(f,T_r)W_2(f) \tag{8}$$

In this case, $\hat{C}(f,T_r)$ is the frequency-domain cross-correlation matrix with components $\widehat{C_{i,j}}$ comprising of the cross-correlation of signals $Y_i(f)$ and $Y_j(f)$ collected at sensors i and j. The corresponding time-domain beamformer output, $B(t,T_r)$, is found by taking the inverse fast fourier transform of $B(f,T_r)$. A pictorial representation of the beamforming equation above is shown in Figure 10 for the case with three sensors (M=3).

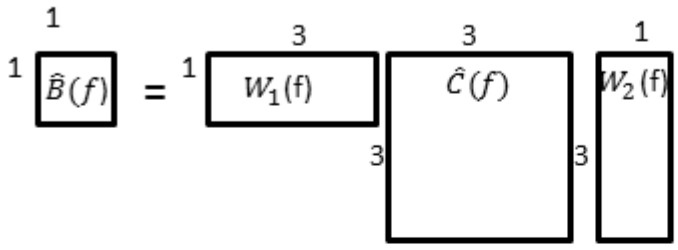

Figure 10: Pictorial representation of the beamforming equation, with matrix sizes represented.

Adaptive beamforming involves implementing a beamforming algorithm where the steering vectors (i.e. weight vectors) W_1 and W_2 are derived from a reference cross-correlation

matrix, *R(t),* rather than derived based on a plane-wave propagation assumption. *R(t)* is based on *a priori* knowledge of the average structure of the coherent waveforms between the receiver arrays. In previous studies [8] [26], *R(t)* was derived from the expected value of the cross-correlation matrix averaged over a long period of time. The weight vectors $W_1(f)$ and $W_2(f)$ are the left and right SVD singular vectors of *R(f)*, where *R(f)* is the fast fourier transform of *R(t):*

$$R(f) = U(f)\Sigma(f)V(f)^H \tag{9}$$

$\Sigma(f)$ is a diagonal matrix whose non-zero entries are the singular values $\sigma_i(f)$ of *R(f)*. By definition, the first principal component matrix, $\sigma_1 U_1(f) V_1(f)^H$, defines the largest portion of the variance of *R(f)*. Consequently, the empirical weight vectors are $W_1(f)=U_1(f)$ and $W_2=V_1(f)$.

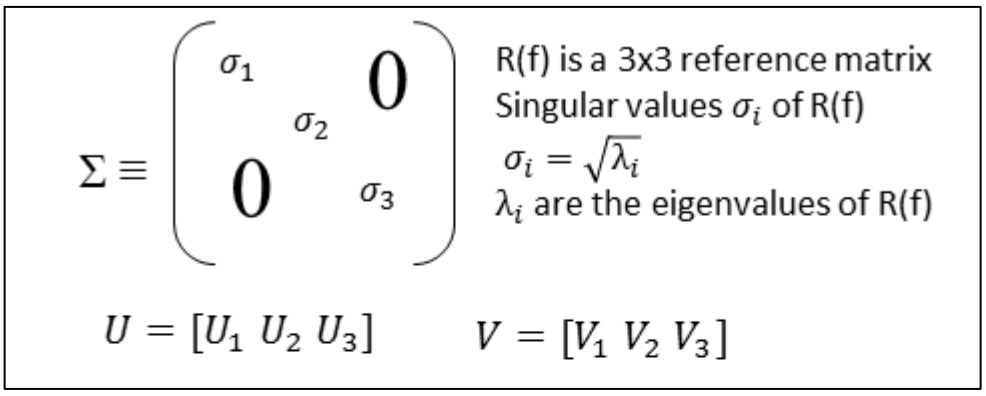

Figure 11: Pictorial representation of the singular value decomposition of the reference matrix, R(f).

Chapter 3: Monitoring Deep Ocean Temperatures using Acoustic Ambient Noise

3.1 Introduction

This chapter applies the passive monitoring method to study the coherence of ambient low-frequency noise over long ranges (~130 km) and long time periods (5-9 years) [27]. The goal of this chapter is to track arrival time fluctuations of coherent waves, which correspond to changes in deep water temperature. Spatial coherence is examined at two stations: Wake (H11) and Ascension (H10), as shown in Figure 12 . These stations are selected because data are available over long time periods and because these stations each consist of two triads, making it possible to examine propagation between the triads. Data at the Ascension Island station are available continuously from 2006 – 2014. While data at the Wake Island station are available from 2008-2014, the 2008 and 2009 data had many days missing from the data set. Therefore, Wake Island data is only analyzed from 2010-2014.

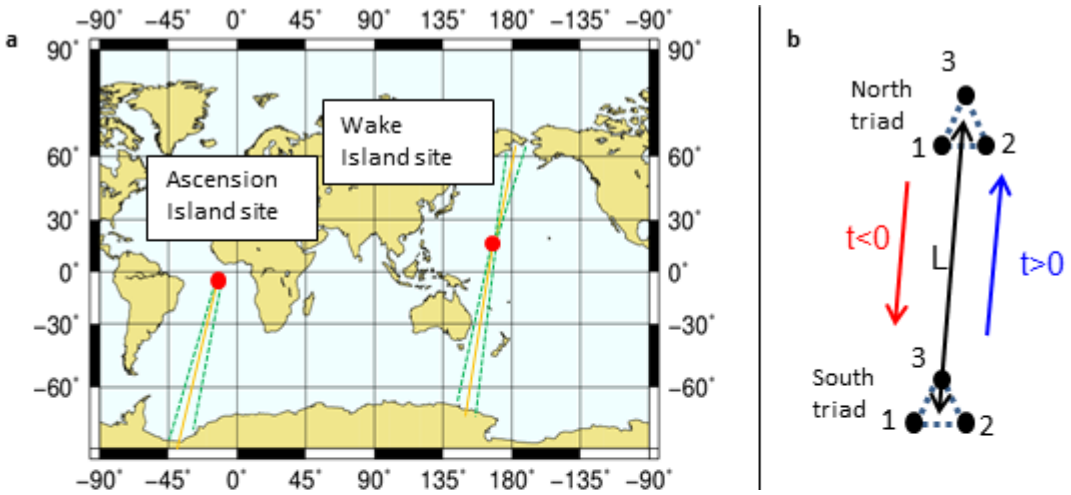

Figure 12: (a) Locations of the two hydroacoustic stations (red dots) near Ascension and Wake Islands. (b) Zoomed-in schematic of the hydrophone array configurations for the Ascension and Wake Island sites. Each hydroacoustic station consists of a northern and southern triangle array of three hydrophones (or triad), with each triangle side having a length ~ 2 km. The distance L between triad centers is equal to 126 km and 132 km for the Ascension Island and Wake Island hydroacoustic stations, respectively.

3.2 Cross Correlations

To establish the feasibility of passive monitoring over long ranges, data from two monitoring stations were used: H10 (Ascension Island) and H11 (Wake Island). Both of these stations have two hydrophone triads, allowing cross-correlations to be computed between the triads. Additionally, both stations had good-quality data that were available over time periods of greater than two years, which was necessary to track seasonal coherent arrival time fluctuations.

The continuous data collected at the hydroacoustic stations were segmented into 1 day intervals. Following the procedure outlined by Sabra *et al*. [12], these segmented data were then amplitude-clipped and frequency-whitened in the most energetic frequency band (1 – 40 Hz) in order to reduce the influence of episodic high amplitude transient events while preserving the

overall phase information of these time series. For each of the two hydroacoustic stations, cross-correlations were then computed between each pairwise combination of the three hydrophones of their North and South triads, thus resulting in a total of 9 cross-correlation waveforms per hydroacoustic station. Specifically, given each 1-day long frequency-whitened and amplitude-clipped time series $Y_i^{(S)}(t,k)$ and $Y_j^{(N)}(t,k)$ recorded, respectively, by the *i*th and *j*th hydrophone of the south and north triad (i,j=1,2,3) during the *k*th day of the recording cycle, the energy-normalized cross-correlation function $C_{i,j}(t,k)$ was computed using

$$C_{i,j}(t,k) = \int Y_i^{(S)}(\tau,k) Y_j^{(N)}(\tau+t,k) d\tau \Big/ \sqrt{\int \left(Y_i^{(S)}(\tau,k)\right)^2 d\tau} \sqrt{\int \left(Y_j^{(N)}(\tau,k)\right)^2 d\tau}. \quad (10)$$

$C_{i,j}(t,k)$ was formatted as a cross-correlation time series averaged over all hours for a single day, *k*. Based on the time-lag convention set in Equation 8, coherent arrivals occurring in the waveform $C_{i,j}(t,k)$ at positive time-delay *t* are generated by coherent noise traveling from the south triad to the north triad , and conversely arrivals occurring at negative time-delay correspond to coherent noise traveling in the opposite direction (i.e. from the north triad to the south triad).

Applying a running average to the cross-correlation (i.e. smoothing the correlation) over the days *k* enables the coherent Green's function estimate to emerge. The smoothed cross-correlation $\hat{C}_{i,j}(t,k;N)$ is

$$\hat{C}_{i,j}(t,k;N) = \frac{1}{N} \sum_{k-N/2}^{k+N/2} C_{i,j}(t,k) \quad (11)$$

where *N* is the window length (in days) of the running average.

$\hat{C}_{i,j}(t,k;N)$ ensemble averaged over $k=1$ to $k=30$ days over the month of January for the Ascension Island and Wake Island locations are shown in Figure 13 (a) and (b), respectively. The same waveforms are also shown averaged over the month of August in Figure 13 (c) for Ascension Island and (d) for Wake Island. While the SNR of the arrival fluctuates over the different years and seasons, the coherent arrivals are stable and deterministic throughout the entire data set. The SNR for each smoothed cross-correlation waveform at each day k is defined in Equation 12 [8].

$$SNR_{i,j}(k) = \frac{\max(\hat{C}_{i,j}(t,k;N))}{3std(\hat{C}_{i,j}([t_{n1}\ t_{n2}],k;N))} \tag{12}$$

The denominator in Equation 12 represents three times the standard deviation of the running-averaged correlation $\hat{C}_{i,j}(t,k;N)$ taken during the lag time $[t_{n1}\ t_{n2}]$ which is considered to be outside of the range of possible physical water-borne arrivals.

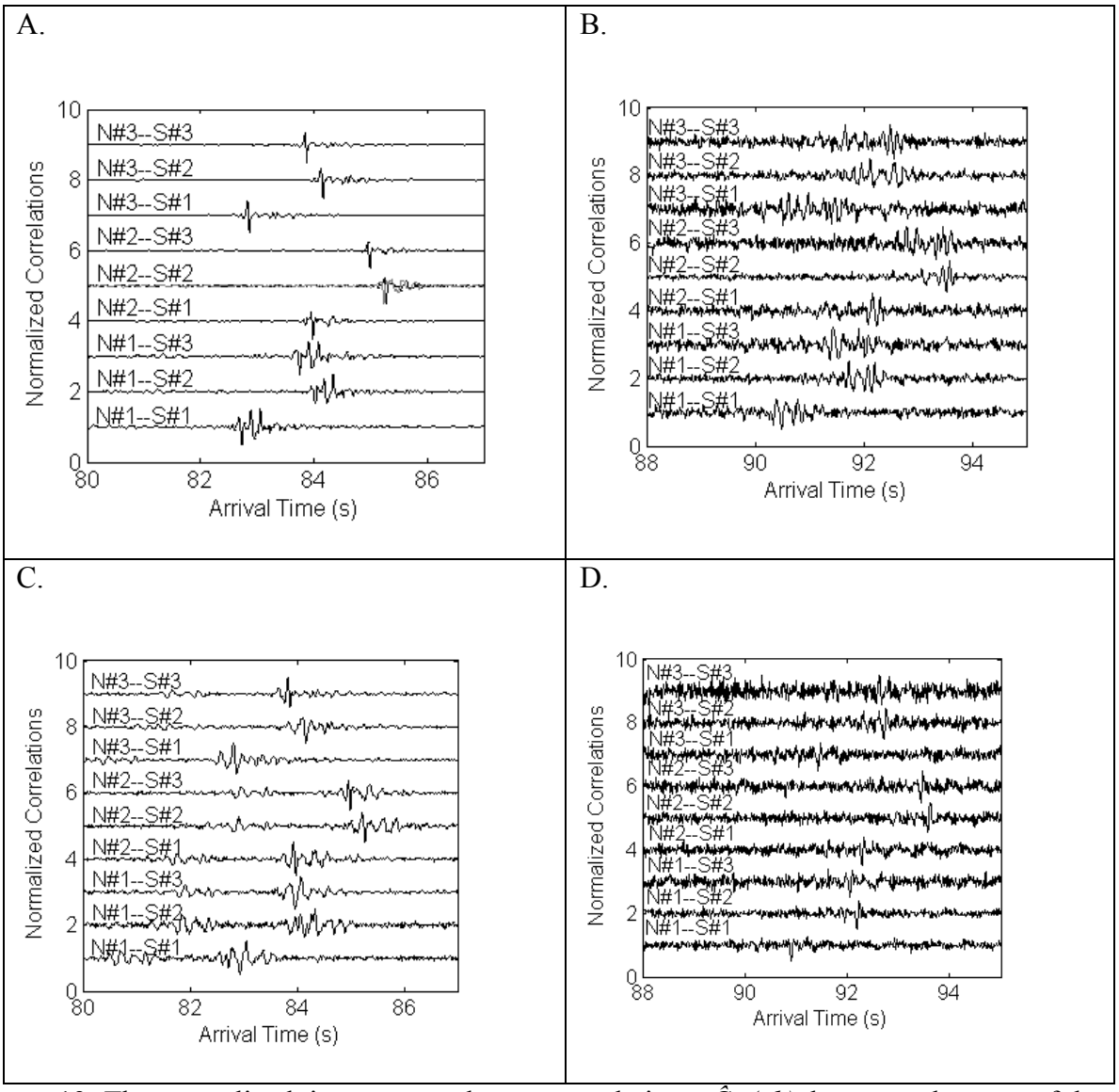

Figure 13: The normalized time-averaged cross-correlations, $\hat{C}_{i,j}(t,k)$, between elements of the north and south triads for the (a) Ascension Island and (b) Wake Island sites averaged over the month of January ($N=30$) in 2010. These cross-correlations are plotted for a positive time lag, which corresponds to coherent waves traveling from south to north. The $\hat{C}_{i,j}(t,k)$ are also plotted at (c) Ascension Island and (d) Wake Island for the month of August ($N=30$) for a summer/winter comparison.

Coherent wavefronts can be extracted across all the years where data exists for both sites. Once a minimum averaging time is achieved, the coherent arrivals are stable. Figure 14 illustrates the fact that increasing the averaging time past the 1 month duration does not change

the coherent waveforms except for perhaps a slight increase in SNR. This figure emphasizes the stability of the coherent arrival structure that travels from the south triad to the north triad. Figure 15 shows the stability of the cross correlations over different years, and Figure 14 shows the stability of the cross-correlations at Ascension and Wake over different seasons. The arrival structure is preserved regardless of season and year at both locations.

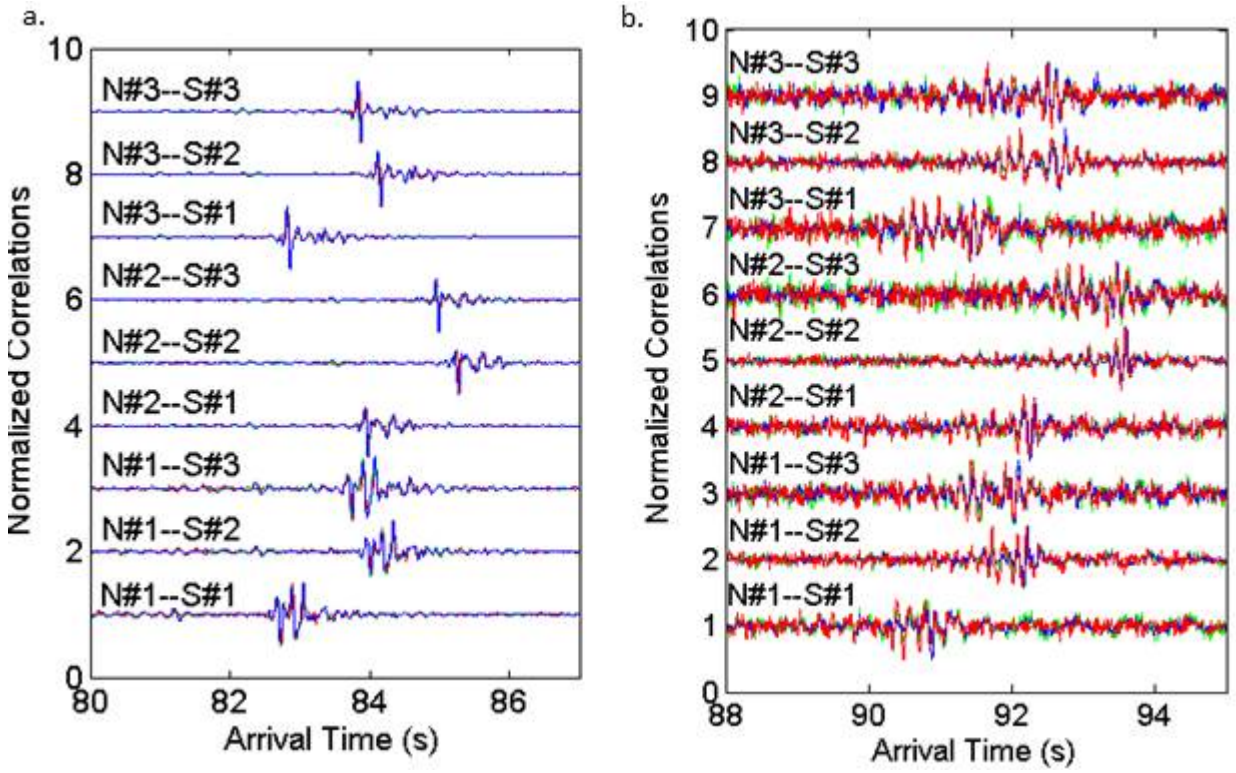

Figure 14: a) Coherent arrivals $\hat{C}_{i,j}(t,k)$ ($i=1,2,3$ $j=1,2,3$) for the Ascension Island site over different averaging times k. Red is 1 month average, green is 6 month average, and blue is 12 month average. b) Same plot as part a, but for Wake Island site.

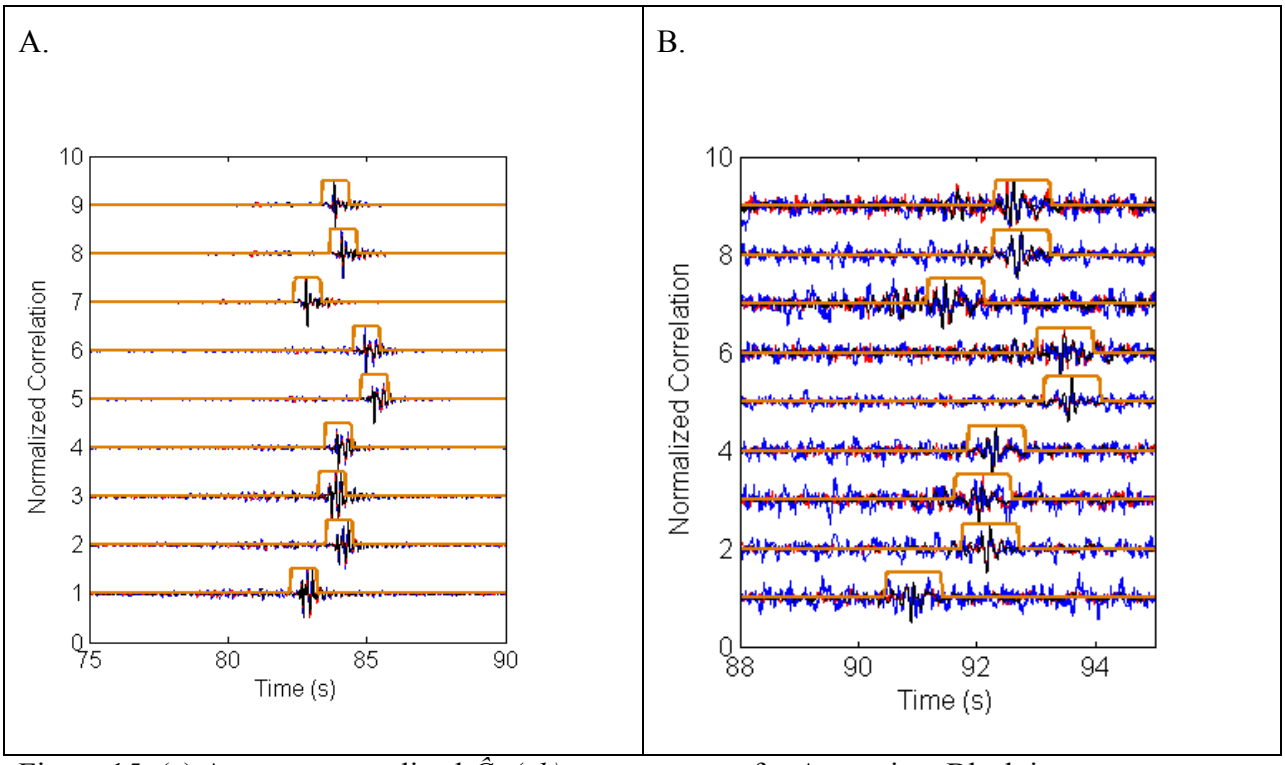

Figure 15: (a) Average normalized $\hat{C}_{i,j}(t,k)$ over one year for Ascension. Black is average over 2006, red is average over 2009, and blue is average over 2012. (b) Average normalized $\hat{C}_{i,j}(t,k)$ over one year for Wake. Black is average over 2010, red is average over 2011, and blue is average over 2012. The window used when finding $R_{i,j}(t)$ is overlaid on the correlations.

3.3 Tracking changes in acoustic arrival time

We have established that coherent acoustic waves can be extracted from background noise at two independent locations. The travel times of these acoustic waves are consistent with the distances that the waves travel between sensors. To track sound speed (and therefore temperature) fluctuations in the water between the sensors, we need to be able to track the change in travel time with a high degree of precision. Increasing the SNR of the signal enables us to track travel times with a higher degree of temporal resolution. The SNR is related to the root-

mean-square error, σ_t , of the travel time estimate as discussed in Appendix C. To have arrival errors on the order of less than a millisecond, it is necessary to have an *SNR* of at least 20 dB. For example, a center frequency of 10 Hz and *SNR* of 20 dB yields an arrival time error of 0.2 ms. A lower SNR will result in larger arrival time errors.

The 9 coherent waveforms of $\hat{C}_{i,j}(t,k)$ are beamformed using a split-beam coherent adaptive beamformer as a spatio-temporal filter to increase SNR and minimize the necessary averaging time [26], [8], [28], as discussed in Section 2.4. The steering vectors $W_1(f)$ and $W_2(f)$ are set as the first left and right singular vectors at each frequency bin f of the 3x3xN_f reference cross-correlation matrix, $R(f)$, where the $(i,j)^{th}$ element of $R(f)$ is the Fourier transform of the reference cross-correlation $R(t)$ (dimensions 3x3xN_t) over t:

$$R(f) = U(f)\sum(f)V(f)^H \tag{13}$$

$R_{i,j}(t)$ is equivalent to $\hat{C}_{i,j}(t,k;N)$ ensemble-averaged over a long time (i.e. the first year of each data set), and then windowed to remove off-axis arrivals. The 3x3 orthonormal matrices $U(f)$ and $V(f)$ are formed by concatenating column-wise the 3 singular vectors $U_i(f)$ and $V_i(f)$, respectively. $\sum(f)$ is a 3x3 diagonal matrix whose non-zero entries are the 3 singular values, $\sigma_i(f)$ (i=1,2,3), sorted in descending order. By definition, the first principal component matrix, $\sigma_1(f)U_1(f)V_1(f)^H$, comprises the largest portion of the variance of the reference cross-correlation matrix, $R(f)$. Therefore, the array weight vectors $W_1(f)$ and $W_2(f)$ are equivalent to $U_1(f)$ and $V_1(f)$, respectively.

The justification for only selecting the first principal component vectors when forming the beamformer weights is illustrated in Figure 16, which shows the normalized singular values of $R(f)$. The first singular value, $\sigma_1(f)$ is much larger than the other two singular values across most frequencies shown.

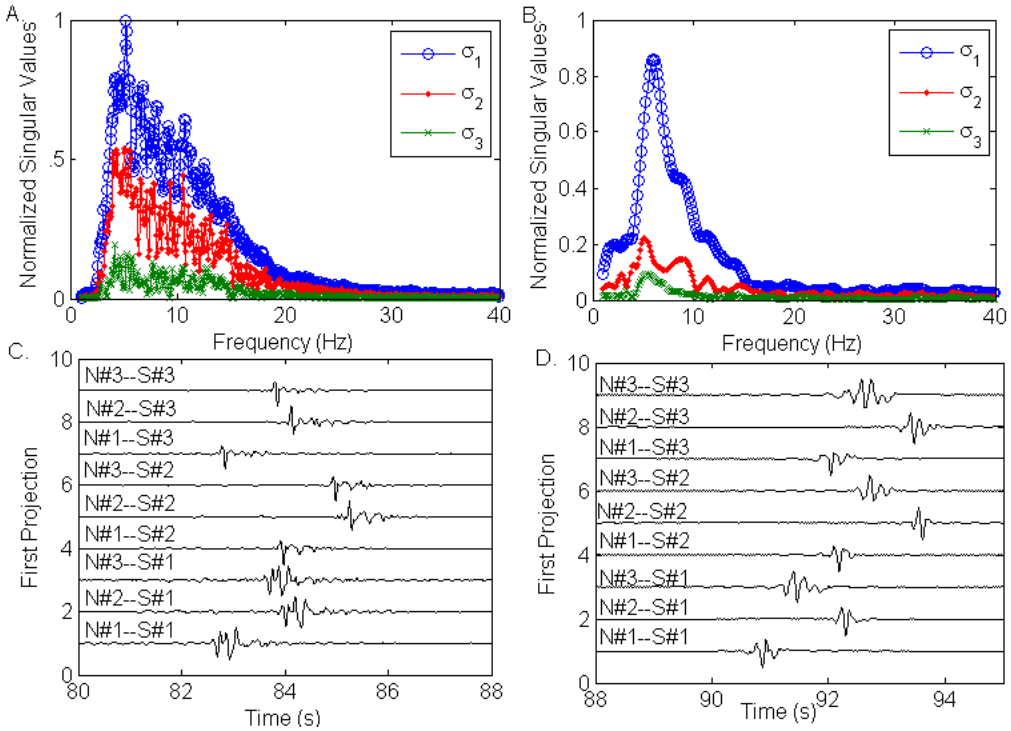

Figure 16: Normalized singular values $\sigma(f)$ of $R(f)$ from 1 to 40 Hz for (a) Ascension and (b) Wake. The inverse Fourier transform of the first projection $\sigma_1(f)U_1(f)V_1(f)^H$ is also shown for (c) Ascension and (d) Wake.

Figure 17 tracks the peaks of the time-domain beamformer output over the years, with the SNR superimposed for each day, for the Wake and Ascension sites. The SNR of the time-domain beamformer output is defined below. 30 day averaging was applied to the beamformer output, and the individual cross-correlations were not averaged before beamforming.

$$SNR_B(k) = \frac{max_t B(t,k;N)}{3std_t(B([t_{n1}\ t_{n2}],k;N))} \tag{14}$$

In this formula, $B(t,k;N)$ is the time domain beamformer output, which is equal to the inverse fast fourier transform of $B(f,k)$, smoothed over N days. The numerator refers to the peak value of the time-domain beamformer over time t at each day k . The denominator is equal to 3 times the standard deviation over time t for the time-domain beamformer in the domain $[t_{n1}\ t_{n2}]$, which is

considered to be the portion of the waveform purely comprised of noise. The time limits for the noise domain are t_{n1} = 115 seconds, t_{n2} = 120 seconds.

The Wake Island positive cross correlations were beamformed using a reference that is the average of the positive cross correlations for 2010. The Wake Island negative cross correlations were beamformed using a similar reference containing an average of the negative cross correlations. A Tukey window with a width of +-0.5 seconds was applied to the peak arrival of each reference correlation waveform to zero out some of the noise. Figure 17a shows the beamformer output and peak arrival times for the positive Wake Island cross correlations beamformed using the positive reference. Note that SNR is lower in this case than in the Ascension Island case. It is also important to note that Figure 17 shows the change in arrival time, not the absolute arrival time. Seasonal arrival time fluctuations are on the order of 40 ms. The overall downward trend in arrival times means that the sound speed is increasing, corresponding to an increase in temperature.

There is a high degree of similarity of the arrival-time variations obtained for both the positive and negative time-delay arrivals (Figure 17c) which result from different noise events propagating northward or southward, respectively, along the same paths linking the south and north triads for the Wake Island site. This symmetry of the arrival-time variations demonstrates that they are due to reciprocal changes in the environment, such as ocean sound speed fluctuations induced by temperature changes, rather than non-reciprocal changes, such as currents, clock drift, or other signal-processing artifacts [12], [8]. The symmetric arrival-time variations could not be established for the Ascension Island site due to the African continent blockage of ocean noise propagating southward.

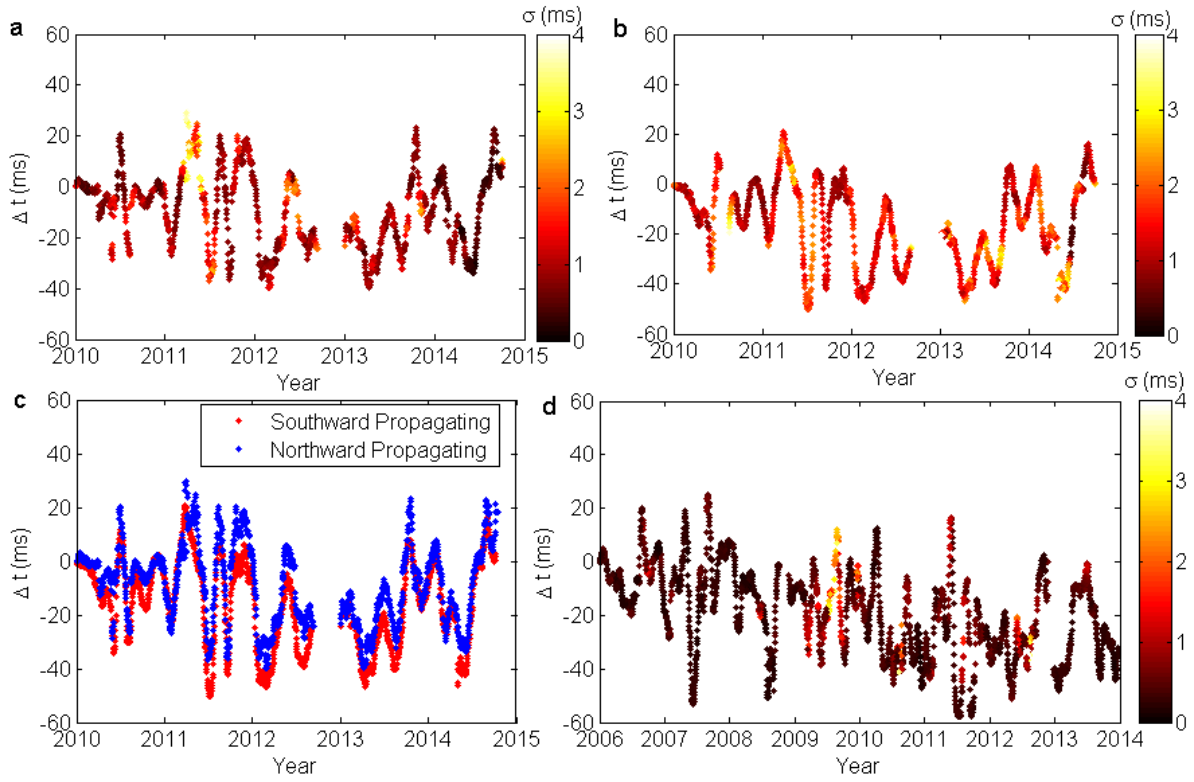

Figure 17: (a) Temporal variations over 5 years of the coherent SOFAR arrivals at the Wake Island site extracted from ambient noise correlations using a one-week moving average. These arrivals primarily result from ice-noise propagating northward towards the Wake Island site successively from the South triad to the North triad. The estimated measurement error σ (ms) associated with each data point is indicated by the colorbar. (b) Same as (a) but using instead noise propagating southward towards the Wake Island site in the reciprocal direction i.e. successively from the North triad to the South triad. (c) Comparison of the arrival time fluctuations for northward-propagating and southward-propagating noise at the Wake Island site, corresponding to the same values shown in (a) and (b) respectively. (d) Same as (a) but for noise propagating northward towards the Ascension Island site successively from the South triad to the North triad. Due to hydroacoustic data availability, a longer observation period of ~8 years was used for the Ascension Island site. Data collected during 2014 discarded at this site for quality reasons.

3.4 Parametric Study of SNR vs. Averaging Time

A parametric study is included to illustrate the seasonal dependence of arrival SNR and to emphasize the SNR gain obtained from beamforming. Figure 18 shows the SNR_B and the SNR of one representative correlation waveform $\hat{C}_{2,2}(t,k;N)$ is (averaged over 7 years for Ascension and 3 years for Wake) versus running average window length N and the day k. Comparison of Figure 18 (a) and (b) with Figure 18 (c) and (d) shows that SNR is increased by 5-9 dB by beamforming. This slightly less than the theoretical limit of an increase of $9.5\text{dB}=20\log_{10}(\sqrt{9})$ (due to 9 correlation waveforms) because of a lack of spatial coherence between sensors. There is a seasonal drop in SNR in the middle of the year at both locations because of a decreased level of coherent ambient noise that travels between the south and north triads. At Ascension, an SNR_B of 20 dB can be achieved with N=10 days. However, then the arrivals would track fluctuations that occur on the order of 10 days, which may include the effects of sensor drift [7] rather than the seasonal trends that are the focus of this study.

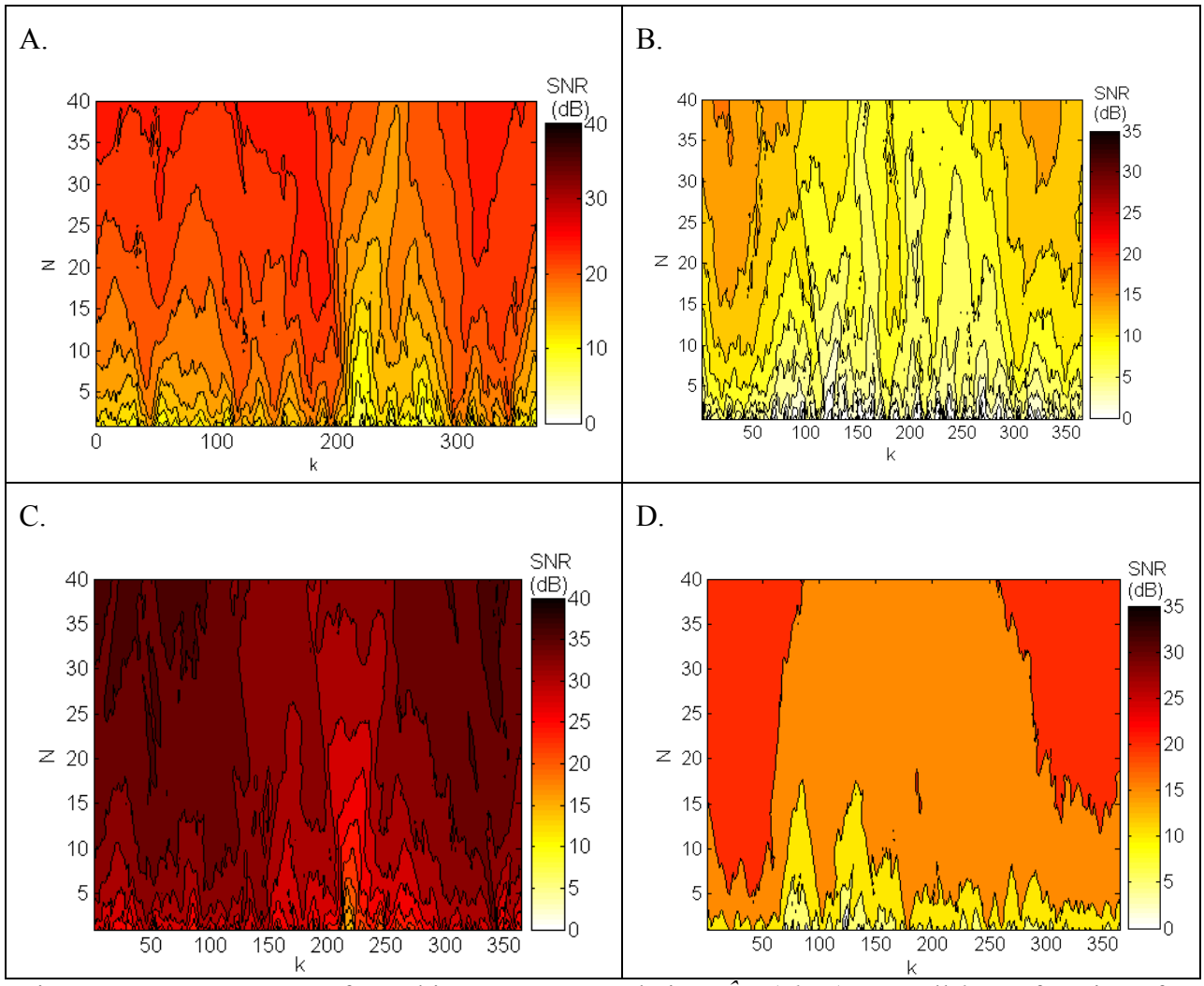

Figure 18: Average SNR of an arbitrary cross-correlation, $\hat{C}_{2,2}(t,k;N)$, over all k as a function of N for (a) Ascension and (b) Wake. Average SNR_B over all k as a function of N at (c) Ascension and (d) Wake.

3.5 Tracking Temperature Shifts; Comparison with Temperature Measurements

The arrival-time variations shown in Figure 17 were used to infer ocean temperature variations at the Wake Island and Ascension Island sites over several years (see Figure 19). These temperature variations are averaged over the effective depth extent of the SOFAR channel for several years at both sites (see Appendix E). Furthermore, overlaid dots on Figure 19

correspond to the monthly temperature changes independently estimated from Argo oceanographic floats measurements [5] -in the aforementioned effective depth of the SOFAR channel - between the north and south triads of each hydroacoustic stations over the same time period. The Argo program consists of thousands of autonomous floats that are distributed throughout the world's oceans [5]. These floats sink down to 2000 m depth, sampling temperature and salinity at periodic depths. Then the floats surface and relay their measurements and location via satellite. The Argo program has stored thousands of temperature profiles from the late 1990's through the present day, with more floats added to the program each year. Argo data is the current method used by climate scientists to track worldwide ocean temperature fluctuations below the surface.

The magnitude of the temperature variations in the SOFAR channel estimated from passive thermometry and from Argo data are in good agreement over the whole observation period, and the measurement error bars from passive thermometry are significantly smaller than those from the Argo data. An important methodology difference between these two estimates is that, given the sparseness of the global coverage by the Argo floats, the coarse Argo measurements are spatially and temporally interpolated to provide monthly estimates of global temperature variations [5]. On the other hand, passive thermometry measurements are not interpolated, as they result from a true sampling of the ocean by propagating sound; they thus integrate temperature variations over the whole propagation path (here ~130km between the North and South triads of each site) and averaging duration (one week). At the Wake Island site, where data are measured only over 5 years, the Argo and thermometry data are found to be 54% correlated. Both data indicate a very small upward trend. The Argo data shows a trend of 0.003

°C /year ± 0.001 °C/ year, for 95% confidence interval, and the thermometry data shows a trend of 0.007 °C /year ± 0.002 °C/ year, for 95% confidence interval. On the other hand, for the Ascension site, the SOFAR channel temperature variations measured over a longer duration of eight years from passive thermometry and Argo data are found to be significantly correlated, with a 0.8 correlation coefficient. This indicates that this passive thermometry method is sufficiently accurate to detect mesoscale variations of ocean temperatures occurring within the effective depth of the SOFAR channel. Furthermore, Figure 19b indicates a warming of the SOFAR channel in the Ascension area, as inferred from the similar upward trend of both passive thermometry (0.013 °C /year ± 0.001 °C/ year, for 95% confidence interval) and Argo (0.013 °C/ year ± 0.004 °C/ year, for 95% confidence interval) temperature variation estimates, this finding is consistent with the recent report of increased heat transport to the deep southern Atlantic ocean over the same time period [1].

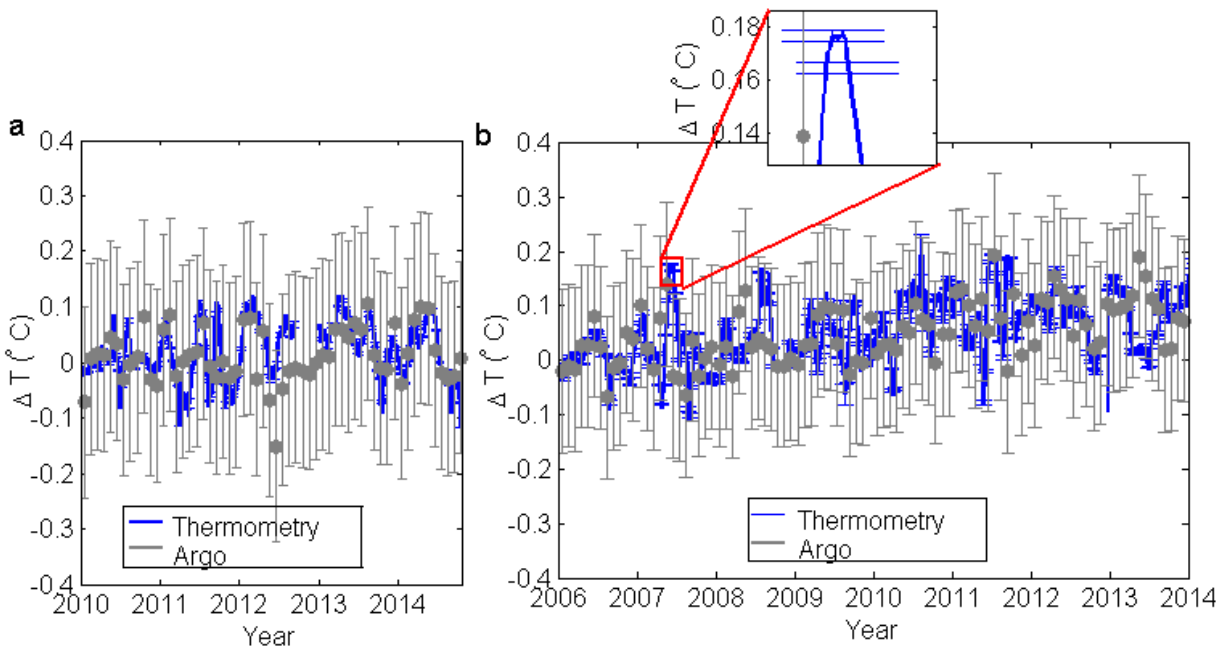

Figure 19: (a) Comparison of the deep ocean temperature variations at the Wake Island site estimated from passive thermometry (blue line) -using the SOFAR arrival-time variations- with free-drifting profiling oceanographic Argo float measurements (grey dots), along with corresponding error bars. (b) Same as (a), but for the Ascension Island site. Each ΔT data series is normalized so that a linear fit on the data would have a y-intercept at zero.

For the passive acoustic tomography data, change in sound speed is found from change in arrival time by the finite difference formula below, similar to the formula in the previous section, where t_0 is the absolute arrival time between array centers and c_0 is the absolute sound speed.

$$\frac{\Delta t}{t_0} = \frac{-\Delta c}{c_0} \tag{15}$$

Change in sound speed is related to change in temperature by the linearized relationship: [6]

$$\frac{\Delta c}{c_0} = \alpha \Delta T \left(1 + \frac{\mu \beta}{\alpha}\right) \tag{16}$$

The numerical values for α and μβ/α come from the equation for a typical sound speed profile, and are assumed to be $\alpha=3.19 \times 10^{-3}$ $(°C)^{-1}$ and $\mu\beta/\alpha=0.03$ [6]. Combining these two equations yields the following formulation.

$$\frac{\Delta t}{t_0} = -\alpha \Delta T\left(1 + \frac{\mu\beta}{\alpha}\right) \tag{17}$$

A localized linear relationship between temperature and salinity is assumed in the derivation of μ. The derivation of α and β is shown below. The Mackenzie equation for sound speed (c in m/s) as a function of temperature (T in $°C$), salinity (S in parts per thousand), and depth (D in meters) is [29]:

$$\begin{aligned} c(T,S,D) = {} & 1448.96 + 4.591T - 0.05304\,T^2 + 2.374 \times 10^{-4}\,T^3 \\ & + 1.340\,(S-35) + 1.630 \times 10^{-2}\,D + 1.675 \times 10^{-7}\,D^2 \\ & -1.025 \times 10^{-2}\,T(S-35) - 7.139 \times 10^{-13}\,TD^3. \end{aligned} \tag{18}$$

Following the procedure outlined in [30], taking partial derivatives of c with respect to T and S yields:

$$\frac{\partial c}{\partial T} = 4.591 + higher\ order\ terms \tag{19}$$

$$\frac{\partial c}{\partial S} = 1.340 + higher\ order\ terms \tag{20}$$

Dividing the first-order approximations of the partial derivatives by a nominal sound speed, c_0 = 1480 m/s, gives the alpha and beta parameters.

$$\alpha = \frac{\partial C}{\partial T} \cdot \frac{1}{C_0} = 3.1\ x\ 10^{-3}\ (℃)^{-1} \quad \textbf{(21)}$$

$$\beta = \frac{\partial C}{\partial S} \cdot \frac{1}{C_0} = 0.9\ x\ 10^{-3}\ (‰)^{-1} \quad \textbf{(22)}$$

3.6 Confidence in Arrival Time Shift Causation

It is important to understand the physical reason for the arrival time shifts to establish confidence that these shifts are indeed due to water temperature change. Are the arrival time shifts due to drifting sensors, currents, clock drift, or temperature? Currents would not cause a symmetric shift in positive and negative correlations unless the north and south arrays were exposed to currents flowing in opposite directions, which is very unlikely. Because the sensors are hardwired to the monitoring station on land, the setup ensures precise timing. Therefore, clock drift is also unlikely. The remaining causes of the time shifts have been eliminated to drifting sensors and / or temperature variations. To understand the effect of sensor drift, it is necessary to use the finite difference formula below, relating change in arrival time and change in array position. In this equation, d_0 is the distance between array centers, t_0 is the absolute travel time between array centers, Δd is the change in distance, and Δt is the time shift.

$$\frac{\Delta d}{d_0} = \frac{\Delta t}{t_0} \quad \textbf{(23)}$$

Because the maximum arrival time shift is approximately 40 ms, this equation can be used to find the corresponding potential sensor drift. Absolute values from the Wake Island site will be used, where d_0 = 138 km, t_0 = 93 s. From these values, a time shift of 40 ms would correspond to a position change of ~50 m. Because the hydrophones are on the SOFAR channel (~800 m depth) and anchored to the sea floor, it is unlikely that they would drift so much. This means that changes in arrival time are most likely due to changes in temperature. This conclusion is confirmed in the next section.

3.7 Comparison with a Normal Modes Propagation model

In order to understand the relationship between arrival-time variations of the SOFAR arrival and estimated temperature change of the SOFAR layer, independent normal mode numerical simulations were run to predict the SOFAR arrival time shifts based on the known characteristics of sound propagation in the SOFAR channel. A similar model using the parabolic equation technique is discussed in Appendix F. Given the sensor separation and north-south orientation between the two triads at both sites, noise detected at both sites largely emanates from the Polar regions and propagates in the SOFAR waveguide up to the mid-latitude regions where the hydroacoustic stations are located. Because the noise propagates over such long ranges and has a low center frequency of around 10 Hz, most of the acoustic energy trapped within the SOFAR channel reaching the hydroacoustic stations is primarily carried via the first (i.e. lowest) normal mode of the ocean water column [6]. Hence, the arrival-time variations of the SOFAR arrival were calculated numerically for the actual instrument positions and depths using a normal-modes acoustic propagation code [31]. For each month of a climatological year, the

relationship between temperature perturbations and arrival time perturbations was calculated by propagating in the perturbed temperature and salinity fields derived from Argo data. The propagation model was run for three cases to estimate the arrival time of the main SOFAR arrival (for the sound propagating between the centers of the North and South triad arrays at H10 using either:

1. A single frequency of 10 Hz (i.e. the center frequency) and computing the group speed of the lowest mode 1. SOFAR arrival times were estimated from the ratio of the separation distance between the centers of the North and South triad array and the group speed value of mode 1.

2. A broadband (1-40 Hz) Fourier synthesis of the SOFAR arrival assuming that only mode 1 propagates. Arrivals were weighted by a spectral amplitude equal to the averaged spectral amplitude of SOFAR arrivals shown in Figure 14 a.

3. A full broadband solution, via Fourier synthesis (1-40 Hz), including the contributions of both mode 1 and mode 2. The same spectral amplitude weightings were used as stated above.

The arrival structure at H10 is shown in Figure 20 for three separate normal mode models. The arrivals are derived using the Argo sound speed profile for the month of January in 2006. The figure illustrates how most of the propagating energy is located in the SOFAR channel, centered on the SOFAR axis depth of ~900 m. Arrivals are spectrally weighted by the reference waveform at H10 for the cross-correlation between south hydrophone #1 and north hydrophone #1.

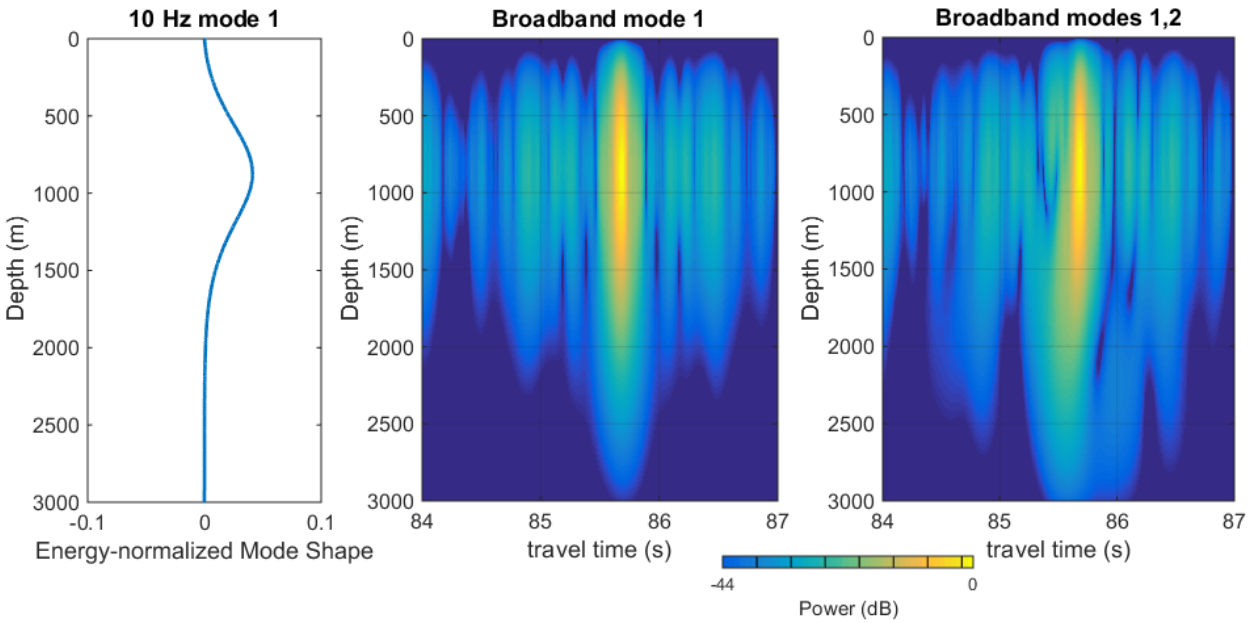

Figure 20: Arrival structure as a function of depth (z in meters), predicted for January 2006 at H10, using a center frequency mode 1 approximation, a broadband mode 1 model, and a broadband combination of modes 1 and 2.

Arrival time shifts predicted by the normal-mode simulations are shown in Figure 21 and compared to predicted arrival-time fluctuation obtained from Equation 19 using Argo temperature data. The results found from the normal-mode simulations differed from those calculated by Equation 16 by a mean of 3 ms and a standard deviation of 3 ms with a minimum of 0.4 ms and a maximum of 10 ms. This good agreement justifies using Equation 16 to provide a simple, yet accurate, conversion between arrival-time variations of the SOFAR arrival and estimated temperature changes of the SOFAR layer. This good agreement also demonstrates that for Equation 16 (i.e. the Munk Equation) to be valid, acoustic propagation must be largely limited to the SOFAR channel, where energy from mode 1 is dominant. The PE model in the previous section failed to limit acoustic energy propagation to the SOFAR channel, and therefore resulted in predicted arrival time fluctuations that were overly-influenced by the large sound

speed variations in the upper portion of the water column.

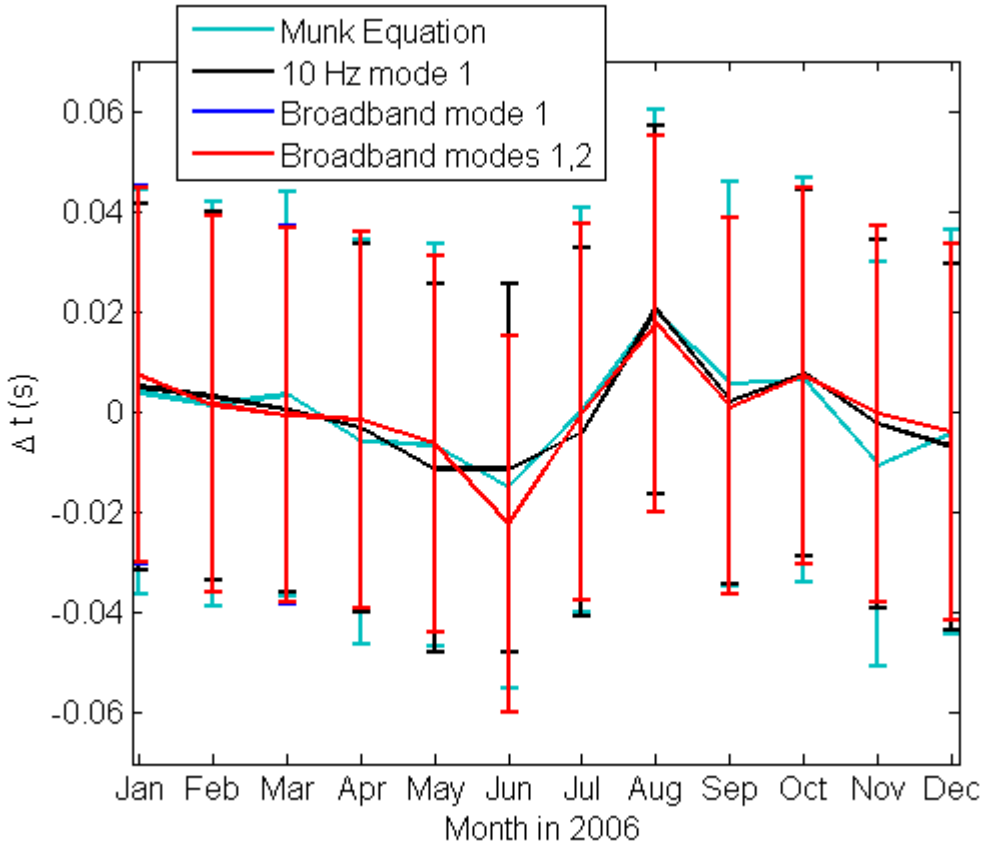

Figure 21: Comparison of the predicted arrival time changes using Equation 16 or a full normal mode propagation model using three different cases (single frequency vs. broadband excitation). Results are scaled to have a mean arrival time of zero over the 12 month period. Arrival time fluctuations predicted by Equation 16 are virtually identical to those predicted by the various implementation of the normal mode propagation model.

3.8 Chapter Conclusions

This chapter demonstrated the tracking of deep ocean temperatures using a completely passive acoustic tomography method. Temperature fluctuations were tracked over a time period of up to 9 years, spatially averaged over the length (~130 km) and depth (see Appendix E) of the SOFAR channel between the north and south triads. Results show an order of magnitude

increase in precision over the current state-of-the-art (i.e. Argo floats) while predicting temperature changes that are consistent with the float measurements. Ultimately, the results of this section provide substantial evidence that passive acoustic tomography is a viable alternative to current active acoustic methods.

Chapter 4: Variability of the coherent arrivals extracted from low-frequency deep ocean ambient noise correlations

4.1 Introduction

The ability to extract coherent arrivals from ambient noise correlations as rapidly as possible and in a robust fashion is subject to the physical characteristics of the ambient noise field (e.g. degrees of anisotropy, statistics of the noise sources, etc.) [32], [19], [33]. For instance, an anisotropic noise field will make it more difficult to extract the coherent components of the noise propagating along the ray-paths connecting the sensors. Additionally, these passive ocean monitoring methods require a sufficiently rapid emergence rate of persistent coherent arrivals (over residual temporal fluctuations of the correlation waveforms) to mitigate the effects of environmental fluctuations during the recording interval. Hence the focus of this chapter is the spatially-diffuse, temporally-consistent component of the low-frequency ($f<40$ Hz) background ambient noise propagating in the deep water sound channel, as this is best suited for practical implementation of passive ocean monitoring techniques, such as passive acoustic thermometry [27]. Low-frequency ocean noise is of particular interest because it tends to generally be more energetic than higher-frequency noise [15] and because low-frequency acoustic propagation can be fairly stable especially in the deep water sound channel. Furthermore, deep oceans play a major role in absorbing atmospheric heat; hence measurement of their temperature variations is necessary to quantify air-sea heat exchanges that play a role in assessing global warming trends [1] and in calibrating climate change models [2], [3].

In contrast to the episodic events studied in previous work (e.g. isolated ships, whale calls) [34], this work focuses on the spatially-diffuse, temporally-consistent component of the ambient noise field, which is referred to as the background component. Overall, this homogeneous background component is what contributes to the emergence of the coherent arrivals from ambient noise correlations [32], [33], [10]. Hence, it is the spatial and temporal variability of this background component that primarily affects the ability to extract coherent arrivals that are useful for passive ocean monitoring purposes. The exception to this is when episodic events are located in-line with a sensor pair and contribute to the coherent arrival [11], [9], [26], [35]. Furthermore, while previous work on passive monitoring has focused on simple shallow water environments [8] or a single deep water site [28], this chapter systematically compares multiple deep water sites around the world to arrive at some broad recommendations for the deployment of future deep water hydrophone arrays for passive monitoring applications. In this chapter, the variability of the coherent components of the low-frequency ambient noise is examined for all five sites located in the Atlantic, Indian and Pacific oceans (see Figure 5).

4.2 Data Processing

Continuous recordings of underwater noise used in this study were collected during the year 2010 (except for the site H03, where only recordings from the year 2009 was available for this study), with a sampling rate of 250 Hz, at the hydroacoustic stations shown in Figure 5. Ambient noise recordings were filtered in the low-frequency band $1\ \text{Hz} < f < 40\text{Hz}$. The 1 Hz limit was set by the roll-off of the hydrophone response at the lower frequency end of the spectrum, and most of the energy of the background component of the ambient noise field occurred below 40 Hz [28], [36].

The recorded data at all the sites contains a combination of episodic events and more diffuse background noise. Noise sources in the very low ($f < 50$ Hz) frequency range include shipping [37], whale [38], [39], ice-noise [40], [41], earthquake [42], and seismic air-gun sources [43]. Following the data processing steps outlined in section 3.2, data recorded at each hydrophone are truncated into hour-long segments, then amplitude-clipped and frequency-whitened in order to reduce the influence of episodic high-amplitude transient events while preserving the overall phase information of the original time series. Thus, these pre-processing steps are used to emphasize the background component of the ambient noise field. Quantitatively speaking, the level of spatial coherency of the background ambient noise field between a given hydrophone pair (i,j) can be measured by the peak amplitude of the cross-correlation waveform of the ambient noise recorded by these two hydrophones: this peak amplitude is a measure of the similarity of signals collected at two spatially-separated hydrophones [21]. Cross-correlations are performed on each hour-long segment between hydrophone i and hydrophone j (i,j=1,2,3) at each array, and energy-normalized by the energy measured by both hydrophones over that hour, as shown in Equation 10. The difference between the cross-correlations used in this section and those described in Equation 10 is that in this chapter, cross-correlations are computed between hydrophones at each triad, not between triads. Autocorrelations (where $i=j$) are not used in this study. The hour-long cross-correlation segments are averaged over a period of 24 hours, providing a final result of one averaged cross-correlation for each day (k) in the year 2010 between each hydrophone pair at each site.

In theory, the time-averaged cross-correlation provides an estimate of the arrival-time structure of the Green's function between sensor pairs [22]. Because $C_{i,j}(t,k) = C_{j,i}(-t,k)$, there are

three unique short-range cross correlations corresponding to distinct pairwise combinations of the hydrophone elements of a single triangular array for each site: $C_{1,2}(t,k)$, $C_{1,3}(t,k)$, and $C_{2,3}(t,k)$. The SNR of the cross-correlation waveform for each sensor pair (i,j) at each day k is defined in the equation below. The numerator is the peak amplitude of the main arrival of the cross-correlation waveform $C_{i,j}(t,k)$, and the denominator represents three times the standard deviation of the correlation taken during the lag time interval $[t_{n1}\ t_{n2}]$, which is considered to be far outside of the range of possible physical water-borne arrivals. For all sites, t_{n1} = 115 s and t_{n2} = 120 s.

$$SNR_{i,j}(k) = \frac{\max\left(C_{i,j}(t,k)\right)}{3std\left(C_{i,j}([t_{n1}\ t_{n2}],k)\right)} \tag{24}$$

For a given sensor pair, the SNR value can be used to estimate the precision of the arrival-time measurement of the corresponding peak arrival [6], [27]. For instance, a higher SNR means that arrival time variations can be tracked with more precision. For acoustic thermometry purposes, this means that ocean temperatures variations, which are related to sound speed variations, can ultimately be tracked with more precision.

4.3 Directionality of the Low-Frequency Coherent Noise

As discussed by Roux *et al*. [11], when coherent noise travels along the sensor pair axis within a pair of beams centered on the endfire direction, that noise is the portion of the ambient noise field that contributes to increasing the SNR of the coherent arrivals. The coherent arrivals correspond to actual Green's functions arrivals. This portion of the noise field is referred to hereafter as coherent on-axis noise. Conversely, off-axis noise sources primarily generate the residual fluctuations of the cross-correlation waveform. Loud, episodic, off-axis noise sources

(hereafter called loud interferers) also generate spurious arrivals. Hence, extracting an accurate Green's function estimate from ambient noise correlation primarily depends on the amount of coherent noise sources that line up with the sensor axis relative to the amount of incoherent noise sources located away from the sensor axis. Additionally, for passive acoustic thermometry purposes, it is best if the generation mechanism of the on-axis sources is relatively stable over time in order to extract persistent coherent arrivals corresponding to on-axis noise propagating along the path between sensor pairs. Hence the azimuthal variability and directionality of the spatial coherence of the noise field is investigated hereafter for all five hydroacoustic stations in order to understand which hydrophone orientation pair is best suited for passive acoustic thermometry at the selected sites.

First, windowing was applied to the cross-correlations at each site to isolate the coherent arrivals that most correspond to the direct SOFAR arrival between selected sensor pairs. The purpose of the windowing was to isolate the coherent arrivals that most closely correspond to the direct SOFAR arrival between the sensor pairs. The direct arrival along the SOFAR channel axis is of particular interest for passive acoustic thermometry because it is loud, energetic, and stable [6]. The selected window here is a 200 ms-wide Tukey window centered on the last water-borne arrival along the sensor pair axis. The window centers are based on the average SOFAR channel sound speeds (known within ± 2 m/s for 95% confidence) estimated at each location using data from the Argo floats measurements [5], i.e. respectively 1484 m/s at H10, 1489 m/s at H11, 1481 m/s at H03, 1489 m/s at H08, and 1482 m/s at H01. The window width was determined by the estimated sensor positioning errors (on the order of ~200m) due to motion of the moored arrays to be 200 ms; this choice was further confirmed by visually inspecting the persistent part of the

cross-correlation stacked over the year to remove clearly-identifiable off-axis interferers. To illustrate this process, Figure 22 shows the average cross correlation between sensors 1 and 2 at H11 North array for each day of the year in 2010 (see array configuration in Figure 5). There are a few off-axis interferers at this location, but it is easy at this site to distinguish between the actual SOFAR arrival corresponding on-axis coherent noise and a spurious arrival caused by off-axis noise sources. At other sites, distinguishing between the SOFAR arrival and arrivals generated by slightly off-axis noise sources can be more difficult due to imprecise knowledge of the SOFAR sound speed [5]. Figure 22 shows the year-averaged cross-correlation waveform along with the window limits that more clearly isolate the SOFAR arrival of interest.

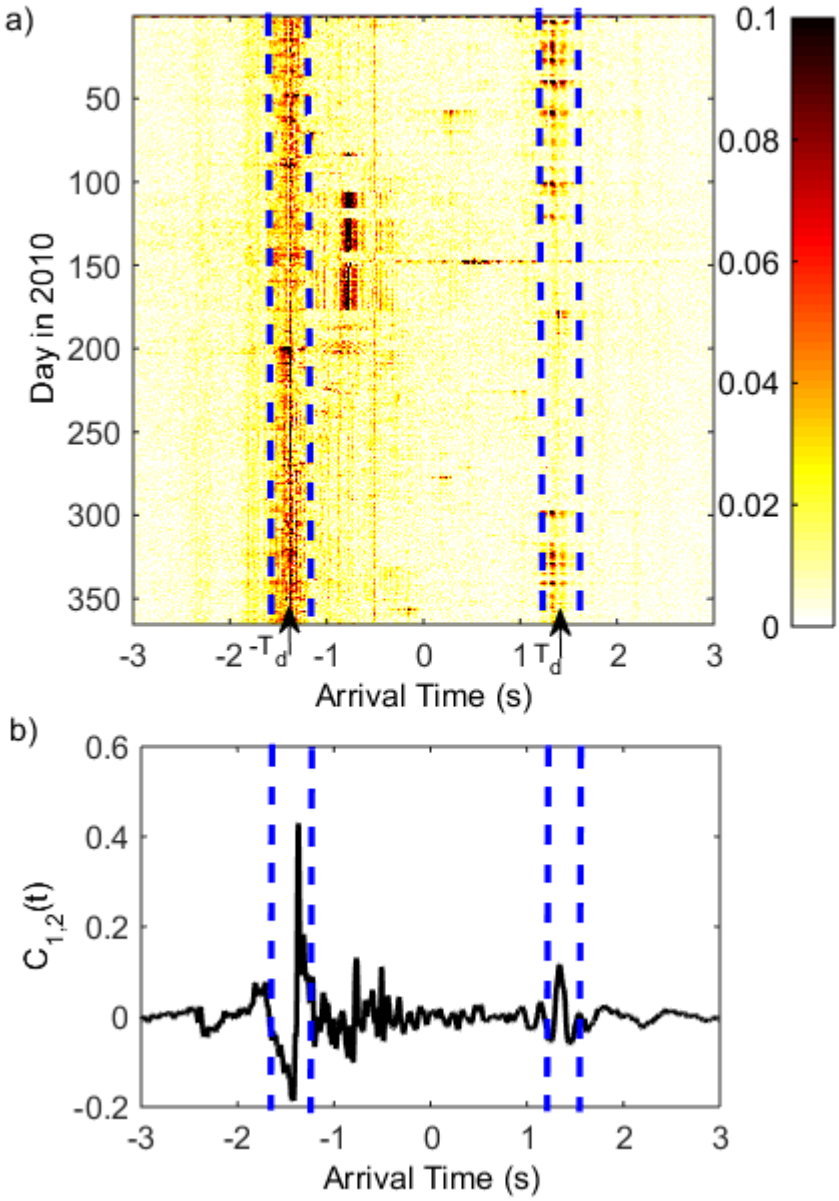

Figure 22: a) Evolution of the magnitude of the cross-correlation waveform (averaged over each successive 24h interval) between sensors 1 and 2 of the H11N array over the year 2010. By definition, the positive (resp. negative) time delays correspond to noise traveling from sensor 1 to sensor 2 (resp. sensor 2 to sensor 1). Note the presence of loud off-axis interferers. The 200 ms-wide time-gated window is shown with the dashed lines, centered on the expected arrival time Td = 1.37 s of the SOFAR arrival for the selected hydrophone pair for both positive and negative time delays. b) Average of all the cross-correlation waveforms shown in (a) over the entire year 2010. The limits for the time-gated window used in this reminder of this study are shown as dashed lines.

The peak amplitude of the main arrival (at both positive and negative time delays) of the time-windowed cross-correlation waveform averaged over the whole year 2010 (Fig 2(b)) is used to generate a vector plot that quantifies the ambient noise directionality and spatial coherence at each hydrophone triad (see Figure 23). The longest vectors show the predominant direction where the low-frequency noise is coming from at each site. Furthermore, Figure 24 shows the frequency content of coherent noise arrival corresponding to the dominant direction at each site. These spectra were computed by taking the Fast Fourier Transform (FFT) of the half (i.e. the positive or negative time domain) of the windowed cross-correlation, averaged over one year, that corresponds to the dominant coherent noise. These spectra can be used to infer the most likely generation mechanism for the dominant noise detected at each site

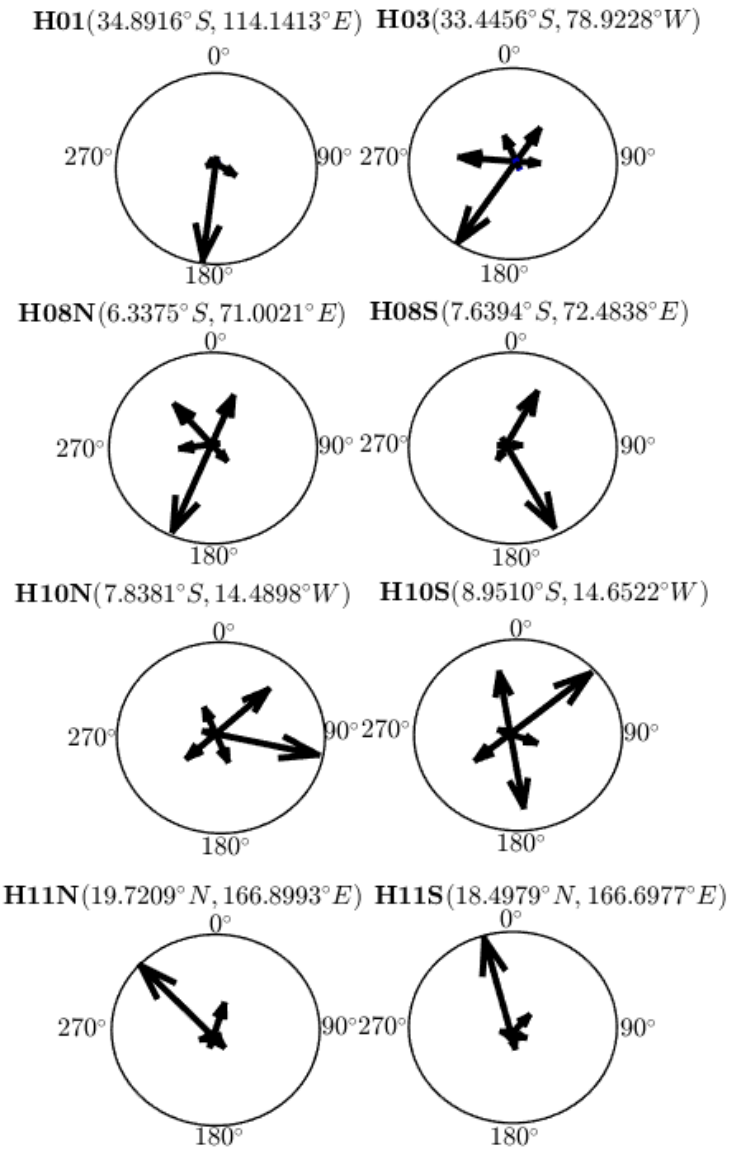

Figure 23: Normalized main arrival of the windowed cross-correlation averaged over one year for each short-range sensor pair at each array. The peak is plotted as the vector magnitude, and the vector direction corresponds to the sensor pair orientation, pointing in the direction the noise is coming from. At each site, the magnitude of the vector pointing in the dominant direction is normalized to one, and the magnitudes of the remaining vectors are linearly scaled from the dominant vector.

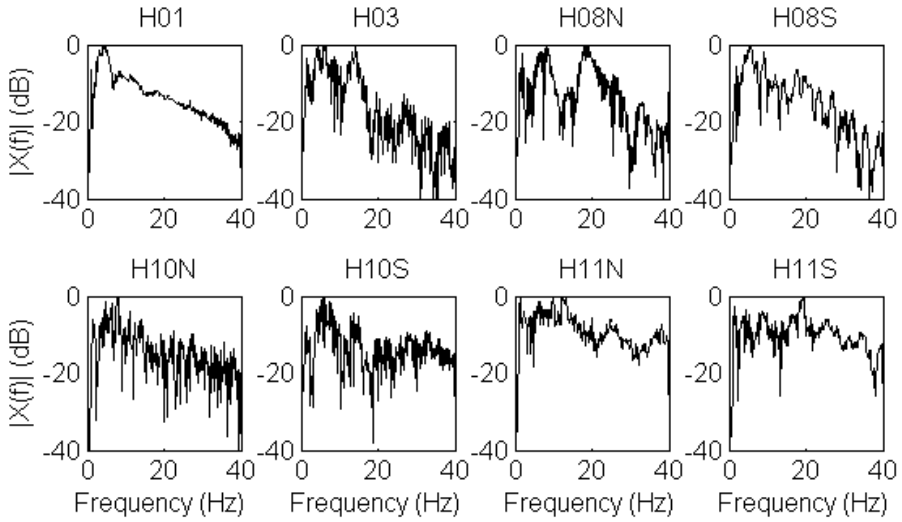

Figure 24: 1-40 Hz spectrum of the spatially-coherent noise traveling from the dominant direction (see Figure 23) at each site, calculated from the time-windowed cross-correlation averaged over the year 2010.

As stated at the beginning of this section, sources that are in-line (± a maximum of 20°, given the 200 ms Tukey window) with a given sensor pair contribute to the coherent SOFAR arrival of interest [11]. To investigate the spatial origin of the coherent noise propagating between the sensor pairs, a conventional time-delay beamforming procedure (see section 2.3), [28], is applied to the coherent SOFAR arrivals (extracted from windowed, time-averaged cross-correlations), using the same nominal sound speeds in the SOFAR channel previously stated. For instance, Figure 25 displays the azimuthal variations of the peak value of the time-domain plane-wave beamformer output $|B(\theta)|$ = for H01 and H08N, which can be used to estimate the angular width of main lobe associated with the dominant noise direction.

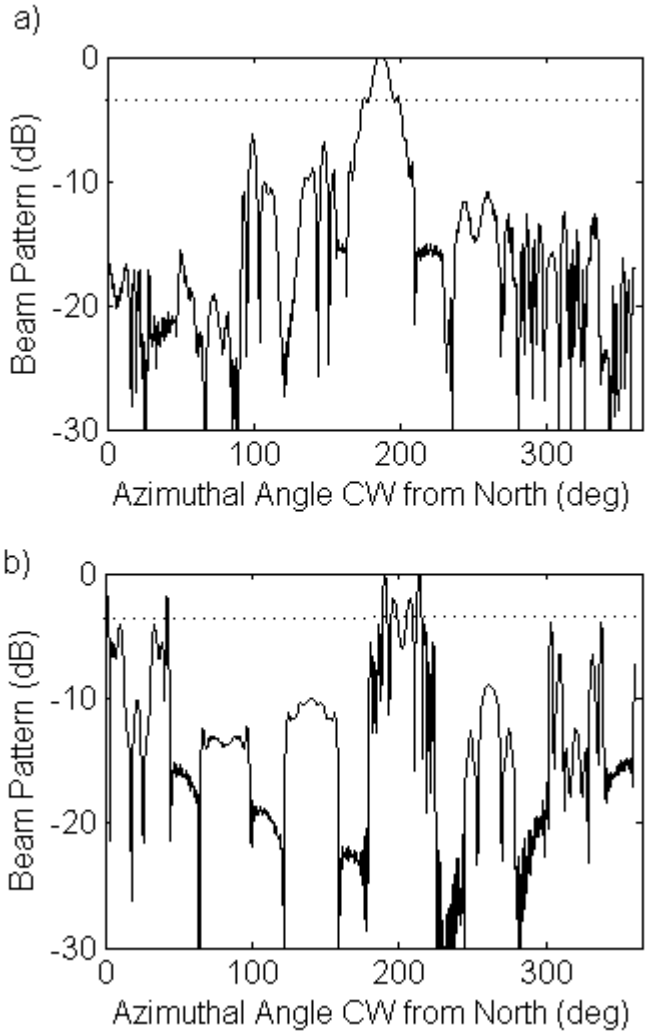

Figure 25: a) Beam pattern for plane wave beamformed H01 array cross-correlations time-windowed by 200 ms (corresponding to a beam width of 40°) in the band 1-40 Hz. b) Same but for H08 North array. The half-power (-3 dB) point is marked with a dashed line, and the azimuthal search angle is measured clockwise from the north.

The main lobe of the beamformer output is defined as the angular interval for which the normalized beamformer output remains above -3dB, which was found to correspond to an angular beamwidth of between 20° and 40°, depending on the site. At these Indian Ocean sites, the main lobe of the beamformer output is, as expected [28], aligned along the line axis (i.e. end-

fire direction) joining the sensor pairs oriented north-south and pointing towards the Antarctica coastline. Assuming a 2-dimensional geodesic propagation path, these angular beams are projected on Figure 26 to show the potential spatial locations of the noise sources that contribute to the most energetic coherent arrivals at each site (i.e. the longest vectors in Figure 23).

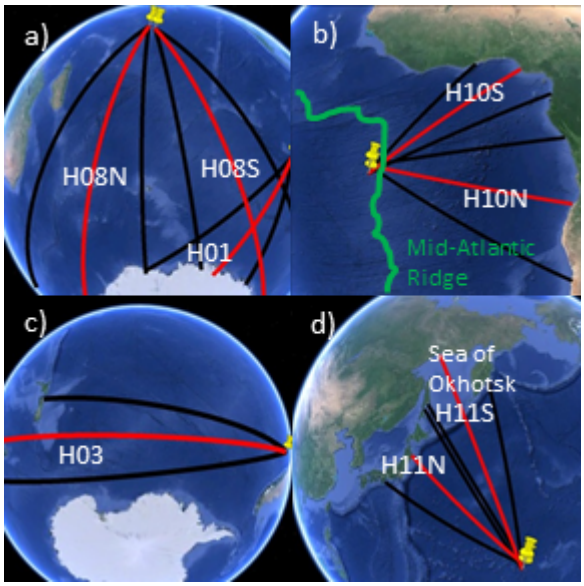

Figure 26: The possible locations of coherent noise sources detected by the dominant pair at each site. The beam widths (black lines) are shown around the projection of the main north-south array axes (center lines) in the a) Indian Ocean b) Southern Atlantic Ocean, c) Southern Pacific Ocean and d) Northern Pacific Ocean (Maps from ©Google Maps)

To examine how the noise field directionality changes over season and frequency band, the cross-correlation waveforms computed between all sensor pairs were averaged over successive one month intervals and then filtered into consecutive 5 Hz bands from 1 Hz to 40 Hz. This yielded 12 sets of cross-correlation waveforms (one per month) over the year 2010 per frequency band at each site. Then the directionality plot (similar to the ones shown on Figure 23) for each case (i.e, frequency band, month, and site) were examined to determine the monthly

variability of the dominant direction. Here, the variability is quantified as the number of times the largest vector for each site (e.g. as shown in Figure 23) changes direction (out of 6 possible directions) over the 12 months, i.e. from a minimum of zero times to a maximum of 11 times. The variability is color-coded and displayed in Figure 27(a) for each frequency band and site, and provides a metric of the temporal stability of the dominant noise direction at each site across frequency. Furthermore, Figure 27(b) provides a measure of the degree of directionality at each site by counting the number of months that the amplitude of the largest side lobe exceeds 0.8 times the amplitude of the main lobe (the factor 0.8 being selected as an arbitrary threshold). Figure 27 shows that the H11 north and south arrays, the H08 south array, and the H01 array contain relatively stable noise field directionality over time and frequency. Conversely, the rest of the sites (H10 north and south arrays, H08 north array, and H03 arrays) exhibit highly variable directionality across both time and frequency band.

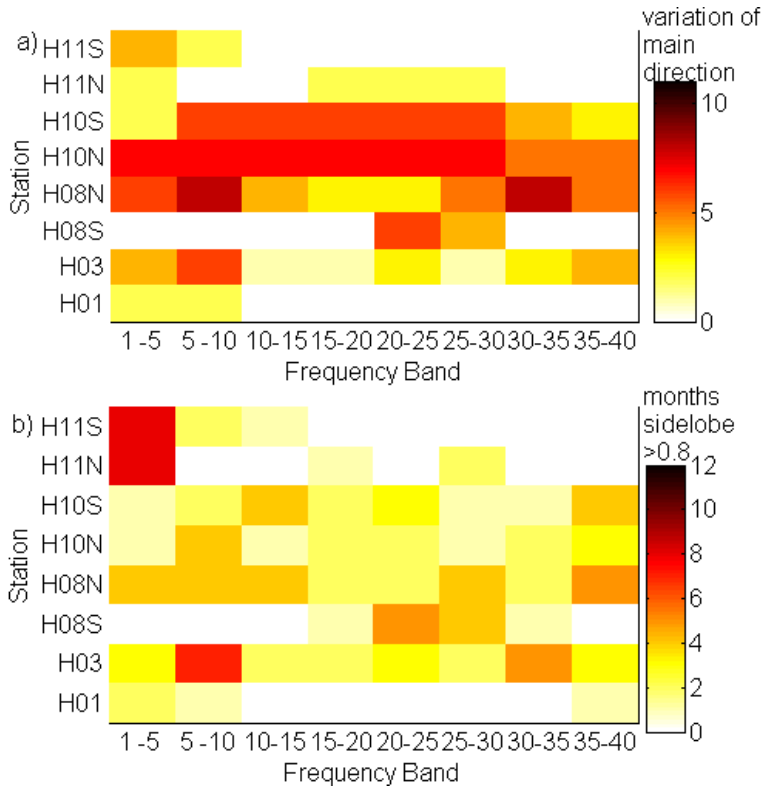

Figure 27: a) The number of times the dominant direction (large arrows in Fig. 3) changes over the months in each 5 Hz band at each site. b) The number of months that the sidelobe (second largest arrow in Figure 23) height is >0.8 for each 5 Hz band at each site.

4.4 Influence of Noise Source Characteristics on the Emergence of Coherent Arrivals

This section examines the noise source mechanisms that are likely to be predominant at each hydroacoustic station and discusses how each noise source type affects the temporal stability of the dominant coherent noise component. Additionally, this section discusses how each noise source type affects the noise field stability over time and the ability to extract high-SNR coherent arrivals from ambient noise correlations for passive ocean monitoring applications. For instance, an SNR higher than 20 dB would be necessary to measure arrival time variations on the order of milliseconds for acoustic tomography or thermometry purposes [6].

As discussed previously, there are a variety of potential noise sources in the 1-40 Hz band. The degree to which sound from each of these sources is present varies from site to site, but for the purposes of passive ocean monitoring, consistent sources are preferable to episodic ones. Figure 26 can be used to infer the likely origin of the dominant noise sources at each site (i.e. located within the angular beam delimited by each black lines at each site).

When the directionality and seasonality of the noise field at these sites are both taken into account, it is likely that sound generated by sea ice and icebergs off the coast of Antarctica is dominant at sites located in the Indian Ocean (H08N, H08S, and H01 arrays). Additionally, Figure 24 shows peaks at 6 Hz for H01 and H08S, likely corresponding to the fundamental frequency of structural resonances of icebergs off the coast of Antarctica [44], [45]. Overall, these findings are consistent with previous studies [46], [41], [36] that demonstrated the prevalence of low-frequency ice-generated noise in the Indian Ocean. On the other hand, at the H10 arrays, the angular beam intersects the Mid-Atlantic ridge, an area with seismic noise that has been shown to couple to the SOFAR channel [42]. Additionally, Figure 27 shows a high level of variability in the directionality of the noise field at H10 across time and frequency, implying that the noise field at H10 is generated by a summation of several types of sources distributed across space and frequency. Therefore it is hypothesized that seismic air gun survey noise, which are predominant in the Atlantic ocean, is the predominant sound source for these short-range Green's function estimates [43], [41]. H03 is difficult to interpret because the projection of the main noise direction does not intersect with obvious noise sources, as the Juan Fernandez Island blocks ice-generated noise coming from Antarctica, and the coastline of South America blocks most westward-propagating noise. The field is not very stable or directional over

time or frequency, exhibiting high sidelobe levels and a high degree of variability in the orientation of the dominant direction. At this site, sound is likely generated from a combination of many distant noise sources of various origins. Finally, the dominant noise direction for both H11 arrays continually point northwest towards Japan and the Sea of Okhotsk. The Sea of Okhotsk contains year-round sea ice [47], and the trench off the coast of Japan is an area of high seismic activity. I hypothesize that both of these sources combine to create a highly directional, consistent broadband (1-40 Hz) coherent noise field at this location across seasons, as mentioned previously.

To maximize the SNR of the SOFAR coherent arrivals of interest - and therefore maximize the precision of arrival time estimates for passive tomography purposes - hydrophone pairs should be oriented to have a direct line-of-sight with consistent noise sources over time and frequency. To quantify the emergence rate of the coherent arrivals used in this study, the SNR of the cross-correlations between selected hydrophone pairs was computed for increasing number N of averaging days for the year 2010 (N=1…365). First, Figure 28 shows SNR variations at sites that were found to have a dominant noise field directionality that is consistent over time and frequency (based on Figs. 3, 5, and 6). Hydrophone pairs were selected to align with the dominant noise direction.

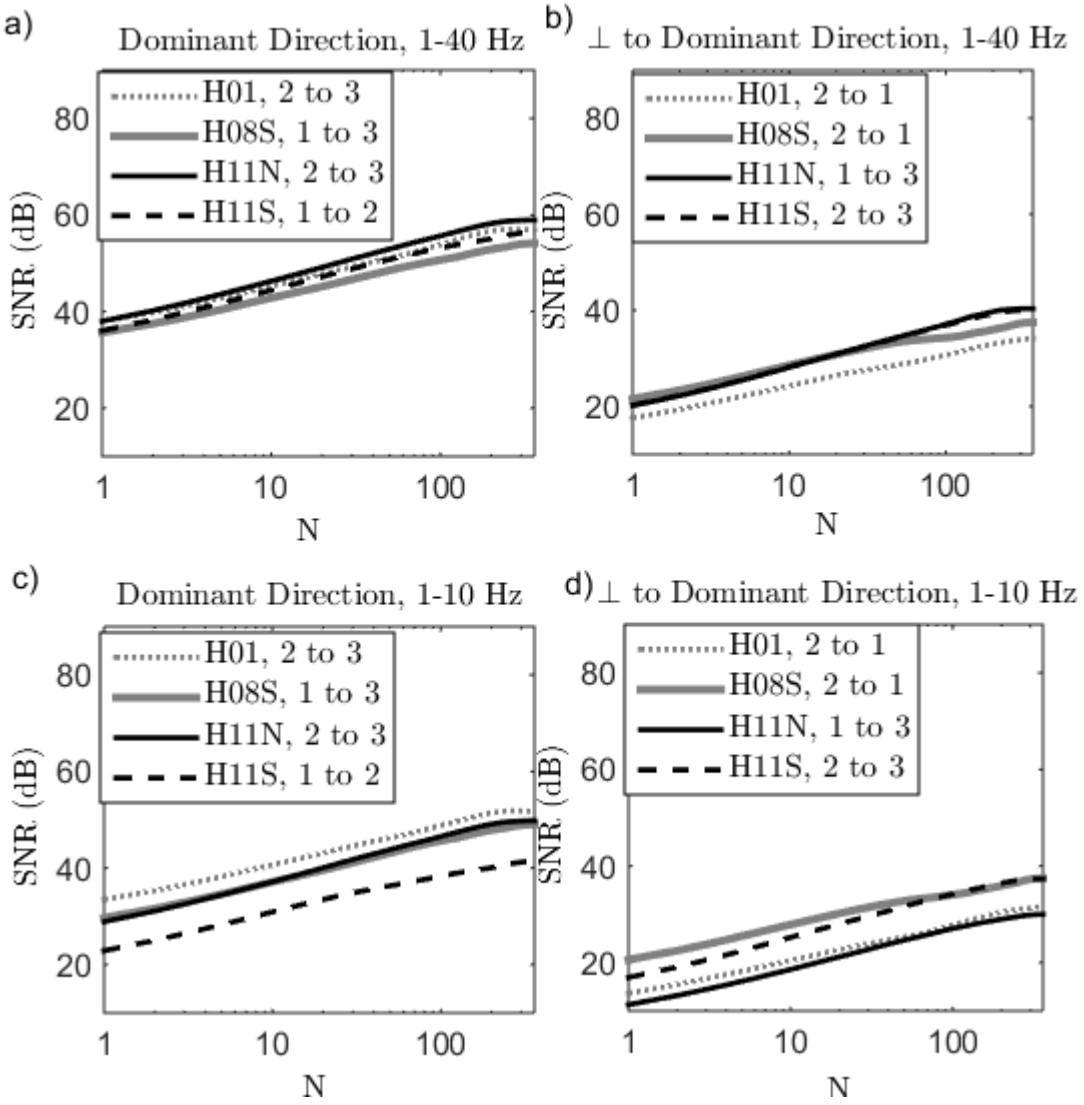

Figure 28: Evolution of the SNR of the coherent arrivals band for increasing number N of averaging days for the year 2010 (N=1…365), between hydrophone pairs at sites with a high degree of noise field stability over time for pairs a) pointing towards the dominant noise directions (Antarctica or the Sea of Othotsk) computed in the 1-40 Hz frequency b) Same as (a) but for hydrophone pairs pointing roughly perpendicular from dominant noise direction. c) Same as (a) but using the 1-10 Hz frequency band. d) Same as (b) but using the 1-10 Hz frequency band.

At H01 and H08S, the pairs were oriented approximately along the the north-south direction, thus pointing towards Antarctica. At the H11N and H11S arrays, the pairs were oriented to detect

ice-generated noise emanating from the Sea of Othotsk and seismic noise from the trench off the coast of Japan. Specifically, these hydrophone pairs correspond to the longest vectors shown in Figure 23, and are namely hydrophones 2 and 3 at H01, hydrophones 1 and 3 at H08S, hydrophones 2 and 3 at H11N, and hydrophones 1 and 2 at H11S. As expected, Figure 28(a) indicates that high SNR values (i.e., >> 20 dB) are rapidly achieved for the coherent SOFAR arrivals obtained between these selected hydrophone pairs in the frequency band 1-40 Hz. For comparison, Figure 28(b) displays SNR variations for the same sites, but using instead hydrophone pairs that are aligned roughly perpendicularly to the dominant noise direction. These arrivals correspond to noise propagating from hydrophone 2 to 1 at H01, hydrophone 1 to 3 at H11N, hydrophone 2 to 3 at H11S, and hydrophone 2 to 1 at H08S. These hydrophone pairs were selected to capture the fewest transient events in order to obtain coherent SOFAR arrivals that are mostly generated by the consistent background noise field. Figure 28(b) confirms that lower SNR values - when compared to Figure 28(b) - are achieved for the coherent SOFAR arrivals in the frequency band 1-40 Hz. These lower SNR values are obtained because the hydrophone pairs used in Figure 28(b) have a less favorable orientation: pointing away from Antarctica or seismically active regions.

Furthermore, Figure 28(c)-(d) shows SNR variations for the same parameters used to generate Figure 28(a)-(b), except that a narrower frequency band 1-10 Hz was used. Theoretically, the SNR scales with the square-root of the effective frequency bandwidth of the coherent noise field component [11], [48]. Comparing Figure 28(a)-(b) with Figure 28(c)-(d) shows that the consistent and dominant noise sources tend to be more broadband than noise sources which are located in the non-dominant direction. This can be deduced because the SNR

values only drop significantly when the frequency bandwidth is reduced for the hydrophone pairs pointing towards the consistent and dominant noise sources (compare Figure 28(a) to Figure 28(c)). We hypothesize that this is because the consistent noise sources examined in this study tend to be sea ice or seismic sources, which continuously generate broadband impulse-like sound [41], although sea ice also generates tonal noise in addition to the broadband components. Furthermore, the broadband nature of these noise sources is emphasized by the frequency-whitening preprocessing step applied to the recorded noise data prior to computing the noise correlations. Conversely, sporadic interferers – such as ships or whales- that are prevalent in the East-West direction tend to be more narrow-band with lower frequency content [43]. The exception to this is broadband noise generated by seismic air guns. Nevertheless, hydrophone pairs oriented perpendicular to the dominant noise source direction in stable noise fields can still obtain relatively high-SNR arrivals. However, these hydrophone pairs require a longer averaging duration N to achieve a high SNR due to their less favorable orientation.

Finally,

Figure 29(a) and (b) show similar results to Figure 28(a) and (c) (i.e., using hydrophone pairs pointing towards stable, ice-generated or seismic noise), but for the sites H10N and H10S where the dominant direction was found to be unstable. Overall,

Figure 29(a) and (b) show similar SNR variations and frequency-dependence as in Figure 28(a) and (c).

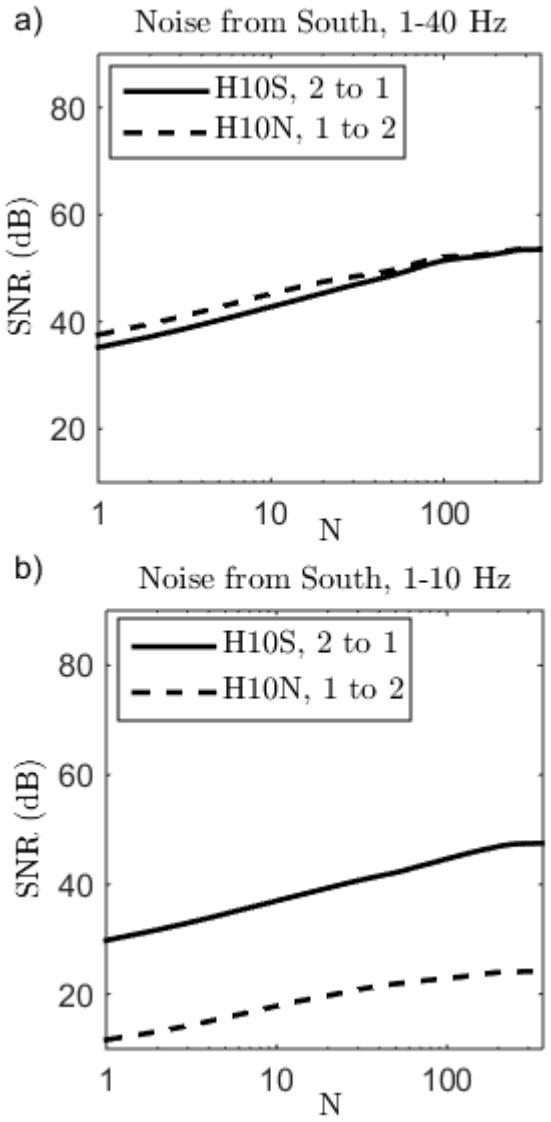

Figure 29: Average SNR in 2010 as a function of N averaging days for the hydrophones pairs oriented towards stable ice-generated noise, but in the unstable fields at H10N and H10S a) in the 1-40 Hz band, and b) in the 1-10 Hz band.

Hence, the results in this study, obtained for several hydroacoustic stations around the world, confirm that a good strategy to obtain high-SNR coherent SOFAR arrivals from low-frequency ambient noise correlations is to use hydrophone pairs aligned pointing towards the stable ice-noise sources (e.g. Antarctica). At most sites, this alignment is roughly along a North-

South direction. The exception to this strategy applies when it is beneficial to orient hydrophone pairs so that they point towards other areas potentially containing stable ice-noise or seismic sources (e.g., the Sea of Okhotsk or Japanese Trench), as used for the H11 station.

4.5 Predicting the Emergence Rate of the Coherent SOFAR Arrivals for the Long-Range Hydrophone Separations

Theoretically, the SNR of a coherent SOFAR arrival is expected to decrease with increasing range separation r between hydrophone pairs [11], [48], and, for 2-D propagation, is proportional to

$$SNR \propto \sqrt{\frac{N*B_\omega}{r*k_c}} \quad \textbf{(25)}$$

In the equation above, N is the number of averaging days, B_ω is the frequency bandwidth of the coherent SOFAR arrival (and thus of the coherent noise component generating this SOFAR arrival), and k_c is the wavenumber at the center frequency. Note that this equation assumes an isotropic distribution of noise sources. Furthermore, the dependency of the SNR as $1/\sqrt{r}$ is justified by the fact that the propagation of the coherent noise of interest here (originating from the polar regions or seismically active regions) is primarily two-dimensional as this propagating noise is guided within the SOFAR channel and dominated by mode 1 for the low frequency band used [27], [49], [50].

Hence, using the equation above, the SNR values displayed in Figure 28 and Figure 29 - obtained between existing hydrophone pairs of the IMS network for small separation

distances (r~2km)- can theoretically be used to predict SNR values between hypothetical hydrophone pairs aligned along nearly the same orientation but separated by much larger distances (e.g. r >100km). These SNR predictions could be used to estimate the feasibility of a long-range passive thermometry experiment. The goal of this section is to verify this hypothesis.

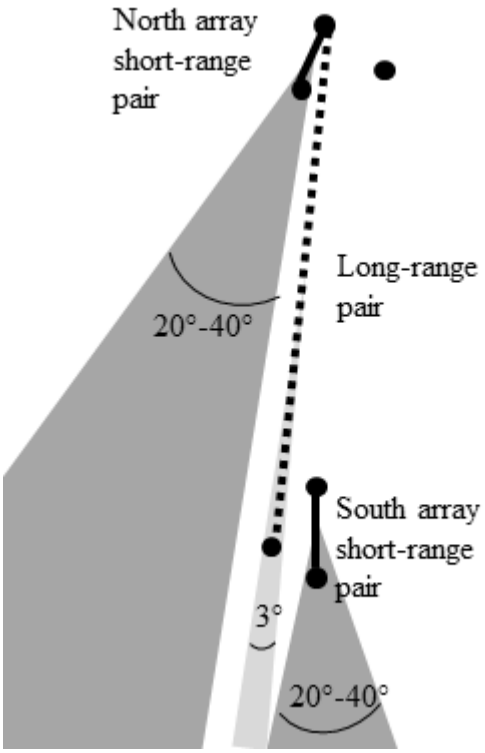

Figure 30: Schematic of the relative orientation of the broad endfire beams (angular width ~20°-40°) for hydrophone pairs separated by r~2 km (labeled by solid black line), and orientation of the narrower endfire beams (angular width 3°) for hydrophone pairs separated by r~130 km (labeled with the dashed black line). Note that for visualization purposes, this figure is not a to-scale depiction of the actual hydrophone geometries at H10 or H11 sites.

There is a line-of-sight path between the north and south triangular hydrophone arrays (or triads) for the H10 and H11 hydroacoustic stations. This configuration makes it possible to extract coherent SOFAR arrivals corresponding to coherent noise traveling between the north

and south arrays at each site. Between the arrays, the coherent noise propagates across a range of ~130 km along the SOFAR channel. However, a direct comparison of the SNR values for the coherent SOFAR arrivals between hydrophones separated by r~130 km (e.g. between hydrophones of the north and south arrays) and between hydrophones separated by r~2 km (e.g. between hydrophones of a single triangular array) is limited because the hydrophone pair axes are not exactly collinear (see Figure 30). Additionally, the beam widths of the main lobe, which are here primarily determined by the sensor separation distance, are not the same. Differing beam widths of the main lobe mean that different noise sources contribute to the long-range vs. short-range coherent SOFAR arrivals. Nonetheless, by selecting hydrophone pair combinations that are oriented north-south and have as much endfire beam overlap as possible, it is possible to compare long-range and short-range sensor pairs that at least share the same noise source type: in this case, consistent ice-generated noise emanating from the Antarctica region. The short-range (r~2km) hydrophone pairs used in this analysis are pair 2-3 at H10N array, pair 1-3 at H10S array, pair 1-2 at H11N array, and pair 1-2 at H11S array (see Figure 5). The long-range (r~130 km) hydrophone pairs at H10 and H11 -selected to have the closest possible alignment with the short-range hydrophone pair axes- are the pair South #2 – North #3 at H10 and the pair South #2 – North #2 at H11. Furthermore, in order to compare directly the SNR values for the long-range and short-range coherent SOFAR arrivals, the cross-correlation waveforms are all filtered in the same frequency band [3-14 Hz] where their frequency spectra overlap the most (see Figure 31 (a)-(b)). Assuming this band-pass filter operation yields similar frequency bandwidths for all correlations regardless of sensor separation distance (r), we can predict the following ratio between SNR values for the long-range (r~130 km) and short-range (r~2 km) cross-correlations.

$$\frac{SNR_{short}}{SNR_{long}} = \sqrt{\frac{r_{long}}{r_{short}}} \quad \textbf{(26)}$$

The long-range distance r_{long} between hydrophone pairs is respectively 132 km (for H11) and 127 km (for H10), and the short-range distance r_{short} between hydrophone pairs is ~2 km, thus yielding theoretical SNR ratios SNR_{short}/SNR_{long} ~8 . Figure 31 compares the measured SNR ratios for H10 and H11 (using either the South array or the North array as for the short-range correlations) to this predicted theoretical ratio for increasing averaging duration (*N=1,2,…200).* To improve statistical significance, SNR values obtained for a given number *N* of averaging days were further averaged over three years 2010-2012.

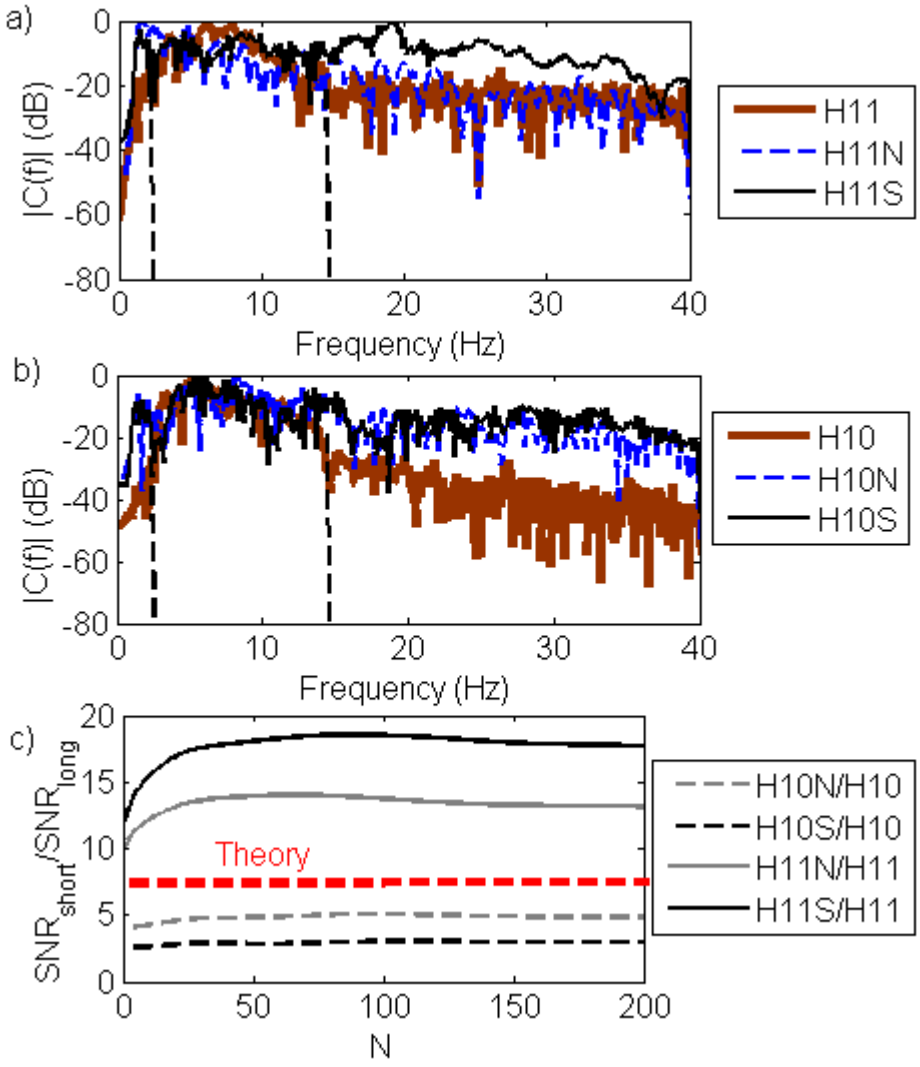

Figure 31: a) Amplitude spectrum of the time-gated cross-correlation waveform. Each waveform is averaged over one year and computed between hydrophone pairs all pointing in a similar direction towards ice-noise sources (see Figure 30). The cross-correlations were either computed between two hydrophones of the same North triad (H11N) or South triad (H11S) with a short separation distance of ~2km, or between one hydrophone of the north triad and one hydrophone of the south triad (H11) separated by a large distance of ~130km. b) Same as (a), but for the hydroacoustic station H10. c) Measured ratio between SNR values for the long-range and short-range cross-correlations filtered in the 3-14Hz band and computed for increasing number of averaging days *N*. Ratio values were further averaged over the years 2010-2012 for a given value of averaging days *N*. For comparison, the theoretical prediction of this ratio is indicated by a thick dashed line.

Overall, Figure 31 shows that the values for measured SNR ratios for H10 and H11 are within a factor of two of the theoretical ratio (i.e. a reasonable 3 dB difference). The differences between experimental and theoretical values are primarily due to 1) bandwidth differences and 2) and variations in hydrophone pair alignments. First, even though all cross-correlations were bandpass-filtered in the same frequency band 3-14 Hz, their spectra were not flat within that band. These non-flat spectra thus yield a different effective frequency band, based on the spectrum amplitude variations of the long-range and short-range correlations across frequencies. Second, the theoretical SNR ratio assumes that the noise field is isotropic (and therefore the same) for the long-range and short-range hydrophone pairs. However, in practice, the low-frequency deep ocean noise field is fairly directional and thus not isotropic as shown in the previous sections. Furthermore, since the hydrophone pair axes for the long-range and short-range pairs at each site (H10 or H11) are not completely aligned, different noise sources generate the coherent components of the noise detected at each hydrophone pair. Using a similar procedure to the ones used to generate Figure 23, back-projected endfire beams are shown in Figure 32 for H10 and H11 for the corresponding long-range and short-range hydrophone pairs. Based on the plane-wave beamforming procedure shown in Section 2.3, the beam widths for the short range hydrophone pairs were estimated to vary between 20°- 40° (depending on the spatially coherent noise spectrum at each site). The beam widths for the long-range hydrophone pairs were found to be ~3°. Figure 32 shows that the endfire beams for the long-range and short-range hydrophone intersect with different areas of the Antarctica region. Therefore, noise sources with different spatial origins and statistics contribute to the variations in amplitude of the coherent SOFAR arrivals obtained between the long-range and short-range hydrophone pairs.

Finally, Figure 32 shows that the SNR ratio is higher at H11 than it is at H10. Recall that H10 is in the southern Atlantic while H11 is in the northern Pacific, and both sites are likely detecting ice-noise coming from Antarctica. It is hypothesized that this difference is because H11 will pick up much less ice-generated noise coming from the south than H10 (due to distance from these sources), and that a broader beam width is more effective at detecting these distant sources.

Nevertheless, given these differences in beam widths and noise field variability at each site, the theoretical formula for SNR ratio is likely to provide a good approximation of the SNR as a function of increasing range for low-frequency ambient noise correlations when the dominant mechanism is ice-noise. Hence, results from Figure 31 validate the initial hypothesis that the SNR values -obtained between existing hydrophone pairs of the IMS network for small separation distances (r~2km) could theoretically be used to predict SNR values between hypothetical hydrophone pairs aligned along nearly the same orientation (north-south orientation here) but separated by much larger distances (e.g. r >100km), assuming the same bandwidth and noise source type as those in this study.

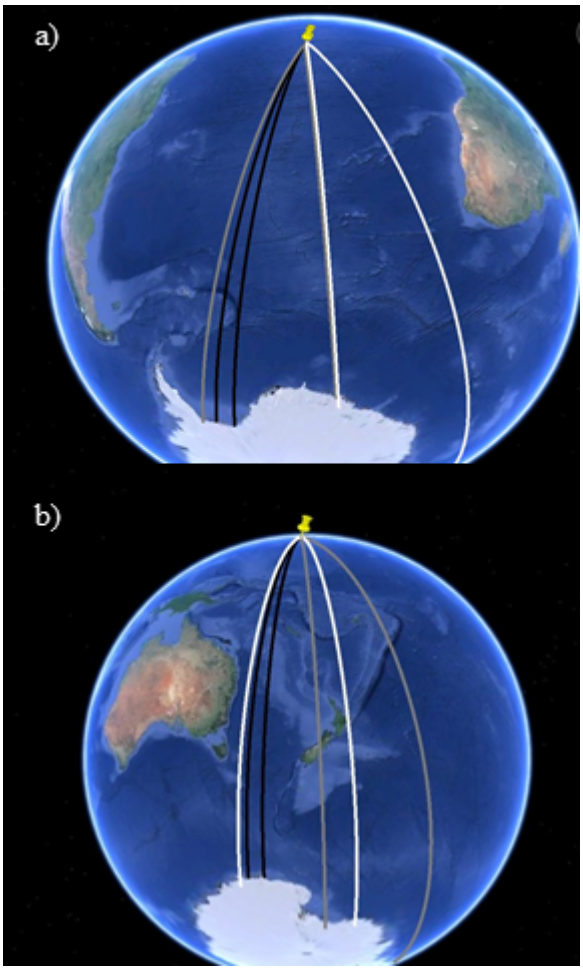

Figure 32: Comparison of the back-projected endfire beams for the same long-range hydrophone pair (delimited by black lines), short-range hydrophone pair for the North triad (delimited by white lines), and short-range hydrophone pair for the South triad (delimited by grey lines) used in Figure 31 for a) H10 and b) H11 hydroacoustic stations. (Maps from ©Google Maps)

4.6 Chapter Conclusions

This chapter investigated the frequency dependence, seasonal variability and emergence rate of the coherent components of the background low-frequency ambient noise (1-40 Hz) propagating between hydrophone pairs. These pairs were located at five IMS hydroacoustic stations distributed around the globe and were located in the SOFAR channel. The results

demonstrated how the variability of the ambient noise recorded by these five hydroacoustic stations affects the ability to extract coherent SOFAR arrivals - useful for passive acoustic thermometry purposes - in a robust and consistent fashion throughout the year. Ideally, hydrophone pairs should be oriented in-line with stable, loud, broadband sources (such as ice-noise or seismic noise). This favorable orientation minimizes the averaging time necessary to extract coherent SOFAR arrivals from ambient noise correlations. Minimizing the averaging time is important to obtain a high-SNR arrival as fast as possible to mitigate the effects of environmental fluctuations occurring during the averaging duration (i.e. integration time for the cross-correlation). Furthermore, even in a highly variable noise field (such as the South Atlantic), it was shown that hydrophone pairs could still be oriented with stable noise sources. This orientation is contingent upon the existence of an un-obstructed, direct, and geodesic propagation path between the ice noise and the sensors (i.e., in the absence of land blockages).

Additionally, it was shown that simple theoretical predictions can be used to estimate the SNR values for coherent arrivals obtained from hydrophone pairs separated by a long range, (>100 km) given ambient noise data collected from sensors separated by a shorter range (~2 km here). This prediction is possible as long as both the long-range and short-range sensor pairs have similar orientations, thus pointing towards distant noise sources with similar statistics and frequency bandwidth. Hence, the analysis of the emergence rate of coherent components of the low-frequency ambient noise performed at existing hydrophone pairs of the IMS network could potentially be used to assess the feasibility of low-frequency passive acoustic thermometry of the

deep ocean between hydrophone arrays oriented in-line with polar ice and separated by larger (possibly basin-scale) distances.

Chapter 5: Optimization of Averaging Duration when Tracking Arrivals

5.1 Introduction

The major limitation of the ambient noise correlation method is that the averaging duration of the cross-correlations must be smaller than the time scale of environmental fluctuations of the medium. For example, to track phase changes in a Green's function estimate that occur over a given time span, it is generally necessary to average the ambient noise recordings over a smaller time span. Phase fluctuations occurring on a time scale smaller than the averaging time applied to the cross-correlations become blurred.

This chapter proposes a method that uses genetic algorithm optimization [51] to track Green's function fluctuations (in this study, we examine only phase fluctuations) over a smaller time span than the averaging duration of the ambient noise recordings. The genetic algorithm is a stochastic search algorithm based on the principles of evolution. Section 5.2 explains the optimization method and demonstrates some simulated results. Section 5.3 parameterizes the accuracy of the optimization method, and Section 5.4 applies the optimization method to experimental deep ocean acoustics data.

5.2 Methodology and Numerical Simulation

As an illustration, the proposed optimization methodology for tracking arrival times fluctuations of a coherent arrival with low SNR is first applied in Figure 33(a) to a simple scenario consisting of a stack of N=10 noisy waveforms $Y_i(t)$ $(i=1..N)$:

$$Y_i(t) = \tilde{Y}(t - \hat{\tau}_i) + N_i(t) \quad , \tag{27}$$

Each $Y_i(t)$ waveform is composed of a time-shifted replica (by an shift $\hat{\tau}_i$) of the same baseline waveform $\tilde{Y}$(t) which is buried in additive bandlimited Gaussian noise $N_i(t)$. Here, the shifted waveform $\tilde{Y}(t-\hat{\tau}_i)$ can be thought to correspond to the coherent arrival of interest (e.g. the estimate of the Green's function arrival in free space between a pair or receivers) whose arrival time $\hat{\tau}_i$ rapidly varies across the N=10 epoch intervals. The peak-to-peak Signal-to-Noise Ratio (SNR) of each waveform $Y_i(t)$ is defined as:

$$SNR = \frac{\max(\tilde{Y}(t))}{3*std(N_i(t))} \tag{28}$$

where *std* refers to the standard deviation [8]. This chapter focuses on the case when the SNR for each individual waveform $Y_i(t)$ is less than 1, such that the time-delay ($\hat{\tau}_i$) across the N=10 successive epoch intervals are not trackable with a conventional peak amplitude detection algorithm.

For passive monitoring applications, a typical way to increase the SNR of the coherent arrival of interest is to using longer averaging durations when computing the ambient noise cross-correlation waveform. Doing so, the low-SNR cross-correlation waveforms $Y_i(t)$ $(i=1..N)$ computed over short durations (a single epoch interval here) would be simply summed to yield the long-time average waveform $\bar{Y}(t)$ (here across N epoch intervals):

$$\bar{Y}(t) = \sum_{i=1}^{N} Y_i(t) \tag{29}$$

However, because of the arrival-time fluctuations (given by the time shifts $\hat{\tau}_i$) this summation is sub-optimal and the resulting coherent arrival extracted from $\bar{Y}(t)$ is smeared out, thus with low amplitude (see Figure 33b).

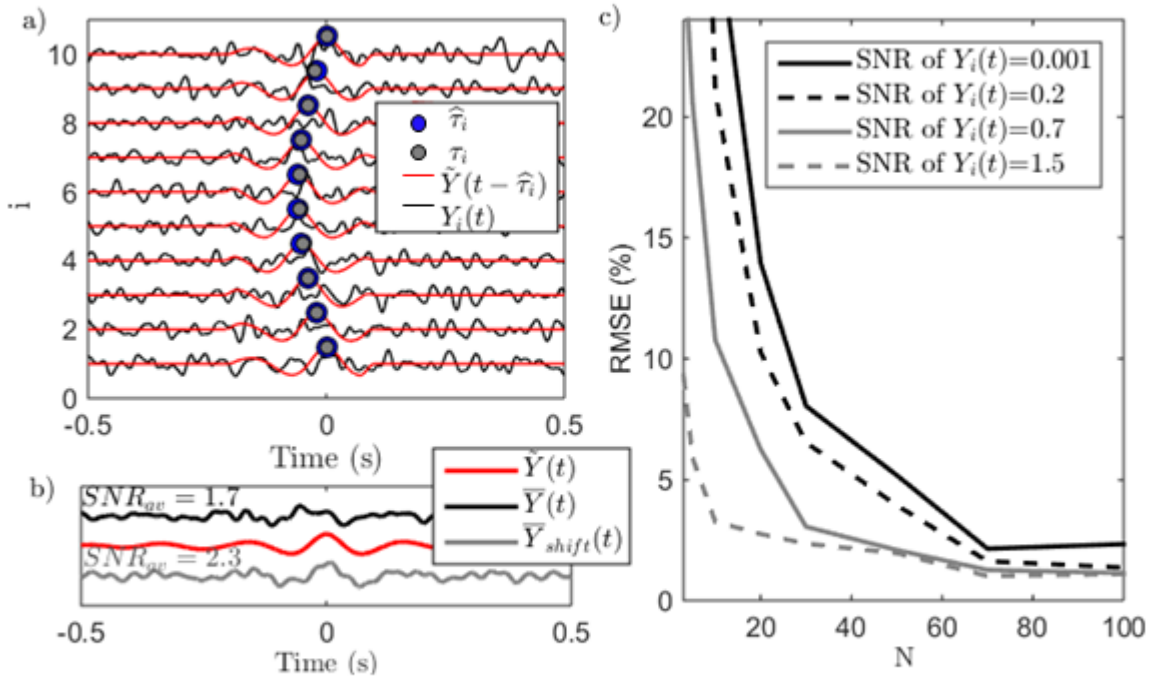

Figure 33: a) Stacking of the waveform $Y_i(t)$ over $N=10$ realizations, with a different time shift $\hat{\tau}_i$ applied to each realization. The SNR of each waveform is 0.2. The underlying shifted waveform $\tilde{Y}$(t-$\hat{\tau}_i$) with no added noise is also shown. Both waveforms are normalized. The time shifts τ_i are found by the genetic algorithm optimization. b) $\bar{Y}(t)$, corresponds to $Y_i(t)$ averaged over all i, as shown in (a). $\bar{Y}_{shift}(t)$ corresponds to the $Y_i(t)$ realizations shifted by τ_i and then averaged over all i. $\tilde{Y}(t)$ is the original, unshifted waveform without added noise. c) RMSE as a function of total number N of $Y_i(t)$ for various SNR of $Y_i(t)$.

However, because of the significant arrival-time fluctuations (given by the time shifts $\hat{\tau}_i$) this summation is sub-optimal and the resulting coherent arrival extracted from $\bar{Y}(t)$ is smeared out, thus with low amplitude (see Figure 33b). Instead, we propose to directly recover the time delay variations ($\hat{\tau}_i, i = 1..N$) using an optimization procedure relying on a stochastic search algorithm. Here we used the MATLAB™ genetic algorithm optimization function [52], [53] to

solve for set of time shifts τ_i (i=1..N) that maximize the following objective function $P(\tau_i, i = 1 \ldots N)$.

$$P(\tau_i, i = 1 \ldots N) = \left([\sum_{i=1}^{N} Y_i(t - \tau_i)] * \tilde{Y}(t) \right)\Big|_{t=0} \quad (30)$$

In Eq. (30), the star symbol denotes a cross-correlation operation, which acts as a matched filter between the noisy shifted waveforms $Y_i(t - \tau_i)$ and the baseline waveform $\tilde{Y}(t)$ which represents the a-priori known (or estimated) coherent arrival waveform. Note that the matched filter is evaluated at zero lag (t=0) since the shifts τ_i are deviations from a zero-lag peak arrival time for the reference waveform $\tilde{Y}(t)$. When the set of time shifts τ_i (i=1..N) obtained from this optimization procedure are closest to the actual time shifts $\hat{\tau}_i$, the objective function $P(\tau_i, i = 1 \ldots N)$ is maximized since the noisy shifted waveforms $Y_i(t - \tau_i)$ now add up coherently to provide an optimized-average correlation waveform $\bar{Y}_{shift}(t)$

$$\bar{Y}_{shift}(t) = \sum_{i=1}^{N} Y_i(t - \tau_i) \quad , \quad (31)$$

which has a higher SNR_{av} (see Eq. 33) than the smeared out conventional average $\bar{Y}(t)$ (Eq. (29)), as shown in Figure 33b. To quantify the error in the optimization results, the following Root-Mean-Square-Error (RMSE) metric is used and expressed as a percentage hereafter

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\hat{\tau}_i - \tau_i)^2}{N}} f_c \times 100\% \quad , \quad (32)$$

where f_c denotes the center frequency in Hz of $\tilde{Y}(t)$ (here f_c = 10 Hz) . Multiplying by this frequency term is equivalent to normalizing the RMSE by the period of the center frequency. The SNR of the shifted waveforms averaged over all i is quantified below, where here the "noise" portion of the SNR is taken over the interval $[t_1,t_2]$=[-3s,-1s] of the averaged waveform.

$$SNR_{av} = \frac{\max(\bar{Y}_{shift}(t))}{3*std(\bar{Y}_{shift}([t_1,t_2]))} \quad (33)$$

Figure 33 displays numerical results for this optimization procedure using a reference waveform $\tilde{Y}(t)$ (shown in Figure 33b) having a center frequency of 10 Hz and a bandwidth of 19 Hz. Here the bandwidth is defined by the -20 dB point. $\tilde{Y}(t)$ is derived by applying a 1 second-wide Tukey window to a 1 year time average of the experimental data presented later. The additive Gaussian noise $N_i(t)$ waveforms were filtered in the same frequency band as $\tilde{Y}(t)$ and their standard deviation was adjusted such that the SNR of the individual noisy $Y_i(t)$ ($i = 1..N = 10$) waveforms was SNR= 0.2. The arbitrarily selected time-shifts $\hat{\tau}_i$ are shown in Figure 33(a) (blue dots) with a maximum shift $\hat{\tau}_{max}$= 0.06s. Because of the very low SNR, the actual time-shifts $\hat{\tau}_i$ cannot be directly measured using conventional tracking algorithms in Figure 33(a). Applying the stochastic search algorithm yields a set of estimated N=10 time-shifts τ_i (shown in Figure 33(a) as grey dots) which closely match the actual time-shifts variations $\hat{\tau}_i$ with a RMSE of 1.6%. These optimization results are shown for an idealized case where there is no change (other than a phase shift) of the underlying reference waveform $\tilde{Y}(t)$ of interest for each of the *N* waveforms $Y_i(t)$. In the optimization algorithm, the search interval for the estimated time shift was set to $-0.065s \leq \tau_i \leq 0.065s$. Furthermore, because of the stochastic nature of this GA, other parameters were selected to minimize residual misfits in arrival times and optimize run times as listed in Table 1. Parameters not listed are kept at MATLAB's default values; a more detailed description of these GA parameters is given by Conn *et al.* [53].

Table 1: Genetic Algorithm Parameters

Genetic Algorithm Parameter	Function Tolerance	Population Size	Initial Population Range	Maximum Generations	Elite Count	Seed for random number generator
Value	1e-6	100	-1 ms to 1 ms	200	2	45

For a given reference waveform $\tilde{Y}(t)$, the matched filter operation in the objective function (i.e., the correlation operation in Eq. (30)) yields a higher output if the SNR_{av} of the optimized-average correlation waveform $\bar{Y}_{shift}(t)$ (Eq. 31) is high; this would in turn reduce the RMSE for the estimated time shifts. This SNR_{av} is expected to increase as $\sqrt{N}$, where N is the total number of waveforms being optimized, and increase proportionally with the mean SNR value of the individual waveforms $Y_i(t)$. To better quantify the accuracy of the optimization procedure using the GA, Figure 33(c) displays the evolution of the RMSE as a function of N and the SNR of the individual $Y_i(t)$. In Figure 33(c), the same reference waveform $\tilde{Y}(t)$ shown in Figure 33(b) (red trace) was used to construct the N waveforms $Y_i(t)$. The applied time shifts $\hat{\tau}_i$ were determined by interpolating the same half-sine-shaped delay-law (varying between 0 and $\hat{\tau}_{max}$= 0.06s) shifts shown in Figure 33(a) for increasing value of N = 3, 5, 10, 20, 30, 50, 70, and 100. To increase statistical significance of the results, the GA algorithm was run K times (with the random number generator being unseeded to be non-repeatable and truly random) for each value of N to obtain a constant number of generated waveforms NK=100. This product was kept constant so that the RMSE was found over a constant number of data points (time shifts). For instance, for N=5, the GA algorithm was ran K=20 times and the resulting set of estimate

time-shifts τ_i (for $i=1,2...N=5$) were each averaged over the K =20 runs. As expected, the algorithm is better able to predict the correct set time shifts (i.e. yield a smaller RMSE) when either the SNR of individual of waveforms or the total number N of waveforms increases, although computational complexity is rapidly increased for large values of N. It should be noted that since the genetic algorithm is stochastic in nature, the lines in Figure 33(c) represent average values of the overall trend: the outcome for a single run of the algorithm may give better or worse results.

5.4 Experimental Results

In this section, the GA optimization technique is used to track arrival-time fluctuations of coherent arrivals extracted from cross-correlations of low-frequency (1Hz<f<40Hz) deep water noise. This noise is dominated by ice-noise originating from the Antarctica region and propagates nearly horizontally along the SOFAR channel between two horizontal triangular arrays separated by ~132km. These arrays compose the International Monitoring System (IMS) hydroacoustic station H11 near Wake Island in the Pacific Ocean. The center frequency of this coherent SOFAR arrival was ~10Hz with a -20 dB bandwidth of 19 Hz. The experimental set-up and data processing are described in more detail by Woolfe *et al.* [27]. This data set is used as a proof-of-concept to illustrate the applicability of the proposed optimization technique to potentially enhance the resolvable time scales of the measurable sound speed fluctuations for passive thermometry purposes, specifically by extracting coherent arrivals from noise correlations computed using a minimal amount of averaging time. Recordings of ambient noise (propagating northward) made on the two triangular hydrophone arrays of the H11

hydroacoustic station were coherently processed to obtain a single beamformed correlation waveform $Y_i(t)$ for each of the 40 consecutive Julian days $i=181,182,\ldots220$, of the year 2010. The average SNR of these 40 daily waveforms is SNR ≈2.2 (the time interval $[t_1, t_2]$=[-3 s, -1.5 s] was used to estimate the standard deviation in the denominator of Eq. (33)). This particular 40 days interval was arbitrarily chosen because it represented a challenging time interval for passive ocean monitoring purposes, as the peak arrival time of the coherent SOFAR arrival (i.e. true time shifts $\hat{\tau}_i$ indicated by blue dots on Fig. 34a) varied rapidly during this interval with an overall shift of ~50 ms. This shift constitutes half of a period at the center frequency of 10Hz for this SOFAR arrival.

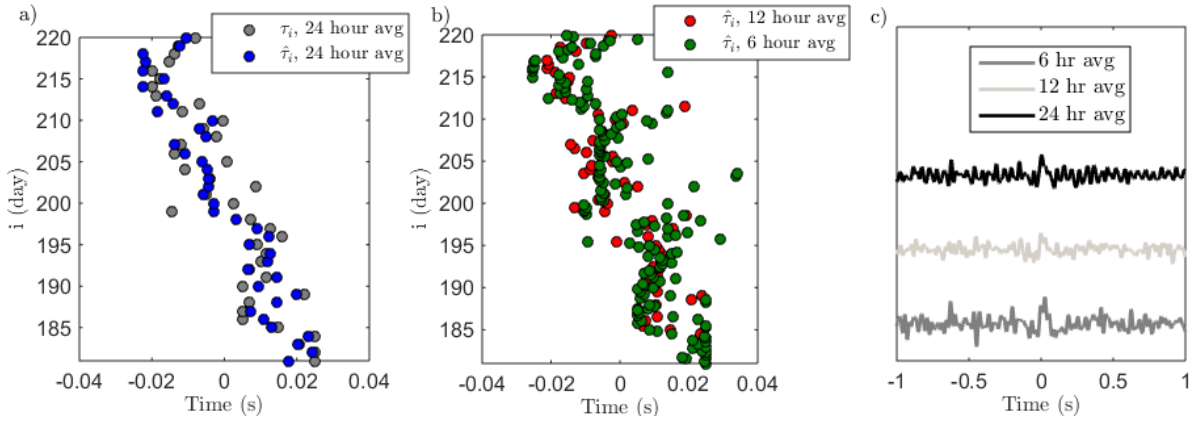

Figure 34: a) Optimization results for the experimental data, $Y_i(t)$, where i= day 181,181…220 in the year 2010. In this case the time shifts $\hat{\tau}_i$ are the true peaks of each day-averaged waveform, not artificially applied shifts. b) Optimization results for when $Y_i(t)$ is averaged over 12 and 6 hours rather than one day as shown in (a). c) Typical $Y_i(t)$ waveforms for each averaging duration case.

In the optimization algorithm, the baseline reference waveform $\tilde{Y}(t)$ was obtained by averaging $Y_i(t)$ over $i=1,2,\ldots365$, and then applying a 2.4 s wide Tukey window centered at the main peak of the averaged waveform at $t=0$. A key difference from the numerical simulations

presented earlier is that the 24h-averaged correlation waveforms $Y_i(t)$ exhibit some distortions compared to the reference waveform $\tilde{Y}(t)$, likely due to daily fluctuations of the ambient noise spectral content and characteristics. There is an average correlation coefficient of 0.9 between each $Y_i(t)$ and $\tilde{Y}(t)$, which is expected to reduce the output of the matched-filter used in the objective function. First, Figure 34a shows that the estimated time shifts τ_i (grey dots) obtained from the GA optimization algorithm can correctly match (RMSE=12%) the actual time shifts $\hat{\tau}_i$ (blue dots) obtained from a conventional peak-amplitude detection procedure applied to the 40 correlation waveform $Y_i(t)$. This result was expected since each waveform had a relatively high SNR (mean value of SNR~2.2), as discussed in Figure 33c; indeed this first quality check was only done to ensure that a correct parametrization was used for the GA search algorithm. To obtain this result, the GA search algorithm was applied to four blocks of 10 consecutive days at a time (i.e. $N=10$) over the 40 day interval. For the first 10 day block (i.e. $i=[181,190]$), the search domain for the estimated time-shifts τ_i was constrained to be $0.005s \leq \tau_i \leq 0.025s$ (i.e., only 50% of the total domain) to provide a good starting point for the search algorithm. For the subsequent three 10 day blocks, the search interval for the time shifts τ_i was constrained more loosely to be within ± 20 ms (i.e. covering 80% of the total search domain) of the last estimated time-shift τ_i in the previous 10 day period. The SNR_{av} of the conventional average $\bar{Y}(t)$ of the 40 unaligned $Y_i(t)$ waveforms is 4.8. The SNR_{av} of optimized-average correlation waveform $\bar{Y}_{shift}(t)$ increases to 6.3, giving a 23% increase in SNR_{av} provided by the optimization procedure.

In Fig 34b, the GA optimization procedure was then implemented to track arrival-time fluctuations of the coherent SOFAR arrival obtained from ambient noise correlations computed over smaller time spans (i.e. here 6h and 12h averaging time instead of 24h used in Fig 34a).

These time spans are smaller than the averaging time it would take to extract a coherent SOFAR arrival with a sufficiently high SNR. In other words, those time spans of 6h and 12h are too short to use a simple peak detection algorithm (as was done to obtain the blue dots in Fig 34a) to reliably estimate the peak amplitude of the coherent SOFAR arrival. The same search bounds and other optimization parameters used in Fig. 34(a) were used to estimate the time-shifts τ_i, using the same 10 days block at a time. Therefore, the same optimization process was used to estimate the time-shift τ_i over *N*=20 (resp. N=40) correlation waveforms when using a 12 (resp. 6) hour-averaged correlation waveforms, as shown in Fig. 34(b). Overall, the trend of the estimated time shifts in Fig 34b is consistent with the trend shown in Fig. 34a, thus potentially indicating that this GA optimization procedure yields genuine fluctuations of the arrival times of the coherent SOFAR arrivals over these shorter durations of 12h (red dots) and 6h (green dots). The SNR_{av} of the optimized-average correlation waveform $\bar{Y}_{shift}(t)$ (Eq. (5)) for the 12h (resp. 6h) average is 5.9 (resp. 6.1). The SNR_{av} for both cases is slightly less than the 6.3 obtained for 24h average, indicating that the optimization works less well for lower averaging times (i.e., lower SNR of $Y_i(t)$). However, since no independent ground truth was available for the time-shifts $\hat{\tau}_i$ over these very short averaging durations, the RMSE could not be computed. Furthermore, as the averaging time decreases from 12h to 6h, it can be observed that larger deviations are obtained when estimating the time-shifts τ_i. This may be indicative of nonstationarity of the noise field and/or transient anisotropy in the spatial distribution of the noise sources, for instance caused by the presence of loud transient interferers (e.g. shipping events), which briefly bias the estimated arrival times. Or alternatively, these deviations may occur as the SNR of each of the N individual correlation waveforms $Y_i(t)$ is getting too small

(i.e. SNR<1 for the several of these waveforms) to allow the GA algorithm to converge towards the optimal solution (as discussed for Fig. 33c). Indeed, the SNR of the coherent SOFAR arrival theoretically grows as $\sqrt{T}$, where T is the ambient noise recording (i.e. averaging) duration [11]. Hence since the mean SNR values for the correlation waveforms $Y_i(t)$ averaged over 24h is SNR~2.2, the expected mean SNR values for the correlation waveforms $Y_i(t)$ averaged over 12h (resp. 6h) should be SNR~2.2/$\sqrt{2}$=1.5 (resp. SNR~2.2/$\sqrt{4}$=1.1). Again the reader should note that the actual SNR of several of these 12h-averaged or 6h-averaged waveforms was smaller than 1 and thus the mean SNR value could not be directly computed.

Consequently, in order to quantify the RMSE of this optimization procedure for estimating the time shifts τ_i for the low-SNR cases (SNR <1), band-limited noise (1Hz<f<40Hz) with increasing standard deviation was artificially added to the 40 correlation waveforms averaged over 24h. This is akin to the simulation approach used to generate Fig 33c, with the difference being that Fig. 33c shows the simulation results when noise is added to the shifted reference waveform $\tilde{Y}(t)$ while Fig. 35a shows the results for adding noise to the day-averaged $Y_i(t)$. Fig. 35a shows the estimated time shifts τ_i –using the same optimization procedure and settings used for Fig. 34a- when the SNR of the 40 correlation waveforms was reduced to 0.7.

Overall, the trend of the estimated time shifts in Fig. 35a for this low SNR case (SNR=0.7) is consistent with the trend shown in Fig. 34a for the higher SNR case (SNR=2.2), with an RMSE of 16%. This potentially indicates that this GA optimization procedure can be used to estimate the arrival time fluctuations of the coherent SOFAR arrival for these relatively low SNR cases (or equivalently for the shorter averaging duration of 12h and 6h previously used). Using a similar approach, Fig. 35b displays the variations of the RMSE for increasing

values of SNR, showing that an averaging duration down to 1.8 hours (corresponding to an SNR of 0.6) could potentially be used while maintaining a RMSE of 20% or less. Furthermore, the RMSE values plotted in Fig. 33c are lower than the RMSE values shown in Fig. 35c for similar values of SNR and number of waveforms N=40. This is likely due to the fact the measured ambient noise field is nonstationary because it is frequently dominated by transient interferers especially when using very short averaging durations. Another potential contributing factor to this shift in RMSE is that the coherent SOFAR arrival present in each individual correlation waveforms $Y_i(t)$ is not an exact time-shifted replica of the reference waveform reference waveform $\tilde{Y}(t)$ (as assumed in the simulations plotted in Fig. 33c).

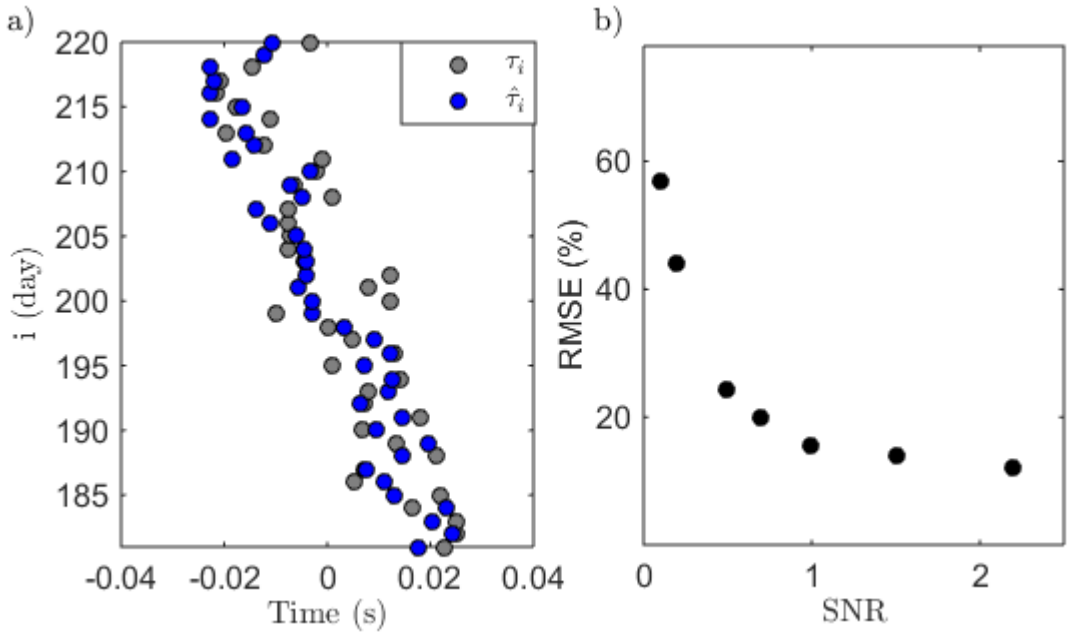

Figure 35: a) Optimization results adding bandlimited noise for an SNR of 0.7 to each $Y_i(t)$ obtained from averaging the experimental data over 24 hours. b) RMSE of the optimization shown in (a), but for varying SNR.

5.6 Chapter Conclusions

This letter provided a proof-of-concept that an optimization algorithm can be used to track rapid arrival time variations of Green's function estimated from noise correlations that occur over time periods smaller than the conventional averaging duration that would be necessary to extract a high-SNR Green's function estimate. To do so, this tracking problem was recast as an optimization problem using a stochastic search algorithm and a suitable objective function based on a-priori knowledge of the baseline-or reference- Green's function estimate (e.g. which can be estimated from a conventional long-time averaging process). Because the stochastic nature of the genetic algorithm used in this study does not guarantee convergence to the exact solution (i.e. here the actual time-shifts of interest), other more robust optimization techniques could be applied to maximize the proposed objective function in future work.

Chapter 6: Conclusions

6.1 Summary

This work demonstrated the use of passive acoustic tomography in a deep ocean environment. Specifically, Chapter 3 showed how this method could be successfully used to track deep ocean temperatures - with an unprecedented degree of temporal resolution and precision – over almost one decade. Chapter 4 developed the recommendation that future monitoring sensor pairs should be aligned with sea-ice or seismic noise sources to take advantage of the nearly-constant low-frequency sound sources. Chapter 5 proposed a method to track acoustic medium fluctuations that occur on a time scale shorter than the averaging time necessary to extract the useful coherent components of the ambient noise.

6.2 Contributions to the Literature

For completeness, the following contributions are repeated. These are the major contributions of this work to the scientific literature. This list is meant to be brief; precise contributions can be found in the cited papers. The major contributions are listed below. This thesis culminated in several journal papers and conference presentations. The work may be extended to future endeavors, outlined in the next section.

Major contributions:

1. This is the first deep ocean passive acoustic tomography study. Previous work has been limited to shallow water, where the acoustic propagation physics and frequency range are very different.

2. Previous studies in shallow water have collected coherent arrivals over time periods of up to a month. This work examines up to 9 continuous years of data.

3. Previous work in passive acoustic thermometry has stopped at extracting the Green's function. This work completes the inverse problem, converting from acoustic travel times to average temperature shifts along the acoustic travel path. The resultant temperature shifts are compared with independent direct temperature measurements.

4. Although there have been many studies on noise coherence in the ocean, this research presents the first comparison of the coherence of multiple sites around the world, with an emphasis on the effects of noise coherence on passive acoustic tomography.

5. Previous studies have discussed one of the major limitations of using noise correlations for passive monitoring: that it is difficult or impossible to monitor events that occur on a time scale that is smaller than the emergence time of the coherent components of the

ambient noise. This work presents a method to overcome this limitation, enabling the potential to monitor events that occur on a shorter time scale than the averaging time.

6.3 Future Work

One possible outcome of this work is the development of a worldwide network of autonomous hydrophone arrays to conduct passive acoustic tomography. A major item that must be investigated before this happens is the maximum range (obviously greater than the 130 km used in this study) over which trackable arrivals can be extracted, particularly when sensors are oriented in-line with consistent, loud noise sources. Over these long ranges, long averaging times will be necessary to extract the coherent waveforms. These averaging times may be large enough to require the optimization method presented in Chapter 5 in order to track seasonal temperature fluctuations if the sensor separation is too large to quickly obtain a high-SNR coherent arrival. If the optimization method is required, then future work should include finding a suitable replacement for the genetic algorithm component of the optimization (but keeping the overall objective function the same).

Appendix A: Selection of References for Adaptive Beamforming

Selection of the best reference correlations is crucial when applying adaptive beamforming. The reference correlations should represent the best or most stable cross correlations. For this reason, reference correlations are taken from long time averages of the cross correlations. The "best" reference correlations are those which generate the highest SNR in the beamformer output with no big abrupt jumps in the beamformer output peak arrivals. As long as there are no transient sound sources (i.e. whales or submarines) on the SOFAR axis transmitting sound in the frequency spectrum of interest (1 – 40 Hz) and the ambient noise remains diffuse and at a relatively constant level, it is possible to apply one set of reference correlations to all the years of data. Otherwise, some sort of moving reference would have to be used.

The same general process was followed in determining the appropriate references for the Wake and Ascension Island sites:

1. The cross correlations for each year were examined. The goal was to find a long time average of cross correlations that had a high SNR. It was found that in general, averaging the cross correlations over an entire year worked the best. Ascension cross correlations were averaged over 2006, and Wake cross correlations were averaged over 2010.
2. Anomalies such as off-axis noise events from transient sources were discarded by skipping the days containing those anomalies in the averaging.
3. The peaks of the time-averaged cross correlations were carefully selected by hand. In the Wake Island case it was useful to apply a +-0.5 s sharp Tukey window around the

peaks to zero out additional noise. The Ascension cross correlations already had a high enough SNR that applying a window around the main peak had no effect.

Figure 36 shows the positive and negative reference correlations for the Wake Island site. It is interesting to note that the positive references are more narrowband than the negative references.

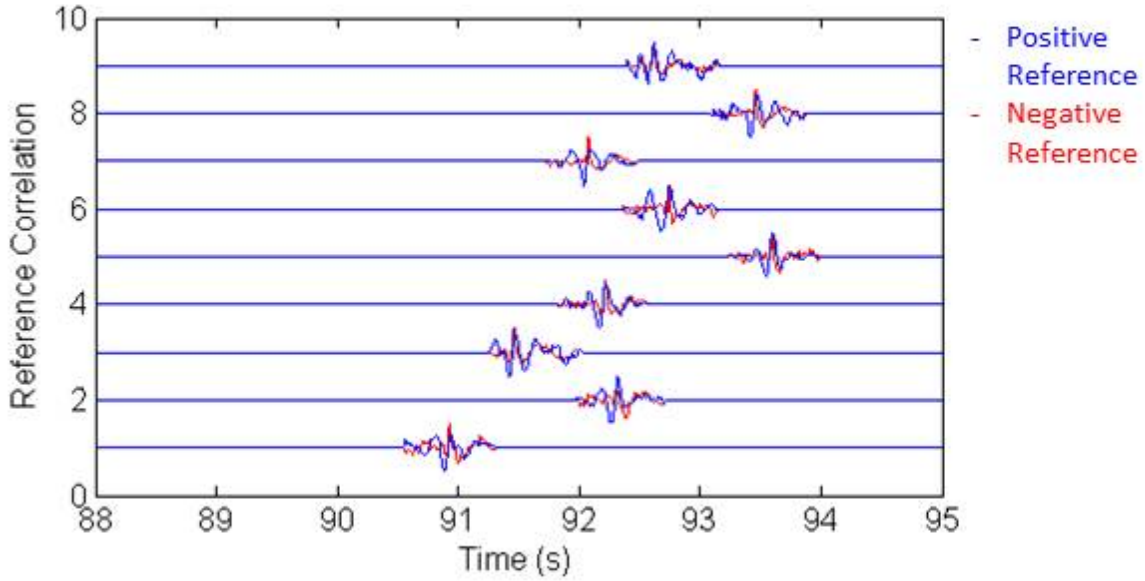

Figure 36: Positive and negative reference correlations for the Wake Island site.

When the reference is more narrowband, the first singular values in the frequency domain of the reference cross-correlation matrix will form a narrowband filter if the data is multiplied by them. Ultimately, it was decided not to use this filter because it made the peak arrivals wider in the time domain, resulting in a lower SNR and less precision in the beamformed data.

Appendix B: Ice-generated Noise

The main lobe width (defined as 6dB from the peak) of the plane wave beamformed H10 cross-correlations can be used to determine the direction of the source of the ambient noise that is used to compute the cross-correlations. The projection of the main lobe width upon the globe is shown in Figure 37. This projection shows that the sound moving from south to north that is picked up by the array comes from anywhere within the boundaries of this projection, shown by the dashed lines. This projection raises the following question: is the sound that travels in the SOFAR channel and is captured in the cross-correlation between the north and south arrays primarily seismic or caused by ice in Antarctica. There are ample reasons for either sound source, and indeed the sound may come from a combination of the two sources. The hydrophones are located on the mid-Atlantic ridge, which is a hotbed of constant seismic activity. Additionally, the center of the main lobe projection passes right through a volcanic archipelago of islands that connect South America with Antarctica. Therefore, it is impossible to rule out seismic sound sources.

Figure 37: The projection of the main lobe of the plane wave beamformer upon the map. The green line is the center of the main lobe, and the dashed lines show the width of the lobe.

However, some insight into ambient sound caused by ice tremors may be gleaned from comparing the average SNR of the beamformer output for each month of the year with the ice extent in the zone enclosed by the plane wave beamformer main lobe width, displayed in Figure 38. Images of the sea ice extent around Antarctica were obtained from the National Snow and Ice Data Center (NSIDC). Figure 38 displays the ice extent for a random day in 2006, where the pink line is the median ice boundary for that day. The surface area in km^2 covered by ice was calculated by calculating the number of white pixels in the red box for each day from January 1, 2006 – December 31, 2012. Each pixel corresponds to 25km x 25km.

Figure 38: This figure shows the area of interest in Antarctica where the ice coverage is calculated. Data is obtained from the National Snow and Ice Data Center [54].

Figure 39 illustrates the comparison between ice extent and average SNR per month, averaged over the years 2006-2012. There is a distinct drop in SNR throughout the colder months where there is more sea ice coverage. It is hypothesized that this is because the warmer months enable sea ice to move with more freedom, enabling icebergs to scrape each other rather than hold in place. It is important to note that the geographical region of interest is named the Weddell Sea, which commonly serves as an iceberg "parking lot" where icebergs remain trapped by the Weddell Gyre current and the Antarctic peninsula. It is likely that iceberg motion is a

source of the noise that is detected by the Ascension Island Site. At any rate, there is a definite relationship between sea ice coverage and SNR.

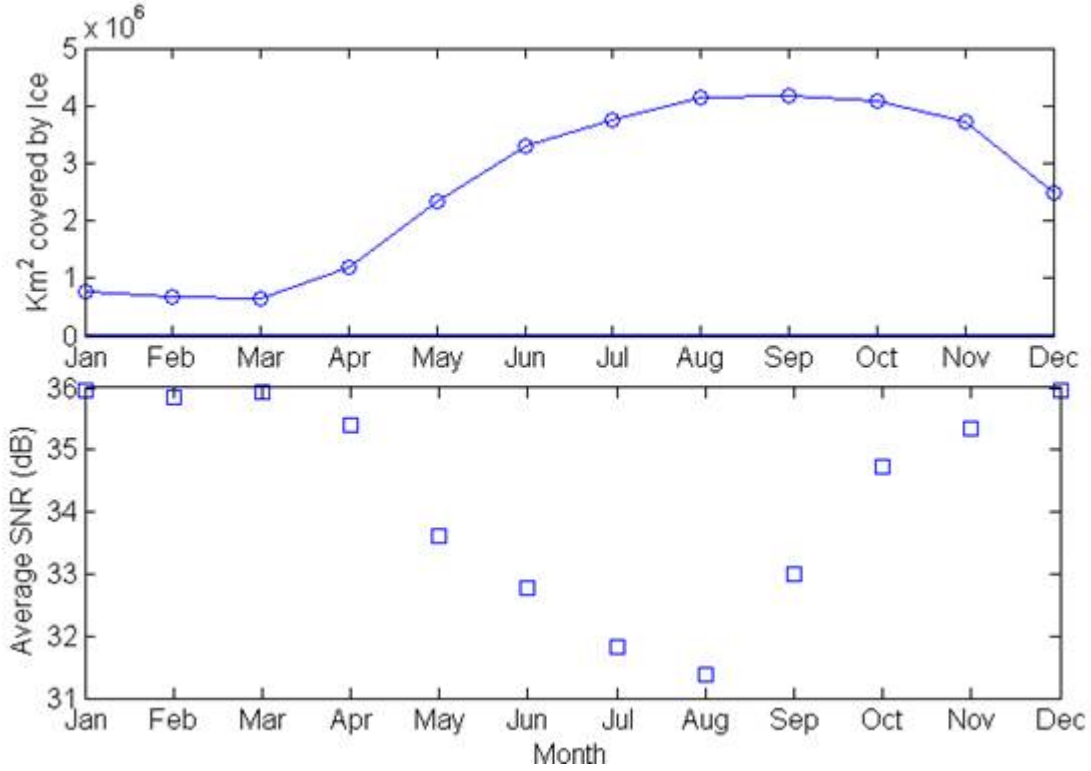

Figure 39: Comparison between the average SNR of the beamformed cross-correlations for the Ascension Island arrays for each month with the surface area covered by ice within the red box shown in Figure 38.

The same analysis can be repeated at the Wake Island site for noise coming from both the Arctic and Antarctic regions. As shown in Figure 40, there is a clear seasonal pattern of the *SNR* of the beamformed cross-correlations (denoted as SNR_B) at both sites, meaning that the acoustic sources that create the coherent noise that propagates between the north and south triad array have a seasonal dependence. SNR_B average values for each month are found by finding the average beamformer output for each month over 3 years at Wake Island and over 7 years at Ascension Island. Given the low frequency band (1-40 Hz), the receiver positions in the SOFAR

channel, and the spatial location of the potential noise source, it is likely that a dominant portion of this noise is generated by complex seasonal ice cracking and/or shifting events in the Polar regions [41], [46], [55].

The SNR_B of northward-propagating noise detected at the Wake Island site shows a similar trend, albeit with weaker variability and a slightly different seasonal dependence compared to the Ascension Island site. These differences may be due to unknown geographic factors and local variations of the noise generation mechanism. It should be noted that the absolute values of SNR_B are always higher for the Ascension Island site than for the Wake Island site. This is likely because the Ascension triads have a direct line of sight with the Antarctica coastline, while land blockages are more pronounced between the Wake Island site and the Polar regions. Furthermore, the number of icebergs –which determines the ice noise source density and thus the amount of coherent noise propagating along the SOFAR channel up to the hydroacoustic stations- is typically higher in the South Atlantic than in the South Pacific [41].

Because there is no land mass near the Wake Island site blocking noise coming from the North pole (contrary to the Ascension island site), it is possible to compare coherent noise levels coming from the Arctic area with noise coming from the Antarctica area in Figure 40. Notably, Figure 40 indicates that the SNR_B of northward-propagating noise detected at the Wake Island site drops as the Arctic ice becomes more dynamic and melts during the warmer months (June - September). It is hypothesized here that these differences between the coherent noise components generated in the Arctic vs. Antarctica could be driven by a complex interaction of weather patterns, lack of a polar landmass, and other factors. For instance, in the Arctic, the absence of a continental land mass means that the ice noise generated can efficiently be

generated by the whole extent of the ice mass. This implies that ice noise may couple into the water column over the whole Arctic area, especially during the winter months when the sea ice extent and thickness are maximal. This in turn causes more radiating pressure waves in the ice, resulting in more intense ocean ambient noise generation. On the other hand, in the Antarctic, the land continent blocks sound from ice lodged on land. Thus noise is primarily generated by ice scraping against the coastline in addition to the noise generation mechanisms such as iceberg cracking.

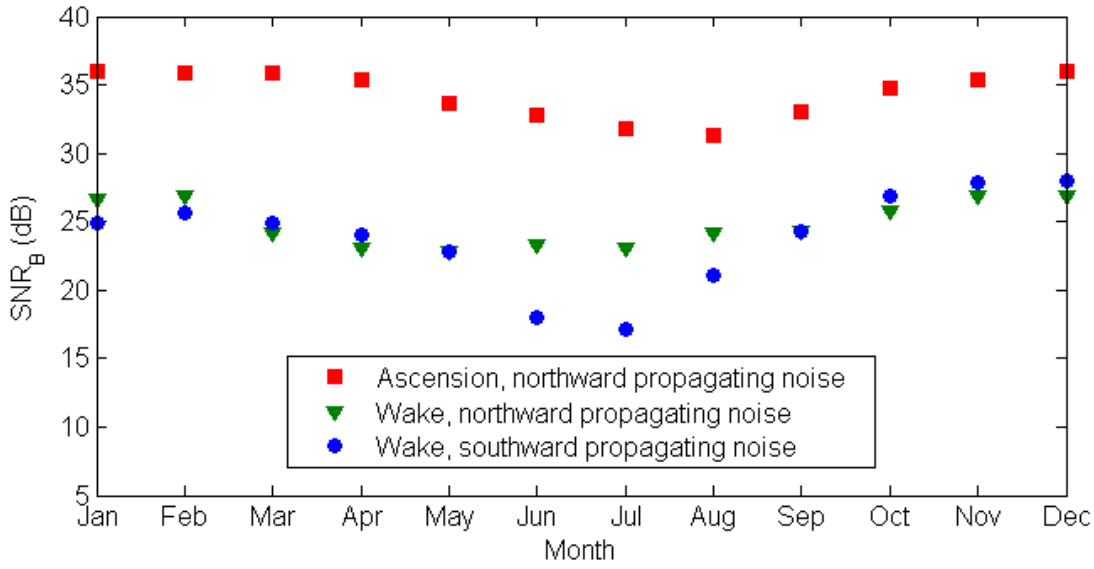

Figure 40: Comparison of the average monthly values of SNR_B (dB) for the beamformed coherent arrivals at the Wake and Ascension Island sites.

Appendix C: Arrival Time Error

It is necessary to establish confidence in the precision of the arrival time measurements of passive acoustic tomography. In the ray approximation, the received signal is the sum of delayed replicas of the transmitted signal. Matched-filter processing is the optimum procedure for estimating the arrival time of a single, resolved arrival embedded in Gaussian noise [56]. We use the time at which the envelope is a maximum and assume that the signals are merely delayed replicas of each other and that phase is proportional to travel time. This assumption means that the propagation is low-frequency over a relatively short range, and that internal scattering is negligible. In this case, the rms error of the travel time estimate is given by

$$\sigma_t = \left[\omega_0 \sqrt{2E/N}\right]^{-1} , \tag{34}$$

where ω_0 is the center frequency of the signal and 2E/N is the SNR obtained with a matched filter:

$$\frac{2E}{N} = \frac{[y_s(\tau)]^2}{<y_n^2>} \tag{35}$$

In this case, $y_s(\tau)$ is the peak of the signal and $< y_n^2 >$ is equal to 9 times the variance of the noise (i.e. peak noise level).This formula is applicable to any arrival that meets the previously mentioned conditions, and therefore works for the beamformed correlations.

Appendix D: Hydrophone positions

Station	Hydrophone	Latitude	Longitude	Depth (m)
H01 (Cape Leeuwin)	W1	-34.89300	114.15400	1063
H01 (Cape Leeuwin)	W2	-34.89850	114.13380	1046
H01 (Cape Leeuwin)	W3	-34.88320	114.13610	1056
H03 (Juan Fernandez)	N1	-33.44120	-78.91200	757
H03 (Juan Fernandez)	N2	-33.43990	-78.93240	745
H03 (Juan Fernandez)	N3	-33.45560	-78.92400	761
H08 (Diego Garcia)	N1	-6.34210	71.01430	1248
H08 (Diego Garcia)	N2	-6.32510	71.00010	1243
H08 (Diego Garcia)	N3	-6.34540	70.99190	1182
H08 (Diego Garcia)	S1	-7.64530	72.47440	1413
H08 (Diego Garcia)	S2	-7.64530	72.49330	1356
H08 (Diego Garcia)	S3	-7.62750	72.48380	1359
H10 (Ascension)	N1	-7.84570	-14.48020	847
H10 (Ascension)	N2	-7.82780	-14.48750	845
H10 (Ascension)	N3	-7.84090	-14.50170	850
H10 (Ascension)	S1	-8.94120	-14.64840	865
H10 (Ascension)	S2	-8.95910	-14.64530	852
H10 (Ascension)	S3	-8.95270	-14.66290	863
H11 (Wake)	N1	19.71360	166.89110	731
H11 (Wake)	N2	19.73110	166.89680	721
H11 (Wake)	N3	19.71790	166.90990	729
H11 (Wake)	S1	18.50830	166.70030	750
H11 (Wake)	S2	18.49050	166.70540	742
H11 (Wake)	S3	18.49490	166.68730	726

Appendix E: Effective Depth of the SOFAR Waveguide

Given the sensor separation and north-south orientation between the two triads at H10 and H11, noise detected at both sites largely emanates from the Polar Regions and propagates in the SOFAR waveguide up to the mid-latitude regions where the hydroacoustic stations are located. Because the noise propagates over such long ranges and has a low center frequency of around 10 Hz, most of the acoustic energy trapped within the SOFAR channel reaching the hydroacoustic stations is primarily carried via the first (i.e. lowest) normal mode of the ocean water column [6]. The depth of the SOFAR waveguide is approximated here as containing 95% of the energy in the first propagating mode at the center frequency of 10 Hz, which determines the effective propagation velocity (i.e. group velocity) of the coherent SOFAR arrivals used in this study [57]. Specifically, this approximation is derived from the following relationship,

$$\frac{1}{c_g} = \frac{\omega}{\rho k_r} \int_0^D \frac{\Psi^2(z)}{c^2(z)} dz \tag{36}$$

where c_g is the modal group velocity, ω is the radial frequency, D is the local ocean depth at the receiver, ρ is the density of the seawater (here assumed to be constant through depth), and $c(z)$ is the sound speed profile. k_r is the modal wave number and $\Psi(z)$ is the normalized mode amplitude for mode 1 over the ocean depth. The limits of the integral $z1$ and $z2$ are adjusted to find the depth integration limits (symmetric about the SOFAR channel axis) that contribute to 95% of the total integral, as shown below.

$$\int_{z1}^{z2} \frac{\Psi^2(z)}{c^2(z)} dz = 0.95 \int_0^D \frac{\Psi^2(z)}{c^2(z)} dz \tag{37}$$

The normalized mode shape $\Psi(z)$ is found at 10 Hz using a sound speed profile $c(z)$ obtained from NOAA 2009 World Ocean Atlas [58], [59] temperature and salinity data input to

the Mackenzie sound speed equation [29]. The mode shape is found using the Kraken normal mode code package. The sound speed profile, normalized mode shape, and normalized quantity $\frac{\Psi^2(z)}{c^2(z)}$ with *z1* and *z2* are shown in Figure 41 for both H10 and H11.

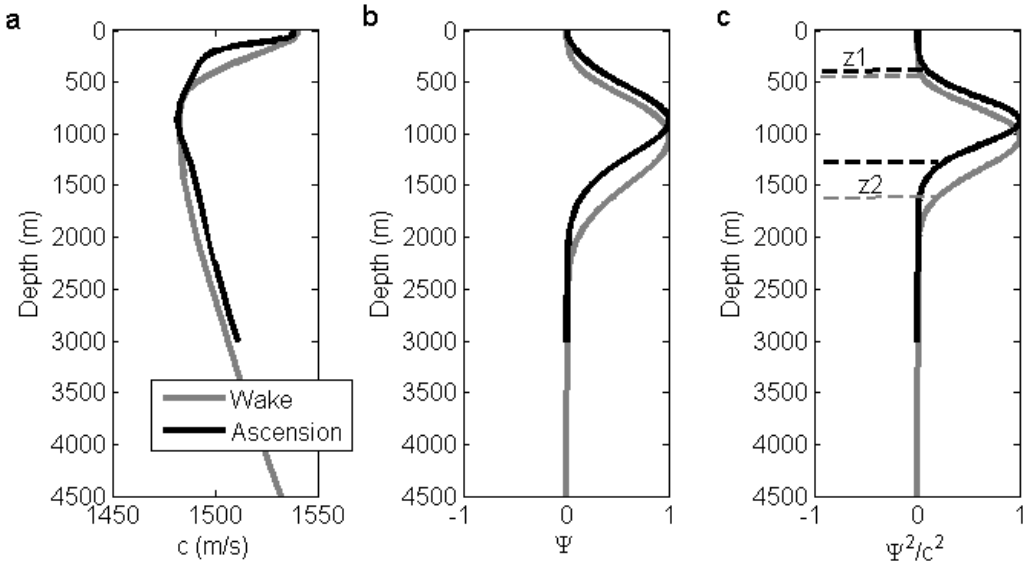

Figure 41: (a) Sound speed profile c(z), (b) normalized for mode 1 mode shape Ψ(z), and (c) integrand $\frac{\Psi^2(z)}{c^2(z)}$ at the Wake and Ascension Island sites (see Eq.(S8)). The depths z1 and z2 determining the effective depths of the SOFAR channel are labeled for both sites in (c). The modeled ocean depths are 4500 m at Wake Island, and 3000 m at Ascension Island.

The depths z1 and z2, determining the effective depths of the SOFAR channel for the Ascension Island site, are *z1*=390 m and *z2*=1356 m. The corresponding depths for the Wake Island site are *z1*=460 m and *z2*=1600 m. These depth limits are also consistent with the turning points of the acoustic ray paths propagating along the SOFAR channel, defined as the maximum and minimum depths containing SOFAR-refracted acoustic ray paths [57]. These depth limits not only provide boundaries for determining Argo temperature changes, but they also give an

estimate of which parts of the ocean are being sampled by the coherent ambient noise traveling between sensors (i.e. which parts of the water column are primarily contributing to the overall temperature variations obtained from the passive thermometry method).

Appendix F: Long-Range Parabolic Equation Propagation Model

While the temperature change comparison provides a good first verification of the passive monitoring method, it fails to take into account several important phenomena that happen in the ocean in the low-frequency regime. Because the frequencies are so low (1-40 Hz), the coherent waves cannot be confidently modeled as rays that only travel along the SOFAR channel. Instead, the frequency range necessitates a combination of a normal-mode approach (where the entire water column, not just the SOFAR channel, plays a role in propagation) and a ray approach, both of which can be captured in a 2-dimensional Parabolic Equation (PE) model. The PE model uses a split-step Padé approach with 4 Padé terms. The code for this model is called the Range-Independent Acoustic Model (RAM), and was developed by Mike Collins at the Navy Research Lab [60] , [61]. Bathymetry data input to the model is taken from Google Earth, which sources its bathymetry data from many government and academic institutions. However, it should be noted that modeling the Wake location using a flat bathymetry gave pressure wave forms that match better with the actual cross-correlations. At Ascension Island, the bathymetry is roughly modeled, with one data point every 30 km, to speed computation time. The code uses a self-starter at the source location and runs over intervals of 0.2 Hz for the frequency range 1-40 Hz. While the physical scenario makes use of diffuse noise sources, the PE code substitutes a source for the array element in the direction the noise is coming from to calculate the pressure waveform between sensor pairs. There is no attenuation modeled in the water, although attenuation in the sediment is assumed to be 1 dB/m starting at 100 m below the bottom of the water and 5 dB/m starting at 300 m below the bottom. This increasing attenuation removes any fictitious reflections from the bottom of the sediment layer. The model is run for

each South-North sensor pair. The pressure arrivals are filtered by the magnitude of the reference cross-correlation wave forms in the frequency domain to ensure equal frequency weighting as the real-world scenario. This helps to account for the spectral difference between a self-starter, which gives equal weight to all frequency bands, and ambient noise sources which are physically complex and have many unknown properties. The phase of the arrival at each frequency is solely determined by the PE model with no post-processing.

The sound speed profile used in the PE model is range-independent, and is estimated by applying Del Grosso's equation for sound speed to the Argo averaged temperature, salinity, and depth measurements [62]. Errors exist in the averaging process for the Argo data, and the differences between the Argo data and an independent data set at Ascension are shown in Figure 42. Note that the averaging process drastically smooths the profile with respect to depth. Yet the averaged data gives the necessary temporal resolution (1 month) to be able to track arrival time fluctuations with the PE model. The Ocean World Atlas data has only one sound speed profile over several years, and point Argo measurements are sporadic.

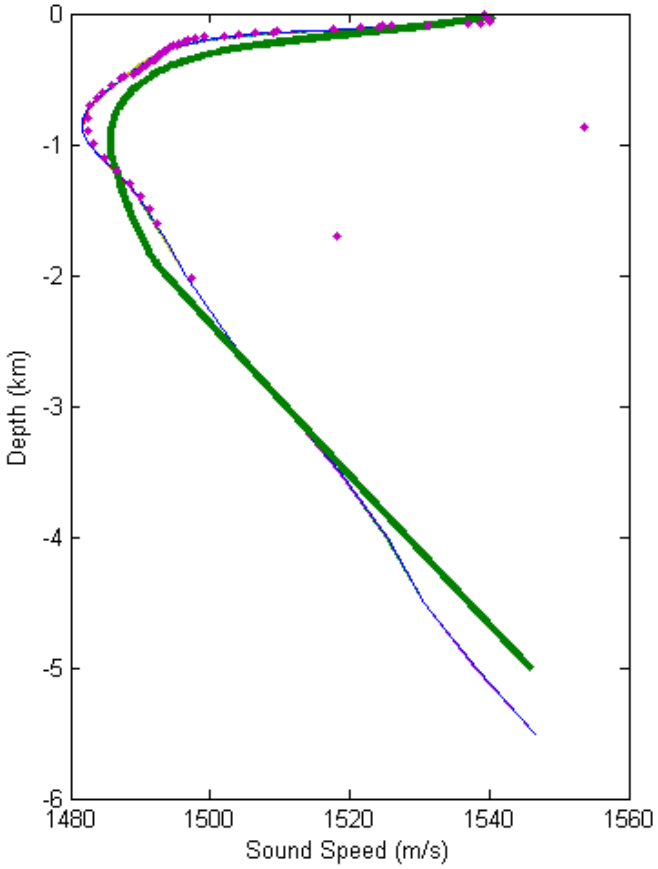

Figure 42: This is a comparison of the sound speed profile as determined by the averaged Argo data (green), a single, non-averaged Argo measurement dive (purple), and the NOAA Ocean World Atlas (blue). The Ocean World Atlas measurements are independent of the Argo data. Note that the averaged profile is a rough estimate of the other two profiles.

Looking at snapshots of the propagation over various ranges in Figure 43 and Figure 44 shows high-angle wave fronts that propagate outward. Reflections from the surface and sea floor are modeled, with the changing bathymetry at Ascension decreasing travel time between wave fronts.

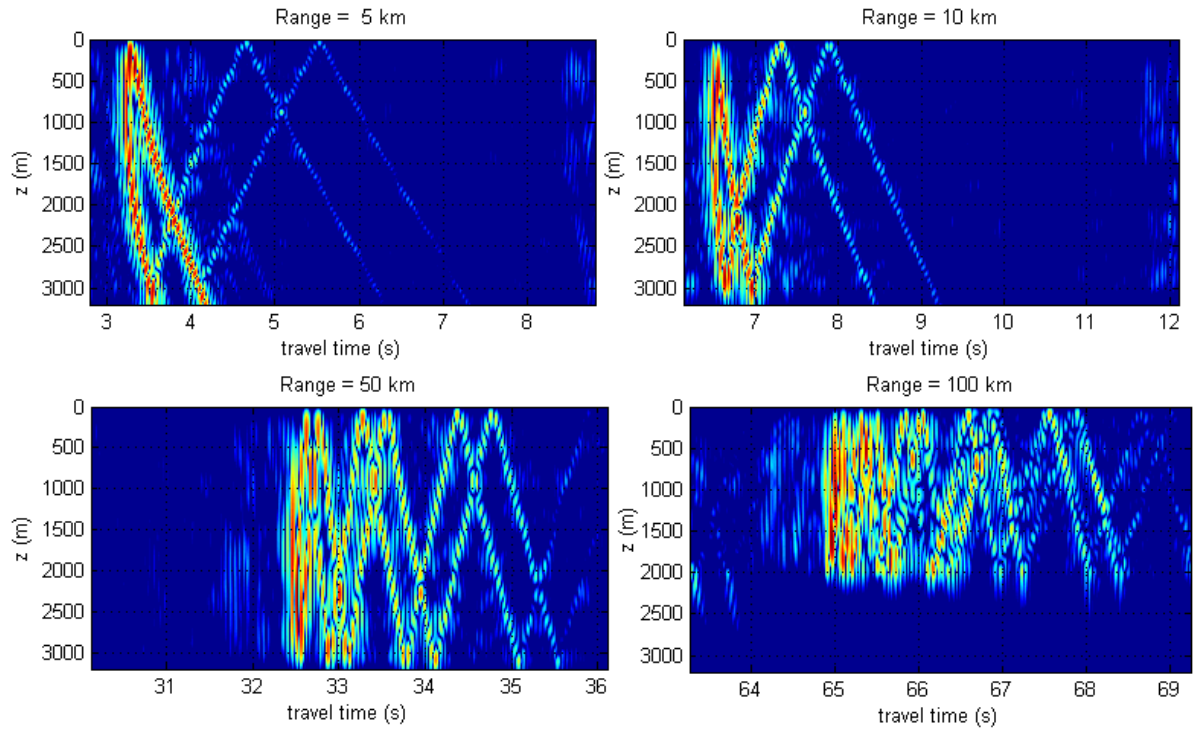

Figure 43: Propagated intensity arrivals at Ascension traveling from south array element 1 to north array element 1 at various ranges. This figure gives some insight into how the changing bathymetry at Ascension affects the propagation by closely spacing the arrivals. This is plotted for the first month in 2006.

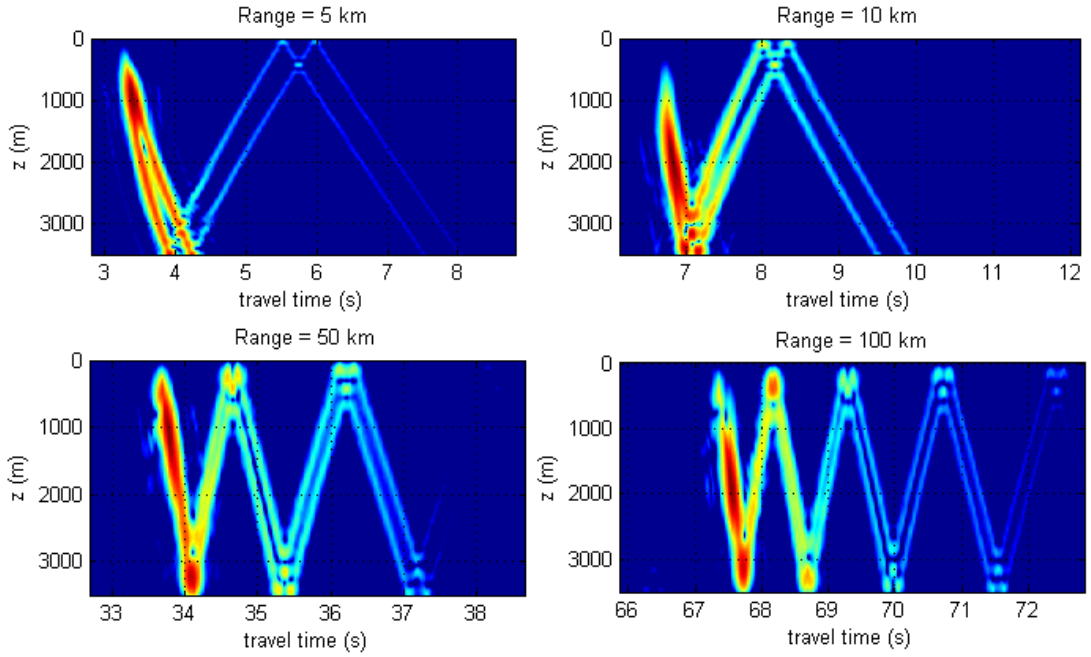

Figure 44: Propagated intensity arrivals at Wake between the south array first hydrophone and the north array first hydrophone using a constant bathymetry. This is for this first month in 2010.

The intensity at the receiver locations for Wake and Ascension is plotted in Figure 45 and Figure 46. The arrival spread in time indicates that there are significant multipaths of propagation, meaning that most (if not all) of the water column is sampled by the traveling waves. Because arrival times are dependent on the sound speed of the entire path that a wave travels, this path-integral effect means that shifts in arrival times in the PE model are due to changes in temperature in the entire water column, not just around the SOFAR channel.

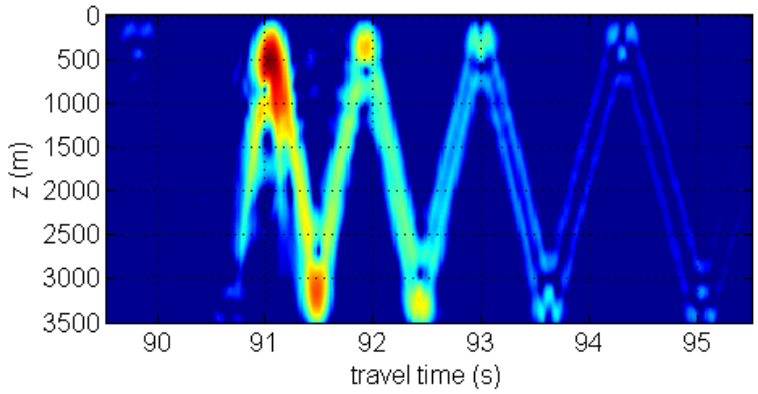

Figure 45: Intensity of arrivals arriving at the range of the north first hydrophone at Wake Island. These arrivals are plotted for the January 2010 Argo data, and the source is located at the south first hydrophone.

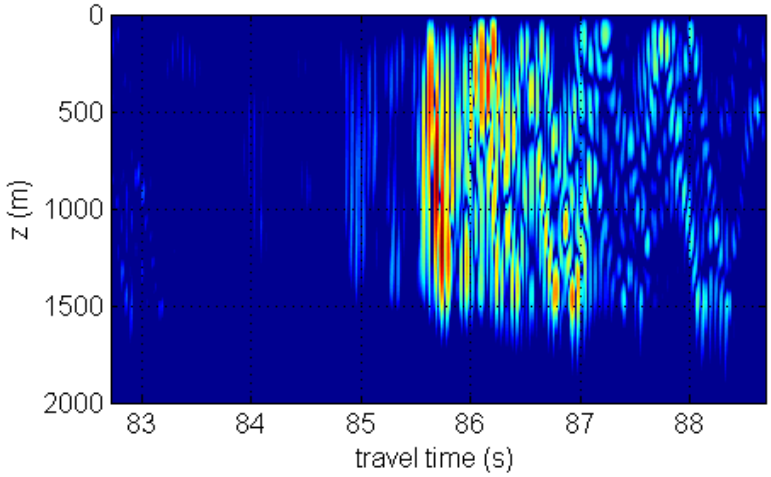

Figure 46: Intensity of arrivals at the range of the first hydrophone in the Ascension north array. The source is located at the first hydrophone in the south array, and this intensity is plotted for January 2006. The wave fronts of the arrivals are closely spaced and hard to distinguish, although reflected paths are clearer as the wave dies out.

Running the PE model for all sensor pairs at both sites yields a model of the arrival waveforms between sensor pairs. The PE model does not capture all the potential arrival paths, and this fact is demonstrated in Figure 47B. The physical environment around Wake Island

supports arrivals that arrive earlier than the main arrival, but these arrivals are not present in the model. Similarly, there are late arrivals in the first three cross correlation waveforms in Figure 47A that are not accurately reflected in the PE model. These discrepancies are likely due to 3-D effects as well as a too-rough bathymetry profile.

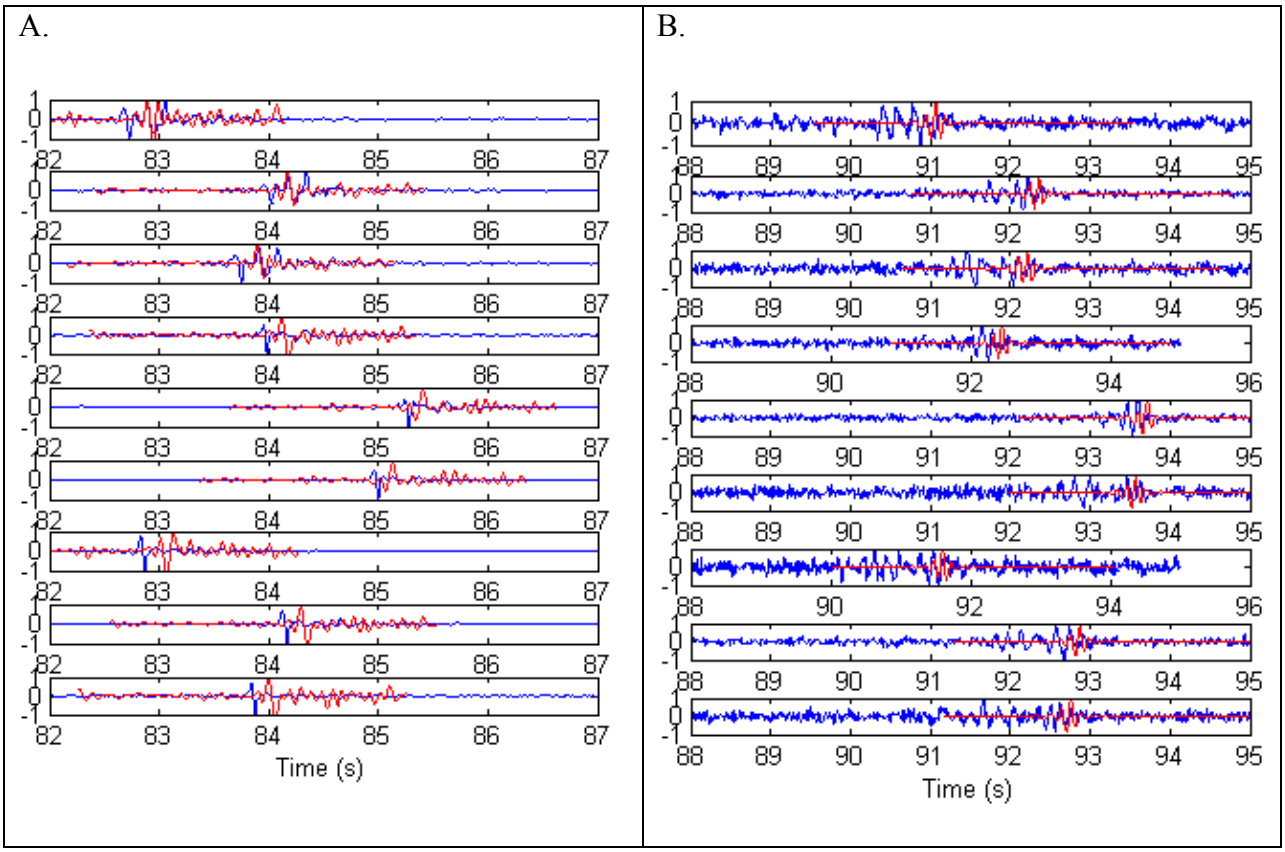

Figure 47: The 9 sensor pair normalized arrival wave forms for A) Ascension and B) Wake. The blue wave forms are the non-windowed reference cross correlations used in the beamforming. The red wave forms are the PE model result.

To understand the error associated with the arrivals, the NOAA Ocean World Atlas sound speed profile is input to the PE model. Arrival waveform results for this model are displayed only for Ascension Island in Figure 48. Clearly the errors associated with the Argo profiles are large. The published temperature errors associated with the Argo data are a variance of 0.55 $°C^2$. Applying this variance to the entire temperature profile and feeding the profile through the PE model yields an arrival time standard deviation of ±0.2 s. The cross-correlations are within the

errors of the PE model. The NOAA arrivals are closer to the cross-correlations in arrival time than the Argo data, but the Argo data are needed for tracking peak arrivals over time.

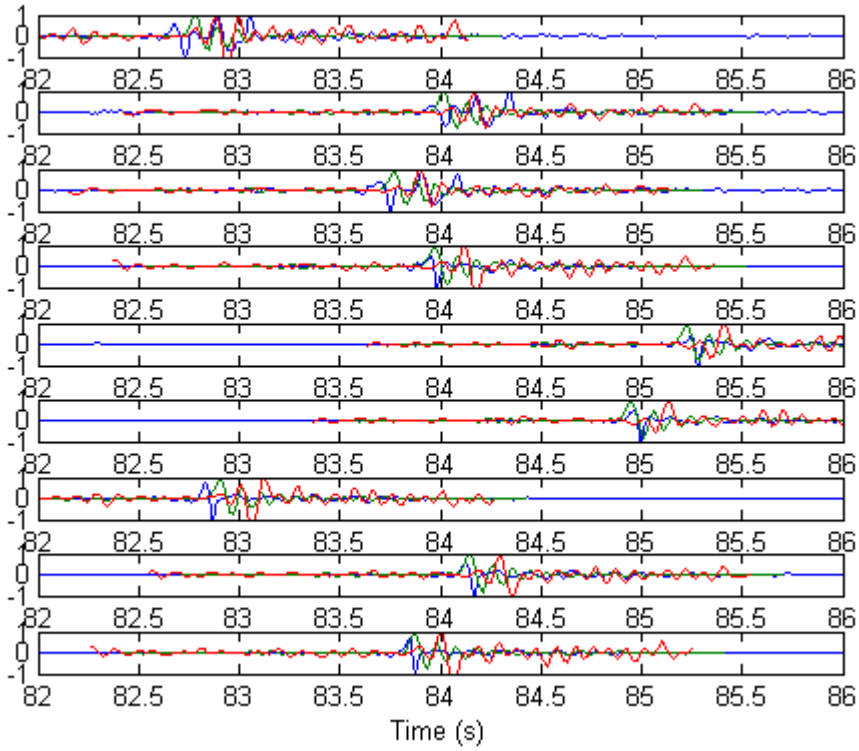

Figure 48: Normalized arrivals between all 9 sensor pairs at Ascension using the cross-correlation (blue), PE model with averaged Argo sound speed profile (red), and PE model with Ocean World Atlas sound speed profile (green).

To model the arrival time shifts over months, one north-south sensor pair at each site was selected. Then the model was run for each month with the averaged Argo-derived sound speed profile for that month. The waveform positive peaks are tracked over all the months where data exists for the IMS sites. Arrival time variance is determined by inputting maximum estimated temperature errors into the sound speed profile, then running the model for that profile. Figure 49 illustrates the arrival time shift for Wake Island from the model compared with the beamformer

output, and Figure 50 plots the same for Ascension Island. The modeled peak arrivals do not match up as well with the beamformed arrivals as the Argo temperature comparison. This is partially due to the fact that the PE model assumes a wide angle source, captured by the self-starter. However, propagation in the physical ocean is likely more spatially restricted to the few hundred meters around the SOFAR channel because the real sources are so distant (i.e. located at the Polar Regions). It is difficult to capture this source directivity with the RAM PE model. Attempts to create the effects of a more directive source by inserting a line segment of point sources in lieu of a single point source were fruitless.

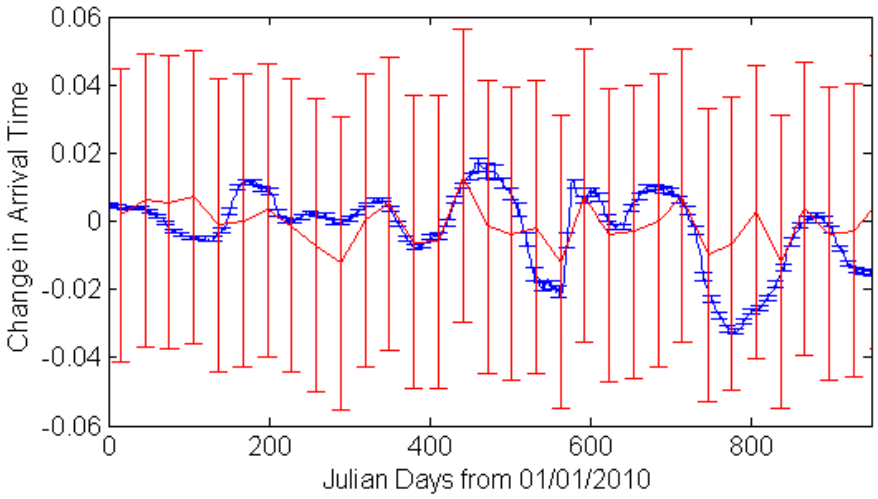

Figure 49: Peak arrival time shifts for the south first and north first sensor combination at Wake Island from January 2010 through December 2012. The PE model is in red, and the beamformer output peak arrivals are in blue.

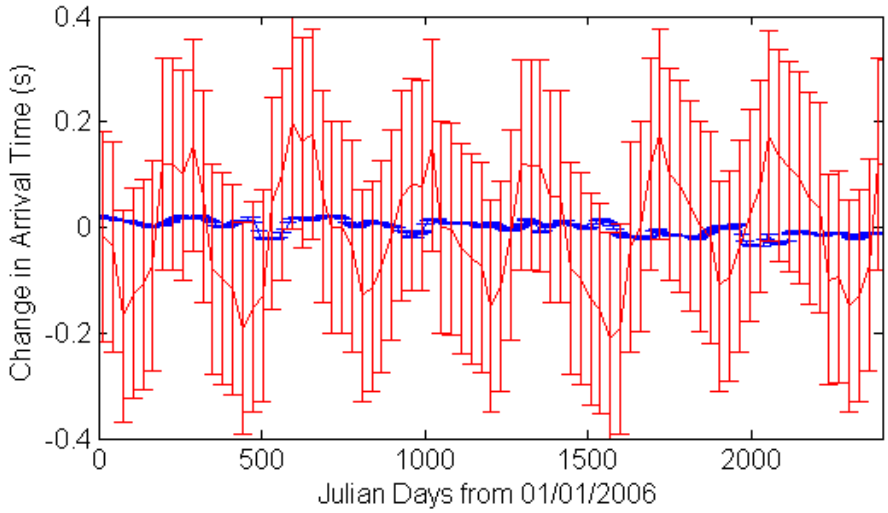

Figure 50: Peak arrival time shifts for the south first and north first sensor combination at Ascension Island from January 2006 through December 2012. The PE model is in red, and the beamformer output peak arrivals are in blue.

Bibliography

[1] X. Chen and K. Tung, "Varying planetary heat sink led to global-warming slowdown and acceleration," *Science,* vol. 345, pp. 897-903, 2014.

[2] S. Häkkinen, P. Rhines and D. Worthen, "Atmospheric Blocking and Atlantic Multidecadel Ocean Variability," *Science,* vol. 334, pp. 655-659, 2011.

[3] T. Barnett, D. Pierce, K. Achuta-Rao, P. Gleckler, B. Santer, J. Gregory and W. Washington, "Penetration of Human-Induced Warming into the World's Oceans," *Science,* vol. 309, pp. 284-287, 2005.

[4] The ATOC Consortium, "Ocean Climate Change: Comparison of Acoustic Tomography, Satellite Altimetry, and Modeling," *Science,* vol. 28, pp. 1327-1332, 1998.

[5] D. Roemmich and J. Gilson, "The 2004-2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program," *Prog. Oceanogr.,* vol. 82, pp. 81-100, 2009.

[6] W. Munk, P. Worcester and C. Wunsch, Ocean Acoustic Tomography, New York: Cambridge University Press, 1995.

[7] P. Worcester, R. Spindel and B. Howe, "Reciprocal acoustic transmissions: instrumentation for mesoscale monitoring of ocean currents," *IEEE J. Ocean. Eng.,* no. 2, pp. OE-10, 1985.

[8] S. Lani, K. Sabra, W. Hodgkiss, W. Kuperman and P. Roux, "Coherent processing of shipping noise for ocean monitoring," *J. Acoust. Soc. Am.,* vol. 133, p. EL108, 2013.

[9] O. Godin, N. Zabotin and V. Goncharov, "Ocean tomography with acoustic daylight," *Geophys. Res. Lett. ,* vol. 37, p. L13605, 2010.

[10] O. Lobkis and R. Weaver, "On the emergence of the Green's function in the correlations of a diffuse field," *J. Acoust. Soc. Am.,* vol. 110, pp. 3011-3017, 2001.

[11] Roux, P., Kuperman, W., and the NPAL Group, "Extracting coherent wave fronts from acoustic ambient noise in the ocean," *J. Acoustic. Soc. Am,* vol. 116, no. 4, pp. 1995-2003, 2004.

[12] K. Sabra, P. Roux and W. Kuperman, "Emergence rate of the time-domain Green's function from the ambient noise cross-correlation function," *J. Acoust. Soc. Am,* vol. 118, pp. 3524-3531, 2005.

[13] W. Munk, R. Spindel, A. Baggeroer and T. Birdsall, "The Heard Island Feasibility Test," *J. Acoust. Soc. Am.,* vol. 96, pp. 2330-2342, 1994.

[14] A. Einstein, "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen (On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat," *Annalen der Physik,* vol. 17, pp. 549-560, 1905.

[15] G. M. Wenz, "Acoustic ambient noise in the ocean: Spectra and sources," *J. Acoust. Soc. Am.,* vol. 34, pp. 1936-1956, 1962.

[16] M. Balmaseda, K. Trenberth and E. Källén, "Distinctive climate signals in reanalysis of global ocean heat content," *Geophys. Res. Lett.,* vol. 40, pp. 1754-1759, 2013.

[17] N. Bindoff, N. Willebrand, V. Artale, A. Cazenave, J. Gregory, S. Gulev, K. Hanawa, C. Le

Quéré, S. Levitus, Y. Nojiri, C. Shum, L. Talley and L. Unnikrishnan, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor and H. Miller, Eds., New York, 2007.

[18] B. Dushaw, P. Worcester, W. Munk, R. Spindel, J. Mercer, B. Howe, K. Metzger, T. Birdsall, R. Andrew, M. Dzieciuch, B. Cornuelle and D. Menemenlis, " A decade of acoustic thermometry in the North Pacific Ocean," *J. Geophys.,* vol. 114, p. C07021, 2009.

[19] R. L. Weaver and O. I. Lobkis, "Ultrasonics without a source: Thermal fluctuation correlations at MHz frequencies," *Phys. Rev. Lett.,* vol. 87, 2001.

[20] M. Campillo and A. Paul, "Long-range correlations in the diffuse seismic coda," *Science,* vol. 299, no. 547-549, 2003.

[21] J. Bendat and A. Piersol, Random data: analysis and measurement procedures, 3rd Ed, New York: Wiley, 2010.

[22] P. Roux, K. Sabra, W. Kuperman and A. Roux, "Ambient noise cross-correlation in free-space: theoretical approach," *J. Acoust. Soc. Am.,* vol. 117, pp. 79-84, 2005.

[23] D. Johnson and D. Dudgeon, Array Signal Processing, Upper Saddle River: PTR Prentice-Hall, Inc., 1993.

[24] S. Stergiopoulos and A. Ashley, "An experimental evaluation of splitbeam processing as a broadband bearing estimator for line array sonar systems," *J. Acoust. Soc. Am.,* vol. 102, pp. 3556-3563, 1997.

[25] S. Autrey, "Ambient noise field edge effects in product array processing," *J. Acoust. Soc. Am.,* vol. 55, p. 102, 1974.

[26] C. Leroy, S. Lani, K. Sabra, W. Hodgkiss, W. Kuperman and P. Roux, "Enhancing the emergence rate of coherent wavefronts from ocean ambient noise correlations using spatio-temporal filters," *J. Acoust. Soc. Am.,* vol. 132, pp. 883-893, 2012.

[27] K. Woolfe, S. Lani, K. Sabra and W. Kuperman, "Monitoring deep ocean temperatures using acoustic ambient noise," *Geophys. Res. Lett.,* 2015.

[28] K. Sabra, S. Fried, W. Kuperman and M. Prior, "On the coherent components of low-frequency ambient noise in the Indian Ocean," *J. Acoust. Soc. Am.,* vol. 133, p. EL20, 2013.

[29] K. Mackenzie, "Nine-term equation for sound speed in the oceans," *J. Acoust. Soc. Am.,* vol. 70, pp. 807-812, 1981.

[30] W. Munk, " Sound channel in an exponentially stratified ocean, with application to SOFAR," *J. Acoust. Soc. Am.,* vol. 55, pp. 220-226, 1974.

[31] M. Porter, "The KRAKEN normal mode program," 1997. [Online]. Available: http://oalib.hlsresearch.com/Modes/kraken.pdf. [Accessed August 2014].

[32] O. Godin, "Accuracy of the deterministic travel time retrieval from cross-correlations of non-diffuse ambient noise," *J. Acoust. Soc. Am.,* vol. 126, no. 6, p. EL183, 2009.

[33] R. Weaver, B. Froment and M. Campillo, "On the correlation of non-isotropically distributed ballistic scalar diffuse waves," *J. Acoust. Soc. Am.,* vol. 126, p. 1817, 2009.

[34] W. Carey and R. Evans, Ocean Ambient Noise: Measurement and Theory, New York: Springer, 2011.

[35] S. Fried, S. Walker, W. Hodgkiss and W. Kuperman, "Measuring the effect of ambient noise directionality and split-beam processing on the convergence of the cross-correlation function," *J. Acoust. Soc. Am.,* vol. 134, pp. 1824-1832, 2013.

[36] M. Prior, D. Brown and G. Haralabus, "Data features from long-term monitoring of ocean noise," in *Proceedings of the 4th International Conference and Exhibition on Underwater Acoustic Measurements*, Kos, Greece, 2011.

[37] M. McDonald, J. Hildebrand and S. Wiggins, "Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California," *J. Acoust. Soc. Am.,* vol. 120, pp. 711-718, 2006.

[38] K. Stafford, S. Nieukirk and C. Fox, "Low-frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific," *J. Acoust. Soc. Am.,* vol. 106, pp. 3687-3698, 1999.

[39] M. McDonald, J. Hildebrand and S. Webb, "Blue and fin whales observed on a seafloor array in the Northeast Pacific," *J. Acoust. Soc. Am.,* vol. 98, pp. 712-721, 1995.

[40] J. Miksis-Olds, C. Smith, R. Hawkins and D. Bradley, "Seasonal soundscapes from three ocean basins: what is driving the differences?," in *ECUA 2012 11th European Conference on Underwater Acoustics* , Edinburgh, Scotland, 2012.

[41] H. Matsumoto, D. Bohnenstiehl, J. Tournadre, R. Dziak, J. Haxel, T. Lau, M. Fowler and S. Salo, "Antarctic icebergs: A significant natural ocean sound source in the Southern Hemisphere," *Geochem. Geophys. ,* vol. 15, pp. 3448-3458, 2014.

[42] R. Dziak, D. Bohnenstiehl, H. Matsumoto, C. Fox, D. Smith, M. Tolstoy, T. Lau, J. Haxel

and M. Fowler, "P- and T- Wave Detection Thresholds, Pn Velocity Estimate, and Detection of Lower Mantle and Core P-Waves on Ocean Sound-Channel Hydrophones at the Mid-Atlantic Ridge," *B. Seismol. Soc. Am.,* vol. 94, no. 2, pp. 665-677, 2004.

[43] S. Nieukirk, K. Stafford, D. Mellinger, R. Dziak and C. Fox, "Low-frequency whale and seismic airgun sounds recorded in the mid-Atlantic Ocean," *J. Acoust. Soc. Am.,* vol. 115, pp. 1832-1843, 2004.

[44] C. Müller, V. Schlindwein, A. Eckstaller and H. Miller, "Singing Icebergs," *Science,* vol. 310, p. 1299, 2005.

[45] J. Talandier, O. Hyvernaud, D. Reymond and E. Okal, "Hydroacoustic signals generated by parked and drifting icebergs in the Southern Indian and Pacific Oceans," *Geophys. J. Int.,* vol. 165, pp. 817-834, 2006.

[46] E. Chapp, D. Bohnenstiehl and M. Tolstoy, " Sound-channel observations of ice-generated tremor in the Indian Ocean," *Geochem. Geophys. Geosyst.,* vol. 6, p. Q06003, 2005.

[47] J. Comiso and F. Nishio, "Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data," *J. Geophys. Res.,* vol. 113, pp. 2156-2202, 2008.

[48] K. Sabra, P. Roux, A. Thode, G. D'Spain, W. Hodgkiss and W. Kuperman, "Using Ocean Ambient Noise for Array Self-Localization and Self-Synchronization," *IEEE J. Ocean. Eng.,* vol. 30, no. 2, pp. 338-347, 2005.

[49] P. Gouedard, L. Stehly, F. Brenguier, M. Campillo, Y. Colin de Verdiere, E. Larose, L. Margerin, P. Roux, F. Sanchez-Sesma, N. Shapiro and R. Weaver, "Cross-correlation of random fields: Mathematical approach and applications," *Geophys. Prospect.,* vol. 56, pp.

375-393, 2008.

[50] N. Zabotin and O. Godin, "Emergence of acoustic Green's functions from time averages of ambient noise," *Acta. Acust. U. Acust.,* vol. 97, pp. 44-53, 2011.

[51] J. Holland, Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, University of Michigan Press, 1975.

[52] D. Goldberg, Genetic Algorithms in Search, Optimization, & Machine Learning, Addison-Wesley, 1989.

[53] A. Conn, N. Gould and P. Toint, "A Globally Convergent Augmented Lagrangian Algorithm for Optimization with General Constraints and Simple Bounds," *SIAM J. Numer. Anal.,* vol. 28, pp. 545-572, 1991.

[54] F. Fetterer, K. Knowles, W. Meier and M. Savoie, "Sea Ice Index," National Snow and Ice Data Center, Boulder, Colorado, 2009.

[55] G. Kinda, Y. Simard, C. Gervaise, J. Mars and L. Fortier, "Under-ice ambient noise in Eastern Beaufort Sea, Canadian Arctic, and its relation to environmental forcing," *J. Acoust. Soc. Am.,* vol. 134, pp. 77-87, 2013.

[56] C. Helstrom, Statistical Theory of Signal Detection, 2nd Edition, London: Pergamon Press, 1968.

[57] F. Jensen, W. Kuperman, M. Porter and H. Schmidt, Computational Ocean Acoustics, 2nd Edition, New York: Springer, 2011.

[58] R. Locarnini, A. Mishonov, J. Antonov, T. Boyer, H. Garcia, O. Baranova, M. Zweng and

D. Johnson, World Ocean Atlas 2009 Volume 1: Temperature, Washington, D.C.: U.S. Government Printing Office, 2010.

[59] J. Antonov, D. Seidov, T. Boyer, Locarnini, A. Mishonov, O. K. Garcia, Baranova, M. Zweng and D. Johnson, World Ocean Atlas 2009, Volume 2: Salinity, Washington, D.C.: U.S. Government Printing Office, 2010.

[60] M. Collins, "A split-step Padé solution for the parabolic equation method," *J. Acoust. Soc. Am.,* vol. 93, pp. 1736-1742, 1993.

[61] M. Collins, "An energy-conserving parabolic equation for elastic media," *J. Acoust. Soc. Am,* vol. 94, pp. 975-982, 1993.

[62] V. Del Grosso, "New equation for the speed of sound in natural waters (with comparisons to other equations)," *J. Acoust. Soc. Am.,* vol. 56, pp. 1084-1091, 1974.

[63] Baggeroer, A., Scheer, E., and the NPAL Group, "Statistics and vertical directionality of low-frequency ambient noise at the North Pacific Acoustic Laboratory site," *J. Acoust. Soc. Am.,* vol. 117, no. 3, pp. 1643-1665, 2004.

[64] G. D'Spain, W. Hodgkiss and G. Edmonds, "Horizontal directionality of the ocean infrasonic sound field: The effect of bottom topography," *J. Acoust. Soc. Am.,* vol. 87, no. S111, 1990.

[65] K. Curtis, B. Howe and J. Mercer, "Low-frequency ambient sound in the North Pacific: Long time series observations," *J. Acoust. Soc. Am.,* vol. 106, p. 3189, 1999.

[66] W. Carey and R. Wagstaff, "Low-frequency noise fields," *J. Acoust. Soc. Am,* vol. 80, no. 5, pp. 1523-1526, 1986.