Optimized Acoustic Sensing for Fixed-Wing Uncrewed Aerial Vehicles

Rowe Bradley Brookfield, II

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
https://digitalcommons.memphis.edu/etd/3088

This Thesis is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.
OPTIMIZED ACOUSTIC SENSING FOR FIXED-WING UNCREWED AERIAL VEHICLES

by

Rowe Bradley Brookfield, II

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

Major: Electrical and Computer Engineering

The University of Memphis
May 2023
First and foremost, this thesis is dedicated to my heavenly Father in whom I have faith, hope, love, and salvation. I pray daily that the work I do is not my own, but rather it is His work performed through me because He is the reason my life has meaning and purpose. Also, I dedicate this to my immediate family and friends who have always supported me and believed in me, but also loved me and provided me an outlet in the midst of a high-stress season, in addition to Maddie who has done all the same things, but I have also had the privilege of sharing the journey of developing a thesis alongside her.
ACKNOWLEDGMENTS

To begin, I would like to thank everyone on my thesis committee. I thank Dr. Eddie Jacobs for walking me through every step of developing a thesis, helping me through every obstacle I have stumbled upon, setting me on track to graduate effectively and efficiently, and being an outstanding mentor overall having extreme patience and consideration. I thank Dr. Daniel Foti for his extreme selflessness in guiding me in the areas of fluid mechanics and aeroacoustics (areas in which I was extremely lacking expertise, but was also in desperate need of), as well as his incredible patience with me as I slowly chipped away at formulating my thesis. Finally, I thank Dr. Alfredo Ramirez for all of his ideas and support in the experimental execution stages of the project.

To all of them collectively though, I thank them again for showing great interest in my proposal and bouncing off ideas with each other. It has been an incredible work environment with every one of them. Another special thanks goes to every member involved in the MUAS project as well as the Army Research Laboratory (ARL) for being the main sponsor. Had this project not been launched with every member included, this thesis may not have even become an idea worth being entertained. Additionally, I thank the composers of every open-source package used within my Python characterization web app. I also thank the Agricenter staff for giving us the opportunities for UAVs flights on their property. I especially thank Kevin and Ethan, both of whom were outstanding helpers in collecting data, building preamps and recorders for the microphones, and adjusting the 3D models for the wing mounts and other miscellaneous containers. Lastly, I thank every friend and family member who has helped me review this thesis and make it the best it possibly could be. Without any of these contributors, I would not have been able to finish as quickly as I did.

My list of gratitude to others in the workings of this thesis could be close to never-ending, but for the sake of simplicity, I thank everyone brought up in the dedication section as well for all reasons that are specified there. I have been very blessed with large spheres of influence in this
season, and without them, none of this work would have been possible. An incredibly large thank you to everyone!

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-21-2-0294. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.
ABSTRACT

Acoustic sensors are devices that are not commonly used on autonomous uncrewed aerial vehicles (UAV). Obtaining a usable signal-to-noise ratio (SNR) is challenging. Given the most problematic noise is the flight-induced wind noise, one way of approaching the problem is to stop the wind noise at the source by designing a mount for the acoustic sensors to reduce the wind component before the signal and noise enter the microphone. Subsequently, signal processing stages can be added to improve the SNR further. We begin by formulating an atmospheric attenuation model using both point and line acoustic sources. The model predicts the frequency spectrum and how it reacts to changes in atmospheric conditions. This model is used to predict the SNR over the frequency range of interest as measured at the UAV for various wind speeds for a given acoustic source sound pressure level (SPL) as well as predict the SNR as a function of distance. Multiple fixed-wing UAV mounting strategies are then developed based on the predicted airflow during flight with each analyzed with respect to SNR. Based on predicted SNRs, various signal processing algorithms are evaluated for their improvement of detection statistics. Finally, the SNR of the processed signal is evaluated for usability. Particular instantiations of the acoustic sensing wing mounts are evaluated in the lab using a wind tunnel as well as in some physical UAV test flights. Data collected from these flights is processed offline using different signal processing approaches. Based on the model predictions and the results of the limited field measurements, conclusions regarding the feasibility of acoustic sensing on a UAV are discussed.
# TABLE OF CONTENTS

**List of Tables**  viii  
**List of Figures**  ix  
**List of Abbreviations**  xi  

## 1 Introduction  
- Background ........................................ 1  
- MUAS Project ..................................... 2  
- Acoustic Sensor Taxonomy ....................... 3  

## 2 Literature Review  6  

## 3 Methodology  10  
- Overview ......................................... 10  
- From Fluid Dynamics to Propagation of Acoustic Waves 10  
- Determining Atmospheric Conditions ............... 16  
- Linear Propagation of Acoustic Sources .......... 18  
- Spectral Atmospheric Propagation Model ......... 21  
- Doppler Effect ................................... 24  
- A, C, and Z Acoustic Weightings ................. 25  
- Wind Speed and SPL Relationship ............... 26  
- SNR Estimation .................................. 27  
- Filtering Techniques and Data Transformations 29  
  - Matched and Wiener Filtering .................. 29  
  - Harmonic Spectral Transform ................... 31  
- Detection Scheme ................................ 34  
- Data Collection Strategy ....................... 37  
  - Atmospheric Conditions and Experiment Setting Data 38  
  - Audio Samples ................................ 39  
- Fluid Dynamics Inspired Wing Mount Design .... 41  

## 4 Results  43  
- Lab Experimentation Results ...................... 43  
  - Characterization of the Mounts ............... 43  
  - SNR Performance ................................ 45  
- In-Field Flight Results ......................... 49  
  - SNR and Distance ................................ 49  
  - Detection Performance ......................... 53  

## 5 Conclusion  61  

## 6 Future Work  62  
- Alternative Mounts and Acoustic Sensing Devices 62
# LIST OF TABLES

(1) Sound Pressure Level vs. Sound Intensity Level vs. Sound Power Level .......................... 19

(2) Applicable types of statistical averaging to Harmonic Spectral Transform (HST) ........... 33

(3) Flight results for the first 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude. .............................................. 54

(4) Flight results for the middle 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude. .............................................. 55

(5) Flight results for the last 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude. .............................................. 56

(6) Wing mount style variations. .................................................................................... 72
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Concept of long-range anomaly detection and short-range interrogation UAVs.</td>
</tr>
<tr>
<td>2</td>
<td>Albatross UAV on runway awaiting pre-flight checks.</td>
</tr>
<tr>
<td>3</td>
<td>Electret condenser microphone structure [1].</td>
</tr>
<tr>
<td>4</td>
<td>Noise cancellation active noise control (ANC) [2].</td>
</tr>
<tr>
<td>5</td>
<td>Eye Lift 60 foam UAV mic array implementation [3].</td>
</tr>
<tr>
<td>6</td>
<td>Directivity factors due to propagation spread. Top left: $Q = 1$, top right: $Q = 2$, bottom left: $Q = 4$, bottom right: $Q = 8$ [4].</td>
</tr>
<tr>
<td>7</td>
<td>Attenuation coefficient varying humidity for all frequencies per atmosphere of pressure at a distance of 100 meters.</td>
</tr>
<tr>
<td>8</td>
<td>Attenuation coefficient varying temperature for all frequencies per atmosphere of pressure at a distance of 100 meters.</td>
</tr>
<tr>
<td>9</td>
<td>A moving bell to the right of the image and its respective changes in wavelength [5].</td>
</tr>
<tr>
<td>10</td>
<td>A, C, and Z weighting SPL curves [6].</td>
</tr>
<tr>
<td>11</td>
<td>HST of flat mount diesel recording in $20 \text{ m/s}$ winds.</td>
</tr>
<tr>
<td>12</td>
<td>Angel VTOL UAV used for flight data collection.</td>
</tr>
<tr>
<td>13</td>
<td>Weather station sensing device capturing atmospheric data at the Agricenter.</td>
</tr>
<tr>
<td>14</td>
<td>Acoustic sensor wing mount being tested in the wind tunnel.</td>
</tr>
<tr>
<td>15</td>
<td>A few different variations of the tested acoustic sensor wing mounts.</td>
</tr>
<tr>
<td>16</td>
<td>Comparison of spectral response for all multi-channel mounts.</td>
</tr>
<tr>
<td>17</td>
<td>Comparison of spectral SNR for single-channel ramp and flat mounts in $13 \text{ m/s}$ winds.</td>
</tr>
<tr>
<td>18</td>
<td>Comparison of spectral SNR for all multi-channel mounts in $13 \text{ m/s}$ winds.</td>
</tr>
<tr>
<td>19</td>
<td>Comparison of spectral SNR for the flat mount with various wind speeds.</td>
</tr>
<tr>
<td>20</td>
<td>Comparison of spectral SNR for all multi-channel mounts in $20 \text{ m/s}$ winds.</td>
</tr>
<tr>
<td>21</td>
<td>Mission path for the Angel VTOL UAV.</td>
</tr>
<tr>
<td>22</td>
<td>Measured Angel SNR as a function of distance compared to ideal conditions with Albatross.</td>
</tr>
</tbody>
</table>
(23) RMSE plot as a function of distance from Angel flight data collection.

(24) Error metric plot as a function of distance from Angel flight data collection.

(25) A comparison of original, harmonic, and percussive spectrograms of noisy signals taken at the wind tunnel with a 50 to 2000 Hz exponential sweep signal and 15 m/s winds.
LIST OF ABBREVIATIONS

ANC: Active Noise Control
AOP: Acoustic Overload Point
AQBNE: Adaptive Quantile Based Noise Estimation
ARL: Army Research Laboratory
AVS: Acoustic Vector Sensors
CDF: Cumulative Density Function
DOA: Direction of Arrival
DTFT: Discrete Time Fourier Transform
EMI: Electromagnetic Interference
FFT: Fast Fourier Transform
GHST: Generalized Harmonic Spectral Transform
HPS: Harmonic Product Spectrum
HPSS: Harmonic and Percussive Source Separation
HST: Harmonic Spectral Transform
ICA: Independent Component Analysis
IMCRA: Improved Minima Controlled Recursive Averaging
ISR: Intelligence, Surveillance, and Reconnaissance
JFET: Junction Field Effect Transistor
LMS: Least Mean Squares
MAE: Mean Absolute Error
MEMS: Microelectromechanical System
MSE: Mean Square Error
MSEN: Multi-Sensor
MUAS: Multi Uncrewed Aerial System
NLNS: Nonlinear Noise Subtraction
PCB: Printed Circuit Board
PDF: Probability Density Function
PSD: Power Spectral Density
RMSE: Root Mean Square Error
ROC: Receiver Operating Characteristics
SIL: Sound Intensity Level
SNR: Signal-to-Noise Ratio
SPL: Sound Pressure Level
STFT: Short Time Fourier Transform
SWL: Sound Power Level
TFS: Time-Frequency Spatial Filtering
UA: University of Arizona
UAV: Uncrewed Aerial Vehicle
UCF: University of Central Florida
UM: University of Memphis
VTOL: Vertical Takeoff and Landing
WSS: Wide-Sense Stationary
1 INTRODUCTION

Background

Embedding acoustic sensors on autonomous uncrewed aerial vehicles (UAVs) and other forms of aircraft has been an area of research that is less traveled compared to other sensors due to the challenges of common aerial system noises like engine noise, rotor noise, electronic noise, electromagnetic interference noise, and wind noise. However, once the research is done to mitigate noise as much as possible, there may be great benefits that will be unlocked. For military and other investigative purposes, many believe that they can be used for long-range vehicle detection. This is a valid achievement to consider, but much work needs to be done to get to that point. An alternative use that makes this idea much more feasible is using these sensors for closer-range target identification. This is the intent of this project.

In some sense, complex electronic and digital filtering is absolutely necessary for reducing these noises, but there are other means of filtering and improving the signal-to-noise ratio (SNR) from mechanical, fluid dynamic, and acoustic standpoints. Stopping or reducing the most problematic noise from the source before it enters the electronic system is an essential first step. The work performed for this thesis exploits this idea by analyzing the movement of the wind around and within the apparatus, the types of acoustic sensors used, how they are mounted, and how they can be best filtered and used for detection after the fact. This is done by considering a sound source, observing the atmospheric conditions, calculating the acoustic propagation characteristics, and adjusting the aerodynamic mounting designs in order to improve the signal-to-noise ratio. The evaluation consists of characterizing the microphones, preamps, and respective mounts, testing them with normal UAV traveling wind speeds in a wind tunnel, performing in-field flight tests, optimally post-processing and filtering the data, and using extensive spectral, temporal, and statistical analysis to break down the performance, with SNR being the characteristic of most interest.
MUAS Project

This thesis ultimately seeks to contribute to the Multi-Uncrewed Aerial System Multi-Sensor Intelligence, Surveillance, and Reconnaissance project. This will hereafter be referred to as the MUAS/MSEN project or the ISR project. The effort was proposed by the University of Memphis (UM), the University of Arizona (UA), and the University of Central Florida (UCF) and sponsored by the Army Research Laboratory (ARL) with the goal of designing autonomous UAVs to identify threats and items of interest for security, agricultural, and military applications. To further elaborate, the project can be broken down into two phases: phase one, which involves UAVs from longer distances to determine anomalies or inconsistencies in the explored terrain, and phase two, which then uses UAVs for short-range interrogation to better identify and confirm the target anomalies. An idea of this is depicted in Figure 1.

Figure 1: Concept of long-range anomaly detection and short-range interrogation UAVs.

Essentially, a long-range UAV identifies the area of interest and notifies a common host, and then an interrogation UAV is notified and sent for target confirmation and verification. Phase two is where these acoustic devices have a high potential. If the SNR can be enhanced, this will lead to much higher detection results and can simplify electronic filtering. Acoustic sensors are suited for
the military application of this project because they can use acoustic signals to help distinguish which targets are true targets and which are not threats or are decoys. An improvement in these sensors can allow military forces to separate targets from decoys at a quicker rate, thus, giving them an advantage on the battlefield.

Because this research is based around phase two of the MUAS project, the UAV that this work focuses on is the Albatross UAV supplied by Applied Aeronautics. It is an affordable commercial fixed-wing UAV that is lightweight, robust, and has good range and distance traveling capabilities [7]. It commonly flies at average speeds around 15-18 \( \frac{m}{s} \), with increased turning and transit speeds near 20 \( \frac{m}{s} \) and can slow down to speeds near 13 \( \frac{m}{s} \). UAVs with slower speeds are desirable as wind noise is reduced at lower speeds resulting in an improved signal-to-noise ratio (SNR).

![Figure 2: Albatross UAV on runway awaiting pre-flight checks.](image)

**Acoustic Sensor Taxonomy**

The broadest categories of acoustic devices are dynamic, carbon, and condenser microphones. Dynamic microphones are ones that use electromagnetism to convert the variations of sound pressure level (SPL) to electrical signals [8]. Due to this principle, they tend to be large and rugged and can have various polar patterns for certain directivity indices, leaving them to be best used for musical purposes. Then there are carbon microphones that use carbon particles to
cause resistance variations with respect to the changing sound pressures. Given this is an older type of microphone, they are now used for few applications due to poor signal quality, poor signal response, and high noise. Thirdly, there are condenser microphones, which move capacitive plates with the sound pressure changes, causing variations in voltage that can later be amplified.

Stemming from these broad categories, there are acoustic sensors that are designed for small, printed circuit board (PCB) level purposes. Technology in this category has come a long way due to many designs made with various filters and amplifiers and with intentions for different acoustic environments. The most notable ones include piezoelectric microphones, microelectromechanical system (MEMS) microphones, and electret condenser microphones.

Piezoelectric sensors use piezoelectric crystals that directly convert mechanical pressure to electric charges \[8, 9\]. They often come in either a ceramic material that is friendly to high frequencies or a film material that is more efficient overall in delivering the signal. One disadvantage they have is a high output impedance, so a specialized amplifier is required.

Another PCB-level sensor worth noting is MEMS microphones. These are microphones that contain both the mechanical diaphragm and electronic components on one semiconductor wafer and are protected by one physical cover \([10]\). They can withstand high temperatures, have great response characteristics, consume little power, and are compact. A few are designed specifically with optimal noise-filtered characteristics and are optimal for noisy environments. These were considered for the design, but the ones desired with high acoustic overload points and differential noise reduction filters were subject to supply chain issues. As will be later specified in section \(6\), if these are obtainable to future designers, these are highly recommended and deserve further consideration.

Electret condenser microphones are ones with dielectric materials that hold a permanent polarization and are constructed with a low-noise JFET common source amplifier for immediate signal growth \([8, 9, 11]\). An example of this structure is given in Figure\(3\).
They have a relatively low cost, have a good response, and come in various sizes for any embedded design. Their diaphragms can be sensitive to heat, so resoldering them can easily lead to damaged signal quality. Fortunately, they can be deployed redundantly and are easily replaced because of the low cost and high supply. For all of these reasons along with dust and humidity durability, these were the sensors used for this design.

One final sensor to note is the acoustic vector sensor (AVS). It is equipped with four sensors in one package, one is a microphone and the other three are particle velocity sensors that are orthogonal to one another. AVS sensors are well suited for purposes such as direction of arrival (DOA) estimation, observing turbulence, and sensing differential airspeed. Because it technically only has one microphone, these were not chosen due to the higher focus on filtering that it would require. Also, the problem that this thesis addresses is merely a detection problem; however, these devices would be excellent choices for acoustic source localization and for adapting to turbulent fluid flow.
2 LITERATURE REVIEW

Numerous recent works have made efforts to optimize acoustic sensing in noisy environments such as UAVs. UAVs are notorious for generating lots of noise simply from the motors and rotors so much so that it has been dubbed "ego-noise." This is especially prevalent with multi-rotor UAVs, which has led this to be a focus of research for these types of aircraft \[2,13,16\]. In turn, attempts have also been made to address the reduction of wind noise with fixed-wing UAVs \[3,17-20\]. There are large differences in reducing the noise between the two types of aircraft because on one hand multi-rotor drones are dominated more by the drone-generated ego-noise, but then on the other hand, fixed-wing UAVs have less of a problem with this and more of a problem with the high-speed wind turbulences that are constantly engulfing the signal. Other scientists have attempted to predict the direction of arrival (DOA) of distanced acoustic sources \[12,18,19,21-23\]. Additionally, supportive jumps have been made such as improving model behaviors through Bayesian inferencing, improving UAV reactions to certain acoustic signatures, mechanically bettering windscreens, developing more durable microphone designs for harsher environments, and of course, many new filtering techniques \[18-20,23-29\].

To elaborate on some of these advances, some topics that were exploited based on multi-rotor drones include more durable microphone designs, methods that reduce ego-noise, and DOA estimation. Since multi-rotor UAVs are designed for closer-range missions that may traverse various terrains, some have sought to make acoustic sensor designs more unbreakable. As described by recent sources, rugged designs may include making them waterproof, making them more resilient to sudden impacts, and optimizing them to have utmost reliability for wireless communication, all while still performing well with DOA estimation \[26\]. As for addressing the problem of ego noise, filtering schemes have been diverse. To begin, engineers tried measuring the noise that most multi-rotor UAVs generate and their respective polar patterns at which the noise has a stronger presence \[16\]. A few scientists attempted to build from spectral subtraction and hybridize this with an active noise control (ANC) (Figure 4) least mean squares (LMS)
algorithm, and achieved a high degree of ego noise cancellation [2].

![Noise cancellation active noise control (ANC) (2)](image)

Figure 4: Noise cancellation active noise control (ANC) [2].

Also stemming from spectral subtraction, another applied a nonlinear noise subtraction (NLNS) algorithm and compared this with two other methods: adaptive quantile-based noise estimation (AQBNE) and improved minima controlled recursive averaging (IMCRA) [14]. Another filtering maneuver is spatial filtering of the microphone channels followed by both a histogram-based and a kurtosis-based spatial likelihood calculation to improve the spatially informed filtering [13]. From the same researchers, a blind source separation framework was also proposed so that better data preprocessing for a clustering algorithm could be achieved, and this involved independent component analysis (ICA) and time-frequency spatial filtering (TFS) to align the ICA permuted outputs across the spectrum [15].

As those are some of the most notable recent multi-rotor filtering techniques, projects regarding DOA estimation are also numerous, and many of them have built upon coprime microphone array theory. It can increase degrees of freedom for the detection of multiple targets. This idea was first attempted and validated by Bush and Xiang [21]. Improvements were made through Bayesian inferencing and likelihood estimation. This was followed by a technique that uses rotational invariance to provide an alternative way to spatial smoothing that normally results in inefficient performance and a heavy computational burden [22, 23]. One should also note that
acoustic vector sensors were also developed in an attempt to optimize localization and DOA estimation [12].

With more relevance to this project, fixed-wing UAV acoustic sensing has come a long way over the past couple of decades as well. In humble beginnings, many were merely trying to see if it would be feasible and were trying to mount sensors on even foam airframes with extremely quiet motors in hopes of achieving a decent baseline of properties for detection performance [17]. After this, some attempted to embed mics within the UAV airfoil for improved wind, engine, and rotor noise reduction, and to explore the possibilities of microphone arrays [3].

Figure 5: Eye Lift 60 foam UAV mic array implementation [3].

This was also a foam glider so that the drone-induced noise could be kept to a minimum, and the work done for this thesis attempts to extend this in some ways. Research in this area has also extended beyond noise reduction as one group of scientists did use UAVs and their microphones as intelligence seekers and tackled using Bayesian inferencing to get the next moves of the UAV so that it can avoid dangerous situations [20]. Most recently, another has worked on a gas-powered fixed-wing UAV and tried to use a Harmonic Spectral Transform (HST) technique to give the detection system improved detection rates at low SNRs and at longer distances [18].

Regarding further noise reduction techniques, work has been done to study the porosity of microphone windscreens and how much it affects wind blockage [25]. The research previously mentioned for multi-rotor drones has high applicability to this project as does the HST transformation done by Harvey [2, 14, 15, 18, 19, 30]. Moreover, some have attempted to extend Wiener filtering to multiple channels and applying it in conjunction with microphone
beamforming [27]. Another has applied a recursive estimation method to find the spatial coherence matrix of the wind when given the microphone observations [28]. This spatial coherence matrix could then, in turn, be applied to predict the wind strength and determine how much of each component could be filtered across the spectrum. Other scientists have been developing and testing multiple concepts including single-channel methods like decision-directed SNR estimation and recursive gain spectral subtraction along with multiple-channel methods like coherence weighting, coherence-based noised estimation and subtraction, differential array noise suppression, and innovative methods for various applications like partial-speech synthesis [24].
3 METHODOLOGY

Overview

As stated in the introduction, the goal of this project is not only to confirm or deny targets of interest that are possible threats but also to see how far away detections can be made within a reasonable interval of tolerance. Moreover, atmospheric propagation is an important factor to ensure proper calculation of signal-to-noise ratios (SNR) on which detections can be based. There are many steps to predicting atmospheric propagation, which include deriving a model from the wave equation, finding and calculating several atmospheric parameters, breaking down the chemical characteristics of the atmospheric gases, and applying a point source acoustic model in conjunction with a spectral model from previous derivations. To fully quantify SNR and establish a detection decision system, other steps include finding the sound pressure and intensity from one point to another, quantifying the noise of the wind and determining the changes at different speeds, applying stages of signal processing, and evaluating regression error metrics on the harmonic spectrum to establish the detection boundary.

From Fluid Dynamics to Propagation of Acoustic Waves

To begin this mathematical atmospheric propagation model, it is assumed that all gases are assumed to be incompressible when objects are traveling at a low speed (less than Mach 0.3, which translates to 100 m/s) [31]. The term incompressible fluids in fluid dynamics implies that the density of the fluid, which is the density of air in this case, does not change [32]. This is a beneficial assumption to make if the proper conditions are met because this in turn leads to the air density and speed of sound being constant in the model’s calculations. However, this does not imply that it will not change long term due to atmospheric conditions; it simply means that it is stationary in the wide-sense (WSS) and will remain constant given the atmospheric conditions over a time consistent with the scenario being modeled. Additionally, note that if this were a substantially faster aircraft than the lightweight, moderate-speed, fixed-wing UAV that this design
focuses on, this assumption of incompressible flow would likely no longer hold. Because our aircraft move at speeds close to 20 m/s, we are justified in making this assumption.

Another assumption to be made that simplifies this model is that the viscosity of air can be ignored due to the air having low shear viscosity forces. Whether or not viscosity for a fluid can be considered a negligible force is determined by the Reynolds’ number, which is the ratio of the object’s momentum within the fluid to the shear viscosity forces of the fluid \([32][33]\). If the Reynolds’ number is much greater than 1, it is considered safe to be ignored. A relationship for this is given by

\[
\text{Re} = \frac{\rho u L}{\mu} = \frac{u L}{\nu},
\]

(1)

where \(\rho\) is the density of air, \(\mu\) is the viscosity coefficient of the fluid, \(\nu\) is the kinematic viscosity of the fluid (which can also be represented by \(\nu = \frac{\mu}{\rho}\)), and \(u\) and \(L\) are characteristic velocity and length scales of the fluid element flow \([34]\). In this case, the viscosity coefficient is approximately \(1.8 \times 10^{-5}\ \text{kg m}^{-1}\text{s}\), the density of air is close to \(1.2\ \text{kg m}^{-3}\) on average, and the minimum flight speed any of our UAVs can fly and still remain stable is \(13\ \text{m s}^{-1}\). In other words, no matter what is flown or what speed, it will be safe to ignore viscosity in every case because the numerator in the Reynolds’ number will always be far greater than the viscosity coefficient in the denominator.

Noting these assumptions, we start with the acceleration of a fluid element given by \([35]\)

\[
a = \frac{du}{dt} = \frac{\partial u}{\partial t} + u \cdot \nabla u,
\]

(2)

where \(u\) is the velocity field vector. The term \(\frac{\partial u}{\partial t}\) resembles the rate of change of velocity in reference to a locally fixed point, and the term \(u \cdot \nabla u\) resembles the changes in fluid element velocity as a result of its changing position in space. The next step involves applying Newton’s second law, but it is applied where acceleration and force are considered as acceleration per unit of volume and force per unit of volume for the fluid element. Upon taking the next step, two more variables are considered being fluid element density \(\rho\) and pressure \(p\). Density is more broadly defined as mass per unit of volume, while pressure is defined as force per unit of volume.
Looking at the convective term $\mathbf{u} \cdot \nabla \mathbf{u}$, the gradient of velocity is considered much less than the velocity vector in linear theory, so after applying Newton’s second law, the term is considered negligible and becomes the linearized equation of momentum given in equation (3).

$$\rho_0 \left( \frac{\partial \mathbf{u}}{\partial t} \right) = -\nabla p,$$

where $-\nabla p$ denotes the negative gradient of pressure (a force per unit volume term), and $\rho_0$ is the density of air and can be assumed constant because it is considered an incompressible fluid.

Another famous compressibility equation is the equation of continuity given by [35]

$$\frac{\partial \rho}{\partial t} + \mathbf{u} \cdot \nabla \rho + \rho \nabla \cdot \mathbf{u} = 0.$$  (4)

Because air is assumed incompressible, this implies that the density of air does not change in 3D space, which causes the $\mathbf{u} \cdot \nabla \rho$ term to vanish. It is then left with

$$\frac{\partial \rho}{\partial t} = -\rho_0 \nabla \cdot \mathbf{u},$$  (5)

where $\rho_0$ denotes a constant air density due to it being incompressible. Conceptually, this means that the change in air density over time is equal to the divergence of the fluid velocity negatively scaled by a density constant.

One variable that behaves well is the rotational part of velocity, called vorticity, which is given by the curl of velocity as shown in equation (6) [35].

$$\Omega = \nabla \times \mathbf{u}$$  (6)

Referring to equation (3), when the curl of velocity is taken, the vorticity does not change with respect to time because taking the curl of the gradient of pressure causes it to vanish. The reason this is beneficial is because that only leaves the irrotational part of velocity changing in time, and
that allows the velocity vector to be rewritten as

\[ \mathbf{u} = \nabla \phi, \]  

(7)

the gradient of a velocity potential \( \phi \) (or gradient of a linear displacement). Then, equation (8) can be gathered from equations (3) and (7) because the gradients of both sides equate to one another.

\[ p - p_0 = -\rho_0 \frac{\partial \phi}{\partial t}. \]  

(8)

In this expression, \( p - p_0 \) resembles the excess pressure caused by the acoustic source by subtracting the static atmospheric pressure \( p_0 \) from the total measured pressure \( p \). Additionally, equation (5) can now be rewritten as

\[ \frac{\partial \rho}{\partial t} = -\rho_0 \nabla^2 \phi. \]  

(9)

Pressure can be noted as a function of air density and can be expanded to a Taylor series. If the series is simplified to the first two terms in the series, after substituting relationships from equations (8) and (9), we are eventually left with

\[ \frac{\partial^2 \phi}{\partial t^2} = c^2 \nabla^2 \phi, \]  

(10)

where \( c \) is the speed of the wave (the speed of sound in this case). This is what many recognize as the wave equation. A simple solution to this for a plane wave in the \(+x\) direction could be modeled by

\[ \phi = f (x - ct) \]  

(11)

where \( f (x) \) is the waveform shifted by \( ct \). Conceptually, the derivative of this function would be the velocity of the traveling wave. If the derivative of equation (11) is evaluated at \( t = 0 \), this
would lead to equation (8) being simplified to

\[ p - p_0 = \rho_0 cu, \]  

(12)

However, to model a point source more precisely, the velocity potential is a function of 1D space and time (i.e. \( \phi = \phi(t, r) \)). If the Laplacian is taken of this, it is found to be

\[ \nabla^2 \phi = \frac{\partial^2 \phi}{\partial r^2} + \frac{2}{r} \frac{\partial \phi}{\partial r} = \frac{1}{r} \frac{\partial^2 (r \phi)}{\partial r^2}, \]

(13)

and after applying this to equation (10), the wave equation becomes

\[ \frac{\partial^2 (r \phi)}{\partial t^2} = c^2 \frac{\partial^2 (r \phi)}{\partial r^2}. \]

(14)

The new wave equation above has the general solution

\[ r \phi = f(r - ct) + g(r + ct), \]

(15)

but the concern, in this case, is the acoustic waves propagating outward from the source, so the \( g \) term goes away \[35\]. Comparing this with the spherically symmetric flow of incompressible fluid, \( \phi \) now takes the form

\[ \phi = \frac{-m(t)}{4\pi r}, \]

(16)

where \( m(t) \) is the rate of volume outflow, which is determined by the outflow cross-sectional area as given by

\[ m(t) = 4\pi r^2 \frac{\partial \phi}{\partial r}. \]

(17)

Because this simply refers to acoustic point sources, the cross-sectional area takes on the form of the surface area of a sphere (i.e. \( 4\pi r^2 \)). To account for the delay due to the speed of sound, the
velocity potential solution then delays the volume outflow by a factor of $\frac{t}{c}$ as given by

$$
\phi = -m \left( t - \frac{x}{c} \right) \frac{1}{4\pi r}.
$$

(18)

Although the above expression does provide a solution for the acoustic wave equation, velocity potential is not a quantity that is easily measured [35]. Excess pressure is a much more physically achievable quantity. With this in mind, excess pressure can be related to the velocity potential and volume outflow by

$$
p - p_0 = \frac{\dot{q} \left( t - \frac{x}{c} \right)}{4\pi r},
$$

(19)

$$
q(t) = \rho_0 m(t).
$$

(20)

The functions $q(t)$ and $\dot{q}(t)$ refer to the mass outflow and rate of change of mass outflow, respectively. The rate of change of mass outflow closely follows the excess pressure, so it is common for it to be referred to as the strength of the source. Noting that $\frac{\partial \phi}{\partial r}$ is another way of writing velocity $u$, if we set $4\pi r^2 = A$, equations (11) and (12) can then be used where $q(t)$ can then become

$$
q(t) = A \rho_0 u = A \frac{p - p_0}{c},
$$

(21)

and when the time lag is accounted for, this simplifies to

$$
p - p_0 = c \frac{q \left( t - \frac{x}{c} \right)}{A}.
$$

(22)

From the above model of excess pressure, the acoustic intensity is easily found because the intensity is defined as the rate of work done per unit area [35]. In other words, this also means it is an acoustic source with a given power over a certain area because a point source disperses energy over a sphere’s surface area. The intensity would normally be calculated simply by multiplying $pu$, but in the context of acoustics, this is performed on the excess pressure $p - p_0$, where the static pressure $p_0$ is removed and the acoustic dynamic pressure is all that is left. Other
relationships can be used as follows in equation (23).

\[ I = (p - p_0)u = \frac{(p - p_0)^2}{\rho_0c} = \frac{c q^2 (t - \frac{x}{c})}{A^2} \]  

(23)

The acoustic power is very similar to this as given in equation (24).

\[ W = \frac{c q^2 (t)}{\rho_0 A} \]  

(24)

A key difference is that it does not include the dissipative time lag because it is a source of power rather than an energy quantity like intensity that dissipates over time as it propagates. Moreover, intensity is power over a unit of area, which leads to a reduction of degree for the area in the denominator for the power quantity. Despite the derivation of acoustic intensity and source power for a point source being a long procedure, it leads to all formulas necessary for acoustic propagation being linearized to 1D linear equations. As seen in equations (23) and (24), the main variables that constitute the acoustic propagation characteristics are air density, the speed of sound, and the cross-sectional area, and this will be elaborated upon in the next few subsections.

**Determining Atmospheric Conditions**

From the equations above, it is seen that the differences in sound pressure vary with air density and the speed of sound. The atmosphere has been deemed incompressible when data is taken and the model is applied, but these constants do vary with atmospheric conditions. First, the pressure at which the atmosphere is completely saturated must be calculated, and this is a constant that varies with the temperature as given in equation (25) [36].

\[ \frac{p_{sat}}{p_{ref}} = 10^C \]  

(25)

\[ C = -6.8346 \left( \frac{T_{01}}{T} \right)^{1.261} + 4.6151 \]  

(26)

In this equation, \( T \) is the absolute atmospheric temperature in Kelvins, \( T_{01} \) is the triple-point
isotherm temperature constant 273.16 K, \( p_{s0} \) is the reference pressure, and \( p_{sat} \) is the saturation vapor pressure. Another pressure quantity needed is the measured barometric pressure, which we were able to measure with a weather station, but if one did not have access to a sophisticated weather station, many phone apps are available that are capable of measuring this adequately. To find the true air density and avoid as many approximations as possible, the barometric pressure \( p_s \) must be split up into its water vapor and dry air components. The partial pressure of water vapor can be found using the relative humidity (%) given by \[37\]

\[
P_v = h_r p_{sat}.
\] (27)

Then the dry air pressure component can be found using

\[
P_d = p_m - p_v.
\] (28)

A few more conditions must be calculated as well before the air density and speed of sound can be estimated. Equation (29) gives the absolute humidity (%) gathered from the saturation vapor pressure, barometric pressure, and the relative humidity (%) \[36,38\].

\[
h = h_r \left( \frac{p_{sat}}{p_{ref}} \right)
\] (29)

The absolute humidity \( h \) (now converted from a percentage to a decimal) is used to estimate the binary combined heat capacity ratio mixture \( \gamma_{mix} \) and the binary mean molecular mass mixture \( M_{mix} \left( \frac{kg}{mol} \right) \), both of which are essential for the speed of sound and air density \[36\]. The two relations are given below in equations (30)(31) with constants being the molar mass of dry air \( M_d = 0.02897 \ \frac{kg}{mol} \), the molar mass of water vapor \( M_v = 0.018016 \ \frac{kg}{mol} \), the heat capacity ratio of dry air \( \gamma_d = 1.4 \), and the heat capacity ratio of water vapor \( \gamma_v = 1.33 \).

\[
M_{mix} = h M_v + (1 - h) M_d
\] (30)
\[ \frac{1}{\gamma_{\text{mix}} - 1} = \frac{h}{\gamma_v - 1} + \frac{1 - h}{\gamma_d - 1} \] 

(31)

Finally, the density of the air \( \rho_{\text{mix}} \) and the speed of sound \( c_{\text{mix}} \) can be calculated using the universal gas constant \( R = 8.314462 \frac{\text{J}}{\text{mol-K}} \) and the measured temperature (converted to Kelvins) \[8, 36\].

\[ \rho_{\text{mix}} = \frac{p_d M_d + p_v M_v}{RT} \] 

(32)

\[ c = \sqrt{\frac{\gamma_{\text{mix}} \cdot R \cdot T}{M_{\text{mix}}}} \] 

(33)

Note that using \( M_{\text{mix}} \) is not allowed for air density but it is for the speed of sound. This is because the speed of sound is not a function of pressure. Also, if the gas of interest were not a mixture, the numerator in equation (32) would reduce to the single barometric pressure term \( p_s \) multiplied by the single gas’ molar mass \( M_{\text{gas}} \).

**Linear Propagation of Acoustic Sources**

As noted above, air and sound pressure behave in a nonlinear fashion when considering its velocity gradient in two or three-dimensional space. Again, we can ignore this due to the incompressibility behaviors of air along with ignoring viscosity due to a high Reynolds’ number. On the other hand, because pressure by itself has a nonlinear behavior, applying an attenuation model would be difficult. The issue can be simplified substantially by converting this one-dimensional acoustic pressure value to a more predictable quantity being acoustic intensity and acoustic power \[8\]. Both of these variables do not depend on the changing pressure, thus, making them more predictable. This is performed as shown in equation (34) below \[8, 39, 40\]

\[ I = \frac{(p - p_0)^2}{\rho_{\text{mix}} c}. \] 

(34)

This equation is recognizable from equation (23) from the wave equation derivation. Sound power, intensity, and pressure all have a range that behaves in a logarithmic manner. It is convenient to convert all of these variables to decibels (dB) for ease of interpretation. Sound
pressure level meters often display their readings in this manner as well. Conversions and
descriptions for all of these can be gathered from Table 1 [8][40].

Table 1: Sound Pressure Level vs. Sound Intensity Level vs. Sound Power Level

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Pressure</th>
<th>Intensity</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Units</td>
<td>SPL → dB_{SPL}</td>
<td>SIL → dB_{SIL}</td>
<td>SWL → dB_{SWL}</td>
</tr>
<tr>
<td>Reference Value</td>
<td>20 × 10^{-6} Pa</td>
<td>10^{-12} W/m^2</td>
<td>10^{-12} W</td>
</tr>
<tr>
<td>Conversion Formula</td>
<td>20log\left(\frac{p}{p_{ref}}\right)</td>
<td>10log\left(\frac{I}{I_{ref}}\right)</td>
<td>10log\left(\frac{W}{W_{ref}}\right)</td>
</tr>
<tr>
<td>Quantity Type</td>
<td>Field</td>
<td>Energy</td>
<td>Power</td>
</tr>
<tr>
<td>Proportionality</td>
<td>(p^2 \propto I \propto W), (p \propto \frac{1}{r}), (I \propto \frac{1}{r^2})</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Intensity is the preferred quantity for estimating the acoustic propagation characteristics
because it is an energy quantity at a certain point rather than a field quantity that is used for
multi-dimensional purposes like pressure. For this reason, it best identifies the loudness of a
signal universally as it propagates through the atmosphere. The modeling of acoustic point and
line sources rebound from this idea. The equation for the intensity of a point source as it
propagates through the atmosphere generally speaking is given in equation (35) [8][39][40].

\[
I = \frac{W \cdot Q}{4\pi r^2}
\]  

(35)

\(W\) denotes the source’s output acoustic power, \(Q\) denotes the directivity factor, and \(r\) denotes the
distance from the source. Intuitively, this can be seen more generally as power over some unit of
area. In the case of a point source, the area at which the sound propagates is spherical where its
area is given by \(4\pi r^2\). The directivity factor for an uninterrupted point source is 1, but this is
hardly ever the case in real-world applications. Oftentimes, the source propagates from the open
ground, from a wall, from a corner, or even from a structure like a hallway or a tunnel. Figure 6
gives a visual of how the directivity factor $Q$ is quantified in these cases.

Figure 6: Directivity factors due to propagation spread. Top left: $Q = 1$, top right: $Q = 2$, bottom left: $Q = 4$, bottom right: $Q = 8$ [4].

Additionally, if an acoustic source is behaving as a line source rather than a point source (ex. a highway or train), then the area at which it propagates begins to reduce by one dimension. In other words, the inverse square law of point source propagation gets reduced to strictly being inversely proportional (i.e. $\frac{1}{r^2} \rightarrow \frac{1}{r}$). The intensity is then modeled by the circumference from the line source rather than the spherical area and is described in equation (36) [8, 39, 40].

$$I = \frac{W}{2\pi r} \quad (36)$$

All of these relationships are valid estimators in short distances, but as the distance from the acoustic source is increased, other factors still need to be considered.
Spectral Atmospheric Propagation Model

The frequency spectrum becomes more important if the distance from the source is to be maximized. Spectral propagation is one of the most recent atmospheric propagation concepts present in this work and has been developed since 1990, has been closely analyzed and corrected, and recently left the focus of journals and made a presence in textbooks in the past few years [36,41–43]. It is well known that lower frequencies travel through solid, dense objects well while higher frequencies tend to either get absorbed or reflected. Atmospheric conditions play a major role in this because long-term changing variables like barometric pressure, temperature, and relative humidity change the air density and the speed of sound, which in turn affects how each frequency component behaves. The atmosphere is made up of 78% nitrogen, 21% oxygen, and 1% other gases on average [44–46]. These gases that make up the atmosphere have characteristic relaxation frequencies that vary with temperature and humidity. Generally speaking, relaxation frequencies are frequencies at which propagation losses are at a maximum, so this would have a significant impact on how certain frequencies are attenuated. The relaxation frequencies of nitrogen and oxygen (in Hz) can be found using equations (37–38) below [36]. Note that $h$ is once again expressed as a percentage, $T_{ref}$ is the reference atmospheric temperature of 293.15 K, and all other arguments assume the same units as previously mentioned.

$$f_{rN} = \frac{p_m}{p_{ref}} \left( \frac{T}{T_{ref}} \right)^{-\frac{1}{2}} \left( 9 + 280h \cdot \exp \left\{ -4.17 \left[ \left( \frac{T}{T_{ref}} \right)^{-\frac{1}{2}} - 1 \right] \right\} \right)$$ \hspace{1cm} (37)

$$f_{rO} = \frac{p_m}{p_{ref}} \left( 24 + 40400h \frac{0.02 + h}{0.391 + h} \right)$$ \hspace{1cm} (38)

As these are the biggest variables altering the spectral propagation, the attenuation coefficient (given in $\frac{Np}{m}$) can now be obtained using this and other previously quantified atmospheric
conditions.

\[ \alpha_{\text{air}} = f^2 \times 1.84 \times 10^{-11} \left( \frac{p_m}{p_{\text{ref}}} \right)^{-1} \left( \frac{T}{T_{\text{ref}}} \right)^\frac{1}{2} + \left( \frac{T}{T_{\text{ref}}} \right)^{-\frac{3}{2}} \]

\[ \times \left( 0.01275 e^{-\frac{\text{frO}}{f_r^2 + f^2}} + 0.1068 e^{-\frac{\text{frN}}{f_r^2 + f^2}} \right) \]  

(39)

In close distances, the additional attenuation across the frequency spectrum can arguably be deemed negligible, but as distance increases, this spectral attenuation model becomes much more necessary. The end goal is to be able to make detections as far as possible from the acoustic sources, so as scientists seek to achieve the furthest distances, this model assumes a more vital role.

As one observes the spectral attenuation model and relaxation frequency terms given in equations (37-39), it is clear that the relaxation frequencies are highly dependent upon the relative humidity, and the model as a whole is dependent upon temperature. The differences can be noticed in the graphs depicted in Figures 7 and 8.

Figure 7: Attenuation coefficient varying humidity for all frequencies per atmosphere of pressure at a distance of 100 meters.
Figure 8: Attenuation coefficient varying temperature for all frequencies per atmosphere of pressure at a distance of 100 meters.

The attenuation coefficient appears to increase at a linear rate when examined on a logarithmic decibel scale. Both humidity and temperature have similar effects on propagation across the spectrum. Humidity has a much more noticeable effect on the attenuation coefficient being flat across a large band of frequencies. It shifts to higher frequencies when the humidity increases. This is mainly because the relaxation frequencies of nitrogen and oxygen substantially depend on the absolute humidity, and the relaxation frequencies of both set the frequencies at which the attenuation is at a maximum.

One final fact worth mentioning is that extremely low humidities and temperatures have a negative effect on the lower frequencies. As a reminder, this project addresses seeking out targets of interest and identifying their threat status using acoustic sensors. Targets of interest for this application are going to be heavy vehicles and machinery usually possessing diesel engines. The dominant frequencies of these vehicles are rumbling noises as low as 50 Hz all the way to as high as 1000 Hz depending on the engine with a most common resonant frequency around 250-300
Hz. If machinery is so heavy that its dominant frequencies are on the lower side of the spectrum, this system could be prone to detection failures. With this in mind, the detection scheme or threshold should be chosen carefully for applications such as this; the climate and setting are always worthy of consideration for acoustic propagation.

**Doppler Effect**

Another topic to consider for filtering and drawing the detection boundary in future steps is the Doppler effect because it will affect what frequencies the acoustic sensors on the UAV will perceive as it flies near or over the acoustic source at a high speed. Ultimately, the Doppler effect describes a difference in an acoustic source wavelength and a perceived wavelength at the observer when either one or both of the parties are moving \[8\]. An equation that explains this behavior is

\[
f_o = \frac{c + v_o}{c + v_s} f_s,
\]

(40)

where \( f \) denotes frequency, \( v \) denotes velocity, and subscripts \( o \) and \( s \) denote the observer and the source respectively. A visual of this wavelength movement is given in Figure 9.

![Figure 9: A moving bell to the right of the image and its respective changes in wavelength](image)

Despite the speed of sound changing with atmospheric conditions, it commonly stays close to the mean value of 343 m/s. A T-72 tank on average travels at a speed close to 18 m/s \[47\], which is also
the average speed of the Albatross UAV as previously mentioned in section I. Assuming the objects are moving towards each other in a one-dimensional space, and the source is generating a 250 Hz wave, the observer would actually hear a 278 Hz wave. Then, this would transition to a 225 Hz wave as they crossed and began moving away from each other. Though this difference may not seem that large, it does help quantify the required tolerance of the filters and detection scheme that is applied. Too large of a tolerance on the system would lead to increased false positive detections while too small of a tolerance would lead to increased false negative detections.

**A, C, and Z Acoustic Weightings**

Before addressing the later steps of acoustic propagation, it is important to recognize the different acoustic weightings that are applied to SPL measurements. The C-weighting is the SPL filter at which humans hear and perceive a flat frequency spectrum at high volumes (i.e. environments approaching 100 dB\text{SPL} or greater), and its frequency response is given by \[ W_C = 20 \log \left( \frac{A_1 (f F_4)^2}{(f^2 + F_1^2)(f^2 + F_4^2)} \right). \] (41)

The A-weighting is like it, but it is the filter that simulates the human ear at lower normal daily volumes and is given by

\[ W_A = W_C + 20 \log \left( \frac{A_2 f^2}{\sqrt{(f^2 + F_2^2)(f^2 + F_3^2)}} \right). \] (42)

Constants for both of these equations include factors \( A_1 = 1.007152 \) and \( A_2 = 1.249936 \) along with cutoff frequencies \( F_1 = 20.598997 \) Hz, \( F_2 = 107.65265 \) Hz, \( F_3 = 737.86223 \) Hz, and \( F_4 = 12194.217 \) Hz. Values for \( A_1 \) and \( A_2 \) were picked such that the weighting was equivalent to 0 dB\text{SPL} at 1000 Hz. There is an intermediate weighting between these two called the B-weighting, but it has become obsolete and unused in recent years. Lastly, the Z-weighting is
technically no filter at all but is merely the name given to a flat 0 dB across the entire spectrum; this is what most electronics and circuits perceive at their inputs. A complete graph containing each filter’s frequency response is given in Figure 10.

![Figure 10: A, C, and Z weighting SPL curves](image)

From all of these weightings, most of the upcoming propagation relationships utilize the A-weighting when referring to SPL quantities or any other attenuation quantities. Also, note that from now on in this paper, whenever a value is referred to by the units of dB(A), dB(C), or dB(Z), it is referring to SPL.

### Wind Speed and SPL Relationship

Another challenging task is predicting how the SPL changes as the wind speed changes. Thanks to scientists specializing in wind and renewable energy, it is possible to calculate the total power of the wind traveling at a certain speed through a given cross-sectional area. This relationship is defined by \( P = \frac{1}{2} \rho A v^3 \), \((43)\)

where \( P \) is the total power obtainable from the wind (\( W \)), \( \rho \) again is the air density \((\frac{\text{kg}}{\text{m}^3})\), \( A \) is the cross-sectional area that the wind is passing through \((\text{m}^2)\), and \( v \) is the wind velocity normal to the cross-section \((\frac{\text{m}}{\text{s}})\). Alternatively, if for any reason only the difference in wind speed is known
along with an initial SPL level, a relatively scaled equation is given in equation (44) for this purpose \cite{50,51}

\[
SPL_2 = SPL_1 + 50 \log \left( \frac{N_2}{N_1} \right),
\]

(44)

where \(SPL_1\) and \(SPL_2\) are the old and new SPL levels of the wind, and \(N_1\) and \(N_2\) are the old and new wind speeds respectively. In order to calculate a prediction of what the overall signal-to-noise ratio (SNR) would be as distance changes, these relationships can be used to quantify the loudness of the noise in comparison to the signal that the UAV is trying to detect.

Then, to predict the loudness of the signal at a certain distance from the source, we can apply the source’s SWL \(L_W\), cross-sectional area \(A\), the attenuation coefficient \(\alpha \left( \frac{\text{dB}}{\text{m} \cdot \text{atm}} \right)\), and the distance \(r\) from the source to estimate the A-weighted SPL of the signal that it should be measuring. This is done by using the following formula in equation (45) \cite{8,52}.

\[
L_p = L_W - 10 \log A - \alpha r
\]

(45)

If this is evaluated at different distances, the average SNR can be predicted for any source that emits a certain power level with a certain directivity index in the midst of known wind speeds.

**SNR Estimation**

Estimating the SNR in audio applications is very difficult compared to doing so for other types of signals. In many other cases, SNR can be found by simply taking the mean over the standard deviation of the data or the squared mean over the variance. With audio though, despite the signal and noise being defined the same, it is interpreted differently. To elaborate, the signal is defined as the wanted part of the noisy signal, and the noise is defined as the unwanted part of the noisy signal. Intuitively, it is more difficult to estimate because what is considered the unwanted part of the noisy signal can vary by application and by each user. Also, this is especially more difficult because audio data is recorded as a one-dimensional time series; due to the lower dimensionality, it is even harder to break apart the core characteristics of the signal. Another way
that SNR can be defined though is by taking the ratio of the power of the signal over the power of
the noise given by [53]

\[
\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}} = \frac{s^2}{E[N^2]} \tag{46}
\]

\[
\text{SNR}_{\text{dB}} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) = P_{\text{signal, dB}} - P_{\text{noise, dB}} \tag{47}
\]

Note that if the noise is considered a random variable, the expected value of the square of the
noise (or the variance) is assumed to be the noise power. Doing this allows SNR to be found
experimentally where the user can define what they consider to be part of the signal and part of
the noise through the data samples that they collect.

In the case of this project, the characterization of the SNR specifically was done at a wind
tunnel. To ensure that every part of the noise was accounted for, three different kinds of signals
were obtained, which were a noisy signal with the signal source playing and wind engaged, a
noise recording of just the wind, and a signal recording of no wind but the signal was playing. All
three different algebraic combinations of ratios were taken for comparison and optimal estimation
because SNR is difficult to obtain in applications like this when the noise is much greater than the
signal due to the extreme amplified randomness of the noise. These three variations are given in
equation (48)

\[
\text{SNR} = \frac{S}{N} = \frac{1}{\frac{S+N}{S} - 1} = \frac{S+N}{N} - 1, \tag{48}
\]

where \( S, N, \) and \( S+N \) can be substituted for the power of the measured signal, noise, and noisy
signal respectively.

Using the ratio of the measured signal to the measured noise explains the SNR purely based
on the wind being the dominant noise of interest, which is beneficial due to the fact that wind is
the main noise that these acoustic sensor mounts seek to alter. Using the signal and noisy signal
variation explains the exact same thing; however, taking the SNR variation involving the noise
and noisy signal accounts for all noises in the process (being wind, electromagnetic interference
(EMI), electronically generated noise, and noise from echos in the room) because it does not
define a signal ground truth recording. In other words, it captures all the noises in the noise
recording and compares this to the noisy signal where the only difference should be the signal component, so in this case, the signal should theoretically not contain the other recorded noises like room echo, EMI, and electronic noise. Taking the system-based approach as in equation (49) acknowledges this as well as it only compares the true signal component that was sent through the source to the noisy signal.

\[ P_{\text{noise}} = P_{\text{output}} - P_{\text{input}} \]  

(49)

\( P_{\text{output}} \) denotes the power of the system output (the noisy signal in this case), \( P_{\text{input}} \) denotes the power of the system input (the true signal being sent through the source), and \( P_{\text{noise}} \) is the proposed estimate of the noise power. All four of these SNR approaches were taken and all estimations (aside from occasional spectral SNR estimations that seemed like extreme outliers) were averaged together to achieve the best result. The signal and noise variation and the signal and noisy signal variation consistently gave trustworthy results, but there were occasions for the other two variations where the spectral SNR estimation was far from the other results. The reason for this will be elaborated upon in the results section when SNR is modeled with distance.

Filtering Techniques and Data Transformations

Matched and Wiener Filtering

Many noise reduction techniques were applied to the noisy signal to get an improved version, but two methods that performed the best spectrally speaking are matched filters and Wiener filters. Both of these are in many textbooks and are common in optimal filtering theory. The main goal that matched filters seek is to maximize SNR and are commonly used for detection problems, especially with binary channels [54]. On the other hand, Wiener filters seek to minimize the mean square error (MSE) between the estimated and desired processes, so they are used often for reducing noise. Both of these would be beneficial in making the spectrums more separable so that accurate detections can be made, and both share similar principles in that they
assume the same noisy signal model given by

\[ y(t) = x(t) + n(t), \quad (50) \]

where \( y(t) \) is the noisy signal, \( x(t) \) is the signal, and \( n(t) \) is the noise.

Wiener filtering is a useful strategy for denoising signals due to its principle of minimizing MSE \[55\]. The derivations for estimating MSE are covered throughout literature, but simply speaking, if the signal plus noise model from equation \( (50) \) is assumed and the signal and noise are independent of each other, then the following transfer function in equation \( (51) \) is obtained to optimize the output spectrum.

\[ H(\omega) = \frac{S_{xy}(\omega)}{S_{yy}(\omega)} = \frac{S_{xx}(\omega)}{S_{xx}(\omega) + S_{nn}(\omega)} \quad (51) \]

In this transfer function, \( S_{xx}(\omega) \) is the power spectral density (PSD) of the signal, \( S_{nn} \) is the PSD of the noise, and \( S_{yy}(\omega) \) and \( S_{xy}(\omega) \) are the PSDs of the system input and output respectively, or the PSD of the noisy signal \( y(t) \) and the PSD of the filtered noisy signal that should closely resemble the signal \( x(t) \).

Matched filtering can be an effective approach to signal detection if the signal is known on the receiver’s end and an additive Gaussian white noise model is assumed \[54\]. Essentially, using the same noisy signal model in equation \( (50) \), matched filtering uses the known signal or an estimate of the signal, and it uses the time-reversed conjugate of it to match the response of the known signal so that only the signal is present when being extracted from the noisy signal. There are many books that derive this in detail, but to be brief, the impulse response and transfer function for the filter are given in equations \( (52) \) and \( (53) \) respectively.

\[ h(t) = kx^*(t_0 - t) \quad (52) \]

\[ H(\omega) = kX^*(\omega) \exp(-j\omega t_0) \quad (53) \]
In these characteristic equations, $k$ is a scaling factor, $x^* (t_0 - t)$ denotes the conjugate of the time-shifted signal in the time domain, and $X^* (\omega)$ is the conjugate of the Fourier Transform of the signal. Then, these filters would be applied by either convolving $x(t)$ with $h(t)$ in the time domain or by multiplying $X(\omega)$ with $H(\omega)$ in the frequency domain. One downside to this as it stands though is that it assumes white Gaussian noise. With that said, this can still be used on any signal as long as it is first spectrally whitened, which tends to lead to a slight degradation of the noisy signal [56]. Spectral whitening makes the noise spectrum flat across the entire spectrum (as is the form of white Gaussian noise) and is a procedure that closely resembles finding a unit vector. The difference is that the Fourier spectrum is multiplied by its complex conjugate instead of just being squared as indicated by equation (54).

$$X_{\text{white}}(\omega) = \frac{X(\omega)}{\sqrt{|X(\omega)X^*(\omega)|}}$$

(54)

An alternative to this though is if a colored noise is assumed. However, this is not always an effective approach compared to spectral whitening because it is uncommon for an exact colored noise to be known. With this approach, the transfer function can now be defined as

$$H(\omega) = k \frac{X^*(\omega)}{S_n(\omega)} \exp(-j\omega t_0)$$

(55)

where the transfer function is almost the same but is divided by the PSD of the noise $S_n(\omega)$ [55].

The noise in the case of this thesis does keep a constant dB per decade roll-off rate like most colored noises, but it is an extreme decline steeper than brown noise because there is a much greater presence in the lower end of the spectrum associated with wind turbulence. The principles of estimating the PSD of the noise though for this updated transfer function should still apply.

**Harmonic Spectral Transform**

After optimally filtering the noisy signal, the Harmonic Spectral Transformation (HST) was then applied to make the signal and noise more differentiable. The HST was proposed by Harvey
to exploit spectral peak periodicity to better locate the signal within uncorrelated noise, and his work builds upon the Harmonic Product Spectrum (HPS) and Harmogram presented by Hinich and Schroeder, respectively, to develop multi-channel and multi-window realization signal processing techniques applicable to spectral characteristics [19, 57, 58]. The HST is very similar to a Fourier Transform, but instead of strictly transforming to each frequency component, in its simplest form, it takes the $R$ number of harmonics of each frequency component and performs a form of statistical averaging over the harmonic peaks, and the result of each of those iterations would become the transformation’s ending values as a function of the input Fourier fundamental frequencies [19]. A formula for this is given in equation (56)

$$
\mathbb{H}_a[X(f)] = \left[ \frac{1}{R} \sum_{r=1}^{R} |X(f \cdot r)|^a \right]^{\frac{1}{a}},
$$

where $R$ is the number of harmonics to be averaged (inclusive of the fundamental frequency) and $a$ is the form of statistical mean applied, both of which are determined by the user beforehand. Each fractional harmonic peak that is averaged can be given in the set given by

$$
f_{a,b} = \frac{af_0}{b}, \quad a, b \in \{1, 2, \ldots, R\}
$$

where $f_0$ is the fundamental frequency of interest. Also, because HST operates on harmonic multiples of the fundamental frequency, the resulting spectrum that can be observed after the transformation is limited by the sampling frequency $f_s$ such that $[0, \frac{f_s}{2R}]$ is the maximum observable range of the spectrum. Different types of statistical averaging are capable of being applied to this transformation, all of which are given in Table 2.
Table 2: Applicable types of statistical averaging to Harmonic Spectral Transform (HST).

<table>
<thead>
<tr>
<th>Harmonic</th>
<th>Geometric</th>
<th>Standard</th>
<th>RMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = -1$</td>
<td>$a = 0$</td>
<td>$a = 1$</td>
<td>$a = 2$</td>
</tr>
<tr>
<td>$\left[ \frac{1}{R} \sum_{r=1}^{R}</td>
<td>X(f \cdot r)</td>
<td>\right]^{-1}$</td>
<td>$\prod_{r=1}^{R}</td>
</tr>
</tbody>
</table>

Although this may vary with different signals and their respective spectral periodic structure, using the standard mean has proven to give the best results in the case of this project.

The simplified HST can be effective alone, but it can be made more precise if this is applied over multiple signal channels [19]. Beyond that, it can be improved even further by applying the same thing over multiple time realization windows. By applying this again for both of these improvements, equation (56) then expands to the generalized harmonic spectral transform (GHST) given by

$$
\overline{H}_{a,b,c}[X_{s,w}(f)] = \left[ \frac{1}{W} \sum_{w=1}^{W} \left[ \frac{1}{S} \sum_{s=1}^{S} \left[ \frac{1}{R} \sum_{r=1}^{R} |X_{s,w}(f \cdot r)|^{a} \right]^{\frac{1}{b}} \right]^{\frac{1}{c}} \right],
$$

where $S$ denotes the number of signal channels, and $W$ denotes the number of time realization windows. Note how different statistical means can be applied at each step. To explain this in more applicable terms, the first step of averaging over the harmonics would be similar to taking the simple HST of a single periodogram of a single-channel signal obtained by using Welch’s method (but leaving the result in terms of magnitude rather than power). Then, expanding this to multiple signals would be similar to averaging that result over the multiple periodograms obtained from multiple channels. Lastly, to expand that to multiple time-realized windows would be similar to expanding those multiple periodograms into spectrograms that each show multiple time instances of the instantaneous spectrum, and then averaging the HSTs of each of those frequency spectrums for each windowed time realization. A finer result can be obtained from expanding out to these
forms.

**Detection Scheme**

The above three noise reduction techniques were used to help process the data, but then for the detection scheme to follow, this thesis takes an uncommon approach to statistically model the detection predictions. Because the signal of interest is a diesel engine, it has a rather widespread complex harmonic spectrum versus just a pure sine wave tone, meaning that using a peak detection algorithm on this representation is not going to work well, and receiver operating characteristics (ROC) are not easily quantified. Noting that a diesel engine varies across the spectrum, detection hypotheses will be made based on regression metrics such as mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE) by comparing how the noisy signal harmonic spectrum compares to the signal spectrum and the noise spectrum. One major advantage to doing this is that not only would a system like this be able to detect a diesel engine, but it will also be able to specify specifically what vehicle’s engine is present because no two engines’ harmonic spectra are alike. Then for further evaluation, the results will cover how these metrics behave as a function of distance from the acoustic source. A benefit that arises from this approach is that this compares the noisy signal to a specific characteristic harmonic spectrum. In theory, this detection system can be taken a step further and be able to identify specific targets according to their characteristic harmonic spectra, rather than simply identifying that something is present.

To elaborate on the detection process, each frequency sample in the harmonic spectrum of the noisy signal would be compared to that same frequency sample from the signal spectrum and the noise spectrum and the error from each of those points would be calculated. Signals can vary by application though and some things that need to be considered are behaviors such as whether or not the spectrum is widespread or has focused tones, if the noise has a low or high variance, or if the spectrum is very dense. All three of the previously mentioned regression metrics are considered because picking a metric for a certain signal is not a one size fits all application.
Mathematical equations for these regression cost functions are given in equation 59.

\[
\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|, \quad \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}, \quad \text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2
\]  

MSE looks for outliers in the data and amplifies that error to correct the system more for the points being far off, whereas MAE develops an error with more inertia that observes smaller errors and outliers on a closer scale, and RMSE is the middle ground for these other two metrics. Keeping the signal of interest under consideration, if the objective signal for this project were to be a sine wave, MSE would be a great choice because the error would draw a better boundary between the few spectral peaks, but because a diesel engine is of focus, either RMSE or MAE in some cases is a better choice because the differing peaks are going to be widespread.

Building upon the regression metrics, the decision principles are rather simple but require some intuition and data analysis for an effective decision estimate. As a background, an example harmonic spectral transform plot is shown in Figure 11 where a diesel engine recording was taken at the wind tunnel, and the data has been normalized along with the noisy signal being processed through a Wiener filter.

![Harmonic Spectral Transform](image)

**Figure 11**: HST of flat mount diesel recording in 20 m/s winds.
It should be pointed out from the plot that in addition to having a low-frequency presence of around 150 Hz, the diesel engine had more high-frequency presence than expected around 1-1.2 kHz as well. Moreover, the noise generated from the wind has a mostly low-frequency presence that dwindles off as frequency increases. This is indicative of a great opportunity for signal and noise separability, and it demonstrates that signals that show any form of periodicity are prone to having a higher realizable harmonic presence, even if the signal is notorious for having a low-frequency presence.

Additionally from Figure [F11] the overall magnitude of each spectrum resides at a very low level with few harmonic peaks arising wherever present. This is important to take notice of because if one of these mean error metrics were taken of the entire spectrum, this could hurt the detection performance because it would give too much weight to the oscillating lower magnitude values across each signal’s spectrum when they are all in all unnecessary and just resemble noisy, percussive, and residual components. With that said, the decision is based on whether or not the chosen mean error metric of the noisy signal and signal is less than that of the noisy signal and noise; however, an additional constraint is added where only frequency components that go above the standard deviation of the noisy signal are considered. The magnitude spectrum for each has been normalized and has a mean of 0, so the standard deviation is the only necessary characteristic to account for in the random process and improve detection rates. An algorithm describing this decision function is given by the following.
Algorithm 1 RMSE Detection Hypothesis Algorithm

Ensure: $\sigma \leftarrow$ noisy signal standard deviation
Ensure: $\alpha \leftarrow$ scaling factor
Ensure: $y, x, n \leftarrow$ noisy signal, signal, and noise HST
Ensure: $L \leftarrow$ length of $y$
Ensure: $c \leftarrow$ mean counter
Ensure: $\varepsilon_x, \varepsilon_n \leftarrow$ noisy signal error with signal and with noise

$\varepsilon_x, \varepsilon_n, c \leftarrow 0$

for $i \leftarrow 1, L$ do
  if $y_i > \alpha \sigma$ then
    $\varepsilon_x \leftarrow \varepsilon_x + (y_i - x_i)^2$
    $\varepsilon_n \leftarrow \varepsilon_n + (y_i - n_i)^2$
    $c \leftarrow c + 1$
  end if
end for

$\text{RMSE}_x \leftarrow \sqrt{\frac{\varepsilon_x}{c}}$

$\text{RMSE}_n \leftarrow \sqrt{\frac{\varepsilon_n}{c}}$

if $\text{RMSE}_x < \text{RMSE}_n$ then
  detect $\leftarrow$ True
else
  detect $\leftarrow$ False
end if

The scaling factor $\alpha$ is one that is optimally chosen for the user’s signal of interest and noise level and is estimated through experimental detection trials. For this procedure, a scaling factor of 1 was optimal for Wiener filtering and a factor of 4 for matched filtering. When these filtering techniques were cascaded, this scaling factor became less critical with a general range of 1-4.

Data Collection Strategy

The strategy for collecting data for this thesis was threefold. In one setting, the microphones and mounts were characterized within an enclosed 4’ × 2’ × 2’ box where the walls were 2” sound absorbing fiberglass acoustic panels. In another setting, a wind tunnel was used for testing the response and SNR for every acoustic sensor mounting design. This was done without wind as well as at various wind speeds so that true signals, true noise, and true noisy signals for different wind speeds were obtained. Filtering and the detection algorithm were also practiced in this stage,
but this is not an entirely realistic setting to test whether or not detections can be made at various
distances. The last setting was in actual test flights at Agricenter International, which is an
international agricultural research center. Unfortunately, the Albatross fixed-wing UAV that was
the aircraft of choice for this application was not available at the desired time of data collection.
An Angel vertical takeoff and landing (VTOL) UAV provided by Dunlevy Consulting was used as
an alternative. An image of the UAV is given in Figure 12.

Figure 12: Angel VTOL UAV used for flight data collection.

The Angel is an easy to set up yet dependable design that takes off with four rotors and then
transitions into a fixed-wing mode with a propeller located at the rear. Average air speeds for this
UAV reside near 25 $\text{m/s}$, and engine, servo, rotor, and electronic noises are comparatively louder on
this aircraft than the Albatross UAV. Despite these drawbacks, it is still a remarkable aircraft for
many applications, and it was more rugged and easier to maneuver in the case of this project. The
mounts which seemed to have the most potential were mounted to the Angel and were used for
test flights.

Atmospheric Conditions and Experiment Setting Data

As can be seen from the atmospheric model in previous sections, there are a few atmospheric
conditions that need to be gathered. The data that is needed for this includes temperature, relative
humidity, average wind speeds, average wind gusts, average wind direction, barometric pressure
(or elevation instead, but this is less preferred because it would be an unnecessary estimation), and
if possible, reference pressure. Our team has a portable weather station that records all of this data at the intervals specified; however, if one were to replicate this experiment, only average values at the time of taking flight (or making any recordings) are needed, so most temperature/humidity gauges or phone weather and barometric pressure measurement apps would suffice. All of this information was gathered for both the wind tunnel data collections and for all UAV test flights.

Figure 13: Weather station sensing device capturing atmospheric data at the Agricenter.

Another important piece of data needed for every experiment setting is location data. This is absolutely necessary to calculate the propagation losses as shown in equations (35) and (45). For the wind tunnel, this is a simple task because only the distance from the speaker at the front of the wind tunnel to the microphone needs to be known. When flying any UAV, GPS location data was needed for the UAV as well as for the location of the acoustic source so that the distances at which the UAV detected the signal of interest could be known.

Audio Samples

Planning the recordings that needed to be made for each mount required structure to achieve the best results. The two main goals of this project were to improve the SNR and to improve detection performance. Beginning with the improvement of the SNR, as mentioned in earlier sections, measuring this for audio can be a difficult process compared to other applications. The best approaches to obtaining the SNR were by getting recordings of the signal by itself, the noise by itself, and the noisy signal and taking the ratio of their power spectral densities (PSD).
Technically only two of these are required to estimate the SNR, but we gathered all three wherever possible to take all three ratio combinations and ensure that the results agreed with each other.

Figure 14: Acoustic sensor wing mount being tested in the wind tunnel.

In taking all three of these recordings, this meant that recordings of the signal with no wind, the wind by itself, and the wind-infused signal needed to be recorded. This was done for every test signal and level of noise desired to be analyzed. The types of signals that were tested started with individual sine waves at 50 Hz, 250 Hz, and 1000 Hz (with 250 Hz being the largest frequency of interest because of its close resemblance to a diesel engine), then frequency sweeps were done to observe any changes in PSDs and to take the SNR across the spectrum, and then diesel engine recordings were gathered to test detection capabilities. A 120-watt recording studio speaker with a flat frequency response was used for playing all of these signals. This was all simulated at different wind speeds corresponding to the specifications of the Albatross, including average speeds of 15 m/s and 18 m/s, the above average speed of 20 m/s, and the slower speed of 13 m/s for temporary SNR improvement. Afterward, all signals were extensively characterized using a homemade web app made with the Python programming language and its many open-source signal processing packages.

As for the UAV test flights, making detections was of the highest interest. The signals used in the field consisted of the 120-watt studio speaker playing the same diesel engine recording mentioned before along with air horns to have some signal variety in case the diesel engine was hard to detect in the sub-optimal conditions.
Fluid Dynamics Inspired Wing Mount Design

When in a moving system such as a UAV flying at a high speed, changes in pressure are constantly being made. Pressure can be affected by the atmosphere via air density and speed of sound as mentioned in equation (12). To relate fluid mechanics to electricity, pressure shares a high similarity with voltage. Just as a difference in voltage causes an electrical current, a change in pressure causes the volumetric flow of a fluid element. In both situations, the flow or current will continue to occur until a state of equilibrium is reached in the difference in pressure or voltage. Not only is this given in fluid dynamics by the Hagen-Poiseuille equation that describes volumetric flow \( Q \) as

\[
Q = \frac{p_1 - p_2}{R}
\]  

(60)

where \( R \) is the resistance to flow (viscosity in this case), but it is also supported in chemistry by Le Chatelier’s Principle, which also declares that the differences in atmospheric pressure will constantly be trying to reach a state of equilibrium [34][62]. With this in mind, where high pressure and low pressure are quickly divided, it is possible that a flow inrush of fluid elements could occur in these abrupt differences. The microphone wing mounts designed for this project puts this idea to the test to observe in what configurations the wind will mostly be avoided to allow a stronger signal to enter the sensor.

Each mounting idea that came to mind was 3D modeled, 3D printed, and then tested in the wind tunnel, and a select few of them were flown on UAVs later on. The acoustic sensors were embedded on each of these with only a small hole to allow the signal through so that the power of the wind would be minimized at the input as given by equation (43). After this, different variations were made around the input holes such as putting a wind-blocking ramp right by the input hole, sinking the holes into a cavity, making those cavities tighter, wider, and deeper, and keeping the mount flat with no surrounding mechanical differences at all. Some examples of these designs are given in Figure [15] but all of them are depicted more in detail in the appendix.
Figure 15: A few different variations of the tested acoustic sensor wing mounts.

Involving Helmholtz resonators on the interior was also a consideration; however, if this were attempted to have a resonant frequency close to 250 Hz (i.e. the average frequency of a diesel engine), the acoustic sensor mounts would become large and would take up more space than is available on the UAV. Wind foam filters were another major consideration. For some tests, pieces of cotton balls were used to try to serve as a windscreen for the microphones.

The holes on the mounts that allow sound through to the acoustic sensor do require special consideration. As was described in equation (43), the power that the microphone will extract from the wind is directly proportional to the cross-sectional area. Therefore, if the acoustic entry points are made smaller, the noise will proportionally decrease. Although this does help reduce the wind noise, it will not help as much as slowing down the UAV. Once again, the power of the wind is proportional to the cube of velocity, so this is where the greatest improvement will be seen.

Upon designing these shapes, the flat mount was anticipated to be the best-performing design. With all of the other designs, despite attempts to block the wind, cavities and ramps cause abrupt changes in pressure. These abrupt changes in pressure are likely to seek a state of equilibrium. In other words, the cavities and ramps would cause fluid elements of air with lower pressure right behind their walls, and the air with high pressures passing by would approach these low-pressure fluid elements and more flow could be introduced into the inputs due to these differing pressures trying to stabilize.
4 RESULTS

As there were three different settings where data was collected, being an acoustic box, a wind tunnel, and by flying a VTOL UAV, the results can be broken down and analyzed in a similar manner. Data samples were collected in an enclosed fiberglass acoustic box of certain signals such as sine waves and frequency sweeps and will be briefly discussed to analyze all necessary characteristics of each mount. The performance of the wind tunnel data will be evaluated by observing the SNR improvements for each mounting design, and then the flight data will be evaluated by applying the statistical detection scheme and examining at what distances the vehicle detections are able to be made and under what statistical constraints. Atmospheric conditions were also gathered such as humidity, temperature, barometric pressure, and GPS data so that all distance, propagation, and acoustic intensity decay relationships could be fully evaluated and map SNRs as a function of distance. In terms of statistical performance, this thesis seeks a different approach to evaluating regression metrics (such as mean square error) as a function of distance.

Lab Experimentation Results

Characterization of the Mounts

Beginning with the characterization details of these mounts, data collected for this was done either in an acoustically optimal box or in a wind tunnel depending on whether spectral characteristics were being mapped or temporal characteristics of the wind were being exploited. For signals being played from the acoustic source, a diesel engine is the target signal of interest and usually has an SPL close to 100 dB with more presence in the lower end of the spectrum around 250 Hz. With this in mind, all signals (including those beyond a diesel engine like frequency sweeps and sine waves) were played at an SPL of 100 dB for consistency. Additionally, all measurements from the acoustic box were done from 0.75 meters away whereas the acoustic sensors and mounts were around 2.5 meters away when being tested at the wind tunnel. All distances were taken into account when evaluating the signal characteristics and atmospheric
propagation model.

Many of the mounts spectrally share similar traits with only marginal differences as shown in Figure 16.

![Figure 16: Comparison of spectral response for all multi-channel mounts.](image)

In terms of keeping a flat spectrum, the flat mounts are the best at maintaining this behavior. When adding cotton to the mount as a turbulence filter for the mic, this applies a squelching all the way across the spectrum, which is not as intuitive as many would expect. It is fascinating that it does not allow more lower frequencies where only the higher end is reduced when applying this, but the thickness of the cotton filter is a large factor in this as the range that it reduces is in the 200-1000 Hz low-mid range. All of the mounts stay close in the high-mid range of 1000-2000 Hz; where the other mounts really differ is in the low and higher end of the spectrum. Many of them have a small variance in magnitude, which could be a result of errors in data collection or noise, but all of the mounts aside from the flat one behave as a band pass filter allowing more of the mid-range in comparison to the highs and lows. Some of these mounts apply this filter differently than others by slightly increasing the order or quality factor, but this is due to the level of obstruction as it does appear that the mounts with the deeper cavities apply this reduction more
as opposed to the shallow ones. Overall, the lesser the obstruction is, the more flat response will be obtained, and in regards to this project, more of the lower end will be returned.

**SNR Performance**

Considering the spectral characterization in previous sections, SNR does follow somewhat closely to the same behavior, but there are some differences. The ramps were expected to have the poorest performance, and this is reflected in the spectral SNR as given in Figure 17.

![Figure 17: Comparison of spectral SNR for single-channel ramp and flat mounts in 13 m/s winds.](image)

The reason for this expectation is that adding a wind-blocking ramp causes sudden low pressure where the mic would be located, and that would lead to lots of turbulence near the mic acoustic entry point due to the pressures trying to reach a state of equilibrium as discussed before. A wall or ramp would eventually lead to a stagnance in the air-fluid element flow right behind it if the wall was big enough, especially at higher speeds where the wind would have more inertia. In reality though, coming up with such a large blocking cross-sectional area would not be feasible to mount on a UAV unless the microphone wind-blocking device took the form of a bullet and the mic was embedded so deep that the wind would have too much inertia to reach the sensor. This
would also be difficult to pick up the signal if it were embedded in this way and would require optimization. However, one positive to the lower, more subtle ramp mount is that it achieved marginally better performance in the higher frequencies. In the case of this project, this is not worth the trade-off of poorer lower frequency performance, but for other applications where other higher frequency signals are being sought out, it could be useful. It is important to note that only having a single channel is much less reliable because it leads to higher variance as compared to multiple channels and gives less information in the post-processing and detection stage.

Moving on to the multi-channel mounts, Figure 18 shows the lower end of the spectrum linearly centered around 500 Hz is treated as roughly unity gain. It is seen that there is a boost in performance around 1.2-1.7 kHz, and then the clarity rolls off beyond that while having a high variance in every case.

![Figure 18: Comparison of spectral SNR for all multi-channel mounts in 13 m/s winds.](image)

It is also evident that the shallow cavities obtain better results than the deeper cavities as do the wide mounts over the tight mounts with borders closer to the acoustic entry hole. In fact, the tight cavities either perform roughly the same or worse than having no mount at all, whereas the wide cavities marginally perform better, and the flat mounts clearly show the best performance by
nearly a 10 dB difference all across the spectrum. Again, as one may hypothesize, this is explained by the abrupt changes in pressure due to the obstructions of airflow as mentioned before. Due to the boost in the mid-range frequencies, mount designs such as these would clearly perform better if the signal of interest’s presence was slightly higher in the spectrum. One intriguing characteristic that might be hard to spot is that there is less low-end spectral roll-off when cotton is not used. For the entire SNR spectrum, cotton does not change the SNR within the range of interest as opposed to the characterization response spectrum. The only difference is a harder low-end roll-off, which is not in favor of the signals that this project seeks to detect. For this purpose, the acoustic sensors would be better off without it. However, if one wanted to exploit this further, the pieces of cotton used were very dense, so finding a less dense screen material could possibly serve as a stronger turbulence filter and improve the SNR.

Naturally, if the SNR is considered on the dB scale, the SNR does decrease linearly as the wind speed increases. This trend occurs for every acoustic sensor mount as well, no matter what form of wind obstruction was used. An example of the flat mount with cotton is given in Figure [19] for wind speeds that are characteristic of the Albatross UAV and its fastest average speed all the way down to its slower idle speed.
From the figure, it is clear how beneficial it is to have flexibility in flight functions. Being able to slow all the way down to 13 $\frac{m}{s}$ can result in approximately a $+5$ dB difference in SNR purely based on the wind speed contribution to the noise. Furthermore, the Albatross is much quieter in terms of engine and rotor noise, so its contribution to the noise would be even less as well, thus leading to a stronger hit probability detection result. Building on this, results similar to Figure 18 are given in Figure 20 where the wind speed is increased to 20 $\frac{m}{s}$. Though the SNR across the spectrum does assume basically the same shape, achieving positive SNRs is almost impossible for the desired frequency bin of the signal of interest.
Figure 20: Comparison of spectral SNR for all multi-channel mounts in 20 m/s winds.

In-Field Flight Results

SNR and Distance

As spectral SNR is a valuable characteristic that was evaluated in the wind tunnel for each mount, taking the average SNR and mapping that to distance to help quantify how far away detections can be made is an important next step. When coming from taking data from the wind tunnel and engaging in an actual flight, additional elements of complexity are introduced because the wind can no longer be considered a single component going in a single direction and the distance from the source is no longer fixed. The procedure for estimating the SNR is similar to what was mentioned before in the methodology in equation (48). The unfortunate fact for flights is that only the noise and noisy signal can be easily obtained, so the noisy signal to noise variation has to be taken. The reason this is not ideal is that the noise extremely overpowers the signal in this application, and dividing the noisy signal by the noise (which is close in power across the spectrum) can result in a poor estimate of the signal and SNR. The goal here is to take the average SNR and compare it to distance, so taking the mean of the SNR across the spectrum will lead to
an estimated SNR with a lesser standard deviation. Then, by using formulas from the spectral distance propagation model previously discussed, the values for SNR at every other distance from the source can be expanded from that.

Atmospheric conditions are necessary for mapping this relationship. Some of these conditions include relative humidity, barometric pressure, reference pressure, temperature, wind speeds, wind direction, and GPS location data to be able to locate where the UAV is in reference to the source. Many of these are constants that are able to seamlessly be plugged into equations, but it is necessary to consider 2D and 3D vectors when considering the net wind speed and the distance from the acoustic source from the perspective of the sensor. After these considerations, they can be converted to scalar quantities by taking the net wind speed passing by the microphone and the Euclidean distance from the source. An approach to determining the total power of the wind that the microphone should be experiencing is by taking the air speed of the UAV and adding the necessary velocity vectors of the wind speed and direction and finding the magnitude. Then, this can be applied to equation (43) where the wind power can be calculated and converted to an SWL as can the propagated signal using the intensity and cross-sectional area principle.

Though the flight conditions were not ideal, calculations were made easy due to the flight patterns that the Angel took. Taking a look at Figure 21 near the top of the image in the brown plot of land where the orange and green highlighted paths meet, the UAV took off from there moving along the orange path directly against the $6 \frac{m}{s}$ wind that was coming out of the NNW (the right side of this bird’s eye view image).
The UAV then found its path and loitered in the thick orange circular clockwise pattern for 10 loops before it took the sharper circle and returned to the point from where it took off. The acoustic source that was propagating the sounds of a diesel engine was located on the dirt road at the very top of the loitering circle. As mentioned before, the UAV took off directly against the wind, and as it passed by for every loop, it was passing the source right as it was directly flying into the headwinds, so this led to an easy net wind velocity vector calculation. As the Angel passed the source, it held the air speed of $25 \text{ m/s}$ and an altitude of $40 \text{ m}$, and the direct opposing wind speed was $6 \text{ m/s}$ leading to a net wind speed of $31 \text{ m/s}$. With this net wind speed and altitude, the recorded SNR values before any electronic filters were applied are plotted in Figure 22 along with a trend curve for these points, the mathematical model prediction of the SNR for the Angel flight, and the model prediction of the SNR if ideal conditions with the Albatross were assumed.
Figure 22: Measured Angel SNR as a function of distance compared to ideal conditions with Albatross.

Looking at Figure 22, there are some curious trends between these few curves. One thing many may notice are the placements of the points with distance; this is due to how the flight data was sampled. The recordings were chopped into 1, 5, and 10 second (nonoverlapping) intervals to see how the detection system behaves with distance. Reasons for this will be further explained with the detection results in the next section. Referring to the decaying curves, the one in orange is what the model in the methodology predicts for the Angel when given the flight conditions that we were in, the measured data points and their trend are in blue, and what the model predicts for the Albatross in ideal conditions is given in green. The measured data points do clearly like to reside in SNR levels close to the predicted model, and the trend is decreasing with distance, but it is not exactly the same as the model predicted shape. The cause for this is explained simply due to the fact that SNR is being estimated in low SNR conditions. Because the SNR is so low being embedded on a UAV and the noisy signal and noise SNR estimation variation (referring to equation (48)) is the only one obtainable with flights, this leads to high measurement uncertainty. Taking a ratio where the noise, a highly random variable, is present in both the numerator and the denominator and the signal is known to be low, the model trend is going to be difficult to match in
these conditions.

Furthermore, wind speeds up to $6 \frac{m}{s}$ is approaching the maximum tolerated winds for this UAV to fly in and remain stable; the preferred ideal conditions would have been flying the Albatross, a much quieter UAV in much calmer winds (about $1-2 \frac{m}{s}$ would be reasonable) and slowing down over the acoustic source at its speed of $13 \frac{m}{s}$ so that a substantially greater SNR could be obtained. Note that the reason SNR would see such a great improvement from this is because of the extracted power of the wind relationship in equation (43). The power obtainable from the wind is proportional to the cube of the fluid element velocity, so the increasing wind will increase the power of the noise exponentially causing a degradation in the SNR. It is still important to note that signal power is measured on the logarithmic decibel scale with its reference value of $10^{-12} \text{W}$ as given in Table 1. Even though power may be proportional to the cube of velocity, the noticeable gain will not increase close to a cubic increase as is shown in the wind power formula in equation (43). Nevertheless, the SNR as a function of distance was still attainable and it will provide the user a better idea of how far away it is feasible for their respective UAS to make a detection.

Detection Performance

Audio file processing time intervals of 1, 5, and 10 seconds were processed to evaluate performance differences. Processing the detection distances in 1-second intervals was going to lead to a stronger detected signal presence because it would not be considering the outer time samples where the UAV is far from the source (which would hurt the estimates). However, at the same time, 1-second intervals have fewer samples than 5 or 10-second intervals, meaning that more information is present in the others due to the higher number of samples. This is where the conditions with the Albatross UAV could be extremely useful, because not only would it have less noise, but its temporary slowing capabilities could allow it to linger over the source, thus providing a stronger signal for a larger duration. Nonetheless, the Angel still was able to make detections.
As the hypothesis detection algorithm is described before in algorithm 1, the decision is based on taking the error of the peaks that are outside the standard deviation multiplied by a scaling factor so that the noisy oscillating low magnitude components do not interfere with the error metrics of the harmonic peaks. Based on that decision process, detections from the flight were able to be made for distances up to 40 meters somewhat consistently with 10-second intervals (7 out of 9 recorded passes to be exact). The results for the 9 flight passes assuming 10-second intervals are given in tables 3, 4, and 5.

Table 3: Flight results for the first 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude.

<table>
<thead>
<tr>
<th>Passes</th>
<th>Time Shifts</th>
<th>Signal RMSE</th>
<th>Noise RMSE</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pass 1</td>
<td>$t = -20$</td>
<td>5.04</td>
<td>3.38</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>3.67</td>
<td>3.24</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 0$</td>
<td>4.11</td>
<td>3.72</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>3.8</td>
<td>3.57</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>3.91</td>
<td>3.63</td>
<td>No Threat</td>
</tr>
<tr>
<td>Pass 2</td>
<td>$t = -20$</td>
<td>4.02</td>
<td>2.77</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>3.16</td>
<td>4.16</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 0$</td>
<td>4.19</td>
<td>3.19</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>3.89</td>
<td>3.87</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>4.04</td>
<td>3.3</td>
<td>No Threat</td>
</tr>
<tr>
<td>Pass 3</td>
<td>$t = -20$</td>
<td>3.93</td>
<td>2.57</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>4.13</td>
<td>4.21</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 0$</td>
<td>3.32</td>
<td>4.13</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>4.3</td>
<td>3.23</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>3.78</td>
<td>3.51</td>
<td>No Threat</td>
</tr>
</tbody>
</table>
Table 4: Flight results for the middle 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude.

<table>
<thead>
<tr>
<th>Passes</th>
<th>Time Shifts</th>
<th>Signal RMSE</th>
<th>Noise RMSE</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = -20$</td>
<td>3.57</td>
<td>3.58</td>
<td>Threat</td>
</tr>
<tr>
<td>Pass 4</td>
<td>$t = -10$</td>
<td>3.87</td>
<td>3.38</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 0$</td>
<td>3.72</td>
<td>4.04</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>3.75</td>
<td>4.42</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>3.24</td>
<td>3.53</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -20$</td>
<td>3.55</td>
<td>3.25</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>3.87</td>
<td>3.29</td>
<td>No Threat</td>
</tr>
<tr>
<td>Pass 5</td>
<td>$t = 0$</td>
<td>3.35</td>
<td>3.26</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>3.8</td>
<td>4.07</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>4.62</td>
<td>3.43</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -20$</td>
<td>3.92</td>
<td>3.33</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>2.7</td>
<td>3.14</td>
<td>Threat</td>
</tr>
<tr>
<td>Pass 6</td>
<td>$t = 0$</td>
<td>4.25</td>
<td>3.57</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>4.25</td>
<td>2.91</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>4.76</td>
<td>3.73</td>
<td>No Threat</td>
</tr>
</tbody>
</table>
Table 5: Flight results for the last 3 passes of the diesel engine source, where time $t = 0$ indicates the relative time when the UAV was estimated to be directly over the source at a 40 m altitude.

<table>
<thead>
<tr>
<th>Passes</th>
<th>Time Shifts</th>
<th>Signal RMSE</th>
<th>Noise RMSE</th>
<th>Hypothesis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$t = -20$</td>
<td>4.01</td>
<td>3.08</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>3.76</td>
<td>3.32</td>
<td>No Threat</td>
</tr>
<tr>
<td>Pass 7</td>
<td>$t = 0$</td>
<td>3.6</td>
<td>3.86</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>4.17</td>
<td>3.66</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>4.67</td>
<td>2.73</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -20$</td>
<td>3.89</td>
<td>4.07</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>3.67</td>
<td>4.25</td>
<td>Threat</td>
</tr>
<tr>
<td>Pass 8</td>
<td>$t = 0$</td>
<td>4.03</td>
<td>3.59</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>4.66</td>
<td>3.82</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>3.95</td>
<td>2.83</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -20$</td>
<td>3.68</td>
<td>3.42</td>
<td>No Threat</td>
</tr>
<tr>
<td></td>
<td>$t = -10$</td>
<td>4.54</td>
<td>4.14</td>
<td>No Threat</td>
</tr>
<tr>
<td>Pass 9</td>
<td>$t = 0$</td>
<td>3.12</td>
<td>3.37</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 10$</td>
<td>3.56</td>
<td>3.88</td>
<td>Threat</td>
</tr>
<tr>
<td></td>
<td>$t = 20$</td>
<td>3.77</td>
<td>3.11</td>
<td>No Threat</td>
</tr>
</tbody>
</table>

The signal and noise RMSE listed show the RMSE gathered for the noisy signal with respect to the signal and noise using the detection algorithm proposed before. The hypothesis is shown that is based on the two errors, and the time shift value is indicative of the UAV location. As these are the results for the flight using 10-second intervals, $t = 0$ indicates that the UAV is directly above the acoustic source, $t = -10$ is the UAV approaching the source from 10 seconds in the past with respect to $t = 0$, $t = 10$ is the UAV leaving the source at 10 seconds in the future, and so on. From these results, it is clear that there is a consistent trend in the detections being made.

Although the results do seem promising, there are also common flaws in the detection results...
that have plausible explanations. For instance, as some of those detections were made in neighboring intervals rather than the interval directly over the source, this could be explained by not perfectly aligning the audio with the flight pattern, but it still was close enough to be within reason of only leaking detection results to the neighboring intervals. There is no doubt that this system performs better with higher samples when more information is given though, because distances covered with 1-second intervals did technically predict up to 190 meters, but there appeared to be lots of oscillating decisions, so it is likely that 1 second at a 48 kHZ sample rate is not enough information to make a practical estimate. The 5-second intervals seemed to get most distances at 100 meters, but detections outside that range did not seem consistent. It should be noted that this flight involved the Angel merely going in a circle at a 200-meter radius, so the variance in detections could also be explained by random turbulences and sharp turns it abruptly had to take to stay on the path, or by random gusts happening every once in a while. Gusts could especially be a plausible explanation too because as it is given in equation (43), the power of the wind is proportional to the cube of the wind velocity. A swift change in wind speed could throw it off enough at those distances and traveling speeds.

The scaling factor $\alpha$ that helps determine how many of the harmonic peaks to be weighed in the error estimate behaved differently for the various intervals. The data that took intervals up to 10 seconds performed much better with a scaling factor of $\alpha = 5$, but for the intervals of 1 second each, $\alpha = 3$ was a more trustworthy scaling factor. The logic behind this behavior is explained by duration and sample size because as the 10-second intervals have a much higher sample size, the HST would bring out and separate the harmonic peaks more from the noise because more information about the harmonic presence and their amplitudes is gathered which further separates the harmonics from the noise floor. On the other hand, for intervals as short as 1 second, the noise floor will not be as low because less information is gathered, and the spectrum will vary more in the residual components. In other words, the 10-second intervals should have better-separated peaks while the 1-second intervals have a wide spectrum, which explains why $\alpha = 5$ is a good scale for the 10-second intervals to ensure that the noise is not interfering, while more peaks
should be considered for the 1-second intervals to get a better estimate.

As a final performance evaluation of this system, error metrics have been plotted as a function of distance to express the trend in the error difference as the UAV gets further away from the source. Upon detection decision making, the error that seemed to perform the best with the most consistency is RMSE, which was used for the decision algorithm and is plotted by itself in Figure 23 comparing the error of the noisy signal between the signal and noise.

![RMSE Trend with Distance](image)

**Figure 23:** RMSE plot as a function of distance from Angel flight data collection.
Figure 24: Error metric plot as a function of distance from Angel flight data collection.

As can be seen from the RMSE plot in Figure 23, the RMSE (which was the metric used for the detection hypothesis) takes a somewhat positive linear trend as the distance is increased between the source and the UAV. The plotted data points do include all distances recorded for the 5 and 10-second slices of the flight data, but not the 1-second intervals as those demonstrated the most inconsistent behaviors in terms of error difference and threat detections. The error variance is higher for the 40, 210, and 310-meter measures than the rest. These are the distances corresponding to the 10-second intervals, which demonstrates that more samples allow the error to approach more specific values. Moreover, this spread can also be expected if the wind speeds change constantly during the flight. Higher winds will lead to a higher error in the 10-second intervals, especially since the UAV is moving far beyond the target. A way to minimize the error would be to slow down the UAV (as is possible with the Albatross) and allow more signal power as it lingers over the target longer while the wind component is also reduced.

It is likely that the approximation curve’s bend in the beginning is a result of the combination of measurement error and of the polynomial fitting estimation for this trend. A couple of characteristics worthy of noting are what the distances are where the error of the signal is less
than the error of the noise and overall how close the error curves are to each other. To begin, Figure 24 does mark x’s where the curves of each metric cross. Firstly, the x’s that are on the far left close to the distance of 50 meters are likely a result of the polyfitting estimation, and with that said it is safe to assume that the error of the signal before this point is probably also less than the noise in most cases. The second crossings for each show the point at which the hypothetical boundary is drawn where the system can detect a signal and predict that something is there. Note that this is for the roughly given net wind speed of $31 \text{ m/s}$ (25 m/s travel speed plus the oncoming average wind component of 6 m/s as discussed before) strictly speaking. At higher wind speeds, the noise error curve would shift slightly downwards indicating that there is less of an error between the noisy signal and the noise, and the signal error curve would shift slightly upwards.

Another fact about these estimations is how close they are to each other. It makes sense that the MAE is the lowest due to its high inertial weighting, RMSE is a step above this in error penalty, and MSE is much higher as it punishes harder for the outliers, which are common in this HST transformation. But for each metric, the signal error keeps very close to the noise error. The reason for this is because of how intensity dissipates with $\frac{1}{r^2}$. The distance from the source in this data collection was high, so the signal has already lost energy as it propagated through the atmosphere, and its rate of decrease in energy becomes less at a logarithmic rate as it continues to go further. In this case, the SNR ratio was so low and the signal was so far away that moving much further would not make a large difference.
5 CONCLUSION

To conclude, fixed-wing acoustic sensing mounts were developed for the optimal minimization of air fluid flow around the UAV acoustic sensors, data was collected from three different locations for characterization and performance tests, and a flight with a VTOL UAV was conducted evaluating if detections could be made from a moderate distance. The wing mounts were each spectrally characterized to see how each mount design affected the response. They were then tested in a wind tunnel with signal, noise, and noisy signal recordings to analyze the improvement of SNR due to the wind. Results said that a decrease in wind can make a large difference as the power of the wind is proportional to the cube of the velocity. The mounts were flown on a VTOL UAV to investigate the feasibility of confirming detections with them and from what distances.

Being able to make a consistent detection in suboptimal flight conditions from at least 40 meters away was demonstrated. Seven out of nine of the flight passes made consistent detections when the 10-second intervals were considered. More data should be collected to solidify the results of these methods as more detections were claimed to be made from further distances and with lesser intervals of recorded data, but the detection results were not entirely consistent. Nevertheless, progress has been made, and this area of research is still in its early stages. Furthermore, if this system is later fully verified and improved, the capabilities of the detection system will be able to detect specific vehicles’ engines as opposed to just detecting an engine in general. For this research to only be an investigation of the feasibility of detections, it shows that the area is worth further investigation.
6 FUTURE WORK

Alternative Mounts and Acoustic Sensing Devices

There are other areas of this project that could be expanded upon. Beginning with the mounts developed in this project, it is clear that not every mounting design was tried, so further shapes and mounting styles could be explored. Microphone filters like foam and wind muff filters to reduce the wind even further should be investigated. As covered in Section [1], the mics used in this design were electret condenser microphones, but the initial hopes of this project were to compare MEMS microphones with high acoustic overload points (AOPs) alongside this, but supply chain issues kept them from being used before this thesis was completed. Implementing mics like these has the potential to be beneficial, and other microphones with more focused polar patterns would be a great idea as well to only allow in the signal and block out the crossing winds.

Furthermore, for the problem of DOA estimation, acoustic vector sensors (AVS) are excellent sensors to look into due to their functionality of measuring differential particle velocity on each axis. In addition, it was mentioned in the literature review that coprime arrays of microphones are said to open up more degrees of freedom. This is an area worth exploring because it could make DOA estimations more precise and reduce sensor costs when fully optimized.

Filtering Techniques and Harmonic & Percussive Source Separation

Additionally, harmonic and percussive source separation (HPSS) through median filtering should be examined. Once this idea began in 2010, many works have been done trying to improve this technique [63-66]. Signals can be broken down into harmonic and percussive components, where harmonic components are those associated with waves, pitches, and musical tones, and then percussive components are those associated with clashing objects and causes the frequency spectrum to be uniform in those instances. This is exactly how the wind noise normally behaves. For a fixed-wing UAV, it is a saturation across all frequencies as it flies at a certain speed, and unfortunately in the case of this project, it does still have a stronger presence in the lower end of
the spectrum. Another way to identify these though is by looking at a spectrogram of a signal (with time on the x-axis and frequency on the y-axis). Observing this, the percussive components will have intensity lines trending vertically, while harmonic components like a pure sine wave, would be trending horizontally. An example of how these components differ in spectrograms is displayed in Figure 25 which demonstrates some noisy signal samples of a frequency sweep taken at the wind tunnel used in this project. If this separation median filtering technique is optimized, applications such as making vehicle detections could see an improvement.

Figure 25: A comparison of original, harmonic, and percussive spectrograms of noisy signals taken at the wind tunnel with a 50 to 2000 Hz exponential sweep signal and 15 $m/s$ winds.

Taking HPSS further, referring to equation (43), another useful task would be characterizing the acoustic impedance of the acoustic entry points with respect to their cross-sectional area and determining what its effect is on the attenuation of the percussive components versus the harmonic components. If this is characterized, it would give better insight into how much it is helping reduce the percussive components while still allowing the harmonic components when
the cross-sectional area is adjusted. Moreover, as the project continued being developed, applying a CNN over characteristic graphs such as spectrograms, onset strength graphs, PSDs, autocorrelations, or HPSS spectrograms may have potential. There are no large published datasets for this though, and recording all the data needed to train such a model would take a large amount of time. As CCNs are used in many areas of research though regarding audio, they have been noted to perform very well for this application and would be worth seeking out.

Lastly, one filtering technique that could have been exploited is active noise control (ANC). As specified in the literature review, it does remarkably reduce the noise as it recursively calculates an anti-noise and applies it to cancel the noise in real-time. Some studies have shown that this can still provide substantial noise reduction and speech intelligibility, especially when cascaded with other algorithms [2].
REFERENCES


APPENDIX

Detailed Wing Mount Pictures

Table 6: Wing mount style variations.

<table>
<thead>
<tr>
<th>Wing Mount Style Abbreviation Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR  High ramp mount</td>
</tr>
<tr>
<td>TD  Tight deep cavity mount</td>
</tr>
<tr>
<td>WD  Wide deep cavity mount</td>
</tr>
<tr>
<td>FM  Flat mount</td>
</tr>
<tr>
<td>LR  Low ramp mount</td>
</tr>
<tr>
<td>TS  Tight shallow cavity mount</td>
</tr>
<tr>
<td>WS  Wide shallow cavity mount</td>
</tr>
</tbody>
</table>