Winter 12-18-2021

Developing Iot-Based Geophysical Micro-Observatories Utilizing Cloud Computing

Joshua Sylvester
Southern Methodist University, jsylvester@smu.edu

DEVELOPING IOT-BASED GEOPHYSICAL MICRO-OBSERVATORIES

UTILIZING CLOUD COMPUTING

Approved by:

_____
Dr. Eric Larson
Professor of Software Eng.

_____
Dr. Theodore Manikas
Professor of Computer Eng.

_____
Dr. Matthew J. Hornbach
Professor of Earth Sciences

DEVELOPING IOT-BASED GEOPHYSICAL MICRO-OBSERVATORIES

UTILIZING CLOUD COMPUTING

A Thesis Presented to the Graduate Faculty of the

Bobby B. Lyle School of Engineering

Southern Methodist University

in

Partial Fulfillment of the Requirements for

Master of Science

With a Major in

Computer Science

by

Joshua Henry Sylvester

B.A. Geology, Southern Methodist University, 2020
B.S., Computer Science, Southern Methodist University, 2020

December 18, 2021

ACKNOWLEDGMENTS

Sylvester, Joshua Henry   B.S., Southern Methodist University, 2020

<u>Developing IoT Based Geophysical Micro-Observatories Utilizing Cloud Computing</u>

        Instrumentation for collecting geophysical data, specifically heat flow, in lake and
marine environments has been in existence for over fifty years. Despite this, the costs
associated with data collection and the technological limitations of existing
instrumentation can be preventative when conducting geophysical studies. Furthermore,
the success rate of such studies is limited by the lack of real time data transmission
capabilities when instruments are deployed for extended periods of time. As a solution to
this problem we have created cost-effective and lightweight IoT-driven instrumentation
and combined it with cloud computing technology facilitating real time data transmission
to the cloud. Furthermore, we have also explored a new data analysis technique
employing Granger causality-based time series clustering to investigate relationships
between atmospheric data and water column temperature data.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: INTRODUCTION

1.1 Introduction

Geophysical instruments for collecting heat flow data have been in existence for the last 50 years, however, few if any provide real-time measurements for dynamic geological systems. Long-term heat flow data is of scientific and industrial interest because it can offer insight into how subsurface fluid and heat flow reacts to regional and environmental changes. In addition to this need for long-term heat flow data, there is also a need for long-term water column data of the water body itself to correct for seasonal temperature variability that reduces the accuracy of spot heat flow measurements. The operation of standard heat flow collection technology is typically cumbersome and expensive, often requiring large research vessels with A-frames or ROVs (Remotely Operated Vehicles) in both ocean and lake environments. Currently, there are few reliable methods for collecting high-resolution, continuous, real-time heat flow data for extended periods of time in dynamic systems (like seep and vents in volcanic systems). To address these problems, we have created the HyLO-Mo (*Hy*brid *L*ister *O*utrigger – *Mo*nitoring System) and the WaCoMS (*Wa*ter *Co*lumn Monitoring System). Once deployed, these instruments operate for months at a time (or indefinitely with solar) and are capable of transmitting continuous data in real-time at a cost two- to three- orders of magnitude less than ocean-bottom cable observatories.

1.2 Motivation for Real-Time Heat Flow Instrumentation Development

Heat Flow, put simply, is the movement of heat within the Earth. This heat is a product of conduction and convection ultimately sourced from the decay of radioactive elements in the Earth's crust and interior. In its simplest form, the 1-D equation used for calculating heat flow can be written as follows:

$$Q = \lambda * \delta T / \delta z$$

Where λ is the thermal conductivity (here assumed constant) of the medium and $\frac{\delta T}{\delta z}$ is the thermal gradient with depth.

An understanding of a region's heat flow and access to heat flow measurements is important for both researchers and industry alike. Companies conducting oil and gas exploration rely on regional subsurface temperature profiles to assess if an area's hydrocarbons are thermally mature. Additionally, companies looking to move into geothermal energy use heat flow to identify areas that have high subsurface temperatures necessary for geothermal power production. Researchers and scientists also use heat flow as a way to better understand the Earth's evolution and the geologic processes that played a role in it. This is especially the case for volcanic systems, where researchers use heat flow to assess, quantify, and monitor volcanic activity and volcanic hazards [1].

Despite the relevance of heat flow data to both industry and academia, there are few efficient and cost-effective ways of collecting it. One common way of collecting heat flow data is by measuring it in drill holes. These measurements, however, are often rife with uncertainty as there are typically large amounts of fluids being pumped in and out of these holes disturbing the thermal equilibrium. Furthermore, these measurements

often consist of only a few bottom hole temperature measurements [2]. Another way of measuring heat flow is through the use of ~1-2 ton Lister or coring probes consisting of multiple thermometers on a metal lance that is inserted several meters into marine sediments [3][4]. Temperature measurements between thermometers are used to estimate a thermal gradient. This is a reliable form of measuring shallow heat flow and is significantly more cost-effective than drilling but is generally only performed in deep marine environments where seasonal bottom water fluctuations are minimal. However, these are spot measurements only, and therefore offer only a snapshot of a dynamic area's heat flow potential. It has been recognized for decades that it is very difficult, if not impossible, to determine a region's true heat flow in shallow subsurface environments without time series heat flow data collected on longer timescales [5][6]. Long-term data should be in the form of both long-term near-surface measurements and long-term subsurface temperature measurements so that both surface effects and subsurface thermal properties (e.g. sediment diffusivity) can be accounted for properly. With long-term time-series heat flow data, researchers can remove noise introduced into the measurements from seasonal temperature fluctuations of the water body overlying the sediments, resulting in a more accurate picture of the heat being dissipated from the Earth's interior. These fluctuations have become exaggerated due to the effects of climate change that can extend boundary conditions fluctuations for decades, further necessitating long-term heat flow measurements [7][8][6].

Besides monitoring drill holes (which is often cost-prohibitive due to required drilling and well access), few methods exist for sustained long-term heat flow data collection [7]. Of the instruments and probes which can be used to collect long-term heat

flow data, only a very expensive few (for example, ocean-bottom cable systems) have real time data transmission capabilities [9]. Alternatively, some systems on land and at sea collect continuous data but require that the instrument be manually retrieved before the data can be accessed [8]. This presents several problems: 1) instruments are deployed for several months if not years at a time and often fail. When a failure occurs, it is unknown until the device is inspected, resulting in significant data and time lost. 2) Data cannot be analyzed in real-time, meaning that when the instrument is finally collected, there is a backlog of analysis. Because of this, subtle changes in the data record that might be associated with small earthquakes or other potentially minor geological events could be overlooked. 3) There is often clock drift due to real-time links to GPS atomic clocks, making correlations of thermal response with other geological phenomena difficult to access. An additional problem exists in the form of feasibility: current technology used for long-term heat flow detection is cumbersome and cost-ineffective to deploy and recover— for example, often requiring expensive ROVs and large ships in marine/lake environments. As a result, for more than a decade, the heat flow community has challenged itself to develop more cost-effective, real-time heat flow monitoring systems as was noted in the 2007 NSF-sponsored workshop report, "*The Future of Marine Heat Flow: Defining Scientific Goals and Experimental Needs for the 21st Century).*

In the Spring of 2018, our group set out to develop a lightweight and versatile instrument to complement existing instruments by providing on-the-fly heat flow measurements. This development resulted in the creation of the *Hy*brid *L*ister-*O*utrigger (HyLO), a probe that streamlines the collection of spot measurements. The research

presented in this thesis represents a continuation of our group's prior heat flow instrumentation research by combining the newly developed HyLO probe with advancements in cloud computing to construct a new instrument that can be deployed for months at a time while transmitting data in real-time. Such a probe will completely eliminate the hours-to-years of downtime traditionally spent waiting for instrument and data recovery.

1.3 Chapter Organization

- In Chapter 2 the history of heat flow instrumentation used for both spot measurements and long-term geological monitoring is discussed.

- In Chapter 3 the *Hy*brid *L*ister *O*utrigger (HyLO) Probe is introduced and an overview of this thesis's precursory work is presented.

- In Chapter 4 the *Hy*brid *L*ister *O*utrigger *Mo*nitoring System (HyLO-Mo) and the *Wa*ter *Co*lumn *M*onitoring *S*ystem (WaCoMS), two novel approaches to collecting long-term geophysical time series data are presented.

- In Chapter 5 the HyLO-Mo and WaCoMS systems' first deployments are discussed and preliminary results are shown.

- In Chapter 6 data collected by the WaCoMS probe is analyzed using a novel Granger causality approach to time series clustering.

- In Chapter 7 discusses the implications for the newly developed micro-observatory technologies and potential future work.

CHAPTER 2: THE HISTORY OF HEAT FLOW INSTRUMENTATION

2.1 Introduction

The process of collecting heat flow measurements on the sea bottom or lake floors

has complex requirements and the means for collecting such data vary. Instruments

designed for such purposes must be able to 1) withstand the pressures associated with

ocean bottom or lake bottom operations 2) penetrate one or more meters of sea bottom or

lake bottom sediments to minimize seafloor temperature seasonal variability and 3)

record and/or transmit the data they measure. Since the mid-1900s, several tools have

been created to satisfy these requirements for heat flow data collection. Though designs

vary, heat flow probes have several important commonalities: they all include a series of

thermistors incased in or attached to a lance that acts as a vehicle to insert the thermistors

into the subsurface. These thermistors, spaced at intervals, will record the respective

temperatures in the sediments. A thermal gradient can then be calculated with these

measurements. Because ocean and lake bottom sediments are typically very similar in

composition and porosity, a thermal conductivity of ~1 W/mK can be assumed in most

cases so that a heat flow calculation can be obtained.

To collect heat flow data these probes are typically lowered close to the ocean- or

lake-floor where the probe is momentarily stationary so that the measured temperatures

equilibrate. Then the probe is dropped several meters and penetrates the water bottom

sediments. Because of frictional heating associated with the penetration, the probe is

typically left in the sediments for several minutes and in some cases up to an hour so that equilibration can occur.

The first instrument created to operate in such a manner was the Bullard probe [10]. However, today the two instruments most prevalently used to collect marine heat flow measurements include the "violin-bow" Lister probe and the Ewing-type probes, pictured below. The "violin-bow" Lister probe is a standalone instrument while the Ewing temperature probes can be outrigged to a gravity corer and record data while core samples are being collected[11][12]. As discussed in our group's previous work, each of these probes has its advantages and disadvantages, and arguably the most significant disadvantage to all heat flow measurement probes is that none provide either rapid or real-time data transmission so that scientists can make on-the-fly measurements of subsurface temperature [3]. This disadvantage can lead to hours or even days of delay (as data recovery sometimes requires Remote Operated Vehicle (ROV) deployments and can cost hundreds of thousands of dollars in ship time). In this chapter, I outline both the history and progression of heat flow instrumentation development up to this stage to provide the reader with context for the new real-time HyLO system that our group has developed.

2.2 History of Heat Flow Spot Measurement Instrumentation
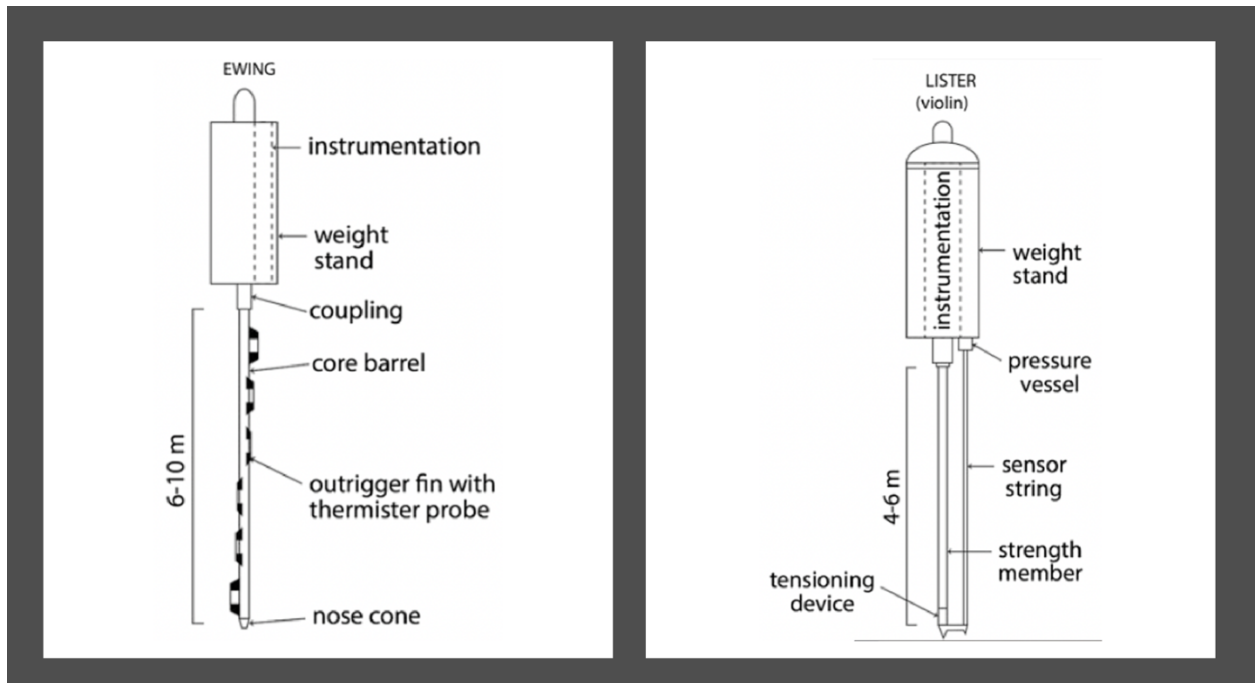
Figure 1: Diagrams of EWING (left) and Lister-style (right) heat flow probes. Diagrams originally published in J. Sclater et al.'s "Marine Heat Flow" [4]

Since the creation of the Lister (violin-bow) probe in the 1970s, it has since become the heat flow collection workhorse when operating in lake and ocean environments [13][3]. These probes are extremely durable and have the added benefit of being able to collect in situ thermal conductivity measurements through the use of a heating element attached to the probe which generates a known temperature pulse [13]. It is also capable of transmitting some real-time basic and preliminary real-time data. The rigid and durable nature of the Lister probe does come at the cost of maneuverability.

These instruments typically weigh over half a ton and can be 3-6 meters long depending on the length of the thermistor lance, making them cumbersome to deploy.

Figure 2: Lister probe on the deck of a ship before deployment via A-Frame. Image courtesy of Matthew J. Hornbach, taken during heat flow study in the Beaufort Sea.

The EWING temperature probe contains a single thermistor and is designed so that multiple probes can be attached to a gravity corer [4]. A thermal gradient cannot be calculated until several, if not all, of the individual probes have been recovered. Because of their individual nature, the EWING probes are extremely portable but the requirement that they are outrigged to a gravity corer system does limit their versatility. They are very applicable when wanting to collect additional environmental data when heat flow and thermal gradients are not the primary objectives of the study. The biggest downside

to these probes is that they do not have any real time transmission capabilities and each individual probe must be opened up to retrieve the recorded data.

Due to the cumbersome nature of the previously mentioned instruments, they are used for spot measurements and are not left for long periods of time on lake bottoms or the seafloor.

2.3 History of Long Term Deployment Tools

Long-term deployments of instruments designed to monitor heat and fluid flow are limited due to either limitations of existing technology, cost, or both [3]. In most cases, long-term heat flow experiments have been conducted at significant costs primarily due to complex installment/recovery logistics.

One method for collecting long-term temperatures measurements is by instrumenting boreholes, installing instruments in open boreholes to obtain time-series measurements, but doing so is a technologically complex and expensive endeavor. The most recognized approach for this involves the use of CORKS, or borehole observatory systems, which essentially seal off a portion of the borehole so that instruments can collect pressure and temperature data from a particular subsurface formation [14]. These instruments can collect data for extended periods of time, usually up to four years. However, they are extremely costly due to the need for a drillship to drill and case the holes as well as the use of ROVs to retrieve data or service the CORK. In several cases, the sealed-off capsules have leaked, compromising data quality [14][15][16].

Figure 3: A CORK system. Picture originally published in *"Design, deployment and status of borehole observatory systems used for single-hole and cross-hole experiments IODP Expedition 327, eastern flank of Juan de Fuca Ridge"* [14]

In inland lake environments, studies such as one conducted at Yellowstone Lake in 2017, use long-term 1-m long thermistor lances to collect temperature and various chemical data [17]. In the Yellowstone Lake study, two probes were deployed using an ROV system which first inserted casing into each deployment site to protect the instrument during insertion. Data from these probes were not recovered until a year later when the probes were retrieved by the ROV. A more recent study in Yellowstone utilized instruments that made improvements to existing designs [18]. However, the technology was still dependent on an ROV for deployment/recovery and lacked real time data transmission capabilities as well as had a limited precision of 0.05° C resolution for a 1 m long probe.

With each of these instruments, the costs associated with drilling or ROV deployments can exceed hundreds of thousands or even millions of dollars, and catastrophic failure (where no useable data are recovered due to instrument failure, or leaking) are common—rarely is a deployment without at least one minor failure resulting in data loss. In a system with real time data transmission capabilities, system failures are less of a concern because data will have already been transmitted. Real-time systems, therefore, eliminate the risk of complete data loss that can occur in long-term deployments, where failure even at the very end of the experiment or during recovery can result in complete data loss. Furthermore, real-time data feeds provide insight into the mechanical health of the instrument by providing insight into onboard temperature, pressure, humidity, and orientation values that allow scientists to diagnose real-time instrument mechanical integrity.

CHAPTER 3: INITIAL HyLO DESIGN AND DEVELOPMENT (2018-2020)

3.1 Introduction

Following an expedition in the Artic in which heat flow data was collected using the previously mentioned Lister probe, my group set out to develop and construct a new heat flow spot measurement probe, the HyLO (Hybrid Lister Outrigger) [6]. The engineering of the HyLO probe is documented in our previous work [3]. The HyLO was constructed to be a spot measurement tool with two goals. It would 1) have the capability to "collect near-real-time temperature gradient data and 2) [be] a low-cost, lightweight, and highly portable tool" [3]. After several design iterations, a working prototype was finished in the Spring of 2019. It was then tested extensively through sea trials onboard the R/V Meteor in May of 2019 where it collected heat flow data off the coast of Montserrat. In this chapter, I briefly discuss the HyLO design as the probe is an integral component of the long-term monitoring probes which are the subject of this thesis.

3.2 Hylo Design:

The HyLO consists of an interchangeable stainless-steel lance which can vary in length from 1-3 meters and has an outer diameter of 0.95 cm. Inside the lance is 8-12 analog thermocouples which are incased in vacuum-pumped high thermal conductivity epoxy, to ensure rapid response times in measurements following probe insertion. The epoxy has the added benefit of increasing the structural rigidity of the probe lance, decreasing the likelihood of a bending event during insertion. The lance can be attached to the polycarbonate probe head with a Swagelok connection. This probe head has 113

cm^3 of space for onboard instruments and is rated to a depth of ~2,100 meters below sea level (mbsl).

After attaching the lance to the probe head, data lines for each thermocouple along with shared power and ground lines extrude from the top of the lance. These lines are accessible inside the probe head. The data lines are fed into an 18-bit analog-to-digital converter which is then connected to the onboard ATmega32u4 microcontroller for data storage or transmission [3]. Accompanying these two devices are an accelerometer/tiltmeter, a humidity sensor, and an internal pressure sensor. These instruments can be run via batteries also stored within the probe head, or via shipboard power which is run down directly to the probe.

3.3 HyLO Deployment Configurations:

The entire system weighs less than 40 kilograms and was designed to be deployed in one of two ways, a schematic for each deployment method is shown in figure 4 . The first is through attachment to a strength member cable that can double as a data line for real-time data relay to the shipboard deployment crew. This approach requires adding an extra 10-20 kg of weight to ensure the insertion of a 2-meter probe lance into the water bottom sediments. This system is designed to work as a standalone approach that can be conducted with a crew of 2 in a small vessel (20ft outboard skiff) when rapid measurements are desired.

The second deployment method is to outrig the HyLO probe to a gravity corer, or other large ocean bottom instrument. This deployment approach is applicable when large-scale experiments requiring A-frame ship infrastructure are being conducted but

heat flow is not the main objective of the study. The outrigged HyLO approach allows

for the easy collection of heat flow data which can supplement and support other results

collected during the study. To conduct an outrigged deployment of the HyLO probe,

brackets are attached to a gravity corer's core barrel via band straps. The probe is then

attached to the brackets and put under tension to avoid bending the probe lance. During

an outrigged deployment, data are recorded on an onboard SD (Secure Digital) card.

When the gravity corer is removed from the water the data can be transmitted to the

shipboard crew using Bluetooth. This enables the crew to quickly recover data and move

onto the next deployment site without having to physically remove the HyLO probe

outrigger system from the gravity corer.
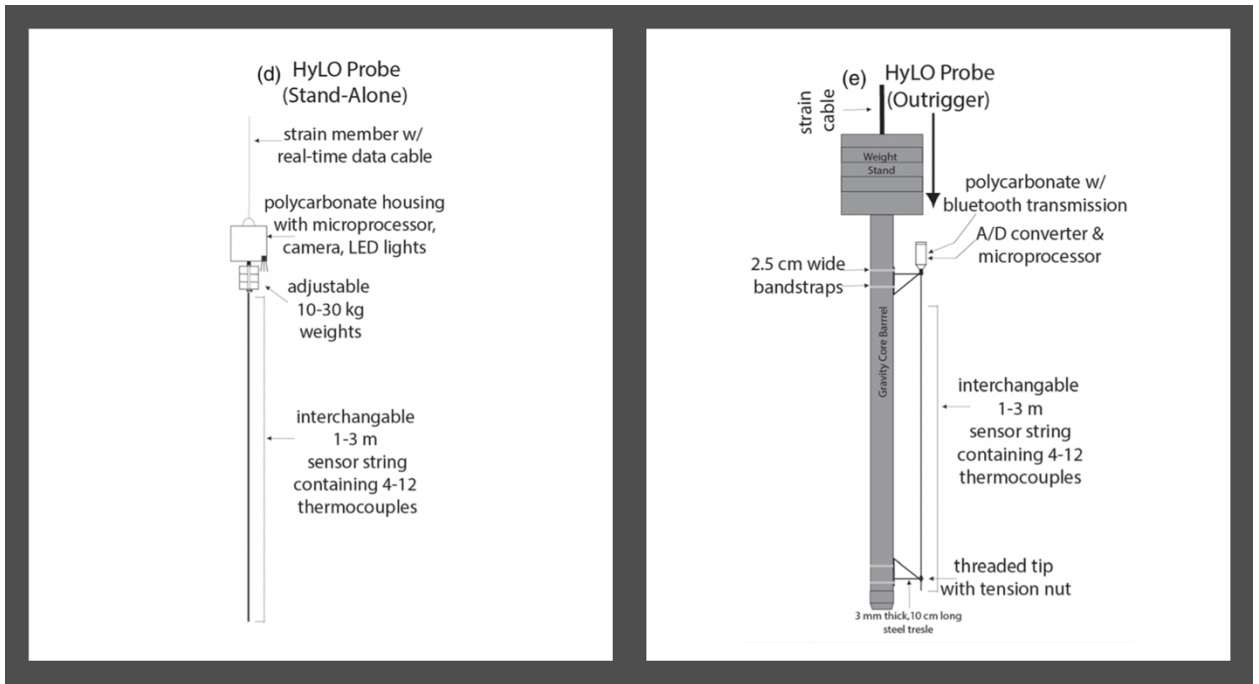
Figure 4: HyLO Deployment Configurations. The left diagram shows a standalone
deployment configuration and the right shows an outrigger deployment configuration.

Figures originally published in *"A Hybrid Lister-Outrigger Probe for Rapid Marine Geothermal Gradient Measurement"* [3].

CHAPTER 4: HyLO LONG TERM DEPLOYMENT SYSTEM

4.1 Introduction

The creation of the HyLO-Mo (*Hy*brid *L*ister *O*utrigger –*Mo*nitoring System) and the *Wa*ter *Co*lumn *Mo*nitoring *S*ystem (WaCoMS) is the focus of this thesis. The HyLO-Mo consists of three primary components. 1) The first is the data collection device, which will be a repurposed HyLO probe. 2) The second is the float platform and equipment case which contains solar panels, batteries, and microcontrollers used to communicate with nearby cell towers. 3) The third is the cloud computing resources which receive, process, and store the data transmitted via cell towers, allowing researchers to access the data in real-time from anywhere with an internet connection. The design for the WaCoMS follows the same specifications as the HyLO with the only difference being that instead of using a HyLO probe for data collection, there is a 36-meter strand of several analog thermocouples and HOBO temperature dataloggers used to measure temperature fluctuations in the water column. Figures 5 and 6 show diagrams of both designs.

Schematic of HyLO-Mo Micro-Obseravtory Deployment
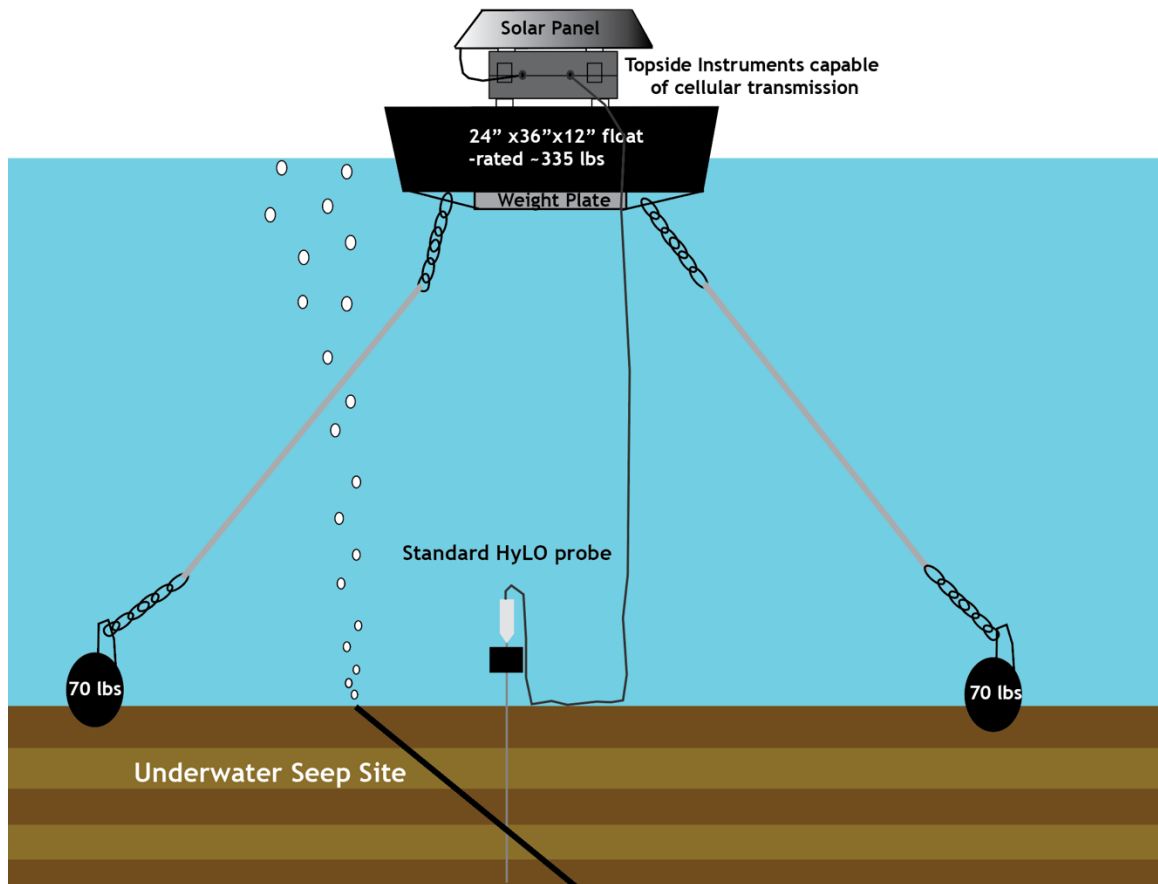
Figure 5: Diagram of the proposed Hybrid Lister Outrigger Monitoring System (HyLO-Mo)
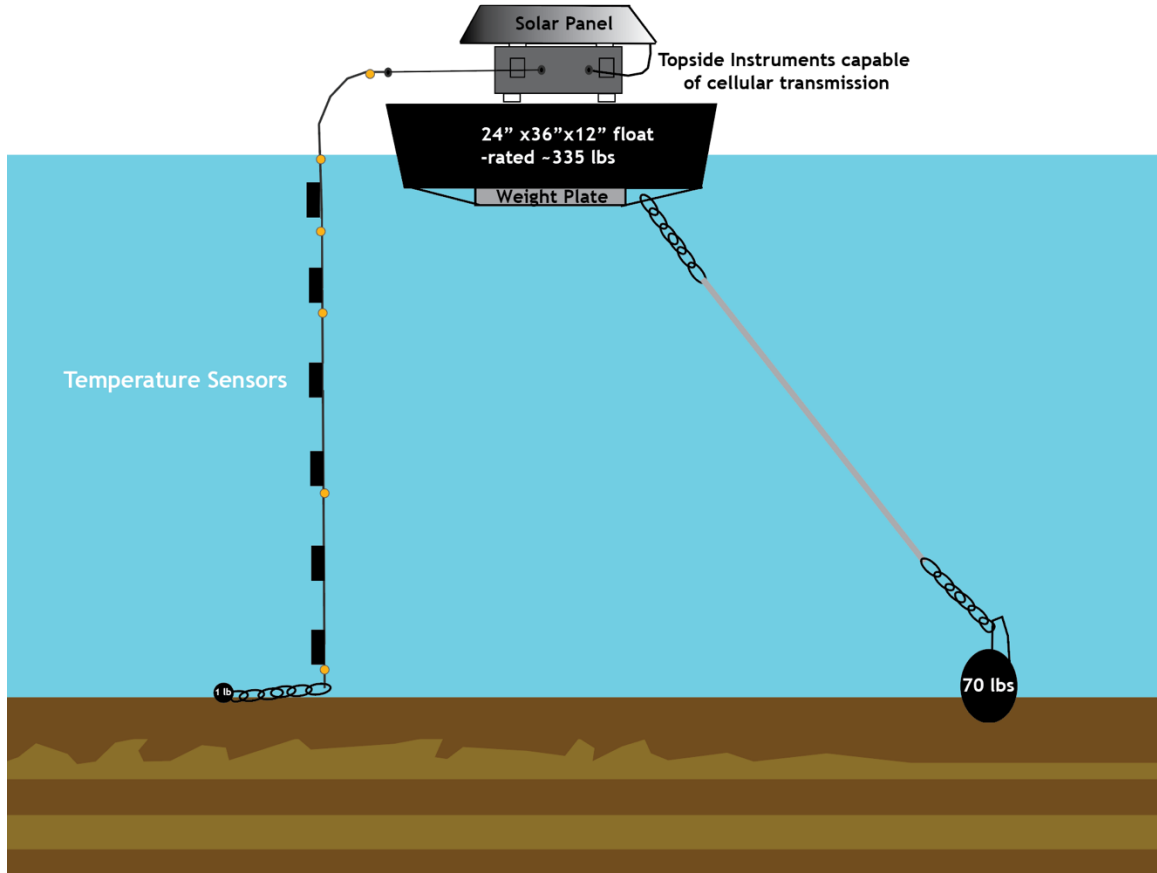
Figure 6: Diagram of the proposed Water Column Monitoring System (WaCoMS)

4.2 Primary Component 1: Data Collection Instrument

As previously mentioned both the HyLO-Mo and the WaCoMS systems have two

different primary data collection instruments. The HyLO-Mo utilizes the previously

developed HyLO probe for the collection of subsurface temperature data. The WaCoMS

probe utilizes several thermocouples spread out in intervals along a cable strand for the

collection of water column data.

4.2.1 HyLO-Mos HyLO Component

The HyLO probe utilized with this system is deployed via a strength member

which doubles as a data and power transmission cable. This cable transmits power from

the topside float to the HyLO probe inserted into the water bottom sediments and

transmits data from the probe to the topside microcontrollers. Possible lance

configurations for the HyLO include 1-meter, 1.5-meter, and 2-meter lances, each with

up to 8 thermocouples. In softer sediments the longer lances can be utilized to facilitate

the collection of deeper data points that are less impacted by seasonal bottom water

temperature variations. Additionally, the thermocouples strings used in the 2-meter

lances are designed so that the thermocouples are more heavily concentrated at the

bottom of the lance to ensure sufficient data collection in the case that an incomplete

insertion occurs.

4.2.2 WaCoMS Cable Strand

The need for long-term time-series water column temperature data to reduce heat

flow uncertainty was heavily discussed in *"Heat Flow on the U.S. Beaufort Margin,*

Arctic Ocean: Implications for Ocean Warming, Methane Hydrate Stability, and

Regional Tectonics" and has already been mentioned in this thesis [6]. The creation of

an instrument to collect such data is simple, with the design consisting of a cable with

several thermocouples attached in intervals. Initially, a single cable system was

developed with several DSB18B20s attached to it, the benefit to such an approach was

the ability to have only one data line and use sensor addresses to differentiate between

data streams. However, the DSB18B20s only had a precision of 0.1 C and to collect

more precise data we opted to utilize TMP36s. When using an 18-bit analog-to-digital

converter the TMP36s can record measurements within 0.0018 degrees C. The one

downside to the use of TMP36s was that there had to be one data line per thermocouple,

resulting in a cable bundle for several thermistors; however, we felt that it was worth the

extra precision in the measurements. Because of the separate data line requirements of

the TMP36s, five separate cables were created. Each cable had a TMP36 encased in

epoxy attached to the end. The cables were then zip-tied to each other resulting in one

long strand with thermocouples spaced at 5-, 10-, 15-, 25-, and 35-meter lengths. A small

(2lb) lead weight was then attached to the end of the stranded cable to keep it vertical in

the water column. Figure 6 shows one of the TMP36 thermocouples attached to a 35-

meter cable and encased in epoxy.

These thermocouples all transmit their analog data back up to the topside float

where the data is digitized and transmitted to an AWS EC2 instance (Amazon Web

Services Elastic Compute Cloud Instance) for processing and storage. In addition to the

real-time TMP36s, I included 11 Onset HOBO temperature data loggers along the cable

strand in 3-meter increments. These data loggers were included to ensure system

redundancy if the TMP36 data transmission went down. These data loggers were set to

have a one-hour sample rate and log all data to a local disk. The data could then be

recovered at a later date when the physical devices were recovered. Following are

several pictures of the WaCoMS data cable strand. Interestingly, in terms of cost

comparison, the HOBO systems (standard temperature technology used by scientists)

each cost $600 (a total of $6600 for 11), have a precision of 0.1 deg. C, a 1-hour sample

rate only, and provide no real-time data, so it is unclear if these systems will fail over the course of the year and provide any data. In contrast, the HyLO system currently provides data to us continually at an instrument construction cost of ~$2000 with real-time temperature at five locations (easily upgradeable to 8), every ~30 seconds to the cloud at 1 order of magnitude better precision.
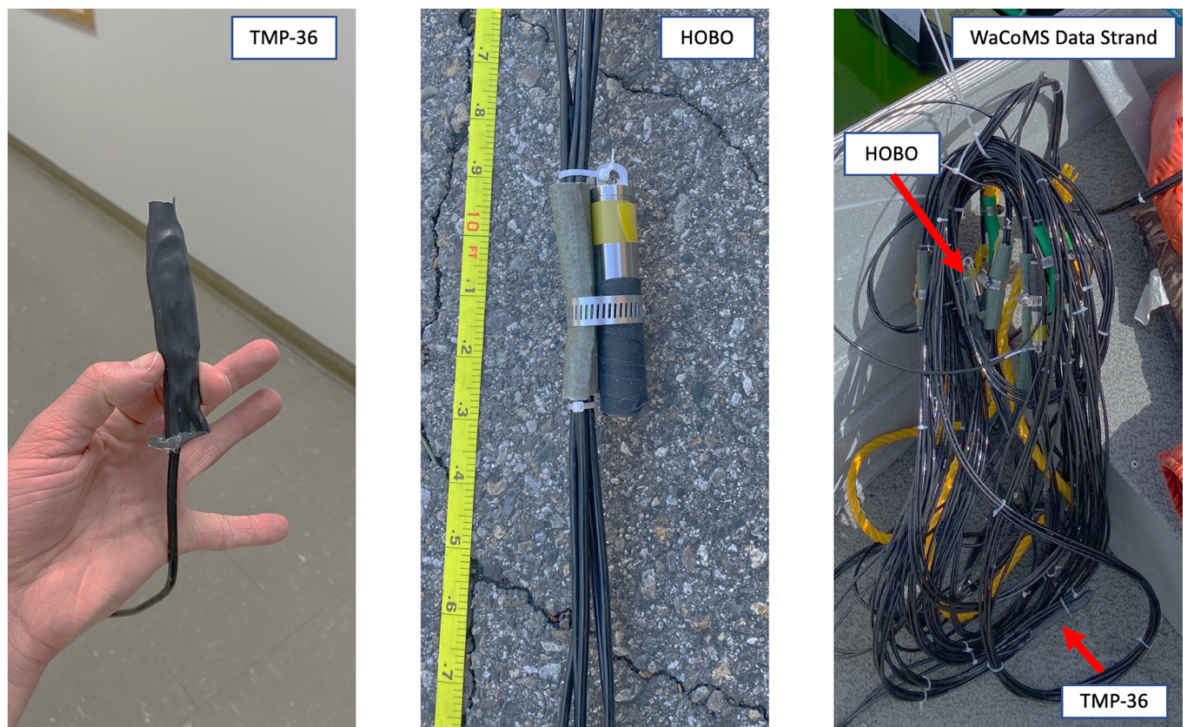
Figure 7: The leftmost image shows a TMP36 data cable with the thermocouple encased in high thermal conductivity epoxy and heat shrink. The middle image shows an Onset HOBO data logger attached to the TMP36 data cable strand via zip ties and hose clamps. A segment of garden hose was wrapped around the TMP36 data cable strand to prevent chafing against the HOBO data logger. The rightmost image - taken just before deployment in Mono Lake, CA - shows the whole cable with both TMP-36 thermocouples and HOBO dataloggers attached.

4.3 Primary Component 2: The Float and Topside Equipment

A watertight Rigid case was used to enclose all of the topside microcontrollers and batteries. The configuration of the onboard instruments differs slightly between the HyLO-Mo and WaCoMS.

The Rigid case for the HyLO-Mo has a 60-meter cable attached to it which is tethered to the HyLO probe. The cable provides power, ground, and data to and from the probe to the float. The single data line within the cable allows for TX/RX (Transmit and Receive) communication between the *wet-end* microcontroller contained within the deployed HyLO probe and the *dry-end* GSM (Global System for Mobiles) modem and data transmitter capable microcontroller in the Rigid case. The GSM microcontroller is responsible for taking data it receives on its RX line, packaging it, and sending it to an AWS EC2 instance for processing. Included in the data stream are 8 digitized thermocouple data points, a probe head humidity data point, a probe head temperature data point, a pressure data point, and three accelerometer data points. Due to the fire-and-forget nature of TX / RX serial communication, delimiters were included in the *wet-end* microcontrollers data stream to delineate between each successive set of measurements.

In addition to the GSM microprocessor is an Arduino 32u4 used to write the measured data to an onboard SD card. This was done to ensure redundancy in the HyLO-Mo system so that in the event of GSM failure, a backup of the data would still be retrievable. The Arduino 32u4 also receives its data stream via RX / TX serial communication. As an additional redundancy measure, hourly power cycling was added

to all components via a watchdog clock that auto-cycles power each hour. During initial experiments, several instances were encountered where fluctuations in the power supply or GSM connection issues led to stalls in the GSM microcontroller operations. The implementation of routine power cycling ensured that in the case of a microcontroller stall, the system could reset and recover within an hour. A UCTRONIX MODEL U6030 watchdog was used to implement this hourly power cycling. The watchdog was rated to 10,000 total recycles; a year-long deployment only requires ~8700 recycles.
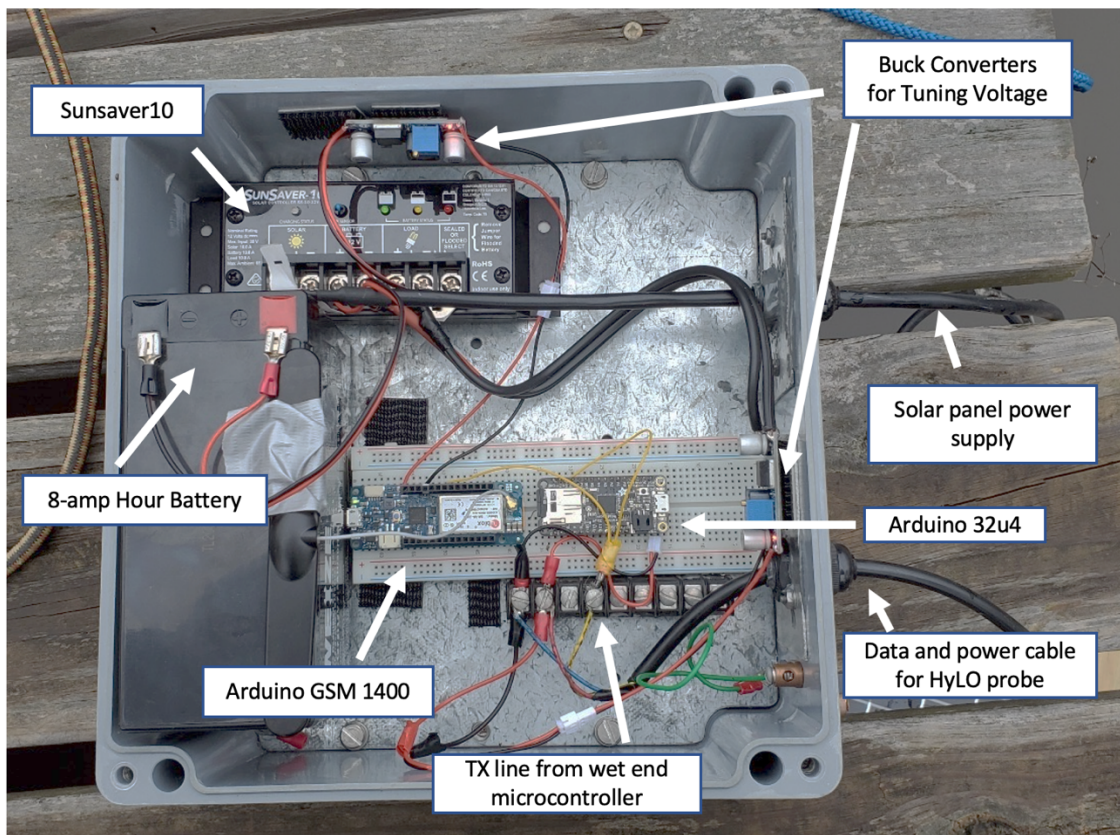
Figure 8: This image shows a condensed version of the HyLO-Mo equipment case with an 8-amp hour battery. Not pictured here is the UCTRONIX Watchdog.

The Rigid case for the WaCoMS has a similar setup to the HyLO-Mo with a few minor variations. In the WaCoMS system, the data cable strand essentially acts as the thermocouple lance does in the HyLO-Mo system. Therefore, the *wet-end* microcontroller used with the HyLO-Mo system is now moved topside and placed in the Rigid case. It is still responsible for digitizing the analog TMP36 thermocouple data and packages it for transmission via RX / TX serial communication to the GSM microcontroller which then transmits the data to an AWS EC2 running a NodeJS web server. Unlike the HyLO-Mo, an onboard Arduino 32u4 was not included for backup data storage. The reasoning behind this design decision was that the Onset HOBO data loggers attached to the thermocouple strand could serve as such a backup given that they are writing data on their local disks. Therefore, an additional datastore on an Arduino 32u4's SD card was deemed unnecessary.
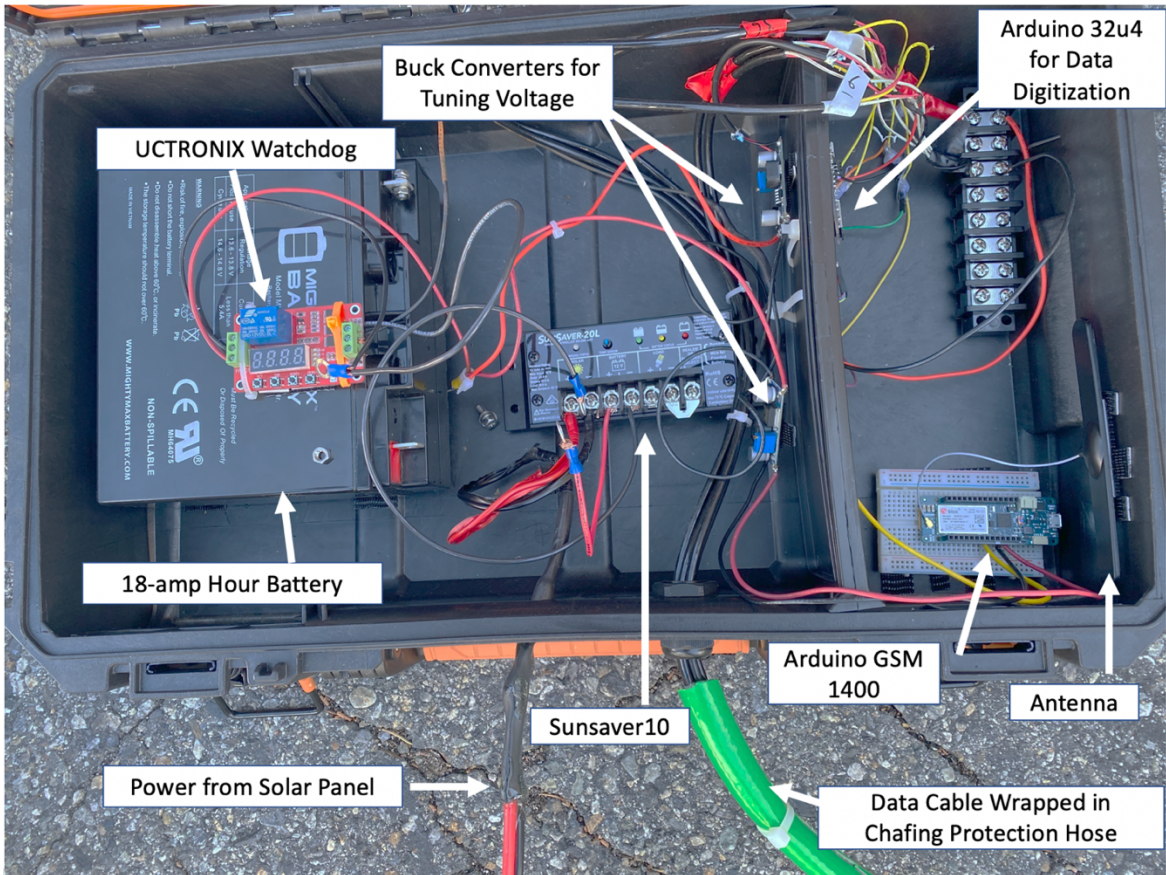
Figure 9: The WaCoMS topside instrument Rigid Case

With two microcontrollers being topside in both the HyLO-Mo and WaCoMS probe and a third *wet-end* microcontroller being included in the HyLO-Mo system, the peak power requirements during data transmission could reach 2 amps depending on the range to the closest cell tower. After experimenting with several different battery configurations, a 12-volt, 18-amp hour AGM (Absorbent Glass Mat) battery was chosen. It should be noted though that the system is functional with an 8-amp hour battery, but 18-amp hours provides for more operational days (approximately 5 dark/cloudy days)

before battery depletion. The 18-amp hour battery weighed approximately 15 pounds which adds a significant amount of top-heaviness to the float. However, due to the improvements in the float design, which will be discussed later, the luxury of a battery with an extra 10-amp hours was affordable.

Solar power was deemed the appropriate choice to maintain the battery's charge and allow the probe to theoretically operate indefinitely. The panel needed to have a peak amperage which was no more than 25% of the battery's total amp-hours. Therefore, a 4-amp, 12-volt, 50-Watt solar panel was chosen. A sunsaver10 controller was used, rated up to 10-amp draw, to regulate the charge going into the battery from the panel. Once the battery was fully charged the controller would stop the flow of charge from the panel to the battery. It was estimated that in the event of absolute darkness, the battery will be able to sustain probe operations for approximately 5.5 days. However, during experiments with heavy cloud cover, the solar panel has still been shown to produce ample power, around 12 watts. Additionally, icing of the panel will be of little concern in saltwater environments, further increasing confidence in the system. We initially planned to use Bird Away Flex Steel to prevent birds from landing and defecating on the panel, however, this plan was abandoned for fear of damaging the panel when attaching the steel spikes. The one concern remaining with the panel is that when deployed in saltwater environments, sea spray residue can accumulate on the panel and hinder its performance. However, rain events should help to clean such residue.

The Rigid case and solar panel discussed above were designed with the vision that they would be attached to a floating platform at the water's surface. However, we quickly realized that creating such a float is a difficult problem. The design required

there to be approximately 30 pounds of equipment kept safely above the water line and the float carrying this equipment would need to resist capsizing even at high degrees of list.

The first prototype consisted of connecting several lobster buoys together via galvanized iron pipe. The equipment was then band-strapped to the top of the constructed float as pictured below. However, upon placing this device in still water it would capsize almost immediately.
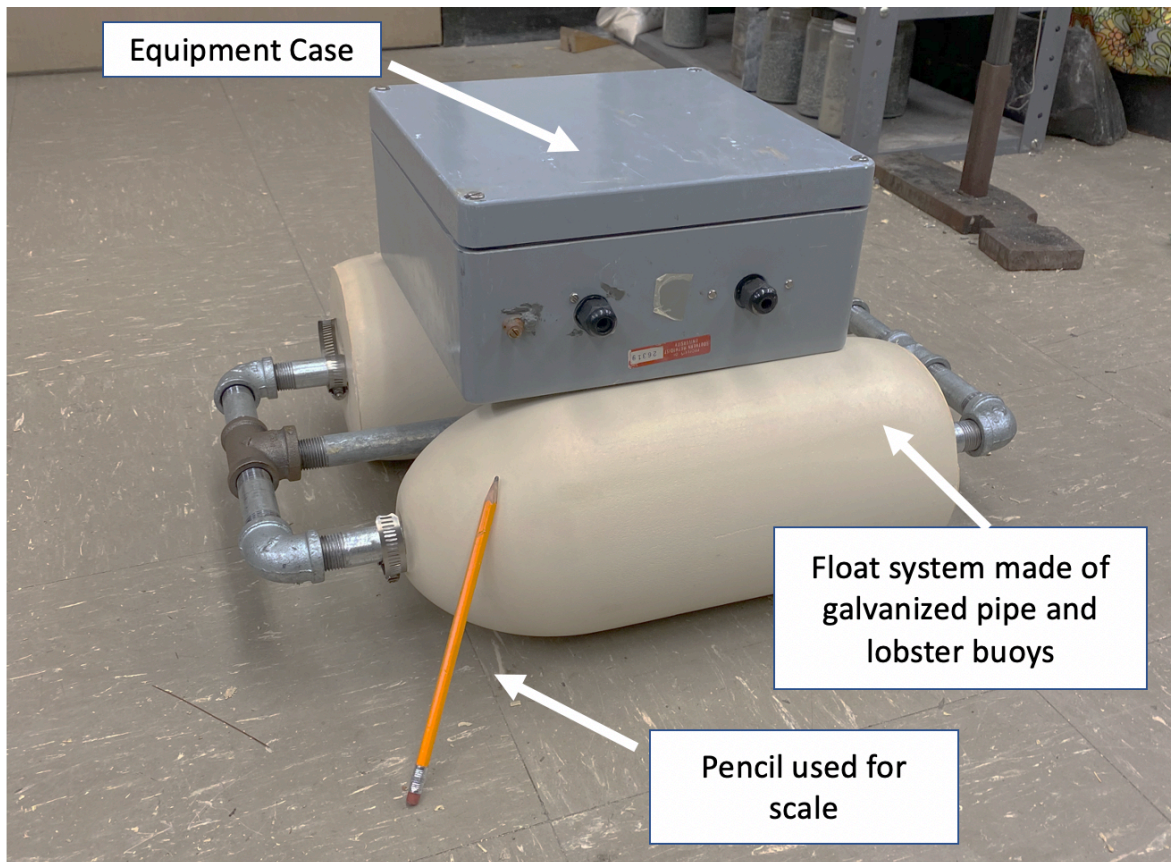
Figure 10: Failed lobster buoy configuration

As an alternative approach, we turned to using dock floats. A dock float is a hard plastic box filled with Styrofoam and is typically attached to other dock floats to construct temporary floating docks. One float is a 36in x 24in x 12in Styrofoam-filled float with a buoyancy rating of close to 400 pounds. This extra spacing and high buoyancy rating meant that a bigger solar panel could be attached and allow for bigger batteries with more amp hours.

Despite the improvements in the float's design, the results of the first test with this system did not go much better than the first. Due to the float's high buoyancy rating, most of the float would remain above the waterline when placed in still water. Adding equipment to the top of the float made it top-heavy and susceptible to capsizing with more than a 20-degree list. To counteract this instability additional weight was added to the bottom of the float, this weight acted as a ballast stone. This helped to stabilize the float and make it less top-heavy once the Rigid case and solar panel were strapped to the roof. The extra weight (two 25-pound weight plates) was band-strapped to the bottom of the float as pictured below. This addition meant that even with 30 pounds of equipment strapped to the roof, the float could roll close to 70 degrees before a capsize event would occur.
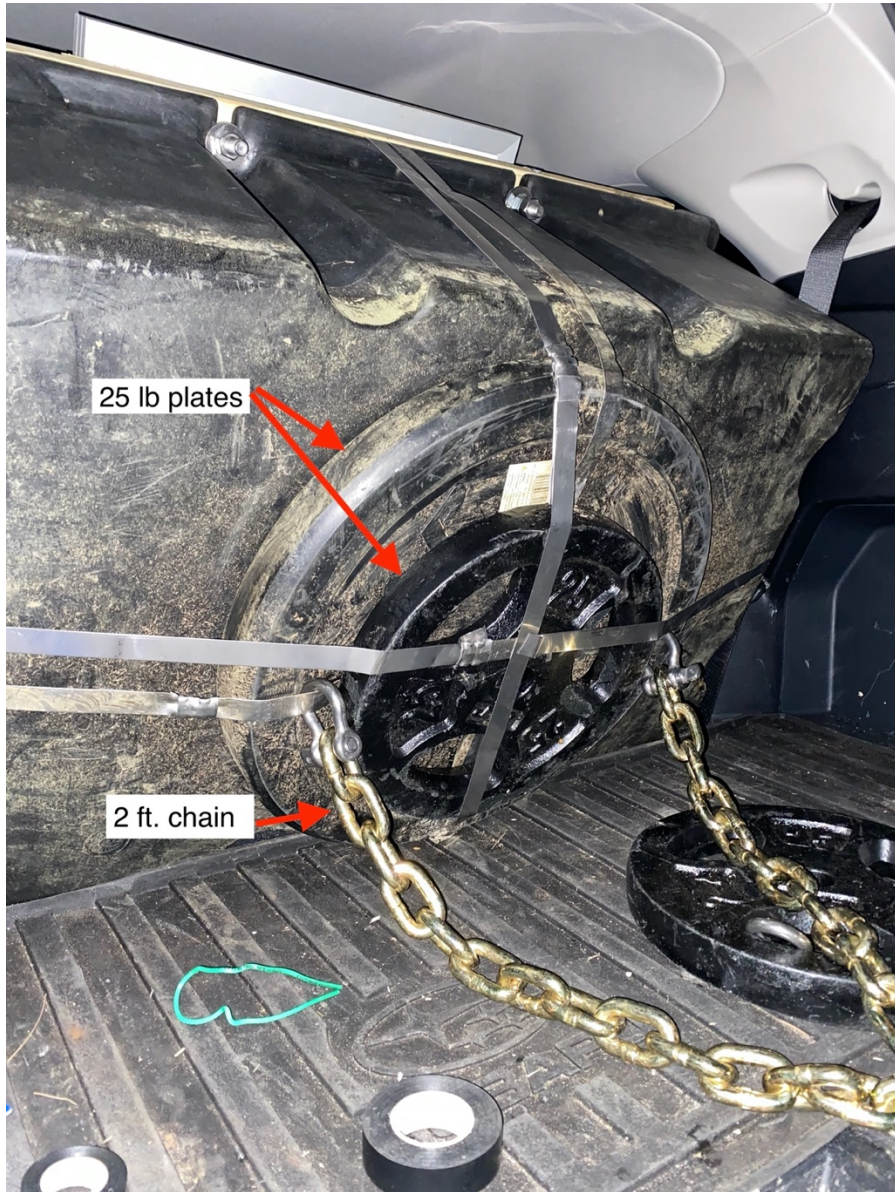
Figure 11: An image of the underside of one of the top-side probe floats. This image shows the weight plates used as ballast stones and the chain used for attaching the anchor system.

Attached to the band-straps was a 2-foot chain used for attaching one or more anchors to the float. Anchors would be necessary to keep the float in place during storms and high sea events. The anchor line was attached to the 2-foot chain hanging under the float so that the anchor point would be in the center of the float and not off to one side which could destabilize the float. Also having the anchor point two feet below the waterline helped to lower the float's center of gravity, further preventing capsizing.

The anchor line itself consisted of a boating anchor attached to a 20-pound chain which would sit on the water bottom. This chain was then attached to a 45-meter rope. The rope was then secured to the float's 2-foot chain with a shackle and chafing gear. Below is a picture of the rope, chain, and anchor. In addition to the anchor and weight of the chain, kettlebells and weight plates were dropped down the line at the time of deployment to further anchor the float in place.
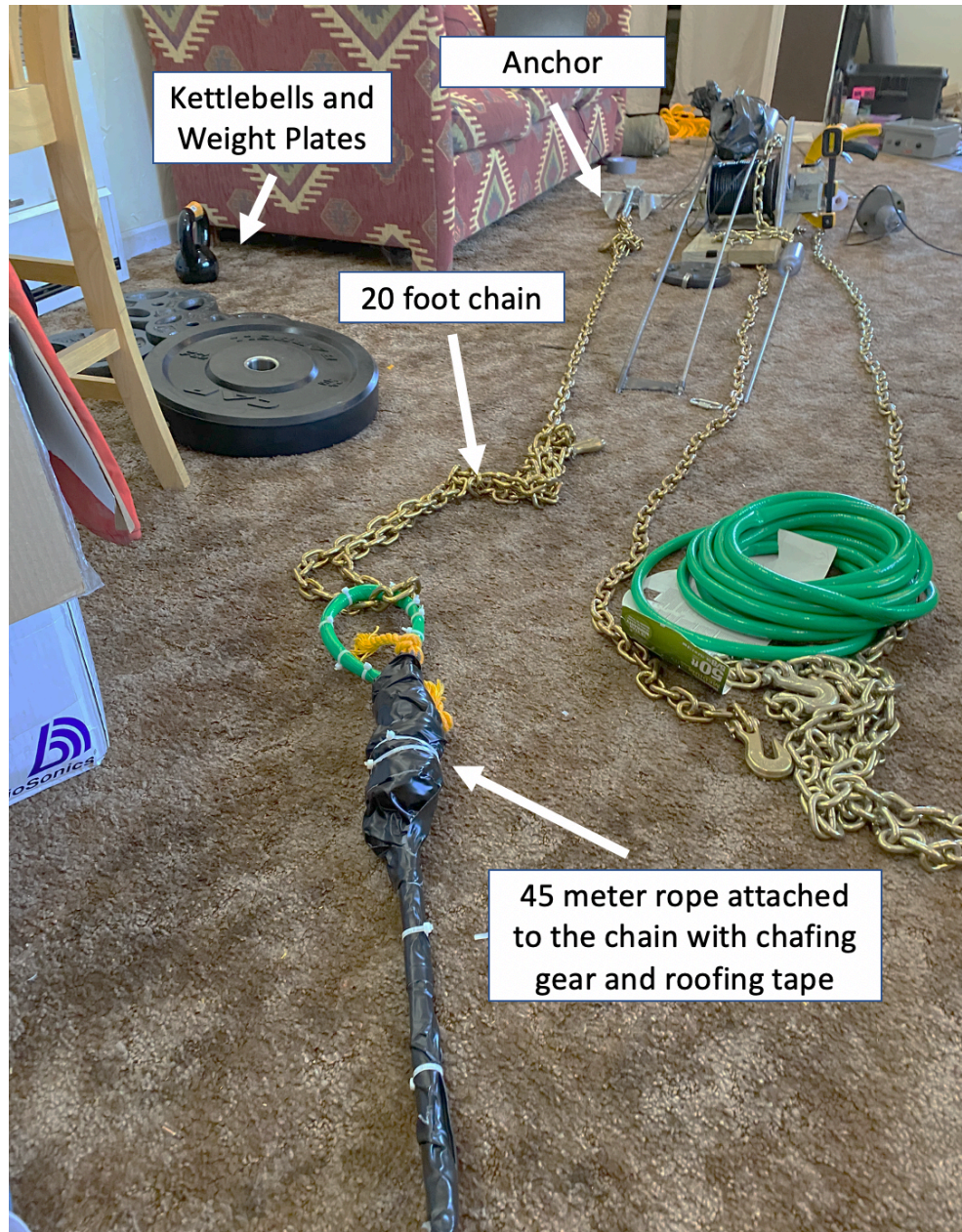
Figure 12: Construction of the probe's anchor system

In the case of the HyLO-Mo, two anchor lines were used. This was done to prevent the float from spinning and prevent the inserted HyLO probe from being dislodged from the water bottom. The WaCoMS is deployable with only one anchor.

A wooden brace was attached to the roof of the float using stainless steel ½ inch bolts and nuts. This brace was used for mounting the solar panel and securing the Rigid case. The brace consisted of 2 2x4s and a single 2x6. A hose clamp was then screwed into each of the 2x4s. The Rigid case was then placed between the two boards and tightening the hose clamps down locked in the Rigid case into place. The solar panel was then screwed into the 2x6 board and had the added benefit of shading the Rigid case to prevent overheating (we later realized overheating was still an issue and spray-painted the Rigid cases white to help reduce the heat absorption from sunlight).
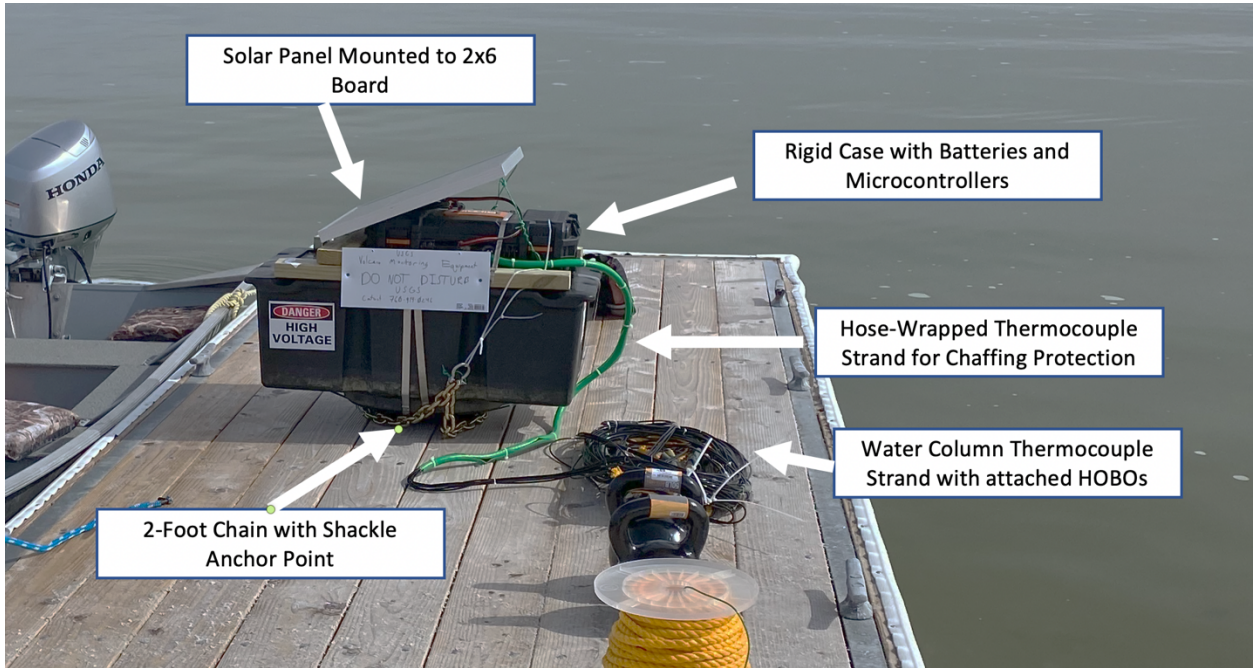
Figure 13: The complete WaCoMS system shown on a dock prior to deployment in Mono Lake, CA
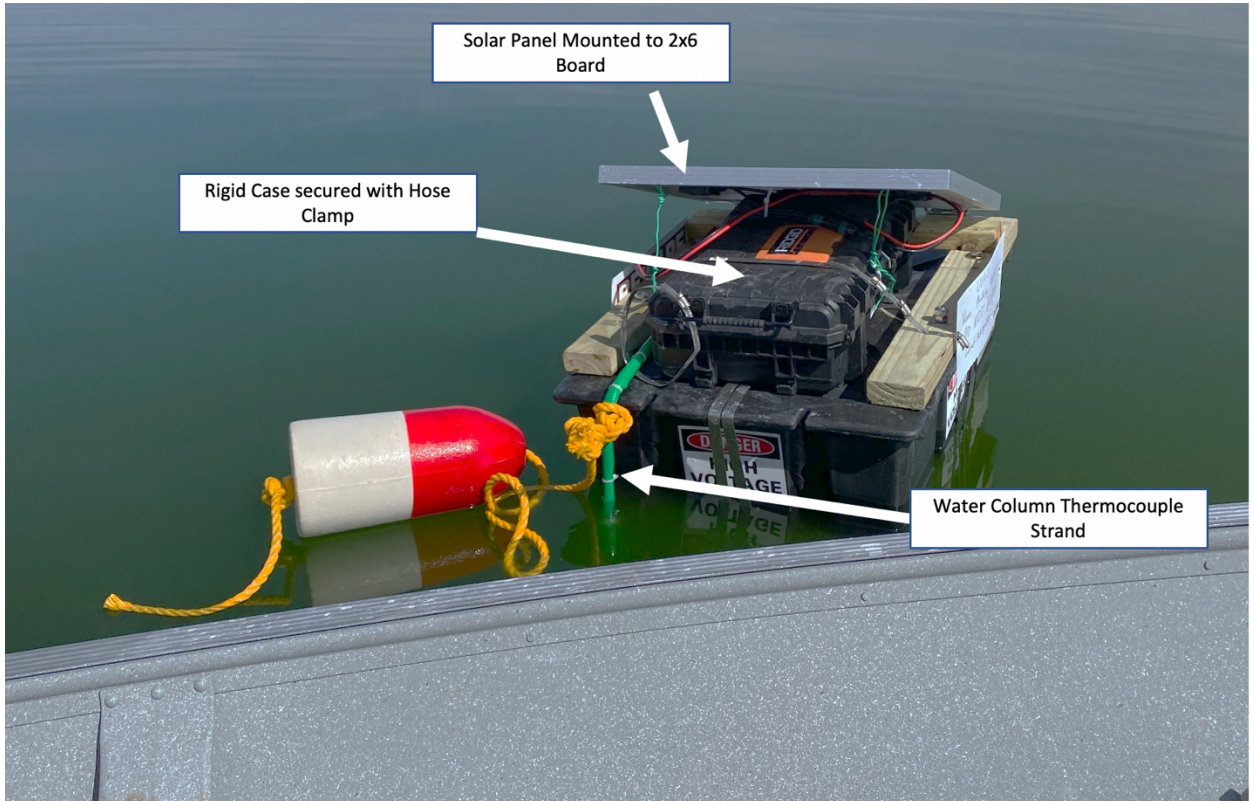
Figure 14: Deployed WaCoMS device in Mono Lake, CA

4.4 Primary Component 3: AWS Data Processing and Storage

Recall that there are two sources of data collection on two separate real-time probes: The HyLO-Mo's HyLO probe inserted into lake bottom or sea bottom sediments, and the WaCoMS's water column thermocouple strand.

Once a measurement has been transmitted to the float-board microcontrollers from either the HyLO probe or the water column thermocouple strand, the GSM microcontrollers will transmit this data to AWS for processing and storage. Traffic is directed to a domain hosted on AWS Route53 which then sends the traffic to an AWS

EC2 instance running a NodeJs webserver. The Express backend handles the incoming

requests and stores the relevant measurements into an RDS database instance. The

recorded measurements are timestamped on the server-side when the request is received.

This server-side timestamp is stored with the data in the DB entry. The below figure

illustrates the flow of data. Additionally, the EC2 instance running the Express server

also runs an Angular website, www.hylomonolake.com. This website allows individuals

to view the most recently transmitted data from the deployed probe. This provides

interested parties with near real-time data access.

While such an implementation is quite simple, it has never been applied to such a

use case and therefore, revolutionizes heat flow data collection. The utilization of the

Arduino GSM microcontrollers in conjunction with AWS Cloud Resources is the key to

receiving real time data from deployed instruments.
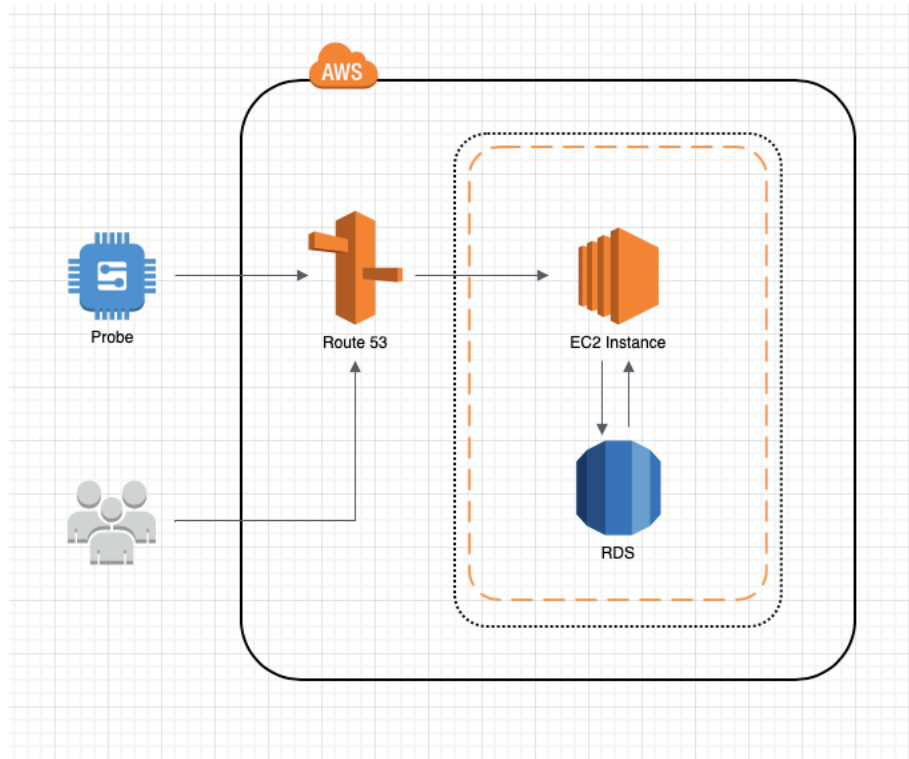
Figure 15: AWS Architecture Diagram for Arduino Data Processing and Storage

CHAPTER 5: DEPLOYMENTS AND RESULTS

5.1 Initial Field Test

An initial test of the HyLO-Mo was conducted in a pond in East Texas to test the system for an extended period of time in a dynamic environment. The deployment is pictured in figure 16. A 1.5-meter HyLO lance was inserted 0.5 meters into the pond sediments with the remaining 1 meter of the lance in the water column. The probe was deployed for a period of 14 days and transmitted data continuously. The deployment period coincided with a large snow/freeze event, which occurred in late February, 2021. The below temperature-depth profile shows that the freeze event led to a slow cooling in the pond sediments which was recorded by our HyLO probe. This was the first time that a HyLO instrument was successfully deployed for an extended period of time while still having access to the real time data through AWS resources. Furthermore, the probe did so in a below-freezing environment.

Figure 16: Field deployment of the HyLO-Mo System
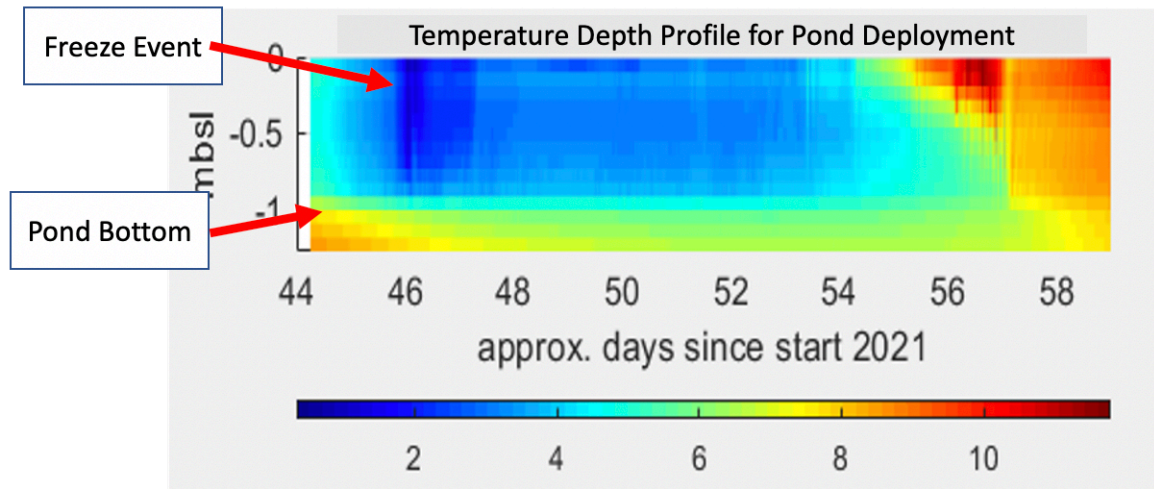
Figure 17: Temperature Depth profile over time for the East Texas pond deployment

5.2 Mono Lake

After the successful two-week field test in East Texas, our group participated in a USGS and NSF-funded study to collect heat flow data in Mono Lake, California as a part of a volcanic study. Our role in the USGS project was twofold. Our primary objective was to collect heat flow data in the form of spot measurements throughout the lake using the standard HyLO probes. Our secondary objective was to deploy the HyLO-Mo and WaCoMS probes to collect time series data on a longer timeframe to offer further insight into the heat flow and fluid flow processes occurring underneath the lake on a longer

timescale to help reduce uncertainty in spot measurements. To deploy the long-term monitoring systems our group utilized a ~20-foot skiff provided by the Mono Lake Committee. Deploying either the HyLO-Mo or WaCoMS micro-observatories took no more than an hour, considerably less time than the several days or weeks required for the drilling and installation of traditional CORK observatories or for conducting ROV deployments.

The WaCoMS micro-observatory was deployed on March 24th, 2021. The micro-observatory transmitted data up until October 18th, 2021. During the nearly seven months of up time the probe transmitted data to the cloud with a sample rate of ~30 seconds. There were occasional disruptions in WaCoMS data transmission where the data flow was be halted for several minutes. During our time in Mono Lake, it was noticed that there would be intermittent ATT connection issues; therefore, the breaks in WaCoMS data transmission could be attributable to such events.

The HyLO-Mo system was deployed on March 27th, 2021. A 1-meter thermocouple lance configuration was used and data from the deployment indicate that the HyLO lance achieved complete 1-meter penetration into the lake bottom. A temperature-depth profile obtained from the HyLO-Mo micro-observatory shows a general warming trend which is expected as the days have been getting warmer since March 27th. Several vertical jumps are also observable in the data. We attribute these to the release of bubbles from gas seeps. These seeps are common throughout Mono Lake and on calm days, large bubble fields are readily observable on the lake's surface. Additionally, seismic data collected during a USGS survey 12 years ago suggests the overwhelming presence of columns of bubbles rising from the lake bottom.
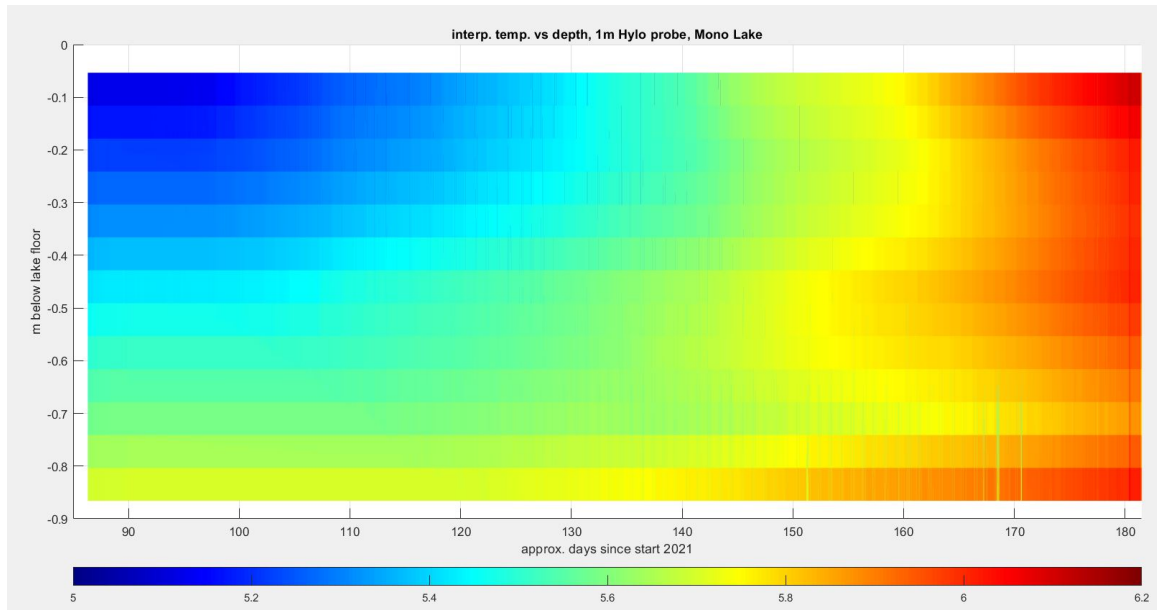
Figure 18: Temperature depth profile of the data collected by the HyLo-Mo system

On April 9th, 2021 the HyLO-Mo micro-observatory stopped transmitting in the night. From then until June 30th 2021, the HyLO-Mo micro-observatory only transmitted data during the daylight hours; therefore, we hypothesize that the probe experienced a battery failure of sorts. From April 9th to June 30th, onboard instruments were still operational using solar power and would transmit data periodically throughout daylight hours, cloud coverage permitting.

It was noted earlier in this thesis that failure events are expected in the deployment of geophysical instruments. It is reassuring that even after experiencing a failure event, the HyLO-Mo system was still capable of recording and transmitting some

data. Because long-term time series trends are the interest, the impact of being only 50%

operational is negligible.

Despite the eventual failures of both the WaCoMS and HyLO-Mo, our group was

still able to retain the previously collected data thanks to the micro-observatory's real

time data transmission capabilities and the data being stored in AWS RDS. This would

likely not have been possible in the previously discussed existing systems where

measurements are only recorded to onboard dataloggers.

Figure 19: The WaCoMS on a skiff shortly before deployment by a three-man team

HyLO-Mo

WaCoMS

Figure 20: The successfully deployed HyLO-Mo and WaCoMS systems in Mono Lake

CHAPTER 6: GRANGER CAUSALITY-BASED TIME SERIES CLUSTERING

6.1 Introduction

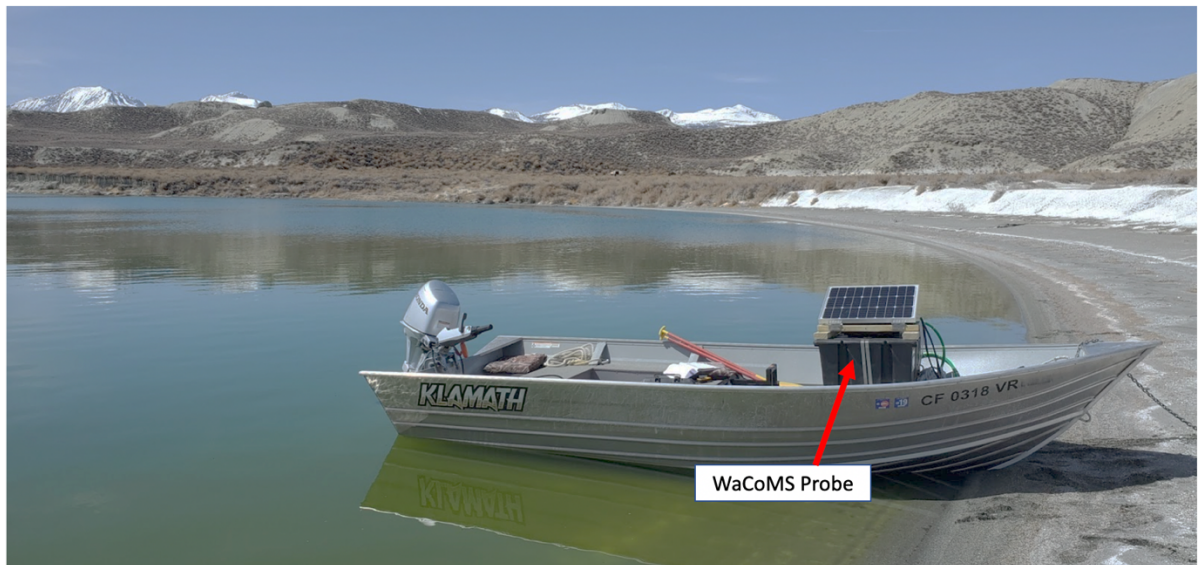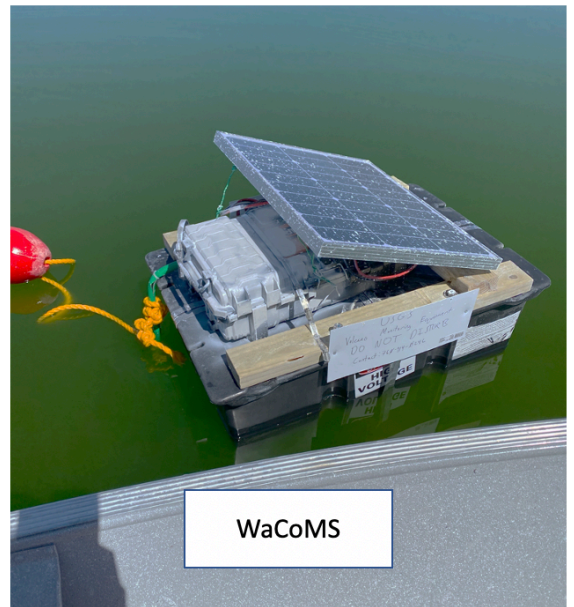At the conclusion of the deployment and after several months of data collection, a novel Granger Causality time series clustering approach was used to analyze the WaCoMS data from Mono Lake, CA. One application of the WaCoMS data not previously touched on is that water column time series data can be used to identify areas of dramatic density change. Differing salinity levels drive these density changes and as a consequence sudden changes in temperature can be observed at such boundaries. Therefore, drastic temperature changes observable in the dataset can be associated with drastic density changes. In lake and marine environments, the primary area of density change in the water body is defined as the thermocline.

In addition to the temperature data from the water column obtained by the WaCoMS, I also had access to several sensor data including surface barometric pressure and surface accelerometer data which can serve as a proxy for wave action. I wanted to investigate the influence between these external sensor data and fluctuations in the water body temperature data collected by the WaCoMS. Clustering was used to assess which thermocouple datastreams are similarly affected, but instead of using a traditional correlation method as the basis for clustering, a novel Granger causality approach was chosen in an effort to better capture the forecastability of external factors on the measured thermocouple data. Knowing these relationships can offer insight into the depths of density boundaries (where drastic temperature changes are observable) and the drivers of the amplitudes, and variability of internal waves which occur along such boundaries.

The results of this Granger-based time series clustering approach on the WaCoMS data will be cross-referenced with low-resolution historical Mono Lake data collected by the Los Angeles Department of Water and Power (LADWP). The LADWP dataset will act as the ground truth dataset when it comes to determining the lake's historical thermocline location.

6.2 Granger Causality Based Time Series Clustering

Time series clustering is a data mining method that groups time series based on their similarities such that time series clustered together have higher similarities and time series across clusters have less similarity [19]. It is common for clustering to utilize correlation as the grounds for determining similarity. However, the method utilized in this thesis instead uses Granger causality such that time series are grouped based on their level influence on one another whether than their correlation.

Granger causality was first introduced as a method for determining how well one time series forecasts another, essentially looking to see if influence exists between two time series [20]. The method looks at a bivariate autoregressive model and tests whether the variance of the model's residuals is smaller in the bivariate case than in the case when one term is forced to zero and the model becomes univariate. For example, say you have two variables X and Y, and a bivariate model incorporating them. If the Y variable is forced to zero by the model coefficients and the variance of this univariate model's residuals is greater than the variance of the bivariate model's residuals then it is said that Y Granger Causes X. In layman terms if a model performs better with two variables than with one, then this additional variable must Granger cause the first. The amount of Granger causality present, henceforth referred to as the inference statistic, can be

estimated by calculating the logarithm of the F-statistic resulting from an F-test comparing the variance of the model residuals.

To combine Granger causality with time series clustering, we run Granger causality tests on several different time series and aggregate the resulting inference statistics into a matrix referred to as **G**. The inference statistics in this matrix are then inverted such that small p-values which indicate a high magnitude of Granger causality are mapped to large values and larger p-values are mapped to small values. This transformation is performed to make the remaining analysis more intuitive as larger numbers now represent a greater magnitude of Granger causality. The transformation is given below, where $\hat{G} = [\widehat{g_{ij}}]$. Then hierarchical agglomerative clustering with complete linkage is performed on $\hat{\mathbf{G}}$ before a visualization and dendrogram are generated [21]. The resulting visualization can be visually examined for levels of Granger causality within the dataset as well as determining which time series cluster together.

$$\hat{g}_{ij} = 1 - \frac{1}{1 + e^{-\gamma(g_{ij} - t)}}$$

6.3 WaCoMS Granger-Based Clustering Results

I chose to run the Granger causality-based clustering method on the dataset in monthly windows instead of on the entire dataset at once. Seasonal changes occur throughout lake and marine environments, which can lead to shifts in trends from month to month. Such behavior can introduce nonlinear relationships to the model. Because

46

Granger requires linear models, the dataset was subsetted into smaller windows to accommodate this requirement.

Throughout this experiment, external sensor data in the form of barometric pressure and vertical acceleration are the time series being investigated for the magnitude of Granger causality on the water column temperature time series. The initial sample rate on the raw data was approximately every 20 seconds. To alleviate the computational burdens of the Granger method, all data were downsampled to a rate of ~6 minutes and a lag of ~8 hours was used.

The figures below show monthly clusterings from April-July of 2021. All data came from the WaCoMS probe deployed in Mono Lake, CA. When interpreting the results, a point in the clustering with a higher value means that there was higher Granger causality between the external sensor data on the x-axis and the thermocouple at the

specified depth. A point in the clustering with a lower value exhibited lower Granger causality.

Figure 21: Granger Causality Clustering from April 2021 WaCoMS Data

Results from the April experiment indicate that barometric pressure exhibits high Granger causality in the upper sections of the water column and becomes less Granger

causal with depth. The vertical acceleration column indicates this sensor data is highly

Granger causal with the temperature data streams at the 10 and 15-meter depths and has
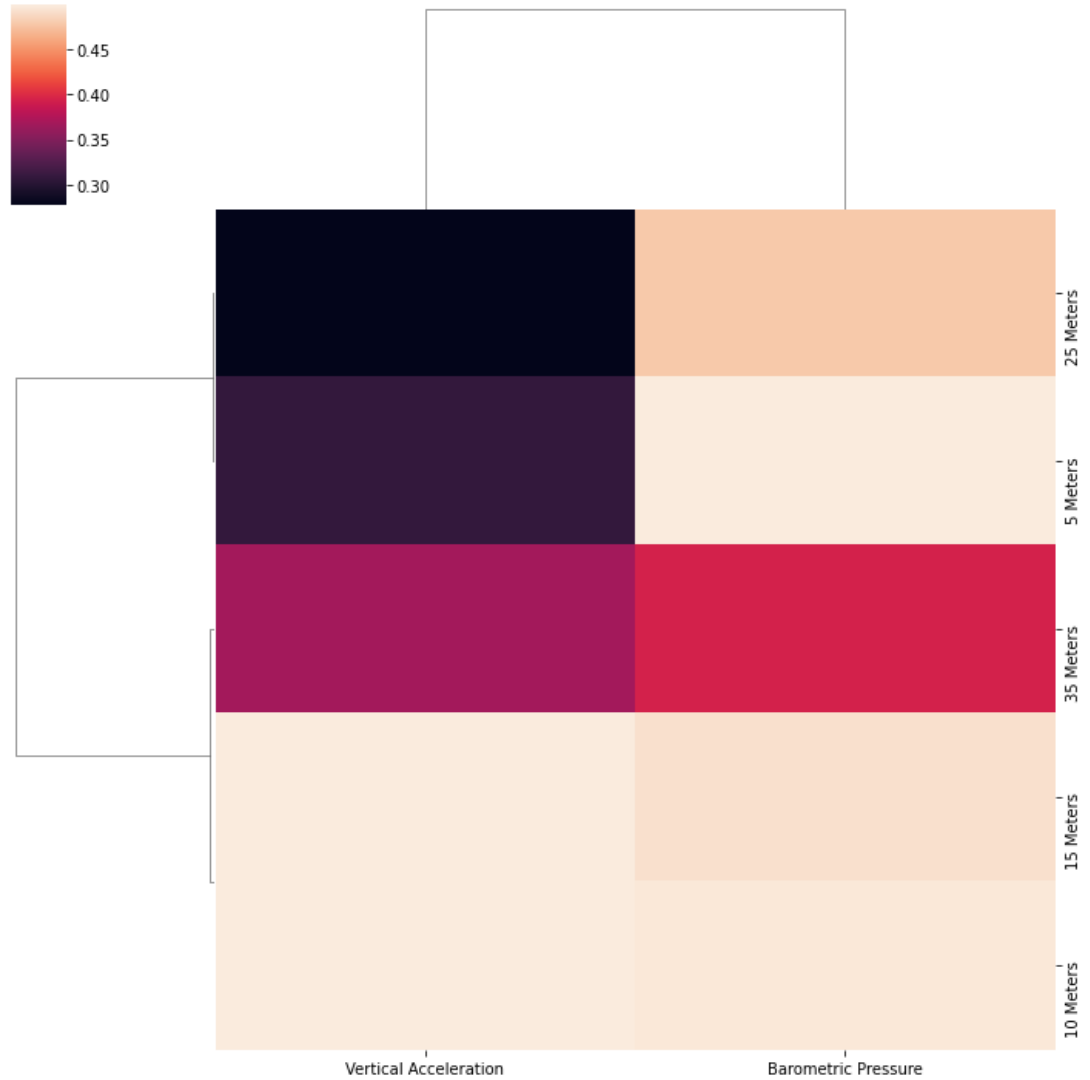
low Granger causality elsewhere.
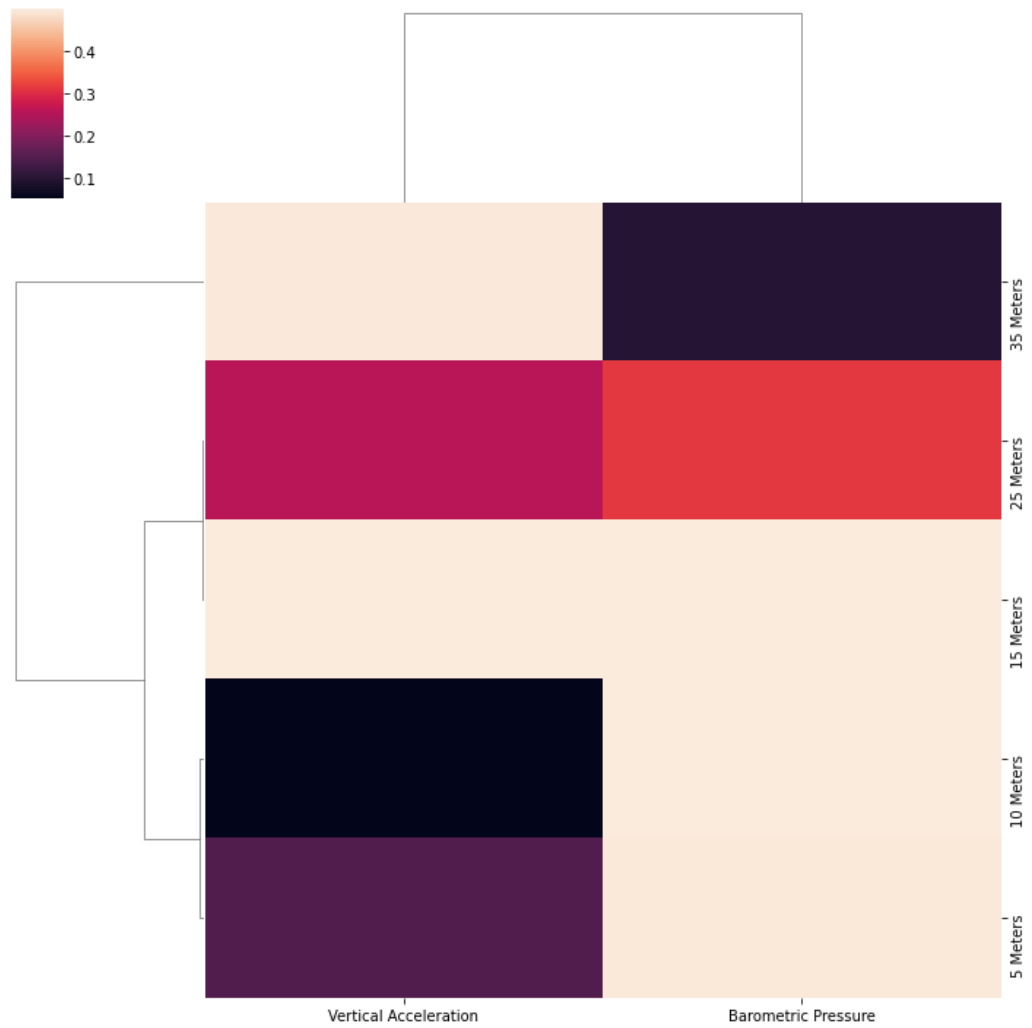
Figure 22: Granger Causality Clustering from May 2021 WaCoMS Data

The clustering from May's experiment indicates that barometric pressure has high Granger causality with temperature data in the upper 15 meters of the water column, but the Granger causality drastically decreases at depths greater than 15 meters. Vertical acceleration shows higher Granger causality with temperature data at 15 and 35-meter water depths and lower Granger causality with temperature data at 5, 10, and 25-meter water depths.
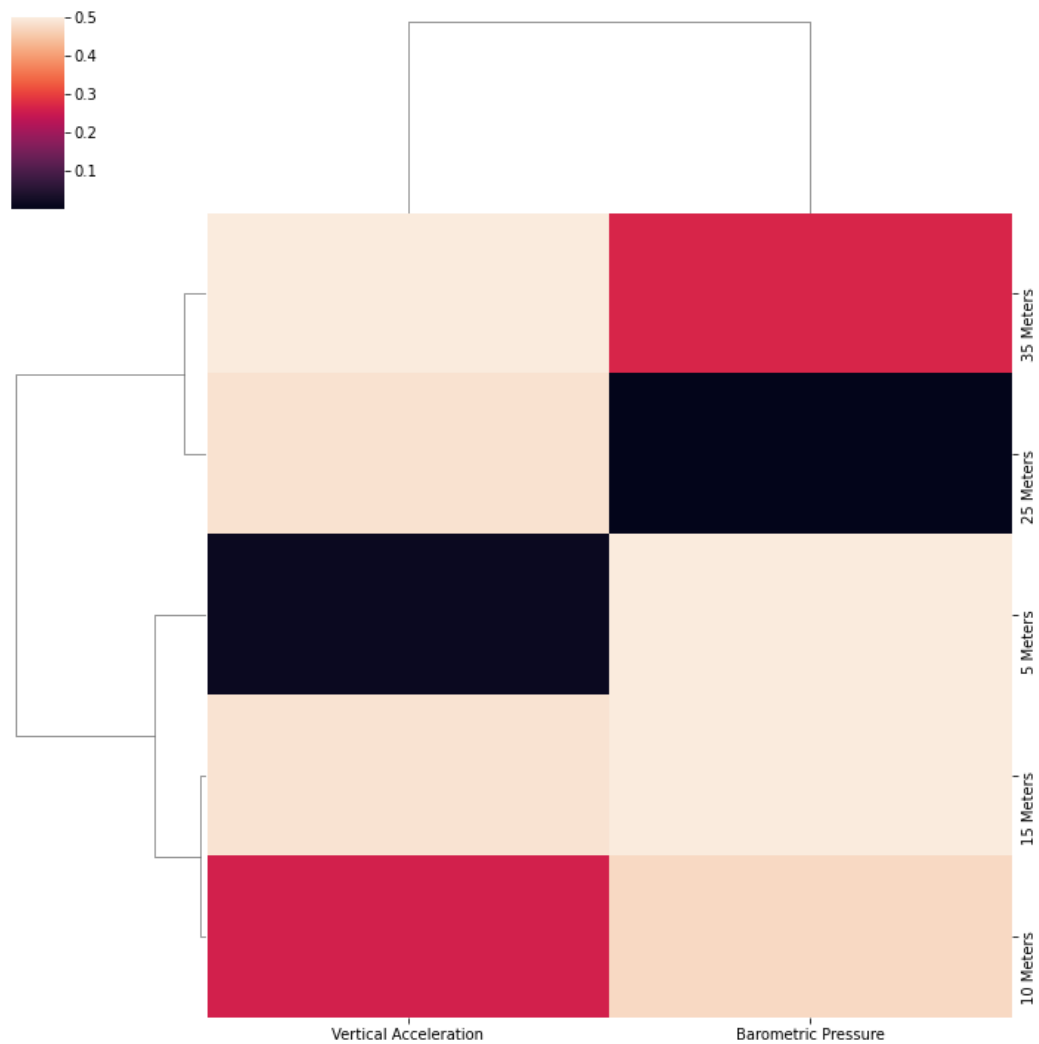
Figure 23: Granger Causality Clustering from June 2021 WaCoMS Data

June's clustering shows that barometric pressure once again exhibits high Granger causality at depths of 15 meters and above with lower Granger causality beneath this depth. The vertical acceleration shows the inverse of this with higher Granger causality at 15 meters and below with lower Granger causality at the temperature data streams from 5 and 10 meters.
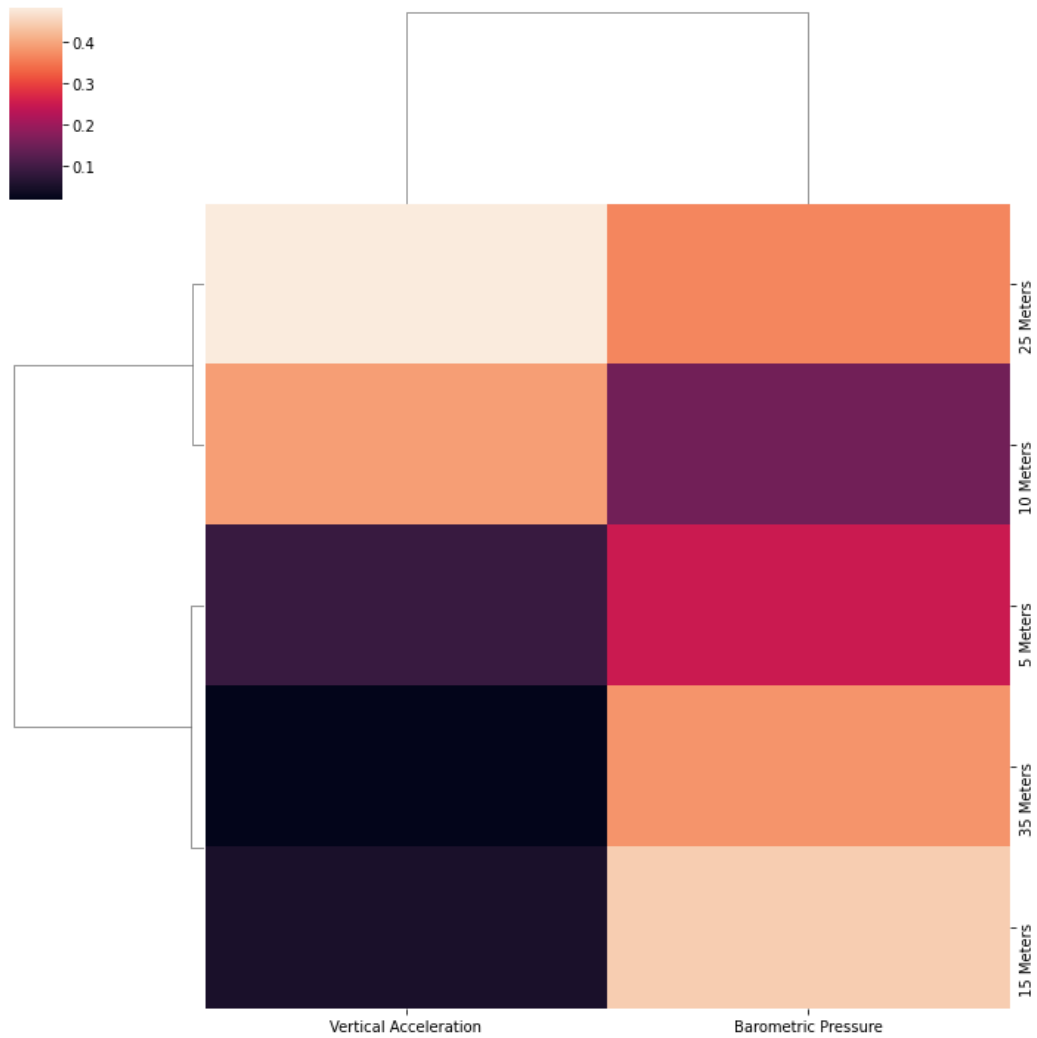
Figure 24: Granger Causality Clustering from July 2021 WaCoMS Data

In the barometric pressure column for July's clustering visualization, there is still high Granger causality with temperature data at the 15-meter depth. However, the above datastreams are less Granger causal than before and the temperature datastreams (while not highly Granger causal) exhibit higher Granger causality than previous months. The vertical acceleration column shows low Granger causality with the temperature data at 5, 15, and 35-meter water depths and higher Granger causality at 10 and 25 meters water depth.

6.4 LADWP Historical Low-Resolution Dataset

The Los Angeles Department of Water and Power (LADWP) regularly collects temperature and conductivity data at Mono Lake at 1-meter intervals in the water column. These data are routinely collected once per month. Therefore, they are only spot measurements and do not provide a holistic picture of lake behavior over the entirety of the month. However, the measurements do provide a rough historical reference point about lake conditions throughout the year. Furthermore, lake conditions would need to be stable on days when measurements were recorded otherwise the boat and crew would have probably not been on the water. If this is the case, then the spot measurements could be treated as a reliable snapshot.

The temperature data is taken with a precision of 0.1 degrees C. The conductivity measurements are in MilliSiemens per centimeters (mS/cm) and the precision is unknown. The conductivity can be used as a proxy for the salinity levels in the lake. Salt is a conductor that allows electricity to flow more easily. As salinity increases, so will the conductivity of the water; however, it is important to note that this relationship is nonlinear. There is little to no information available regarding the actual measurement

method, the instrument being used, or the medium the measurements were made across. Therefore, we are more interested in the general trends in the data rather than the raw data itself. Additionally, it would be reasonable to assume that there is a 2-3 meter margin of error in the depth measurements due to operator error as well as the fact that the spacing between the conductivity and temperature instruments on the line is unknown.
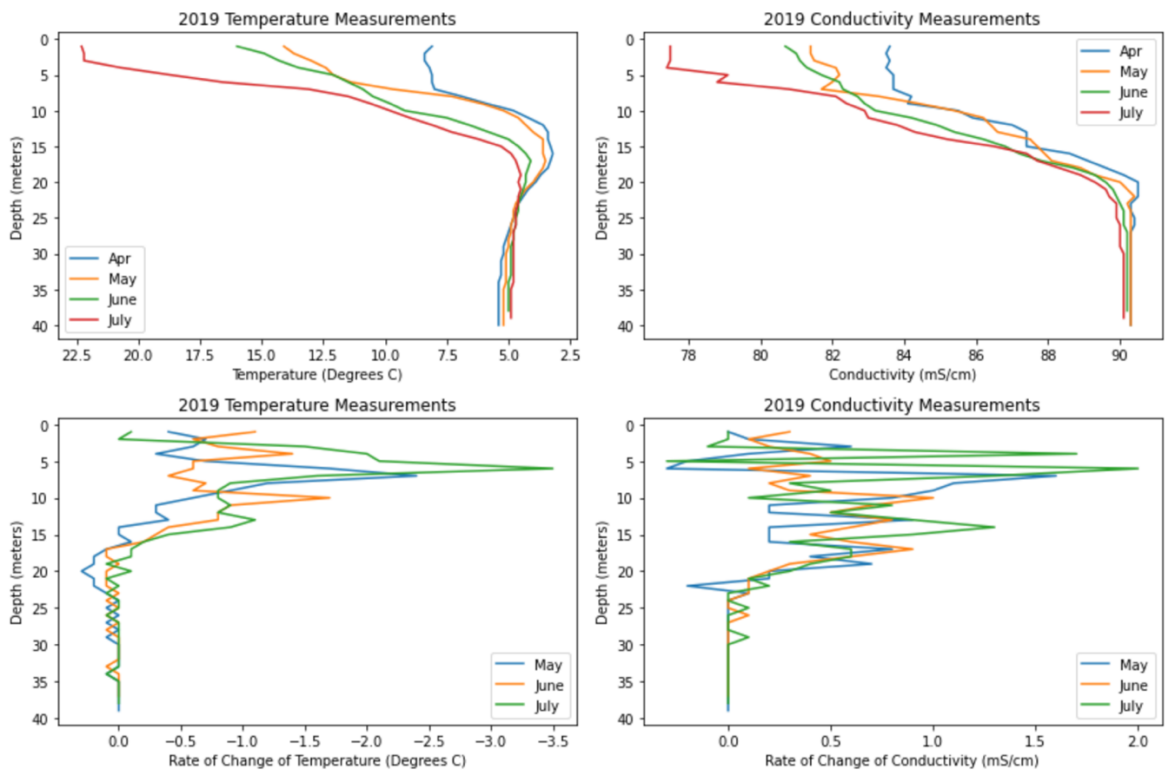
Figure 25: 2019 LADWP Low-Resolution Mono Lake Data
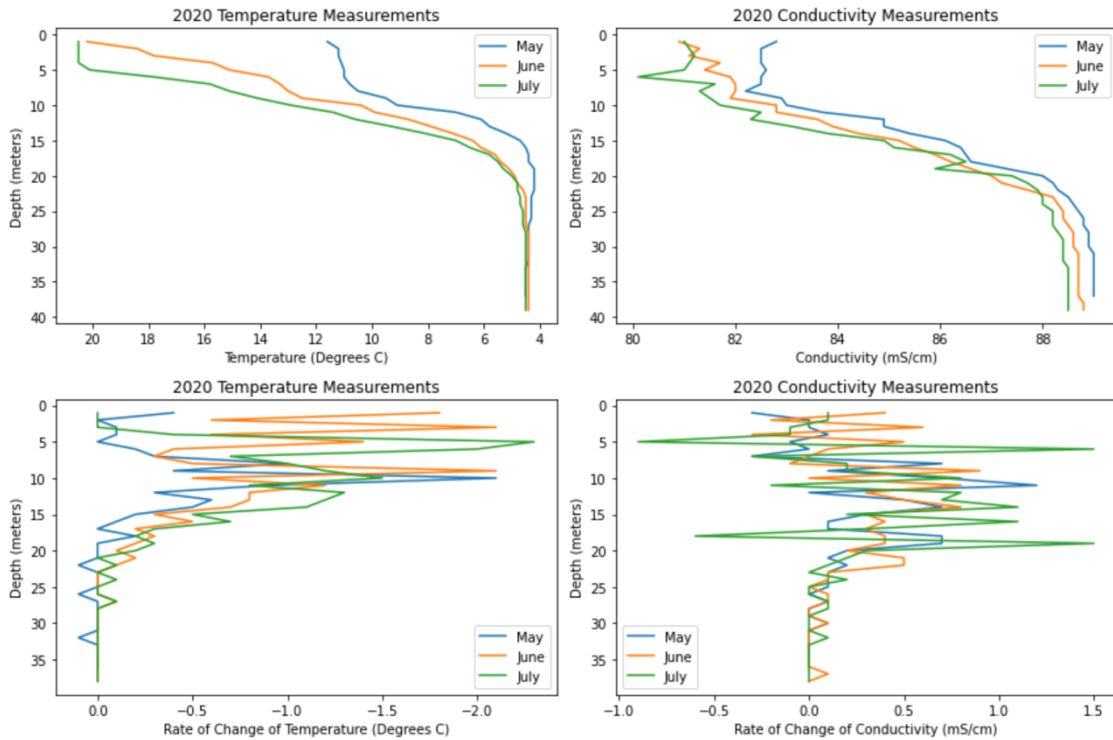
Figure 26: 2020 LADWP Low-Resolution Mono Lake Data

6.5 Discussion

The results shown in section 6.3 help to visually illustrate the different levels of influence that barometric pressure and lake-surface vertical acceleration data have on temperature time series data with depth. With regards to barometric pressure, it is shown to consistently exhibit high Granger causality with shallower temperature time series data in all months except for July. Furthermore, barometric pressure is highly Granger causal with the 15-meter temperature time series in every month, including July. Beneath that point, the Granger causality is consistently lower. This could suggest the presence of some boundary at or below 15 meters which changes the water properties such that

whatever relationship existed at and above 15 meters between the atmospheric barometric pressure and water column temperature data is no longer present. An additional observation is that the 10-meter and 15-meter temperature time series cluster together in every month except for July, indicating that these two time series are affected similarly by the time series on the x-axis.

When analyzing the vertical acceleration column, the results are less conclusive than with barometric pressure. A granger causal relationship between vertical acceleration and water column temperatures is shown to be present at different depths. throughout the experiment. However, the lack of consistency between observations from month-to-month leaves me unsure of the nature of the relationship. Further work will need to be done in order to understand the mechanisms at play with regards to this relationship.

The LADWP low-resolution data is included in section 6.4 to offer some context about the lake's historical temperature and conductivity depth profiles. In both years it is clear that a boundary exists close (within 2-3 meters) to the 15-meter depth which leads to dramatic drops in both water temperature and conductivity. In 2020 this boundary is slightly deeper than in 2019; however, given the previously mentioned uncertainty in the measurements on the order of 2-3 meters, this difference is negligible. This result indicates that during the summer months Mono Lake's thermocline lies close to, if not at, 15 meters water depth.

Recall that the highest Granger causality in the barometric pressure column consistently occurred with the 15-meter temperature time series. This result seems to

suggest that barometric pressure is an extremely good forecaster of temperature swings along thermocline boundaries. It is also possible that barometric pressure could have a Granger causal relationship with internal waves happening along this thermocline boundary, however, it is difficult to know without more data.

The methodology and results presented in this chapter are scheduled to be presented at the AGU Fall 2021 meeting.

CHAPTER 7: CONCLUSIONS & FUTURE WORK

Historically, costs and logistics have been preventative in the collection of long-term heat flow and other geologic time series data. The successful deployment and performance of the HyLO-Mo and WaCoMS instruments in Mono Lake, CA presents a novel cost-effective, and lightweight method for collecting long-term geophysical data. Advancements in IoT and Cloud Computing technology facilitated the real-time transmission of these probes' measurements, completely mitigating the risks of loss of data due to probe failure events. Additionally, performing Granger causality-based time series clustering on the WaCoMS dataset demonstrated proof of concept that this method can be successfully applied to complex geological environments.

Future work on the micro-observatory systems will incorporate improvements in sample rates and data transmission techniques. Additional sensors will be added to collect seismic, acoustic, and CO_2 data. With improvements in probe design and transmission capabilities (transitioning from cellular to satellite), the micro-observatories will become more dynamic and capable of deploying into more demanding environments.

BIBLIOGRAPHY

[1] S. Hurwitz, C. D. Farrar, and C. F. Williams, "The thermal regime in the resurgent dome of Long Valley Caldera, California: Inferences from precision temperature logs in deep wells," *J. Volcanol. Geotherm. Res.*, vol. 198, no. 1, pp. 233–240, Dec. 2010, doi: 10.1016/j.jvolgeores.2010.08.023.

[2] J. Gallardo and D. D. Blackwell, "Thermal Structure of the Anadarko Basin1," *AAPG Bull.*, vol. 83, no. 2, pp. 333–361, Feb. 1999, doi: 10.1306/00AA9A84-1730-11D7-8645000102C1865D.

[3] M. J. Hornbach *et al.*, "A Hybrid Lister-Outrigger Probe for Rapid Marine Geothermal Gradient Measurement," *Earth Space Sci.*, vol. 8, no. 1, p. e2020EA001327, 2021, doi: 10.1029/2020EA001327.

[4] J. Sclater, D. Hasterok, B. Goutorbe, J. Hillier, and R. Negrete-Aranda, "Marine Heat Flow," in *Encyclopedia of Earth Sciences Series*, 2014. doi: 10.1007/978-94-007-6644-0_112-1.

[5] P. Morgan, D. D. Blackwell, R. E. Spafford, and R. B. Smith, "Heat flow measurements in Yellowstone Lake and the thermal structure of the Yellowstone Caldera," *J. Geophys. Res. 1896-1977*, vol. 82, no. 26, pp. 3719–3732, 1977, doi: 10.1029/JB082i026p03719.

[6] M. J. Hornbach, R. N. Harris, and B. J. Phrampus, "Heat Flow on the U.S. Beaufort Margin, Arctic Ocean: Implications for Ocean Warming, Methane Hydrate Stability, and Regional Tectonics," *Geochem. Geophys. Geosystems*, vol. 21, no. 5, p. e2020GC008933, 2020, doi: 10.1029/2020GC008933.

[7] A. H. Lachenbruch and B. V. Marshall, "Changing Climate: Geothermal Evidence from Permafrost in the Alaskan Arctic," *Science*, vol. 234, no. 4777, pp. 689–696, 1986.

[8] J. F. Batir, M. J. Hornbach, and D. D. Blackwell, "Ten years of measurements and modeling of soil temperature changes and their effects on permafrost in Northwestern Alaska," *Glob. Planet. Change*, vol. 148, pp. 55–71, Jan. 2017, doi: 10.1016/j.gloplacha.2016.11.009.

[9] J. Delaney and Chave, Alan, "NEPTUNE: A Fiber-Optic 'Telescope' to Inner Space," *https://www.whoi.edu/*. https://www.whoi.edu/oceanus/feature/neptune-a-fiber-optic-telescope-to-inner-space/ (accessed Oct. 09, 2021).

[10] E. Bullard, "The Flow of Heat through the Floor of the Atlantic Ocean," *Proc. R. Soc. Lond. Ser. Math. Phys. Sci.*, vol. 222, no. 1150, pp. 408–429, 1954.

[11] S. Nagihara and C. R. B. Lister, "Accuracy of Marine Heat-Flow Instrumentation: Numerical Studies on the Effects of Probe Construction and the Data Reduction Scheme," *Geophys. J. Int.*, vol. 112, no. 2, pp. 161–177, Feb. 1993, doi: 10.1111/j.1365-246X.1993.tb01447.x.

[12]   M. Pfender and H. Villinger, "Miniaturized data loggers for deep sea sediment temperature gradient measurements," *Mar. Geol.*, vol. 186, no. 3–4, pp. 557–570, Jul. 2002, doi: 10.1016/S0025-3227(02)00213-X.

[13]   C. R. B. Lister, "The pulse-probe method of conductivity measurement," *Geophys. J. Int.*, vol. 57, no. 2, pp. 451–461, May 1979, doi: 10.1111/j.1365-246X.1979.tb04788.x.

[14]   "Design, deployment, and status of borehole observatory systems used for single-hole and cross-hole experiments, IODP Expedition 327, eastern flank of Juan de Fuca Ridge1," *Integr. Ocean Drill. Program*, no. 327, Sep. 2011, doi: 10.2204/iodp.proc.327.2011.

[15]   A. T. Fisher, E. E. Davis, and K. Becker, "Borehole-to-borehole hydrologic response across 2.4 km in the upper oceanic crust: Implications for crustal-scale properties," *J. Geophys. Res. Solid Earth*, vol. 113, no. B7, 2008, doi: 10.1029/2007JB005447.

[16]   A. T. Fisher, T. Urabe, A. Klaus, and and the Expedition 301 Scientists, Eds., "Scientific and technical design and deployment of long-term subseafloor observatories for hydrogeologic and related experiments, IODP Expedition 301, eastern flank of Juan de Fuca Ridge," *Integr. Ocean Drill. Program*, vol. 301, Oct. 2005, doi: 10.2204/iodp.proc.301.2005.

[17]   C. Tan, A. P. G. Fowler, A. Tudor, and W. E. Seyfried, "Heat and mass transport in sublacustrine vents in Yellowstone Lake, Wyoming: In-situ chemical and temperature data documenting a dynamic hydrothermal system," *J. Volcanol. Geotherm. Res.*, vol. 405, p. 107043, Nov. 2020, doi: 10.1016/j.jvolgeores.2020.107043.

[18]   R. A. Sohn and R. N. Harris, "Spectral Analysis of Vertical Temperature Profile Time-Series Data in Yellowstone Lake Sediments," *Water Resour. Res.*, vol. 57, no. 4, p. e2020WR028430, 2021, doi: 10.1029/2020WR028430.

[19]   S. Aghabozorgi, A. Seyed Shirkhorshidi, and T. Ying Wah, "Time-series clustering – A decade review," *Inf. Syst.*, vol. 53, pp. 16–38, Oct. 2015, doi: 10.1016/j.is.2015.04.007.

[20]   C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969, doi: 10.2307/1912791.

[21]   D. Defays, "An efficient algorithm for a complete link method," *Comput. J.*, vol. 20, no. 4, pp. 364–366, Jan. 1977, doi: 10.1093/comjnl/20.4.364.