# ENABLING UNDERWATER BIOLOGICAL CONTACT SENSING SYSTEMS FOR SEA LAMPREY DETECTION THROUGH CARBON-BASED INTERDIGITATED ELECTRODES

By

Ian Gonzalez-Afanador

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Electrical and Computer Engineering—Doctor of Philosophy

2025

# ABSTRACT

Accurate monitoring of sea lamprey populations is critical to enabling the deployment of more targeted and effective control measures to minimize the impact associated with this species. This dissertation demonstrates the development of an automated sea lamprey detection system and the effects of algal-based biofouling on its voltage response.

The system is built around a sensor composed of two exposed carbon-based planar interdigitated electrodes (IDE) functioning as an underwater biological contact sensor. A microcontroller-based DC measurement system for the detection of lamprey attachment underwater is presented; measuring voltage instead of impedance reduces cost and signal processing complexity, making the device more attractive for field deployment. The system is calibrated to a baseline output voltage, and deviations from this baseline occur when objects touch the IDE. Validation was done through testing on live adult sea lampreys using video recordings to correlate lamprey attachments to the sensor response. Three response types were identified corresponding to different attachments: sustained, short, and sliding-sustained (video demonstrations are included in this work). The response to sustained and sliding-sustained attachments showed a characteristic exponential decay, whereas the response due to short attachments was indistinguishable from measurement noise. Sea lamprey size was found to have a weak linear correlation with both response parameters, positive for the voltage drop and negative for the time constant of the voltage drop. A representative circuit for the lamprey-sensor interaction is proposed and simulated using element values calculated from the response parameters. The response of the model shows agreement with experimental data.

Characterization of the IDE sensor's response to sea lamprey attachment allowed for the development of detection algorithms, to automate the process of detecting sea lamprey from the sensor output. Inherent limitations on the computing power of the microcontroller unit used to measure the sensor motivated the exploration of low-complexity models for the task: single-layer artificial neural networks, logistic regression, Gaussian Naïve-Bayes, decision trees, random forest, and Scalable, Efficient, and Fast classifieR (SEFR). Threshold models tuned using a multi-objective optimization formulation were also considered. Models were trained/tuned with a data set generated through live animal testing and presented accuracies between 80-86%. The models were deployed on an Arduino microcontroller platform and compared in classification accuracy, detection performance, time complexity, and memory size using real-time detection testing. Classification accuracies between 65-75% were observed during validation. Models demonstrated capture rates of 63-85% for sea lamprey attachments and average detection delays of 9-36 seconds. A video demonstration of a real-time validation test is also presented. Sensor robustness, sensitivity, and detection speed were identified as areas of improvement for the system, so electrodes were updated to more durable, conductive carbon-based 3D-printed electrodes, and an updated measurement system combining a resistive bridge topology and an instrumentation amplifier was developed to allow for scaling the sensor response.

The effects of biofilm accumulation on the baseline sensor response were also studied to simulate field conditions. Sensors were exposed to optimal biofilm growth conditions in a photobioreactor running a culture of *Chlorella sorokiniana MSU*, a robust algae species native to the Great Lakes region. Two responses to biofilm formation on the sensors were observed: (1) a consistent rise in voltage following the increasing form of the exponential decay function, and (2) a similar rise in baseline but followed by a period of approximately exponential decay. These responses were generally related to changes in sensor resistance, with (1) corresponding to a decrease and (2) to an increase. A study relating these changes in sensor resistance to biomass accumulation was performed, but did not produce a conclusive relationship. The sensors' ability to pick up biological contact after fouling was observed to be inconsistently affected. The implications of these results for sea lamprey detection and potential ways to address them are discussed. Overall, this research presents the first step toward an electronic sea lamprey monitoring system that can provide a detailed view of sea lamprey activity, enhancing control and conservation efforts across its entire range.

To science.

# ACKNOWLEDGMENTS

I would like to begin these acknowledgements by recognizing the person who put me on the path to becoming a researcher, Dr. Gerson Beauchamp. Dr. Beauchamp was my research advisor during my undergraduate career, and through his rigorous approach to mentoring helped me develop my confidence as a scholar and engineer. It was also through him that I was able to meet another instrumental figure in my development, my current advisor, Dr. Nelson Sepúlveda. It is thanks to Dr. Sepúlveda's continuing support and mentorship that I have been able to make it this far in my academic career. He has provided me with a nurturing environment in which I've been able to grow not only as a scientist but also as a collaborator and peer mentor. I would also like to thank my friends and Applied Materials Group alumni Dr. Henry DŚouza, Dr. José Luis Figueroa-Soto, and Dr. Juan Pastrana who welcomed me into the group with open arms and made me feel at home at MSU. To the current AMG group members Gerardo Morales-Torres, Yusuf Bin Yaacob, and Luis Colon Santiago I extend my thanks as well for the good times spent together, the productive conversations regarding our research projects, and for trusting me to help guide you through the beginning of your graduate school journeys like the group alumni did for me.

I would also like to thank the members of my committee for making this document possible: Dr. Xiaobo Tan, the first faculty member to join my graduate committee, whose insight has been proven invaluable when facing roadblocks in my research, Dr. Wei Liao, who inspired me to adapt my sensors for biofilm detection in the first place and who has supported me in those endeavors, and Dr. Saiprasad Ravishankar, who opened my eyes to the possible applications of the signal processing field to my research.

There are also people from my personal life who have supported me through this odyssey whom I would like to acknowledge, beginning with my parents, Angel and Lourdes, who have always pushed me to be the very best I can be while providing a comforting environment if and when I did not live up to expectations. My siblings, Angeliz, Chaterly and Yansel, have also been instrumental in my development, serving as confidants and sources of unexpected inspiration. My long-time friends Alejandra and Carlos Nahyr, have been a constant source of emotional support, encouragement, and laughs throughout my time in graduate school. My close friend Alexandra, who always pushes me to take care of myself and take time to enjoy life in the moment. Lastly, I would be remiss if I did not mention my cat, Isabel, who brings much needed joy and tenderness into my life.

To all those mentioned here, who have helped me become who I am today, I extend my most sincerest thanks.

# TABLE OF CONTENTS

# CHAPTER 1

# INTRODUCTION

This work focuses on the development of automatic sensing systems for detecting and quantifying interactions between aquatic organisms and underwater surfaces. The effort is targeted towards detecting invasive sea lamprey (Petromyzon marinus), parasitic fish that present surface-attaching behaviors, in the Great Lakes region. In addition to its application in invasive wildlife monitoring, a system capable of monitoring biological interactions with underwater surfaces has potential applications that also motivate this study.

## 1.1 Invasive Sea Lamprey Detection

Sea lampreys (*Petromyzon marinus*) are an invasive fish species in the Laurentian Great Lakes region. They live as parasites for a large part of their life cycle, feeding on other fish by attaching to them using their suction cup-like mouth (as seen in Fig.1.1-A) and scraping a hole through the skin to suck on the prey fish's blood and other bodily fluids [3] leaving wounds like those shown in Fig.1.1-B. While it is debated whether sea lampreys are native to Lake Ontario, they have been reported in the lake since at least 1888 [4]. It is believed they spread from Lake Ontario to Lake Erie, and later to the rest of the Great Lakes, through the

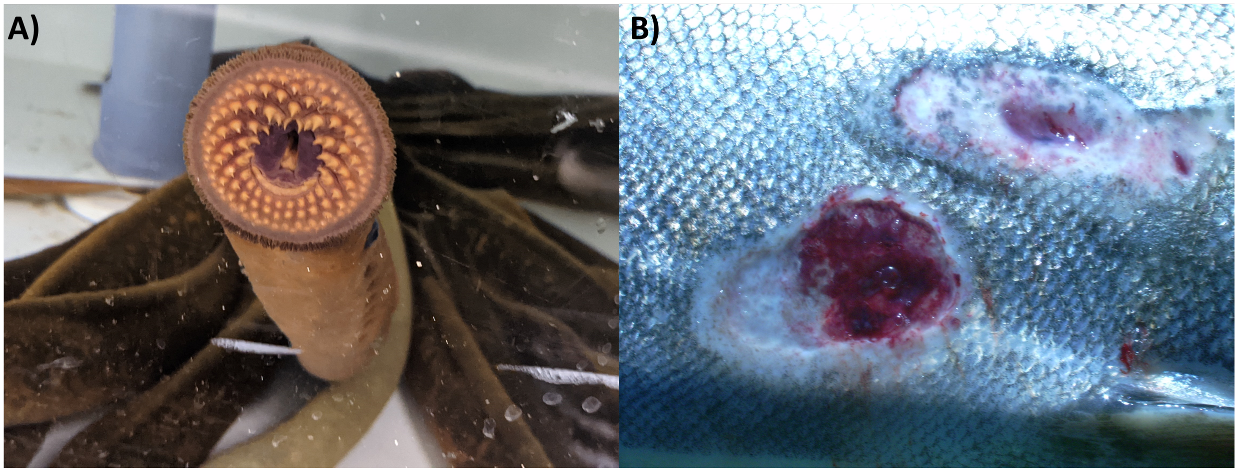


Figure 1.1 A) Sea Lamprey attaching to view pane of a holding tank at Hammond Bay Biological Station B) Sea Lamprey wounds on a Chinook salmon by Wikimedia Commons user Fungus Guy[1], reproduced here without alterations under the Creative Commons Attribution-Share Alike 3.0 Unported license.[2].

Welland Canal which began operation in 1829. The presence of sea lampreys in Lake Erie was first confirmed in 1921 [5], and by 1938 they had spread to the remaining Great Lakes [6]. They quickly exploded in number and contributed to a decline in large fish stocks causing a devastating decrease in commercial fishing yields, which resulted in the effective collapse of this industry in many parts of the region [7]. Lake trout (*Salvelinus namaycush*), lake whitefish (*Coregonus clupeaformis*), and burbot (*Lota lota*), which were lampreys' preferred prey, were depleted or extirpated from much of the Great Lakes [8]. As these were the main predators in the region, their disappearance opened the door for other invasive species such as rainbow smelt (*Osmerus mordax*) and alewife (*Alosa pseudoharengus*) which created intense competition for species not directly threatened by sea lampreys [7], further destabilizing the ecological balance of the region. Through aggressive control measures, including deploying physical [9] and electrical barriers [10, 11] to limit access to spawning areas, lamprey-specific pesticide (lampricide) application [12], and trapping and sterilizing male sea lamprey [13], sea lamprey numbers have been effectively reduced. Accurately monitoring sea lamprey populations is critical to adequately address the catastrophic threat they pose to the region and its fishing industry which employs around 75,000 people and is valued at USD 7 billion annually [14].

Currently, the sea lamprey control program assesses its effectiveness via an adult lamprey abundance index generated by the capture of adult sea lamprey from tributaries around the Great Lakes. Adult lamprey traps in these tributaries are staffed for daily operations from April through June (sometimes through July) [15]. Because of operation cost and poor efficacy in many streams trapping is conducted in only twenty-nine tributaries across all the Great Lakes. Other tools have been considered to improve or expand this assessment effort but have limitations. For example, statistical models have been trained to forecast the probability of meeting control targets using abundance index information and historic measures of sea lamprey control effort [16]. While this provides a useful metric for planning control strategies, it is still limited by the available data. Environmental DNA (eDNA) approaches, where water samples from the field are analyzed to determine the concentration/presence of genetic material from the target species, have been demonstrated to be able to detect the adult and larval sea lamprey in lab and field settings, though it is not possible to distinguish between the two [17]. Further, while there is evidence that the concentration of genetic material is positively correlated with adult sea lamprey population densities in the lab [18], in practice samples can provide presence or absence data for an entire river but do not provide location, density, or even timing. Dual-frequency identification sonar (DIDSON) has also been used to successfully track lamprey entry into a tributary of Lake Huron [19], though the high cost of the equipment (upwards of 25,000 USD per camera) and specific requirements for deployment (e.g., water depth, non-turbulent environment) make it suited for limited deployment only. A low-cost, effective tool to assess the presence or relative abundance of adult sea lamprey in streams could benefit the lamprey control program assessment efforts and provide a means to better understand the timing of lamprey migration and distribution of adult sea lamprey throughout the Great Lakes tributaries.

Remote lamprey sensing systems present an alternative surveillance technique; these systems could factor in environmental disturbances in their measurements, enhancing operational capabilities by allowing for continuous population monitoring. Moreover, by exploiting the behavior of the species during the design of these systems, it results in a less disruptive assessment technique than tagging as well as reduces dependence on traps. In addition to workload reduction for conservation agencies like the SLCP, trap-free detection could provide greater insights into the behavior of these fish at times and locations where capture or direct observation has been historically limited. However, the lack of commercially available sensors that can detect lampreys and the necessary scale of deployment for these systems to be effective are major hurdles in the development of such systems. A review of the state of the art in electronic lamprey detection is included as part of **Chapter 2**.

The automated sea lamprey detection system presented in this work is meant to be deployed in rivers and tributaries of the Great Lakes region. Any time a system or structure is submerged in (or used to contain) water, its surfaces can interact with the myriad of microorganisms that call that water home. This can lead to the undesired accumulation of biological matter on these surfaces, a process known as biofouling [20], which can cause structural and operational deficiencies. An example of the biofouling formation process as it occurs in a marine environment is presented in Fig. 1.2. Biofouling like this will likely be an issue in the deployment of an automated sea lamprey detection system based on this work, motivating the study of the response of these sensors to the accumulation of biofilm on their surfaces. Extending the capabilities of the measurement system to the sensing of accumulation of biological material can expand the possible applications of these sensors as biofouling is a pervasive issue in many fields including food production [**?**], medicine [21], industrial processes [22] and algal cultivation systems [23]. Current approaches to biofilm monitoring rely on staining techniques for visual confirmation of total biomass film or differentiating between live and dead cells as well as various assays analyzing cultivation of biofilm samples [24]. These techniques provide good characterization of biofilms at discrete time instances but do not allow for online monitoring of the build-up. The operating principle behind the sea lamprey contact detection sensors could be an alternative to traditional methods, providing real-time information on the accumulation of biological material on underwater surfaces that can be used as a feedback signal for prevention and control measures.

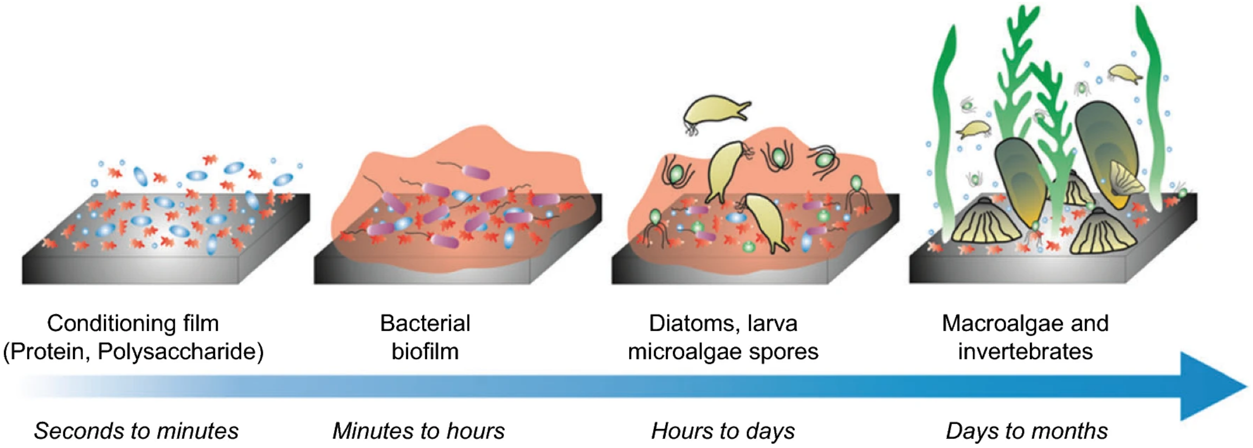


Figure 1.2 Typical growth process of marine biofouling. Reproduced with permission from [25].

## 1.2 Problem description and motivation

Differentiating between contact sources poses a challenge for typical contact sensing mechanisms such as mechanical switches and pressure sensors. Comprehensive characterization of the contact source using purely mechanical or pneumatic sensors is simply not possible, as the physical principles behind these sensors will result in identical responses regardless of source. Approaches that do offer some selectivity by monitoring the electrical properties of the sensor, such as capacitive touch sensing, are typically vulnerable to false readings or noise in the presence of water. The problem of developing a selective underwater contact sensing system is addressed in this work through:

- The design of low-cost carbon-based interdigitated electrodes for contact sensing using relative conductivity differences as a driver for contact source characterization.
- Development of a measurement system taking advantage of these principles and the characterization of the system's response to sea lamprey attachment.
- Exploring the correlation between characterized response parameters and sea lamprey size parameters.
- The refining of the measurement system through signal conditioning and data processing techniques for improved performance.
- The integration of the system with machine learning algorithms for automating the detection of sea lamprey using the sensing system.
- Determining the effect of biological accumulation on the baseline response of the sensor.

To this end, an iterative design process including fabrication, simulations, experiments, and analysis was carried out to present solutions to these problems. The following chapters provide fabrication, coding, and data processing guidelines that can be used to build similar systems as well as a methodology to incorporate machine learning approaches in similar applications through embedded systems.

## 1.3 Thesis Statement

The major contribution of this work is the development of an underwater contact sensing system that can distinguish between underwater biological and non-biological contact sources and automatically detect sea lamprey attachment on these surfaces, as well as study on the impact of biofilm accumulation on the response of the system

## 1.4 Outline

The structure of the document is as follows: **Chapter 2** provides a review of the state of the art in electronic sea lamprey detection as well as the rationale behind the selection of the IDE structure for the task. This is followed by a discussion of the design methodology for the sensors including the operating principles behind the IDE sensor, material selection for the device and the differences with traditional IDE applications. The chapter ends with an overview of the fabrication process for the first two iterations of interdigitated electrode sensors presented in this work **Chapter 3** presents the characterization of the response of the first-generation sensor design to sea lamprey attachment. Modeling of the response is performed and is used to undertake a correlation study investigating the relationship between sea lamprey size parameters and the response parameters. The result of the correlation study are discussed and a representative circuit for the sea lamprey sensor interaction is proposed and simulated. **Chapter 4** covers the use of learning models deployed on embedded systems to automate the detection of the sea lamprey based on the characterized response and contains a discussion on the model selection, how they were trained and deployed as well as the evaluation of their performance during a real-time sea lamprey detection test. **Chapter 5** briefly presents the third-generation sensing system design, including an updated measurement approach to allow for amplification of the sensor response. It also includes the design and fabrication of the final iteration of the IDE sensor based on conductive 3D-printed electrodes. **Chapter 6** covers the attempted characterization of the response of the sensing system to algal biofilm accumulation to better understand how the system will perform in the field. It also includes a discussion on the implications of these results for the deployment of the sensors in the field and some suggestions for future research directions. The document ends with a summary of the contributions and main takeaways of this work.

[1]

---

[1] Some content in chapters 2-4 is adapted from modified versions of *Invasive Sea Lamprey Detection and Characterization Using Interdigitated Electrode (IDE) Contact Sensor* - IEEE SENSORS (2021)[26] © 2011 IEEE; *Real-Time Invasive Sea Lamprey Detection Using Machine Learning Classifier Models on Embedded Systems* - Neural Computing and Applications (2024)[27] © 2024 Springer Nature and is reproduced here with the permission of the copyright holders. In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Michigan State Universities's products or services. Internal or personal use of this material is permitted.

# CHAPTER 2

# BACKGROUND, DESIGN METHODOLOGY AND DEVICE FABRICATION

This chapter provides a review of the current state of the art in electronic sea lamprey detection, as well as a review of interdigitated electrode sensors and their applications. It also features the design methodology used to develop the contact sensors, including the working principle, material considerations, and design parameters. Finally, the sensor fabrication process is presented for both generations of sensors.

## 2.1 Background

### 2.1.1 Electronic Sea Lamprey Detection Sensors

Recent advances in electronic sensors for sea lamprey detection [28, 29, 30, 26] show potential for use in wide-scale electronic monitoring systems that could bypass some of the highlighted challenges in the expansion of assessment capabilities. These sensors look to exploit sea lamprey behaviors to enable the use of transducers to detect their presence. As sea lampreys commonly attach to structures using their suction cup-like mouths, pressure and contact

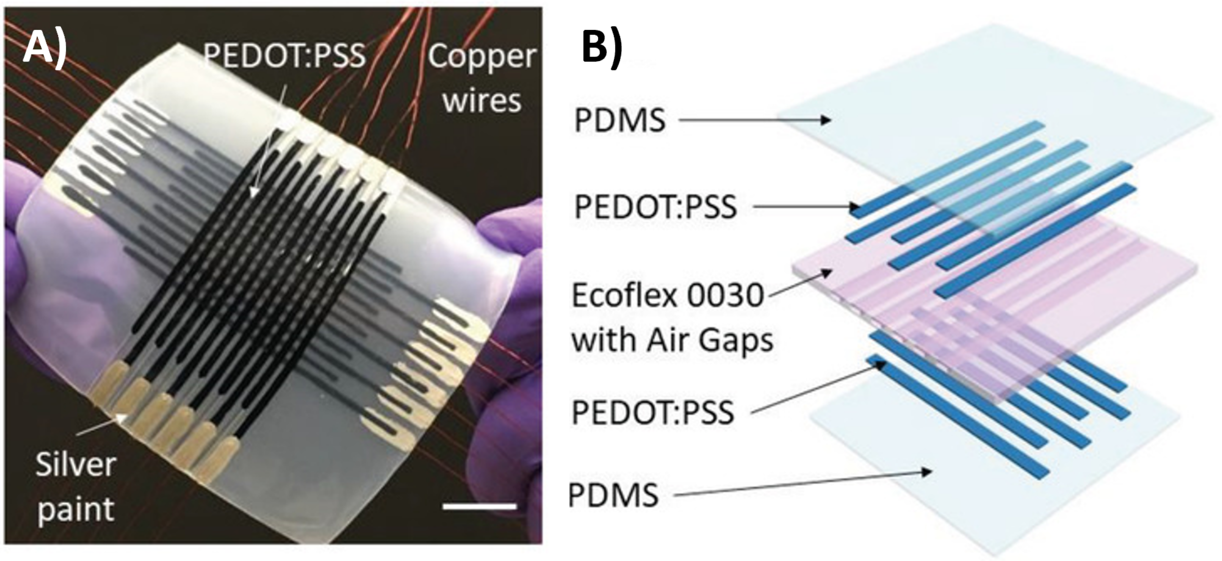


Figure 2.1 . A) Schematic illustration of the soft capacitive pressure sensor array with air gap channels and conductive polymer PEDOT:PSS electrodes. B)Photograph of a 12 × 12 capacitive pressure sensor array. Scale bar: 1 cm. reproduced with permission from [28] © 2019 John Wiley & Sons.

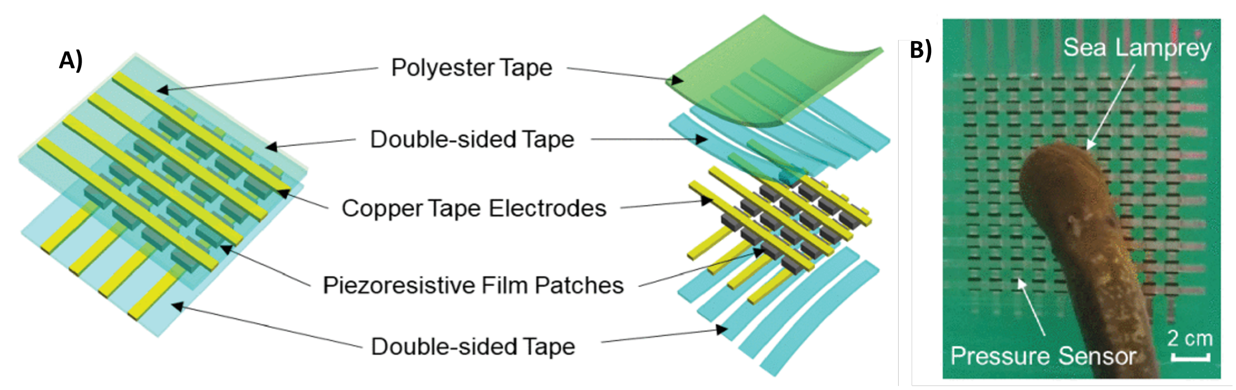


Figure 2.2 A) Schematic of a 4-by-4 piezoresistive pressure sensing matrix. B) Sea lamprey interacting with the sensing array. Reproduced with permission from [29] © 2022 IEEE.

sensing are natural avenues to explore for their detection.

Flexible pressure sensing arrays are a promising candidate for sea lamprey attachment detection thanks to their ability to generate profiles of the pressures exerted at their surface [31], from which the suction action of the lampreys can be identified. Several transduction mechanisms can be used to create these sensors, including capacitance, piezoresistivity, and piezoelectricity. Shi et al. [28] reported a soft capacitive pressure sensing array, shown in Fig.2.1 A, capable of detecting positive and negative pressures which could potentially be used for sea lamprey detection. The device was manufactured almost completely from flexible polymers (as seen in Fig.2.1B) and laboratory testing of these sensors using a polymer suction cup produced promising results; however, the quickly changing pressure profile exerted by the lamprey [32] proved difficult to capture with the implemented measurement system [33]. Additionally, capacitive interference due to the dielectric nature of biological material may have also contributed to difficulties with these measurements. These issues are less prevalent in piezoresistive sensors, which utilize materials that react to changes in pressure with a corresponding change in resistance as the sensing mechanism instead of relying on changes in capacitance driven by deformation under pressure. Indeed flexible encapsulated piezoresistive array structures, whose construction is shown in Fig.2.2A, reported by Shi et al. [29] have been shown to be successful at capturing sea lamprey attachments on their surfaces (as seen in Fig.2.2B). These sensors are not without their drawbacks, as they suffer from crosstalk and hysteresis issues that complicate the processing of their output. Additionally, the manufacture of these devices can be a time-consuming process, tough advances have been made in streamlining their fabrication [34]. Finally, flexible encapsulated polypropylene ferroelectret (PPFE) devices, which present a quasi-piezoelectric effect due to charged voids in the film as seen in Fig 2.3, have also been shown to be capable of measuring positive and negative pressures underwater[30] and could be deployed in array structures to generate similar pressure profiles. The output of these types of sensors could be analyzed using a computer vision algorithm to detect and characterize the profile generated by sea lamprey attachment [35]. While promising, the use of computer vision as the underlying detection algorithm does limit the system due to the computational resources these algorithms require.

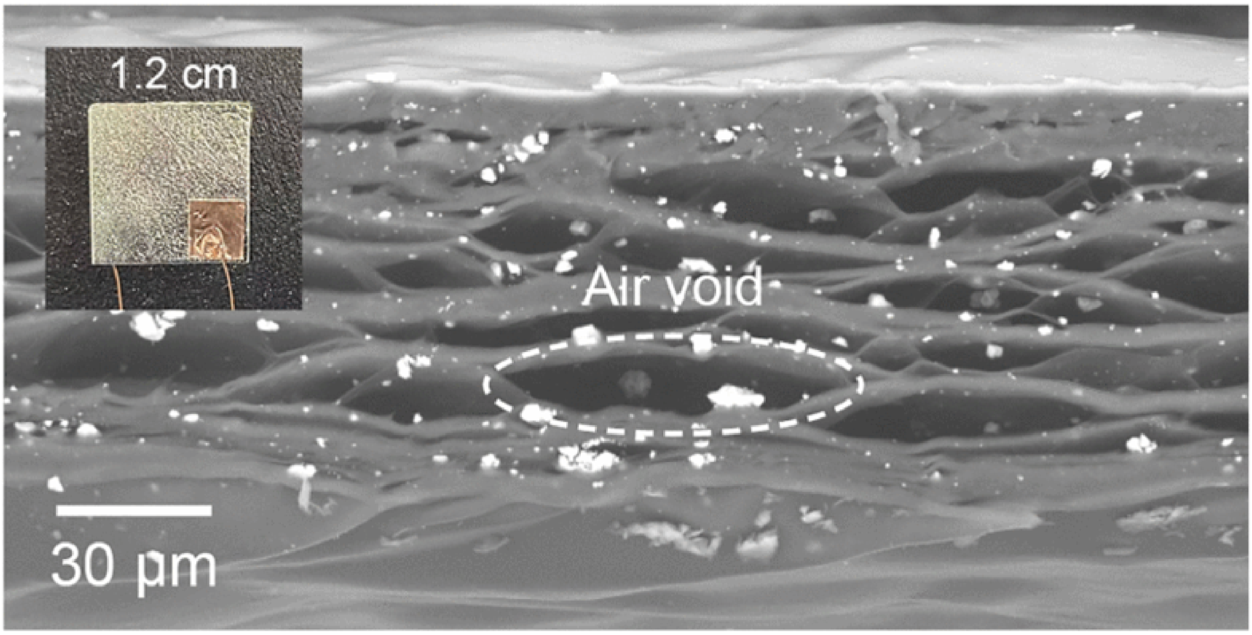


Figure 2.3 SEM image of the cross-sectional view of a PP-based ferroelectret device. The inset shows a PPFE sensor. Reproduced with permission from [30] © 2022 IEEE.

### 2.1.2 Interdigitated Electrodes as Electrochemical and Biosensors

The development of the sensing system presented in this work took inspiration from the interdigitated electrode (IDE) electrode structure, which is widely employed in electrochemical sensors. The working principle behind these devices uses the change in impedance seen across the electrodes, caused by the surface accumulation of the compounds, as the sensing signal [36]. These sensors have been widely utilized as chemical detectors [37] and biosensors [38, 39] using a technique known as electrochemical impedance spectroscopy (EIS) analysis to perform label-free detection and concentration measurements. As the surface accumulation that drives the sensing signal could be described as the aggregate of multitudes of incredibly small discrete "contact" events [40], it is possible this sensing principle could be adapted on a larger scale for the task of underwater biological contact sensing.

## 2.2 Design Methodology

### 2.2.1 Operational principle of IDE Contact Sensing System

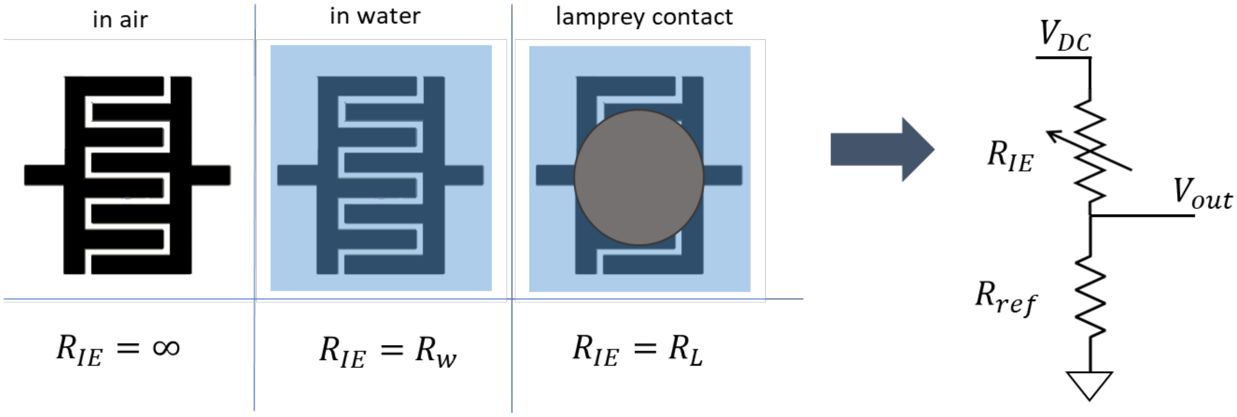


Figure 2.4 Visual representation of IDE contact sensor operating principle.

One of the main contributions of this work is presenting a novel use of the IDE structure as an underwater contact sensor for invasive sea lamprey detection. Figure 2.4 provides a visual showing the operating principle used to enable contact sensing underwater with the IDE devices. The sensors are composed of two separate planar electrodes with interlocking fingers in close proximity but electrically isolated when in air. They are encapsulated in a way that only the sensing electrode area is exposed to water. To measure the sensors, a direct current (DC) probing voltage is applied to the sensor and a reference resistor connected in series, forming a voltage divider circuit. The voltage across the electrodes is monitored and any abrupt changes to electrode impedance caused by attachment/detachment of the sea lamprey will be reflected in this output. This approach was selected over more sophisticated impedance measurement schemes, which require additional components and signal processing because the precision requirements for simple contact detection are lower than those for electrochemical characterization and detection. Furthermore, the additional components would increase the cost and complexity of the device making it less suitable for the large-scale deployments necessary for the system to be effective at monitoring sea lamprey populations.

### 2.2.2 Differences from traditional IDE applications

While electrochemical IDE sensors use an alternating current (AC) probing signal to directly determine the impedance/capacitance (or its frequency response) during EIS [41], the approach taken here is akin to an indirect impedance measurement; since the probing signal is DC, the steady state value of the response is assumed to be due to the effective resistance between the electrodes with any capacitive behavior being captured in the transient response of the sensor. When the sensor is submerged in water, an electrical connection is formed between both electrodes resulting in a measurable baseline, which can be disturbed during interaction with a foreign object, e.g., lamprey contact. This mechanism introduces a degree of selectivity in the contact sensing, as the system will react differently to contact sources with different conductivities, relative to the conductivity of the surrounding medium —in this case, lake water. This behavior opens the door to characterization of the contact source which is important as the American brook lamprey (Lampetra appendix) [42], northern brook lamprey (Ichthyomyzon fossor)[43], chestnut lamprey (Ichthyomyzon castaneus) [44] and the silver lamprey (Ichthyomyzon unicuspis) [45] are all native to the region and share behavioral patterns with sea lamprey that may lead to their interaction with these sensors. Electrochemical IDE sensors enforce selectivity by functionalizing the electrode surface to have an affinity for bonding with the target molecule [46]. In the case of the device presented here, selectivity is partially enforced simply by the detection mechanism, as non-lamprey species are not likely to interact with the sensor for a sustained time. During the adult life stage targeted by this work, the major distinction between native and invasive species of lamprey is their size, with adult sea lamprey typically being twice as large (35 – 60 cm) as the native species (10.3 – 32.6 cm) [47]. The need to determine if there is a correlation between the size of the lamprey and the response parameters motivates the correlation study in **chapter3**.

### 2.2.3 Sensor Materials and Design Process

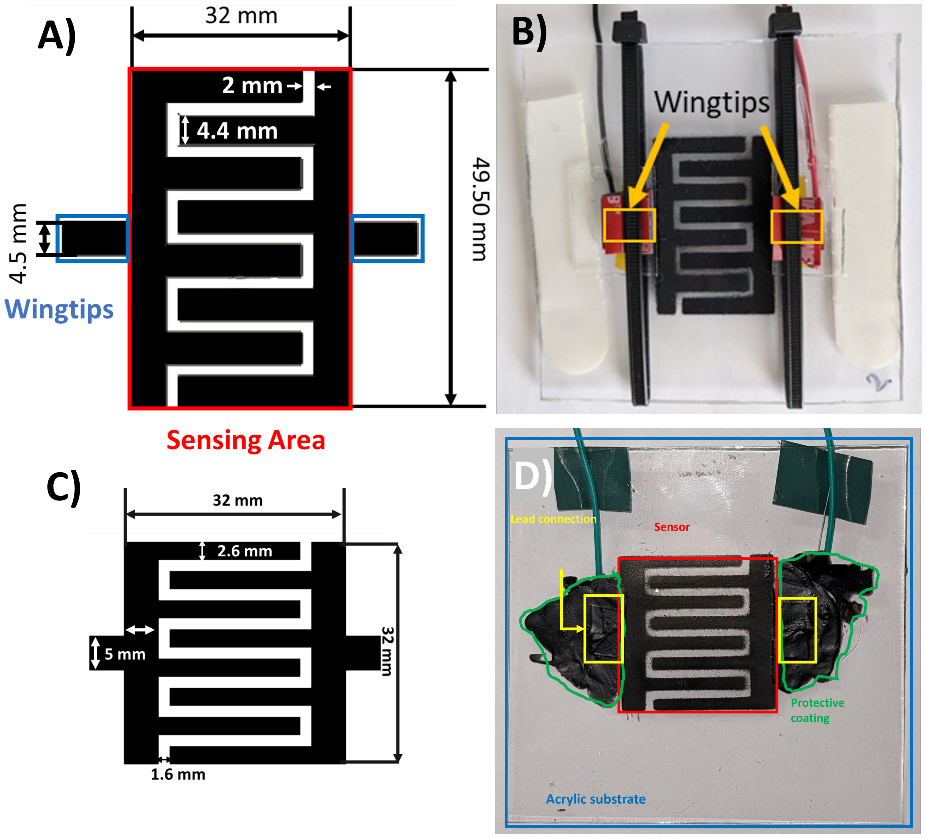


Figure 2.5 A) Schematic of first-generation IDE sensor geometry and dimensions. B) Example of a manufactured device with a first-generation design. C) Schematic of second-generation IDE sensor geometry and dimensions. B) Example of a manufactured device with a second-generation design. Reproduced with permission from [26] © 2021 IEEE.

Clear acrylic panels were used as the substrate for the IDE sensor to allow visual monitoring of lamprey attachment events and correlate these to the electrical measurements of the system. The specified operating conditions of the sensor preclude utilizing any conductive materials that are prone to corrosion, as this would not only hamper its functionality but also lead to unintended contamination of the freshwater systems where it is to be deployed. Marine-grade conductive carbon spray paint (838AR - Total Ground Carbon Conductive

Coating by MG Chemicals) was selected as the electrode material for our prototypes due to its low cost, relatively high conductivity, and ease of application. The electrodes were patterned onto the acrylic substrates using 0.6 mm thick 3D-printed stencils developed in AutoDesk's Solidworks 3D modeling suite. This fabrication process allowed for an iterative approach at the device design stage, where we tested different geometries, overall size, electrode finger length and width, and distance between the electrical lead attachments and noted the effect on the magnitude of sensor response to contact from a conductive probe and a human hand. Multiple prototype generations were tested until reaching the final design which showed the best response magnitude under test conditions.

A schematic of the first iteration of the IDE device, which was used in the characterization study presented in **3**, is shown in Fig.2.5A. The design consists of a small rectangular interdigitated electrode (sensing area of 32 mm x 49.50 mm) with wingtips to allow for attachment of electrical leads away from the sensing area. During preliminary experiments, it was observed that the sensitivity of the sensor increased with decreasing distance between the electrical leads (see Fig. 1-E); and the sensor design/configuration used (Fig. 1-A) allowed for the smallest possible distance that will still allow enough sensing area to include the lamprey's oral disk. Wingtips were added to the design to facilitate the protection of the sensor-lead connection without having to sacrifice the sensing area. It was found that a gap of 2 mm between electrodes would facilitate stencil fabrication and any post-processing of the deposited traces, while also providing a detectable change in resistance upon the lamprey's attachment. Trace width was selected to optimize the fill factor of the device under the area and separation constraints. Fig.2.5B shows an example of a final manufactured device used from the first generation. After the characterization study, the design was updated into a smaller design, shown in Fig.2.5C, to further optimize the response magnitude of the devices. Fig.2.5D shows a manufactured device example from the second generation. This iteration was utilized when the system was integrated with learning models for automatic detection as presented in **4**.

## 2.3 Sensor Fabrication Process

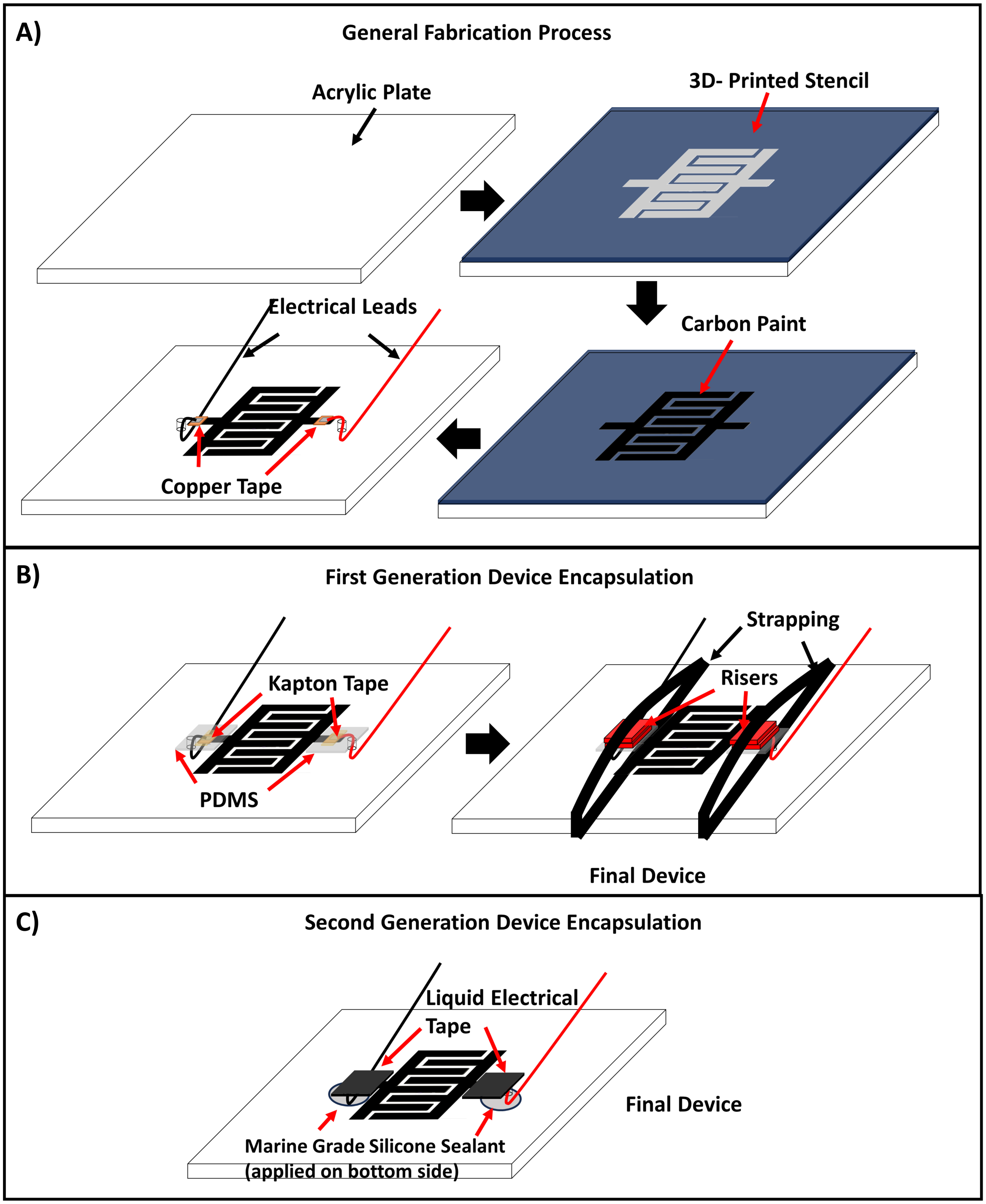


Figure 2.6 A) Device manufacturing process flow B) Encapsulation and lead protection for first-generation sensors. C) Encapsulation and lead protection for second generation sensors. A and B reproduced with permission from [26] © 2021 IEEE.

Fig.2.6-A shows the device fabrication process beginning with a square acrylic panel sub-

strate (approximately 102 mm x 102 mm), on which the 3D-printed stencil was attached using repositionable craft adhesive. A coat of conductive carbon paint was then deposited, allowed to cure for at least one hour, and followed by a second coat which was cured for at least 12 hours. By following the manufacturer's instructions each coat should be approximately 25 $\mu$m thick, giving the device a total thickness of approximately 50 $\mu$m. After the paint was completely cured, the stencils were separated from the acrylic, and any excess paint between the traces was removed using a cotton swab applicator with small amounts of isopropyl alcohol. Holes were then drilled just outside of the wingtips, which served as feedthrough vias for the electrical leads. A small piece of copper tape was soldered onto the leads and used to create a conductive contact between the carbon trace and the leads. Fig. 2.6B shows how this connection was secured for the first generation of sensors, first by fixing it to the acrylic plate using Kapton tape, followed by encapsulation using polydimethylsiloxane (PDMS) for waterproofing, finishing with strapping and use of high-strength tape for a robust attachment to the acrylic plate. To fix the position of the leads on the backside of the plate, small amounts of hot glue were utilized. This encapsulation process was improved for the second generation, as seen in Fig. 2.6 C, where liquid electrical tape was used to protect the lead connection, and marine grade silicone sealant was used to fix the leads on the bottom instead. Finally, two velcro tabs were attached to the back of the acrylic panel, so that it could be attached to a larger panel housing the measurement system for testing.

After the devices were manufactured, the sensor response to open and short-circuit tests was confirmed in air using a handheld multimeter. For the open-circuit test, an infinite resistance across the two separate IDEs was measured. For the short circuit test, both electrodes were connected using a piece of copper tape; and the test confirmed a very low resistance ( 100 Ohms) between both copper tape leads. After validating the device's operation ex-situ, their integration in a system for underwater deployment and testing followed which will be covered in the next chapter.

# CHAPTER 3

# IDE SENSOR SEA LAMPREY CONTACT RESPONSE CHARACTERIZATION AND MODELING

This chapter covers the characterization of the response to sea lamprey attachment for the first generation of the underwater IDE contact sensor. This includes a description of the measurement system and an overview of the testing methodologies used to validate the device through experiments with live sea lampreys. Data from the experiments were analyzed to characterize the disturbance caused by lamprey attachment on the sensor and observe whether it was distinguishable from measurement noise. The experimental results are presented along with a correlation study between the identified response parameters and lamprey size parameters to enhance the utility of the sensor by exploring the characterization of lamprey size from the sensor response. The results of a FEM simulation of the sea lamprey-sensor interaction are used to confirm the experimental sensor behavior. Finally, a simple representative circuit modeling the lamprey-sensor interaction is proposed and simulated using component values extracted from the sensor response data.

## 3.1 Sensor Measurement Approach

The schematic of the system used to measure the sensors is presented in Fig.3.1-A. A DC voltage divider circuit was used to convert the change in impedance across the electrodes to a change in voltage which can be more easily measured. An Arduino microcontroller platform (highlighted in yellow in Fig.3.1-A and B) was utilized to provide the 5 V DC bias ($V_{DC}$) to the voltage divider circuit and measure the output voltage at a 10 Hz sampling frequency using its integrated analog-to-digital converter (ADC) module. This configuration also allows for the calibration of the measurement system in the presence of water through the reference resistor ($R_{ref}$). To this end, the device is submerged in water and the reference resistor is adjusted until the output voltage of the circuit is approximately half of the DC bias, around 2.50 V. This results in an approximate value of the baseline resistance between the electrodes ($R_{IDE}$) due to water and enables a relationship between the output voltage and

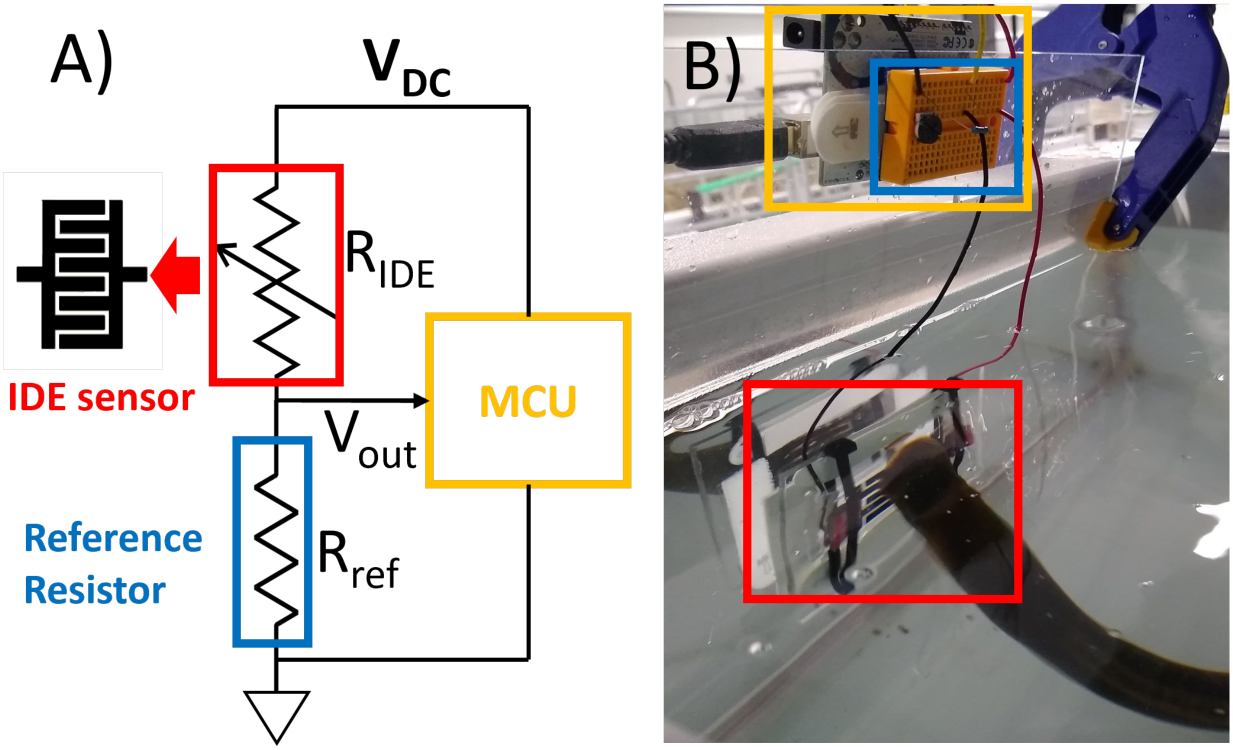


Figure 3.1 A) Measurement system schematic. B) Implemented measurement system attached to an acrylic panel. Reproduced with permission from [26] © 2021 IEEE.

the change in resistance. Using a voltage divider analysis for the circuit shown in Fig.3.1-A, the output voltage ($V_{out}$) can be calculated as:

$$V_{out} = V_{DC} * \frac{R_{ref}}{R_{ref} + R_{IDE}} \tag{3.1}$$

This equation only holds at all time instances if the circuit is purely resistive, and since water and animal flesh are dielectric materials, it is reasonable to expect some capacitive behavior from the IDE sensor. However, since the probing signal is DC, the capacitance will come into play only in the transient behavior of the response, meaning that (3.1) will still be a valid description of the steady state of the signal measured by the MCU unit. It should be noted that $R_{ref}$ has a fixed resistance value, whereas $R_{IDE}$ has a variable resistance that depends on any contact between the two isolated electrodes. In the experiments performed, the background $R_{IDE}$ value (i.e., resistance before lamprey's contact) represents the water resistance between the electrodes ($R_{water}$). By selecting $R_{ref} = R_{water}$ and expressing $R_{IDE} = R_{water} + R_{\Delta}$ (where $R_{\Delta}$ represents the magnitude of the change in resistance across the electrodes and its sign indicates whether $R_{IDE}$ is increasing or decreasing) we can re-write (3.1) as:

$$V_{out} = V_{DC} * \frac{1}{2 \pm \frac{R_{\Delta}}{R_{water}}} \tag{3.2}$$

It should be noted that the addition or subtraction of the $\frac{R_{\Delta}}{R_{water}}$ term occurs when $R_{IDE}$ increases or decreases, respectively. Thus, an increase in the resistance across the electrodes ($+R_{\Delta}$) will correspond to a drop in the measured output voltage ($V_{out}$ ); and, correspondingly, a decrease in $R_{IDE}$ ($-R_{\Delta}$) will increase $V_{out}$. This allows for the distinction between two general types of contact: 1) when $R_{\Delta}$ is negative (i.e., $R_{IDE} < R_{water}$, and 2) when $R_{\Delta}$ is positive (i.e.,$R_{IDE} > R_{water}$). Further characterization within those two categories is possible by examining the magnitude of the change in $R_{\Delta}$. Fig.3.1-B shows how the system was implemented for testing. The microcontroller platform and a small circuit proto-board housing the reference resistors were attached to the top of a larger acrylic panel measuring 304.8 mm x 304.8 mm using velcro tabs for easy replacement/modification if necessary. The sensor was then attached at the bottom of the acrylic panel, to allow for submerging the system without exposing the rest of the system to moisture.

### 3.2 Experimental Procedure

Experiments were performed at U.S. Geological Survey, Great Lakes Science Center's Hammond Bay Biological Station (HBBS) in Millersburg, MI. HBBS staff provided 20 male sea lamprey specimens, each tagged with a unique identifier for tracking purposes between experiments. The length, weight and mouth diameter of the lamprey were measured and recorded as they were tagged (see Table 3.1). As explained later in Section 3.4, the physical characteristics of the lampreys shown in this table were used to determine the strength of correlation between the sensor response and each lamprey size parameter. Such correlation is of particular interest because it would enhance the monitoring capabilities of the developed system from simple detection to characterization of an attached lamprey's size, which may

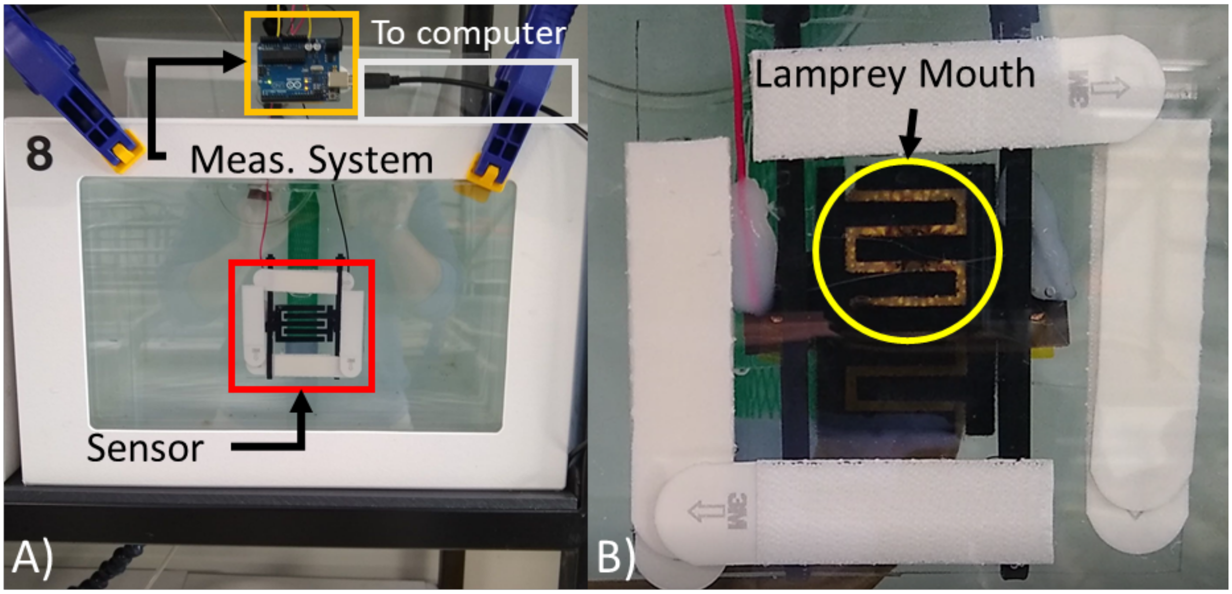


Figure 3.2 A) Sea lamprey holding tanks with sensing panel installed. B) Lamprey attaching onto sensor during testing; view of the lamprey is obstructed by sensor traces as the image was taken from the outside of the testing tank. Reproduced with permission from [26] © 2021 IEEE.

allow distinction between species or life stages.

Experiments were conducted in a 200 L rectangular aquarium tank (Fig.3.2-A) supplied with aerated water from Lake Huron. The sensing panel was submerged in the tank and fastened using clamps before tests started. The sensor was calibrated to an output voltage of approximately 2.5 V, and a baseline measurement was recorded for several minutes before testing to ensure this baseline did not drift significantly due to water turbulence. The test procedure consisted of placing a lamprey into the tank and coaxing it (manually) into attaching to the sensor. A closeup of how the sea lamprey's mouth looks while attaching to the sensor during testing can be seen in Fig.3.2-B; as the image was taken from the opposite side of the sensor, through the tank wall, the lamprey is partially obstructed by the electrodes. Video and voltage measurements were recorded until the lamprey detached from the sensor. An audio cue was triggered in the video to synchronize the recording and the data offline. On the first day of testing, 20 experiments were completed on the sensor, one with each lamprey of the cohort, with additional baseline measurements taken after halfway through testing.

Table 3.1 Sea Lamprey Size Parameters

| Tag | Weight (g) | Length (cm) | Mouth Diameter (mm) |
|---|---|---|---|
| blue-018 | 255 | 44 | 44.8 |
| blue-009 | 246 | 50 | 40.6 |
| orange-002 | 176 | 42 | 40.4 |
| orange-017 | 241 | 48 | 39.6 |
| white-041 | 200 | 45 | 39.6 |
| orange-025 | 161 | 39 | 38.0 |
| white-028 | 117 | 38 | 37.9 |
| blue-002 | 158 | 39 | 37.9 |
| orange-004 | 177 | 40 | 37.8 |
| orange-040 | 177 | 42 | 37.2 |
| pink-034 | 144 | 39 | 36.9 |
| pink-027 | 174 | 45 | 36.5 |
| white-043 | 118 | 39 | 36.0 |
| pink-035 | 143 | 40 | 34.7 |
| orange-026 | 132 | 40 | 34.3 |
| white-037 | 141 | 40 | 34.2 |
| pink-029 | 124 | 38 | 33.9 |
| blue-024 | 143 | 40 | 33.3 |
| pink-028 | 128 | 40 | 33.0 |
| orange-035 | 87 | 35 | 29.3 |

The baseline measurements were taken again after all tests were completed for validation purposes. On the second day, 10 additional tests were completed with the lamprey cohort for validation purposes. Voltage data were then processed in OriginLab using a low-pass filter with a cut-off frequency of 0.25 Hz to filter out high-frequency measurement noise associated with ADC readings such as input-referred noise and voltage regulation noise. Finally, a time offset was applied to the data to synchronize voltage measurements and video recording. Voltage data and video recordings from these tests were then cross-referenced to determine which disturbances in the baseline were caused by lamprey attachment events (see links to videos of the test in Section 11 of the appendix)

### 3.3 Characterization of Sensor Response to Sea Lamprey Attachment

Three distinct types of responses were identified from the sensor, each corresponding to a different type of attachment. The following discussion is focused on three tests that

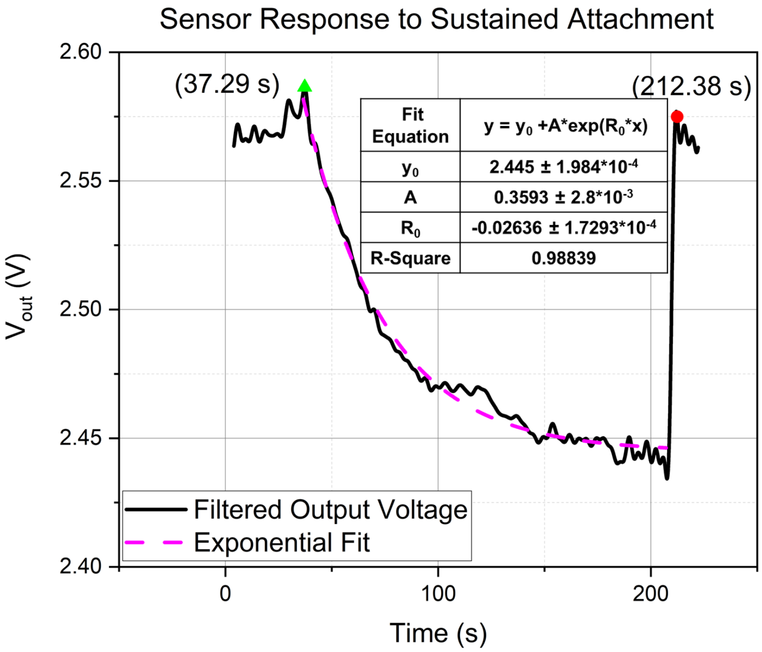


Figure 3.3 Sensor response to a sustained sea lamprey attachment event. Reproduced with permission from [26] © 2021 IEEE.

are representative of these responses, but the remaining experimental results are included in Section 11 of the appendix. Fig.3.3 shows a representative measurement of the sensor response to a sustained lamprey attachment. It can be noticed in the video for this test (link 11 in Section 11 of the appendix), that the attaches to the sensor at t = 37.29 s and detaches at t = 212.38 s, denoted on the plot through colored markers with green indicating attachment to the panel and red indicating release. The output voltage of the sensor ($V_{out}$) shows a clear exponential decay to a DC value below the original baseline when the lamprey is attached, corresponding to an increase in the resistance observed between the IDEs ($R_{IDE}$), with a sudden change back to baseline when it detaches. The observed response suggests that the system is not a purely resistive circuit. Instead, it includes reactive components that result in a time-dependent response characteristic of a first-order transient circuit, which is to be expected as water and living tissue are known dielectric materials. This behavior presents an opportunity to model this type of interaction using discrete circuit components, which will be discussed in Section 3.4.1.

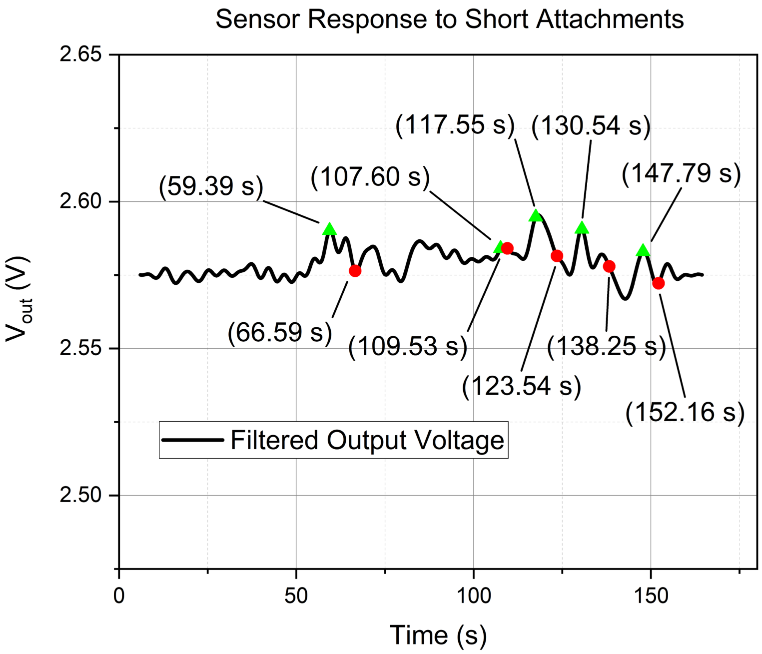


Figure 3.4 Sensor response to short sea lamprey attachment events. Reproduced with permission from [26] © 2021 IEEE.

The next plot shown in Fig.3.4 is representative of the sensor response to brief attachments. During this test, the lamprey was visually observed to be attached to the sensor multiple times for less than 10 seconds (link 11 in Section 11 of the appendix). Each of these attachment periods is highlighted on the plot using colored data markers. While there appear to be transients in the data at these instances, the magnitude of the response of the sensor is not enough to be clearly distinguished from measurement noise. This observed behavior suggests that a minimum attachment time is needed for the sensor response to be detectable, which supports the time dependency of the exponential decay response represented in Fig.3.3.

Fig.3.5 shows the sensor response when the lamprey slid onto the sensor traces after attaching partially or completely outside of the sensing area. In the video recording of this test (link 11 in Section 11 of the appendix), two quick attachments directly onto the sensing area are observed, before a third attachment occurs where the lamprey only partially covers the sensor. At this point, the handler slid the lamprey across the device to bring it to a central

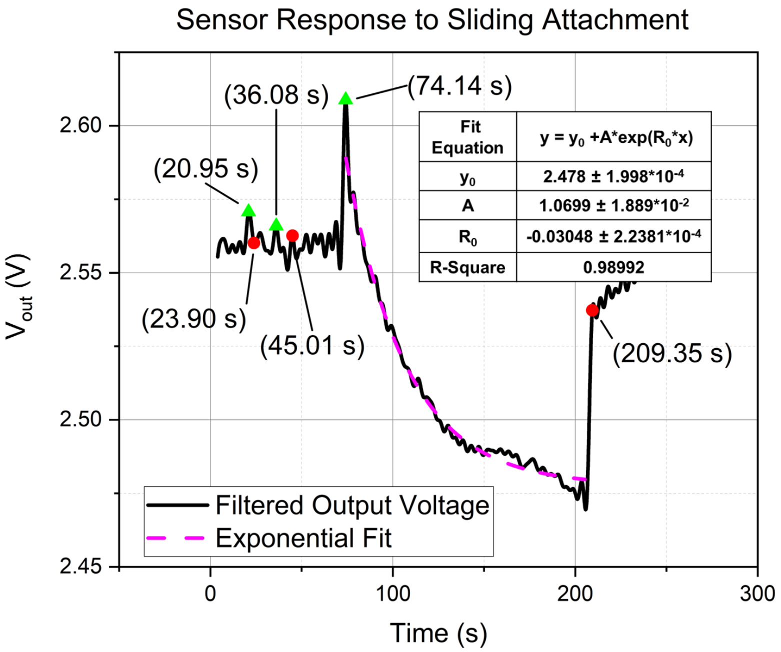


Figure 3.5 Sensor response to a sliding attachment event. Reproduced with permission from [26]© 2021 IEEE.

position on the panel. As seen in Fig. 3.4 the quick attachments did not generate a response distinguishable from measurement noise. However, while the characteristic exponential decay of a sustained attachment seen in Fig. 3.3 is still observed, it is now preceded by a large transient spike caused by the lamprey sliding on the sensor. This transient indicates that this attachment method somehow reduces the resistance seen across the IDE sensor for a very brief period before it begins to increase – resembling the observation for a direct, sustained lamprey attachment. The mechanism behind this phenomenon is still not fully understood; though it may be that the movement of the lamprey is contributing to a quick discharging of the double-layer capacitance that naturally forms between a conductor and ionic liquid in contact when subjected to an electric field, briefly creating a current path with resistance smaller than the ionic solution's bulk resistance.

### 3.3.1 Baseline Validation

Since the responses of the sensor to the three attachment types that were identified during testing have been presented, it is also necessary to examine the baseline measurements of

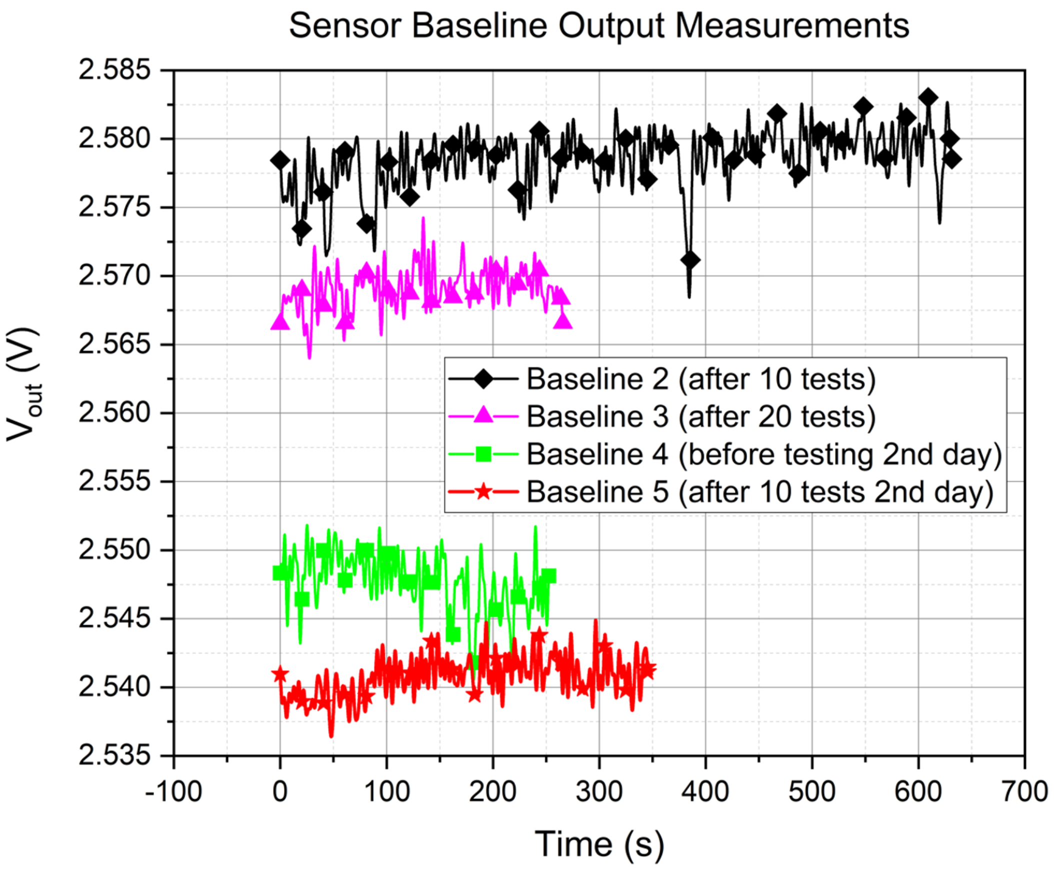


Figure 3.6 Compilation of baseline measurements taken throughout testing. Reproduced with permission from [26] © 2021 IEEE.

our device to determine its reliability and susceptibility to false positives. The plot shown in Fig. 3.6 contains the results of each baseline measurement test taken throughout our experiments. While there exists measurement noise in the baseline, there is a clear and stable average $V_{out}$ value for each of the tests. This value does vary slightly from test to test, indicating some drift, but these could be attributed to variances in water temperature as the water in the tanks is taken directly from Lake Huron, which experiences some slight temperature variations throughout the day. Regardless of the cause, these variations not only appear to be much smaller than the sensor response to a lamprey attachment event, but they also occur over a timescale of hours meaning they could be easily distinguished from the sensor's characteristic exponential decay response. Additionally, this stability of

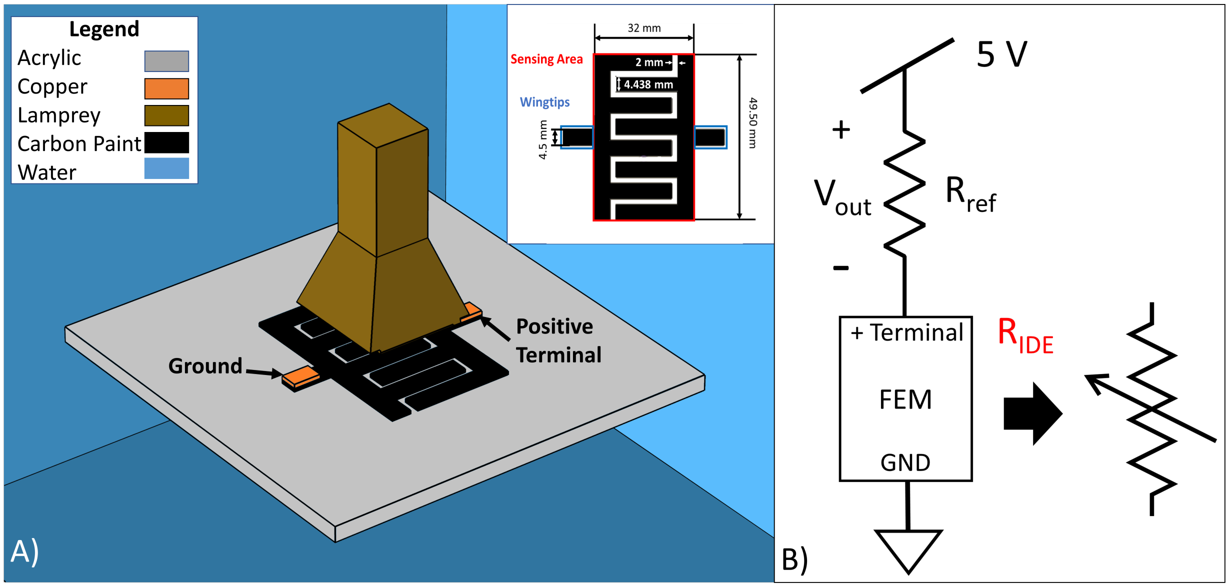


Figure 3.7 A) Simplified geometry used in FEM static and transient simulations B) Electric circuit schematic for COMSOL transient simulation. Reproduced with permission from [26] © 2021 IEEE.

the measured baseline was maintained even after a total of 30 experiments across two days, thus providing strong support to the robustness of the device.

### 3.3.2 FEM Sea Lamprey-Sensor Interaction Simulation

The exponential decay suggests the lamprey-sensor interaction can be represented by an RC circuit lumped model. A finite element model (FEM) simulation of the lamprey-sensor interaction was also performed using COMSOL Multiphysics 5.3a. The 3D geometry used in the FEM simulation is shown in Fig. 3.7-A.

In order to facilitate meshing and obtain a solvable model, the height of the traces was scaled by a factor of 20 (from 50 $\mu$m to 1 mm) and a pyramid shape followed by a rectangular prism body was used to represent the lamprey in the simulation. This geometric simplification was necessary because the considerable disparity in sizes between the smallest and largest features in the model geometry and the very small regions generated when conforming a conic lamprey mouth to the traces resulted in an extreme amount of mesh nodes, making the problem very computationally expensive to run. The rest of the geometry is to scale with the sensor schematic shown as an inset in Fig. 3.7-A and everything is surrounded by a domain representing water (the back boundaries of this domain are light blue in the figure). It is important to mention that the length of the lamprey object was selected as such because this parameter did not impact the resistance results significantly when increased beyond that value. The electric currents module was used to determine the resistance between the electrodes of the simulated device under static conditions. A parameter sweep of the distance between the lamprey and the IDE traces was conducted to see how the resistance changes as the lamprey gets closer to the sensor. The FEM was coupled to external circuit elements using the electric circuit module, and Fig. 3.7-B shows the connection diagram. The electrical parameters used in the simulation, $\sigma$ the electrical conductivity and $\epsilon_r$ the relative permittivity, are shown in Table 3.2.

Table 3.2 Electrical properties of materials used in COMSOL simulation

| Object | $\epsilon_r$ | $\sigma(S/m)$ |
|---|---|---|
| Lamprey | 97.75 | $2.5*10^{-3}$ |
| Carbon Paint | 4.5 | 14.492 |
| Water | 80 | $2.17*10^{-2}$ |
| Acrylic | 3.73 | $4* 10^{-14}$ |
| Copper | 1 | $5.998* 10^{7}$ |

The objective of the static FEM simulation was to examine the conditions under which it would be possible to recreate the steady state behavior of the sensor, meaning the observed increase in resistance across the IDE when the lamprey came into contact with the sensor. The value of conductivity utilized for water is the measured value for the lake water utilized in the experiments. Since there are currently no values reported for the conductivity of the lamprey, this parameter was modulated until the resistance of the device in the contact condition was greater than when the lamprey was away from the sensor. For the carbon paint, the value of the conductivity as reported by the manufacturer was initially utilized but this resulted in resistance values that were much smaller than what was experimentally measured through the voltage divider circuit during calibration (43.17 $k\Omega$). Upon measuring the conductivity of the traces of the device it was found that the actual conductivity was around

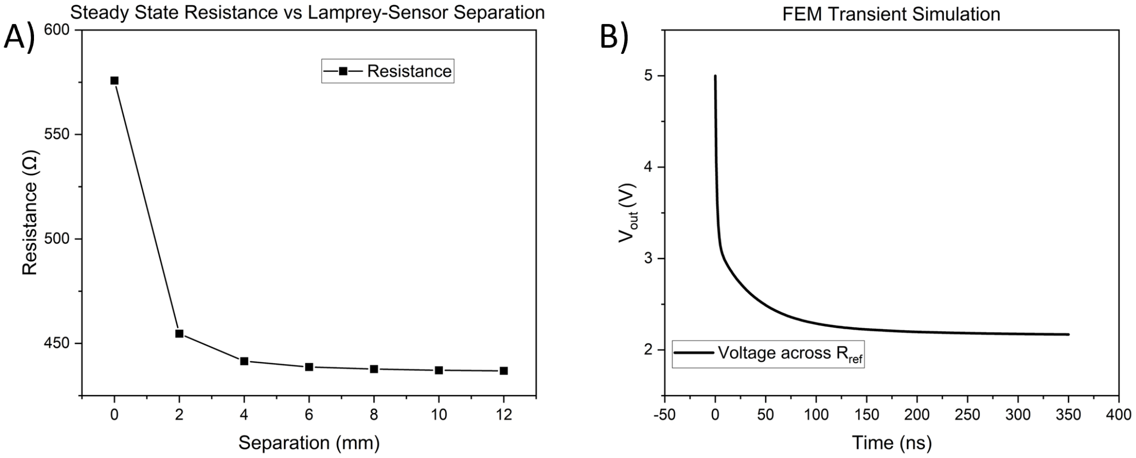


Figure 3.8 A) Simulated steady-state resistance of the IDE sensor as a function of separation between lamprey and sensor B) Simulated transient response of the sensor voltage output. Reproduced with permission from [26] © 2021 IEEE.

10% of the manufacturer specification, so it was adjusted accordingly for the simulation. The values of resistance obtained from the separation parameter sweep are presented in Fig 3.8.

There is a definite increase in the resistance seen across the terminals of the device as the lamprey gets closer under the simulated conditions. However, this behavior was observed only when the conductivity of the lamprey object was less than the conductivity of the water. Even after utilizing the measured conductivity for the carbon paint, the simulated resistance was still smaller than the measured value. This discrepancy could be attributed to the differences in geometry between the simulation and the actual device. Once the conditions for the static behavior were established, the transient behavior of the simulated device was examined. It is important to note that the simulated circuit represents a slightly different situation than what happened during the experiments. The simulation represents the case where the lamprey is attached to the sensor before the device is powered and stays attached after the power is switched on. This means the voltage across the reference resistor will not reach a baseline corresponding to the resistance of the water before decaying further because of the lamprey attachment. Instead, the decay that will be observed will be from the supply voltage to the steady state voltage level associated with the lamprey attachment. The results of the transient simulation are presented in Fig. 3.8-B.

An exponential decay of the output voltage, taken across the reference resistor, which then settles out to a baseline can be observed for the transient simulation of the system. This behavior is reminiscent of a parallel RC circuit, just as what was observed experimentally, and strengthens the case for modeling the lamprey sensor interaction as such a circuit. However, there is a significant difference in the time constant observed for the simulation, which is in the order of tens of nanoseconds, and the time constants observed during the experiments, which were in the order of tens of seconds. One factor that contributes to this discrepancy is the smaller resistance values obtained for the simulation, which will affect the time constant for an RC circuit, but this difference is not enough to account for the discrepancy. One important difference between the simplified model and the actual lamprey-sensor interaction is that the lamprey object in the simulation does not have the longitudinal cavity which comprises the lamprey's digestive tract and breathing apparatus. It may be that this cavity causes the lamprey to behave almost like a cylindrical capacitor when attached to the sensor, which can have considerably higher capacitance values than other capacitor geometries. The capacitance in cylindrical capacitors is also affected by the length of the coaxial cylinders which is neglected in this model for ease of simulation. However, as the objective of this simulation is to corroborate the general steady-state and transient behavior observed experimentally rather than perfectly recreate the interaction, these results were deemed sufficient.

### 3.4 Response Modeling and Correlation Study

As explained in the previous section, whenever there was a sustained lamprey attachment on the developed sensor (whether by direct attachment or sliding) an exponential decay was observed at the output voltage of the sensor ($V_out$). This type of response can be characterized by the time constant of the decay along with the steady-state value of the response. Considering the sea lamprey attachment period, the response voltage can be modeled as:

$$V_{out} = V_{baseline} - V_{\Delta}(1 - \exp\frac{-t}{\tau}) \tag{3.3}$$

where $\tau$ is the time constant of the decay and $V_{baseline} - V_{\Delta}$ is the value of $V_{out}$ at steady state. These parameters can be extracted from an exponential fit to the measurements where a sustained lamprey attachment was observed and then used to explore their correlation to sea lamprey characteristics, such as length, weight, and mouth diameter. The exponential fit tool from OriginLab was used, which uses the following exponential relationship:

$$y = y_0 + A \exp R_0 t \tag{3.4}$$

Seventy percent of the first day's tests presented the characteristic exponential decay corresponding to the sustained attachment and sliding attachment cases. With a few exceptions, detailed in Section 11 of the appendix, these tests were used for the correlation study. The other remaining tests showed behavior that resembled that of intermittent contact – i.e., not sustained mouth attachment – to the sensor (see Fig. 3.4). The baseline voltage ($V_{baseline}$) was obtained on a test-by-test basis, by taking the average of the response before the exponential decay occurs. The change in voltage ($V_{\Delta}$) was calculated by subtracting the steady-state value ($y_0$) from the exponential fit from the baseline value. Fig. 3.9-A shows the result of the correlation study between $V_{\Delta}$ and the lamprey size parameters. In order to reduce spread on the x-axis, the size parameters are normalized with respect to the median value of each parameter of the entire 20 lamprey cohort. value of each parameter of the entire 20 lamprey cohort. The values of $V_{\Delta}$ are plotted along with the corresponding error bars obtained from the fit. The plots shown in Fig. 4-A show a weak positive linear correlation between all size parameters and the $V_{\Delta}$ of the measured response, indicating that larger lamprey had a larger response magnitude than smaller lampreys. The next parameter that was studied is the time constant, which was obtained from the fit equation as $\tau = \frac{-1}{R_0}$ seconds. The results of the correlation study between the lamprey size parameters and $\tau$ are

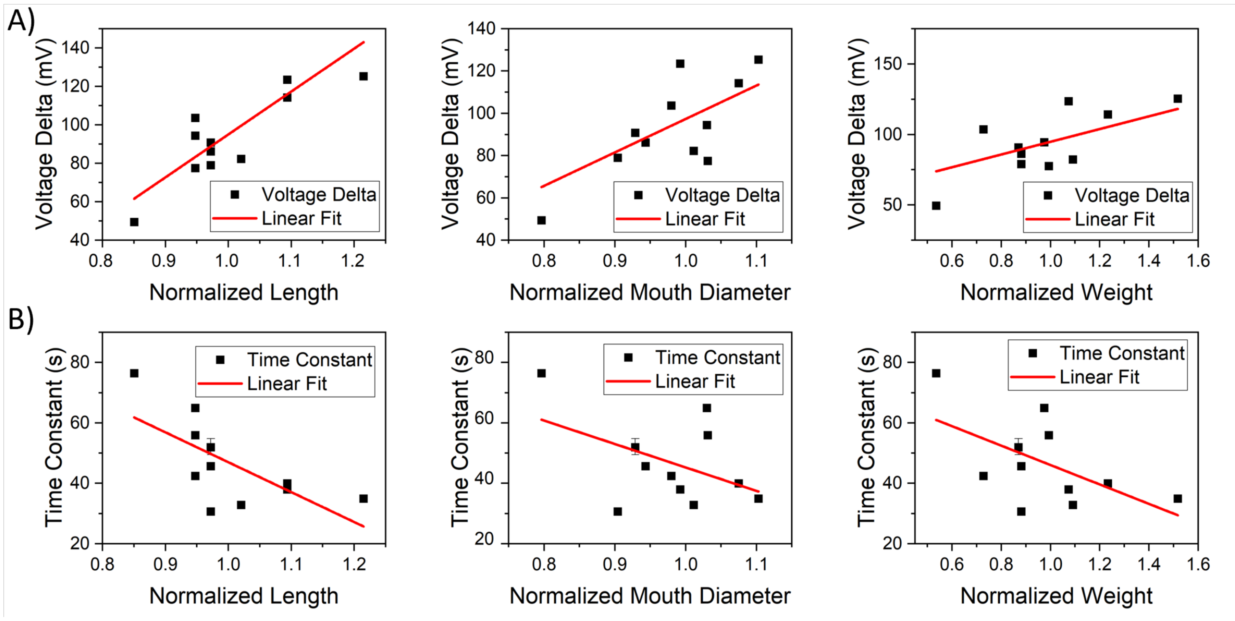


Figure 3.9 A)) Correlation study between sea lamprey size parameters and voltage delta response parameter. B) Correlation study between sea lamprey size parameters and time constant response parameter. Reproduced with permission from [26] © 2021 IEEE.

shown in Fig. 3.9-B. The plots show a weak negative correlation between the time constant and the lamprey size parameters, indicating that larger lampreys within our sample produce sensor responses with smaller time constants, implying smaller capacitance, which corresponds to a faster decay to the attachment baseline. Although these are promising results toward a contact sensor that can distinguish between invasive and native lamprey species, it is important to recognize that the relatively small sample size (20 male sea lamprey) used in this study is not sufficient to establish a generalized relationship between the lamprey size parameters and the response parameters. Additionally, we are not considering if there exists any cross-correlation between the lamprey size parameters themselves. Because of the pattern observed between oral disk size and sensor response parameters, partial attachment cases where the lamprey is not completely covering the sensor could also be problematic; in these cases, the sensor could respond as if a smaller lamprey was attached allowing for possible misclassification. However, this effect may be mitigated by using an array of these sensors, rather than a single unit.

### 3.4.1 Lumped Circuit Model

The simplest circuit structure that would result in the observed electrical behavior is shown in Fig. 3.10-A. In this circuit, the attachment of the lamprey is represented as a normally-closed switch (NC Switch in schematic) that opens upon lamprey attachment, forcing the series connection of a simple parallel RC circuit to the conduction path. When the switch is triggered (i.e., switch is opened), the capacitor initially behaves as a short circuit, causing the current to bypass the parallel resistor ($R_{Lamp}$); but after that initial transient behavior, the capacitor begins to accumulate charge and behaves more like an open circuit, thus sending more current to the parallel resistor. This has the effect of gradually introducing the resistor into the current path, which results in a waveform like the one observed during testing. The effective values of the electric lumped model shown in Fig. 3.10-A can be extracted from the exponential fit shown in Fig. 3.3. Since a sudden change in $V_{out}$ is not observed at the instant the lamprey attaches, it is reasonable to assume that the baseline resistance somehow stays in the conduction path. Therefore, the steady-state voltage is the result of the following voltage divider:

$$V_{out} = y_0 = \frac{V_{DC} * R_{ref}}{R_{ref} + R_{water} + R_{Lamp}} \tag{3.5}$$

Note that the value of the resistance due to water is different from $R_{ref}$, as the baseline output voltage before the lamprey attaches was not exactly 2.50 V. However, $R_{water}$ can be readily obtained by solving for $R_{IDE}$ in (3.1) using the measured values of $V_out$ before any lamprey activity is initiated. Once the value of $R_{Lamp}$ is obtained, it can be used to determine the corresponding capacitance value, since $\tau = R_{Lamp}C_{Lamp}$ .

Simulation of the proposed circuit using the values obtained from this analysis was performed using LTspice XVII and the response of the circuit is shown in Fig. 3.10-B, along with the experimental measurement. The simulation shows the exponential decay behavior when the switch opens (corresponding to the lamprey attachment) and a sudden return to baseline when the switch is closed (corresponding to the lamprey releasing from the panel).

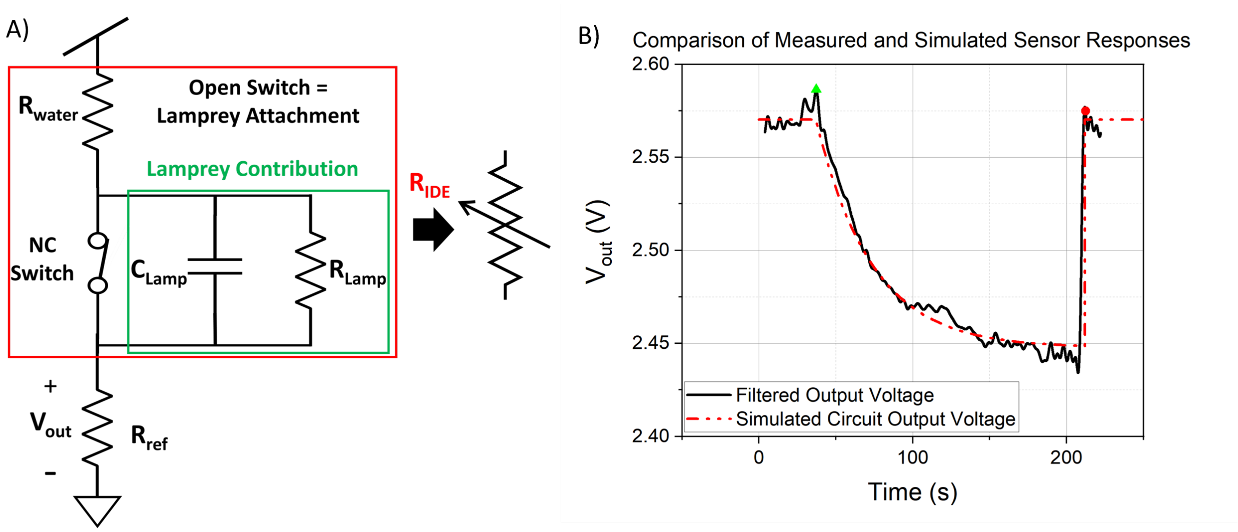


Figure 3.10 A) Schematic of the representative electrical circuit. B) Comparison of simulated circuit output with measured response. Reproduced with permission from [26] © 2021 IEEE.

The simulation, however, does not capture the behavior of the sensor response to a sliding attachment, which could be viewed as a ramp input to the system and would require a more complex circuit lumped model.

## 3.5 Summary

A simple and effective IDE sensor for lamprey attachment detection was demonstrated. Three different categories of sensor response were identified, corresponding to different attachment types: sustained, short, and sliding-sustained. Whereas the response due to short attachments was not distinguishable from measurement noise, the other two response categories showed a characteristic exponential decay shape. The measured voltage at steady state and the time constant of the response were identified as parameters that may be used to obtain information about the lamprey's physical characteristics, which could allow classification of the attached lamprey, although confirmation would require further testing with larger sample sizes. The correlation study performed between the response parameters and the lamprey size parameters showed a weak linear correlation between them. The voltage change appeared to have a positive correlation with all the lamprey size parameters studied (length, weight, and mouth diameter), i.e., the bigger lamprey generated a sensor response with a larger change in voltage, whereas the time constant had a negative correlation with every lamprey size parameter. Expansion of the system into an array of these devices may help mitigate possible issues that arise in the case of partial attachments, by expanding the effective sensing area without sacrificing device sensitivity. A representative circuit to model the lamprey-sensor interaction was proposed and simulated using element values extracted from the parameters of the measured experimental response. The simulated circuit matched the behavior observed in the experimental data for the case of sustained lamprey attachment. The case of sliding-sustained attachment would require a more complex circuit lumped model that can account for a ramp-type response. Further improvements to the system hardware will be necessary, including additional ruggedization of the sensor, expansion into sensor arrays, waterproofing for supporting electronics, and the addition of data storage/transmission modules. Software enhancement of the system through the use of machine learning to improve the identification of invasive sea lamprey is also possible. The device presented in this work is the first step toward a remote sensing system to continuously monitor lamprey populations in the tributaries of the Great Lakes without the need to physically capture individual animals.

# CHAPTER 4

# REAL-TIME AUTOMATIC SEA LAMPREY DETECTION SYSTEM USING MACHINE LEARNING CLASSIFIER MODELS ON EMBEDDED SYSTEMS

This chapter includes an overview of the second-generation sensing system including a signal conditioning circuit and additional hardware modules bringing it closer to a stand-alone system. The methodology used to build and process the data sets for training machine learning classifier models to detect the sensor response to sea lamprey attachment is presented. The validation of these models through live-animal validation testing is shown. Classification results on the testing fraction of the training set are presented to show the expected performances and are compared to the classification results for the validation tests. The performance of each algorithm at the detection task is quantified and presented along with the execution time and memory size for each algorithm. Additional details on the methods used to deploy ANN models to a microcontroller device are included in the appendix.

## 4.1 Second Generation Sensing System Design

The Arduino Mega platform was chosen as the control unit for the system because of its memory size (256 KB of flash memory) and number of I/O ports. SD card and real-time clock (RTC) modules were included in the system to enable independent data collection with accurately time-stamped data files to move closer to a stand-alone system fit for field deployment. Communication between the peripherals is done with Serial Peripheral Interface (SPI) for the digital potentiometers and the SD card reader and Inter-Integrated Circuit (I2C) for the RTC module. A peripheral board with two buttons and an LED served as the user interface of the system during testing. One button was used to toggle system data collection on and off, and the other was used during data collection to indicate when a sea lamprey is attached to the panel, allowing for real-time labeling of the sensor output data. The LED was used as a visual indicator for the state of data collection. An additional peripheral board with a 2 x 2 LED array was used during validation testing as visual indicators for the output of a selection of the detection models. The Arduino board 5 V supply is used

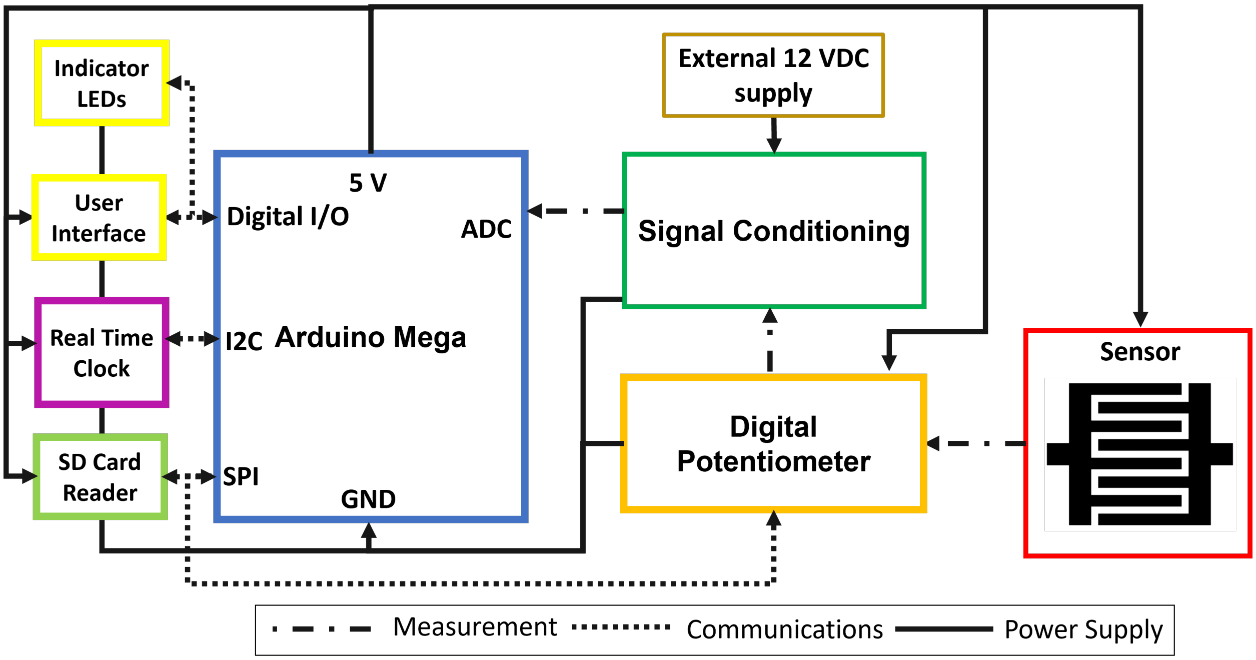


Figure 4.1 Complete sensing system schematic diagram showing connections for signal measurement, data communication and power supply.

to provide power to most elements of the sensing system, though the signal conditioning circuit is powered directly from an external supply. The schematic for the system can be seen in Fig. 4.1.

### 4.1.1 Signal Conditioning

In this updated sensing system a digital potentiometer in rheostat configuration is used as a programable reference resistor in the voltage divider. This allows the system to self-calibrate to a voltage baseline of around 2.5 Volts (within a tolerance of +/- 25 mV) using proportional feedback tuning at the start of each test. Though the sensing area of a single sensor is limited, an array configuration using individual sensors as “pixels” can effectively increase the sensing area at the cost of additional components and processing. However, whenever a multiplexed input (1 sample-and-hold circuit used for all channels) analog-to-digital converter (ADC) is used to measure multiple channels, care must be taken to avoid “crosstalk” noise, where one channel’s measurement is reflected on the next channel being measured. In this system, op-amps in a low-pass voltage follower configuration are used to decouple the voltage dividers from the ADC input, ensuring a small source impedance that

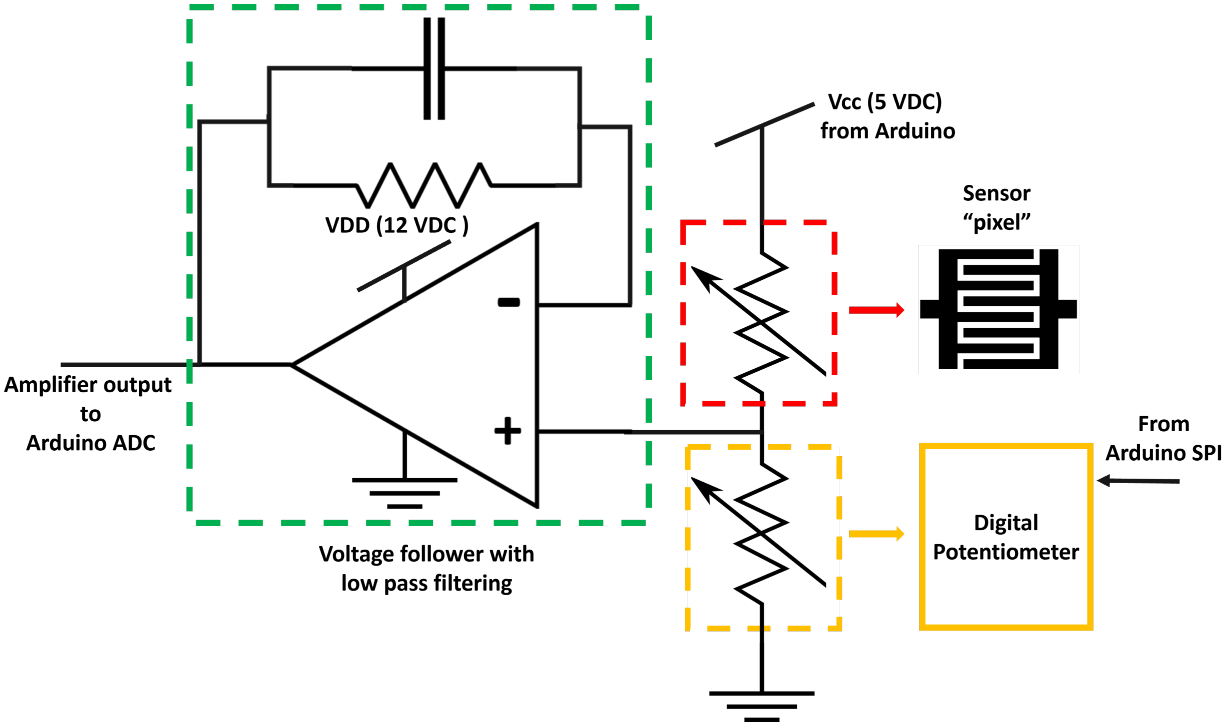


Figure 4.2 Circuit schematic showing the measurement and signal conditioning circuitry for a single sensor pixel. The output of a voltage divider between the sensor and the programmable reference resistance is passed through a filtering voltage buffer before being sent to the ADC module for capture.

allows the sample and hold circuit to discharge properly between measurements. The circuit schematic representing the measurement of a single “pixel” in the sensor array is shown in Fig. 4.2.

### 4.1.2 Device Preparation

Several sensors consisting of a single IDE ‘pixel’ were manufactured to be used in experiments to build a dataset of the sensor response to sea lamprey attachments for this generation of design. System components were integrated onto a large acrylic panel for testing as shown in Fig. 4.3A).

## 4.2 Data Collection, Labeling, and Processing

Large amounts of data measurements (sensor response to lamprey attachment) are necessary to train a model that will accurately determine when a lamprey has attached to the panel. This data was obtained from live-animal experiments at the U.S. Geological Survey, Great Lakes Science Center's Hammond Bay Biological Station (HBBS) in Millersburg, MI. HBBS staff provided 66 lamprey specimens (33 males and 33 females), each tagged with a unique identifier for tracking purposes between experiments. The length, weight, and mouth diameter of the lamprey were measured and recorded as they were tagged for future correlation studies between lamprey size parameters and the sensor response parameters as any relationships found would be useful in differentiating between lamprey species or life stages of the animals attaching to the panel. Data collection was performed during two trips: in the first trip (June 2022) 26 sea lamprey (13 male and 13 female) were tested across three days on a single pixel sensor to begin building a data set; in the second trip (July 2022), 40 sea lamprey (20 male and 20 female) were tested to enhance the dataset.

### 4.2.1 Experimental Setup for Sensor Response Data Collection and Labeling

Experiments took place in a 204 L rectangular aquarium tank supplied with aerated water from Lake Huron. The sensing panel was submerged in the tank and fastened using clamps before testing began. At the beginning of each test, the Arduino was restarted, which generated a new timestamped file that would hold the data for the test. The system then proceeded to self-calibrate itself, indicating when it finished through a serial connection so that the system operator could begin data recording. Once the operator started the recording, a handler introduced a lamprey to the tank and attempted to coax it to attach to the sensor. During the test, the operator pressed the second button on the peripheral any time the sea lamprey attached to the panel, effectively manually labeling the data in real-time. Once the sea lamprey is attached to the sensor uninterruptedly for at least 30 seconds) it is coaxed off the sensor and taken out of the tank, ending the test. The tests were video recorded to corroborate the manual label if needed. A view of a sea lamprey attachment from a sample video recording during one of the single pixel tests, where the lamprey can be seen attaching to the sensor through the gaps in the conductive carbon traces, can be seen in Fig. 4.3D).

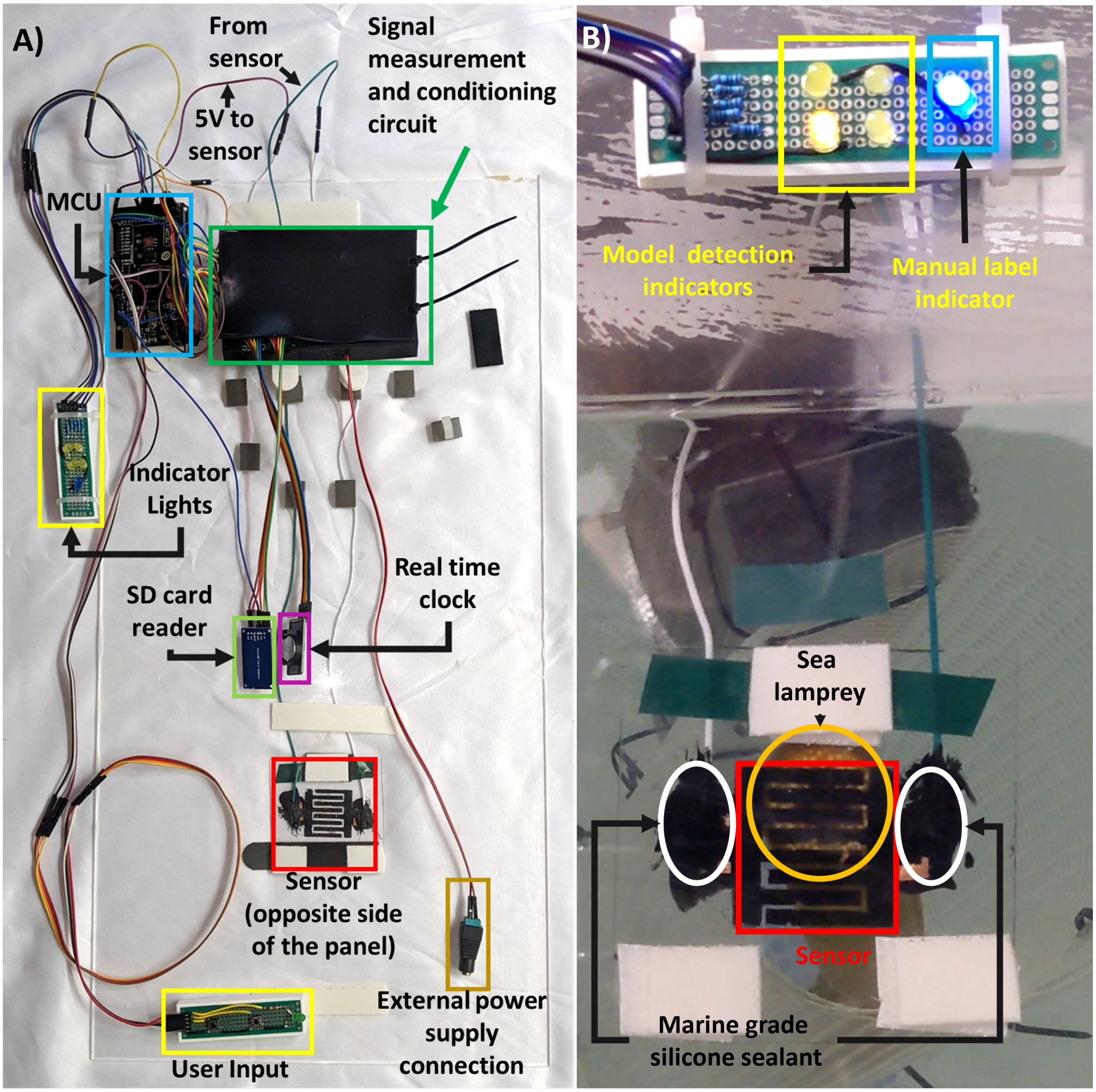


Figure 4.3 A) Full view of the sensing system with all components. B) Indicator lights responding to lamprey attachment during validation testing as seen through the glass tank wall and the IDE sensor traces.

### 4.2.2 Data Processing

Once testing was completed, the data for all tests were compiled into a single set for use in training the detection algorithms. Normalization of the training data was performed to equalize the influence of each input variable on the model since the voltage level of the sensor is about two orders of magnitude larger than the rate of change. Although the sensor response to lamprey attachment is generally a pure decay and it is assumed that the baseline will hold over time due to the self-calibration algorithm, response values greater than the baseline are possible if the lamprey slides on the sensor. Because of this, the voltage was normalized to a maximum expected value of 3 volts (baseline + 0.5 V buffer zone), to completely capture the expected values in a range between [0,1]. For the values of the rate of change, the selection of the normalization factor is not immediately clear from previous research. This is because the rate of change is expected to stay relatively small most of the time, even during the decay, but will have very large positive spikes when the response suddenly goes back to baseline after the sea lamprey detaches. Additionally, an offset will be necessary to fit the values of the rate, which can be positive or negative, to the range [0,1]. The statistical information for the rate data was obtained to aid in selecting the normalization factor and the offset. Since the distribution of the data points looked to follow a normal distribution profile centered about 0, the normalization factor was selected as 3 times the standard deviation of the set. This captured  98% of the data between [-1,1] after which the data was mapped to a range of [0,-1] using a point-slope equation. Values for the average voltage and rate were computed using a rolling average with a window size of 5 on the normalized data.

## 4.3 Classifier Models for Detection Algorithms

There is a consistent trend of combining machine learning models with sensors to create ”smart” systems that can react to or infer things from measured data [48]. While many complex state-of-the-art classifier models exist, this application requires a model simple enough to be deployed on a microcontroller unit (MCU), where memory and power constraints must be considered, for wide-scale deployment of the sensing system to be economically viable.

Implementations of ML algorithms for resource-constrained embedded systems have been explored and proven successful [49, 50]. Embedded machine learning (EML) applications have been demonstrated in many fields, from enabling self-powered chemical oxygen demand sensors [51], to performing tasks such as gesture recognition [52], urban sound classification [53] and injury prediction [54]. This motivates the exploration of ML algorithms for the problem of sea lamprey attachment detection, which can be reduced to classifying the sensor response measurements as corresponding to baseline or attachment. This study focused on ML classifier models that were computationally simple and did not require much memory storage: single-layer artificial neural networks, logistic regression, naive-Bayes, decision trees, random forests, and the scalable, efficient, and fast classifieR (SEFR) [49].

A straightforward thresholding approach (i.e., indicating there is a sea lamprey on the panel if the input falls within a certain region in the input variable space) was also considered due to its low computational complexity, making it well-suited for deployment in low-resource systems. However, manual tuning of the four boundary values to achieve the best detection performance is not practical. This was resolved by posing the tuning as a multiple objective optimization (MOO) problem with the goal of maximizing accuracy measures and using an MOO solver to effectively automate the selection of the thresholding algorithm's parameters.

Previous work on the IDE sensors has shown they present a characteristic decay response to a sustained sea lamprey attachment. Thus, an algorithm that can recognize these decays will be able to detect their attachment onto the panel. The sensor response behaves differently during baseline than it does during the lamprey attachment-induced decay, both in terms of the voltage level and rate of change. This means it might be possible to use a machine learning classifier model on these measurements to identify when the response has decay characteristics and when it has baseline characteristics, effectively identifying sea lamprey attachments. This can be visually confirmed by looking at the scatter plot showing the normalized average voltage levels and normalized average rate of change for all the data points in the training data shown in Fig. 4.4. The data points corresponding to baseline

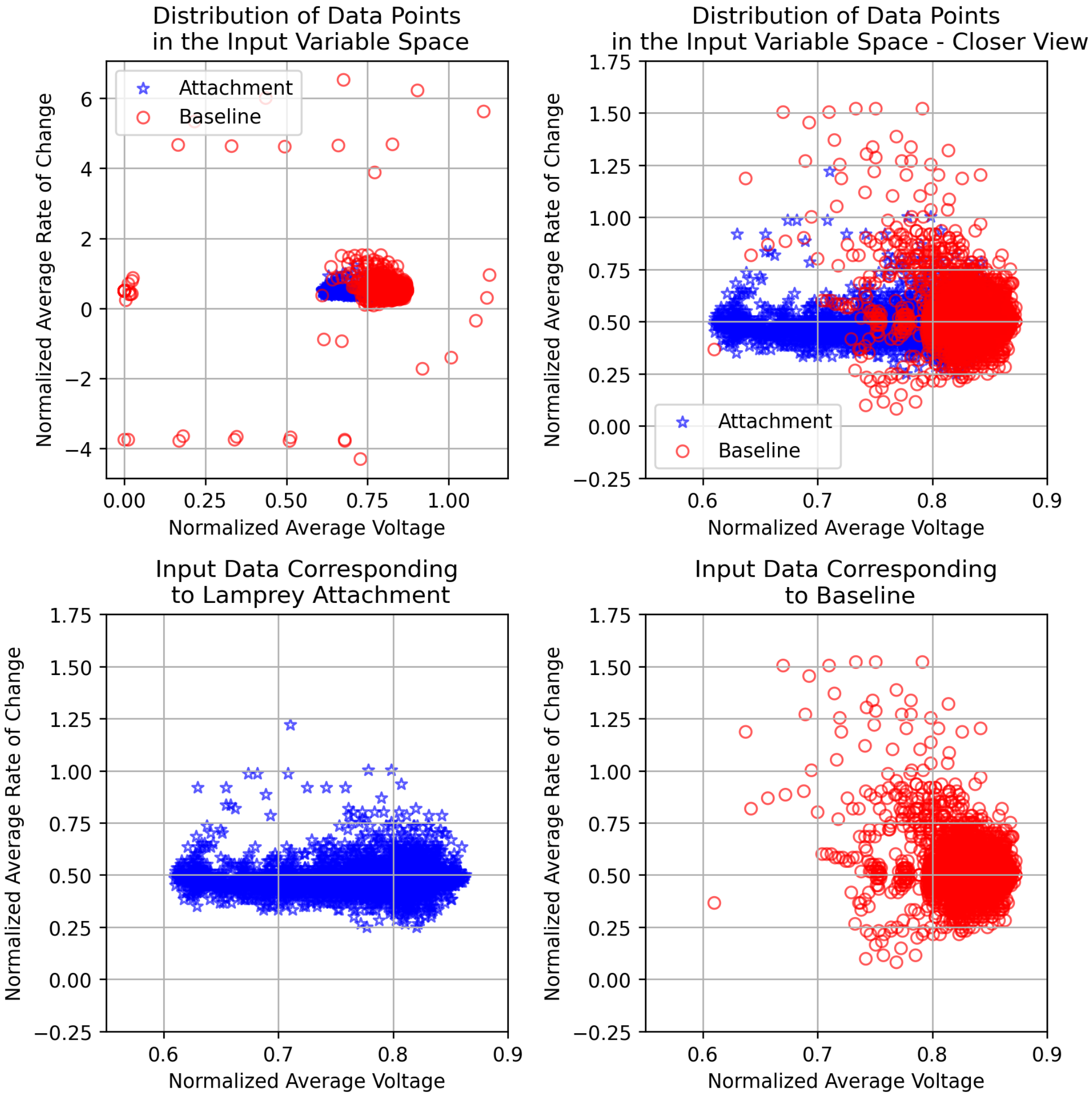


Figure 4.4 A) (Top left) Distribution of training set data in the input variable space. B) (Top right) Close-up of the distribution in the highest density region in the input space. C) (Bottom left) Distribution of data points that correspond to lamprey attachment only. D) (Bottom right) Distribution of data points corresponding to baseline only.

measurements are colored red, while the blue points correspond to sea lamprey attachment. The clustering of the data points around particular regions is a promising indicator that classifier models could differentiate between them. The shape of the clusters also reinforces the potential viability of a simple threshold approach (i.e., indicate there is a sea lamprey on the panel once the signal decays beyond a specified threshold) suggesting that a rectangular threshold model with 4 parameters (upper and lower bounds for each variable) could be sufficient for the task and should be explored as a computationally simple alternative to the ML models as well.

All of the machine learning models studied in this work were trained on a desktop computer (Processor: Intel(R) Core (TM) i7-10510U CPU @ 1.80GHz, 2304 MHz, 4 Core(s), 8 Logical Processor(s), RAM: 16 GB) using the sci-kit-learn python package [55]. The rationale behind their selection as well as details on how they were deployed on the micro-controller system are covered in the following subsections.

### 4.3.1 Neural Network Models

A single-layer artificial neural network (ANN) classifier model was chosen because of its straightforward low-resource implementation (in terms of both memory and computation) for inference in an embedded system when compared to other models such as support vector machines (SVMs), k-nearest neighbors (kNN) or deep neural networks (DNN). The model was deployed on the Arduino Mega using a slightly modified version of the open-source C99 library tinn [56]. This library implements a single-layer ANN with a sigmoidal activation function. It was modified to include a bias for each individual neuron in the layer, which matches the ANN model implementation in scikit learn. The model parameters are extracted from the trained networks and stored as float arrays which are loaded into an empty instance of the model. As this is a binary classification problem, the choice was made to have a single output neuron meaning the output of the model is the probability the inputs correspond to a baseline measurement and not a class label directly. If this value falls below 0.5 then that data is labeled as a lamprey attachment; it is labeled as a baseline measurement otherwise.

Four ANNs of increasing hidden layer size (n = 6, 12, 24, 36) were trained and tested to explore if performance varied with complexity.

### 4.3.2 Logistic Regression and Naive-Bayes Models

Logistic regression (referred to as LGR throughout the paper) and Gaussian Naive Bayes (GNB) classifiers were also considered for the classification task due to their ease of implementation and low computational complexity. LGR is equivalent to reducing the sigmoidal-based ANN to a single output neuron with two inputs and no hidden layer effectively being the simplest network possible. The LGR model was deployed by computing the following equation with the parameters obtained from training:

$$P = \frac{\exp(B + w_1 x_1 + w_2 x_2)}{(1 + \exp(B + w_1 x_1 + w_2 x_2)}$$

where P is the probability of the input data $x_1$ and $x_2$ (in this case the average voltage and average rate of change) being classified as baseline, $B$ is the bias for the LGR, and $w_1$ and $w_2$ are the weights for each variable determined during training.

The GNB model uses a similar approach, under the assumption that the data classes follow Gaussian distributions, the log of the probability of the input belonging to each class is computed, and the class with the highest value is selected as the classification. This is implemented as,

$$y_{pred} = \underset{y}{\operatorname{argmax}} \log\Big(P(y) \prod_{i=1}^{2} P(x_i|y)\Big)$$

where $P(y)$ is the class prior (determined during testing) and $P(x_i|y)$ is the class-conditional probability which for the Gaussian distribution is given as:

$$P(x_i|y) = \frac{1}{\sqrt{2\pi {\sigma_y}^2}} \exp\Big(-\frac{(x_i - \mu_y)^2}{2{\sigma_y}^2}\Big)$$

where $\sigma_y$ and $\mu_y$ are the Gaussian distribution parameters found for each class. This model was deployed on the microcontroller using the MicroMLGen python package [57] which enables users to port a variety of models from the scikit-learn package into C header files for easy deployment on such devices. The model file defines the structure, any required operations, and trained model parameters.

### 4.3.3 Tree-Based Models

Another class of models considered for the detection task are tree-based algorithms. These models can be considered as extensions of the thresholding approach. Their simple structure based on branching conditional statements, called decision trees (DT), is very simple to implement and can be constrained to reduce the model size. Ensembles of multiple trees, known as random forest (RF) models, can also be utilized for inference potentially providing greater accuracy at the cost of a larger model size. The decision tree model generated was limited to a depth of 10 levels and 20 nodes. The RF model consisted of 10 of these base units. The header files for these models were also generated using the MicroMLGen package.

### 4.3.4 SEFR

The SEFR model was specifically developed for binary classification on low-resource systems[49], making it ideal for the target application. During training, the algorithm finds the parameters of a hyperplane defined by the following matrix equation:

$$w^T x + b = 0$$

where $w$ is the weight matrix, $x$ is the input vector and $b$ is the bias. Inference is then performed by evaluating $w^T x + b \leq 0$; if the statement holds the input is classified as an attachment. This model is not included in scikit-learn, but there are wrappers available to use the Python implementation with the package and it is compatible with MicroMLGen for deployment on the Arduino.

### 4.3.5 Threshold Models

The boundary values for the threshold models were generated through a multi-objective optimization-based tuning procedure using the NSGA-2 [58] genetic optimization algorithm. A multiple objective optimization (MOO) problem formulation was selected over traditional single-objective optimization tuning approaches to avoid overfitting that may occur from using a single (potentially class-biased) accuracy metric. By using NSGA-II [58], a genetic MOO algorithm with excellent performance solving two-objective optimization problems, two accuracy metrics can be considered during optimization tuning leading to a more balanced model. The algorithm takes a group of randomly generated solutions contained within the specified upper and lower bounds for each variable and modifies them over multiple generations by computing the objective values for each solution in the population, replacing the worst-performing solutions with slightly altered combinations of the best ones. This generates an optimal set of solutions (known as a pareto-front) which the user can then select from depending on the relative importance of the two objectives. This selection process can be used to further reduce the influence of a biased metric. The precision, recall, and F1 scores for each class in the data set (attachment and baseline) were considered as targets for the tuning problem. The most effective pairwise combination of these six metrics was not immediately clear, so all 15 possible combinations were used to generate threshold models under the same conditions and their performance was examined to determine the best-performing pair. As the NSGA-II algorithm is structured to solve minimization problems the accuracy metrics are rewritten as:

$$F_i = 1 - f_{metric} i = 1, 2$$

Where $F_i$ is the i-th objective function to be minimized, and $f_{metric}$ is any of the six accuracy metrics. This forces the algorithm to minimize the difference between the metrics and their ideal value of 1 (i.e., a perfect score). The threshold values were bounded (between 0 and 1.2 for voltage and -2 and 2 for rate) to obtain reasonable values for the generated models as well as improving the speed of the search. A population size of 100 solutions was evolved over 100 generations to obtain models for each pairwise combination of metrics. The performance of the solutions was evaluated each generation by feeding the training set (same set used to train the ML models) into the generated models and using Scikit-learn's

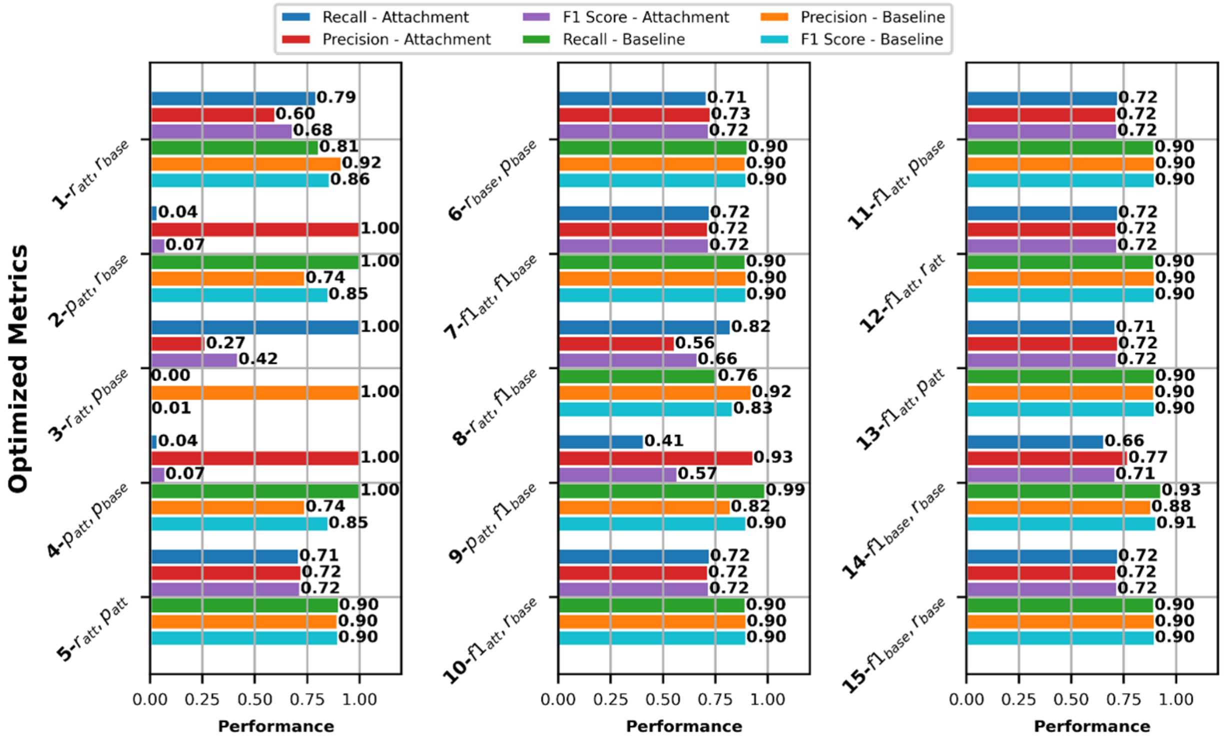


Figure 4.5 Classification performance comparison between the 15 possible accuracy metric combinations on the testing fraction of the training set.

classification report function to obtain the relevant accuracy metrics before comparing them to the target value. The achievement scalarization function (ASF), which uses a weight vector to specify the relevant importance of each objective [59] and extract a solution matching the specified trade-off, was used to select a model from the set of solutions generated by the algorithm. To simply the selection process, equal importance was assigned to each object by using 0.5,0.5 as the weight vector. The optimization and solution selected was facilitated by the use of the PyMOO multiple objective optimization package for Python [60]. The classification performance of the selected models on the testing fraction of the training set for each combination of accuracy metric targets is presented in Fig. 4.5.

Several trends in performance can be identified among the generated models. Three models (identified as 2, 3, and 4 in Fig. 4.5) presented very poor classification performance overall, perhaps due to antagonism between selected tuning parameters or bias toward one class, with critically low scores for one of the two classes and were discarded from consideration for deployment. Eight of the fifteen models (identified as items 5, 6, 7, 10, 11, 12, 13, and 15 in Fig. 4.5) show almost identical performance across metrics. These present recall and precision scores around 0.72 respectively for the attachment class indicating that a large majority of the points corresponding to the class are captured and that most of the data points identified as attachments correspond to that class. Models 1 and 8 presented the best recall scores for attachment of the group, though their precision scores are on the lower end of the group as well showing the tradeoff in performance between the two metrics. The final models have distinct performance profiles with 9 at the opposite end of the spectrum showing the highest precision for the attachment class and the lowest recall score of the viable models and 14 presenting a middle point in performance between the two extremes. After examining the results from the 15 possible models models 1,9 and 14 were the ones selected for real-time deployment testing to explore the range of performance that the thresholding models demonstrated in the training set to represent the range of performance observed. The first of these, referred to as Thresh 1, was tuned using the recall scores of the attachment and baseline classes, which resulted in a higher capture rate for attachment data points at the cost of higher rates of false negatives. Thresh 2 was tuned using the precision for the attachment class and the f1-score for the baseline class, which resulted in a more conservative model with fewer false positives but also less attachment capture. The final model, Thresh 3, was tuned using the f1-score and recall score for the baseline class and demonstrated a performance between Thresh 1 and Thresh 2. The boundary values are presented in Table. 4.1.

Table 4.1 Boundary Values for Threshold Models

| Model | Lower Bound$_{Voltage}$ | Upper Bound$_{Voltage}$ | Lower Bound$_{Rate}$ | Upper Bound$_{Rate}$ |
|---|---|---|---|---|
| Thresh 1 | 0.32421867, | 0.82237967 | 0.23528082 | 1.22225504 |
| Thresh 2 | 0.51726076 | 0.79880256 | 0.32990316, | 0.70216792 |
| Thresh 3 | 0.38311498 | 0.81407758 | 0.27848292 | 0.64207279 |

To implement these models on the Arduino each data point is evaluated against the bounds and if it falls within the outlined region it is labeled as an attachment.

### 4.3.6 Testing the Models

The inference time for the models on the Arduino was verified before validating their performance. The testing fraction data set was saved as a .csv file on an SD card where it could be read by the microcontroller. The data was input into the model implementations and the time to inference was recorded and sent over serial to a PC computer along with the data and the inferred classification.

The performance of the models at identifying sea lamprey attachment was validated at HBBS using live animal testing and baseline data collection during July 2023. Animal tests followed the same procedure as those utilized to build the training set. The raw data was normalized with the same values used for normalizing the training set, though a longer averaging window was used (15) due to considerable noise observed in the sensor output during the validation testing. Baseline data collection is a similar procedure, but no sea lamprey is introduced into the test tank for the duration of the data collection. In total, 105 data recordings comprised of 97 live-animal tests over 80 different lamprey (40 male and 40 female) and 8 baseline collections, were utilized in the validation of the models. Predictions by the models were recorded alongside the input data and the manual label data to .csv files to be analyzed in Python using the scikit-learn package.

## 4.4 Results

### 4.4.1 Memory Requirements and Inference Speed

The memory requirements of the deployed models are presented in Table 4.2. All models required under 3.5 Kb of program memory for their implementation on the microcontroller. Dynamic memory requirements are also low overall, though the implementation of the ANN stores the model parameters as program variables in the flash memory resulting in higher memory usage for these models. The model file column of table4.2 shows the size of the C-header files defining the model structure, required operations, and trained model parameters (for all but the ANN models). Even the largest of these files, the random forest model, does

Table 4.2 Model Memory Requirements

| Model | Program Memory (Bytes) | Dynamic Memory (Bytes) | Model File Size (Kilobytes) |
|---|---|---|---|
| NN-6 | 2698 | 186 | 6 |
| NN-12 | 2794 | 282 | 6 |
| NN-24 | 2988 | 474 | 6 |
| NN-36 | 3182 | 666 | 6 |
| LGR | 786 | 17 | n/a |
| GNB | 1518 | 16 | 3 |
| DT | 934 | 16 | 5 |
| RF | 3454 | 16 | 39 |
| SEFR | 870 | 17 | 1 |
| Thresh | 276 | 17 | n/a |

not exceed 39 Kb, and the LGR and thresholding models did not require them since their structures were simple enough to be included directly in the program code.

The inference time, how long it takes for the models to give a prediction, is another measure of the complexity of the models and could be a key limitation at this scale of computing power. The inference times were measured by having the deployed models perform classification on a preprocessed data set (which the Arduino reads from a CSV file using the SD card adapter) and capturing the time it takes for the model to deliver a classification. Figure 4.6 shows the inference results for all of the models. Most of the models require less than 1 ms to compute their decision, and even the largest model, NN-36, achieved an average inference time of just under 9.4 ms. Since the models are relatively compact, and their combined inference times are considerably less than the 1 second sampling period of the system, multiple models can be deployed on the same microcontroller.

### 4.4.2 Classification Performance

#### 4.4.2.1 Testing Fraction of the Training Set

Table.4.3 shows the classification performance obtained on the testing fraction of the training set for each of the models in the study. The ANN models showed promising performance across all levels of complexity, capturing almost half of the data corresponding to attachments (recall), with a high percentage of true positives among those identified as attachments (precision) and an overall accuracy of 84%. The LGR model appears to be the most conservative one with the lowest recall for the attachment class and tying for the highest precision for that class with the Thresh 2 model. The GNB model's performance is close to that of the ANNs, though it appears to be slightly more conservative at assigning attachment labels and thus captures less of the attachment data. The DT and RF forest models' performance was also promising, with higher attachment recall scores than the ANNs while having similar precision scores. They are also tied with the Thresh 3 model for the highest overall accuracy score at 86%. The SEFR model performed similarly to the ANN models in its recall score but it also delivered one of the lowest precision scores for the attachment class. The thresholding models each showed distinct performance profiles. Thresh 1 was the most aggressive model overall capturing 79% of attachment class data though it also presented the lowest precision score for attachment and the lowest accuracy overall at 80%. Thresh 2 was one of the more conservative models with the second-lowest recall score but it also showed the highest precision for the attachment class and an overall accuracy of 83%. Thresh 3's

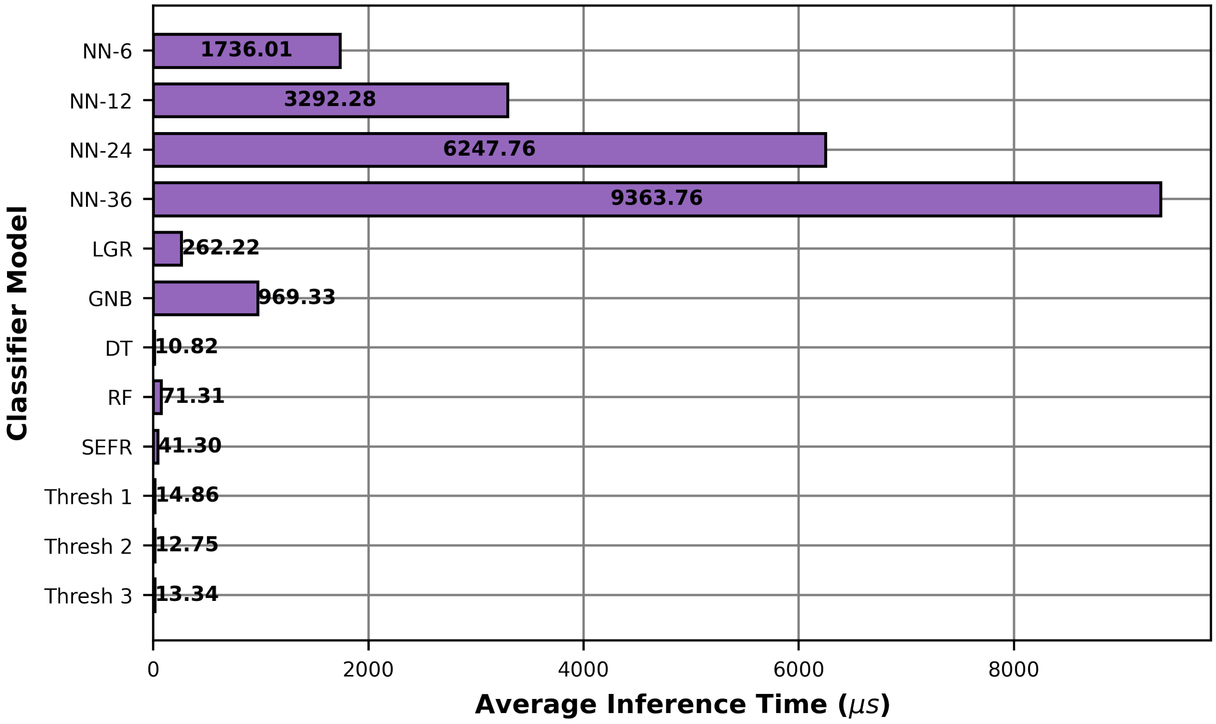


Figure 4.6 Average inference time comparison.

Table 4.3 Classification Performance in the Testing Fraction of the Training Set

| Model | $Recall_{att}$ | $Precision_{att}$ | $F1_{att}$ | $Recall_{base}$ | $Precision_{base}$ | $F1_{base}$ | Accuracy |
|---|---|---|---|---|---|---|---|
| NN-6 | 0.48 | 0.84 | 0.61 | 0.97 | 0.84 | 0.90 | 0.84 |
| NN-12 | 0.49 | 0.83 | 0.61 | 0.96 | 0.84 | 0.90 | 0.84 |
| NN-24 | 0.49 | 0.81 | 0.61 | 0.96 | 0.84 | 0.90 | 0.84 |
| NN-36 | 0.48 | 0.84 | 0.61 | 0.97 | 0.84 | 0.90 | 0.84 |
| LGR | 0.32 | 0.93 | 0.48 | 0.99 | 0.80 | 0.89 | 0.81 |
| GNB | 0.43 | 0.87 | 0.58 | 0.98 | 0.83 | 0.90 | 0.83 |
| DT | 0.67 | 0.77 | 0.72 | 0.93 | 0.89 | 0.91 | 0.86 |
| RF | 0.64 | 0.80 | 0.71 | 0.94 | 0.88 | 0.91 | 0.86 |
| SEFR | 0.49 | 0.70 | 0.58 | 0.93 | 0.83 | 0.88 | 0.81 |
| Thresh 1 | 0.79 | 0.60 | 0.68 | 0.81 | 0.92 | 0.86 | 0.80 |
| Thresh 2 | 0.41 | 0.93 | 0.57 | 0.99 | 0.82 | 0.90 | 0.83 |
| Thresh 3 | 0.66 | 0.77 | 0.71 | 0.93 | 0.88 | 0.91 | 0.86 |

Table 4.4 Classification Performance in Real-Time Validation Testing

| Model | $Recall_{att}$ | $Precision_{att}$ | $F1_{att}$ | $Recall_{base}$ | $Precision_{base}$ | $F1_{base}$ | Accuracy |
|---|---|---|---|---|---|---|---|
| NN-6 | 0.61 | 0.56 | 0.59 | 0.77 | 0.80 | 0.78 | 0.72 |
| NN-12 | 0.61 | 0.56 | 0.59 | 0.77 | 0.80 | 0.78 | 0.72 |
| NN-24 | 0.65 | 0.55 | 0.60 | 0.74 | 0.81 | 0.78 | 0.71 |
| NN-36 | 0.63 | 0.56 | 0.59 | 0.76 | 0.81 | 0.78 | 0.71 |
| LGR | 0.48 | 0.60 | 0.53 | 0.84 | 0.77 | 0.80 | 0.72 |
| GNB | 0.79 | 0.52 | 0.63 | 0.64 | 0.86 | 0.74 | 0.69 |
| DT | 0.74 | 0.56 | 0.64 | 0.72 | 0.85 | 0.78 | 0.73 |
| RF | 0.73 | 0.56 | 0.64 | 0.72 | 0.85 | 0.78 | 0.73 |
| SEFR | 0.63 | 0.55 | 0.59 | 0.75 | 0.81 | 0.78 | 0.71 |
| Thresh 1 | 0.84 | 0.48 | 0.61 | 0.56 | 0.88 | 0.68 | 0.65 |
| Thresh 2 | 0.53 | 0.64 | 0.58 | 0.85 | 0.79 | 0.82 | 0.75 |
| Thresh 3 | 0.76 | 0.56 | 0.64 | 0.71 | 0.86 | 0.77 | 0.72 |

performance sits between these two extremes, mirroring that of the DT and RF models.

#### 4.4.2.2 Real Time Validation Tests

The performance of the models at identifying sea lamprey attachment was validated at HBBS using live animal testing and baseline data collection. Animal tests followed the same procedure as those utilized to build the training set and the raw data was filtered and normalized using the same parameters used on the training data. Baseline data collection is a similar procedure, but no sea lamprey is introduced into the test tank for the duration of the data collection. The fast inference times and low memory requirements of the models allowed for simultaneous deployment of all models on the microcontroller; every time a measurement is taken and processed, each algorithm, in turn, uses this data to infer the status of the sensor. A link (11) to video showing how a selection of four models (NN-24, RF, Thresh 2, and LGR) react through the 2 x 2 indicator array during the validation test (as seen in Fig 4.3C) is included in Section 11 of the appendix. This inference is compared to the manual label generated during testing to see how effective the model is at classifying the sensor state. Table. 4.4 provides a comparison of classification performance for all models during the real-time validation tests performed at HBBS. The relative performance profiles the models demonstrated during validation testing follow the general patterns that were highlighted by the testing fraction classification results. The deployed models were generally successful at correctly identifying the state of the sensor the majority of the time (model accuracy ranges from 69-75%). All models, with the exception of the LGR model, were able to capture more than 50% of the data points corresponding to sea lamprey attachment. The Thresh 2 model had the highest accuracy score among all of the models, followed by the decision tree and random forest models. The ANN models followed closely in accuracy across levels of complexity, as they did for the testing fraction.

Overall, when compared to the classification results in testing fraction, the deployed algorithms fared better at identifying a larger share of the data points corresponding to attachment, noted by the higher recall scores for that class across the board. However, this appears to have come at the cost of a higher amount of false positives as shown by the lower precision scores. There was a similar hit in performance in the recall score for the baseline class, though the precision score for that class stayed near the values for the testing fraction. The overall accuracy of the deployed models, though lower than on the testing fraction, is still considerable.

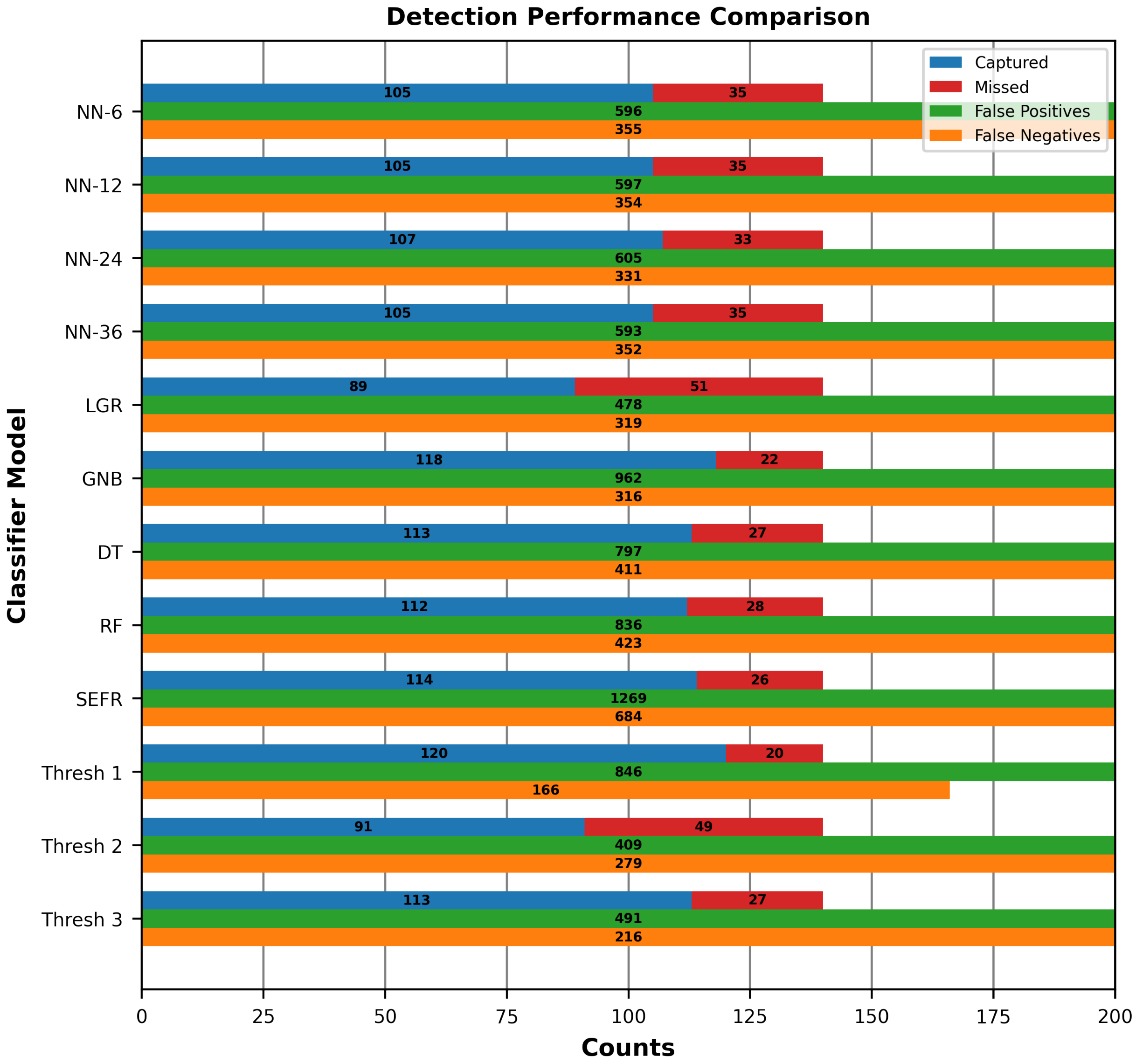


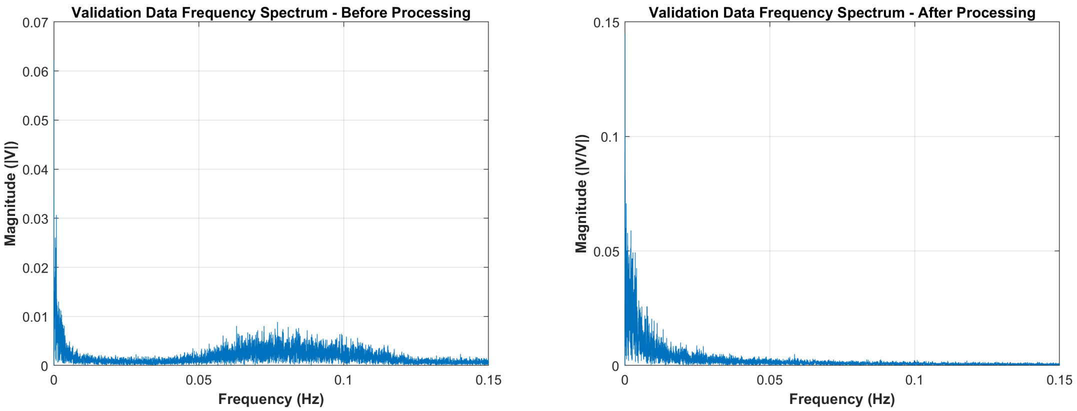
Figure 4.7 Comparison of the detection performance of each model during the real-time validation testing.

### 4.4.3 Detection Performance in Real Time Validation Test

Classification performance measures how well the model sorts individual data points as baseline or attachment. This will give a general idea of the detection performance of the algorithms, however, the task of detection/monitoring is more complicated because it requires sustained correct inferences across time. An algorithm that can measure the model performance at this task was developed by considering the requirements of the detection tasks. The algorithm counts an attachment as detected if the model output correctly identifies it at any point in the attachment. Any continuous sequence of false positives is considered as a single false detection and any sequence of false negatives is considered a single false baseline instance. The algorithm also reports the time it takes for the model to first detect an attachment (detection delay) and the length of missed attachments as well.

Fig. 4.7 provides a comparison of detection performance in the real-time validation tests in terms of the attachment captured and the false positives and negatives generated during the tests. All of the algorithms managed to capture a majority of the attachments that happened during validation testing with capture rates from 63-85%. However, they also generated high numbers of false detections as well as false baseline captures. The models were trained to respond to the decay generated by sea lamprey attachment which is a fairly slow process. Considerable low-frequency noise from 0.05-0.12 Hz Hz (as seen in Fig. 4.8a) of unknown origin was present in the raw measurements across all the validation tests performed and could have contributed to the high amount of false positives observed. The models could be reacting to the decaying half cycle of these slow periodic oscillations and incorrectly identifying them as attachments. Although Fig. 4.8b shows that the noise at that band has been reduced by the data processing, there was a corresponding increase in even lower frequency noise which is more likely to be picked up by the models. Further evidence of this is the fact that the models with the lowest capture rates (LGR and Thresh 2) also have the lowest numbers of incorrect detections, potentially indicating a relationship between the sensitivity to the characteristic decay and the rate of false positives.

Detection time, how quickly a model first detects a sea lamprey attachment, is an important performance metric to consider for this task as it highlights a key limitation for certain applications. Fig. 4.9 shows the timing performance for each model throughout the validation testing. Fig. 4.9a shows the Thresh 1 model as the fastest at identifying sea lamprey attachment, with the SEFR and GNB models close behind. However, these are also the three models with the highest amount of false negatives, again highlighting the trade-off between speed/capture rate and overall accuracy. The models are slow to react to attach-



(a) (b)

Figure 4.8 A) Frequency spectrum (with the constant component removed) of the raw data collected during validation. B) Frequency spectrum (with the constant component removed) of the processed input data for the models.

ment overall; the fastest models required several seconds to react to sea lamprey attachment and the slowest required more than 30 seconds. This is not unexpected as the decay due to sea lamprey attachment is a slow process, with decay constants in the tens of seconds [26], so it takes some time before the sensor response shifts significantly from baseline. Also, because the speed and amplitude of the decay are positively correlated with the amount of sensing area covered by the sea lamprey, factors such as fish size and location of attachment can affect how ”detectable” an attachment is and how long it takes to get there.

For these reasons, it is also useful to look at the length of the attachments that are missed as another limiting factor. Figure 4.9b), shows the average length of missed attachments during the validation testing for all of the models. For the activation function-based models (ANNs, LGR and GNB) and most thresholding models (Thresh 1 and Thresh 2), this value is slightly larger than the average detection delay. This could indicate that the attachments missed by these models were those where the position or size of the sea lamprey caused a slow and low-magnitude response that ended before getting to detectable levels. In the remaining models (DT, RF, SEFR and Thresh 3) the value is lower than the average detection, possibly indicating that these are more sensitive to these types of attachments.

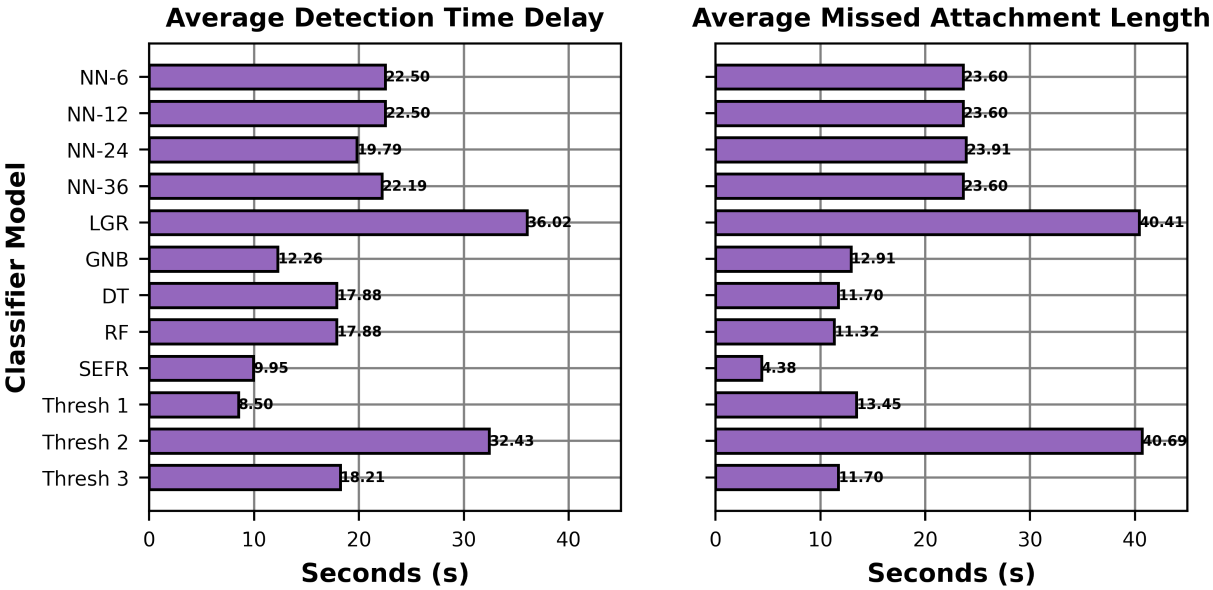


Figure 4.9 A) Comparison of average delay before detection, B) Comparison of average missed attachment lengths.

## 4.5 Discussion

There are several ways a sea lamprey detection system could be utilized to help control the invasive species in the Great Lakes, each with different performance requirements. The least demanding of these would be sea lamprey presence detection, where the system would only be used to determine if sea lampreys are in a particular area. In this application, the output of the detection algorithm would only be scrutinized to determine if any attachments had occurred within a specified time frame. Possible applications include, deployment in places where sea lamprey gather but traps are difficult to deploy to enhance monitoring capabilities; or deployment in streams with seasonal barriers to better inform installation timing. For this application, the system should have a low false positive rate, to avoid directing resources where they are not needed, with detection speed being a secondary consideration. If the output is further analyzed to count the number of attachments in a given period, an activity index could be generated that would provide more detailed information regarding lamprey distribution, enabling more efficient resource allocation. This application would also require a low false negative count to be viable, as false negatives during an attachment could result in it being counted multiple times, which could also lead to improper resource allocation.

All the models in the study presented a relatively large number of false detections (false positives and negatives), ranging from 688 to 1953 counts as seen in Fig. 4.7, apparently making them unsuited for the presence/activity index applications. However, the presence of low-frequency noise in the input data, which exhibits behavior close to the slow decay and sudden increase behavior the algorithms were trained to pick up on, likely led to elevated levels of false positives and negatives during validation testing. This noise could be addressed with additional noise removal techniques like inverse Fourier transform filtering or by expanding the input window of the models so they observe the decay over a larger window. Because multiple models can be deployed at once, it is also possible that a voting scheme can be implemented between them to improve overall system performance. However, this noise was likely generated by equipment in the lab space used for the testing and is unlikely to happen where these devices will be deployed in the first place. Thresh 2, Thresh 3 and LGR models had the lowest combined amount of false detections (positive and negative), making them the leading candidates for the presence/activity index application once noise considerations are addressed. The ANN models presented higher numbers of false detections but they still perform considerably better than the remaining models and could also prove useful in this application. The lower detection delay and higher capture rate of the Thresh 3 model make it the more attractive choice for this application overall.

The system also has potential applications in smart fish passage systems, where barriers can be activated to selectively deny upstream passage in the presence of invasive species. The sensors could also be deployed with an integrated deterrence mechanism (i.e., a stimulus that causes pain/discomfort) in areas of high flow, where sea lamprey are likely to attach to conserve energy during spawning migrations, as a way to make the journey to their spawning grounds harder. These applications place a greater emphasis on fast detection times, to engage the corresponding mechanism before the sea lamprey has a chance to move on upstream, but there is a degree of flexibility with regard to false positives as long as native species are not negatively impacted. Thresh 1, SEFR, and GNB all respond within 15 seconds of attachment, indicating they would be prime candidates for these applications, especially with their high capture rates. The tree-based models, DT and RF, were slightly slower but appear to be better suited for the task than the remaining models.

Several external factors, like the size of the sea lamprey and the location and length of the attachment will affect the likelihood of detection with this system. The magnitude of the response for this type of sensor is heavily dependent on the size of the sea lamprey attaching to the panel as well as the location; the more the sensing area, is covered the larger the magnitude of the response. Indeed, some of the models' average length of missed attachments are longer than the average time to detection, meaning it is likely that the response magnitude was too small to be differentiated from the baseline's natural variation. This may be a considerable problem for a single-pixel sensor with limited sensing area depending on the application, but could be solved by an extension of the detection algorithm for array configurations, which may consider data from neighboring pixels when determining if a sea lamprey is contacting the array. The attachments in the experiments conducted to build the training set used in this paper lasted at least 30 seconds by design, but there is a lack of detailed information on how long sea lampreys attach to surfaces in the wild. However, skillful placement of the sensors may mitigate this issue somewhat; as lampreys generally attach to surfaces to rest before continuing their journey toward spawning grounds, placing the sensors in areas that require considerable exertion to get through could be a way to promote the long attachments required for detection.

### 4.6 Summary

This chapter presents the use of simple machine learning and tunable thresholding models to identify the characteristics of an IDE sensor's response to sea lamprey attachment, enabling detection of the invasive species. The average voltage and rate of change of voltage were selected as the inputs for the detection algorithm since the characteristic exponential decay response will also have a rate profile that is distinct from the baseline rate. A training set of labeled sensor response data was constructed through experiments with live sea lamprey on IDE sensors to train the selected models. A threshold model that could be tuned via optimization approaches was presented as the simplest detection approach, flagging data points within specified bounds in the input space as attachments. Artificial neural networks, logistic regression, Gaussian Naive Bayes, decision trees, and random forest classifier models were also explored. These were successfully deployed on a microcontroller platform interfacing with an IDE sensor and achieved classification accuracy between 65-75% in live animal validation testing. The detection performance of the models was examined and they demonstrated high numbers of false positive and negative detections, though noise observed during testing very likely contributed to this. The models did manage to capture the majority of the attachments with capture rates from 63-85%. The slower, more conservative models (Thresh 2, Thresh 3, LGR, ANNs), which report lower numbers of false detections are better suited for monitoring applications for the sensor, whereas the faster more aggressive models (Thresh 1, SEFR, GNB) would be a better fit for reaction/deterrence applications. Thresh 3, which used x and y as a cost function during tuning, can be considered the best-performing algorithm as it makes the most appearances in the top 5 performers across the performance criteria presented here. It is possible to continue refining the presented models on a powerful workstation and simply follow the same deployment methodology outlined in this paper to deploy the improved algorithms. While attachment time and location are parameters that could affect the performance of these sensors, possible workarounds for both issues are presented.

# CHAPTER 5

# THIRD GENERATION MEASUREMENT SYSTEM AND SENSOR DESIGN

This chapter presents improvements to the sensing system to address the sensitivity/detection speed and sensor robustness issues presented in Chap. 4. This includes the design of the third-generation measurement approach developed to increase the magnitude of the sensor response through amplification to improve sensitivity and detection speed. The implementation of this system using discrete components is presented. The design of the third-generation sensors, which aims to address robustness concerns by replacing conductive carbon paint as electrode material with 3D-printed conductive carbon-based polymer, is also presented. The chapter ends with a presentation of the updated fabrication method.

## 5.1 Third Generation Sensing System Design

A new sensor measurement approach was developed to address the shortcomings of the voltage divider measurement circuit, namely the inability to amplify the sensor response signal independently from the DC bias required to observe both positive and negative changes

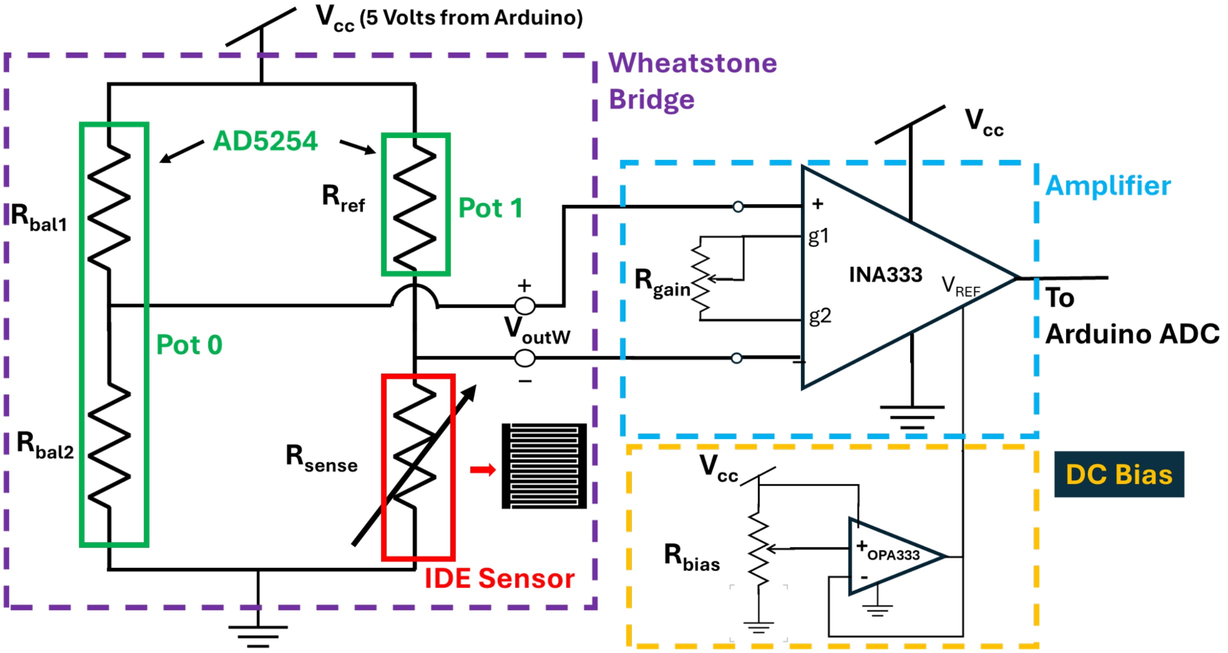


Figure 5.1 Circuit schematic showing the Wheatstone bridge and instrumentation amplifier-based measurement system for a single IDE sensor. The output of the bridge is amplified by the in-amp and offset by a DC bias generated from a voltage follower circuit.

in sensor resistance. Separating the signal response from the bias is required to allow for targeted amplification, which would increase the overall sensitivity of the system while retaining its bidirectional measurement capabilities. A Wheatstone bridge circuit topology, highlighted in purple in Fig. 5.1, was selected to enable this separation. This topology adds another set of resistors, which will be referred to as the reference branch, in parallel with the standard voltage divider that was previously used to measure the sensors, and which will now be referred to as the sensing branch. The output of this circuit is the difference between the node voltages in the middle of each branch, thus, the circuit not only converts the change in impedance across the sensor to a voltage signal but also directly compares it to a reference voltage set by the reference branch. This effectively extracts the sensor output response, allowing it to be amplified. Using a voltage divider analysis for the circuit shown in Fig.3.1-A, the output voltage ($V_{out}$) can be calculated as:

$$V_{\mathrm{outW}} = \left( \frac{R_{bal2}}{R_{bal2} + R_{bal1}} - \frac{R_{sense}}{R_{sense} + R_{cal}} \right) V_{\mathrm{input}} \tag{5.1}$$

The Wheatstone bridge is implemented using Analog Devices' AD5242, which is a digital potentiometer chip with two individually controllable digital potentiometers with 256 possible resistance values going from approximately zero ohms to the nominal resistance value of the device. The resistance across the A and B terminals of each potentiometer is fixed to the nominal value (within tolerances), and the resistance between either terminal and the wiper is modified by sending instructions to the device. The chip is controlled using the I2C protocol, which is compatible with Arduino devices. The first potentiometer on the chip (Pot 0) is used to control the proportion of resistances in the reference branch of the circuit. The second potentiometer (Pot 1) is used in rheostat configuration to adjust the reference resistance in the sensing branch.

Texas Instruments' INA333, a low-power, rail-to-rail instrumentation amplifier, was selected as the amplification stage of the sensing system because of the topology's high input impedance, ensuring minimal disturbance to the measurement circuit. The gain is set through a single external resistor and can vary from 1 to 1000, following $A = 1+100k\Omega/R_{gain}$. A 100 kΩ potentiometer was used as $R_{gain}$ to allow for setting any gain within this range. The INA333 can also apply a bias or reference voltage to the output by providing the corresponding voltage at the $V_{ref}$ pin so that their output follows the following equation: $V_{outA} = A * (V_+ - V_-) + V_{ref}$. This voltage must have a low source impedance to avoid affecting the gain circuit, so it is provided through a voltage buffer circuit using TI's OPA333. The positive input of the buffer is tied to the wiper of another 100 kΩ potentiometer $R_{bias}$ with its main terminals connected between Vcc and ground, allowing for bias voltages across the entire supply range. Additionally, for the purpose of calibrating the system, two more voltage followers are included in the measurement system, allowing for direct measurement of the voltage across $R_{sense}$ and $R_{bal2}$.

## 5.2 Third-Generation Sensor Design and Fabrication

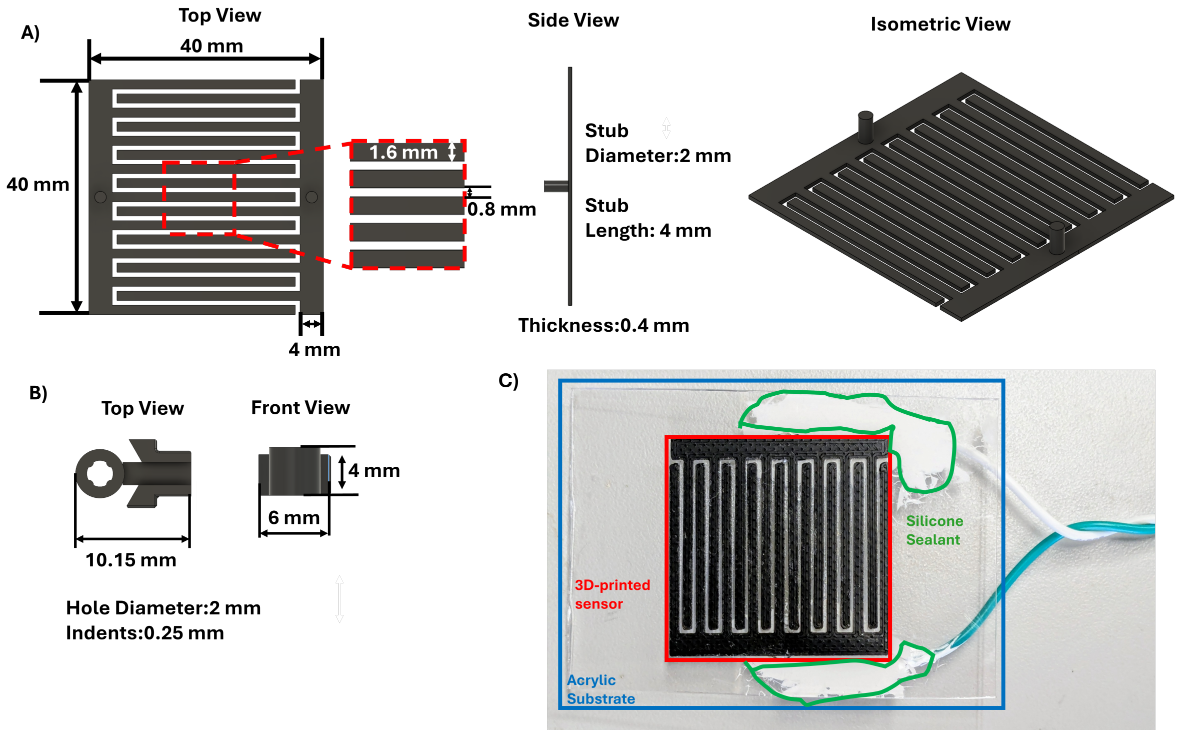


Figure 5.2 A) Schematic view of the third-generation sensor with highlighted dimensions includes different points of view to capture 3D geometry. B) Schematic view of 3D printed connectors for lead attachment, including relevant dimensions. C) Example of a manufactured 3rd-generation IDE device.

Conductive carbon PLA was selected as the electrode material for the third-generation sensing devices due to having increased robustness while maintaining similar electrical characteristics. The schematic for the third-generation design is presented in Fig. 5.2-A. The move towards 3D-printed fabrication also allowed for improvements in the design, including a smaller gap between electrodes and a higher number of fingers per electrode. This effectively increases the fill ratio of the device, which should lead to a corresponding increase in sensitivity. The updated design features a slightly larger sensing area (40 mm x 40 mm) as well as stubs, seen in the side view of Fig.5.2-A, that protrude from the back of the electrodes to facilitate the attachment of electrical leads on the opposite side of the substrate using a conductive 3D-printed connector presented in Fig.5.2-B. The electrode's minimum thickness (0.6 mm) was determined iteratively, reducing the thickness of the design until the 3D printer failed to consistently produce a viable sensor. Trace width was selected to optimize the fill factor of the device under the area and separation constraints. An example of the third generation device is presented in Fig. 5.2-C.

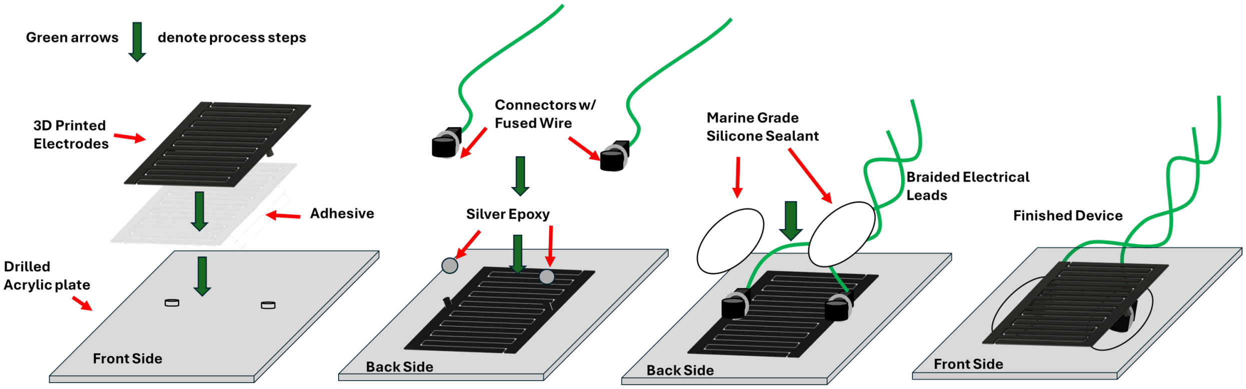


Figure 5.3 Third-generation device manufacturing flow.

Clear acrylic panels were kept as the substrate for the updated IDE sensor to allow light to pass through them and promote algal growth on their surfaces. Conductive carbon-based polylactic acid (PLA) filament (Electrically Conductive Composite PLA by Protopasta) was used as the electrode material, replacing the Marine-grade conductive carbon spray paint. Fig. 5.3-A shows the device fabrication process beginning with the electrodes being 3D-printed on a standard fused deposition modeling printer. The square acrylic panel substrate (approximately 80 mm x 80 mm) is then drilled to allow the sensor stubs to pass through to the opposite side for electrical lead connection. The 3D-printed electrodes are then attached to the substrate using a waterproof adhesive. A soldering iron is then used to melt the acrylic around the stub to ensure a watertight seal around the lead attachment area. The leads are prepared by partially stripping an insulated wire and wrapping the exposed conductor around the conductive 3D-printed connector, highlighted in Fig. 5.3-A, and fusing the wire to the connector using a soldering iron. Conductive silver epoxy is then applied to the stub before attaching the connector to ensure a robust electrical connection. Marine-grade silicone sealant is then used to waterproof the lead attachment. Finally, the leads were braided together to minimize the impact of ambient electrical noise on the sensor response signal. The devices are then subjected to open-circuit and short-circuit verification tests before proceeding to testing.

# CHAPTER 6

# CHARACTERIZATION OF SENSOR RESPONSE TO ALGAL BIOFILM ACCUMULATION

This chapter presents the attempted characterization of the response of the sensing system to algal biofilm accumulation to better understand how the system will perform in the field. The methodology used for the characterization experiments using a live microalgae culture in a controlled setting is presented. The response of the sensor voltage during the biofilm accumulation tests and the change in sensor resistance after each test are analyzed and compared to biomass accumulation measures to determine if there is a relationship between either sensor output and biological accumulation on the sensor. The ability of the sensors to detect biological contact after developing biofilms on their surface is also examined. Finally, the implications of these results for sea lamprey detection and potential ways to address them are discussed.

## 6.1 Experimental Design for Algal Biofilm Accumulation Response Characterization

### 6.1.1 Experimental Set Up

The characterization experiments took place in a UTEX 2L lab-scale photobioreactor (PBR) paired with a UTEX RGB lighting platform. The PBR system consists of three major components: a clear glass tank with a gas inlet connected to an aeration stone; a pump that sends pressurized ambient air to the tank; and a lighting rig that provides full-spectrum light for algal nourishment. To induce flow conditions that are conducive to biofilm formation, the entire system was placed on top of a magnetic stirrer, and a 4-cm stir bar was added to the PBR tank. A picture of the full experimental setup including the PBR, lighting rig, and the magnetic stirrer is shown in Fig. 6.1-A. The microalgae that were cultured during testing are *Chlorella sorokiniana MSU*; it is a robust alga found in the Great Lakes region, which is readily available. The use of native algae species will also make the experiments more relevant to potential biofilm evolution in the field.

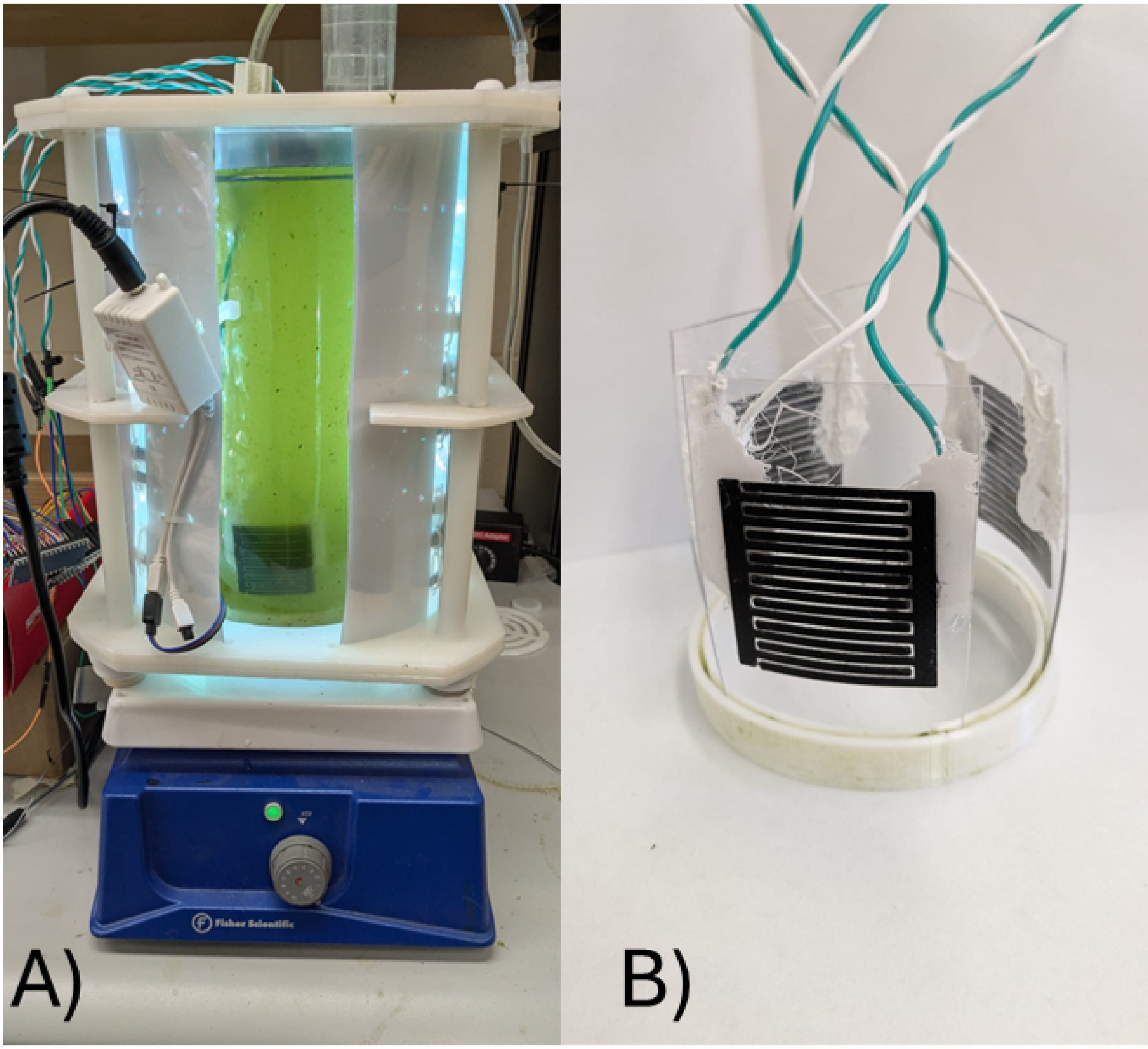


Figure 6.1 A) Experimental setup for biofilm accumulation tests, including the glass PBR tank, RGB LED lighting rig, and magnetic stirrer, a sensor is seen through the gap in the lighting rig. B) Sensors placed in a 3D-printed base before being introduced to the PBR tank.

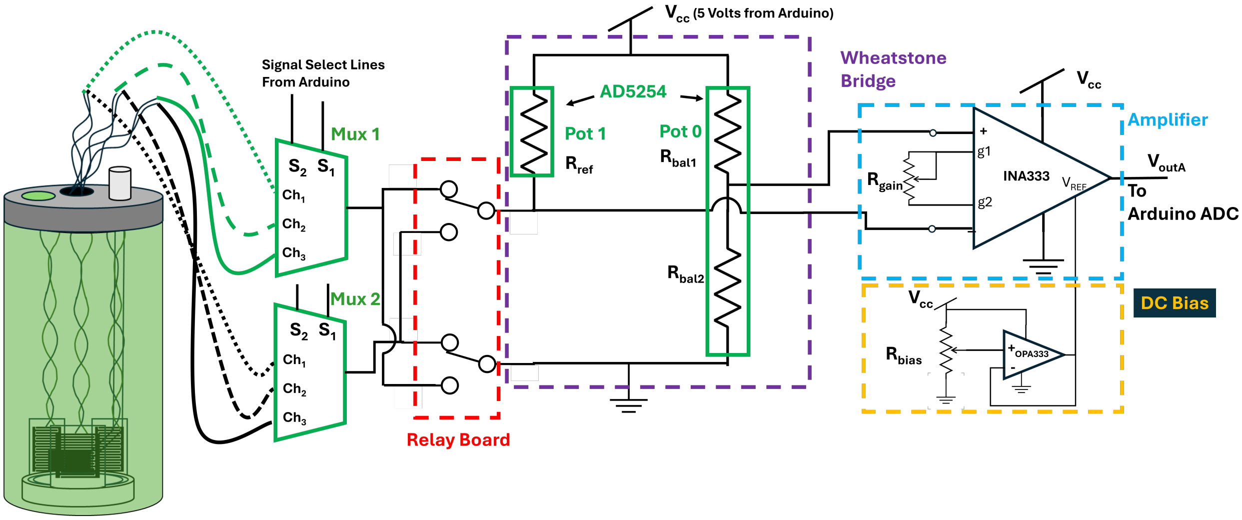


Figure 6.2 Circuit setup for the biofilm measurement experiment.

For each run, 3 sensors were submerged in the active culture using a circular 3D-printed base to hold them at the bottom of the PBR tank, as seen in Fig. 6.1-B. These were connected to the measurement system through two 16-channel multiplexers as seen in 6.2. The multiplexers were controlled simultaneously, as each sensor had one lead connected to each multiplexer on the same channel. This allowed each sensor to be measured individually by addressing the proper channel on the multiplexers (Channel 1 through 3). An SD card module was used as data storage for the system.

For the first round of testing, the polarity of the voltage across the electrodes was fixed throughout the test (each electrode remained connected to the same node in the measurement system), mirroring how the sensors were measured in previous sea lamprey detection experiments. However, as the system uses exposed electrodes with a DC potential in an electrolyte solution, there were concerns that degradation of the sensors due to electrochemical reactions at the electrode surfaces could affect the system output. To address this, a relay board was introduced into the system, highlighted in red in 6.2, to alternate which of the electrodes is connected to the ground node. This effectively alternates which electrode is at the highest electrical potential and could provide some protection from the electrochemical degradation.

### 6.1.2 Experimental Procedure

Before each test, the PBR tank was thoroughly cleaned to remove residue from previous cultures. The sensors were then introduced into the tank along with the stir bar before being connected to the multiplexers. Once the sensors were connected to the measurement system, the algal culture mixture was poured into the tank, and the magnetic stirrer was turned on. The system then followed a two-step self-calibration procedure, first adjusting $R_{ref}$ to achieve a specified baseline voltage across the IDE and then adjusting the proportion between $R_{bal1}$ and $R_{bal2}$ to match. This was done for each sensor individually, and the resulting values of $R_{ref}$ and $R_{bal2}$ were recorded for later comparison. After calibrating, the system begins measuring the sensors every 2 seconds (0.5 Hz sampling frequency). The system connects to each device sequentially to collect a measurement, and once done, it disconnects them from power until the next measurement time. For the second set of tests, the electrode connections are toggled after every measurement (including during calibration). Data are recorded to the SD card in a CSV format, and are also recorded through a serial terminal application for redundancy. Tests were run for approximately 2-5 days to ensure varying levels of biofilm growth on the sensors. The LED lighting rig was kept on at all times throughout each run. After each test, the system is allowed to recalibrate to record the change in $R_{ref}$ values, and an additional hour of data recording is captured to observe how the sensor baseline behaves once the devices have been biofouled. After this, the sensors were extracted from the PBR, and any biofilm that had accumulated on the leads or on the back of the sensing panel was removed to ensure only the biofilm in contact with the sensor traces remained. To quantify the amount of biofilm accumulated on the sensors during the experiment, each sensor was weighed 4 times using a milligram scale to generate an average dry weight measurement before the experiment and again 48 hours after the experiment was complete to allow for all of the moisture in the biofilm to evaporate. Examples of how the sensors looked immediately after the test and after 48 hours of drying are presented in Fig. 6.3-A and B, respectively.

After the monitoring tests were completed and the biofilm accumulation was quantified,

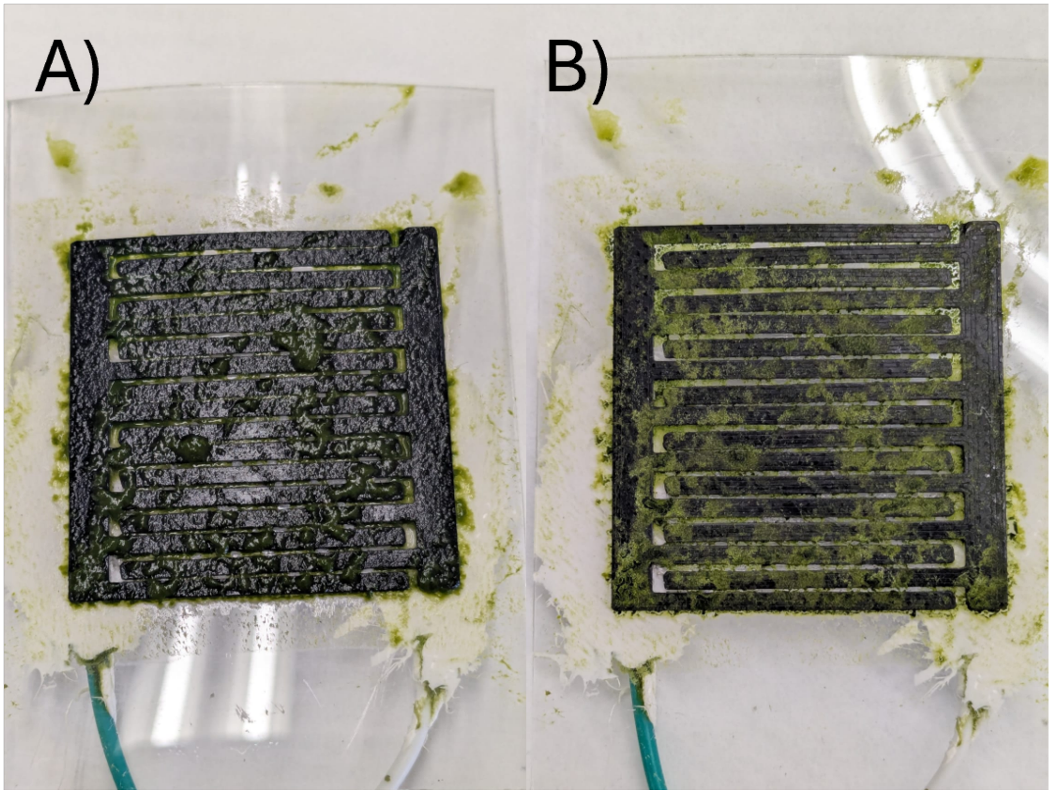


Figure 6.3 A) Representative sample of a sensor immediately after a biofilm accumulation test. B) The same sensor after a 48-hour drying period, ready for the post-test weight measurement.

several sensors were selected and subjected to an additional test to verify the effect of biofouling on the sensor response to biological contact. Three sensors at a time were introduced into a container with a diluted algae culture, the system was allowed to calibrate each sensor, after which baseline data was recorded for 10 minutes. Once this time had passed, the sensors were subjected to a thumb press that was sustained for 1-2 minutes, with a 1-minute wait period before interacting with the next sensor. This interaction was repeated at least twice during the test. After the interaction cycles were completed, the system was allowed to record another 10 minutes of baseline data. Voltage response data were recorded in the same fashion as the biofilm monitoring experiments.

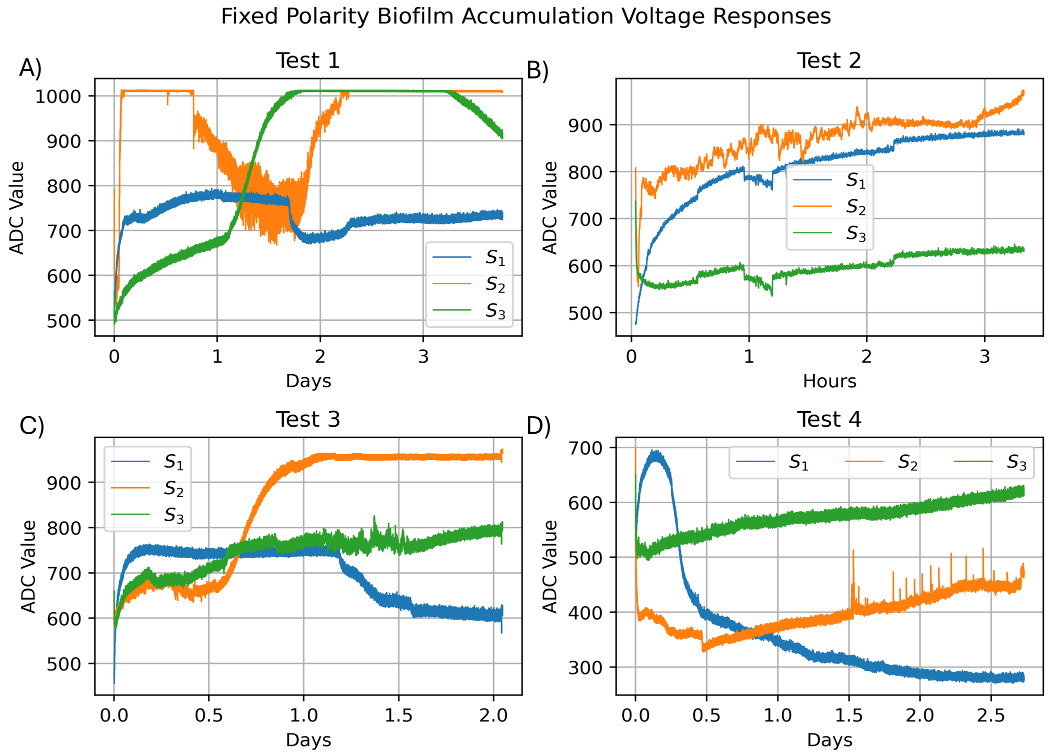


Figure 6.4 Sensor voltage response plots for the 4 tests performed with the polarity of the sensors remaining fixed: A) Fixed Polarity Test 1, B) Fixed Polarity Test 3, C) Fixed Polarity Test 3, and D) Fixed Polarity Test 4.

## 6.2 Sensor Response to Biofilm Accumulation

### 6.2.1 Fixed Polarity Tests

The sensors' voltage responses during the biofilm accumulation tests under fixed polarity conditions are presented in Fig. 6.4. Note that the sensor response varies considerably both during each test and across the set of tests. In Fig. 6.4 - A, which was the longest test in the set, the response of the first sensor ($S_1$) shows an initial increase in the output voltage corresponding to a decrease in the resistance of the sensor, a subsequent decay in the response corresponding to the opposite effect, followed finally by another period of increasing voltage response. The response of the second sensor ($S_2$) saturated very quickly and decayed while presenting a high-frequency noise component of greater amplitude than the other signals before saturating again. The third sensor ($S_3$) response for this test also saturated after approximately 2 days, but did not present the same level of high-frequency

Table 6.1 Resistance Measurements and Biomass Accumulation Results for Fixed Polarity Tests

| Test - Sensor | Initial $R_{ref}$ | Final $R_{ref}$ | $R_{\Delta}$ | Biofilm Accumulation (mg) |
|---|---|---|---|---|
| Test 1 - Sensor 1 | 3906.25 | 1757.81 | -2148.44 | 5 |
| Test 1 - Sensor 2 | 1953.13 | 1562.50 | -390.63 | 9 |
| Test 1 - Sensor 3 | 6601.56 | 2617.19 | -3984.38 | 8 |
| Test 2 - Sensor 1 | 7382.81 | 1171.89 | -6210.94 | 11 |
| Test 2 - Sensor 2 | 3867.19 | 1523.44 | -2343.75 | 0 |
| Test 2 - Sensor 3 | 6562.50 | 2187.50 | -4375.00 | 5 |
| Test 3 - Sensor 1 | 5468.75 | 3828.13 | -1640.63 | 1 |
| Test 3 - Sensor 2 | 5390.63 | 1406.25 | -3984.38 | 9 |
| Test 3 - Sensor 3 | 5546.88 | 3515.63 | -2031.25 | 7 |
| Test 4 - Sensor 1 | 3046.88 | 5781.25 | 2734.38 | 69 |
| Test 4 - Sensor 2 | 3789.06 | 4492.19 | 703.13 | 19 |
| Test 4 - Sensor 3 | 2773.44 | 1992.19 | -781.25 | 44 |

noise components when it decayed back into the ADC range at the beginning of day 3. In Fig. 6.4 - B, which ended up being the shortest test in the set due to a power interruption, all three sensor responses were generally increasing during the duration of the test. There is evidence of crosstalk between $S_1$ and $S_3$ as their responses share sudden transients not observed in $S_2$ at several points in the recording. The third test, taking place over 2 days, is presented in Fig. 6.4 - C. The sensor responses for this test present a combination of the behaviors described for the previous tests; $S_1$ very quickly increases to a stable baseline before slowly decaying after approximately 1.25 days, $S_2$ saturates after 1 day and remains at the saturation point for the remainder of the test and $S_3$ presents a slow increase throughout the length of the test. The final test, presented in Fig. 6.4 - C, was stopped just short of a 3-day duration. During this test, $S_1$ demonstrated another type of response with an initial increase that peaks before entering what appears to be an exponential decay, while $S_2$ and $S_3$ both showed an initial decrease followed by a steady increase for the remainder of the test. The amount of biomass accumulated on each sensor for each test is presented in Table 6.1 together with the initial and final values for $R_{ref}$.

The amount of biomass accumulated on the sensors during these tests was in the order of several milligrams (average value of 15.5 mg), with only the sensors in the final test consistently having more than 10 milligrams of biomass accumulated on their surface; if we don't consider the values for the fourth test the average drops to 6.1 mg. This discrepancy is likely due to differences in the way the sensors were mounted for the final test. Tests 1 through 3 had the sensing surface pointed toward the center of the PBR tank, resulting in considerable biofilm on the back of the devices and not on the sensor traces; for the final fixed polarity test and all subsequent tests, the sensor were placed facing the perimeter of the tank, leading to increased biofilm formation directly on the traces. Except for the final test, the sensors generally calibrated to a lower value of $R_{ref}$ after the tests were completed, indicating a decrease in the sensor resistance. The variety of responses demonstrated during these tests, along with small amounts of biofilm accumulation achieved, suggests it will be difficult to draw a link between biofilm accumulation and the voltage response of the sensors for these tests.

### 6.2.2 Alternating Polarity Tests

The sensors' voltage response during the biofilm accumulation tests under alternating polarity conditions are presented in Fig. 6.5. It is immediately clear that the voltage responses across this set have greater uniformity than those in the fixed polarity test set. The results from the first test of this set, presented in Fig. 6.5-A, show the general response shapes observed throughout the set. The response for sensor 1 ($S_1$) shows a short increase which peaks quickly, followed by a decay that eventually levels off and proceeds into a slow increase in the baseline. This response pattern will be referred to as Type 1-A from here on out. $S_2$ appears to have two distinct phases in its response, the first following the increasing form of an exponential decay function and settling to a stable baseline for around 1.5 days, followed by a decay period. This response pattern will be referred to as Type 1-B for the remainder of this work. In contrast, $S_3$ follows the increasing form exponential decay for the duration of the test, which will hereby be referred to as Type 2. The next test, the results of which are presented in Fig. 6.5-B, mirrors these response shapes with $S_1$ and $S_3$ following the Type 1-B pattern and $S_2$ following the Type 2 pattern. Similarly, the results for test 3, which are

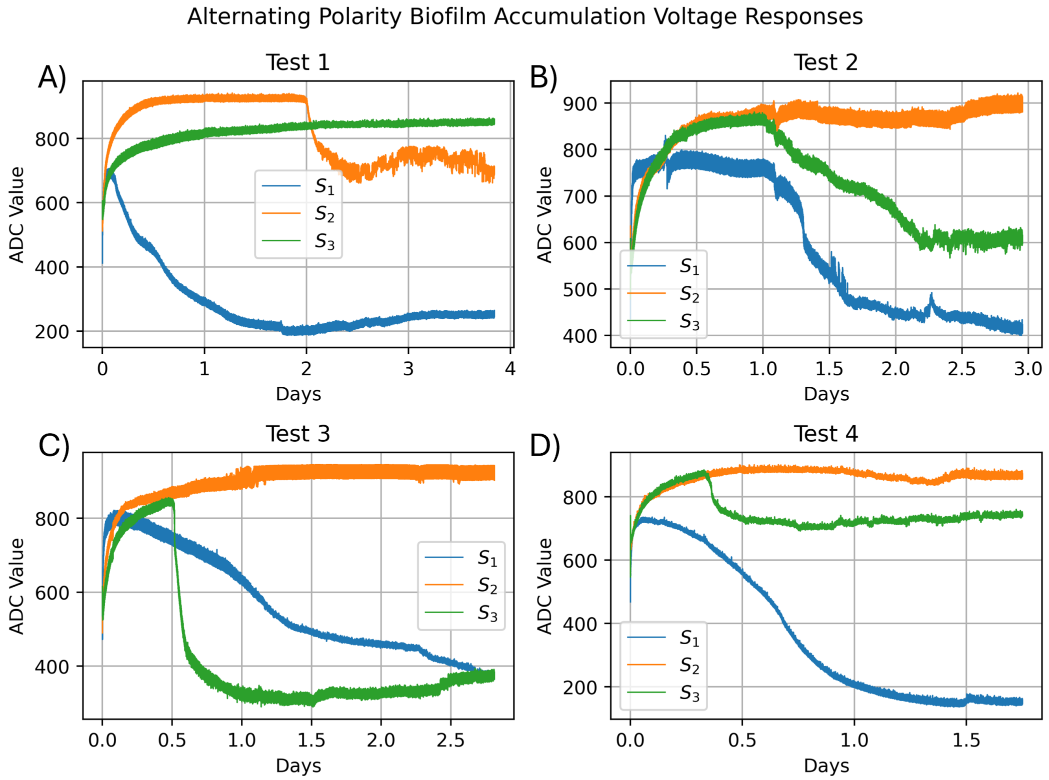


Figure 6.5 Sensor voltage response plots for the 4 tests performed while switching the polarity of the sensors: A) Alternating Polarity Test 1, B) Alternating Polarity Test 3, C) Alternating Polarity Test 3 and D) Alternating Polarity Test 4.

presented in Fig. 6.5-C, continue this pattern with $S_1$ following the Type 1-A response, $S_2$ following the Type 2 response, and $S_3$ following the Type 1-B response. Finally, the results of the last test in the alternating polarity set, presented in Fig. 6.5-D, replicate the pattern seen in test 3 with the exception of $S_3$ having a less pronounced decay. The amount of biomass accumulated on each sensor for each test is presented in Table 6.2 together with the initial and final values for $R_{ref}$.

Greater levels of biofilm accumulation were achieved for the alternating polarity test set than for the fixed polarity set due to the change in sensor mounting position, with an average biomass accumulation of 24.9 mg versus an average of 15.5 mg for the fixed test set. Additionally, unlike the fixed polarity test set, there is an apparent pattern between the voltage response types and whether the sensors calibrate to higher or lower values of $R_{ref}$. The grouping of the sensor voltage responses by whether they correspond to an increase or

Table 6.2 Resistance Measurements and Biomass Accumulation Results for Alternating Polarity Tests

| Test - Sensor | Initial $R_{ref}$ | Final $R_{ref}$ | $R_{\Delta}$ | Biofilm Accumulation (mg) |
|---|---|---|---|---|
| Test 1 - Sensor 1 | 17968.75 | 39062.5 | 21093.75 | 29 |
| Test 1 - Sensor 2 | 29296.875 | 31640.625 | 2343.75 | 20 |
| Test 1 - Sensor 3 | 28906.25 | 14062.5 | -14843.75 | 46 |
| Test 2 - Sensor 1 | 42968.75 | 43359.375 | 390.625 | 6 |
| Test 2 - Sensor 2 | 50000 | 23046.875 | -26953.125 | 35 |
| Test 2 - Sensor 3 | 34375.0 | 46484.375 | 12109.375 | 9 |
| Test 3 - Sensor 1 | 36718.75 | 53906.25 | 17187.5 | 24 |
| Test 3 - Sensor 2 | 44921.875 | 10937.5 | -33984.375 | 29 |
| Test 3 - Sensor 3 | 35546.875 | 59375.0 | 23828.125 | 49 |
| Test 4 - Sensor 1 | 19531.25 | 51953.125 | 32421.875 | 22 |
| Test 4 - Sensor 2 | 33203.125 | 15234.375 | -17968.75 | 38 |
| Test 4 - Sensor 3 | 35937.5 | 29296.875 | -6640.625 | 13 |

decrease $R_{ref}$ and is presented in Fig. 6.6-A and B, respectively.

The response types most commonly associated with increases in resistance are those described by Type 1-A and Type 1-B responses as seen in 6.6-A, while the Type 2 response shape was generally associated with a decrease in $R_{ref}$. The only observed exception to this trend is the $S_3$ response from test 4, which presents a Type 1-B response and the smallest negative change in $R_{ref}$ for the set. This hints at a complex relationship between biomass accumulation, the voltage response, and the change in resistance across the sensors.

## 6.3 Resistance Change versus Accumulated Biomass Study

Given the variety of voltage responses observed for similar amounts of biofouling, it may be useful to study the relationship between the net resistance change occurring in the sensors and the amount of biomass accumulated on them. For the fixed polarity test set, biofilm accumulation as a function of the change in resistance is shown in Fig. 6.7 - A. While there appears to be a weak positive correlation between the change in resistance and biofilm accumulation for this set, with a correlation factor of 0.489, this result is skewed due to the results from the final test having considerably more biomass than those from the first 3 tests. If these values are excluded, the correlation factor drops to 0.210, and the resulting linear

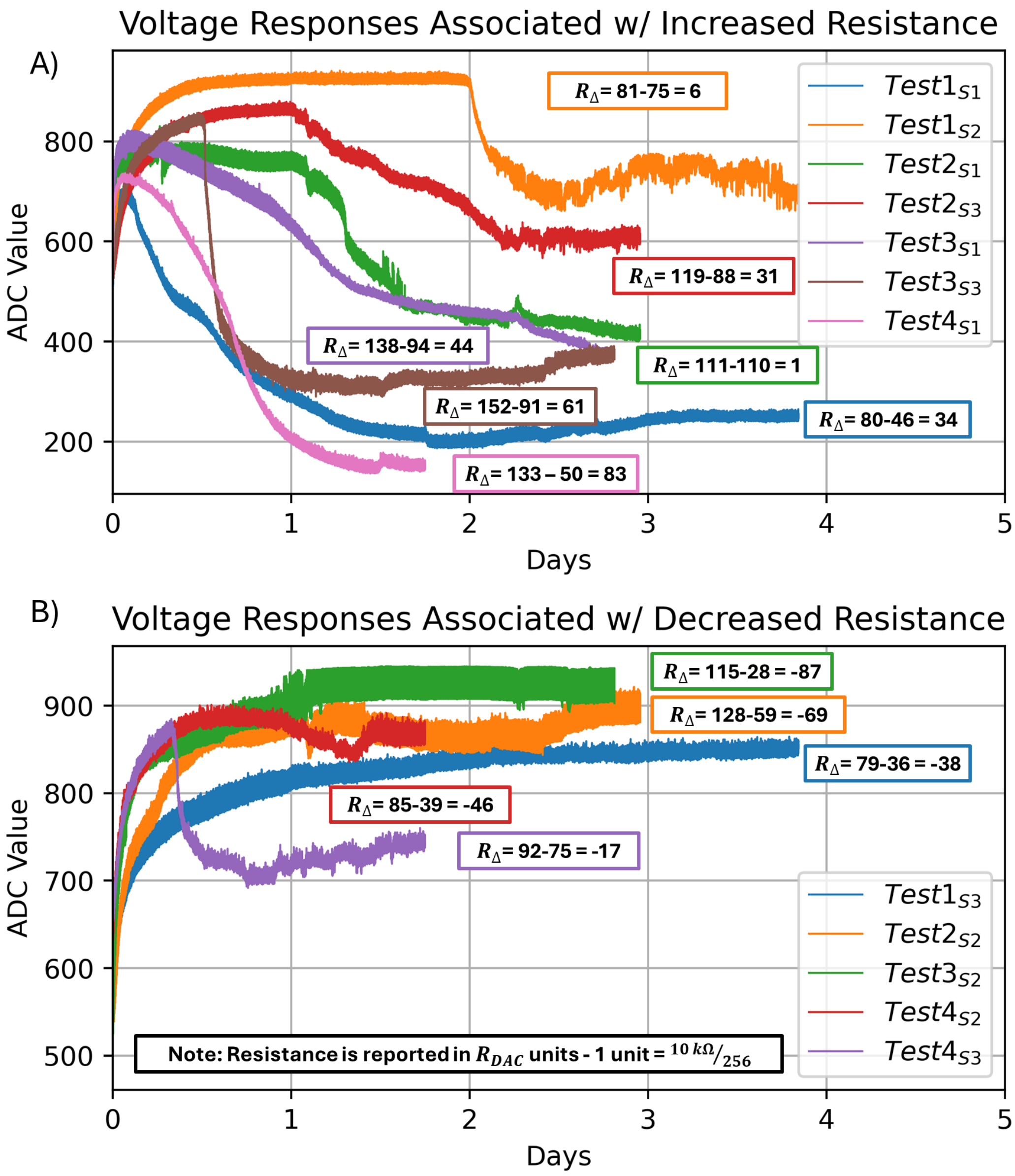


Figure 6.6 A) Combination plot including all the sensor voltage responses for the alternating polarity test set corresponding to an increase in $R_{ref}$. B) Combination plot including all the sensor voltage responses for the alternating polarity test set corresponding to a decrease in $R_{ref}$.

relationship approximates a flat line, indicating little to no relationship between the test variables. Similarly, for the results of the study using the data from the alternating polarity test, shown in Fig. 6.7-B, the linear correlation coefficient is low ($R^2 = 0.151$), indicating a poor relationship between biomass accumulation and change in resistance for this test set. The data for this test set does appear to be grouped into two clusters corresponding to the

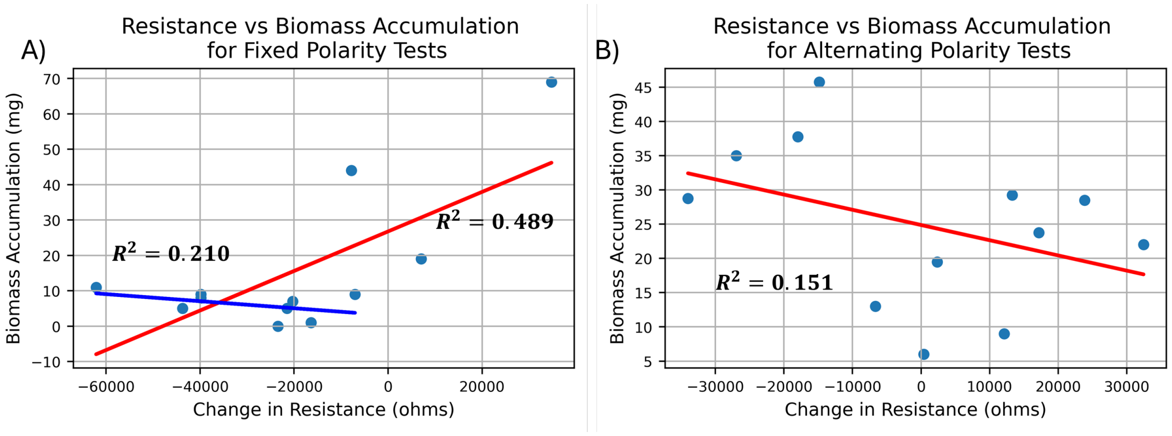


Figure 6.7 A) Change in $R_{ref}$ plotted against the total amount of biomass accumulated on the sensors for the fixed polarity test set. B) Change in $R_{ref}$ plotted against the total amount of biomass accumulated on the sensors for the alternating polarity test set.

response shapes with the Type 1-A and 1-B responses from approximately -700 to 3000 ohms and the Type 2 responses from -3500 to -1500 ohms. However, the small sample size makes it difficult to generalize a relationship from these two clusters.

## 6.4 Post-fouling Biological Contact

The sensors presenting the highest changes in resistance for the alternating polarity test set (increased resistance: Test 3 $S_1$, Test 3 $S_3$, Test 4 $S_1$, decreased resistance: Test 2 $S_2$, Test 3 $S_2$, Test 4 $S_2$) were selected for an additional test to observe how the biofouling would affect their response biological contact. The results from these experiments for the sensors with increased resistance are presented in Fig. 6.8 - A and B. In these plots $S_1$ corresponds to Test 3's $S_1$ ($R_\Delta = 44 R_{DAC}$ units), $S_2$ to Test 4's $S_1$ ($R_\Delta = 83 R_{DAC}$ units) and $S_3$ to Test 3 $S_3$ ($R_\Delta = 61 R_{DAC}$ units). In the case of $S_1$, for both tests, the biological contact is registered in the response as a sudden peak followed by decay with the baseline trending downwards as the test progressed. During both tests, the response from $S_2$ is markedly noisier than the other two sensors, making it harder to distinguish the contact response. This suggests more aggressive filtering may be required as the change in resistance increases. The contact response is also inconsistent for this sensor, presenting as an increase in the response in Contact Test 1 and a decrease in Contact Test 2. On the other hand, $S_3$ presents a sharp increase in the sensor output, which holds through the entirety of the contact. There is a

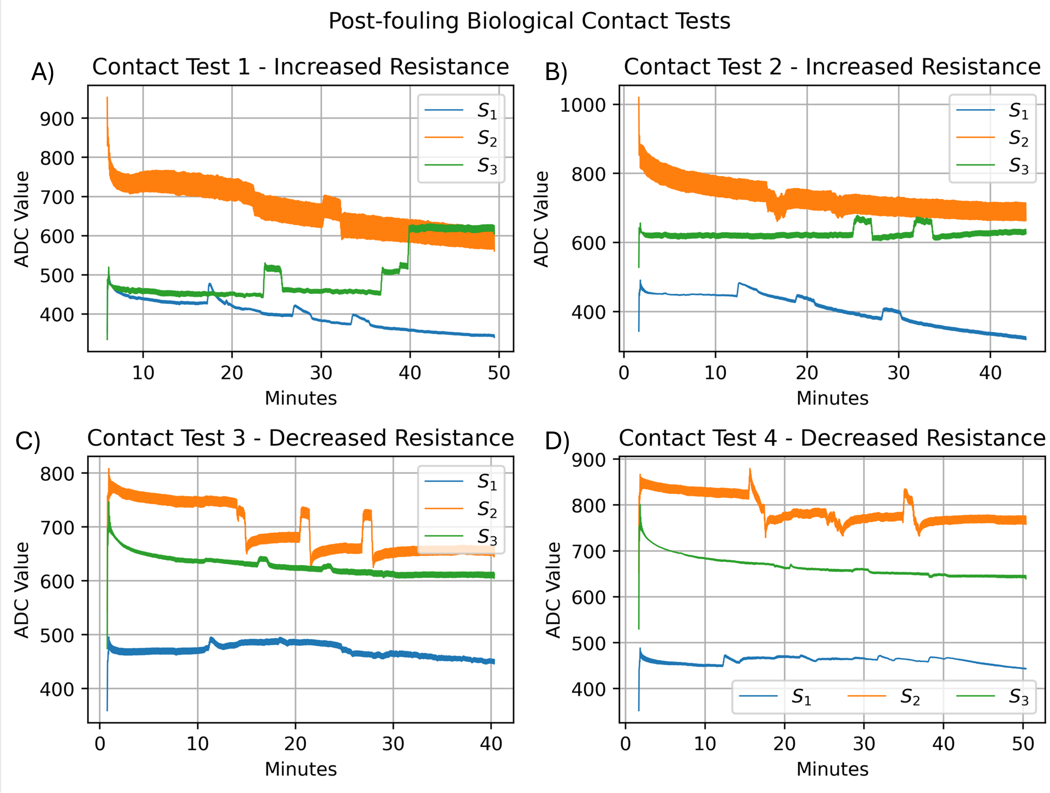


Figure 6.8 A) Sensor responses in the initial biological contact test using the sensors with the highest increased resistance. B) Sensor responses in follow-up biological contact test using the sensors with the highest increased resistance. C) Sensor responses in the initial biological contact test using the sensors with the highest decrease in resistance. D) Sensor responses in follow-up biological contact test using the sensors with the highest decrease in resistance.

sudden transient in $S_3$ towards the end of the first test, which can be explained by established biofilm breaking off as the touch is released.

The results from the experiments using the sensors with decreased resistance are presented in Fig. 6.8 - C and D. In these plots $S_1$ corresponds to Test 2's $S_2$ ($R_\Delta = -69R_{DAC}$ units), $S_2$ to Test 4's $S_2$ ($R_\Delta = -46R_{DAC}$ units) and $S_3$ to Test 3 $S_2$ ($R_\Delta = -89R_{DAC}$ units). The response of the first sensor in these tests, $S_1$, presents a similar response to biological contact as the first sensor in Contact Test 1 and 2, but with a smaller response magnitude. $Contact S_2$ shows the clearest response to biological contact out of the three sensors with decreased resistance, but its response also suffered from inconsistency between Contact Test 3 and 4. Similar to $Contact S_1$, the response of $Contact S_3$ to biological contact presents a small magnitude response to biological contact in Contact Test 3, but it becomes almost imperceptible in Contact Test 4.

## 6.5 Discussion

The targeted deployment environment of the automated sea lamprey sensing system in the tributaries of the Great Lakes will be conducive to the accumulation of algal and bacterial biofilm on the devices. As was shown in this study, environmental conditions that favor biofilm growth will have a clear effect on the baseline of the sensors that will need to be addressed before field deployment is viable. This effect has proven difficult to characterize due to the wide range of responses observed throughout both test sets performed in this work. One potential explanation is the presence of a non-biological contribution to the change in resistance across the sensor. The algal growth medium is much more conductive than the water the sensors are typically tested in ( 4000 $S/cm$ vs 180-230 $S/cm$), which makes it easier for electrochemical reactions to occur on the electrode surfaces. These reactions have the potential to alter the electrode chemistry, depleting the electrode that is held at a positive potential (anode) of its conductive carbon through oxidation. Indeed, for some of the sensors in the fixed polarity test set had one of their electrodes rendered non-conductive, effectively disabling the device. At the same time, it is important to note that this is at odds with the pattern observed for the change in calibrated resistance after the fixed polarity tests, which generally presented a reduction in $R_{ref}$ that should correspond to a decrease in the sensor resistance as well. This decrease should not be possible if one of the sensing electrodes was completely depleted. However, the sensors are slightly porous, which may allow for the highly conductive medium to penetrate the sensors, establishing an electrical connection with the electrical lead that results in a lower resistance. A potential indicator that alternating the polarity of the sensors is protective against electrochemical depletion is given by the fact that while some sensors in the alternating test set also suffered degradation, it was not as extreme as with the sensor tests under fixed polarity.

Even with a reduction in the effects of electrochemical reactions on the device resistance in the alternating polarity tests, the sensor responses for these tests did not present a consistent pattern with regard to algal biofilm accumulation. The composition of the biofilm may contribute to this observed behavior as they have two major components, microalgal cells and extracellular polymeric substances (EPS), which contribute to the physical and electrical properties of the biofilm. While microalgae cells are insulators and should therefore contribute to an increase in resistance as they accumulate on the device, EPS has a wide range of components: proteins, phospholipids, polysaccharides, nucleic acids, humic substances, uronic acids, and some functional groups (e.g. phosphoric, carboxylic, hydroxyl and amino groups) and they can have varying contributions to the impedance of the biofilm. Additionally, where the biofilm accumulates on the sensors may also play a role in the impedance contribution. Deposition directly on the sensor traces is expected to increase the resistance observed by the measurement system, as the insulating components of the biofilm will restrict current flow to the medium, while deposition in the gap between electrodes would likely decrease the resistance if the EPS provides a high conductivity channel between the two electrodes. This indicates that to accurately characterize the sensor response to biofilm attachment, some type of visual tracking method would have to be implemented alongside the dry weight measurement approach.

### 6.5.1 Implications for Field Deployment in Sea Lamprey Detection Application

Looking at the response of the fouled sensors to biological contact, it is clear that biofilm buildup and the changes in sensor resistance observed after the tests affect their response. While the sensors still present a reaction to biological contact when the resistance increases after the tests, it is less consistent than what is expected of the devices, which will require changes in the detection algorithms used alongside this system. In the case of the sensors that presented decreased resistance after the tests, the magnitude of the response was generally smaller, making it harder to distinguish from baseline. This could be addressed by increasing the amplification factor of the measurement system, which was at its lowest value for all of the tests presented in this chapter, in response to changes in $R_{ref}$. Another way to address these impacts is to have the next generation of detection algorithms include responses from deteriorated sensors in their training data, and include the value of $R_{ref}$ as an additional input to the algorithms.

Moving forward, all sea lamprey sensing systems based on the IDE design should adopt the alternating polarity approach to reduce the impact of electrochemical reactions on the devices. Another thing to note is that these experiments were run over a longer timeframe than previous sea lamprey detection experiments, but the system was only calibrated at the beginning and end of each test. Electrochemical reactions are driven by a potential difference, and since the voltage across the sensors is a function of $R_{ref}$ it is possible that by frequently calibrating the system, the effects of these reactions could be further reduced. The rapid change in voltage at the beginning of most tests would indicate that the system should be recalibrated within hours of the previous calibration. Additionally, given the impact that the change in resistance has on the sensor's ability to detect biological contact, $R_{ref}$ should be periodically logged and reported to serve as a measure for determining whether the sensor should be retired. It is also possible to reduce the potential difference between electrodes by altering the bridge circuit used to measure the devices by introducing additional resistors in series before and after the bridge.

### 6.6 Future Work

The complexity of the impedance contribution of the biofilm, along with its geometric dependency, will require a direct impedance measurement with an AC probing signal alongside a visual tracking component to be properly characterized as a function of sensing area coverage. Single frequency impedance analysis has been used successfully before to characterize biofilms [61][62], and suggests that relatively low-frequency signals could be used for this purpose. This measurement can be performed without the use of expensive equipment by using a direct digital synthesis module to generate a sine wave at a particular frequency and using a current sensing amplifier to measure the current across a sensing resistor, but would require a study to be performed to find the optimal probing frequency.

In place of a comprehensive characterization of the system's response to biofilm accumulation, another potential direction of future research is to find ways to reduce the influence of biofilm growth conditions on the sensor baseline. This could be done using the same setup as the biofilm accumulation and following the suggestions laid out in the previous section. For this potential study, and future field deployments, it would be useful to include additional environmental sensors alongside the system to track variables that could also affect the sensor response, such as temperature, pH, and water conductivity. Finally, the response of the sensor to "wild" biofilm accumulation in field deployment conditions should also be studied, as it is possible that the highly conductive environment of the PBR could be a greater contributor to the observed responses than the biofilm accumulation itself.

# CHAPTER 7

# SUMMARY

## 7.1 Summary of Contributions

This work presents the development of an automated sea lamprey detection system that could become a powerful tool for monitoring these invasive fish. The system was built around low-cost conductive carbon electrodes to facilitate deployment at scale. The methodologies used to design, fabricate, and validate the sensors are one of the key takeaways from this work, which could be applied to similar underwater sensing systems. Similarly, the development of measurement and signal conditioning circuits used as a part of this system also have potential applications in the measurement of other resistive sensors. The integration of machine learning models with the system, which is operated by a simple microcontroller unit, to enable automatic real-time sea lamprey detection demonstrates a novel application of embedded machine learning. This effort also resulted in a guide for the deployment of simple machine learning models to embedded systems. This work took the first steps towards characterizing the sensor response to biofilm accumulation, which has the potential to enable monitoring applications in a range of fields.

### 7.1.1 Problems Solved in this Thesis

This work addresses the following:

- The interdigitated electrode structure was shown to be viable for underwater contact sensing using relative conductivity differences as a driver for contact source characterization.
- The response of the sensor to sea lamprey attachment using a simple voltage divider measurement circuit was characterized
- The correlation between characterized response parameters and sea lamprey size parameters was explored

- Machine learning models were trained on sea lamprey response data and deployed to a resource-constrained device to enable automatic detection of sea lamprey
- Shortcomings in the initial measurement system were addressed through an updated measurement approach
- The effect of biofilm accumulation on the sensors was studied, and strategies to deal with this impact were proposed

# BIBLIOGRAPHY

[1] F. Guy, "Sea lamprey wounds (Petromyzon marinus) on a Chinook salmon (Oncorhynchus tshawytscha), St. Mary's River-North Channel, Garden River First Nation."

[2] "CC BY-SA 3.0 Deed | Attribution-ShareAlike 3.0 Unported | Creative Commons."

[3] C. B. Renaud and P. A. Cochran, "Post-metamorphic Feeding in Lampreys," in *Lampreys: Biology, Conservation and Control: Volume 2* (M. F. Docker, ed.), Fish & Fisheries Series, pp. 247–285, Dordrecht: Springer Netherlands, 2019.

[4] R. L. Eshenroder, "The Role of the Champlain Canal and Erie Canal as Putative Corridors for Colonization of Lake Champlain and Lake Ontario by Sea Lampreys," *Transactions of the American Fisheries Society*, vol. 143, pp. 634–649, May 2014. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00028487.2013.879818.

[5] V. C. Applegate, "Natural history of the sea lamprey, Petromyzon marinus, in Michigan," Report 55, 1950.

[6] B. R. Smith and J. J. Tibbles, "Sea Lamprey ( *Petromyzon marinus* ) in Lakes Huron, Michigan, and Superior: History of Invasion and Control, 1936–78," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 37, pp. 1780–1801, Nov. 1980.

[7] W. J. Christie, "Changes in the Fish Species Composition of the Great Lakes," *Journal of the Fisheries Board of Canada*, May 1974. Publisher: NRC Research Press Ottawa, Canada.

[8] A. Muir, C. Krueger, and M. Hansen, "Re-establishing lake trout in the Laurentian Great Lakes: past, present, and future," in *Great Lakes fisheries policy and management: a binational perspective*, pp. 533–588, Jan. 2013. Journal Abbreviation: Great Lakes fisheries policy and management: a binational perspective.

[9] P. J. Hrodey, S. A. Lewandoski, W. P. Sullivan, J. M. Barber, K. A. Mann, B. Paudel, and M. J. Symbal, "Evolution of the sea lamprey control barrier program: The importance of lowermost barriers," *Journal of Great Lakes Research*, vol. 47, pp. S285–S296, Dec. 2021.

[10] N. S. Johnson, H. T. Thompson, C. Holbrook, and J. A. Tix, "Blocking and guiding adult sea lamprey with pulsed direct current from vertical electrodes," *Fisheries Research*, vol. 150, pp. 38–48, Feb. 2014.

[11] N. S. Johnson, B. Snow, T. Bruning, and A. Jubar, “A seasonal electric barrier blocks invasive adult sea lamprey (Petromyzon marinus) and reduces production of larvae,” *Journal of Great Lakes Research*, vol. 47, pp. S310–S319, Dec. 2021.

[12] W. P. Sullivan, D. P. Burkett, M. A. Boogaard, L. A. Criger, C. E. Freiburger, T. D. Hubert, K. G. Leistner, B. J. Morrison, S. M. Nowicki, S. N. P. Robertson, A. K. Rowlinson, B. J. Scotland, and T. B. Sullivan, “Advances in the use of lampricides to control sea lampreys in the Laurentian Great Lakes, 2000–2019,” *Journal of Great Lakes Research*, vol. 47, pp. S216–S237, Dec. 2021.

[13] R. A. Bergstedt and M. B. Twohey, “Research to Support Sterile-male-release and Genetic Alteration Techniques for Sea Lamprey Control,” *Journal of Great Lakes Research*, vol. 33, pp. 48–69, Jan. 2007.

[14] “Great Lakes Fishery Commission - The Fishery.”

[15] J. V. Adams, J. M. Barber, G. A. Bravener, and S. A. Lewandoski, “Quantifying Great Lakes sea lamprey populations using an index of adults,” *Journal of Great Lakes Research*, vol. 47, pp. S335–S346, Dec. 2021.

[16] S. A. Lewandoski and T. O. Brenden, “Forecasting suppression of invasive sea lamprey in Lake Superior,” *Journal of Applied Ecology*, vol. 59, no. 8, pp. 2023–2035, 2022. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2664.14203.

[17] N. Schloesser, “Correlating sea lamprey density with environmental DNA detections in the lab,” *Management of Biological Invasions*, vol. 9, no. 4, pp. 483–495, 2018.

[18] T. D. Gingera, T. B. Steeves, D. A. Boguski, S. Whyard, W. Li, and M. F. Docker, “Detection and identification of lampreys in Great Lakes streams using environmental DNA,” *Journal of Great Lakes Research*, vol. 42, pp. 649–659, June 2016.

[19] E. L. McCann, N. S. Johnson, P. J. Hrodey, and K. L. Pangle, “Characterization of Sea Lamprey Stream Entry Using Dual-Frequency Identification Sonar,” *Transactions of the American Fisheries Society*, vol. 147, no. 3, pp. 514–524, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tafs.10052.

[20] G. D. Bixler and B. Bhushan, “Biofouling: lessons from nature,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 370, pp. 2381–2417, May 2012. Publisher: Royal Society.

[21] J. L. Harding and M. M. Reynolds, “Combating medical device fouling,” *Trends in Biotechnology*, vol. 32, pp. 140–146, Mar. 2014.

[22] F. Di Pippo, L. Di Gregorio, R. Congestri, V. Tandoi, and S. Rossetti, “Biofilm growth and control in cooling water industrial systems,” *FEMS Microbiology Ecology*, vol. 94, May 2018.

[23] O. Zeriouh, J. V. Reinoso-Moreno, L. López-Rosales, M. d. C. Cerón-García, A. Sánchez-Mirón, F. García-Camacho, and E. Molina-Grima, “Biofouling in photobioreactors for marine microalgae,” *Critical Reviews in Biotechnology*, vol. 37, pp. 1006–1023, Nov. 2017. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07388551.2017.1299681.

[24] A. M. Curtin and H. L. Buckley, “Biofouling detection methods that are widely applicable and useful across disciplines: a mini-review,” *Biofouling*, vol. 37, pp. 494–505, May 2021. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08927014.2021.1926998.

[25] Q. Xie, J. Pan, C. Ma, and G. Zhang, “Dynamic surface antifouling: mechanism and systems,” *Soft matter*, vol. 15, pp. 1087–1107, Feb. 2019. Place: England Publisher: Royal Society of Chemistry.

[26] I. González-Afanador, H. Shi, C. Holbrook, X. Tan, and N. Sepúlveda, “Invasive Sea Lamprey Detection and Characterization Using Interdigitated Electrode (IDE) Contact Sensor,” *IEEE Sensors Journal*, vol. 21, pp. 27947–27956, Dec. 2021. Conference Name: IEEE Sensors Journal.

[27] I. González-Afanador, C. Chen, G. Morales-Torres, S. Meihls, H. Shi, X. Tan, and N. Sepúlveda, “Real-time invasive sea lamprey detection using machine learning classifier models on embedded systems,” *Neural Computing and Applications*, vol. 36, pp. 16195–16212, Sept. 2024.

[28] H. Shi, M. Al-Rubaiai, C. M. Holbrook, J. Miao, T. Pinto, C. Wang, and X. Tan, “Screen-Printed Soft Capacitive Sensors for Spatial Mapping of Both Positive and Negative Pressures,” *Advanced Functional Materials*, vol. 29, no. 23, p. 1809116, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.201809116.

[29] H. Shi, I. González-Afanador, C. Holbrook, N. Sepúlveda, and X. Tan, “Soft Pressure Sensor for Underwater Sea Lamprey Detection,” *IEEE Sensors Journal*, vol. 22, pp. 9932–9944, May 2022. Conference Name: IEEE Sensors Journal.

[30] Y. Cao, H. Shi, X. Tan, and N. Sepúlveda, “Enabling Negative Pressure Sensing Through Ferroelectret Device,” *IEEE Sensors Letters*, vol. 6, pp. 1–4, Aug. 2022. Conference Name: IEEE Sensors Letters.

[31] Y. Duan, S. He, J. Wu, B. Su, and Y. Wang, “Recent Progress in Flexible Pressure Sensor Arrays,” *Nanomaterials*, vol. 12, p. 2495, July 2022.

[32] H. Shi, C. M. Holbrook, Y. Cao, N. Sepúlveda, and X. Tan, “Measurement of suction pressure dynamics of sea lampreys, Petromyzon marinus,” *PLOS ONE*, vol. 16, p. e0247884, Apr. 2021. Publisher: Public Library of Science.

[33] X. Tan, “E-Skin Technology Could Aid in Sea Lamprey Fight,” Jan. 2020.

[34] C. Chen, H. Shi, I. González-Afanador, N. Sepúlveda, and X. Tan, “Rapid Fabrication of Flexible Pressure Sensor Array,” *IEEE Sensors Letters*, vol. 7, pp. 1–4, June 2023. Conference Name: IEEE Sensors Letters.

[35] H. Shi, Y. Mei, I. González-Afanador, C. Chen, S. Miehls, C. Holbrook, N. Sepúlveda, and X. Tan, “Automated Soft Pressure Sensor Array-Based Sea Lamprey Detection Using Machine Learning,” *IEEE Sensors Journal*, vol. 23, pp. 7546–7557, Apr. 2023. Conference Name: IEEE Sensors Journal.

[36] J.-G. Guan, Y.-Q. Miao, and Q.-J. Zhang, “Impedimetric biosensors,” *Journal of Bioscience and Bioengineering*, vol. 97, no. 4, pp. 219–226, 2004.

[37] N. Hikmah, H. F. Hawari, and M. Gupta, “Design and simulation of interdigitated electrode for Graphene-SnO2 sensor on acetone gas,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, pp. 119–125, July 2020. Number: 1.

[38] L. Wang, M. Veselinovic, L. Yang, B. J. Geiss, D. S. Dandy, and T. Chen, “A sensitive DNA capacitive biosensor using interdigitated electrodes,” *Biosensors & Bioelectronics*, vol. 87, pp. 646–653, Jan. 2017.

[39] J. Wang, C. Wu, N. Hu, J. Zhou, L. Du, and P. Wang, “Microfabricated Electrochemical Cell-Based Biosensors for Analysis of Living Cells In Vitro,” *Biosensors*, vol. 2, pp. 127–170, June 2012. Number: 2 Publisher: Molecular Diversity Preservation International.

[40] P. Van Gerwen, W. Laureyn, W. Laureys, G. Huyberechts, M. Op De Beeck, K. Baert, J. Suls, W. Sansen, P. Jacobs, L. Hermans, and R. Mertens, “Nanoscaled interdigitated electrode arrays for biochemical sensors,” *Sensors and Actuators B: Chemical*, vol. 49, pp. 73–80, June 1998.

[41] J. S. Daniels and N. Pourmand, “Label-Free Impedance Biosensors: Opportunities and Challenges,” *Electroanalysis*, vol. 19, no. 12, pp. 1239–1257, 2007. _eprint: https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/pdf/10.1002/elan.200603855.

[42] C. Gaul, "Lampetra appendix (American brook lamprey)."

[43] B. Karvel-Fuller, "Ichthyomyzon fossor (Northern brook lamprey)."

[44] T. Acciaioli, "Ichthyomyzon castaneus (Chestnut lamprey)."

[45] C. Blumbergs, "Ichthyomyzon unicuspis (Silver lamprey)."

[46] A. Dizon and M. E. Orazem, "On the impedance response of interdigitated electrodes," *Electrochimica Acta*, vol. 327, p. 135000, Dec. 2019.

[47] S. Cherry, "Petromyzon marinus (Eel sucker)."

[48] N. Ha, K. Xu, G. Ren, A. Mitchell, and J. Z. Ou, "Machine Learning-Enabled Smart Sensor Systems," *Advanced Intelligent Systems*, vol. 2, no. 9, p. 2000063, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aisy.202000063.

[49] H. Keshavarz, M. S. Abadeh, and R. Rawassizadeh, "SEFR: A Fast Linear-Time Classifier for Ultra-Low Power Devices," Nov. 2020. arXiv:2006.04620 [cs, stat].

[50] V. Hurbungs, V. Bassoo, and T. P. Fowdur, "An enhanced binary classifier for Edge devices," *Microprocessors and Microsystems*, vol. 93, p. 104596, Sept. 2022.

[51] F. Shabani, H. Philamore, and F. Matsuno, "An Energy-Autonomous Chemical Oxygen Demand Sensor Using a Microbial Fuel Cell and Embedded Machine Learning," *IEEE Access*, vol. 9, pp. 108689–108701, 2021. Conference Name: IEEE Access.

[52] B. R. Ibrahim, F. M. Khalifa, S. R. M. Zeebaree, N. A. Othman, A. Alkhayyat, R. R. Zebari, and M. A. M. Sadeeq, "Embedded System for Eye Blink Detection Using Machine Learning Technique," in *2021 1st Babylon International Conference on Information Technology and Science (BICITS)*, pp. 58–62, Apr. 2021.

[53] B. da Silva, A. W. Happi, A. Braeken, and A. Touhafi, "Evaluation of Classical Machine Learning Techniques towards Urban Sound Recognition on Embedded Systems," *Applied Sciences*, vol. 9, p. 3885, Jan. 2019. Number: 18 Publisher: Multidisciplinary Digital Publishing Institute.

[54] C. Huang and L. Jiang, "Data monitoring and sports injury prediction model based on embedded system and machine learning algorithm," *Microprocessors and Microsystems*, vol. 81, p. 103654, Mar. 2021.

[55] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[56] G. Louw, "glouw/tinn," Apr. 2023. original-date: 2018-03-27T07:16:59Z.

[57] S. Salerno, "MicroMLGen for Python," Oct. 2020.

[58] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, Apr. 2002. Conference Name: IEEE Transactions on Evolutionary Computation.

[59] A. P. Wierzbicki, "The Use of Reference Objectives in Multiobjective Optimization - Theoretical Implications and Practical Experience," Aug. 1979. Num Pages: 32 Place: IIASA, Laxenburg, Austria Publisher: WP-79-066.

[60] J. Blank and K. Deb, "Pymoo: Multi-Objective Optimization in Python," *IEEE Access*, vol. 8, pp. 89497–89509, 2020. Conference Name: IEEE Access.

[61] M. McGlennen, M. Dieser, C. M. Foreman, and S. Warnat, "Using electrochemical impedance spectroscopy to study biofilm growth in a 3D-printed flow cell system," *Biosensors and Bioelectronics: X*, vol. 14, p. 100326, Sept. 2023.

[62] M. McGlennen, M. Dieser, C. M. Foreman, and S. Warnat, "Monitoring biofilm growth and dispersal in real-time with impedance biosensors," *Journal of Industrial Microbiology & Biotechnology*, vol. 50, p. kuad022, Aug. 2023.

# APPENDIX

## Remaining Experimental Results

This section contains complementary sensor response plots for experiments, which were not presented in the main manuscript. Experiments are named using the lamprey's identifier tag and the test number for that lamprey and are presented in the order they were performed. Plots are accompanied by a brief commentary on the experiment. In all the plots, sea lamprey attachments are identified with a pair of symbols: green triangles indicate when the sea lamprey attaches to the sensor and red dots indicate when it releases; each triangle-circle pair denotes one attachment.

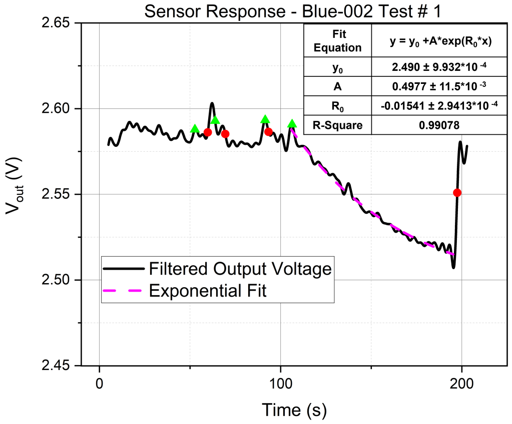


Figure A-1 Sensor response for Blue-002 Test 1.

Three short attachments were observed before a sustained attachment. Noticeable exponential decay observed; response was included in the correlation study.

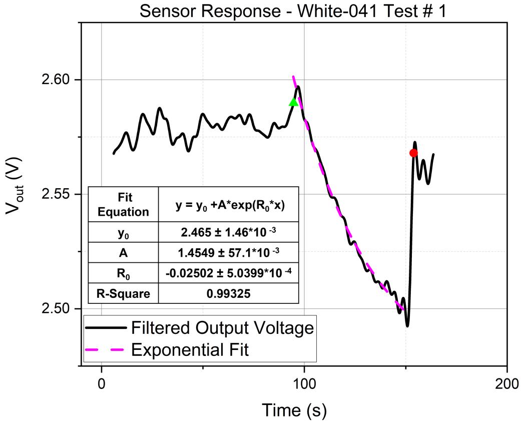


Figure A-2 Sensor response for White-041 Test 1.

Single sustained attachment observed. A transient spike followed by noticeable exponential decay observed; response was included in the correlation study.

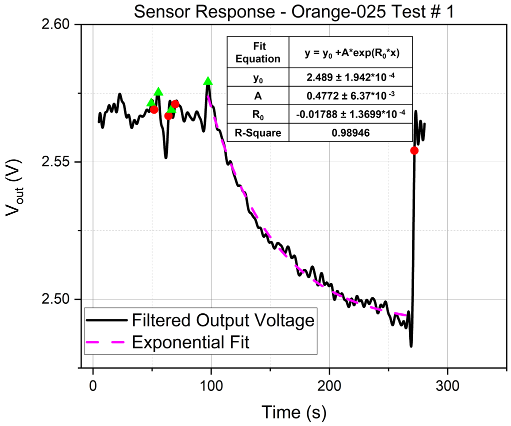


Figure A-3 Sensor response for Orange-025 Test 1.

Three short attachments were observed before a sustained attachment. Transient spike followed by noticeable exponential decay observed; response was included in the correlation study.

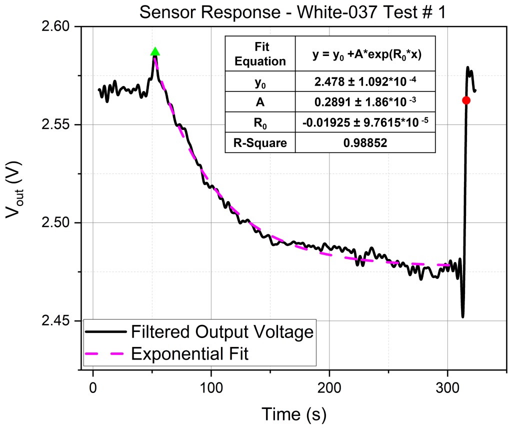


Figure A-4 Sensor response for White-037 Test 1.

Single sustained attachment observed. Transient spike followed by noticeable exponential decay observed; response was included in the correlation study.

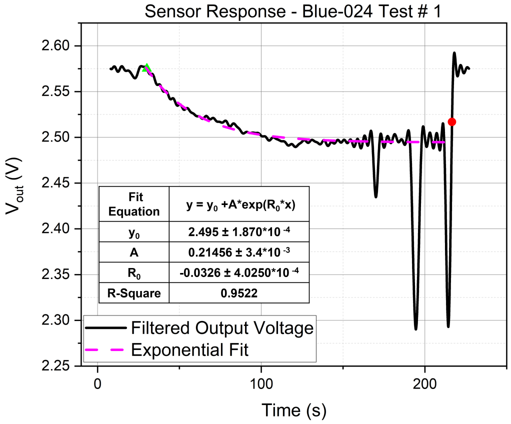


Figure A-5 Sensor response for Blue-024 Test 1.

One attachment with noticeable exponential decay observed; included in the correlation study. Transients from 150- 250 s, possibly due to movement of electric leads, not considered for exponential fit.

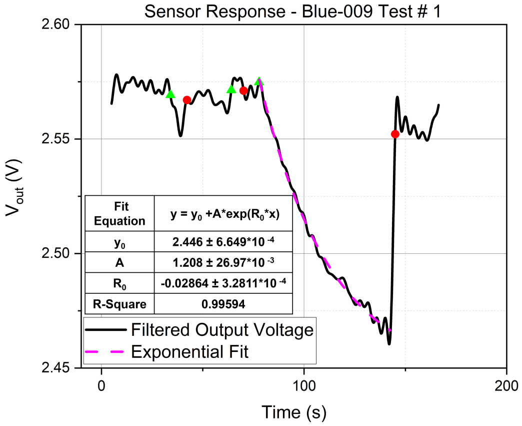


Figure A-6 Sensor response for Blue-009 Test 1.

Two short attachments observed before a sustained attachment. Noticeable exponential decay observed; response was included in the correlation study.

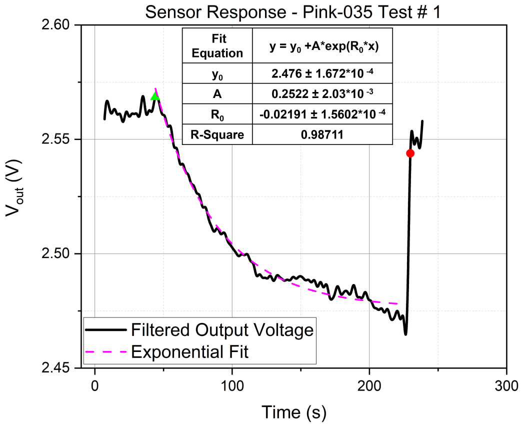


Figure A-7 Sensor response for Pink-035 Test 1.

Single sustained attachment observed. Small transient spike followed by noticeable exponential decay observed; response was included in the correlation study.

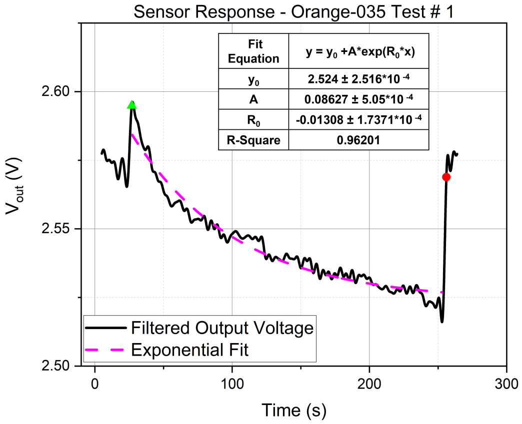


Figure A-8 Sensor response for Orange-035 Test 1.

Single sustained attachment observed. Transient spike followed by noticeable exponential decay observed; response was included in the correlation study.

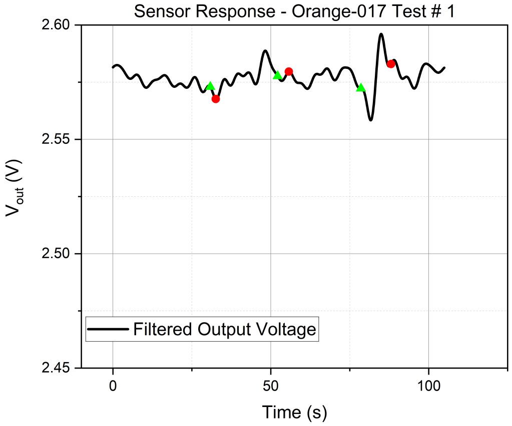


Figure A-9 . Sensor response for Orange-017- Test 1.

Three short attachments observed. Last attachment lasted 10 seconds, observed transient corresponds to lamprey rotating its oral disc on the sensor. Response was not included in the correlation study.

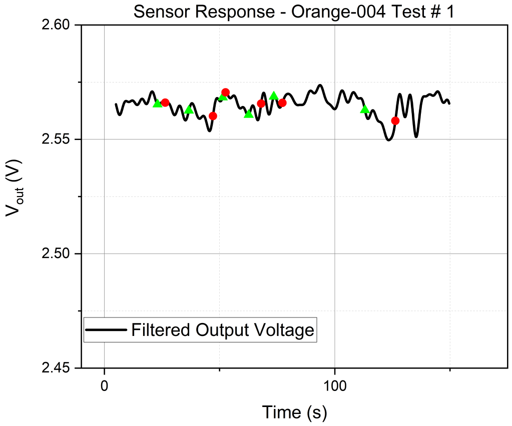


Figure A-10 Sensor response for Orange-004 Test 1.

Six short attachments observed during test. Penultimate attachment lasted 13 seconds, shows apparent decay but falls within noise variation. This response was not included in the correlation study.

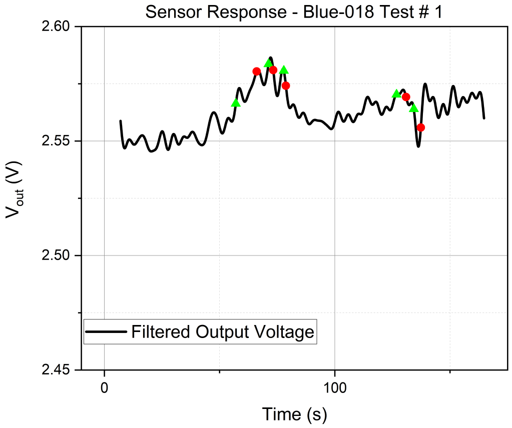


Figure A-11 Sensor response for Blue-018 Test 1.

Six short attachments observed during test. This response was not included in the correlation study.

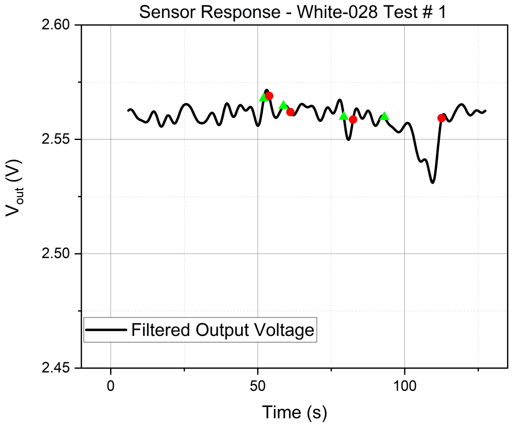


Figure A-12 Sensor response for White-028 Test 1.

Four short attachments observed during test. During last attachment, lasing 19 seconds, lamprey can be observed sliding on sensor. Response shape is not exponential, it was not included in the correlation study.

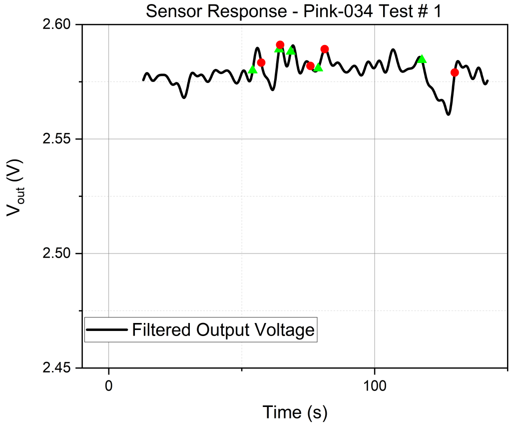
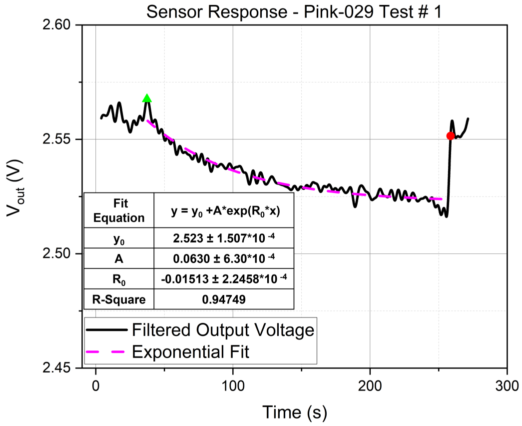


Figure A-13 Sensor response for Pink-034 Test 1.

Six short attachments observed during test. Last attachment lasted 11 seconds, shows apparent decay but falls within noise variation. This response was not included in the correlation study.



| Fit Equation | $y = y_0 + A*\exp(R_0*x)$ |
|---|---|
| $y_0$ | $2.523 \pm 1.507*10^{-4}$ |
| A | $0.0630 \pm 6.30*10^{-4}$ |
| $R_0$ | $-0.01513 \pm 2.2458*10^{-4}$ |
| R-Square | 0.94749 |

Figure A-14 Sensor response for Pink-029 Test 1.

Single sustained attachment observed. Small transient spike followed by noticeable exponential decay observed. Lamprey was only partially attached; response was not included in the correlation study.

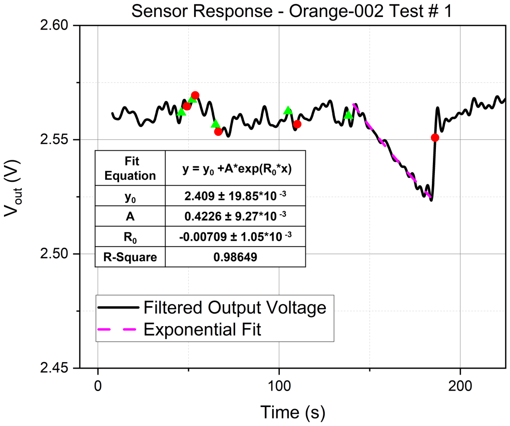


Figure A-15 Sensor response for Orange-002 Test 1.

Four short attachments observed before a sustained attachment. Noticeable exponential decay observed, but error is an order of magnitude greater than every other fit, not included in the correlation study.

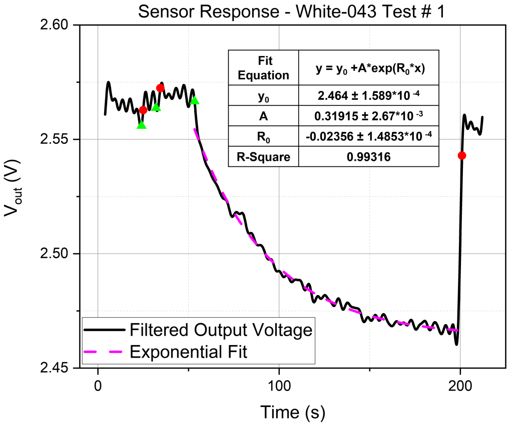


Figure A-16 Sensor response for White-043 Test 1.

Two short attachments observed before a sustained attachment. Noticeable exponential decay observed; response was included in the correlation study.

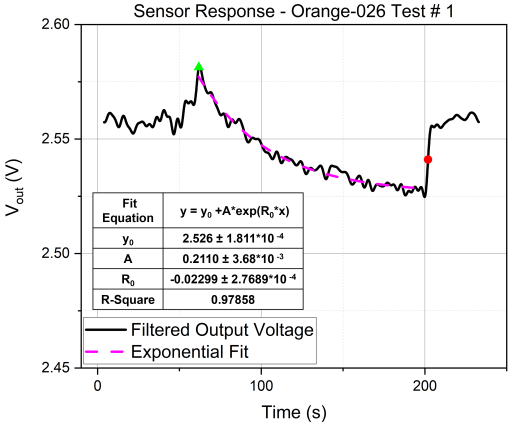


Figure A-17 Sensor response for Orange-026 Test 1.

Single sustained attachment observed. Small transient spike followed by noticeable exponential decay observed. Lamprey was only partially attached; response was not included in the correlation study.

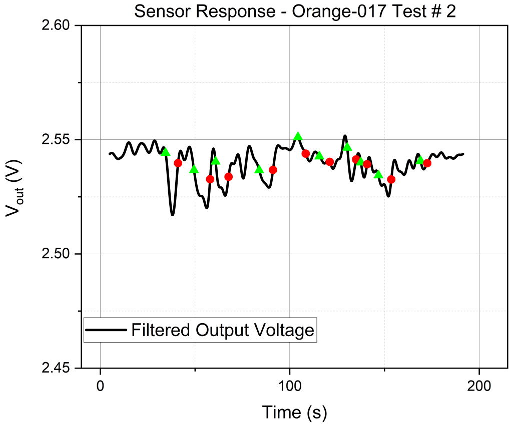


Figure A-18 Sensor response for Orange-017 Test 2.

Ten short attachments observed. First four attachments show apparent decay, but frequency of attachment makes them difficult to distinguish from noise. Response not used in correlation study.

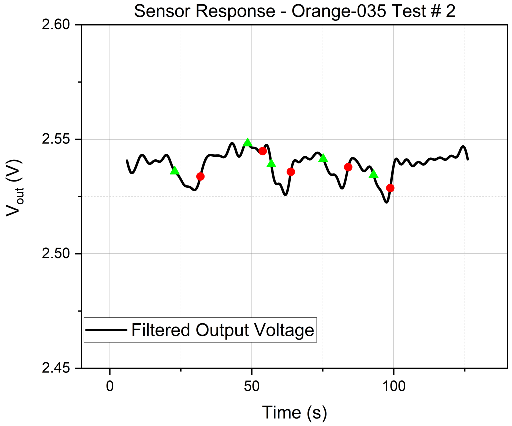


Figure A-19 Sensor response for Orange-035 Test 2.

Five short attachments observed. Most show apparent decay but fall within noise variation. Response was not utilized in correlation study.

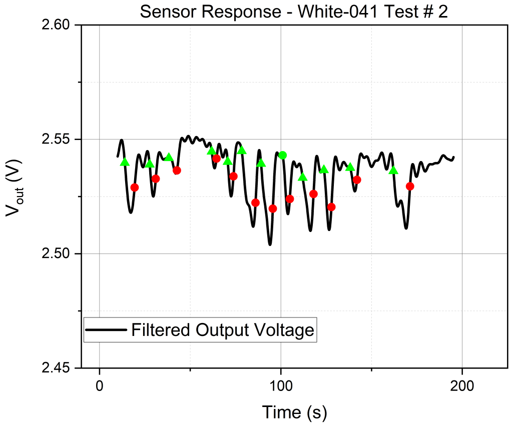


Figure A-20 Sensor response for White-041 Test 2.

Twelve short attachments observed. Most show apparent decay, but frequency of attachment makes them difficult to distinguish from noise. Response was not utilized in correlation study.

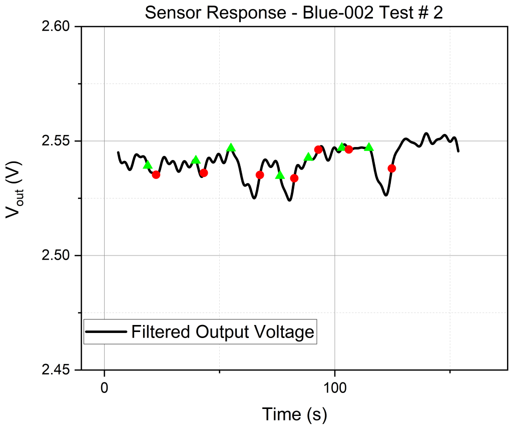


Figure A-21 Sensor response for Blue-002 Test 2.

7 short attachments observed. 3rd and 7th attachment ( 11 and  8 seconds respectively) show apparent decay but fall within noise variation. Response was not utilized in correlation study.

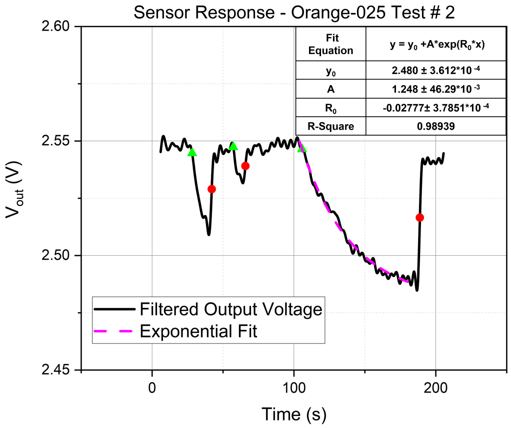


Figure A-22 Sensor response for Orange-025 Test 2.

Two short attachments showing apparent decay before sustained attachment observed. First attachment ( 12 seconds) had a low correlation with the exponential fit, only the 3rd was included in the correlation study.

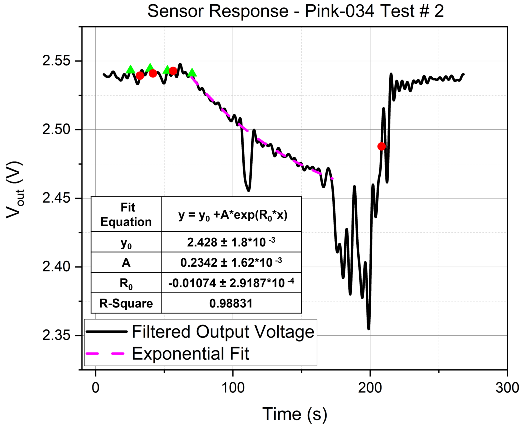


Figure A-23 Sensor response for Pink-034 Test 2.

Three short attachments before sustained attachment. Transients between 100-220s, possibly due to movement of electric leads, not considered for curve fit. Excluded from correlation study for this reason.

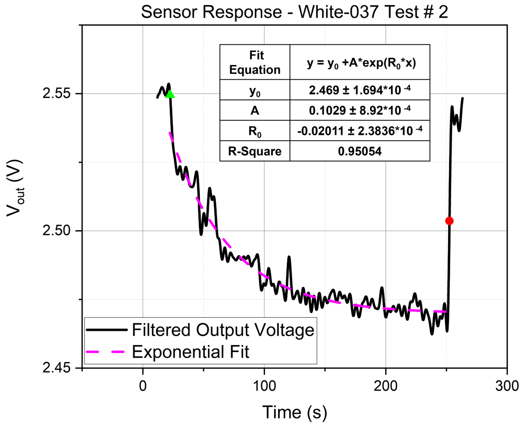
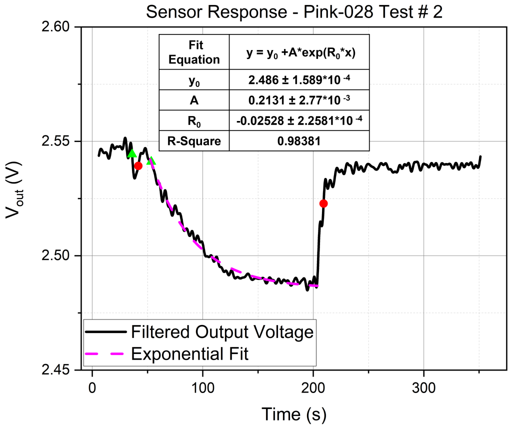


Figure A-24 Sensor response for White-037 Test 2.

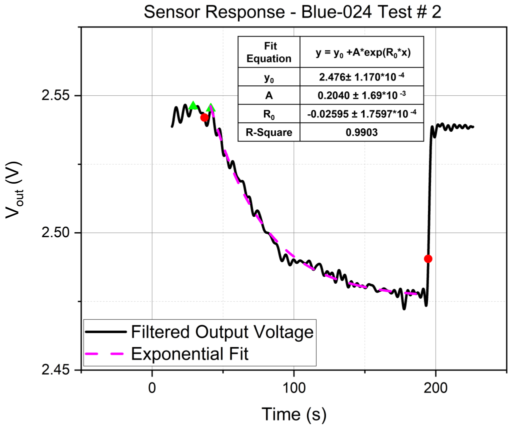
Single sustained attachment observed. Noticeable exponential decay observed; response was included in the correlation study.



| Fit Equation | $y = y_0 + A*\exp(R_0*x)$ |
|---|---|
| $y_0$ | $2.486 \pm 1.589*10^{-4}$ |
| A | $0.2131 \pm 2.77*10^{-3}$ |
| $R_0$ | $-0.02528 \pm 2.2581*10^{-4}$ |
| R-Square | 0.98381 |

Figure A-25 Sensor response for Pink-028 Test 2.

One short attachment before a sustained attachment observed. Noticeable exponential decay observed; response was included in the correlation study.



| Fit Equation | $y = y_0 + A*\exp(R_0*x)$ |
|---|---|
| $y_0$ | $2.476\pm 1.170*10^{-4}$ |
| A | $0.2040 \pm 1.69*10^{-3}$ |
| $R_0$ | $-0.02595 \pm 1.7597*10^{-4}$ |
| R-Square | 0.9903 |

Figure A-26 Sensor response for Blue-024 Test 2.

One short attachment before sustained attachment observed. Noticeable exponential decay observed; response was included in the correlation study.

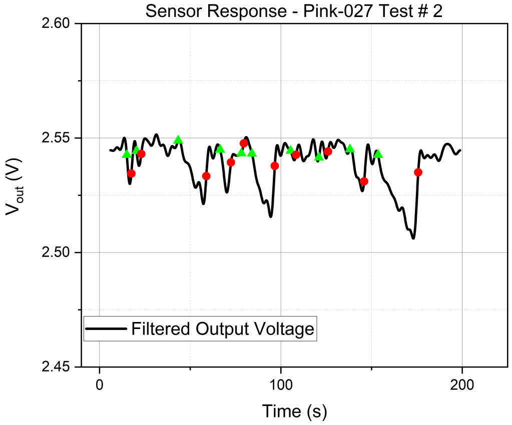


Figure A-27 Sensor response for Pink-027 Test 2.

Ten short attachments observed. 3rd,4th, 6th, 9th and 10th attachments show apparent decay. The largest decay appears almost linear, exponential fit correlation was low. Response not used in correlation study.

### Representative Response Videos

This section contains videos corresponding to the three representative tests presented in the main manuscript. The videos show the test proceedings and include the real-time sensor response plot in the upper left corner, as well as text annotations in the lower left corner indicating attachment status.

Video 1. Sensor Response to Sustained Attachment, Pink-027 Test 1 - Fig. 3.3
Video 2. Sensor Response to Short Attachment, Pink-028 Test 1 - Fig. 3.4
Video 3. Sensor Response to Sliding Attachment, Orange-040 Test 1 - Fig. 3.5

### Example Validation Testing Video

This section contains a video demonstrating the reaction of a selection of the models during validation testing using the 2 x 2 indicator light array included in the system. The models selected were NN-24, LGR, RF and Thresh 2 to get a broad view of the different model types under test.

Video 4. Video demonstration of real-time detection performance for selected models during validation testing

### Artificial Neural Network Deployment Using Tinn

The deployment of most of the models studied in this work through the use of the MicroMLGen python package was very straightforward, however artificial neural network models are not supported by this software at this time. The Tinn library, a dependency free C99 library which implements a single layer ANN[5], was selected due to its incredibly compact implementation of a single layer artificial neural network, making it ideal for deployment on embedded systems. However, the library only allowed for a single bias for the entire hidden layer, so it had to be modified to match the default ANN implementation used in Scikit-learn package used to train the networks which includes a bias for each node in the hidden layer. This modified version of the library files can be found here. To initialize a model from trained parameters Tinn requires they be grouped into two flat arrays with the following structure:

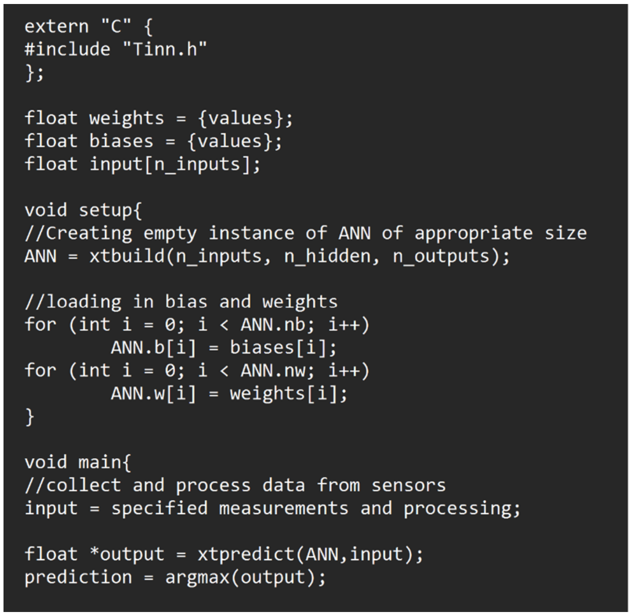
$$NN_{weights} = \{F(w_h^T), F(w_o^T)\} \in R^{1x(n*m+m*p)}$$

$$NN_{bias} = \{B_h, B_o\} \in R^{1xm+p}$$

Here $w_h$ is the n x m matrix containing the weights between the input and the hidden layers, where $n$corresponds to the number of inputs and m to the number of nodes in the hidden layer; $w_o$ is the m x p matrix containing the weights between the hidden and output layers with $p$ corresponding to the number of output nodes; $B_h$ and $B_o$ are vectors containing the biases for the hidden and output layers respectively; and $F(M)$ is an operator defining the row-wise flattening of the input matrix $M$:

$$F : \forall M \in^{nxm} \mapsto [row1, row2, \ldots, rown] \in R^{1+n*m}$$

The model parameters are extracted from the trained model and included in the arduino program as float arrays, though they could be saved to a text file on an SD card and accessed via a card reader module if memory size becomes an issue. After the parameters are declared, an empty instance of a NN with the appropriate dimensions is created using Tinn's xtbuild() method and the values are loaded into the model using for loops in the setup portion of the program. To use the ANN for inference in the main loop, the input is collected and processed as desired and fed into Tinn's forward propagation function xtpredict() alongside the ANN model. The ANN models used in the paper only used a single output node, so the class label was instead determined by comparison to a threshold (0.5). More generally, however, for classification tasks with three or more classes, the output will be a probability array from which the predicted class can be extracted using an argmax function. Figure A-28 shows a general implementation of the ANN deployment on an Arduino system.



```
extern "C" {
#include "Tinn.h"
};

float weights = {values};
float biases = {values};
float input[n_inputs];

void setup{
//Creating empty instance of ANN of appropriate size
ANN = xtbuild(n_inputs, n_hidden, n_outputs);

//loading in bias and weights
for (int i = 0; i < ANN.nb; i++)
        ANN.b[i] = biases[i];
for (int i = 0; i < ANN.nw; i++)
        ANN.w[i] = weights[i];
}

void main{
//collect and process data from sensors
input = specified measurements and processing;

float *output = xtpredict(ANN,input);
prediction = argmax(output);
```

Figure A-28 Rough code snippet showing the deployment of a single layer ANN model on an Arduino microcontroller platform for the general case.