

A Dissertation

entitled

Mitigating Harmful Algal Blooms using a Robot Swarm

by

Adam M. Schroeder

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the
Doctor of Philosophy Degree in Mechanical Engineering

Dr. Brian Trease, Committee Chair

Dr. Alessandro Arsie, Committee Member

Dr. Richard Becker, Committee Member

Dr. Ray Hixon, Committee Member

Dr. Youngwoo Seo, Committee Member

Dr. Amanda Bryant-Friedrichi, Dean
College of Graduate Studies

The University of Toledo

December 2018

Copyright 2018, Adam M. Schroeder

This document is copyrighted material. Under copyright law, no parts of this document may be reproduced without the expressed permission of the author.

An Abstract of
Mitigating Harmful Algal Blooms using a Robot Swarm

by

Adam M. Schroeder

Submitted to the Graduate Faculty as partial fulfillment of the requirements for the
Doctor of Philosophy Degree in Mechanical Engineering

The University of Toledo
December 2018

This research establishes the viability of using a swarm of robots to physically collect harmful algae from a bloom. This was accomplished by performing several sets of algae-collection simulations, measuring swarm performance by quantifying the algae collection rate. The two primary swarm control laws investigated were an approach that assigns each robot to its own region of responsibility, and a random walk biased in the direction of highest algae concentration. Then, an analytical basis was developed to establish how swarm performance changes due to robot-to-robot interference for different robot quantities and sizes. This basis also includes a formulation for robot and swarm cost, which allows performance-cost curves to be generated. Lastly, experiments were conducted where physical robots were used to collect real algae. Two companion technologies were also highlighted. The first such technology is a filter with bioinspired, anti-clogging features. The second technology is an unmanned aerial vehicle with a multi-spectral instrument for observing and quantifying algae concentrations.

Based on these simulations and experiments, it is recommended that robots in the swarm perform an unbiased random walk, which requires minimal robot sensing capabilities, minimal robot-to-robot communication, and therefore, minimal cost. A robot swarm does appear to be a viable solution for collecting harmful algae, but additional work is required to mature this technology.

This work is dedicated to Michelle, Linus, Simon, and Otto and to learning more quickly than we forget.

Acknowledgments

I gratefully, humbly, and most sincerely acknowledge my advisor, Dr. Brian Trease. It has been an honor to work with him and he has greatly supported me in innumerable ways. I also thank all of the members of my committee, who have all given their time to give me helpful feedback.

I have been supported by Dr. Bridgeman's lab group, who have allowed me to accompany them on Lake Erie sampling trips and given me access to other data collected during these sampling trips. I have also been supported by Dr. Becker's lab group, who have collected companion remote sensing data alongside me, and have helped me to analyze the remote sensing data that I've collected.

I am very thankful for funding received through a Small Ohio Sea Grant, which partially funded the multi-spectral instrument used in this work, and a grant from the Rotary club, which funded the hexacopter, improvements to the large algae-collection boat, and the nets that were used on little boats. This work couldn't have been completed without this funding.

I thank several local water treatment professionals for taking the time to show me their facilities, and answer many of my questions. They are Doug Wagner (Oregon), Andrew McClure (Toledo), Mike Sudman (Celina), Mike Fields (BG), and Bill Cody (The Pond Doctor).

I also would like to thank all of the following people for their contributions: Lauren Marshall, Sean Dunphy, Tyler Smith, Robert Longfield, Anna Becker, Nick Michael, Patrick Shemenksi, Cameron McCaskey, Lucas Palumbo, and Asal Soltani.

Contents

Abstract	iii
Acknowledgments	v
Contents	vi
List of Tables	x
List of Figures	xi
List of Abbreviations	xiii
1 Introduction	1
2 Background and Literature Review	6
2.1 Swarm Robotics	7
2.1.1 Taxonomy of Swarm Robotics	8
2.1.1.1 Taxonomy 1: Design and Analysis Methods	8
2.1.1.2 Taxonomy 2: Collective Behaviors	10
2.1.2 Algal Collection - Relevant Examples	12
2.1.3 Swarm Design	13
2.1.4 Unmanned Surface Vehicles (USVs)	15
2.2 Harmful Algal Blooms (HABs)	19
2.2.1 Blue-Green Algae Morphology, Other Major Algae Groups, and Microplastics	19

2.2.2	Existing Treatments and Collection Processes	21
2.2.3	HAB Biomass, Biovolume, and Areal Extent	28
2.2.4	In-situ and Remote Sensing	30
2.2.5	Other Great Lakes Data Streams	35
2.2.6	Cost of Treatment	36
3	Sizing, Scalability, and Control of a Robot Swarm to Mitigate Harmful Algal Blooms	40
3.1	Abstract	40
3.2	Introduction	41
3.3	Background	42
3.4	Problem Formulation and Implementation	45
3.5	Results	47
3.6	Discussion	51
3.7	Conclusion	54
3.8	Supplemental Info	55
3.8.1	Mathematical Problem Formulation	55
3.8.2	Non-Partition Approach: Stochastic Gradient-Following	57
3.8.3	Partition Approach: Modified Voronoi Decomposition	58
3.8.4	Analytical Approaches	60
3.8.5	Semi-Analytical Approach	61
3.8.6	Study Limitations	63
3.8.7	Computational Cost	65
3.8.8	Numerical Techniques	65
3.8.9	Diffusion Coefficients	65
4	Balancing Robot Swarm Cost and Interference Effects by Varying Robot Quantity and Size	67

4.1	Abstract	67
4.2	Introduction	68
4.3	Relevant Literature	72
4.4	Model Development	75
4.4.1	Interference and Swarm Performance	75
4.4.1.1	Application-Specific Model	82
4.4.2	Constant Size and Constant Capacity	83
4.4.3	Swarm Cost	85
4.5	Model Application and Discussion	87
4.5.1	Swarm Designer Perspective	87
4.5.2	Economy of Scale	90
4.5.3	Analytical Basis Comparison	92
4.5.4	General Discussion	93
4.5.5	Application-Specific Discussion	96
4.6	Conclusion and Future Work	97
5	Algae Collection Experiments and Supporting Technology Development	99
5.1	Algae Collection Experiments	99
5.2	Bio-inspired Cross-Step Filtration	104
5.3	UAV-based HAB remote sensing	107
6	Conclusion	112
6.1	Cumulative Conclusions	112
6.2	Major Contributions	113
6.3	Readiness Level Assessment	114
6.4	Future Work	116
6.5	Final Conclusion	117

List of Tables

2.1	Overview of bloom severity, biovolume, biomass, and areal extent	30
3.1	Overview of the main parameters of each simulation set	45
3.2	Partition perimeter to area ratio as robots' partition area varies	51
3.3	Comparison of semi-analytical and full simulation approaches	64
3.4	Diffusion coefficient calculations	66

List of Figures

1-1	Photo of Toledo water intake and location in Lake Erie	3
2-1	Biomachines Lab Aquatic Swarm	16
2-2	Singapore Aquatic Swarm	17
2-3	MIT's Seaswarm	18
2-4	Blue-green algae morphological groups	20
2-5	Overview of bloom diversity in color, patterns, location, and scale	22
2-6	Mobile dissolved air flotation boat	24
2-7	Commercial scum-mitigation solutions	25
2-8	Mobile scum-mitigation solutions	26
2-9	Plankton net structure and material	26
2-10	Microcystis aerial biovolume in western basin of Lake Erie	29
2-11	Spectra for blue-green algae	31
2-12	Cyanobacteria Index for Lake Erie and around Toledo water intake	33
2-13	Location of all drinking water intakes in the western basin of Lake Erie	34
2-14	National Data Buoy Center buoys in the western basin	36
2-15	NOAA continuous monitoring stations in the western basin	37
3-1	Small and large algae collection platforms	42
3-2	Collection performance over a range of diffusivity values	48
3-3	Final algae distributions for varying diffusion rates	49
3-4	Collection performance for a range of robot quantities/collection widths	50

3-5	Final algae distributions for varying robot quantities/sizes	51
3-6	Collection performance for a range of robot densities	52
3-7	Final algae distributions for a range of robot densities	53
3-8	Collection performance for unbiased Lévy collection	54
3-9	Analytical solution assuming average algae always collected	61
3-10	Semi-analytical solution with very coarse discretization	63
3-11	Eddy diffusion coefficient with respect to patch size	66
4-1	71
4-2	77
4-3	80
4-4	84
4-5	86
4-6	88
4-7	89
4-8	91
4-9	95
4-10	96
5-1	Detail of small boat platform	100
5-2	Multi-boat and single-boat algae collection experiments	102
5-3	Steps in the algae concentration process	103
5-4	Cross-step vortical filtration mechanism	105
5-5	Bio-inspired filter geometries and flow characterization	106
5-6	UAV HAB remote sensing platform	108
5-7	Flight deck for UAV flights from LEC boat	109
5-8	Raw remote sensing images from Tetracam and Go Pro	110
5-9	Initial cyanobacteria index scene with true color image	111

List of Abbreviations

ADI	Alternating Direction Implicit
CI	Cyanobacteria Index
CPU	Central Processing Unit
DAF	Dissolved Air Flotation
EPA	Environmental Protection Agency
ESA	European Space Agency
FOV	Field of View
FTCS	Forward-Time Center Space
GAC	Granular Activated Carbon
GLERL	Great Lakes Environmental Research Laboratory
GPU	Graphical Processing Unit
GRC	Glenn Research Center
GUI	Graphical User Interface
HAB	Harmful Algal Bloom
ILS	Incident Light Sensor
LEC	Lake Erie Center
MERIS	Medium Resolution Imaging Spectrometer
MGPD	Million Gallon per Day
MODIS	Moderate Resolution Imaging Spectroradiometer
MSS	Multi-spectral Sensor
ODHE	Ohio Department of Higher Education
NOAA	National Oceanic and Atmospheric Association

NASA National Aeronautics and Space Administration
UAV Unmanned Aerial Vehicle
USV/ASV Unmanned/Autonomous Surface Vehicle
WT Water Treatment

Chapter 1

Introduction

This work investigates the feasibility of employing a swarm of robots to physically collect and monitor harmful algae in open surface-water. Harmful algal blooms (HABs), which are massive populations of toxin-containing algae, present a serious threat to regional and worldwide water supplies. The root cause of HABs are excess nutrients that drain into rivers, lakes, and oceans. Stopping the flow of these nutrients is an enormous technical, legislative, and logistical challenge, which will take many years to accomplish. The robot swarm investigated herein should be viewed as a complementary measure that could more immediately reduce some of the negative impacts of HABs. A very brief introduction to this idea is given below, with subsequent sections providing further details.

Methods for mitigating algae blooms can be classified as physical, chemical, or biological. Physical methods are attractive because, when implemented correctly, they can have a smaller ecological impact. The alga (alga-singular, algae-plural) which have been most problematic in Lake Erie in recent years is taxonomically not a true algae, but a cyanobacteria, or blue-green algae, called *Microcystis aeruginosa*, which may contain a toxin called microcystin. Note that Lake Erie will be frequently cited as an example of a HAB-plagued body of water, which is done because of the wealth of both current and historical literature on HABs in Lake Erie, but this should

not be viewed as a limitation and this work is extendable to other bodies of water. *Microcystis* has some control over their vertical position in the water column and tend to aggregate near the surface during calm water conditions. Thus, they can be physically removed by skimming them from the surface, or by straining them from below the surface. Both of these processes have been implemented at a limited scale. The challenge is to (i) advance the technology for these two processes, and (ii) find a means of scaling these processes to the scale needed to combat massive blooms.

In terms of sheer scalability, swarm robotics is an intriguing option. Swarm robotics is an extension of the field of swarm intelligence, which studies how local interactions between agents following simple rules can lead to an emergent, system-level behavior. Swarm intelligence is largely inspired by observations of biological systems such as ant colonies. This type of system is so scalable because each agent is only following simple rules and not being directed by some central controller, which necessarily has limited computational and communication capabilities. Thus, when compared to a traditional, centralized system, a decentralized or distributed swarm is (i) massively scalable, (ii) robust to the loss of any individual agent, and (iii) composed of relatively unsophisticated agents following only simple rules. Note that using the term swarm implies a relatively large group of agents and that they communicate in a non-centralized manner.

The high-level goal of this research, which will be repeatedly referenced, is to establish the feasibility of constructing a robotic swarm capable of collecting harmful algae. It is envisioned that such a swarm could operate in the area around water treatment intakes, which would facilitate the subsequent conventional water treatment process. The swarm could also be used in other scenarios, e.g. cleaning recreational waters, but it is natural to concentrate on the scenario with very direct human-health consequences. It is helpful to understand this high-level goal in the context of what is state of the art in the field of swarm robotics. At this time, to the author's knowledge,

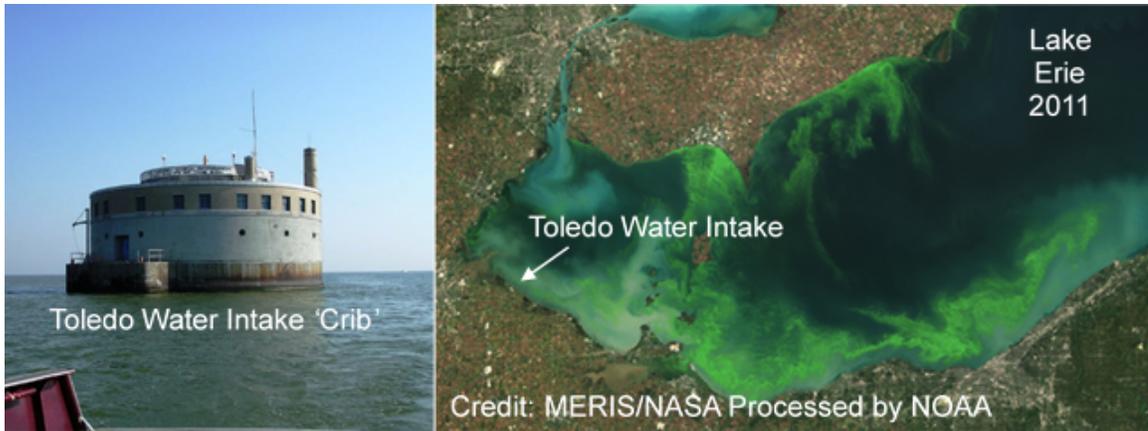


Figure 1-1: (Left) Toledo Water Intake; this 100 foot diameter concrete structure stands 55 feet above the water and draws in water 16 feet below the water's surface (Right) True satellite image of the 2011 bloom, one of the worst blooms on record, from ESA's MERIS satellite

no robotic swarm is operating outside of a controlled environment, utilized in a real-world application.

To facilitate the swarm design, and answer questions about swarm feasibility, several simulation sets were performed to establish a theoretical algae collection rate for different (i) algae diffusion rates, where diffusion is caused by turbulence in the water, (ii) robot quantity and robot size combinations, where the combined collection width of all the robots was held constant, and (iii) robot quantities, where the quantity was varied by several orders of magnitude. Following these simulations, an analytical basis was constructed that allows a swarm designer to predict how frequently robots in a swarm would collide. Finally, physical experiments were performed using up to three robots to collect real algae.

The major contributions of this work, explained in detail in the following sections, include (i) insight into the novel problem of collecting a diffusive substance, (ii) serious consideration of robot size in a swarm as a design input and its impact of swarm performance and cost, (iii) performance scalability across order of magnitude

changes in robot quantities, (iv) an analytical basis for predicting collision frequency, derived from physical phenomena, (v) swarm performance-cost curves, as a new tool for weighing design trade-offs, and (vi) demonstration of real-algae collection using a simple robot platform.

Parallel to these main contributions, several other research threads explored critical-path technologies that would be needed for such a swarm. The first major technology explored open-water filtration, as compared to filtering particles from a closed system, e.g. a pipe. It was recognized early that clogging would be a concern for any such filter. As such, filters were developed inspired by how ram-filter-feeding fish efficiently remove particles from the water. This filtering technique is referred to as 'vortical cross-step, filtration'. Filter development was performed using a water tunnel, and both dye and fluorescent, algae-sized particles were used for flow visualization and observations of particle collection patterns.

The second major technology explored was an unmanned aerial vehicle (UAV) with algae remote-sensing capability. A commercially-available hexacopter (six motors) was built up, and modified to house a multi-spectral camera system. Data was collected from several flights over Lake Erie, which were launched from a research vessel. Test flights were also conducted over a smaller pond, which contained algae. More detail on these two technologies (vortical cross-step filtration and UAV-based HAB remote sensing) is included, but the swarm robot research is the focus of this dissertation.

Starting from this brief, high-level introduction, the following sections provide more-detailed relevant background information and a thorough literature review (Section II), mathematically define the algae collection system, describing and discussing the accompanying algae-collection simulations (Section III), develop an analytical basis for collision frequency and swarm cost (Section IV), describe the algae-collection experiments with physical robots, and recap the open-water filtration and remote

sensing work (Section V), and conclude with an evaluation of the feasibility of an algae-collecting swarm (Section VI), drawing on the work from the previous sections.

Chapter 2

Background and Literature Review

This section gives more-detailed background information and a thorough literature review on (i) swarm robotics and (ii) harmful algal blooms. In contrast to the very concise introduction given in Section I, this serves an extended introduction. Within swarm robotics, the literature of interest is broken down into:

- (a) Taxonomy of Swarm Robotics
- (b) Algal Collection - Relevant Examples
- (c) Swarm Design
- (d) Unmanned surface vehicles (USVs) - Emphasis on Environmental Sensing

Within HABs, the literature of interest is broken down into:

- (a) Blue-Green Algae Morphology, Other Major Algae Groups, and Microplastics
- (b) Existing Treatments and Collection Processes
- (c) HAB Biomass, Biovolume, and Areal Extent
- (d) In-situ and Remote Sensing
- (e) Other Great Lakes Data Streams
- (f) Cost of Treatment

2.1 Swarm Robotics

Observations of social insects have grown into the budding research fields of swarm intelligence and swarm robotics. Swarm intelligence studies how local interactions between many agents can lead to a global, emergent behavior [29, 10] and swarm robotics extends this concept to physical robots [12, 30].

As previously introduced, robot swarms are touted for their (i) massive scalability, (ii) robustness to the loss of any individual agent, and (iii) composition of relatively unsophisticated agents following only simple rules. These attributes have redefined our perception of what robot systems are capable of and robot swarms have been proposed for planetary exploration, in-situ resource utilization, land-mine demining, oil-spill cleanup, disaster relief, surveillance, invasive-species mitigation [81], field cultivation [60], and precision pharmaceutical delivery.

Given this plethora of intriguing applications, it is initially disappointing to discover that no swarm is operating outside of a controlled environment, utilized in one of these real-world applications. This assertion is echoed by four preeminent researchers in the field of swarm robotics, Brambilla, Ferrante, Birattari, and Dorigo, in their highly-cited, recent, and comprehensive review of swarm robotics [12]. This absence can be attributed to the field of swarm robotics still being in its infancy, to the dependence on companion technologies, e.g. drug-delivering nanobots, to slow-moving regulatory processes, e.g. FAA restrictions on the use of quadcopters, and established non-swarm solutions, among other reasons. However, the largest factor holding back the implementation of robot swarms may be economic.

It is difficult to build a strong business case for developing or using robot swarms for many of the previously-listed example applications. For some, it is difficult or nearly impossible to predict where and when they could be deployed (oil-spill cleanup, disaster relief). For other applications, especially applications in the space domain,

there are no direct economic benefits (planetary exploration, in-situ resource utilization). While it may be possible to devise a stronger business case for some of these applications, an application that makes intuitive economic sense could more quickly catalyze the adoption of robot swarms to engineering problems.

HABs appear in the same location (Lake Erie and other eutrophic lakes), during a predictable time of year (June-October), and there are some direct economic costs and losses due to HABs (additional water treatment, recreation-related losses, business losses from water supply interruptions, decreased property values). The economic viability of using a robot swarm to mitigate HABs will be more rigorously established, but this introduction serves to provide the rationale for why literature on the swarm design process is sparse. In other words, many of the open research questions in swarm robotics are open because no one has been sufficiently financially motivated to create a physical swarm.

2.1.1 Taxonomy of Swarm Robotics

This snapshot of the state of swarm robotics can be supplemented by presenting two taxonomies for classifying swarm robotics work to this point, provided by Brambilla et al [12], which is done to establish the context of this new work. Their first taxonomy uses design and analysis methods, and their second taxonomy uses the required collective behavior. For clarity, other literature will be referred to using the common vocabulary of these taxonomies, even when the original authors may have used different terms.

2.1.1.1 Taxonomy 1: Design and Analysis Methods

Swarm design methods can be classified as behavior-based design or automatic design. In behavior-based design, the individual control of each agent is iteratively implemented, tested, and modified until the behavior of the collective swarm is ac-

ceptable. Examples of behavior-based design include using probabilistic finite state machines, which control the probability of a robot switching from one state to another, or virtual physics, such as artificial potential fields, where robots move based on a summation of attractive and repulsive forces. Automatic design methods automatically generate agent behaviors without the designer explicitly intervening. In this way, automatic design may result in a highly effective swarm, but without knowledge of what phenomena are leading to this desired outcome. Reinforcement learning and evolutionary robotics are examples of automatic design methods.

Swarm analysis methods may be viewed using models at the microscopic level, i.e. agent-level, or at the macroscopic level, i.e. collective-level. Microscopic level work may use increasingly complex levels of abstraction where robots are viewed as simple point masses, or as operating in 2D worlds subject to platform-specific kinematics, or as operating in 3D worlds with dynamic physics and detailed actuator and sensor profiles. These microscopic models are typically evaluated and analyzed using simulators. There are many options for robot simulation platforms, e.g. Webots, Gazebo, Stage, but Pincirolli et al developed one of the few that is specific to swarm robotics called ARGoS [77]. Macroscopic models view the system as a whole. These approaches have typically been to model the swarm using either rate and differential equations, including Langevin and Fokker-Planck equations, that can describe stochastic particles and systems, respectively, or classical control and stability theory.

Real, physical robots may also be used to validate behavior, such as the pioneering Kilobot [85].

Since Brambilla et al's review, an alternative analysis method has been proposed which estimates swarm performance by taking the integral of birth-death processes [54]. Another alternative approach developed after this review was to use 'property-driven design' [11]. This approach has four steps which start at the most abstract, and progressively become more realistic, which are well-illustrated using an aggregation

example. In the first step, the desired properties of the swarm are described (most abstract), e.g. robots must aggregate in one of several locations after a set time with a set probability. In the second step, these properties are used to form a model (becoming less abstract), e.g. robots may be inside or outside of the desired location for aggregation and they transition between these states with a certain probability. In the third step, the model is implemented and simulated (becoming more realistic), e.g. these aggregating robots perform a random walk until they are at a desired aggregation location and leave the location with a pre-set probability, depending on the number of other robots at that location. In the fourth step, physical robots are used (most realistic).

Within this taxonomy, the design method chosen for this work uses behavior-based design and a microscopic-level analysis.

2.1.1.2 Taxonomy 2: Collective Behaviors

The second taxonomy classifies work based on basic behaviors, e.g. aggregation, where multiple behaviors are combined to tackle more complex applications, e.g. foraging. These basic behaviors include spatially organizing behaviors, navigation behaviors, and collective-decision making. Spatially organizing behaviors are those that seek to organize and distribute robots and objects in space. These include aggregation, pattern formation, chain formation, self-assembly, and object clustering and assembling. Navigation behaviors seek to organize and coordinate movements. These include collective exploration, which could encompass area coverage and swarm-guided navigation, coordinated motion, also known as flocking, and collective transport. Collective decision making seeks to let a group of robots agree on a decision or allocate themselves among several simultaneous tasks. These include consensus achievement and task allocation. Other collective behaviors that do not fall into these categories include collective fault detection, group-size regulation, and

human-swarm interaction .

This second taxonomy is similar to a classification used in a more recent review of swarm robotic tasks by Bayindir [7]. This related taxonomy lists higher level behaviors like foraging alongside more basic behaviors like aggregation and flocking. It also adds some new behaviors such as odor source localization.

At this point, it is useful to analyze where this algae-collecting swarms fits in this taxonomy. First consider foraging, area coverage, and source localization. *Foraging* is typically used to mean the search for a discrete resource, e.g. an ant may forage for an individual seed. Often these resource fields are static, but they could also be dynamic. Algae, however, is different from a seed in that it is so populous, that is usually measured as a continuous concentration which varies spatially. *Area coverage* is typically used to mean physically moving through an area of interest and either modifying or simply observing the environment. This is the case when the agents' resources are not sufficient to modify or observe the entire domain simultaneously. The algae-collecting swarm could fall within this general definition of area coverage. The term area coverage may also be used to mean a static coverage, as in a network, that, once deployed, can observe an entire area of interest. *Source localization* is usually used to mean searching for a discrete source from which a gradient can be detected. This could be an odor, sound, or chemical plume. Algae is distributed continuously without any discrete sources, but some of the strategies used for following a gradient in a turbulent medium could be applicable.

After considering these existing terms, this application may be best framed using a new term, swarm collection, which will be generically defined as the physical collection of discrete or continuously-distributed materials. From this viewpoint, foraging is a sub-classification of collection, i.e. discrete collection, and continuous collection is a blending of foraging and area coverage. Indeed, the collection of continuously distributed materials, like algae, is an interesting and novel problem. One novel as-

pect of this problem is that continuously-distributed materials are often dynamically distributed because they are subject to physical phenomena such as diffusion or advection. This realization means that one artificial complex system (the swarm) is interacting with another natural complex system (the environment).

2.1.2 Algal Collection - Relevant Examples

Turning the focus now to select literature which have some aspects relevant to algal collection, Aznar et al proposed using a swarm of aerial robots to identify and track oil slicks [6]. The oil slick was modeled using environmental software, called GNOME, to predict the slick's location and size based on ocean currents, turbulent diffusion (simplified as eddy diffusion), and evaporation. When no slick was detected, their control scheme used a finite state machine with a wander behavior. When a slick was detected, they switched to a resource behavior, and when a slick was detected with greater than 80% oil content, they switched to an inresource state. One of the control parameters allowed them to either stay at the perimeter of the slick or move to completely cover the slick and they tested scenarios with a single slick or multiple slicks. They simulated both a microscopic model and a macroscopic model that used the Fokker-Plank equation. In another work, it was proposed to use a swarm equipped with fuzzy controllers to better track the edge of oil slicks, while the edge may not always be so clearly defined [74].

Song and Mohseni investigated how to maintain swarm cohesion when placed in an aquatic environment with a meandering jet, which is often used as an idealized model of ocean currents because it includes both major circulation patterns, that is, currents and vortices [94]. The members of their swarm are dubbed active Lagrangian particles because they drift like simple Lagrangian particles but have some limited self-propulsion. They were able to maintain swarm connectivity and avoid collisions by only actively controlling a small portion of the agents.

Some control schemes of interest may depend on the ability of an agent to sense a local gradient, so it is useful to examine literature that explains the limitations of gradient sensing techniques. Shaukat and Chitre elucidate how a gradient following approach, used by them to localize a source, can either use multiple sensors for each agent (instantaneous gradient detection) or only a single sensor (temporal gradient detection) [92]. They remark that using multiple sensors per agent to instantaneously detect the gradient is subject to the smoothness of the scalar field, the intensity variations over the body of an agent, and the noisiness of the sensor and the environment. This leaves temporal gradient detection, like that performed in bacterial chemotaxis, where the current intensity value is compared with past values.

Other swarm source detection work highlights the difficulties in measuring a gradient introduced due to advection and turbulent diffusion, both of which will be present in the algal environment [65, 43]. These phenomena make the instantaneous gradient time-varying and noisy.

Examples of aquatic swarms, which have not yet been proposed to be used in a swarm collection application, are included in the USV subsection.

2.1.3 Swarm Design

One of the most basic design decisions that a swarm designer needs to make is to choose a swarm size. There appears to be several strong results that show that as more agents are added to a swarm, the individual efficiency of each agent decreases. Adding more agents generally improves the performance of a swarm, but in some cases a point may be reached where adding further agents actually decreases the performance. Hamann perceives this as the result of the combination of cooperation and interference, where each are a function of agent density. Adding more agents simultaneously increases cooperation, while also increasing interference. Interference is mostly due to the increased need for collision avoidance behavior. This realization has

enabled a generic mathematical model to be fit to the curves for group performance and individual efficiency versus agent density across several different applications [44].

The most relevant example studies that examine group and individual performance as more robots are added are foraging and area coverage, with this interference effect demonstrated by Lerman and Galstyan [59] and Gaudiano et al [40], respectively. From Lerman and Galstyan (macroscopic mathematical model/simulations, 3m radius arena, 300 mm s^{-1} speed, 1-10 robots), one scenario did not require robots to deposit foraged objects to a nest, while their second scenario did require the deposition of those objects. The difference was that in the first scenario, adding more robots continued to increase group performance (up to their maximum number of robots tested), whereas in the second scenario group performance peaked when only four robots were used. From Gaudiano et al (physical experiments/simulations, 2000x2000/6000x6000, max velocity $5U/s$, 1000s, 1-10/10-110 robots), group coverage performance only increased as more robots were added.

Hecker and Moses (physical experiments/simulations, 10m x10m/10.5m x10.5m, 1-6/1-768 robots) also performed work studying swarm size for foraging [47]. Unlike Lerman and Galstyan, they did not consider interference in their simulations, but like Lerman and Galstyan, they still saw a sub-linear decrease in individual efficiency as more robots are added. Hamann's work was published before this study, but using his philosophy, this sub-linear decrease in agent efficiency could be explained using only the 'cooperation' function, in the absence of interference.

From this limited literature, it is clear this fundamental relationship, swarm size versus group and individual performance, is not thoroughly understood. This poses a challenge for the swarm designer and is one of the knowledge gaps that will be addressed in this work. A second question that is not addressed in literature is the relationship between agent size/capability and group and individual efficiency. An agent's size/capability may take different forms, depending on the application. This

could include sensing range, payload, max velocity, or if removing algae, the cross-sectional area being processed. A third basic question that arises is more specific to the novel problem of collecting a continuously-distributed, advecting, diffusing substance (algae). In Fickian diffusion, i.e. diffusion that is proportional to flux, substances are classified by their diffusivity. The third question is what the limit of collecting a diffusive substance is, as the diffusivity increases. These questions will be explored in the following chapters.

2.1.4 Unmanned Surface Vehicles (USVs)

Unmanned surface vehicles have not received as much attention as the land-based and air-based autonomous craft, but there are a good number of documented USVs in literature and industry. However, there is much sparser work on using USVs in a swarm. What has been done in this area comes largely from the BioMachines lab at the University of Lisbon. Duarte et al evolved controls leading to the canonical swarm behaviors of homing, dispersion, clustering, and monitoring using a real swarm with up to ten USVs [32], shown in Fig. 2-1. This was performed using a low-cost monohull USV developed within their lab [23]. Duarte et al also investigated how swarm robotics could be applied to aquatic environmental monitoring missions such as water temperature monitoring, where they validated their devised control laws using a small swarm [33]. Christensen et al showed how a swarm of these could use heterogeneous communication capabilities for an intruder detection task [20].

A group in Singapore has developed a swarm of omnidirectional, mobile buoys, dubbed ‘Bunch of Buoys’ [16, 111], shown in Fig. 2-2. They have tested groups of fifty buoys by performing aggregation and leader-follower exercises. Finally, a group at MIT proposed to build a ‘Seaswarm’ to collect oil after a spill, but they only constructed a single robot, shown in Fig. 2-3 [1].

Outside of this group, USV work with a swarm focus, or at least a multi-agent



Figure 2-1: Aquatic swarm created by Biomachines Lab at the University of Lisbon, Portugal [32].

focus, has also been performed by Valada et al, who designed and built a 40-70 cm long monohull, fan-powered USV [101]. They tested USVs by performing environmental monitoring tasks, e.g. monitoring temperature or conductivity, using a (i) random monitoring pattern, (ii) max uncertainty pattern, and (iii) a ‘lawnmower’ movement pattern. Multiple-USV tests used a maximum of 3 USVs. Chamanbaz tested a swarm of mobile buoys, or omnidirectional USVs, which had been designed to perform environmental monitoring [16]. These were part of the ‘Bunch of Buoys’ project, are limited to moving at about 1 ms^{-1} , and have been tested in groups of more than 50 units. In another work, the relative drag force of USVs moving in a V-shaped formation has been considered, where the USV had an omnidirectional profile [63].

There have been many other USVs created outside of the context of a swarm. A 16 foot, solar-powered catamaran was created to autonomously model water quality [34]. It was used alongside stationary water sensing nodes to monitor Lake Wivenhoe in Australia [35]. A different catamaran-style ASV, dubbed Lizabeth [49], was used in conjunction with a YSI fluoroprobe that could be raised and lowered to autonomously



Figure 2-2: Aquatic swarm of mobile buoys created by Singapore group [16].

monitor *Planktothrix* (a cyanobacterium) abundance on Lake Zurich [48]. A 1.1m long, remote-operated catamaran was created as part of the Brooklyn Atlantis project that uses a YSI sonde and can uniquely capture 360° panoramic images alongside collecting water quality data [56]. Unique in that it is powered using sails, a 3.72m long ASV roboat was created and used to acoustically monitor marine life [96].

USVs are also commonly combined with underwater platforms or aerial platforms, which could be aerostats, quadcopters, or fixed-wing aircraft. NOAA funded the development of an open-ocean, long-duration, solar-powered USV, called OASIS, which is 18 ft long and designed to be self-righting in heavy seas. A group of these OA-

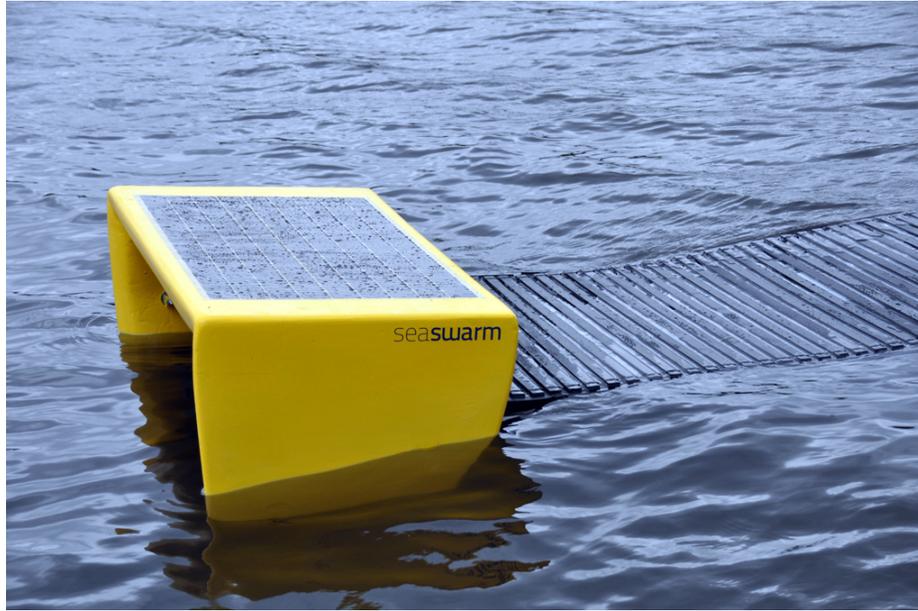


Figure 2-3: Single unit of MIT’s proposed Seaswarm to clean up oil spills [1].

SIS USVs were proposed to monitor HABs as part of a Telesupervised Adaptive Ocean Sensor Fleet (TAOSF), which includes using an unmanned, tethered aerostat [78]. They simulated a HAB using a diffusing, advecting patch of rhodamine dye, and used the fleet to map the dye’s spatial distribution. TAOSF used a pre-existing system for coordinating multiple, heterogeneous USVs called the Multilevel Autonomy Robot Telesupervision Architecture (MARTA) [62]. A second multi-mode robot system focused on coastal measurements, including monitoring algae blooms, and collects data using quadcopters, USVs, and underwater vehicles [102]. Note that one challenge of monitoring algal blooms using underwater vehicles is that they disrupt the water around them as they move, which impacts the sampling. Another example of a multi-mode, heterogeneous robot system pairs a fixed-wing aircraft, a catamaran surface vehicle, and underwater vehicle [93]. They used this system to collaboratively inspect the health of a coral reef.

On the opposite end of the size scale from these larger examples, a 200mm long

row-bot was created that uses a microbial fuel cell to convert organic biomass (could be from algae) to electrical energy [76]. The structure of the row-bot is inspired by the water-boatman beetle.

These examples of USVs by definition exclude ‘drifters’ and autonomous vehicles operating below the water, except when paired with a surface vehicle. Drifters are unmanned surface nodes, which are often used for aquatic monitoring [24] and simply passively move with the current.

2.2 Harmful Algal Blooms (HABs)

This second section introduces different aspects of harmful algal bloom literature which are relevant to the application of swarm collection.

2.2.1 Blue-Green Algae Morphology, Other Major Algae Groups, and Microplastics

Cyanobacteria, or blue-green algae, capable of forming blooms exist in three distinct morphological groups and the following overview of these groups is compiled from Paerl [73]. The first group comprises coccoid cells (spherical), which may exist as solitary cells or in aggregated colonies. *Microcystis* is a member of this first group. The second group comprises filaments of mostly undifferentiated cells. *Planktothrix* is an example genus of this second group. The third and final group consists of filaments of highly differentiated cells. Bloom-forming genera in this group include *Anabaena*, *Aphanizomenon*, and *Cylindrospermopsis*. Genera from all morphological groups are capable of producing secondary metabolites (toxins). Please see Fig. 2-4 for examples of these morphological groups.

For reference, *Microcystis* is the most problematic, dominant cyanobacteria in Lake Erie, but *Planktothrix* blooms frequently occur in Lake Erie’s Sandusky Bay

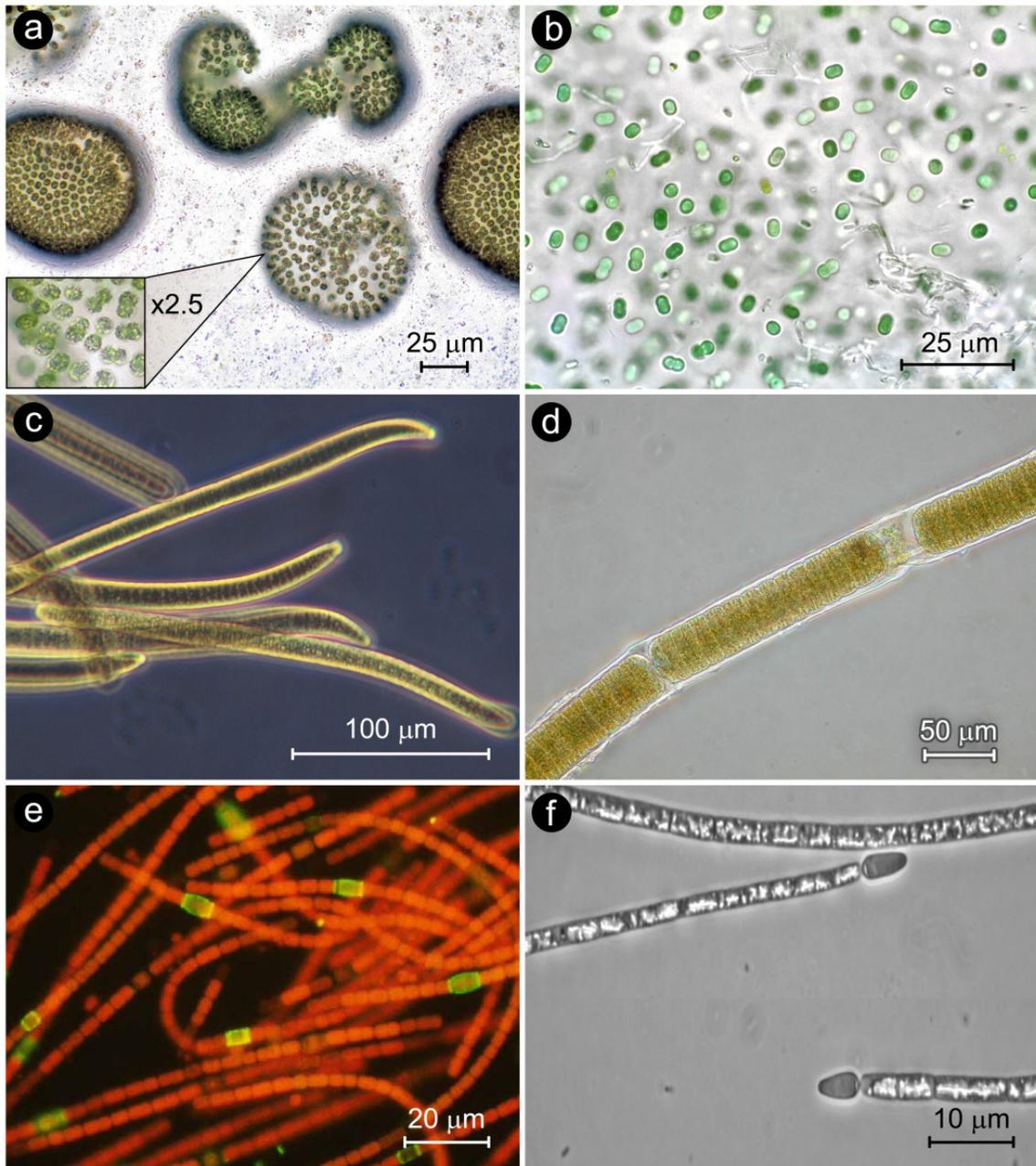


Figure 2-4: Overview of blue-green algae morphological groups from Paerl [73] including coccoidal (a,b), filamentous undifferentiated cells (c,d) and filamentous highly-differentiated cells (e,f). (a) Microcystis, (b) Synechococcus, (c) Oscillatoria, (d) Lyngbya, (e) Anabaena, (f) Cyndrospermopsis.

and in the Maumee River, Lake Erie's main tributary. *Aphanizomenon* caused the very severe bloom in 2010 in Grand Lake St. Mary's, although *Planktothrix* is normally dominant there. These regional examples are a microcosm of blooms that have occurred globally, as shown in Fig. 2-5. This figure, also borrowed from Paerl [73], gives a fine overview of the spatial and visual variation between different blooms.

These cyanobacteria all have some means of controlling their vertical position in the water column. They use gas-filled volumes, called vacuoles, which can be divided into smaller volumes, called vesicles, to control their buoyancy. Outside of cyanobacteria, several groups of non-toxic phytoplanktons, i.e. floating and suspended algae, are also present in the environment. These include green algae, diatoms, and cryptophytes. Unlike cyanobacteria, these other groups of phytoplankton are not able to regulate their vertical position in the water column.

The focus of this work is on collecting harmful algae, but it is useful to briefly draw out the similarities with collecting microplastics, which is a potential extension of this work. Microplastic accumulation is an area of active research [21], including within the Great Lakes [38, 31]. Microplastics are comprised of microparticles and microfibers, which are comparable to the coccoidal and filamentous morphologies of cyanobacteria, as shown in Fig. 2-4. Unlike cyanobacteria, these particles cannot actively control their buoyancy, but like algae, their movement and accumulation is driven by hydrodynamic phenomena.

2.2.2 Existing Treatments and Collection Processes

Methods for mitigating algae blooms can be characterized as physical, chemical, or biological and the following overview of these treatments is largely compiled from the International Guidance Manual for the Management of Toxic Cyanobacteria [71] unless otherwise cited. Biological methods seek to either increase cyanobacterial competition or increase grazing pressure on cyanobacteria, but are very difficult to

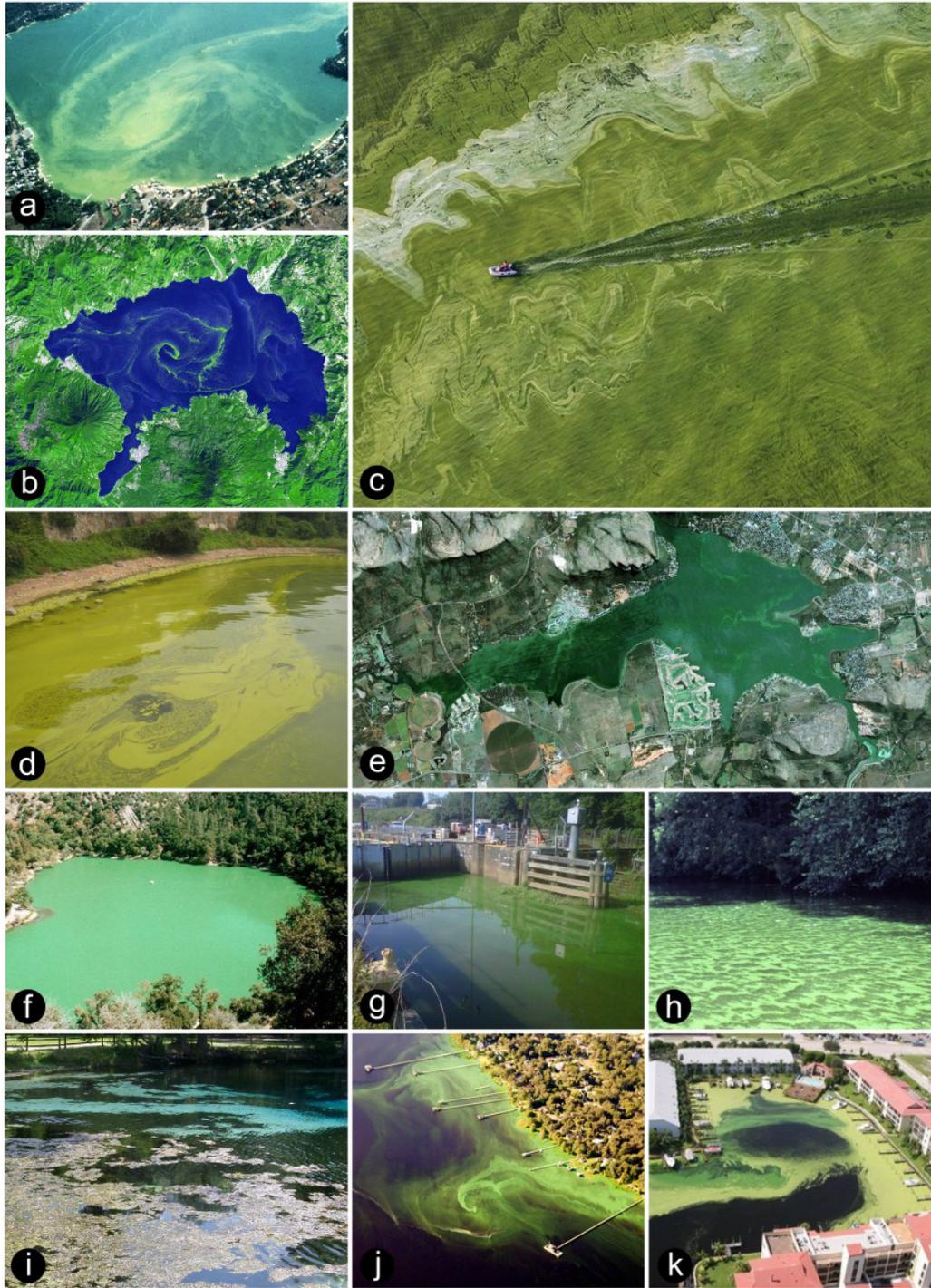


Figure 2-5: Overview of bloom diversity in color, patterns, location, and scale compiled by Paerl [73]. Microcystis-dominated or mixed Microcystis blooms are shown in (a) Liberty Lake, Washington USA, (c) Lake Erie USA, (d) Meiliang Bay, Lake Taihu China, (f) Zaca Lake, California USA, (g) Cape Fear River, North Carolina USA, (h) Neuse River, North Carolina, (k) Indian River Lagoon, Florida USA

implement because of the risk of severely (while perhaps unintentionally) altering the ecosystem. For this reason, they will not be considered for this application.

Physical methods include **(i)** mixing the water, which prevents water column stratification and subsequent release of nutrients from the sediment, **(ii)** drawing down a reservoir such that surface algal mats are deposited on the reservoir's sides, **(iii)** skimming algae from the surface, akin to an oil skimmer, **(iv)** straining algae using a filter, and **(v)** ultrasonically disrupting the gas vesicles used by the cyanobacteria to control buoyancy. The literature for skimming surface scum is quite limited, but Atkins et al did show how oil skimming equipment was used to remove surface scum from the Swan River near Perth Australia [5]. Literature for straining is likewise quite limited, but Chow did successfully demonstrate micro- and ultra-filtration of cyanobacterial cells [17]. It was also noted that micro- and ultra-filtration did not cause any significant release of intracellular toxin. Ultrasonic treatment can successfully cause gas vesicle collapse [109], but may have a deleterious impact on other nearby organisms and, at higher powers, may cause cyanobacterial cell lysis, which releases the intracellular toxin.

Chemical methods include **(i)** hypolimnetic oxygenation, which injects oxygen to reduce stratification in the water column, **(ii)** phosphorous precipitation and capping, which precipitates phosphorous from the water to the sediment where it's capped, **(iii)** coagulation to facilitate precipitation of cyanobacterial cells to the water body floor, and **(iv)** algaecide application which kills cyanobacterial cells. Many algaecides are copper-based which has a devastating ecological impact and must be removed in a subsequent process if used for drinking water. Potassium permanganate, chlorine, and hydrogen peroxide are examples of non-copper-based algaecides. Aside from any other effects the algaecide may have on the environment, algaecides also cause cyanobacterial cell lysis, which releases intracellular toxins into the water.

One method not yet described is a mixture of a chemical and physical processes



Figure 2-6: Dissolved air flotation system mounted on a mobile platform, taken from Mr. Bryan's submission video for the 2016 Barley Prize

called dissolved air flotation (DAF). Bubbles of air are introduced below the surface which attach to cyanobacterial cells and cause them to rise. Then, they can be skimmed from the surface. For the best results, a coagulant should be used upstream of this process [36, 52]. A patent has been filed by Mr. Kent Bryan for a DAF system mounted on a boat, shown in Fig. 2-6, and he is currently testing the design in Grand Lake St. Mary's [14].

In general, a chemical method would be difficult to scale up to the massive size of blooms because of the gross amount of chemical that would be needed and the side impacts it may have on the ecosystem. Some of the physical methods such as skimming and straining could be scaled up in the form of a swarm, are feasible to use on a mobile platform, and do not have the major drawbacks associated with other physical, chemical, and biological methods. Thus, a deeper investigation into skimming and straining is warranted.

Skimming is effective at collecting algae at the surface of the water which could

take the form of either a thicker surface scum, or a more aqueous form when the suspended algae is more concentrated near the surface. Skimming is likely most effective at removing the more solid surface scums which form on the leeward shores of lakes [104], and at the end of the season when the temperature begins to drop and the bloom begins to die off. Products are commercially available to skim nuisance surface plants, like duckweed, leaves, pollen, and other trash and debris. These may take the form of (i) an on-lake floating collection unit from which water is pumped to an on-shore filtration unit (see Fig. 2-7-left and middle), (ii) a manual skimmer that a user casts from a dock and then reels in (see Fig. 2-7-right), or (iii) a gas-powered on-lake surface vehicle with a large conveyor (see Fig. 2-8-left). Although these commercial products are likely to be as effective for cyanobacterial surface scums as for any other surface debris, none of these are marketed specifically for blue-green algae. One reason for this may be the liability incurred by users skimming toxic surface scum. A more analogous example of algae skimming is a remote-operated robot boat constructed in 2013 by students at the University of Waterloo specifically for collecting algae (see Fig. 2-8-right). Using its conveyor, they stated that it collected approximately 300 lbs of algae in 20 minutes.



Figure 2-7: (left) propondandlakes.com’s Proskim system’s in-water floating collection unit; (center) Proskim system’s onshore filtration unit; (right) weedersdigest.com’s hand-cast parachute skimmer

Straining can be used to remove algae either at the surface, or below the surface. Algae is often strained by researchers using planktonic nets made from nylon meshes

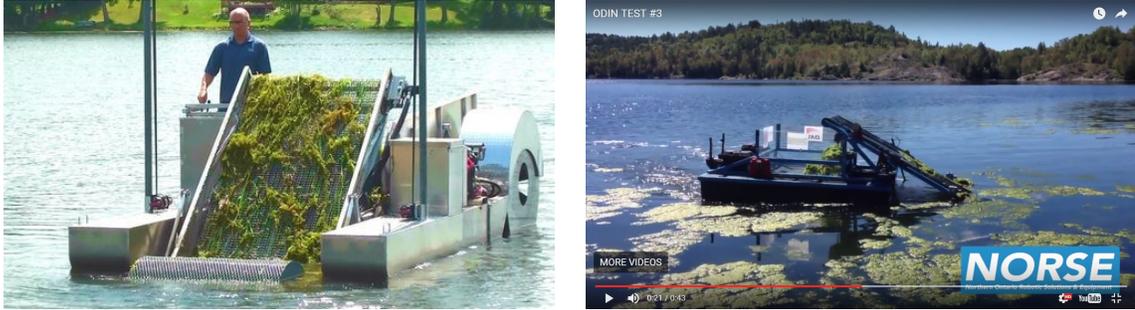


Figure 2-8: (left) weedersdigest.com’s manned Ecoharvester skimmer; (right) remote-operated skimming platform developed at the University of Waterloo

(see Fig. 2-9-left). This type of woven nylon mesh is manufactured by Sefar, shown in Fig. 2-9-right, with mesh openings ranging from 15 μm to 5000 μm . Individual *Microcystis* cells are only 5 μm , but these cells aggregate into colonies with an average size ranging from 20 μm to 700 μm during the year, per a recent study in Lake Taihu [110]. Colonies in Lake Erie are typically greater than 50 μm [103], and it was reported that approximately the same amount of algae was collected with a 64 μm plankton net as a 112 μm net [13].

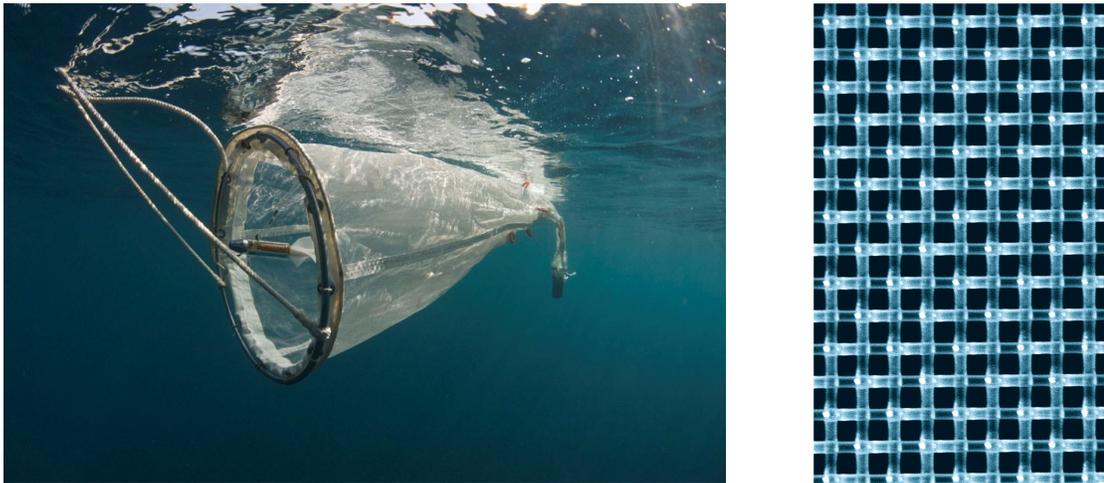


Figure 2-9: (left) Typical plankton net designed to be towed behind a boat [credit Peter Verhoog]; (right) Sefar Nitex woven nylon mesh

The vertical distribution of colonies in the water column, which is important to understand for development of a straining system, is an area of active research (see for a recent example [84]). Although a cohesive understanding has not been achieved, several observations may be given. First, larger colonies have more gas vesicles, and can therefore exhibit stronger buoyancy and faster ascent rates than smaller colonies [70, 69]. Second, strong light exposure can make *Microcystis* colonies less buoyant [69]. Third, wind pressure, and subsequent mixing in the water column may overcome colonies' buoyancy, which is discussed in more detail in the context of remote sensing [107].

Anecdotally, *Microcystis* was believed to rise to the surface at midday to absorb sunlight, and sink deeper in the water column at night to take in nutrients. This simple understanding is called into question by data collected by Dr. Bridgeman's group from the University of Toledo at the Toledo water intake. They performed two, 24-hour-long tests in 2016 where blue-green, diatom, green, and cryptophyta concentrations were measured at 0.1 m depth intervals in the water-column. At least for these two observation windows, where many of the other factors may have been at play, blue-green's concentration did not significantly increase at the surface during the day nor increase below the surface at night.

From a straining system design standpoint, several approaches may be considered. It would be easiest to design a system that always collects from the water's surface to some point at a fixed depth below the surface, e.g. 1 m. A more sophisticated system could sense where the highest concentration of algae was located and adjust the location of the filtering window, e.g. between 1 m and 2 m. One must also consider the other organisms in the environment which are in the size range to be strained by the filter. In the type of waters where this swarm is designed to operate, blue-green algae dominate such that collection of non-blue-green algae is hypothesized to represent only a small fraction of total biomass collected.

2.2.3 HAB Biomass, Biovolume, and Areal Extent

A study by Bridgeman et al [13] does give a good data source for estimating how much biovolume would be collected if all of the algae within a square meter (from surface to bottom) could be removed, the relative biovolume during the entire season, and variation in biovolume from year to year. These biovolume estimates are based on dropping a 112 μm planktonic net to the bottom and drawing it to the surface so that the biomass contained within the entire water column is collected. The authors averaged these numbers over four sites in the western basin of Lake Erie over the bloom season from 2002-2011. Fig. 2-10(left) shows how these biovolumes varied throughout the season and Fig. 2-10(right) integrates the curves from Fig. 2-10(left), giving a pseudo-cumulative biovolume (my term) for the entire season, which is a useful indicator of the severity of the bloom from year to year. The term pseudo-cumulative biovolume is used because it is found by taking the area under the biovolume curve, which is not a true measure of biovolume. However, these pseudo-cumulative biovolumes, determined using only four sampling locations, visually correlate well with more recent data that give a general bloom severity for the years 2002-2017 for the entire lake, shown in Fig. 2-10(right-inset).

Analyzing these data, with pertinent values given in Table 2.1, helps to provide context for the intra- and inter-year variability within which a swarm would be operating. The pseudo-cumulative biovolumes varied from 36 000 $\text{mL m}^{-2} \text{year}^{-1}$ in one of the worst bloom years, 2011, to 15 000 $\text{mL m}^{-2} \text{year}^{-1}$ in a moderate bloom year, 2009, to 3000 $\text{mL m}^{-2} \text{year}^{-1}$ in a non-bloom year, 2007. Thus, the difference between one of the worst years and best years is about an order of magnitude. For each of these years, the maximum biovolume at one sampling date was 1125, 360, and 75 mL m^{-2} , respectively. This wet biovolume can be converted to both wet and dry biomass. Wet biomass and biovolume are more relevant to this collection applica-

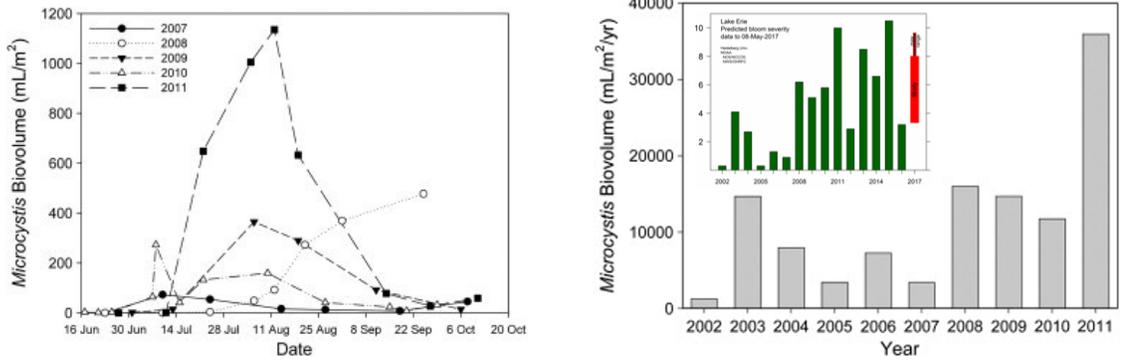


Figure 2-10: (left) Microcystis arial biovolume averaged over four sites in the western basin of Lake Erie [13]; (right) Area under the curves from left-hand image which can be interpreted as a pseudo-cumulative biovolume, allowing bloom severity to be compared from year to year; (right-inset) Heidelberg’s year by year bloom severity where 2011 was set to be the maximum severity (10.0), but was later extended for the 2015 bloom.

tion. The wet biomass values can also be extrapolated to mass per square kilometer, which is a more suitable system level measure. Please see the table notes for how wet biovolume and biomass are determined.

Table 2.1 also includes an entry for the maximum concentration at any one site for the entire span of years in the study. Noting that all of the values in the table are yearly or overall maximum values, this helps establish an upper bound for designing a swarm.

Excellent data is also available to estimate the areal extent of HABs in Lake Erie using remote sensing. Recent work uses the long-duration Landsat program satellites (Landsat 5 1984-2001), in addition to the more recent MODIS (2002-2011) and MERIS (2012-present) satellites, to give an estimate of HAB areal extent [50]. This remote sensing data correlates well with Bridgeman et al’s in-situ data, and is also included in Table 2.1. Note that in the severe bloom year of 2011, the areal extent of the bloom is 1800 km².

Table 2.1: Overview of bloom severity, biovolume, biomass, and areal extent

Bloom Year (Severity)	Pseudo Cumulative Biovolume ¹ [mL m ⁻² year ⁻¹]	Areal Biovolume ¹ [mL m ⁻²]	Areal Wet Biomass ² [g m ⁻²]	Areal Dry Biomass ³ [g m ⁻²]	Areal Wet Biomass [kg km ⁻²]	Areal Extent ⁴ [km ²]
2011 (Severe)	36000	1125	1125	11.963	1125000	1800
2009 (Moderate)	15000	360	360	3.778	360000	1125
2007 (Low)	3000	75	75	0.728	75000	370
Outlier	N/A	2000	2000	21.326	2000000	N/A

¹ Data taken from Bridgeman et al [13]

² Assumes a specific gravity of 1.0 [64]

³ Using relation developed by Bridgeman et al [13]. Note that other literature measures individual cell's diameter to estimate wet biovolume [64, 22], which includes only intracellular water, whereas Bridgeman et al measures the volume of an entire mass of floating cells and colonies, which includes intracellular and some extracellular water. For this reason, the 10:1 wet to dry biomass relation sometimes used in literature [80] cannot be used for Bridgeman et al's data.

⁴ Data taken from Ho [50]

2.2.4 In-situ and Remote Sensing

Sensing cyanobacteria is a critical capability so that a swarm can concentrate its efforts where algae are most abundant. This could either take the form of in-situ measurements, performed by collecting agents, or remote sensing performed by dedicated aerial sensing agents. Cyanobacteria is typically detected and quantified by observing reflectance spectra (from solar irradiance) or induced-fluorescence (when excited by a light source). Two intracellular pigments, chlorophyll-a and phycocyanin, drive these unique spectral responses, shown individually in Fig. 2-11(left). The spectral signature of *Microcystis*, which is clearly influenced by these two pigments, is shown in Fig. 2-11(right). Chlorophyll-a is responsible for the reflectance valley (absorption peak) near 680 nm and phycocyanin is responsible for the reflectance valley (absorption peak) near 625 nm. When measured in-situ, a fluoroprobe or fluorometer, e.g.

the YSI EXO2, excites the sample and the fluorescence intensity is monitored and correlated to cell abundance. When detected remotely, solar irradiation is incident on the algae-containing water and the corresponding radiation can be observed with a multi- or hyper-spectral imaging device.

Of course, algal abundance may be directly, but more laboriously, established by counting the number of cells. Generally, algal abundance, whether directly enumerated or correlated to a spectral response, is given in mg m^{-3} , $\mu\text{g L}^{-1}$, or ppb, which are all equivalent.

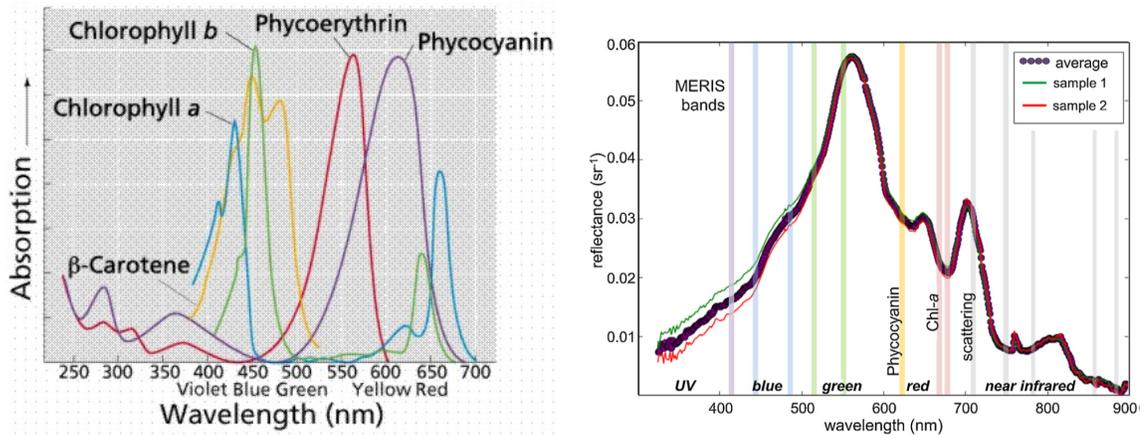


Figure 2-11: (left) Absorption spectra for pigments of interest which help to differentiate blue-green algae from other substances (organic and non-organic) in the water [79]. Fresh-water blue-green algae contain chlorophyll-a and phycocyanin.; (right) Reflectance spectra for *Microcystis*, with illustration of MERIS bands [97].

Several algorithms have been proposed for estimating algal cell abundance from remote-sensed data. One of the algorithms that has been widely adopted is the cyanobacterial index (CI), introduced by Wynne et al [108]. Using the CI for the aerial sensing agents in a swarm is an attractive proposition because the CI (i) requires only three bands, which necessitates a multi-spectral imager, but not a much more sophisticated hyper-spectral imager, (ii) can be correlated to approximate blue-green biomass, and (iii) is already widely used in the remote sensing community. This

algorithm is called a shape algorithm because it calculates how quickly the spectral response is changing its shape at a target wavelength; in this case near the 680 nm chlorophyll-a reflectance peak. The CI value can be correlated to cell count mL^{-1} [107] or biomass $\mu\text{g L}^{-1}$ [2], and then integrated spatially to find biomass. The CI was initially devised for data from the European Space Agency's MERIS satellite with bands centered at 665, 681, and 709 nm. In 2012, contact with MERIS was lost and Wynne et al [106] devised a revised CI for NASA's MODIS satellite, which had been previously used to monitor cyanobacteria [8]. The MODIS bands centered at 667, 678, and 754 nm were used, and they also calculated a correction factor so that data could be compared between both platforms. Fig. 2-12 shows an example MODIS CI image of blooms in the western basin of Lake Erie from September 2013 and September 2014.

Remote sensing is performed not only from satellite platforms, but also from aerial platforms. NASA Glenn Research Center has been providing weekly or biweekly aerial coverage over points of interest in Lake Erie, e.g. the Toledo water intake and Maumee Bay State Park. These aircraft fly hyperspectral imagers (HSI) [58] and can provide a much better spatial resolution than satellites (1 m x 1 m pixel versus 1 km x 1 km pixel). The HSI data is compiled from which the CI is supplied as a standard Quicklook product, which water managers can use to make management decisions [57]. An example CI data field from an August 10 2015 flight over the Toledo water intake is shown in Fig. 2-12.

One complicating factor when monitoring algae remotely is that all algae may not be present near the surface. This could especially be the case when high winds are present to mix the algae into the water column. When studying this possibility, Wynne et al hypothesized that for a wind stress of $<0.05\text{Pa}$, most of the algae is near the surface and the remote sensing CI estimate should be accurate. For wind stress $>0.1\text{ Pa}$, the algae would be more homogeneously distributed throughout the

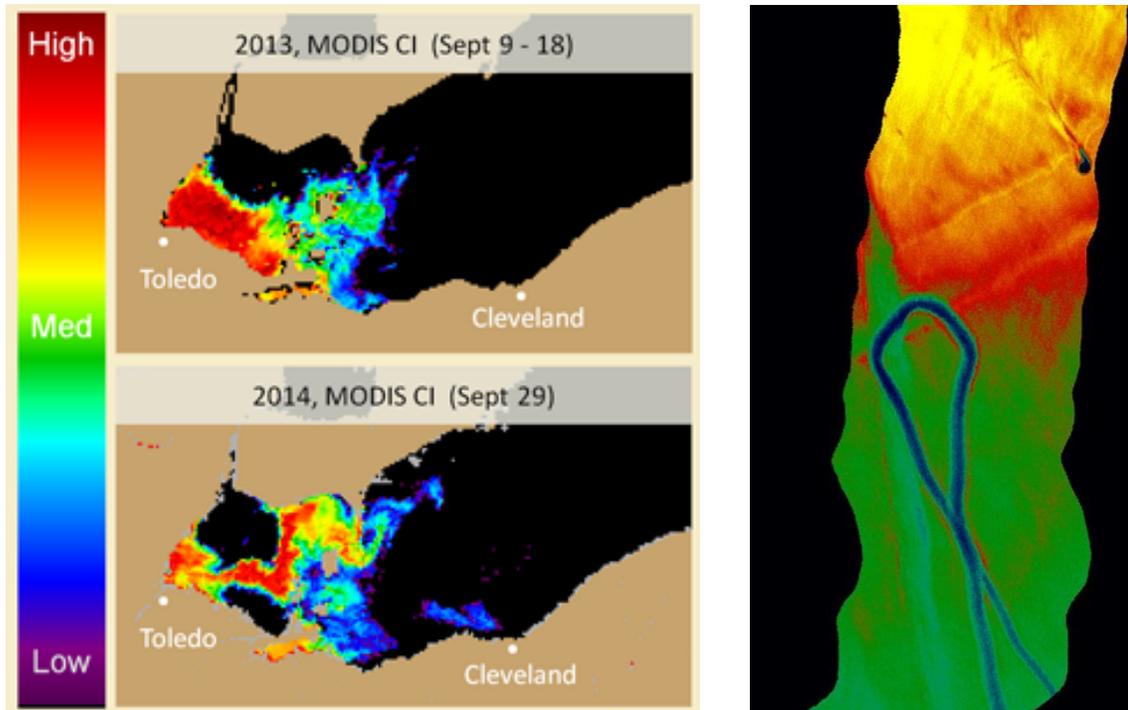


Figure 2-12: (left) Cyanobacteria Index (CI) for the western basin of Lake Erie in September of 2013 and 2014, derived from NOAA MODIS data; (right) CI for a transect over the Toledo water intake in August 2015 derived from NASA Glenn Research Center hyperspectral flyover data. The intake is visible in the upper right and an old boat path is also visible.

water column [107]. Whether due to wind, or due to *Microcystis* controlling their buoyancy, the surface concentration has been shown to vary significantly from hour to hour, which was shown using a geostationary satellite over Lake Taihu in China [51].

CI data was compiled for the years 2002-2014 and used by Wynne et al to assess the threat level to drinking water intakes located in the western basin of Lake Erie, shown in Fig. 2-13 [105]. They concluded that Toledo had the highest frequency of bloom occurrence. Monroe has significantly less bloom activity for most years relative to Toledo, and Carroll and Ottawa County have early blooms while Put-in Bay has a short peak of high frequency blooms in late August. This supports the assertion that

concentrating efforts around Toledo's water intake is appropriate, and that although these other intakes are at a smaller risk, they too could benefit from swarm collection.

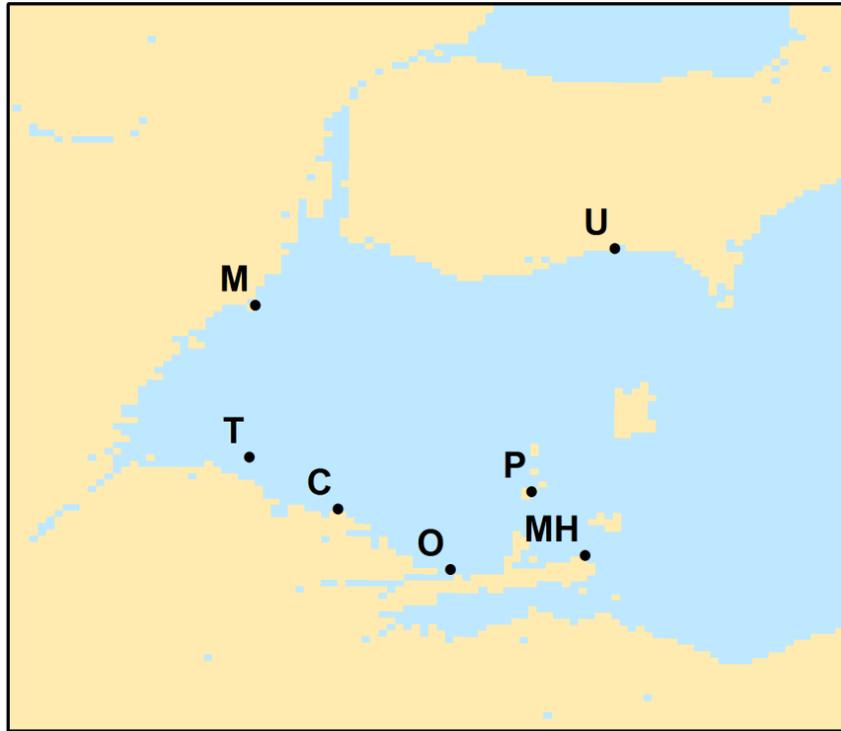


Figure 2-13: Pictorial overview of drinking water intakes of interest in the western basin of Lake Erie [105] (M=Monroe, T=Toledo, C=Carroll Township, O=Ottawa County, P=Put-In-Bay, MH=Marblehead, U=Union County)

Remote sensing data is also valuable to help predict future HAB conditions. NOAA GLERL uses a general hydrodynamical model, coupled with the remote sensing data, to predict how blooms will migrate. Their end data product, called the HAB tracker, gives a five day forecast, which gives water treatment managers advance warning to prepare additional water treatment processes.

Note that cyanobacterial abundance can also be approximated by measuring other water quality properties such as turbidity or dissolved phosphorous concentration. However, these measures are indirect, and for this reason, the focus remains on fluorescence and reflectance-based methods.

2.2.5 Other Great Lakes Data Streams

It is worthwhile to mention several more of the existing HAB-related data streams that exist in the Great Lakes, with an emphasis on Lake Erie. The National Data Buoy Center (www.ndbc.noaa.gov) compiles buoy data from the National Weather Service (NWS), National Ocean Service (NOS), National Oceanic and Atmospheric Association (NOAA), and private companies (e.g. Limnotech), among others (see Fig. 2-14). Note that Limnotech 45165 is placed near the Toledo water intake and has an onboard YSI fluorometer mounted at a depth of 2 ft. Additional NOAA GLERL buoys have been also been placed in the western basin of Lake Erie (see Fig. 2-15). These buoys have varying measurement capabilities, but may be used to measure wind speed and direction, current speed and direction, solar irradiance, water temperature, etc. All of these buoys have data available in real-time.

Much current Lake Erie environmental data, as well as historical data, is available using the interactive map data portal, Great Lakes Observing System (GLOS - <https://www.glos.us/>). NOAA GLERL has also deployed an Environmental Sample Processor (ESP) near the Toledo water intake which can detect microcystin toxins in near-real time. This ESP collects and concentrates particulate from water samples and can run onboard molecular diagnostics.

The data streams presented thus far have focused on measuring algae, but there are also many data sources that measure phosphorus levels and river flow in Lake Erie tributaries, e.g. Heidelberg University's Tributary Loading program and the USGS river gauges. There is also an interactive map data portal called the Western Lake Erie Nutrient Source Inventory (NSI) that includes information on sources of nutrients in the western basin's watershed.

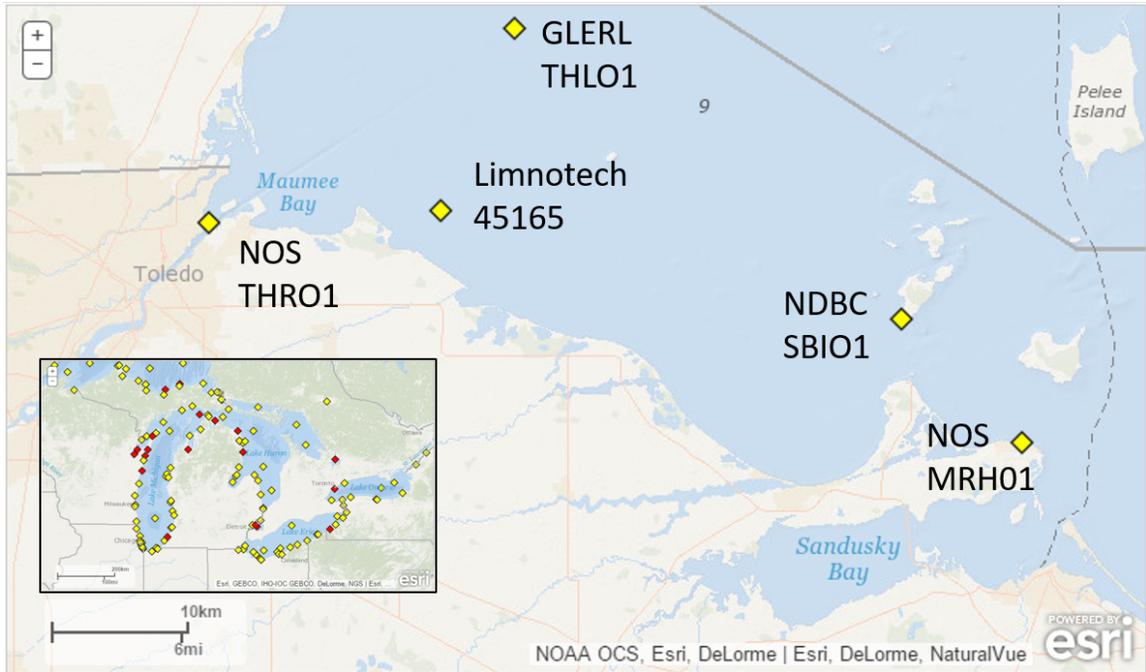


Figure 2-14: National Data Buoy Center (NDBC) buoys in the western basin of Lake Erie and throughout Great Lakes (inset). The Limnotech 45165 buoy is located near the Toledo Water Intake.

2.2.6 Cost of Treatment

A 2015 US EPA report compiled all available published information on the external costs of nutrient pollution impacts [37]. These costs can be classified as either economic losses in the form of tourism and recreation, commercial fishing, and property values, or increased costs stemming from human health, drinking water treatment costs, mitigation costs in lakes, and restoration. Note that comprehensive estimates for these costs cannot be calculated because they are not being documented, so analysis is necessarily limited to the studies that have been performed. The following studies that are presented are limited to fresh-water bodies within the United States.

A pair of studies centered on the economic impact of recent bloom in Grand Lake St. Mary's [25, 26] found that local businesses have lost \$37-47 million in revenues, several local marinas and boat dealers have gone out of business, a nearby state

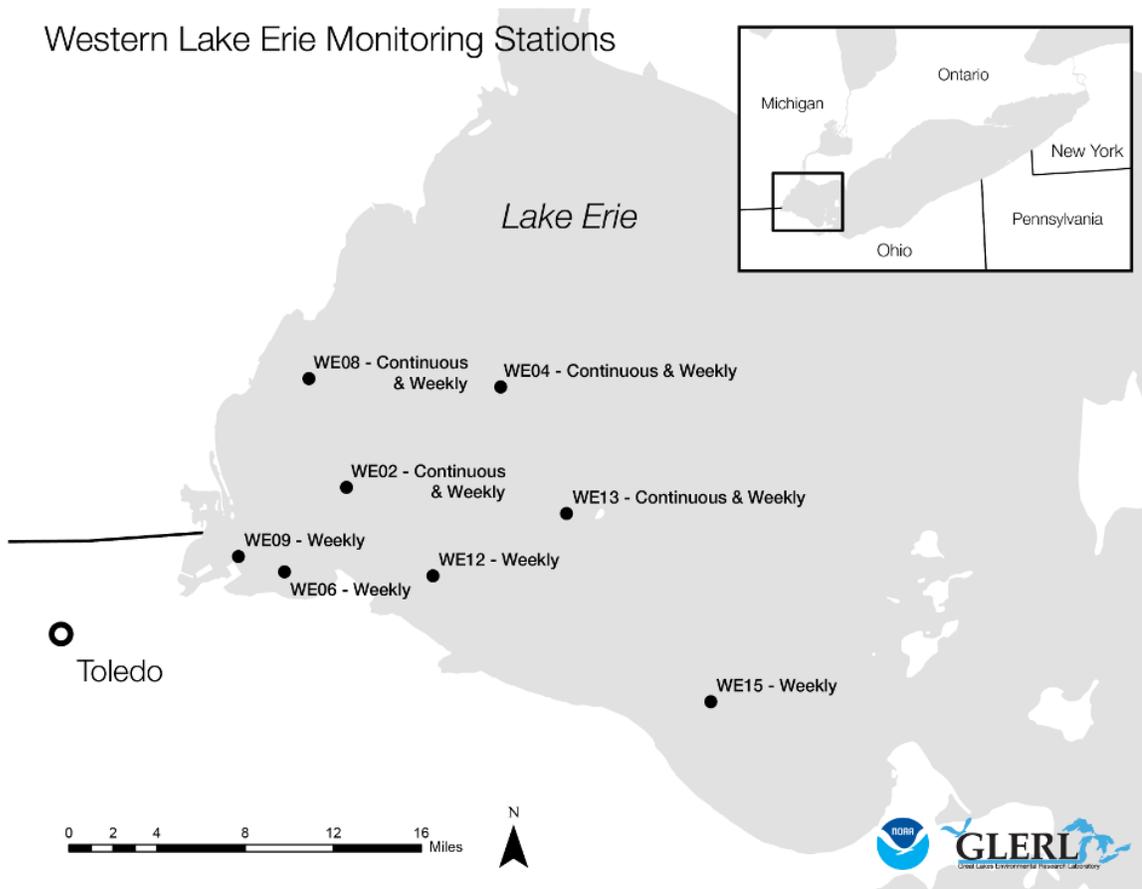


Figure 2-15: NOAA Great Lakes Environmental Research Lab (GLERL) continuous monitoring stations and weekly test sites in the western basin of Lake Erie

park has lost \$260,000 in revenues, a canceled regatta resulted in a \$632,000 loss, and there were 23 reported cases of human illness/dog deaths. The Celina water treatment plant, which draws water from this lake, has conservatively spent \$13.1 million additional dollars, of which \$3.6 million was total operations and maintenance costs for the installation of treatment controls and the setup of toxin testing.

A study of property values along 18 Ohio beaches [4] found that property values decreased by 1.93% for every 1m change in water clarity at homes located the average distance from a beach (13km). As distance to the beach increased, the impact of clarity on value disappeared. Around the headwaters of the Mississippi river in Min-

nesota, a study of lakefront property sales around 37 lakes found a one meter change in water clarity resulted in a price change between \$1,678 and \$84,749 depending on the location/market [55]. A study of two Wisconsin lakes prone to algal blooms compared the value of homes located on these lakes to non-bloom lakes [53]. They found there was a \$128-\$402 decrease in value per shoreline foot from the non-bloom to bloom-prone lakes.

The EPA report included many estimates for algae mitigation efforts in lakes, including applying alum, applying copper sulfate, adding aeration, distributing barley straw around the shore, dredging, herbicide application, biomanipulation, and hypolimnetic withdrawal. As it is difficult to assess if these are only feasibility estimates, rather than costs of actual interventions, these numbers are not listed. Exact numbers are known for alum applications to Grand Lake St. Mary's, which includes \$3.5 million to treat 4,900 surface acres (40% of the lake) in 2011 and an additional \$5 million in 2012.

With this EPA report, and some of these specific studies as a starting point, additional information was gathered on three regional water treatment (WT) facilities which are all impacted by HABs. The Toledo, Bowling Green, and Celina WT plants were chosen and their unique water sources, daily capacities, and treatment options mean that each must be considered individually.

The first WT facility is Toledo's Collins Park WT plant (80 million gallons per day (MGPD)) which pulls water from an offshore water intake in Lake Erie. From there, water runs to a low-pressure pumping station, which pumps the water to the WT plant. The plant's main interventions for high algae levels are potassium permanganate, which is introduced already at the water intake, and powder activated carbon (PAC), which is introduced at the low pumping station. Because of the relatively long distance between the intake, low pressure pumping station, and WT plant, these chemicals have several hours of contact time to begin pre-treating the water.

This environment would place the most technical demands on a robot swarm, because of regularly rough wave conditions and the need to avoid other manned vessels.

The second WT facility is Bowling Green's WT plant (5 MGPD) located on the Maumee River just upstream of Waterville, OH. This facility pumps water from the Maumee river into a reservoir. From the reservoir, the water is pumped into the main treatment facility. This facility has the luxury of a sizable reservoir (25 acres) with a 45 day raw water supply, into which they can apply an algaecide to kill the majority of algae before entering the treatment facility. Starting in early May, about 1000 lbs week⁻¹ of this hydrogen peroxide-based algaecide (PAK-7) is applied, costing \$34,000-\$40,000 per year (2016-2017). The facility also uses granular activated carbon (GAC), which costs \$216,000 per year (2016-2017). Potassium permanganate is also used by the facility, but it is not dosed to the algae concentration, and is only used for mussel control. A robot swarm would be the easiest to implement for this type of facility, with a dedicated reservoir, but operation would have to cost less than the current algae-related consumables, and provide an equivalent algae removal rate to be competitive.

The third and final facility is the Celina WT (1.3 MGPD) plant located on the NW corner of Grand Lake St. Mary's. This plant is one of the most advanced in the state, and it uses this unique series of treatment processes: clarification, ozonate, lime soften, recarb, ozonate (2nd time), sand filtration, GAC, UV, chlorinate. This plant also uses potassium permanganate (\$140,000 annually), and the cost of GAC is \$300,000 annually. A robot swarm could be implemented at this type of facility, but care would need to be taken to segregate the swarm from recreational activities on the lake. Like for the other plants, the cost of the swarm would have to cost less than the current algae-related consumables to be competitive.

Chapter 3

Sizing, Scalability, and Control of a Robot Swarm to Mitigate Harmful Algal Blooms

This section is taken from a manuscript that is being prepared for submission to a journal. It is authored by Adam Schroeder, Brian Trease, and Alessandro Arsie. After the conclusion of this manuscript, a supplemental section has been added with additional information.

3.1 Abstract

A robot swarm is proposed as a novel solution to the global threat posed by harmful algal blooms. To inform the design of a swarm to physically collect the harmful algae, three sets of simulations are performed. The rate of algae collection is studied relative to: the diffusion of the algae in the water (due to turbulence), the composition of the swarm as either few large robots or many small robots, and the quantity of constant-size robots in the swarm. The results are shown as plots of uncollected biomass vs. time, and as maps of the algae distribution after the robots

have begun collection. Both a partitioning and non-partitioning controls approach are taken, which impose different hardware and communication requirements on the robots. This work is being done in parallel with the development of physical robots that can strain and skim algae from the water.

3.2 Introduction

Despite having enormous potential, there are few examples of robot swarms being developed as engineering solutions to real-world applications. This work focuses on designing an aquatic swarm to mitigate harmful algal blooms (HABs), which are a recurring global threat to the world’s water supplies. These blooms can be massive, and they contain toxins which can be harmful to human and non-human health. There are limited physical, chemical, and biological methods for combating these blooms, but a robot swarm that could be scaled up to the size of the bloom would unlock new treatment capabilities.

The development of a robot swarm for this purpose can be divided into control development, highlighted in this work, and development of the physical robot platform. The physical platform is being developed in parallel, including technologies for physically removing algae from the water using straining and skimming processes. An image of both a small prototype for straining algae from water and a larger prototype for skimming surface scum is given in Fig. 3-1.

To aid in the development of this swarm, several sets of simulations are used to inform the swarm designer about (i) the ability of the swarm to collect a diffusive substance, i.e. the algae, (ii) whether many small robots or fewer large robots should be used, and (iii) what scaling effects are present as more and more robots are added to the swarm. Collecting a diffusive substance is an interesting problem, although somewhat specific to this application, but the second and third points are generally

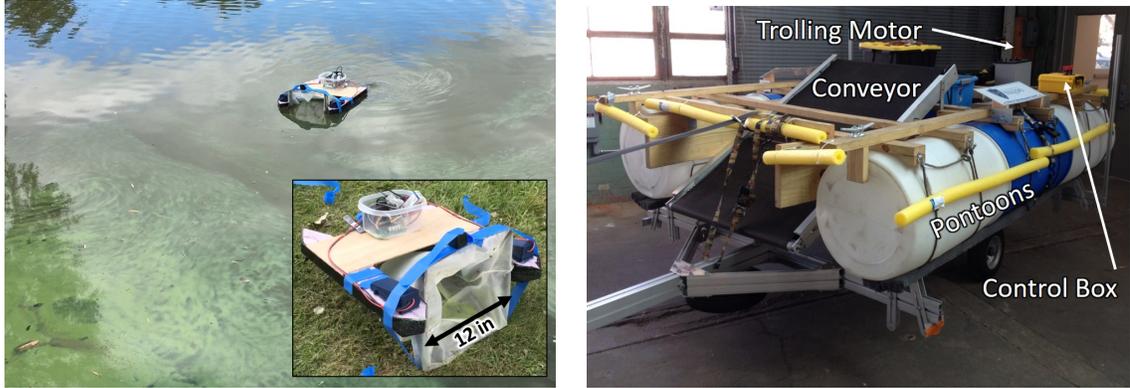


Figure 3-1: (Left) Small platform for straining algae from below the surface using a plankton net. The green harmful algae is visible on the surface and residual algae is visible in the net. (Right) Large platform for skimming algae surface scum using a conveyor. Only single units have been tested thus far as the collection technologies are developed.

applicable to any swarm.

The following sections give a background of harmful algae, techniques for collecting it, and examples of aquatic swarms (Section 2), explain how the problem was formulated and which simulations were performed (Section 3), share the results from those simulations (Section 4), discuss the implications of the results (Section 5), and then offer some concluding thoughts and comments on future work (Section 6).

3.3 Background

Lake Erie serves as a good reference point with respect to harmful algal blooms, as they have been occurring and have been studied for many years. Even in one of the least severe years (2007), the bloom had an areal footprint of 370 km^2 . In one of the most severe years (2011), the areal extent was 1800 km^2 [50]. The type of harmful algae, *Microcystis aeruginosa*, is actually a cyanobacteria which forms spherical colonies ranging in size from $20\text{-}700 \mu\text{m}$ [110]. These colonies are suspended in the water and largely move with any water currents. Their concentration is usually

measured in biomass or biovolume.

The physical collection of algae is a challenging problem in itself. Field biologists collect algae using a plankton net [13], similar to the one mounted on the small platform from Fig. 3-1, but only a fraction of the algae present can be collected. Some colonies go around the net and others are small enough to move through the net. A net with a smaller mesh size would collect smaller colonies but also create a higher drag force. For this work, it is assumed that 90% of the algae within the conveyor width or plankton net width is collected as the robot moves forward and that the robot leaves a defined, cleared path behind it. Turbulence in the water causes algae to slowly diffuse into the previously-cleared area. For simplicity, any additional mixing or hydrodynamics caused by the motion of the boat is not considered.

There are a few examples of aquatic robot swarms, with the first example coming from Duarte et al. [32], which has been demonstrated for environmental monitoring applications. They demonstrated homing, dispersion, clustering, and area coverage using as many as ten robots. A second notable example of an aquatic swarm comes from Zoss et al. [111]. Their ‘Bunch of Buoys’ project used a ‘mobile buoy’ capable of environmental monitoring and other swarm behaviors such as flocking and area coverage. They performed field experiments with up to 45 buoys. The final example of an aquatic swarm, called ‘SeaSwarm’, was intended to absorb oil after an oil spill [1], although it is unclear if more than one robot was constructed. Collecting harmful algae using a robot swarm builds on these works, but also presents its own unique challenges.

There are many controls possibilities for this task, with differing communication, sophistication, and sensing requirements. Two controls approaches are considered in this work. The first approach is for the robots to perform a biased random walk in the direction of highest algae concentration. This does not require any explicit communication, but it does require that the robots have data on their local algae

concentration. In practice, this could be acquired using a fluorometer, which is an in-situ sensing instrument, or by using a remote sensing technology. The second approach is for the robots to partition the domain into individual regions of responsibility. Once partitioned, an individual robot would parse back and forth, exhaustively covering their assigned region. This approach would only require communication at the beginning, when partitions are established, or if partitions need to be adjusted or regenerated. This approach would not require any algae sensing capability. Neither of these controls approaches in themselves are novel, but they will help give the swarm designer intuition about this particular application.

A biased Lévy flight was chosen for the first approach, hereafter referred to as the non-partitioning approach (see supplemental material, Section 3.8.2). This is a type of random walk that uses frequent short steps and occasional very long steps. In general, Lévy flight has been an active area of swarm robotics research for search tasks [28, 98, 72], and area coverage and interception tasks [89, 90]. This biased walk in the direction of highest algae concentration is a form of chemotaxis, or movement in response to a chemical. The robots can indirectly communicate about where they have been, because algae has been cleared from that area.

A Voronoi partitioning scheme was chosen for the second approach, hereafter referred to as the partitioning approach, following the example of Pavone et al [75] (see supplemental material, section 3.8.3). They describe how a domain could be equitably partitioned using Lloyd’s approach [61], used in this work, and other approaches. It has been demonstrated that this process can be decentralized [15].

Referring back to the research questions from the introduction, there is no literature studying the collection of a diffusive substance, although some research studies the localization of a source using a gradient [43, 92]. To the authors’ knowledge, there is not any literature on comparing the performance of few large robots to that of many small robots. Finally, there is literature that studies the relative benefit

gained by adding more robots for many different applications, e.g. foraging [47], area coverage [40], or the amount of interference between robots [59, 83]. However, this study is unique in that it varies the robot quantities over several orders of magnitude.

3.4 Problem Formulation and Implementation

Three sets of simulations were performed, with a list of the most important parameters given in Table 3.1. These simulations address the research questions from the introduction, and are referred to as: *Diffusive Collection*, *Robot Size vs. Quantity*, and *Robot Density*. For all simulations, robots move through water at a speed of one unit per second and collect 90% of the algae that they pass over. Note that algae actually has some control over its buoyancy, and its depth typically varies as part of a diurnal cycle [3]. For these simulations, which have been reduced to a 2D problem, it is assumed that the algae is near enough to the surface to be collected.

Table 3.1: Overview of the main parameters of each simulation set

Study		Domain Size[m]	Spatial Step	Time Step	Diff. Type	Simul. Time	Run Time [hrs] Mean	Run Time [hrs] Set
Diffusive Collection ¹	Part.	100x100	0.1 m	0.1 s	ADI	10 000 s	1.3	5.0
	Not-Part.						1.1	4.3
Size vs. Quantity ²	Part.	100x100	≤ 0.1 m	0.1 s	FTCS	10 000 s	1.8	8.8
	Not-Part.						0.7	4.0
Robot Density ³	Part.	800x800	0.1 m	0.1 s	FTCS	10 000 s	5.2	20.7
	Not-Part.						4.2	16.7

¹ ADI scheme is computationally more expensive, but more accurate than FTCS

² Finer spatial step for many, small robots (0.025 m for 40 robots) takes more time

³ For 10 000 robots, only 1000 seconds simulated time required.

The algae is treated as a diffusive substance, with the turbulence in water causing the diffusion (see supplemental material, section 3.8.1). Diffusivity values can be

estimated based on the size of bloom or algae patch by applying a classic rule for estimating diffusivity, called the ‘4/3 power law’ [82]. Unless noted, a diffusivity of $10^{-4} \text{ m}^2 \text{ s}^{-1}$ was used. Either an explicit forward-time, center space (FTCS) discretization scheme or a Crank-Nicholson alternating direction implicit (ADI) scheme were used to numerically simulate diffusion (see supplemental material, Section 3.8.8), and an NVIDIA K6000 GPU accelerated this portion of the simulation, with total simulation run times given in Table 3.1 (see supplemental material, section 3.8.7).

The initial algae concentration was set to $50 \mu\text{g m}^{-2}$, which represents a ‘relatively high’ concentration. Unless noted, a robot collection width of 1 m was used, robots operated in a 100x100 m domain, 10 000 s of time was simulated, and the algae concentration in the domain was tracked with 0.1 m resolution.

These simulations are necessarily a gross simplification of an environment that is hydrodynamically, biologically, ecologically, and meteorologically complex. This simplification is necessary to make the simulation tractable and to extract higher-level lessons pertinent to the swarm designer.

Recall that for each simulation set, both a biased Lévy flight and a partitioning approach were used. Collisions between robots were not tracked, although the collision rate for an unbiased Lévy flight is being investigated in parallel work that includes the effects of number and size of robots. For brevity, and because the controls approaches themselves are not novel, a detailed mathematical development for each is omitted (see [90, 75] for details). The exact parameters used are given, and for reproducibility, the MATLAB scripts for these simulations are available¹. Details specific to the three individual studies will now be introduced.

Diffusive Collection: Applying the 4/3 power law, diffusivity values ranging from 10^{-2} to $10^{-5} \text{ m}^2 \text{ s}^{-1}$ are possible (see supplemental material, Section 3.8.9). The

¹MATLAB script available at <https://www.mathworks.com/matlabcentral/fileexchange/67071-robot-swarm-harmful-algae-collection>

diffusivity was varied by an order of magnitude within this range for this simulation set. Faster diffusion than this would be caused by very large eddies in the water, which would be powerful enough to move the robot boats along with the algae. Ten robots, each with a one-meter-wide collection width, were used for this study.

Robot Size vs. Quantity: A constant swarm capacity of 10 m was assigned, where capacity refers to the number of robots multiplied by the collection width of each robot. The capacity could be allocated to as few as 2 large robots (5 m-wide), or as many as 80 small robots (0.125 m-wide). The algae concentration was tracked with a resolution of at least 1/10th the width of a robot, e.g. 0.0125 m for the case of 80 robots.

Robot Density: Robot quantities of 10, 100, 1000, and 10 000 were simulated, now within a larger 800x800 m domain so that a larger range of robot densities could be simulated.

3.5 Results

Select results for each of the sets of simulations are shown below, each with a plot of uncollected biomass vs. time. At the outset, the uncollected biomass is the initial concentration ($50 \mu\text{g m}^{-2}$) times the working domain (either 100x100 or 800x800 m^2). For each plot, dashed lines show the results of the partitioning approach, and solid lines show the results from the non-partitioning approach.

Diffusive Collection: Figure 3-2 shows how quickly algae can be collected for diffusion rates that vary by orders of magnitude. Fig. 3-3 shows the final algae distributions for each of the diffusion rates for the non-partitioned case. Algae is collected more quickly when the diffusion rate is low which means that robots only

need to cover an area one time. The partitioning approach clearly outperforms the non-partitioning approach for any rate of diffusion. The partitioning approach is guaranteed to exhaustively cover the area, which is especially beneficial when the rate of diffusion is low. Distinct changes of the line slope in Fig. 3-2 for the partitioning approach indicate when the robots have finished covering their own partition for the first time and begin to re-cover their partition.

The larger lesson from this set of simulations is that a diffusive substance can be collected regardless of diffusion rate, and that even an approach that does not exhaustively cover the area can still collect algae in a reasonable amount of time.

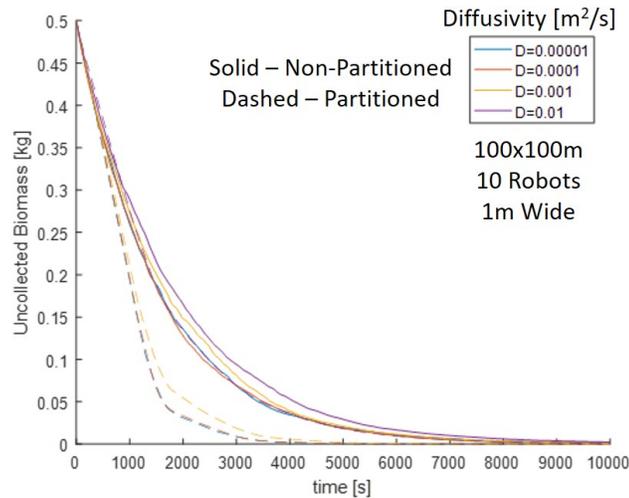


Figure 3-2: Collection performance over a range of diffusivity values for a non-partitioning scheme (solid) and a partitioning scheme (dashed)

Robot Size vs. Quantity: Figure 3-4 shows how quickly algae can be collected by few large robots or many small robots. Figure 3-5 shows the final algae distributions for the non-partitioning case for each robot quantity/size combination. The non-partitioning approach was not very sensitive to the quantity/size combination until the quantity of robots was very low. For very few robots, if even one of the robots is not operating in area with a high concentration of algae, the entire system suffers.

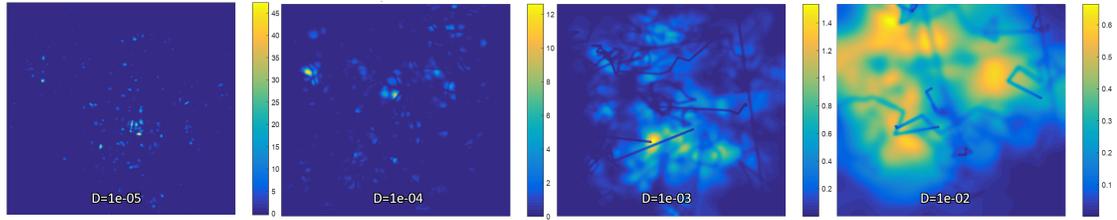


Figure 3-3: Final algae distributions in the non-partitioned case after 10 000 s with diffusion rate (D) increasing by orders of magnitude from left to right. The effect of the faster diffusion rate is evident in how quickly the robot trails become indistinct. Note that the scale for each map is different.

The partitioning approach collected algae more quickly for quantity/size combinations with few robots. Performance decreased for combinations with higher quantities of robots. Subsequent analysis showed that this is likely due to how partitions were formed. The partitioning process, which occurs before any collection is performed, was continued until the difference between the area of the largest and smallest partitions was below a predefined value. This threshold worked well for fewer robots, but for higher quantities of robots, it allowed some unequally-sized partitions to form, which decreased how quickly algae could be collected. It took robots in the largest partitions more time to complete their first pass, which is when the majority of the algae is collected, while robots in the smaller partitions were already re-covering their partition and collecting less algae.

As before, the partitioning approach outperformed the non-partitioning approach. The larger lesson learned here is that the quantity/size combination did not have a significant effect on collection rate, as long as the number of robots was not too low and as long as the size of the partitions was similar.

Robot Density: Figure 3-6 shows how quickly algae can be collected as the number of robots increases by orders of magnitude. Figure 3-7 shows the final algae

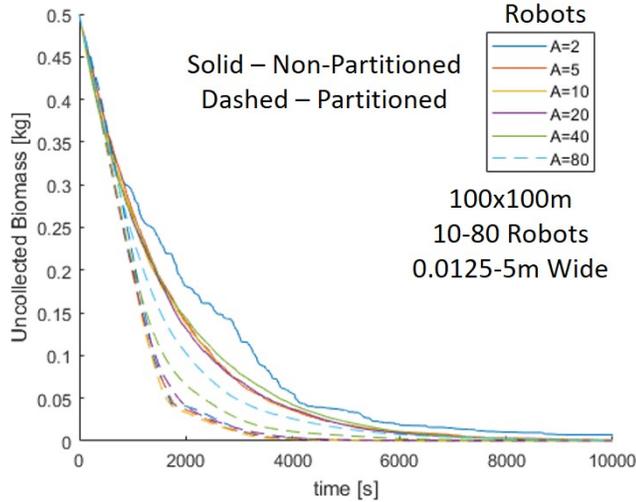


Figure 3-4: Collection performance over a range of robot quantities/collection widths for a non-partitioning scheme (solid) and a partitioning scheme (dashed). A constant system collection width of 10 m was maintained, where, for example, the 5 robot case would use a 2 m-wide robot collection width.

distributions for the different robot quantities for the partitioning case. Unlike the previous study, the size of the robot remains constant as the quantity of robots varies. It is expected that algae is collected more quickly as more robots are added, but it is interesting to see if there are diminishing returns as more robots are added. At the lowest and highest quantities, there is not much difference between the partitioning and non-partitioning approaches. The partitioning approach does collect algae more quickly for more reasonable quantities of robots.

It is more valuable to view these results with the time shown on a log scale, as in Fig. 3-6 (right). This shows how, for the partitioning case, there are diminishing returns as more robots are added. This is attributed to how the robots behave at the boundary of their partitions. Each robot completely covers the boundary of its partition, which results in some overlap into the neighboring partitions. This redundancy becomes more of an issue as the number of partitions increase, because the relative ratio between partition perimeter and area increases as the number of

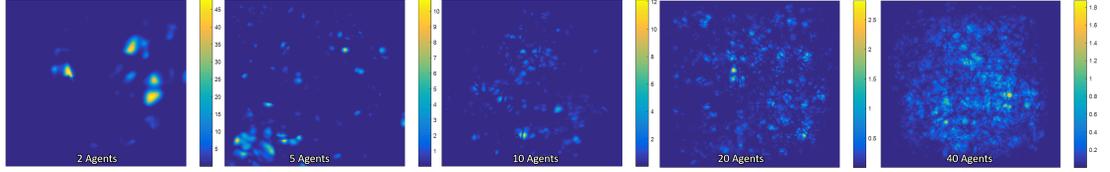


Figure 3-5: Final algae distributions in the non-partitioned case after 10 000 s with number of robots increasing and collection width decreasing from left to right. It is clear that the case of two large robots missed collecting several larger patches of algae. Note that the scale for each map is different.

partitions increases, as shown in Table 3.2. Simply put, as the number of robots increase, and the number of partitions increase, the size of each partition decreases, and each robot spends more time at its partition’s boundary, which causes more redundant coverage.

Table 3.2: Partition perimeter to area ratio as robots’ partition area varies

Robots	10	100	1000	10 000
Area / Robot	64000	6400	640	64
Perimeter ¹	942	298	94	30
Perimeter/Area	0.015	0.047	0.147	0.465

¹ Assumes partition is a regular hexagon. Observed partitions are either irregular pentagonal or hexagonal (Fig. 3-7).

3.6 Discussion

The higher-level takeaways from these simulations are that either a partitioning or a non-partitioning approach can effectively collect a diffusive substance, such as algae. In almost all cases, the partitioning approach is more effective, but it does require robots to know their neighbors’ positions to form the initial partitions or to adjust partitions due to robots entering or dropping out of the swarm. Perhaps even more difficult, the partitioning approach would require the robots to navigate in straight

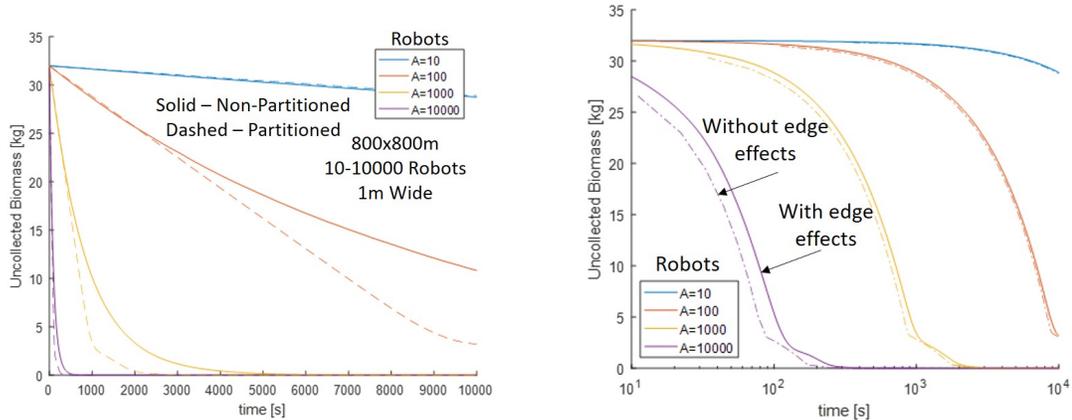


Figure 3-6: Collection performance over a range of robot densities on a linear (left) and log (right) time scale. The log-scale facilitates comparison across orders of magnitude (only the partitioning results shown). The results are compared to a simplified case (long/short dashes) where overlapping coverage across partition boundaries is not considered (see supplemental material, Section 3.8.5).

lines with minimal overlap as they parse back and forth amid wind and waves.

These realizations make the non-partitioning more attractive, but one of the major drawbacks to this approach is the need for each robot to continuously detect its local algae gradient. Shaukat and Chitre elucidated how detecting a gradient [92], in their case to localize a source, can either use multiple sensors on the same robot (instantaneous gradient detection), or only a single sensor (temporal gradient detection). However, due to sensor cost (fluorometer >\$2000), a multi-sensor solution is undesirable. Furthermore, the noisiness of the algae distribution and the disruption from the robot itself moving through the algae may make detecting any gradient a challenge. With these considerations, a robot swarm that performs *unbiased* Lévy flight (no sensors), instead of the previously-used *biased* Lévy flight is of interest. Another simulation set for the diffusive collection was performed, and the new results (dotted lines) overlaid with the previous results, are shown in Fig. 3-8.

There is not a significant degradation in performance even when the robots cannot sense their local algae concentration. This type of relatively unsophisticated control

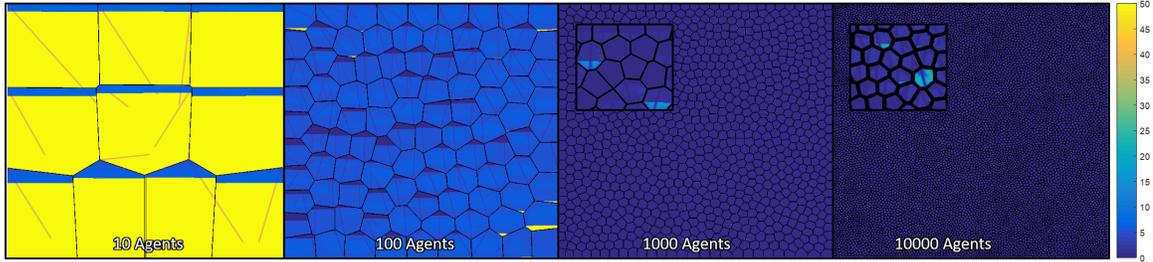


Figure 3-7: Final algae distributions in the partitioned case after 10 000 s with number of robots increasing by orders of magnitude from left to right. For 1000 and 10.000 robots, a magnified portion of the domain is also shown. The larger partitions have not been covered completely or as many times as the smaller partitions.

approach is easily implemented with the existing platforms.

These simulations did not consider the possibility of a collision between robots. Research has shown sublinear performance improvement as the number of robots increases, due to interference between robots [47, 83]. It is expected that the number of collisions between robots increases as the number of robots increases, but it would be interesting to see if the partitioning or non-partitioning approach handles higher robot densities more gracefully than the other. For the partitioning case, it has already been observed how the robot spends more time at its partition's boundaries as more robots are added. This results not only in redundant coverage and decreased collection, but it would also result in a higher rate of collisions.

As previously noted, this simulation environment is a very complex system. These simulations only included the effect of algae diffusion, due to turbulence in the water, but not advection, due to prevailing currents. This work also started with a homogeneous algae distribution. In reality, algae is sometimes distributed homogeneously, but it is often present in heterogeneous streaks and patches. This is evident even in Fig. 3-1. The inclusion of collision tracking, algae advection, and the possibility of a heterogeneous initial algae distribution are three areas targeted for future work.

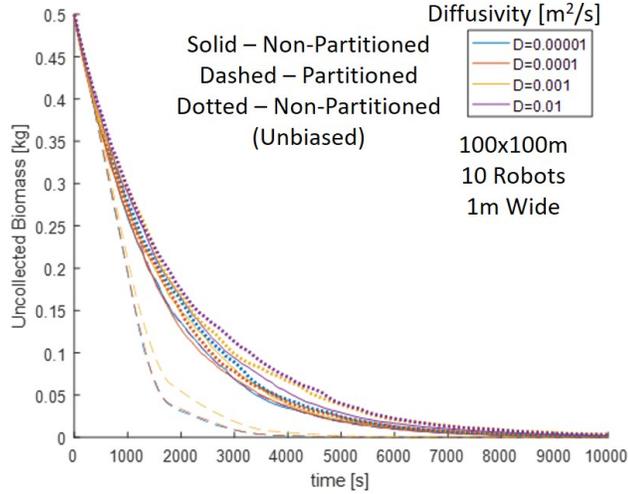


Figure 3-8: In addition to the *biased* non-partitioning (solid) and partitioning (dashed) results shown over a range of diffusivity values (see Fig. 3-2), results are also shown for *unbiased* Lévy collection (dotted). There is little degradation using the unbiased Lévy flight, which does not require any algal sensors.

3.7 Conclusion

Simulations were performed to assist a swarm designer in the task of constructing a robot swarm to mitigate harmful algal blooms by physically collecting the algae. These simulations were targeted at gaining application-specific understanding of collecting a diffusive substance (algae). They were also targeted at obtaining a more generally applicable understanding of the impact of the robot quantity/size combination, and varying the robot quantity while keeping the robot size constant. Two controls approaches were implemented, partitioning and non-partitioning, which impose different hardware and communication requirements. The easiest-to-implement approach would be a non-partitioning control scheme where robots simply collect algae while performing a random walk, without the requirement of sensing their local algae concentration.

The larger lessons learned from the simulations are that **(i)** algae can be collected efficiently, regardless of the diffusivity, **(ii)** an approach that does not exhaustively

cover the domain can still effectively collect algae, **(iii)** the robot quantity/size combination did not strongly influence collection, unless the number of robots was very low, and **(iv)** in the partitioning approach, robots spend more time at their partition boundaries as the partitions shrink, which causes more coverage redundancy and should lead to more collisions.

Overall, using a robot swarm to collect harmful algae merits further study and development. Future work will concentrate on enhancing these types of simulations by including collision tracking, advection due to currents in the water, and non-homogeneous algae distributions. In parallel, physical platforms are already built and will be tested on Lake Erie over the 2018 HAB season.

3.8 Supplemental Info

Additional information that supplements the content of the manuscript being prepared is included here.

3.8.1 Mathematical Problem Formulation

The algae in this system may be modeled as a diffusing and advecting population, as given by (3.1). The first term governs diffusion and the second term governs advection. Here, $a(r, t)$ is the distribution of agents, and $b(r, t)$ is the distribution of algae, which varies in space, $r \in \mathbb{R}^2$, and time, t . D is the diffusion coefficient, but note that the diffusion of algae is not molecular, but rather driven by turbulence in the water, which can be simplified to the same form when assuming eddy turbulence [82]. The prevailing current in the water is given by \vec{v} . Note that in this most general form, the diffusion coefficient and current are not assumed to be constant over the entire field.

$$\frac{\partial b(r, t)}{\partial t} = \nabla^2 D b(r, t) - \nabla \cdot (\vec{v} b(r, t)) \quad (3.1)$$

With this as a basis, an additional term needs to be added to account for the collection of algae by the agents, which is generically defined as function $F_{collect}$, as given in (3.2). The arguments of this generic function include η , the collection efficiency, which concedes that only some fraction of the algae encountered may be collected, a_{width} , the collection width of an individual agent, and the distribution of agents and algae, $a(r, t)$ and $b(r, t)$, respectively.

$$\frac{\partial b(r, t)}{\partial t} = \nabla^2 D b(r, t) - \nabla \cdot (\vec{v} b(r, t)) - F_{collect} \quad (3.2)$$

The collection efficiency, η , in the general case, could capture the complex hydrodynamics of the collection process, but in a more simplified form is assumed to be a uniform ratio of algae collected to algae present, within the treatment area. Another assumption is made that collection is not performed unless the agent is moving through the water. With these assumptions, a more specific formulation of the $F_{collect}$ function is given in (3.3). The term \hat{a} is introduced, which is a binary scalar field set to one in areas where an agent passed in the last time step, and zero otherwise.

$$\frac{\partial b(r, t)}{\partial t} = \nabla^2 D b(r, t) - \nabla \cdot (\vec{v} b(r, t)) - \eta \hat{a}(r, t) b(r, t) \quad (3.3)$$

At this point, it is insightful to perform a dimensional analysis to show how the units from this added collection term are compatible with the diffusion and advection terms. The diffusion coefficient and current have units of $\frac{m^2}{s}$ and $\frac{m}{s}$ in 2D, respectfully, the distribution of algae, $b(r, t)$, is a scalar field with units of $\frac{kg}{m^2}$, and the term \hat{a} has units of s^{-1} . The divergence operator, $\nabla \cdot$, converts a vector field to a scalar field, the gradient operator, ∇ , converts a scalar field to a vector field, and the Laplacian operator, δ , or ∇^2 , which is equivalent to the divergence, $\nabla \cdot$, of the gradient,

∇ , returns a scalar field. These operators are progressively applied to show the dimensional equivalence of these three diffusion, advection, and collection terms. The appropriate units are directly placed in (3.2), which is shown in (3.4), and all vector terms are identified with $[[\cdot]]$ for clarity. In the first step, the gradient portion of Laplacian is applied to the first term and the divergence operator is applied to the second term. In the second step, the divergence operator is applied to the first term which reveals the equivalence of all terms.

$$\begin{aligned}
\left[\frac{kg}{m^2s} \right] &= \nabla^2 \left[\frac{m^2}{s} \right] \left[\frac{kg}{m^2} \right] - \nabla \cdot \left(\left[\left[\frac{m}{s} \right] \right] \left[\frac{kg}{m^2} \right] \right) - \left[\frac{1}{s} \right] \left[\frac{kg}{m^2} \right] \\
&= \nabla \cdot \left(\left[\left[\frac{kg}{ms} \right] \right] \right) - \left[\frac{kg}{m^2s} \right] - \left[\frac{kg}{m^2s} \right] \\
&= \left[\frac{kg}{m^2s} \right] - \left[\frac{kg}{m^2s} \right] - \left[\frac{kg}{m^2s} \right]
\end{aligned} \tag{3.4}$$

One final simplification to (3.3) assumes that the diffusion coefficient and current are constant across the entire field as given in (3.5).

$$\frac{\partial b(r, t)}{\partial t} = D \nabla^2 b(r, t) - \vec{v} \cdot \nabla b(r, t) - \eta \hat{a}(r, t) b(r, t) \tag{3.5}$$

3.8.2 Non-Partition Approach: Stochastic Gradient-Following

In the first approach, agents use a mixture of algae gradient-following and random noise to move through the domain. Equation (3.6) gives the n th agent's velocity, \dot{R}_n , where $\frac{\alpha_n}{\|\alpha_n\|}$ (the normalized contribution from the vector sum of the gradient-following and noise) is then multiplied by a constant magnitude, ϵ . Equation (3.7) is a Lévy flight biased random walk where χ is the sensitivity to the gradient, and σ has a constant magnitude and a uniformly distributed random direction.

$$\dot{R}_n = \epsilon \frac{\alpha_n}{\|\alpha_n\|} \tag{3.6}$$

$$\alpha_n = \chi \nabla b(r, t) \Big|_{R_n} + \sigma \quad (3.7)$$

For a Lévy flight biased random walk, each segment of a path is sampled from a heavy-tailed distribution. To accomplish this, a uniformly distributed random number, u , from $[0,1]$ is first generated. This type of random number is transformed into another distribution using the inverse cumulative distribution function. A cumulative distribution function, Q , is defined for which the path length is $s = Q^{-1}(u)$. For a power-law distribution, which is heavy-tailed, the cumulative distribution function is defined in (3.8) where s_{min} is the minimum path length, set to three, and $\alpha_{Lévy}$ is a value from one to three. The inverse cumulative distribution function in (3.9) will then give the path length.

$$Q(s) = 1 - \left(\frac{s}{s_{min}} \right)^{-\alpha_{Lévy}} \quad (3.8)$$

$$Q^{-1}(s) = s_{min}(1 - u)^{-1/\alpha_{Lévy}} \quad (3.9)$$

3.8.3 Partition Approach: Modified Voronoi Decomposition

In the second approach, agents start by distributively partitioning the domain into individual regions of responsibility. This iterative process takes place before any algae collection is begun and partitions are then fixed. The initial partitions are determined based on the initial random location of each agent. Each point in the domain that is closer to that agent than any other agent is added to that agent's partition, which is called Voronoi partitioning. This is given in (3.10) where P_n is the n th agent's partition, r_0 is every candidate point in \mathcal{R} , d is a function that defines the Euclidean distance between each point and that agent, and $R_{n,0}$ is the initial location of each agent.

$$P_n = \{r_0 \in \mathcal{R} | d(r_0, R_{n,0}) \leq d(r_0, R_{i,0}) \forall i \neq n\} \quad (3.10)$$

After calculating initial partitions, the desire is to modify the partitions so that they all contain an equal amount of algae, building from the work of Pavone et al. [75] which is related to the classic Lloyd's algorithm [61]. Following Pavone et al's approach and nomenclature, each agent is equipped with a virtual 'beacon' that is initially located at the agent's initial location. In subsequent steps, instead of partitioning the area based on the agents' locations, the partitioning will be done based on the beacon's locations, $G_n(x, y)$. The beacons are moved based on the \hat{x} and \hat{y} centroid of each agent's partition, which is calculated using (3.11) and (3.12), respectively. These expressions are first developed in a continuous form and then converted to a discretized form. Here, $b(x, y)$ is the initial algae distribution, and λ is a weighting factor, where λ greater than one creates more downstream partitions and λ less than one creates more upstream partitions. This weighting factor is included because it is hypothesized that concentrating the agents on the upstream or downstream side may improve performance when advection is present. Note that weighting was not applied to (3.12) because drift was constrained to one dimension, but it could be implemented in the same manner as (3.6) in the more general case.

$$\hat{x}_n = \frac{\int_{P_n} x^\lambda b(x, y) dx dy}{\int_{P_n} b(x, y) dx dy} = \frac{\sum_{P_n} x^\lambda b(x, y) \Delta x \Delta y}{\sum_{P_n} b(x, y) \Delta x \Delta y} \quad (3.11)$$

$$\hat{y}_n = \frac{\int_{P_n} y b(x, y) dx dy}{\int_{P_n} b(x, y) dx dy} = \frac{\sum_{P_n} y b(x, y) \Delta x \Delta y}{\sum_{P_n} b(x, y) \Delta x \Delta y} \quad (3.12)$$

Each agent iteratively moves its beacon, $G_n(x, y)$ by performing a gradient descent toward the partition centroid, which is calculated at each step, as given by (3.13) and (3.14), until some end conditions are met. The gradient descent is damped or accelerated using a gain, K . Then they parse back and forth, left and right, while

slowly moving from the bottom to the top of their partition. Once it has been completely covered, they restart the cycle at the bottom of their partition. Note that setting the gain value is akin to the process of tuning a proportional controller, which requires some experimentation. A gain of 0.5 was found to perform acceptably.

$$G_{x,n,new} = G_{x,n} - K (G_{x,n} - \hat{x}) \quad (3.13)$$

$$G_{y,n,new} = G_{y,n} - K (G_{y,n} - \hat{y}) \quad (3.14)$$

3.8.4 Analytical Approaches

The equation developed to describe the time and spatial evolution of the algae distribution, (3.5), is quite complex and not able to be solved analytically. Thus, performing numerical simulations is the preferred methodology. However, some simplifying assumptions can be made and the results later compared to the results from the numerical simulations.

The first such simplification assumes that agents are always collecting an average amount of algae, which has a closed-form solution. This relation is given in (3.15), where b is still the algae concentration, but now just a function of time, A is the total domain area, and the other parameters have been previously defined. A simple dimensional analysis is shown in (3.16) to confirm dimensional compatibility. To simplify presentation, several constants can be combined into a single constant, χ . This is a separable equation, with the closed-form solution given in (3.17), where b_0 is the initial concentration and the total algae in the area of interest, $A_{tot} = bA$.

$$\frac{db}{dt} = -\frac{N\epsilon a_{width}\eta b}{A} = -\chi b \quad (3.15)$$

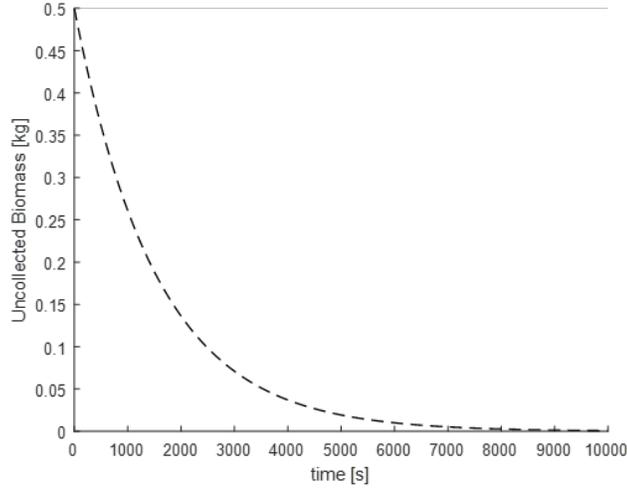


Figure 3-9: Analytical solution of swarm collection problem that assumes agents always collect an average amount of algae

$$\left[\frac{kg}{m^2 s} \right] = \frac{[agent] \left[\frac{m}{s} \right] \left[\frac{m}{agent} \right] \left[\frac{kg}{m^2} \right]}{[m^2]} \quad (3.16)$$

$$b = b_o e^{-\lambda t} \quad (3.17)$$

This analytical solution includes the effect of changing the number of agents, collection width and efficiency, and agent speed but does not include the diffusivity. A representative solution is shown in Fig. 3-9.

3.8.5 Semi-Analytical Approach

A second simplification of the system, applicable to the partitioned case, can be made by solving for only a single partition, using the coarsest spatial and time steps possible, and then extrapolating the results to the larger system. The area of a single partition is set to the area of the entire domain divided by the number of agents and this single partition is assumed to be square. Then a forward time discretization scheme (FTCS) is used with only two nodes for every path segment, each with a

width of the treatment length. Thus, a 10 unit x 10 unit partition with an agent collection width of one unit, would have only 20 nodes (each node 5 units x 1 unit) and diffusion is assumed to only occur from one row to another row (one dimensional). Assuming this one dimensional diffusion is equivalent to saying that diffusion in the same direction as agent movement (along a row) is negligible compared to the agent's velocity, and the corresponding collection. The time step is set to the amount of time needed to move across an entire row, and move down to the next row. Finally, periodic boundary conditions are assumed for the top and bottom boundaries, which would make sense when each partition is bounded by other identical partitions.

Assuming the 10 unit x 10 unit partition, the domain after a single pass is given in Fig. 3-10(left). An agent starts in the upper left and parses through the domain, row by row. Thus, after a single pass, the upper left has the highest concentration, because its had the most time since collection, and the highest gradient (due to the periodic BC), which facilitated the diffusion of algae. This process can be repeated to account for as many passes as desired. It is assumed that it takes no time for the agent to move to the starting position, both at the beginning and after each pass. The differences between this semi-analytical approach and a typical numerical simulation are elucidated in Table 3.3. Care must be taken so that numerical stability is maintained, where the condition for one-dimensional stability is given in (3.18). For the semi-analytical solution given in the table, D could be as large as 0.011 m s^{-1} .

$$\frac{D\Delta t}{\Delta y^2} \leq 0.5 \tag{3.18}$$

A representative solution is shown in Fig. 3-10(right). Unlike the analytical solution, this approach does account for changes due to diffusivity.

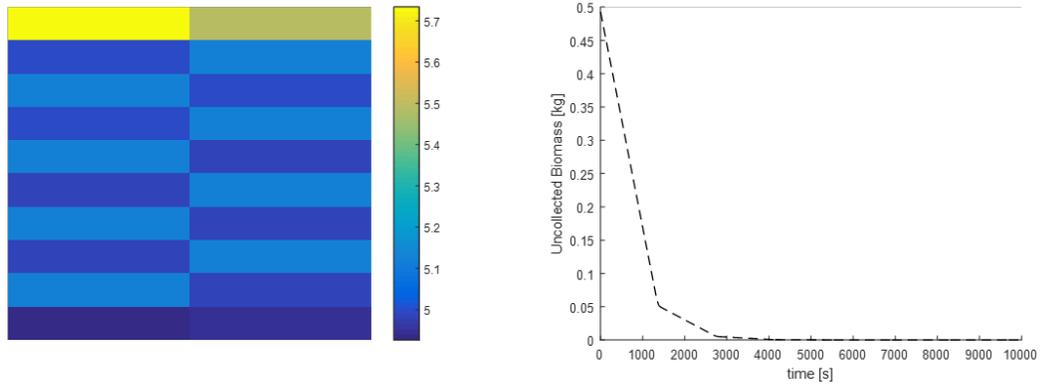


Figure 3-10: (left) Semi-analytical algae field after the initial pass using very coarse forward-time center-space (FTCS) discretization of the partitioned case; (right) Representative solution using the semi-analytical approach where the delineation between passes is clearly evident

3.8.6 Study Limitations

Collision avoidance was not implemented in these studies and will certainly have an impact on the results, so it is difficult to compare the scaling of these swarms. It has been helpful to see some of the scaling effects present, e.g. scaling of partition perimeter to area ratio, which may be masked if collision avoidance is enabled. Collision avoidance could be implemented, but will decrease the computational efficiency of these simulations.

This simulations assume a hydrodynamically complex domain can be modeled with uniform eddy diffusion. A real world domain will be quite different from this, and furthermore, these simulations assumed that the agents themselves do not cause any hydrodynamic disturbance, which is certainly not realistic. In defense of the work done, a simple domain is needed to start gaining some intuition about diffusive collection before introducing this additional complexity. The hydrodynamic disturbance caused by an agent itself will be highly dependent on unknown details of the agent's geometry, for which a dedicated study is needed.

Table 3.3: Comparison of Semi-Analytical and Full Simulation approaches for the partitioned case

Methodology	Spatial Step	Time Step	Diffusion Discretization	Total Nodes	Total Time Steps
Semi-Analytical	1m	43.8s	FTCS (1D)	64	228
Full Simulation	0.1m	0.1s	FTCS or ADI (2D)	1E6	1E5

Note: Assume both methodologies use 10 agents operating in a 100m x 100m domain, with a velocity of 0.7222 m s^{-1} , for a simulated 10,000s. For the semi-analytical case, this would result in a 31.6m x 31.6m domain and each row would require $31.6\text{m} / 0.7222 \text{ m s}^{-1} = 43.8\text{s}$ to traverse.

These studies all started with an initially-homogeneous algae distribution. Algae is sometimes homogeneously distributed, but can take many other forms, the diversity of which is illustrated in Figures 2-5 and 2-12. Acknowledging that there is no ‘typical’ distribution, at a minimum, simulations can be run for a single, non-heterogeneous distribution generated from real-world data.

Thus far, the simulations have only been performed a single time for each set of parameters. This restriction has been mostly due to the amount of time required to run these simulations. If the goal of these simulations were to develop statistically significant estimates of mean performance, than adding additional runs would absolutely be necessary. However, it is fair to draw conclusions with a modest level of certainty from the existing work because (i) the input parameters in some cases are being changed by orders of magnitude, which naturally have a large impact on swarm behavior and consequent performance, (ii) qualitatively, observations of the algae distributions show the impacts of changing parameters, and (iii) the analytical and semi-analytical approximations validate the numerical data, where the root cause of any discrepancies can be elucidated.

3.8.7 Computational Cost

Some steady progress has been made on optimizing the MATLAB code, but even in the optimized form, running the existing simulations as a set requires about 60 hours. There are some options available for achieving more runs, as they are needed. Porting the MATLAB code to a faster language, e.g. Fortran, should give a good improvement in runtime. To this point, the ease-of-implementation of a very high level language like MATLAB has been prioritized. A single GPU is already being well-utilized, but there are both CPU and GPU supercomputing resources available that could enable more runs to be completed.

3.8.8 Numerical Techniques

The numerical techniques used for diffusion (FTCS (explicit) and Crank-Nicholson ADI (implicit)) have limited accuracy in time and space ($O(\Delta t, \Delta x^2)$, $O(\Delta t^2, \Delta x^2)$, respectively), and more accurate techniques are available. Implicit schemes typically have a higher computational cost, but are chosen for their better stability and option of using a larger time step, which can mitigate the higher computational cost. In this case, a larger time step is not an option because agents are interacting with and making decisions based on the algae distribution. To this point, Crank-Nicholson ADI was utilized for the ‘Limits of Diffusive Collection’ study for its higher accuracy, despite its higher computational cost. A higher accuracy explicit scheme would probably be most suitable.

3.8.9 Diffusion Coefficients

Note that due to hydrodynamic heterogeneity, it is very difficult to define a true diffusion coefficient. For practical problems, it is easier to estimate a diffusion coefficient proportional to patch size [82], which can be estimated using Richardson’s

Table 3.4: Diffusion coefficient calculations based on alpha parameter and patch size

Alpha [cm ^{2/3}]	Patch Size [meter]	ϵ [cm ² s ⁻¹]	D [m ² s ⁻¹]
0.01	1	4.64	4.64E-4
0.002	1	0.928	9.28E-5
0.002154	1	1.00	1.00E-4
0.002154	10	10.0	1.00E-3
0.002154	100	100	1.00E-2

Note: highest alpha estimate is 0.01, lowest alpha estimate is 0.002, and a range of patch sizes are considered (1, 10, and 100m). Unless noted, simulations used D=1.00E-4.

well-established ‘4/3 rule’, which states that $D = \alpha L^{4/3}$, where L is a characteristic patch length and α is a constant. Physical experiments have shown that an alpha value of 0.01 to 0.002 cm^{2/3}s⁻¹ is reasonable [39], which is illustrated in Fig. 3-11. Thus, for an alpha of 0.002 (the lower bound), and a patch size of 1m, a diffusion coefficient of 0.0001 m² s⁻¹ can be estimated, with details shown in Table 3.4.

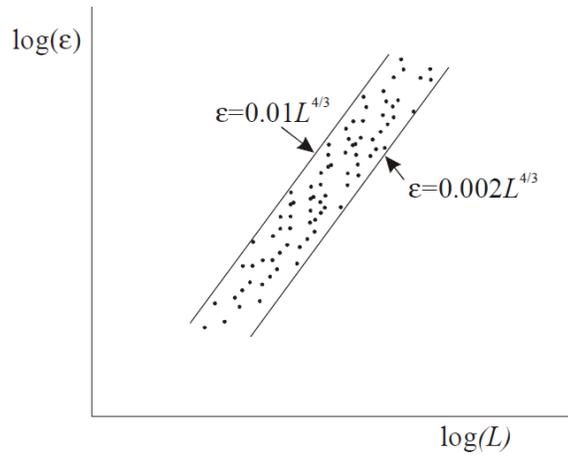


Figure 3-11: Eddy diffusion coefficient (ϵ) with respect to patch size(L) [82]

Chapter 4

Balancing Robot Swarm Cost and Interference Effects by Varying Robot Quantity and Size

This section is a reproduction of a manuscript that has been accepted for publication in the Journal of Swarm Intelligence, authored by Adam Schroeder, Brian Trease, and Alessandro Arsie [91]. It has been modified to conform to the dissertation format.

4.1 Abstract

Designing a robot swarm requires a swarm designer to understand the trade-offs unique to a swarm. The most basic design decisions are how many robots there should be in the swarm and the individual robot size. These choices in turn impact swarm cost and robot interference, and therefore swarm performance. The underlying physical reasons for why the number of robots and the individual robot size affect interference is explained in this work. A swarm interference function was developed and used to build an analytical basis for swarm performance. A swarm cost model was

also developed and used with the analytical basis for swarm performance to generate performance-cost curves for swarms with different numbers of robots and different robot sizes. The swarm designer can use this analytical basis, cost model, and these curves to weigh how the number of robots in the swarm and the individual robot size can be selected to minimize swarm cost and maximize swarm performance. This work is motivated by the desire to engineer a swarm to collect harmful algae from water. In this foraging application, the robots are not required to deposit algae in a central location. Stepping through the design process for this application has exposed several of the knowledge gaps addressed herein.

4.2 Introduction

Interference in multi-robot systems has been studied for some time [67], especially how increasing the number of robots in the swarm affects both the individual efficiency and the swarm performance [59]. For applications with spatial restrictions, e.g. robots returning objects to a central location after they have been found, swarm performance initially improves as robots are added to the swarm. As still more robots are added, a peak swarm performance is achieved and the interference caused by any additional robots now causes performance to decrease. If these spatial restrictions do not exist or are removed, performance continues to improve as additional robots are added, but the gain from each additional robot declines [83]. Swarm performance trends for applications with spatial restrictions have been well-fitted by several models [44], albeit without a comprehensive explanation of the underlying physical phenomena that drive these trends. A detailed examination of this background literature is given in Section 4.3.

This work focuses on applications without these spatial restrictions, and develops an interference function derived from the underlying physical phenomena. This in-

terference function is used to form an analytical basis for swarm performance, which could be utilized by a robot swarm designer in the design of a generic swarm.

Unlike interference, swarm cost, to the authors' knowledge, has not been studied, which is indicative of the current state of the broader field of swarm design. Within the field of swarm robotics, there is a dearth of information on the process of designing a swarm for any specific application. This sentiment is expressed by Brambilla et al [12] in their comprehensive review of the state of swarm robotics. They use the term 'swarm engineering' to describe the emerging field needed to tackle real-world applications and they summarize the state of swarm robotics with "Despite their potential to be robust, scalable and flexible, up to now, swarm robotics systems have never been used to tackle a real-world application and are still confined to the world of academic research".

The swarm designer must have a methodology to follow to make very practical design decisions. One of the primary decisions that must be made is the individual robot size and number of robots in the swarm. In deciding this, two major considerations for any application are the cost of an individual robot (a function of robot size and complexity), and any cooperation or interference effects (a function of robot size and density). Robot size will influence the cost of a robot for any application, but will not equally influence cooperation or interference effects for all applications. Collective transport is an example of an application where robot size will have a strong influence on cooperation. The designer must strike a balance between minimizing swarm cost while also attempting to minimize interference.

The higher level goal of this work is informing the swarm designer in the design of a swarm for a real-world application and the major contributions of this work are aligned with that goal. The first contribution is a generic, analytical basis for swarm performance, with an emphasis on robot size, robot density, and delays incurred due to collisions between robots. This analytical basis is for applications without spatial

restrictions, e.g. foraging when objects do not need to be returned to a central location. This analytical basis for swarm performance is validated numerically. The second contribution is a generic formulation for the cost of a swarm. The third contribution is the application of the swarm performance analytical basis and swarm cost formulation to a typical swarm design process. Two cases are considered where either the robot size is pre-defined and the number of robots varies, or where both the number of robots in the swarm *and* their size are design parameters that can be selected to minimize cost and maximize swarm performance.

This work is best understood by introducing the motivating real-world application. A swarm of aquatic robots is being designed to remove harmful algae from water. These robots may use several different mechanisms for collecting algae, e.g. conveyors for collecting surface scum or plankton nets for collecting sub-surface algae, with a prototype 1st-generation robot shown in Fig. 4-1. For any of these collection mechanisms, the robot width determines both how quickly the algae can be collected and the overall platform cost.

This application can be viewed as a foraging application where the collected algae could be processed aboard each individual robot and would therefore not require deposition in a central location. The algae to be collected are quantified as a concentration, with the added complication that the algae distribution changes due to turbulent diffusion and advection. Furthermore, initial field tests have shown that the robot's collection mechanism, hull, and propulsion system all interact with the algae-laden water, having an overall mixing effect. In other words, despite constantly collecting algae, the robot does not have a well-defined algae-free area in its wake. Due to this mixing effect, it is assumed that robots are collecting the average concentration of algae. Making this assumption allows swarm performance to be estimated for different numbers of robots and different robot sizes.

This paper is not intended to give a complete description of the algae-collecting

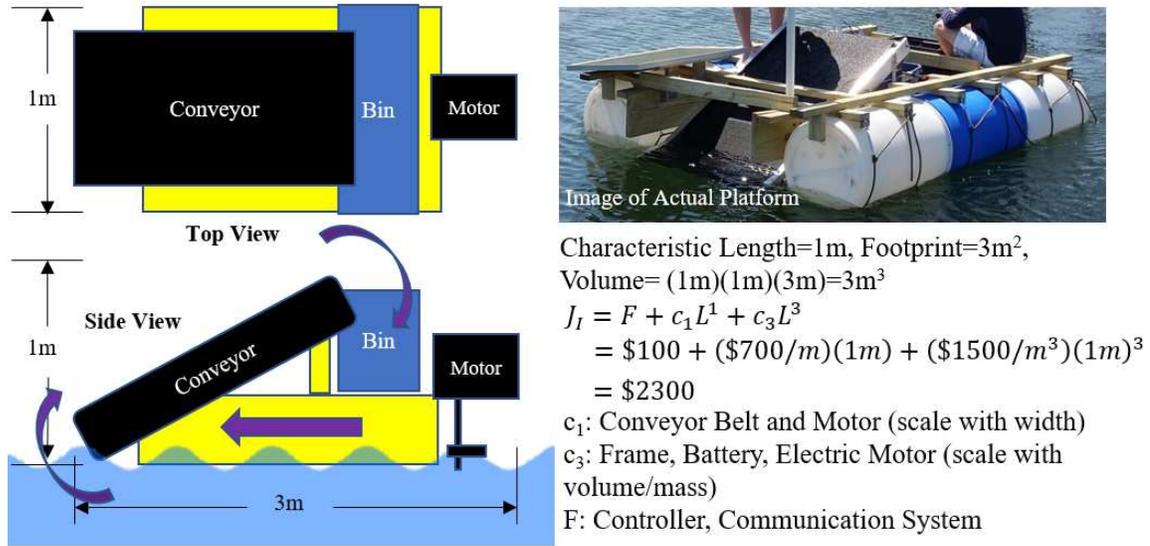


Figure 4-1: Top view and side view of the 1st-generation aquatic platform for collecting algae surface scum with the major components labeled and the basic dimensions given. The fixed and variable costs were estimated based on the build of a single prototype platform, and the components and systems contributing to each type of cost are listed.

robots themselves, which will be described in separate work. Rather, the aim is to provide enough information to understand the application and aspects of the application that influence swarm performance and cost.

The remaining of the paper is organized as follows. Section 2 introduces the relevant literature. Section 3 explores swarm cost, interference, individual efficiency, and swarm performance, both when robot size is held constant and when it is a design choice. Section 4 applies this knowledge to generate several sets of performance-cost curves and discusses the results. Finally, Section 5 concludes and offers thoughts on future work.

4.3 Relevant Literature

The only mention of cost in literature with respect to swarms is the cost of individual swarm platforms as they are developed. For example, the Kilobot requires \$14 in parts [85], the Jasmine robot requires \$130 in parts [68], and the aquatic platform developed by Duarte et al [32] requires \$370 in parts. The effort to develop these demonstration platforms is certainly valuable and necessary, but the total swarm cost is not a focus of the work. For some of these examples, the total cost of the swarm can be calculated because the number of robots is also known. The Kilobot swarm (1000 robots) costs \$14000 (parts only), and the aquatic swarm (10 robots) costs \$3700 (parts only).

This lack of attention to the total swarm cost may be because swarm robotics is a relatively young field of study, where making any platform ‘inexpensive enough’ to construct a demonstration swarm is a significant contribution. However, total swarm cost is a universal swarm design constraint, a better understanding of which is requisite for advancing the field of swarm robotics.

In contrast to total swarm cost, there is a strong basis in literature focused on robot interference going back several decades, e.g. Mataric [67]. One of the works that is often cited is from Lerman and Galstyan [59], who developed a mathematical model for swarm performance that includes interference effects. They considered a foraging task where robots either do or do not need to deposit objects at a central location. In their model, any collision or the need for collision avoidance incurs a delay time, τ . Validated by simulations, they showed that when objects are deposited at a central location, swarm performance initially increased as more robots were added to a swarm. As more robots were added, swarm performance decreased due to interference. When removing the constraint to deposit objects, performance monotonically increased as more robots were added. In contrast to swarm performance, they observed that the

efficiency of an individual robot always decreased as the number of robots increased. This was true both when robots did or did not deposit objects at a central location.

Rosenfeld et al [83] added to the work done by Lerman and Galstyan. They studied swarm performance for different numbers of robots in a simulated foraging task for several different collision avoidance behaviors. When objects were deposited at a central location, they also showed swarm performance peaking, which they called point CP1. After the peak, swarm performance eventually plateaued, which they called CP2. The maximum swarm performance, and the number of robots present when that maximum was achieved, varied for the different collision avoidance behaviors. Consistent with [59], removing the constraint to deposit objects transformed the performance curve to a monotonically-increasing shape. For this less-restrictive case, there was little difference between the different collision avoidance strategies, including one case that allowed robots to simply pass through one another.

Following on this, Hamann [44] formulated a universal swarm performance function, $P(N)$, as a combination of cooperation and interference functions, $C(N)$ and $I(N)$, given in Eqn. (4.1). Here N is the number of robots, and fitting-parameters $c < 0$, a_1 , a_2 , and $b > 0$, and $d \geq 0$.

$$P(N) = C(N) (I(N) - d) = a_1 N^b a_2 \exp(cN) \quad (4.1)$$

This performance function encapsulates positive cooperation effects, N^b , and negative interference effects, $\exp(cN)$. Hamann then fit this general function, using the fitting parameters, to published examples of foraging, collective decision making, aggregation, and swarm motion toward a beacon. Portions of this swarm performance formulation will be referenced throughout this article, and some of the nomenclature retained.

As an alternative to a formulation where cooperation and interference effects are

multiplied, Guerrero et al [41] formulated a swarm’s capacity to perform a task as an ideal capacity minus interference effects.

Performance which decreases as the swarm size increases cannot be solely attributed to interference effects, at least those caused by physically occupying a space. This effect is often referred to as ‘diminishing returns’. Even without including collisions, Hecker and Moses [47] observed sub-linear performance increases as more robots were added in a foraging task simulation. Rosenfeld et al [83] also observed sub-linear performance increases for one of their simulation sets that simply allowed robots to pass through one another without colliding.

Thus, the phenomena of interference, is more complex than just competing for physical space. In a foraging task, any finite, discrete resource collected by a robot is removed from the shared pool of available resources. For every other robot in the swarm, the uncollected resources become sparser and the foraging task becomes more challenging. Hamann’s model can account for sub-linear performance, for $b < 1$, as can Guerrero’s model.

Work has been done to help discriminate between different types of interference. Scharf et al [88] elucidated the difference between exploitive and interference type competition in biological examples of foraging. Interference competition is competing for physical space, and exploitive competition is competing for shared resources. Thus, even when collisions are ignored in simulation, there is still indirect ‘exploitive’ interference as the robots compete for shared resources. Dawson et al [27] concentrated on what they call ‘competing’ and ‘passing’ interference. Competing interference means that robots have proximal goals and are trying to physically occupy the area around their goals. Passing interference means that robots are only interacting in passing and their goals are not close to one another. The following work concentrates on developing a model for the direct interference caused by robots trying to physically occupy the same space. Any indirect interference caused by competing

for shared resources can be accounted for in the cooperation portion of the swarm performance formulation.

It is noted that a robot's physical size has not been considered in this body of literature. The same explanation for why total swarm cost has not been studied could be given here. Robot size has likely not been thought of as a design choice because the swarm designer typically starts at a point when the physical platform has already been selected. As swarm engineering matures as a discipline, it is reasonable that the swarm's physical platform and control strategy will be developed simultaneously.

Another reason why robot size may not have been considered is that its importance is application-dependent. For example, robot size is highly-relevant for foraging for non-discrete resources, collective transportation, and forming chains of robots. Robot size is less relevant for foraging for discrete resources, aggregation, or coordinated motion. Stepping through the swarm design process for the algae collection platform identified physical size as one of the first important design decisions that must be made.

4.4 Model Development

4.4.1 Interference and Swarm Performance

To develop an interference function and analytical basis for swarm performance, two claims are made and developed. These claims are applicable when objects do not need to be returned to a central location.

Claim 1 - The swarm collision rate is proportional to the number of potential collisions, i.e. a mathematical combination, $\binom{N}{2}$, that enumerates the number of potential collision pairs.

Claim 2 - The swarm collision rate is also proportional to the robot's characteristic length.

Together, these claims are expressed in Eqn. (4.2).

$$X_{\Sigma pS} = \lambda \binom{N}{2} L = \lambda \frac{N!}{(N-2)!2!} L = \lambda (N^2 - N) L \quad (4.2)$$

Here, $X_{\Sigma pS}$ can be read as the number of swarm collisions, X_{Σ} , per second, i.e. the swarm collision rate. Throughout this work, a Σ or I subscript is used to refer to the swarm or the individual, respectively. The collision proportionality coefficient, λ , is introduced as a scaling parameter, N is the number of robots in a fixed domain, and L is a robot's characteristic length. For the motivating application of algae collection, the characteristic length is defined as the width of the robot's conveyor.

Note that this relationship will certainly not hold for every possible robot behavior, but will provide a reference point from which to evaluate collision frequency for any behavior.

These two claims are validated numerically by performing two simulation sets¹. Robots performed a Lévy flight-type random walk inside a $100 \times 100 \text{ m}^2$ domain, and collisions were tracked while the number of robots ranged from 10 to 2500. Lévy flight is a type of random walk where robots take frequent short steps and occasional longer steps, and it was implemented similar to previous work by some of the authors [90]. Unlike the previous work, here the Lévy flight is unbiased. This means that robots are moving completely randomly. In general, Lévy flight is an active area of research within swarm robotics for robot search tasks [28, 98, 72] and robot area coverage and interception tasks [89]. Robots moved at a rate of one unit per second for 1000 seconds. For the first simulation set, robots were assigned a physical size of 1 m, and

¹MATLAB script available at <https://www.mathworks.com/matlabcentral/fileexchange/65598-collision-detection> or under Zenodo DOI:10.5281/zenodo.1323875

for the second simulation set, robot sizes of 0.5, 1, and 2 m were tested. A collision was recorded if the centers of any two robots came within the characteristic length of one another. During the simulation, registering a collision did not change a robot's speed or heading in any way. To minimize boundary condition effects, boundaries were selected to be periodic. This means that robots can move freely across boundaries, e.g. passing through the top boundary and re-entering at the bottom boundary. In Fig. 4-2, the swarm collision rate for these simulations is plotted with Eqn. 4.2. The proportionality between the analytical basis and numerical results is evident, and the numerical results were used to determine an appropriate value for the proportionality coefficient, λ .

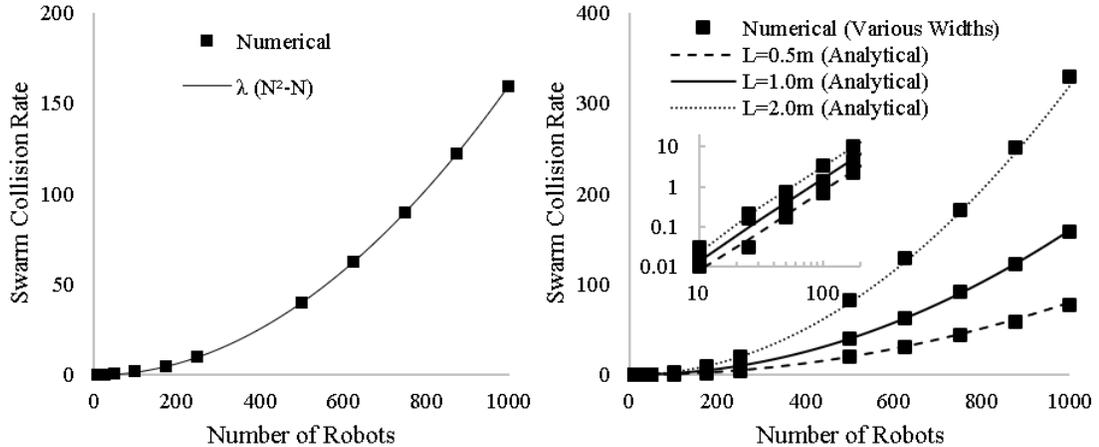


Figure 4-2: Comparison showing the proportionality between an analytical formulation and numerical simulation for swarm collision rate as (Left) the number of robots are varied, and (Right) the number of robots and individual robot size are varied. The inset figure with logarithmic scales shows a closer view of the lower end of the robot range. For this system, a value for the proportionality coefficient of $\lambda = 0.0001566$ fits the analytical formulation to the numerical results. Error bars are not included for clarity, but for reference, the largest 95% confidence interval for these data points is 1.35 swarm collisions per second.

Relations for swarm performance and individual efficiency, P_{Σ} and η_I , respectively,

are developed below. Individual efficiency is defined in Eqn. (4.3), where X_{TpIpS} can be read as the time spent in collision per individual per second. If $X_{TpIpS} = 1$, the individual robot would be spending the entire amount of their time in collision, i.e. 0% efficient.

$$\eta_I = (1 - X_{TpIpS}) \quad (4.3)$$

Swarm performance is defined in Eqn. (4.4) as having a cooperation portion, N^b , and an interference portion, represented by the individual efficiency. Similar to Hamann [44], a_1 is a constant, and $0 < b < 1$ for sub-linear cooperation, $b = 1$ for linear cooperation, or $b > 1$ for super-linear cooperation. The cooperation portion can be understood as the swarm performance that could be achieved if there were no interference. Stating that there is no interference is equivalent to saying that individual efficiency is 100%, i.e. $X_{TpIpS} = 0$ and $\eta_I = 1$. If there are collisions and individual efficiency is less than 100%, swarm performance will be less than this idealized case.

$$P_\Sigma = a_1 N^b \eta_I \quad (4.4)$$

The cooperation portion is equivalent to Hamann's formulation, Eqn. (4.1), wherein he states that mostly values for $b > 1$ are of interest. In very recent work, Hamann [45] explores applications where superlinear scalability, i.e. $b > 1$, can be realized. Examples of these applications include when multiple robots are needed to manipulate an object, or complete a robot chain, or bridge a gap within the swarm. These situations are not expected to apply to the motivating example of algae collection, but this does not exclude the more general application of this analytical basis to superlinear scenarios. The interference portion will not be equivalent to Hamann's formulation, which is exponential.

Calculating X_{TpIpS} is performed by multiplying τ , the average delay time incurred per collision, and X_{pIpS} , the number of collisions per individual per second, as given in Eqn. (4.5):

$$X_{TpIpS} = \tau X_{pIpS} \quad (4.5)$$

The average delay time, τ , accounts for time spent in collision avoidance behavior instead of performing the desired task. The number of collisions per individual per second, X_{pIpS} , is the swarm collision rate, $X_{\Sigma pS}(\tau)$, divided by the number of robots, as given in Eqn. (4.6). The swarm collision rate, $X_{\Sigma pS}(\tau)$, varies with τ and is equal to the product of $X_{\Sigma pS}(0)$, and the individual efficiency, as given in Eqn. (4.7). Note that the swarm collision rate for a delay time of zero, $X_{\Sigma pS}(0)$, has already been formulated in Eqn. (4.2) as $\lambda(N^2 - N)L$.

$$X_{pIpS} = \frac{X_{\Sigma pS}(\tau)}{N} \quad (4.6)$$

$$X_{\Sigma pS}(\tau) = X_{\Sigma pS}(0) \eta_I = \lambda(N^2 - N)L\eta_I \quad (4.7)$$

Explained differently, the swarm collision rate (any delay time) is the maximum swarm collision rate (when no delay is considered), multiplied by the portion of time that robots are not in collision (same as individual efficiency). This approach also implies that no useful work is performed during time spent ‘in collision’. In certain applications, even while engaged in collision avoidance behavior, a robot may still be performing useful work, e.g. providing area coverage, exploring new areas of the domain, or communicating with its partner in collision.

At this point, the impact of τ is validated as before by running numerical simulations, but now forcing robots to pause after a collision. The resulting swarm collision rate is shown in Fig. 4-3 for either no delay or for delay times of 2, 5, and 10s. The analytical model was used with the previously-found value of λ when no delay was

incurred.

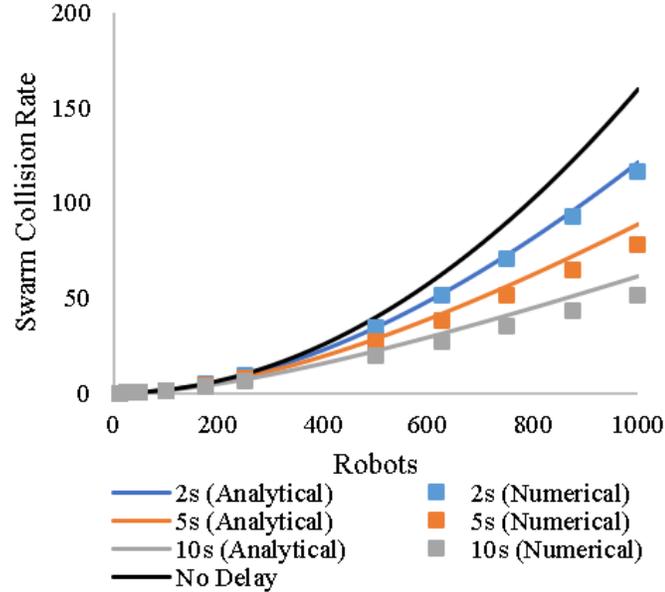


Figure 4-3: Comparison of analytical and numerical swarm collision rates for several different delay times. As the number of robots increases, the analytical model begins to predict higher collision rates than observed in the numerical simulations. This is attributed to the nature of the simulations, which forces robots in collision to pause. In this simplified implementation, robots involved in separate collisions have no chance of colliding with one another.

The analytical formulation does a fair job of representing the measured numerical swarm collision rate, especially for smaller numbers of robots. For higher numbers of robots, the analytical formulation predicts more collisions than actually occurred in the numerical simulations. This result is attributed to how the collision delay was enforced in the simulations, where robots paused for the duration of the delay. Collisions were implemented in this general way to avoid peculiarities stemming from any specific collision avoidance strategy. However, this means that any robots involved in separate collisions have no chance of colliding while paused, which reduces the number of possible collision pairs. These simulations also did not differentiate a robot's physical size from its sensing footprint, if equipped with proximity sensors for

collision avoidance. Taking detection radius into account means that robots have a projected or virtual size even larger than their physical size. This concept would add another level of complexity to the interference function.

Note that a swarm with a shorter delay time actually incurs more collisions than a swarm with a longer delay time. However, when delay time per collision is considered, swarms with a shorter delay time spend less time overall in collision, and are thus more efficient.

In the numerical study, there is some freedom to define how collision delay time is applied to robots. For the results given, if a robot already ‘in collision’ was struck by another robot, its delay time counter was reset to the assigned delay time. Another reasonable implementation would be to simply add additional time to a robot’s remaining delay time from the first collision. Collisions involving more than two robots may need to be considered differently in the analytical model. An enhanced model could penalize these types of collisions with longer delay times.

With the impact of τ now validated, Eqns. (4.2-4.7) can be used to solve for the individual efficiency and swarm performance, given in Eqns. (4.8) and (4.9), respectively.

$$\eta_I = \frac{1}{1 + \lambda\tau(N - 1)L} \quad (4.8)$$

$$P_\Sigma = \frac{a_1 N^b}{1 + \lambda\tau(N - 1)L} \quad (4.9)$$

Using Hamann’s terminology [44], the individual efficiency, Eqn. (4.8), is also an interference function. It is combined in Eqn. (4.9) with Hamann’s cooperation function to generate an analytical basis for swarm performance.

As noted, the analytical basis for swarm performance will not represent every possible control scheme or swarm behavior. It is expected that some control schemes accommodate additional robots or larger robots more gracefully than others. For

example, a scheme which totally partitions the swarm’s domain and assigns robots to work in the individual partitions [75] would likely have fewer collisions than a non-partitioning approach. Another example of a partitioning scheme creates distinct partitions for robots moving toward a common target and for robots moving away from a common target [95]. This analytical basis does serve as a reference point for evaluating observed interference for any control scheme.

4.4.1.1 Application-Specific Model

For the algae collection application, the cooperation portion of the general form of the swarm performance function, Eqn. (4.9) can be modified. This modification is based on the approximation that robots are always collecting the average concentration of algae. Mathematically, the cooperation function, $C(N, L)$, can be expressed as the algae collection rate [kg s^{-1}] divided by the total algae [kg]. This normalized collection rate [s^{-1}] expresses the fraction of total algae present being collected at any point of time. It is derived in Eqn. (4.10), where s is the robot speed, $\eta_{process}$ is the process efficiency, i.e. the capture efficiency of each robot’s algae collection hardware, ρ is the average algae concentration, and A is the area of the domain:

$$C(N, L) = \frac{NLs\eta_{process}\rho}{\rho A} = \frac{[\text{m}] \left[\frac{\text{m}}{\text{s}} \right] \left[\frac{\text{kg}}{\text{m}^2} \right]}{\left[\frac{\text{kg}}{\text{m}^2} \right] [\text{m}^2]} = \frac{\left[\frac{\text{kg}}{\text{s}} \right]}{[\text{kg}]} = \left[\frac{1}{\text{s}} \right] \quad (4.10)$$

A dimensional analysis of this relation is also provided for clarity. This relation for normalized collection rate is further simplified in Eqn. (4.11), which shows that it is constant regardless of the algae concentration, ρ . The application-specific parameter values for s , $\eta_{process}$, and A are also inserted.

$$C(N, L) = \frac{NLs\eta_{process}}{A} = \frac{NL(1\text{ms}^{-1})(0.9)}{(100\text{m})^2} = 0.00009LN \quad (4.11)$$

In terms of the parameters of Hamann’s cooperation function, $a_1 = 0.00009L$ and

$b = 1$, giving linear cooperation. These values are substituted into Eqn. (4.9) to form Eqn. (4.12):

$$P_{\Sigma} = \frac{0.00009LN}{1 + \lambda\tau(N - 1)L} \quad (4.12)$$

For this application, swarm performance, P_{Σ} , is a normalized collection rate. For reference, previous literature has measured swarm performance for a foraging task as a collection rate without normalization. One study used resources collected per hour [47], and a second study used resources collected per nine-minute trial [83].

4.4.2 Constant Size and Constant Capacity

To illustrate the implications of the analytical basis for swarm performance, two specific approaches are introduced which are of particular interest to the swarm designer. The first approach, referred to as ‘constant size’, looks at the performance trend as more robots are added to the swarm. This approach is comparable to how results have been presented in several existing studies. The second approach, referred to as ‘constant capacity’, sets the swarm capacity as the sum of all of the individual robot capacities. In the case of algae collection, this would be the sum of each robot’s conveyor width. This constant capacity, C_s , could then be allocated to either few large robots, or many small robots. Examples of representative performance curves for these two approaches are given in Fig. 4-4, using a collision delay time of 10 s.

The constant size and constant capacity approach are only a subsample of different approaches that may be of interest to the swarm designer. The constant size approach, for example, is suitable for when the physical robot platform has been pre-determined or is not a design choice. This constant size approach remains useful even after an initial swarm has been constructed because additional robots could be added to the existing swarm. The constant capacity approach takes a more system-level view of the

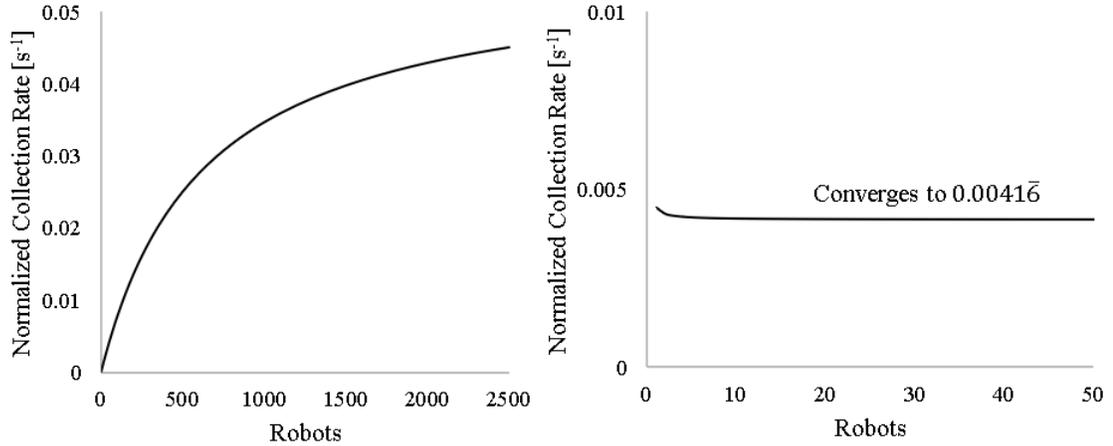


Figure 4-4: Performance curve for (Left) constant robot size of 1 m as more robots are added, or (Right) constant swarm capacity of 50 m, where the capacity can be allocated to either few large robots, or many small robots. The constant size case shows diminishing returns when adding additional robots. The constant capacity case shows a small initial drop in performance as the number of robots increases (and size decreases), but the performance is steady as the number of robots continues to increase.

swarm, where the swarm designer may already know what swarm capacity is needed for a particular application. Additional approaches could also be valuable, e.g. a ‘constant quantity’ approach, where robot quantity remains constant and platform size varies.

For the algae collection application, the mathematical limit of the performance function can be calculated as the number of robots approaches infinity. For the constant size case, the limit of the performance function, Eqn. (4.12), is $\frac{0.00009}{\lambda\tau}$. Filling in values for the λ and τ , a performance limit of 0.05625 is found, which matches Fig. 4-4 (left). The limit of the same performance function for the constant capacity case is $\frac{0.00009C_s}{1+\lambda\tau C_s}$. A swarm capacity of 50 m approaches a performance value of 0.00416 which matches Fig. 4-4 (right).

4.4.3 Swarm Cost

A formulation for individual robot cost, J_I , is developed below, where total swarm cost, J_Σ , is simply obtained by multiplying the individual robot cost by the number of robots. An individual robot's cost comprises fixed costs, F , and several costs that vary with the robot's characteristic length, L . These variable costs can be generally formulated as a scaling coefficient, c_1 , c_2 , or c_3 , multiplied by the characteristic length raised to a power, as given in Eqn. (4.13):

$$J_\Sigma = NJ_I = N(F + c_1L^1 + c_2L^2 + c_3L^3) \quad (4.13)$$

Put plainly, platform costs are assumed to scale with a characteristic length, area (e.g. robot footprint), or volume. This cost model is a simple way to represent how the cost of a robot's systems and components scale, and could be refined as needed. For example, at very small scales, costs would be expected to increase as components need to be further miniaturized. This cost calculation is applied to the constant size and constant capacity scenarios. Corresponding performance-cost curves for these two scenarios are generated and shown in Fig. 4-5.

The swarm capacity, C_s , obtained by multiplying the number of robots by their characteristic length, can be substituted into Eqn. (4.13) to form Eqn. (4.14):

$$\begin{aligned} J_\Sigma &= NJ_I = \frac{C_s}{L} (F + c_1L^1 + c_2L^2 + c_3L^3) \\ &= C_s \left(\frac{F}{L} + c_1 + c_2L + c_3L^2 \right) \end{aligned} \quad (4.14)$$

Differentiating Eqn. (4.14) with respect to the characteristic length and setting the result equal to zero enables an optimal characteristic length to be determined.

$$\frac{dJ_\Sigma}{dL} = C_s \left(-\frac{F}{L^2} + c_2 + 2c_3L \right) = 0 \quad (4.15)$$

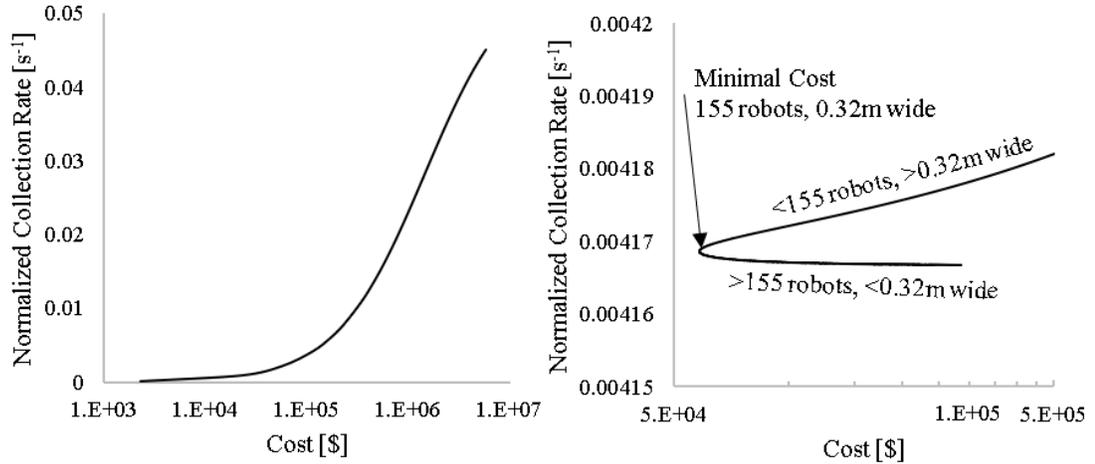


Figure 4-5: Performance-cost curve for (Left) constant robot size of 1 m, and (Right) constant swarm capacity of 50 m. Cost is shown on a log scale in dollars and values for c_1 , c_3 , and F from Fig. 4-1 were used to find cost. Note that order-of-magnitude increases in cost do not translate to order-of-magnitude increases in performance for either scenario. Also note that performance variation for the constant capacity case is much smaller.

Multiplying both sides by L^2 , and dividing by C_s and $2c_3$, puts Eqn. (4.15) in the form of a cubic equation, given as Eqn. (4.16):

$$L^3 + \frac{c_2}{2c_3}L^2 - \frac{F}{2c_3} = 0 \quad (4.16)$$

The roots of this cubic equation can be found using a general method, e.g. Cardano's formula. In general, Eqn. (4.16) could have either one real root and two complex roots, or three real roots. In either case, applying Descartes' rule of signs indicates that there will be only one positive real root. This root is equivalent to the optimal characteristic length. The second derivative of Eqn. (4.14) will always be positive, which indicates that L is a minimum, i.e. is the optimal characteristic length which will minimize J_Σ .

The cost model for the algae collection platform has non-zero coefficients for the

characteristic length (conveyor width, $c_1 = 700$) and volume ($c_3 = 1500$), as detailed in Fig. 4-1, but not the area ($c_2 = 0$). For these coefficient values, Eqn. (4.14), and therefore Eqn. (4.16), can be simplified as given in Eqns. (4.17) and (4.18), respectively:

$$J_\Sigma = NJ_I = N(F + c_1L^1 + c_3L^3) \quad (4.17)$$

$$L^3 - \frac{F}{2c_3} = 0 \quad (4.18)$$

Equation (4.18) may be easily solved without employing a more general methodology, yielding an optimal characteristic length, $L = \left(\frac{F}{2c_3}\right)^{1/3} \approx 0.32\text{m}$. This length matches the minimal cost point on the performance-cost curve in Fig. 4-5 (right). It is also worth noting that the robot size that minimizes cost is independent of the swarm capacity, C_s .

4.5 Model Application and Discussion

4.5.1 Swarm Designer Perspective

From the perspective of a swarm designer, it is useful to assess how the analytical basis for swarm performance could be used, what demands it places on the designer, and what information is still lacking. The constant size and constant capacity scenarios were introduced to illustrate the cost and performance implications of varying the number of robots and the individual robot size. An entire set of performance-cost curves can be generated for a range of sizes and capacities. In Fig. 4-6, a set of curves is shown for the algae collection example. Constant size curves were plotted for 0.05, 0.1, 0.25, 0.5, and 1 m robot widths, and constant capacity curves were plotted for 10, 20, 30, 40, and 50 m capacities.

This set of performance-cost curves allows a swarm designer to weigh the design

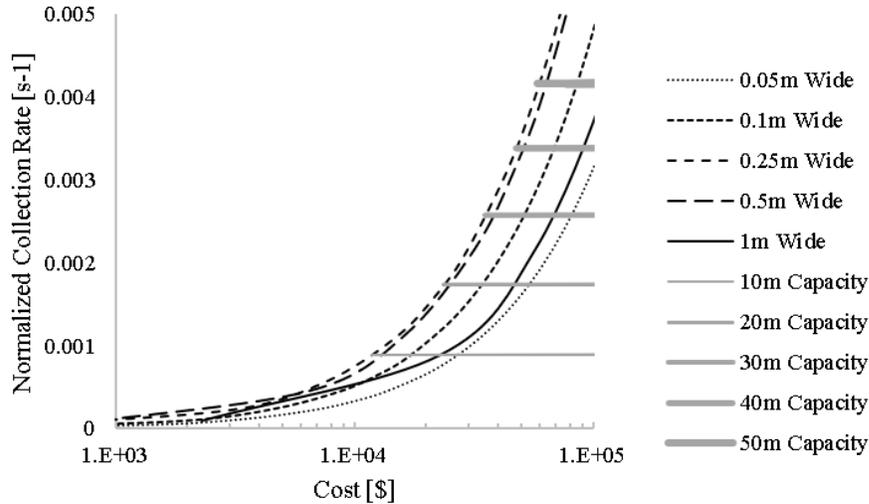


Figure 4-6: Several performance-cost curves where black dashed lines are constant size curves (longer dashes indicate larger robots), and gray solid lines are constant-capacity curves (thicker lines indicate higher capacities).

trade-offs of varying the number of robots, varying swarm capacity, or setting different target swarm costs. For example, for these particular curves, a swarm designer will get the most cost-effective performance with a robot size of approximately 0.25-0.5 m. Deviating from this size while maintaining the same swarm capacity, perhaps to accommodate other design constraints, will have a large cost impact but almost no impact on performance. Note that the constant capacity curves appear to be simply horizontal at this scale, but their full shape becomes apparent at a magnified scale (see Fig. 4-4).

Generating these performance-cost curves makes some demands of the designer, although not dissimilar from any other design process. The designer must settle on an appropriate cooperation function, here accomplished by observing how real, prototype robots performed. Without prototype robots, a designer must turn to published examples of similar applications and extract parameter value estimates. For a demonstration of the latter approach, swarm performance results were analyzed for

a foraging application from Hecker and Moses [47]. Values for a_1 and b were manually tuned until the cooperation function fit their results, as shown in Fig. 4-7 (left). In Fig. 4-7 (right), swarm performance results were also analyzed from Rosenfeld et al [83], a second foraging application. Similar to the previous example, a_1 and b values were manually tuned until the swarm performance function fit their results. Unlike the previous example, Rosenfeld et al [83] included interference, so both the cooperation and interference functions were used. This demonstration of parameter extraction provides a degree of validation for the interference function developed in this work. This interference function is intended for applications without extra spatial restrictions, which was the case for the results from Rosenfeld.

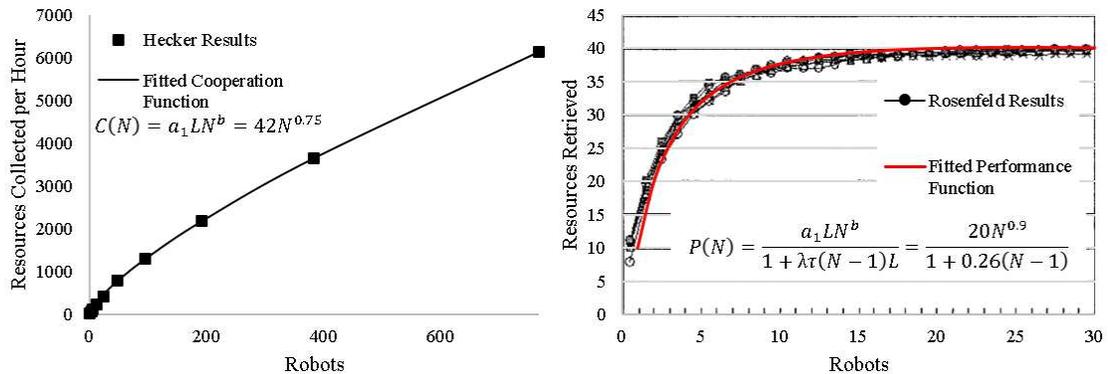


Figure 4-7: (Left) Fig. 11 from Hecker and Moses [47] converted from individual efficiency to swarm performance and then fitted using Hamann’s cooperation function ($a_1 = 42$, $b = 0.75$). Hecker’s results are for a foraging application where objects must be deposited at a central location, and collisions were ignored. Thus, an appropriate cooperation function can be extracted from these results and used in the design process for similar applications. (Right) Fig. 4 from Rosenfeld et al [83] with swarm performance fitted using Hamann’s cooperation function ($a_1 L = 20$, $b = 0.9$) and the interference function developed in this work ($\lambda \tau L = 0.26$). Rosenfeld’s results are for a foraging application where objects do not need to be deposited at a central location and collisions were avoided using several diverse strategies, all with similar performance.

The designer must gather information on how the cost of the robot’s different components are expected to scale with size, or at least know the cost of a single robot if the robot size has already been established. Finally, the designer must estimate the expected delay time for robots to navigate or avoid collisions and find a good estimate for the λ collision proportionality coefficient. A value for this coefficient was found here using simulations in a simplified environment. It is unsatisfying that the designer must currently ‘select’ a reasonable estimate for delay time without good guidelines for that task. Future work will investigate reasonable delay times for specific strategies and test if those delay times are constant for different robot densities.

Several unanswered questions prevent full use of the design approach described in this work. The interference function was developed for applications without spatial restrictions, for which the physical causes of the interference, i.e. potential collision pairs and robot size, were uncovered. The physical causes of interference for applications with spatial restrictions are still unknown, although swarm performance curves for these applications have been well-fitted in previous literature [44]. Furthermore, the interference function developed herein was validated for a random walk. Other swarm behaviors and applications are expected to produce markedly different interference functions, although they may still be a function of potential collision pairs and robot size. The effect of using collision avoidance proximity sensors, which would allow other robots to be detected before an actual collision, is not represented in this model. It is also an open question if the correct λ value could be calculated from known physical parameters, instead of using simulations to find a reasonable value.

4.5.2 Economy of Scale

One of the unexplored points of this work was the potential effect of economies of scale on unit cost, or how an individual robot’s cost is expected to decrease as more and more robots are produced. In economic literature, there is a traditional rule

of thumb called the ‘0.6 rule’ that helps to predict unit cost for different quantities produced [100]. The 0.6 value refers to an exponent for the number of items produced, where any value less than 1 indicates a quantity discount. In the context of swarm robotics, the total swarm cost per the ‘0.6 rule’, $J_{\Sigma, EoS}$, is given in Eqn. (4.19). The effect of applying this rule to the performance-cost curves from Fig. 4-5 is shown in Fig. 4-8.

$$J_{\Sigma, EoS} = N^{0.6} J_I \quad (4.19)$$

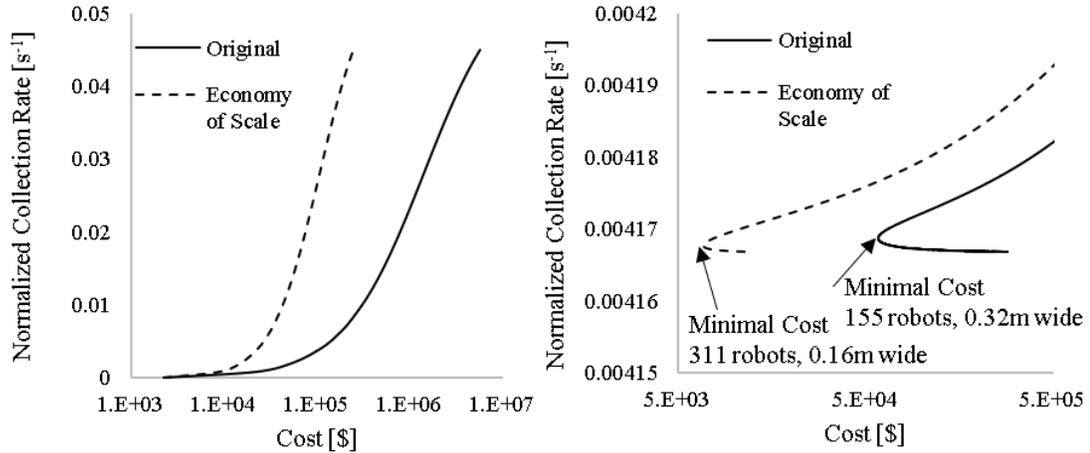


Figure 4-8: Original performance-cost curves from Fig. 4-5, also with a modified curve that includes the cost savings from considering economies of scale for (Left) a constant robot size of 1 m, and (Right) a constant swarm capacity of 50 m. Because of the cost savings in manufacturing higher quantities of robots, the constant capacity case now minimizes cost at a smaller robot width.

Including the estimated impact of economies of scale shifted the minimal-cost robot width for the constant capacity case from 0.32 m (155 robots) to 0.16 m (311 robots). More analysis of the applicability of the ‘0.6 rule’ to the robotics industry would need to be performed, but including the impact of economies of scale should make swarms with many smaller robots more competitive.

4.5.3 Analytical Basis Comparison

Having developed and applied a new analytical basis, it is worthwhile to compare it to Hamann’s example, Eqn. (4.1). His model includes an exponential interference function, which is well-suited for interference trends observed in applications with spatial restrictions. The model developed herein, with a weaker interference function, cannot be fitted to performance trends for those same applications. It can generate a performance curve that peaks and begins to decay, but only for cooperation functions weaker than observed in literature, e.g. with $b \leq 0.2$. Conversely, the new basis can be well-fitted to swarm performance trends in applications with those spatial restrictions removed. The authors have not been able to fit Hamann’s model to these applications, e.g. to the performance curve from Fig. 4-7 (right).

The inability to use the new model for applications with spatial restrictions limits its usefulness and general applicability. It is hypothesized that the physical phenomena used in this new model would also apply to an application with spatial restrictions, but only for fewer numbers of robots, i.e. before the exponential-type term begins to dominate. A more complete model would be usable for applications with or without spatial restrictions, and could partially be achieved by combining these two models. A combined model would still lack a physically-rooted explanation for Hamann’s exponential term that causes the swarm performance to decay.

Outside the explicit field of swarm robotics, Gunther [42] proposed a model for the parallel processing capacity of distributed systems, which was analyzed in the context of swarm robotics in Hamann [45]. Interestingly, one of the models proposed by Gunther matches the form of the swarm performance function independently developed here. His capacity model, which here is being interpreted and presented as a swarm performance function, is given in Eqn. (4.20):

$$P(N) = \frac{N}{1 + \sigma(N - 1)} \quad (4.20)$$

The $(N - 1)$ term was included to represent ‘contention’ over shared resources, and the rationale for this exact form was so that $P(1) = 1$. Given the new claims in this work, it would be useful to re-interpret Gunther’s function in its original context. Specifically, the concept of potential collision pairs could be interpreted in the context of parallel processing.

Note that all of these individual studies involving analytical bases for swarm performance (Gunther [42], Hamann [44], and Hamann [45]) are nicely summarized and contextualized in Hamann [46].

4.5.4 General Discussion

One important open question is how swarm performance varies across different combinations of number of robots and individual robot size for a range of applications. Here, observations from field experiments with real robots were used to estimate an appropriate cooperation function for different combinations. It would be interesting to explore this same research question for different applications. For example, in a collective transport task [86], the transport capability of an individual robot would be expected to increase with its size. Would fewer large robots perform the same as many smaller robots?

Even without knowing the results of these studies, it is clear that a system with many smaller robots would be more robust to failure, and truer to the idea of a swarm, than a system with few large robots. Robustness to failure could be embedded in these analytical bases by including a failure metric such as mean time to failure (MTTF). Fault detection and fault tolerance itself is an active sub-area of research within swarm robotics [9, 99, 18, 19].

The results from these types of studies would also shed light on the potential

advantages of robots operating in formation. Perhaps choosing a small characteristic length could optimize swarm cost, and robots navigating in formation could decrease the interference effect.

One obvious approach for improving swarm performance is to reduce the delay time incurred by the need to avoid collisions. This could be accomplished by adopting an improved collision avoidance strategy, which is only used when a collision is imminent, or by changing to a completely different robot behavior that inherently results in fewer collisions.

The cost function formulated in this work was simple enough that the optimal robot size could be easily found, but a more complex cost function may require more advanced optimization techniques. Another potential outcome of developing a swarm cost model may be the realization that a swarm is not affordable for any combination of robot size and number of robots. In this situation, the effort of generating a platform cost function may help direct efforts in optimizing platform costs, e.g. concentrating on fixed costs or variable costs.

Recall that the cost formulation for the algae collection platform had examples of costs that scale with the robot's width and volume, but not with the robot's area. A solar panel is an example of a component whose cost would scale with area.

Absent thus far has been a discussion of other design considerations which may play a role. There could well be other factors, often conflicting, which restrict the minimum and maximum robot size. For the algae collection example, a larger surface platform would be more resilient to rough water conditions, while a smaller platform would be easier to transport. A designer also does not have a continuum of options for design components. Key components such as motors, sensors, drive wheels, structural parts, etc. are available in discrete sizes which may not match the mathematically optimal size.

Boundary conditions and the behavior of robots at domain boundaries likely affect

swarm performance. A periodic boundary condition was chosen, which allows robots to cross the top boundary and restart at the bottom boundary, as if on a torus. This boundary condition was chosen because it was believed that it would least impact the larger interference trends and it would be representative of the conditions inside the center of a swarm, where boundaries should have a smaller effect. For completeness, the same type of simulation from Fig. 4-2 was repeated with a more realistic boundary. Robots were not allowed to cross boundaries and were forced to choose a new path if they did encounter a boundary. This type of boundary condition is typically referred to as ‘zero-flux’ in the context of mathematics. For the $100 \times 100 \text{ m}^2$ domain, there was no perceivable impact in swarm collision rate when switching to a zero-flux boundary. However, when shrinking the domain to $10 \times 10 \text{ m}^2$, the collision rate for a zero-flux boundary is less than that of the periodic boundary, as shown in Fig. 4-9.

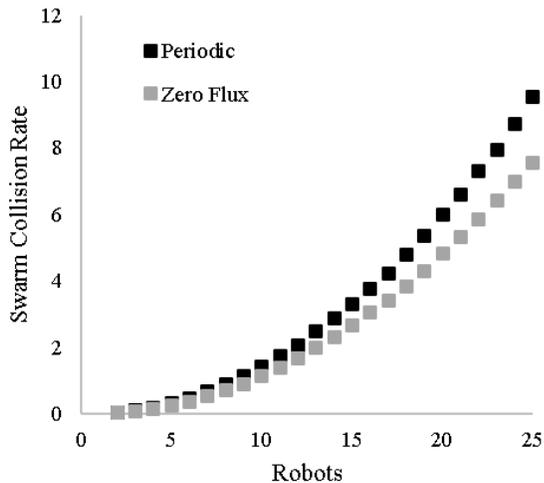


Figure 4-9: Switching from a periodic boundary condition to a zero-flux boundary condition did affect the swarm collision rate for a significantly smaller $10 \times 10 \text{ m}^2$ domain, shown here. It did not affect the swarm collision rate in the original $100 \times 100 \text{ m}^2$ domain.

One final observation is that the swarm design process may be unique in the world of design in that solutions which span many orders of magnitude (10^1 - 10^4 robots, or nanometer-scale to meter-scale robots) could be viable. This applicability across

orders of magnitude enables new, exciting design possibilities, but also introduces new challenges.

4.5.5 Application-Specific Discussion

It is worth noting that the cost formulation for an algae collection robot, Eqn. (4.18), predicted an optimal size of 0.32 m to minimize swarm cost. Based on this cost analysis, and on other design considerations, several new, 2nd-generation robots are being developed and tested at this size, as shown in Fig. 4-10. Using the parameter values for cost obtained during construction of the initial larger platform ($F = \$100$, $c_1 = \$700/m$, $c_3 = \$1500/m^3$), the predicted cost at this smaller size was \$373/robot. These smaller robots actually cost about \$435/robot. In addition to the size change, the conveyor was replaced with a 0.3 m-diameter net, which accounted for three quarters of the total cost.

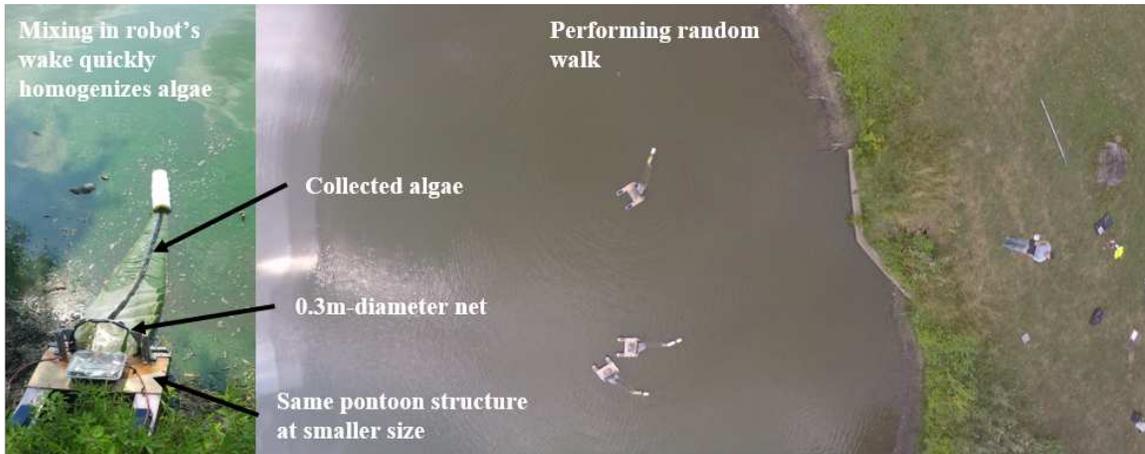


Figure 4-10: Smaller, 2nd-generation algae collection platform whose size was selected to minimize swarm cost. The algae collection conveyor has been replaced with a net towed behind the robot. A total of three robots have been constructed and are currently being tested.

Using a net instead of a conveyor allows both subsurface and surface algae to

be collected. The cost formulation, which was based on a conveyor, predicted that this portion of the platform would cost \$224 at this reduced size. However, the net actually cost \$300. This difference is one reason for the discrepancy between the predicted and actual costs.

The specific results presented for this motivating application included the assumption that the robots are always collecting an average concentration of algae due to mixing induced by the robot. This assumption comes from observing physical collection experiments, although it may be possible to modify the platform to reduce this mixing. If mixing could be reduced, it would motivate more sophisticated collection and foraging behaviors, e.g. a domain partitioning approach or a biased random walk in the direction of the higher algae concentration. With new behaviors, new cooperation and interference functions would also be expected.

4.6 Conclusion and Future Work

This work was motivated by the desire to build a functional swarm of robots to remove harmful algae from the water. Working through the design process exposed knowledge gaps, starting with making the most basic design decisions on how many robots are needed, and what size they should be. The underlying physical phenomena that govern robot interference for this application were suggested and then validated. Starting with these physical phenomena, and including the fundamental parameters of number of robots and individual robot size, an analytical basis for swarm performance was developed. A second knowledge gap in the area of swarm cost was also addressed by making a simple swarm cost formulation that accounted for fixed costs, and costs that vary with individual robot size. Now with bases for both swarm performance and cost, performance-cost curves could be generated for different combinations of the number of robots in the swarm and their individual size.

Two particular scenarios were introduced, referred to as constant size and constant capacity, and performance-cost curves were generated for both. The constant size scenario supposes a constant robot size, where only the number of robots is varied. The constant capacity scenario adds individual robot size as a design choice, where a constant swarm capacity can be allocated to either few large robots, or many small robots. A family of performance-cost curves was generated in a form that facilitated a swarm designer making high-level design decisions about the number of robots and the individual robot size.

This work is an effort toward developing the emerging field of ‘swarm engineering’, which is necessary for robot swarms to be adopted as an engineering solution for real-world problems.

One of the most intriguing questions for future work is establishing how swarms of either few large robots or many small robots perform for the range of swarm applications (not just the particular foraging application in this work). The consequence of viewing robot size as a design choice may open up new possibilities for swarm robotics applications. It would also be worthwhile to expand the investigation of how economies of scale apply to the field of swarm robotics. Finally, the search for underlying physical phenomena that govern robot interference should be expanded to more complex scenarios than the one considered in this work.

Chapter 5

Algae Collection Experiments and Supporting Technology Development

This chapter details the experiments conducted with physical robots to collect real algae. It also gives an overview of the state of supporting technologies (Bio-inspired Vortical Cross-Step Filtration and UAV-based HAB remote sensing). These supporting technologies both remain active research areas, and this is not a comprehensive description of this parallel research, but rather a snapshot of the technology in its current state.

5.1 Algae Collection Experiments

Experiments were conducted at the Scott Park Campus Pond, at the University of Toledo, using up to three prototype robot platforms. The robot platforms had been progressively developed, with the as-tested design shown in Fig.5-1. The robot platforms cost approximately \$100 per platform, and the heart of the platform is a Raspberry Pi Zero W, which has built-in wireless capabilities. A motor control hat

(where 'hat' means a board that sits on top of the raspberry pi) was used to supply power to and control the left and right motors, and a GPS unit was integrated to track robot positions. The robots had no additional sensing capabilities.

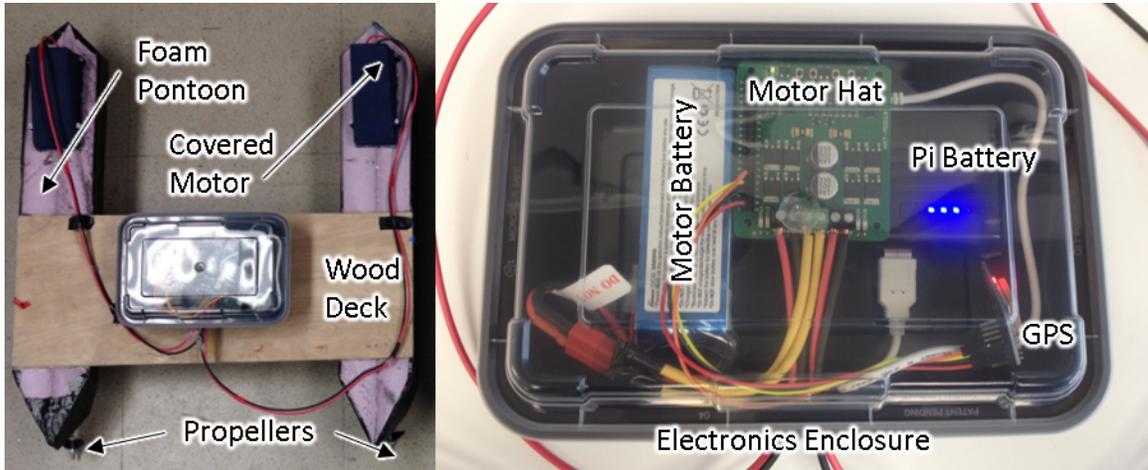


Figure 5-1: (Left) Small boat platform used to perform algae-collection experiments at Scott Park pond and (Right) detailed view of the components inside the electronics enclosure.

The boats use a pontoon structure, with a 300 mm diameter net mounted between the pontoons. The nets (Aquatic Research Instruments) are 1.2 m long, use 100 μm nylon mesh, and terminate in a collection can. The supplied weight for the collection can was removed and replaced with buoyant foam. The net's depth in the water relative to the boat can be manually adjusted, where the shallow Scott Park Pond necessitated that the nets be set to approximately the same depth as the pontoons.

A laptop base station with a wireless router was used to either manually control the robots or initiate an autonomous behavior. A graphical user interface (GUI) for the base station was developed by Tyler Smith, which gives the user the option to control either a single boat, or the entire group.

Preliminary experiments took place in September 2017, and experiments continued through May, June, July, and August 2018. The results described below were

from test days on July 27 2018 and August 8 2018. Fig.5-2 (top) shows an aerial view of three robots performing a random walk. While the pond typically has the highest algae concentrations near the leeward shore , these multiple-boat trials generally took place more towards the center of the pond, where the boats would be less likely to run aground. As such, algae was not collected as quickly for these multi-boat trials.

Single-boat, manually-controlled trials were used close to shore, where the algae was the most concentrated, as shown in Fig.5-2 (bottom). In these areas with high algae concentrations, the net quickly-filled with algae. After the single out and back trial, the boat and net were removed from the water. The algae/water mixture inside the net was agitated so that the water could exit through the net's mesh. This agitation is necessary because the algae collects on the mesh surface, preventing water from passing through. Eventually, a highly-concentrated algae was left in the collection can, with these steps shown in Fig.5-3. This concentration process is similar to the process performed by Dr. Bridgeman's group (Lake Erie Center at the University of Toledo) when they collect vertical plankton net tows during sampling trips on Lake Erie.

These trials demonstrate that algae can be collected, using a net, from a surface platform. For this specific platform and the test conditions, it didn't appear that the robot created a path clear of algae. This could be due to mixing from the robot's hull, net, and propellers all interacting with the water. It could also be due to the relatively shallow water, where algae is present from the surface to the bottom, and the net is only able to capture a portion of the algae. It could also be attributed to colony size, where some colonies are small enough that they may pass through the net. While concentrating algae into a net's collection can, it was observed that some algae passed through the sides of the net.

These robots were not equipped with any algae-detection capability, and the mixing caused by the robot moving through the water suggests that any on-board mea-



Figure 5-2: (top) Aerial view of three boats performing a random walk further from shore, where the algae is less concentrated. The platforms themselves are visible, as well as the collection can at the trailing end of each net. The inset image shows the three boats side by side. (bottom) Single-boat trial in the much denser algae near the shore showing the (left) outbound and (right) inbound legs of the collection run.



Figure 5-3: Steps in the algae concentration process from immediately after the run (top right), to pulling the boat and net from the water (middle right), to concentrating the algae using agitation (left), until the algae is fully concentrated in the collection can (bottom right).

surement of algae concentration or concentration gradient could be quite noisy.

Originally, the nets had been mounted toward the front of the platform, which was hypothesized to minimize the hydrodynamic disturbance from the boat's hull. This never worked well because it caused the front of the boat to sink in the water as the boat's speed increased. Moving the nets to the rear of the boat alleviated this problem.

5.2 Bio-inspired Cross-Step Filtration

Anticipating the negative impact that clogging could have on any algae filtration process, alternative filtration technologies were explored. Sanderson et al [87] originally proposed the idea of a filter inspired by how ram filter-feeding fish can efficiently remove food particles from water. These fish swim for extended periods of time with their mouths open, and must have mechanisms for avoiding clogging.

Initial efforts to develop these types of filters were performed with Lauren Marshall [66]. Collaboration with Dr. Sanderson began in April 2018, with some of the upper-level highlights of this ongoing research described below.

This new filtration process is called vortical cross-step filtration. It is called this because a 'backward-facing' step creates a recirculation area, or vortex, inside a slot as shown in Fig.5-4. A nylon mesh is added between across the slot, which allows water to pass through, but not particles. The recirculation that forms in this slot is the key mechanism that prevents clogging as it scours the center of the mesh. This allows particles to accumulate on the upstream and downstream sides of the mesh, but not in the middle.

Over the course of the filter development, many different filters were tested, but only two are described here. The filters were tested in a water tunnel and flow and collection performance were evaluated qualitatively using both dye injection and by

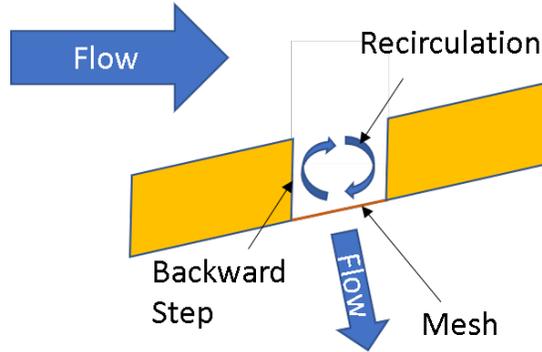


Figure 5-4: Basic cross-step vortical filtration mechanisms. A recirculation pattern forms in the slot which scours the middle of the nylon mesh, delaying clogging.

seeding the tank with microspheres. The neutrally-buoyant fluorescent microspheres (Cospheric) were chosen to be the same size as the algae colonies (106-125 μm).

The first filter, shown in Fig.5-5 (top), is based on Sanderson's original filter. It is different in that the version shown is approximately double the scale, has nine constant-width slots (not five variable width), and is perfectly conical (less representative of original paddlefish-inspiration). It also uses a 100 μm mesh (not 140 μm), which only has 44% open area (instead of 55%).

This filter was able to generate the same type of vortices between the slots as observed in Sanderson's original filters. It is also clear that particles accumulate on the upstream and downstream sides of the slots, which delays clogging. However, clogging will eventually occur, because there is no mechanism to remove the particles from the slots. The second filter remedies this by using a helical slot that extends from the upstream to the downstream side of the filter, shown in Fig.5-5 (middle).

The version shown has eight helical slots, each with an 120 mm pitch. As before, particles accumulate on the upstream and downstream sides of the slot. However, the vortex that forms in the slot also has an axial component that travels along the helical slot. This axial component enables particles to be transported from the upstream side of the filter, to the downstream side. Injecting dye in one of the slots

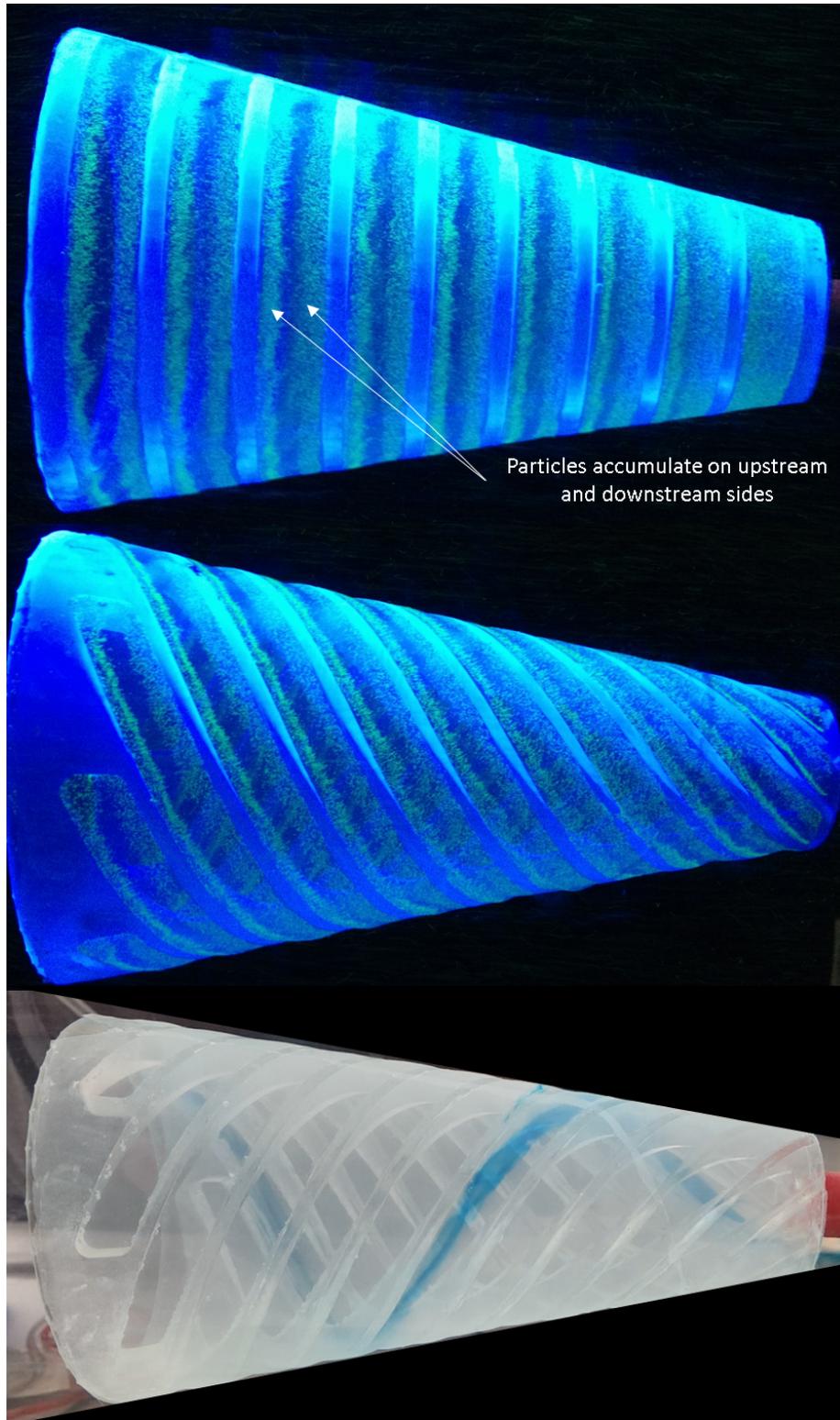


Figure 5-5: Particles accumulate on the upstream and downstream sides of the slots for a filter similar to Sanderson's (top) and for a new helical design (middle). An axial component to the slot vortices (blue dye) transports particles downstream (bottom).

near the upstream side of the filter allows this axial component to be visualized, as shown in Fig.5-5 (bottom). To enable transport, particles that have been deposited on the mesh must be re-suspended. In practice, this has been accomplished by tapping or perturbing the filter.

In addition to tapping, rotating the filters has also had the effect of transporting particles within the helical slots, either upstream or downstream, depending on the direction of rotation.

This is only a glimpse into the filters that have been tested, and the parameter space for this helical filter design is quite large, e.g. conical angle, helix pitch, and slot width to height ratio. The next critical steps in this filter's development are validating that real algae behave similar to the algae-sized, neutrally-buoyant particles, and designing an interface at the downstream side of the filter to capture particles traveling through the helical slots.

5.3 UAV-based HAB remote sensing

Equipping each robot with the ability to measure its local algae concentration would be costly, and the data collected may be too noisy to be useful. For this reason, the possibility of equipping an unmanned aerial vehicle (UAV) with an algae remote-sensing instrument was investigated.

A hexacopter kit (Aqua H20) was purchased and assembled, which included a fiberglass, waterproof body and foam pontoons that allowed it to land on water (see Fig.5-6). The hexacopter comes with a transparent bubble canopy, which was replaced with a custom canopy. The new canopy was needed to give a sufficiently-large field of view for a multi-spectral camera system (Tetracam RGB+3). A 3D-printed bracket was fabricated to house this multi-spectral instrument, as well as a GoPro Hero 4 Black.



Figure 5-6: (Left) Hexacopter body with Tetracam RGB+3 and Go Pro mounted in a custom bracket, and (Right) Hexacopter landing on the water.

The RGB+3 is a four-camera system, where one camera is a standard red/green/blue camera (RGB) and the other three cameras have band-pass filters. Filters were centered at 660, 680, and 710 nm, which were chosen to match MERIS, a former ESA satellite. Choosing these particular wavelengths allows the scenes to be processed to create the cyanobacteria index (CI). This instrument was partially funded by a Small Ohio Sea Grant.

The hexacopter was primarily flown from the Lake Erie Center boat, on six sampling trips from June-August 2018, and at the Scott Park Campus Pond (same time period). A custom, folding landing deck was constructed to launch and recover the hexacopter from the boat, as shown in Fig.5-7.

The RGB+3 was set to capture an image every ten seconds, and an example of a scene captured by the RGB+3's band-filtered cameras is given below in Fig.5-8 (top). For each flight, the GoPro was also set to record video at 30 frames per second (fps) and 4k resolution, with an example still image shown in Fig.5-8 (bottom). Flights were typically conducted up to an altitude of 40 m.

Work is continuing on creating CI-scenes, which requires the Tetracam images to first be converted to reflectance. An example processed scene from the Scott Park



Figure 5-7: Flights from the LEC boat were launched and recovered using a deployable flight deck.

Pond is shown in Fig.5-9. One of the difficulties in processing these images is dealing with sun glint and reflections on the water, which are known issues for remote sensing applications.

After the CI scenes have been generated and vetted, they can be compared to remote sensing data collected by same-day NASA Glenn flyovers and to data collected by Dr. Bridgeman's group at each site (biovolume, fluorometry data, turbidity, etc.)

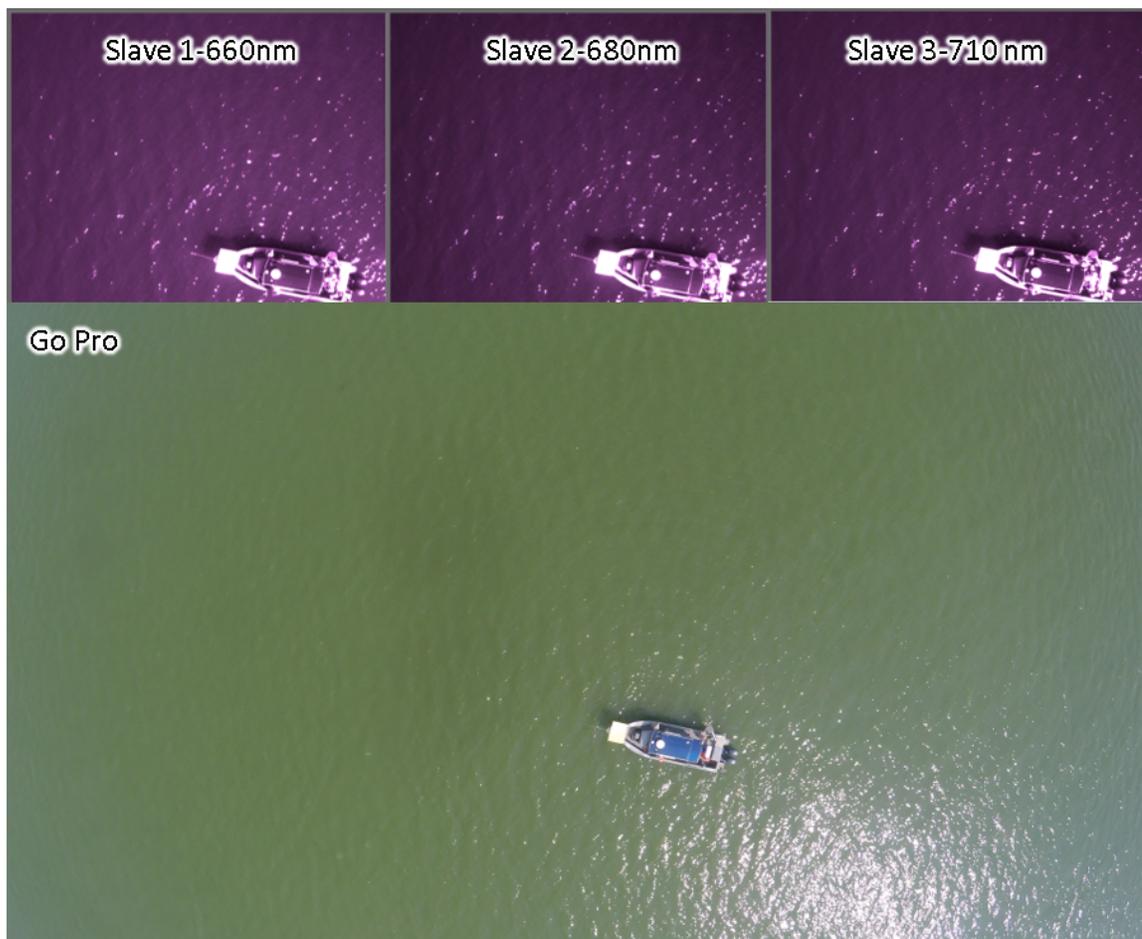


Figure 5-8: Band-filtered images from Tetracam RGB+3 (top) and still frame from GoPro (bottom). The GoPro has a much wider FOV.

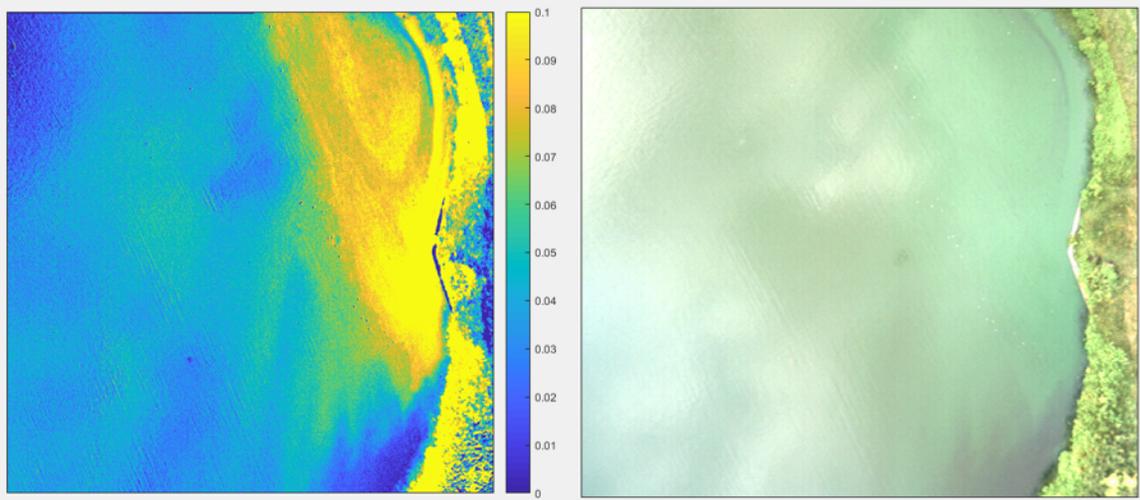


Figure 5-9: Cyanobacteria Index (left) compared to a true color image (right) collected on August 10 2018 at the Scott Park Pond.

Chapter 6

Conclusion

This final section compiles the conclusions from the previous sections, reiterates the major contributions of this work, comments on future work, and concludes with an assessment of the readiness level of this technology and a final conclusion.

6.1 Cumulative Conclusions

The major conclusion from Section III (Algae Collection Simulations) was that performing an unbiased random walk was a viable controls approach to collect a diffusive substance. Although other, more sophisticated approaches (biased random walk, partitioning) could collect algae more quickly, these approaches imposed additional requirements on the robot platform. A biased random walk would need to detect the local algae concentration, which substantially increases robot cost, and a partitioning approach requires robots to travel in controlled, straight paths, which would be difficult in a wavy, turbulent environment.

Also from Section III, it was shown that collection of a diffusive substance is viable for a range of diffusivities, which can be equated to varying levels of turbulence in the water. There was not a significant difference in collection rate between many, small robots and few, large robots, at least until the number of robots became quite small. Also, it was shown how for a partitioning approach, as the number of robots increases,

and the size of each partition becomes smaller, that robots spend more time at their partitions boundaries, which would require more collision avoidance behaviors.

The major conclusion from Section IV (Swarm Interference Simulations and Cost) was that an optimal robot size could be chosen that would minimize swarm cost, and that swarm performance is only expected to increase as more robots are added, despite increasing robot-to-robot interference. This is true as long as robots can process algae on board, and aren't required to deposit algae at a central location. Although adding more robots continues to increase the collection rate, there are diminishing returns as more and more robots are added.

The major conclusion from Section V (Physical Experiments) is that algae can be collected using a robot platform, and that the platform and collection process disrupts the water, having an overall mixing effect on the algae present. It may be possible to modify the collection process to minimize this mixing, which would motivate more sophisticated collection behaviors. The major technological barrier that needs to be overcome is the clogging of the algae-collection nets.

6.2 Major Contributions

The major contributions of this work will be reiterated from most general to most specific.

This work was part of a much larger effort in the field of swarm robotics to transform robot swarms from lab-only demonstrations to field-capable systems. This work contributed to this larger effort **(i)** by exploring swarm performance-cost relationships and tradeoffs, **(ii)** by including robot size as a design input, which influences swarm performance and cost, and by **(iii)** holistically considering both the swarm platform's physical design alongside the control law development. The simple model for swarm cost, which views robot cost as a combination of fixed costs and robot-size-dependent

costs, is a powerful and flexible tool for exploring the swarm design space.

At an application-specific level, this work contributed understanding to the novel problem of swarm collection of a continuously-distributed substance (not just foraging for discrete objects). Furthermore, it introduced the additional wrinkle of collecting a diffusive substance, which could be applied to other materials embedded in turbulent fluids, e.g. microplastics in fresh water and oceans.

At a yet more specific level, it contributed an analytical basis for interference between robots in the swarm, based on the physical parameters of number of potential collision pairs (a mathematical combination) and robot size, for the case where robots do not need to deposit any materials in a central location. It also was able to uncover a relationship between partition's perimeter to area ratio as the number of robots and partitions increases, in a Voronoi partitioning scheme.

Finally, the physical experiments performed (swarm algae collection, bio-inspired cross-step filtration, UAV-based HAB remote sensing) were all important contributions. The swarm algae collection experiments are a proof of concept of the overall idea of collecting harmful algae using a swarm of robots. The bio-inspired cross-step filtration is a promising technology for not only algae-collection, but also for other filtration applications where clogging is an issue. UAV-based HAB remote sensing could be used for directing a swarm of surface robots to problem areas, and can also be used to generate a new data product for environmental scientists and water management professionals.

6.3 Readiness Level Assessment

There are several technological areas that need to be further developed. In the current form, a swarm built after this model could only collect algae for a relatively short amount of time. The robots have a limited battery life, and so would need

either a way of continuously recharging its batteries (e.g. solar panels), or a system for exchanging spent batteries for fresh batteries. In the current form, robots can only collect algae, but can neither process it on-board, nor deposit it somewhere else. There is also not a mature human-swarm interface for communicating higher level commands to the swarm, e.g. using knowledge of bloom locations to direct the swarm where to work.

The robot platform itself would need to be modified to operate in rougher waters, meaning improving resistance to capsizing, or adding the ability to continue operating if capsized, and sufficient motor power to navigate larger currents and waves.

These areas that need further work are certainly not insurmountable, and are engineering problems that, with attention, could be solved. To motivate this further effort, it is worthwhile to reexamine where this type of swarm could be used

It was initially suggested that this swarm could be deployed around a drinking water intake. Since the 2014 water crisis, water treatment plants have upgraded their facilities, adapted their treatment strategies, and now are more capable in early detection of blooms. Given these improvements, they could likely treat almost any bloom. This means that adding a robot swarm would not change whether or not a treatment facility could treat a bloom. There are significant costs with treating these blooms, as were detailed in Section II. Because these costs are known, a cost argument could be made for using a robot swarm, but it would have to be significantly less expensive to replace established treatment technologies.

The strength of using a swarm is the ability to scale up the swarm size to fit the application. To take advantage of this strength, a swarm should be used to tackle a larger area. If a swarm is used in this way, it is offering a capability that conventional technologies cannot match.

6.4 Future Work

Walking through this research process has uncovered many diverse avenues for future work. The areas include extensions of the current lines of research, as well as new lines of research, with both types listed below.

The physical system that was tested only comprised three robots. A swarm with more robots would need to be tested to gain a better understanding of the complex interaction between the swarm and the environment. The three-robot system was tested for a limited duration under close supervision. It would be necessary to run this system for a longer duration test, eventually unsupervised, to better assess system performance. Finally, the surface boats and UAV were used simultaneously, but without communicating with one another. Adding communication between these two robot types would enable more sophisticated collection behaviors without the additional cost of equipping each boat with algae-sensing capability.

One of the most intriguing areas is the application of this approach to similar problems, for example, collecting microplastic particles in oceans. One challenge with this application could be the large range of particle sizes, or additional challenges stemming from also trying to collect microplastic fibers.

Another interesting idea raised was considering robot size as a design input, where a swarm could comprise many, small robots, or few, large robots. This was explored for this specific application, but it would be quite interesting to explore it for other swarm applications, e.g. collective transport.

A major lesson learned from the interference research was how swarm performance curves are different when robots are or are not required to return collected objects to a central location. It would be quite interesting to investigate how performance transitions from one case to another by starting with only a single deposition point, but continuing to add additional deposition points, distributed throughout the domain.

6.5 Final Conclusion

Using a robot swarm to mitigate harmful algal blooms remains an intriguing option and merits further exploration. There remain significant, yet not insurmountable, engineering problems which need to be addressed to create a working system. Bringing this robot swarm to a mature, operational state would require a larger, dedicated team and additional financial resources. Proof of commercial viability would be the next step in justifying this next level of personnel and material investment.

References

- [1] MIT Senseable Lab's Seaswarm. URL <http://senseable.mit.edu/seaswarm/>
- [2] Relating chlorophyll from cyanobacteria-dominated inland waters to a MERIS bloom index. *Remote Sensing Letters* **7**(2), 141–149 (2016)
- [3] Aparicio Medrano, E., van de Wiel, B.J.H., Uittenbogaard, R.E., Dionisio Pires, L.M., Clercx, H.J.H.: Simulations of the diurnal migration of *Microcystis aeruginosa* based on a scaling model for physical-biological interactions. *Ecological Modelling* **337**(October), 200–210 (2016)
- [4] Ara, S., Irwin, E., Haab, T.: Measuring the Effects of Lake Erie Water Quality in Spatial Hedonic Price Models. In: *Environmental and Resource Economists, Third World Congress, October 2015. Kyoto, Japan* (2006)
- [5] Atkings, R., Rose, T., Brown, R., Robb, M.: The *Microcystis* cyanobacteria bloom in the Swan River - February 2000. *Water Science and Technology* **43**(9), 107–114 (2001)
- [6] Aznar, F., Sempere, M., Pujol, M.J., Rizo, R., Pujol, M.J.: Modelling oil-spill detection with swarm drones. *Abstract and Applied Analysis* **2014**, 1–14 (2014)
- [7] Bayindir, L.: A Review of Swarm Robotics Tasks. *Neurocomputing* **172**(8), 292–321 (2015)

- [8] Becker, R.H., Sultan, M.I., Boyer, G.L., Twiss, M.R., Konopko, E.: Mapping cyanobacterial blooms in the Great Lakes using MODIS. *Journal of Great Lakes Research* **35**(3), 447–453 (2009)
- [9] Bjercknes, J.D., Winfield, A.F.T.: On fault tolerance and scalability of swarms. In: Martinoli A. et al. (eds) *Distributed Autonomous Robotic Systems, Springer Tracts in Advanced Robotics*, vol. 83, pp. 431–444. Springer (2013)
- [10] Bonabeau, E., Dorigo, M., Theraulaz, G.: *Swarm Intelligence: From Natural to Artificial Systems*. University Press, New York (1999)
- [11] Brambilla, M., Brutschy, A., Dorigo, M., Birattari, M.: Property-driven Design for Robot Swarms: A Design Method Based on Prescriptive Modeling and Model Checking. *ACM Transactions on Autonomous and Adaptive Systems* **9**(4), 17:1–17:28 (2014)
- [12] Brambilla, M., Ferrante, E., Birattari, M., Dorigo, M.: Swarm robotics : a review from the swarm engineering perspective. *Swarm Intelligence* **7**, 1–41 (2013)
- [13] Bridgeman, T.B., Chaffin, J.D., Filbrun, J.E.: A novel method for tracking western Lake Erie Microcystis blooms, 2002-2011. *Journal of Great Lakes Research* **39**(1), 83–89 (2013)
- [14] Bryan, K.: *Apparatus for Harvesting Algae from Open Body of Water* (2015)
- [15] Cao, M., Hadjicostis, C.: Distributed algorithms for Voronoi diagrams and application in ad-hoc networks. Tech. rep., UIUC Coordinated Science Laboratory (2003)
- [16] Chamanbaz, M., Mateo, D., Zoss, B.M., Toki, G., Bouffanais, R.: Swarm-

- Enabling Technology for Multi-Robot Systems. *Front. Robot. AI* **4**(April), 1–12 (2017)
- [17] Chow, C., Panglisch, S., House, J., Gimbel, R.: A Study of Membrane Filtration for the Removal of Cyanobacteria Cells. *Journal of Water Supply Research and Technology* **46**(6), 324–334 (1997)
- [18] Christensen, A.L., O’Grady, R., Birattari, M., Dorigo, M.: Exogenous Fault Detection in a Collective Robotic Task. In: Almeida e Costa F., Rocha L.M., Costa E., Harvey I., Coutinho A. (eds) *Advances in Artificial Life. ECAL 2007*. LNCS 4648, pp. 555–564. Springer, Berlin (2007)
- [19] Christensen, A.L., O’Grady, R., Dorigo, M.: From fireflies to fault-tolerant swarms of robots. *IEEE Transactions on Evolutionary Computation* **13**(4), 754–766 (2009)
- [20] Christensen, A.L., Oliveira, S.M., Postolache, O., João de Oliveira, M., Sargento, S., Santana, P., Nunes, L., Velez, F.J., Sebastião, P., Costa, V., Duarte, M., Gomes, J.C., Rodrigues, T., Silva, F.: Design of Communication and Control for Swarms of Aquatic Surface Drones. In: *Proceedings of the International Conference on Agents and Artificial Intelligence (ICAART-2015)*, pp. 548–555. SCITEPRESS, Lisbon, Portugal (2015)
- [21] Cole, M., Lindeque, P., Halsband, C., Galloway, T.S.: Microplastics as contaminants in the marine environment: A review. *Marine Pollution Bulletin* **62**(12), 2588–2597 (2011)
- [22] Conroy, J.D., Kane, D.D., Dolan, D.M., Edwards, W.J., Charlton, M.N., Culver, D.A.: Temporal Trends in Lake Erie Plankton Biomass: Roles of External Phosphorus Loading and Dreissenid Mussels. *Journal of Great Lakes Research* **31**, 89–110 (2005)

- [23] Costa, V., Duarte, M., Rodrigues, T., Oliveira, S.M., Christensen, A.L.: Design and development of an inexpensive aquatic swarm robotics system. In: Proc IEEE MTS Oceans Conference, pp. 1–8. IEEE, Shanghai (2016)
- [24] Das, J., Py, F., Maughan, T., O’Reilly, T., Messié, M., Ryan, J., Rajan, K., Sukhatme, G.S.: Coordinated Sampling of Dynamic Oceanographic Features with Underwater Vehicles and Drifters. Springer Tracts in Advanced Robotics **79**, 541–555 (2014)
- [25] Davenport, T., Drake, W.: EPA Commentary: Grand Lake St. Marys, Ohio - The Case for Source Water Protection: Nutrients and Algae Blooms. LakeLine **31**(3), 41–46 (2011)
- [26] Davenport, T., Gibson, R., Mount, T.: Implementing Grand Lake St . Marys Nutrient TMDL. Tech. rep., USEPA/Ohio EPA (2010)
- [27] Dawson, S., Wellman, B.L., Anderson, M.: Categorizing interference in real robot experiments. In: Proceedings of the 2011 IEEE International Conference on Systems, Man, and Cybernetics, pp. 3561–3565. IEEE (2011)
- [28] Dimidov, C., Oriolo, G., Trianni, V.: Random Walks in Swarm Robotics: An Experiment with Kilobots. In: Proceedings of the 10th International Conference on Swarm Intelligence, pp. 185–196 (2016)
- [29] Dorigo, M., Birattari, M.: Swarm Intelligence. Scholarpedia **2**(9), 1462 (2007)
- [30] Dorigo, M., Birattari, M., Brambilla, M.: Swarm Robotics. Scholarpedia **9**(1), 1463 (2014)
- [31] Driedger, A.G.J., Dürr, H.H., Mitchell, K., Van Cappellen, P.: Plastic debris in the Laurentian Great Lakes: A review. Journal of Great Lakes Research **41**(1), 9–19 (2015)

- [32] Duarte, M., Costa, V., Gomes, J.C., Rodrigues, T., Silva, F., Oliveira, S.M., Christensen, A.L.: Evolution of Collective Behaviors for a Real Swarm of Aquatic Surface Robots. *PLOS ONE* **11**(3) (2016)
- [33] Duarte, M., Gomes, J.C., Costa, V., Rodrigues, T., Silva, F., Lobo, V., Marques, M., Oliveira, S.M., Christensen, A.L.: Application of Swarm Robotic Systems to Marine Environmental Monitoring. *Proceedings of IEEE/MTS OCEANS* pp. 1–8 (2016)
- [34] Dunbabin, M., Grinham, A., Udy, J.: An autonomous surface vehicle for water quality monitoring. *Australasian Conference on Robotics and Automation (ACRA)* pp. 2–4 (2009)
- [35] Dunbabin, M., Udy, J., Bruenig, M.: Continuous monitoring of reservoir water quality: the Wivenhoe Project. *Journal of the Australian Water Association* **36**(September), 74–77 (2009)
- [36] Edzwald, J.K.: Developments of high rate dissolved air flotation for drinking water treatment. *Journal of Water Supply: Research and Technology AQUA* **56**(67), 399 (2007)
- [37] EPA, U.: A Compilation of Cost Data Associated with the Impacts and Control of Nutrient Pollution. Tech. rep., US EPA Office of Water (2015)
- [38] Eriksen, M., Mason, S., Wilson, S., Box, C., Zellers, A., Edwards, W.J., Farley, H., Amato, S.: Microplastic pollution in the surface waters of the Laurentian Great Lakes. *Marine Pollution Bulletin* **77**(1-2), 177–182 (2013)
- [39] Fischer, H.B., List, E.J., Koh, R.C.Y., Imberger, J., Brooks, N.H.: Mixing in inland and coastal water. Academic Press Inc, New York (1979)

- [40] Gaudiano, P., Shargel, B., Bonabeau, E., Clough, B.T.: Swarm Intelligence: A New C2 Paradigm with an Application to Control Swarms of UAVs. Tech. rep., Ft. Belvoir Defense Technical Information Center (2003)
- [41] Guerrero, J., Oliver, G., Valero, O.: Multi-robot coalitions formation with deadlines: Complexity analysis and solutions. *PLOS ONE* **12**(1), 1–27 (2017)
- [42] Gunther, N.J.: A Simple Capacity Model of Massively Parallel Transaction Systems. In: CMG National Conference, pp. 1035–1044 (1993)
- [43] Hajieghrary, H., Hsieh, M.A., Schwartz, I.B.: Multi-Agent Search for Source Localization in a Turbulent Medium. *Physics Letters A* **380**(20), 1698–1705 (2016)
- [44] Hamann, H.: Towards Swarm Calculus: Universal Properties. In: Swarm Intelligence. ANTS 2012. Lecture Notes in Computer Science, vol 7461, pp. 168–179 (2012)
- [45] Hamann, H.: Superlinear scalability in parallel computing and multi-robot systems: Shared resources, collaboration, and network topology. In: B. M., B. R., H. H., K. D., P. T. (eds.) Architecture of Computing Systems ARCS 2018. ARCS 2018. Lecture Notes in Computer Science, vol. 10793 LNCS, pp. 31–42. Springer, Cham (2018)
- [46] Hamann, H.: Swarm Robotics: A Formal Approach. Springer, Cham (2018)
- [47] Hecker, J.P., Moses, M.E.: Beyond pheromones: evolving error-tolerant, flexible, and scalable ant-inspired robot swarms. *Swarm Intelligence* **9**(1), 43–70 (2015)
- [48] Hitz, G., Gotovos, A., Pomerleau, F., Garneau, M.É., Pradalier, C., Krause, A., Siegwart, R.Y.: Fully autonomous focused exploration for robotic environ-

- mental monitoring. Proceedings - IEEE International Conference on Robotics and Automation pp. 2658–2664 (2014)
- [49] Hitz, G., Pomerleau, F., Garneau, M.É., Pradalier, C., Posch, T., Pernthaler, J., Siegwart, R.Y.: Autonomous Inland Water Monitoring: Design and Application of a Surface Vessel. *IEEE Robotics & Automation Magazine* **19**(1), 62–72 (2012)
- [50] Ho, J.C., Michalak, A.M.: Phytoplankton blooms in Lake Erie impacted by both long-term and springtime phosphorus loading. *Journal of Great Lakes Research* **43**(3), 221–228 (2017)
- [51] Huang, C., Shi, K., Yang, H., Li, Y., Zhu, A.x., Sun, D., Xu, L.: Satellite observation of hourly dynamic characteristics of algae with Geostationary Ocean Color Imager (GOCI) data in Lake Taihu. *Remote Sensing of Environment* **159**, 278–287 (2015)
- [52] Inosako, K., Yoshida, I., Tarui, Y., Nakagawa, R.: Removal of the suspended solid composed of algae by the dissolved air floatation method. *Tottori Daigaku Nogakubu Kenkyu Hokoku* (1997)
- [53] Kashian, R., Kasper, J.: Tainter Lake & Lake Menomin: The Impact of Diminishing Water Quality on Value. Tech. rep., Department of Economics-University of Wisconsin (2010)
- [54] Khaluf, Y., Dorigo, M.: Modeling Robot Swarms Using Integrals of Birth-Death Processes. *ACM Transactions on Autonomous and Adaptive Systems* **11**(2), 1–16 (2016)
- [55] Krysel, C., Boyer, E.M., Parson, C., Welle, P.: Lakeshore Property Values and Water Quality: Evidence from property sales in the Mississippi Headwaters Region. Tech. rep., Submitted to the Legislative Commission on Minnesota

Resources By the Mississippi Headwaters Board and Bemidji State University.
(2003)

- [56] Laut, J., High, B., Nov, O., Porfiri, M.: A Robotic Vehicle for Aquatic Environmental Monitoring. In: Proceedings of the Asme 8Th Annual Dynamic Systems and Control Conference, 2015, Vol 3, pp. 1–7 (2016)
- [57] Lekki, J., Anderson, R., Avouris, D., Becker, R.H., Churnside, J., Cline, M., Demers, J., Leshkevich, G., Liou, L., Luvall, J., Ortiz, J., Royce, A., Ruberg, S., Sawtell, R., Sayers, M.J., Schiller, S., Dakota, S., Shuchman, R.A., Simic, A., Green, B., Stuart, D., Sullivan, G., Tavernelli, P., Tokars, R., Woude, A.V.: Airborne Hyperspectral Sensing of Harmful Algal Blooms in the Great Lakes Region : System Calibration and Validation From Photons to Algae Information : The Processes In-Between. Tech. Rep. February, NASA (2017)
- [58] Lekki, J., Anderson, R., Nguyen, Q.v., Demers, J., Leshkevich, G., Flatico, J., Kojima, J.: Development of Hyperspectral Remote Sensing Capability for the Early Detection and Monitoring of Harmful Algal Blooms (HABs) in the Great Lakes. In: Proceedings of the AIAA Infotech Aerospace Conference, April, pp. 1–14 (2009)
- [59] Lerman, K., Galstyan, A.: Mathematical Model of Foraging in a Group of Robots: Effect of Interference. *Autonomous Robots* pp. 127–141 (2001)
- [60] Liekna, A., Grundspenkis, J.: Towards practical application of swarm robotics: Overview of swarm tasks. *Engineering for Rural Development* **13**, 271–277 (2014)
- [61] Lloyd, S.P.: Least Squares Quantization in PCM. *IEEE Transactions on Information Theory* **28**(2), 129–137 (1982)

- [62] Low, K.H., Podnar, G.W., Stancliff, S., Dolan, J.M., Elfes, A.: Robot Boats as a Mobile Aquatic Sensor Network. In: Proceedings of the IPSN-09 Workshop on Sensor Networks for Earth and Space Science Applications (2009)
- [63] Majid, M.H., Arshad, M.R.: Hydrodynamic Effect on V-Shape Pattern Formation of Swarm Autonomous Surface Vehicles (ASVs). *Procedia Computer Science* **76**(Iris), 186–191 (2015)
- [64] Makarewicz, J.C.: Phytoplankton Biomass and Species Composition in Lake Erie 1970-1987.pdf. *Journal of Great Lakes Research* **19**(2), 258–274 (1993)
- [65] Marques, L., Nunes, U., De Almeida, A.T.: Particle swarm-based olfactory guided search. *Autonomous Robots* **20**(3), 277–287 (2006)
- [66] Marshall, L., Schroeder, A., Trease, B.: Comparing Fish-Inspired Ram Filters for Collection of Harmful Algae. In: Proceedings of the ASME 2018 International Mechanical Engineering Congress and Exposition IMECE2018, pp. 1–9 (2018)
- [67] Mataric, M.J.: Controlling a Mobile Robot Herd : Theory and Practice. Tech. rep., Applications of Artificial Intelligence to Real-World Autonomous Mobile Robots. Papers from the 1992 Fall Symposium, Technical Report FS-92 (1992)
- [68] McLurkin, J., Lynch, A.J., Rixner, S., Barr, T.W., Chou, A., Foster, K., Bilstein, S.: A low-cost multi-robot system for research, teaching, and outreach. In: Martinoli A. et al. (eds) Distributed Autonomous Robotic Systems, DARS2010, *Springer Tracts in Advanced Robotics*, vol. 83, pp. 597–609. Springer (2012)
- [69] Ming, T., Vanderploeg, H.A., Rowe, M.D., Fanslow, D.L., Strickler, J.R., Miller, R., Johengen, T.H., Davis, T.W., Gossiaux, D.: Effect of light exposure and nutrients on buoyancy of *Microcystis* colonies (2017)

- [70] Nakamura, T., Adachi, Y., Suzuki, M.: Flotation and sedimentation of a single microcystis floc collected from surface bloom. *Water Research* **27**(6), 979–983 (1993)
- [71] Newcombe, G.: International Guidance Manual for the Management of Toxic Cyanobacteria. Tech. rep., Global Water Research Coalition (2009)
- [72] Nurzaman, S.G., Matsumoto, Y., Nakamura, Y., Koizumi, S., Ishiguro, H.: Biologically Inspired Adaptive Mobile Robot Search With and Without Gradient Sensing. In: IROS 2009. Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009, pp. 142–147. IEEE (2010)
- [73] Paerl, H.W.: Mitigating Harmful Cyanobacterial Blooms in a Human- and Climatically-Impacted World. *Life* **4**(4), 988–1012 (2014)
- [74] Pashna, M., Yusof, R., Yazdani, S.: Analysis of Prediction Methods for Swarm Robotic - In the Case of Oil Spill Tracking. In: Proceedings of 2015 10th Asian Control Conference (ASCC), pp. 1–6 (2015)
- [75] Pavone, M., Arsie, A., Frazzoli, E., Bullo, F.: Equitable Partitioning Policies for Mobile Robotic Networks. *IEEE Transactions of Automatic Control* **56**(8), 1834–1848 (2011)
- [76] Philamore, H., Rossiter, J., Stinchcombe, A., Ieropoulos, I.: Row-bot : An Energetically Autonomous Artificial Water Boatman. In: Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3888–3893 (2015)
- [77] Pinciroli, C., Trianni, V., Grady, R.O., Pini, G., Brutschyt, A., Brambilla, M., Mathews, N., Ferrante, E., Caro, G.D., Ducatelle, F., Stirling, T., Gutiérrez, A., Gambardellai, L.M., Dorigo, M.: ARGoS: a Modular, Multi-Engine Simulator for Heterogeneous Swarm Robotics. *Swarm Intelligence* **6**(4), 271–295 (2012)

- [78] Podnar, G.W., Dolan, J.M., Elfes, A., Stancliff, S., Lin, E., Hosler, J.C., Ames, T.J., Moisan, J., Moisan, T.A., Higinbotham, J., Kulczycki, E.A.: Operation of Robotic Science Boats Using the Telesupervised Adaptive Ocean Sensor Fleet System. In: Proceedings - IEEE International Conference on Robotics and Automation (2008)
- [79] Purves, W.K., Orians, G.H., Heller, H.C.R.: Life: The Science of Biology, 4th edn. Sinauer Associates, Sunderland, Massachusetts (1994)
- [80] Raikow, D.F., Sarnelle, O., Wilson, A.E., Hamilton, S.K.: Dominance of the noxious cyanobacterium *Microcystis aeruginosa* in low-nutrient lakes is associated with exotic zebra mussels. *Limnology and Oceanography* **49**(2), 482–487 (2004)
- [81] Reynaud, L., Guérin-Lassous, I.: Design of a force-based controlled mobility on aerial vehicles for pest management. *Ad Hoc Networks* **53**, 41–52 (2016)
- [82] Roberts, P.J.W., Webster, D.R.: Turbulent diffusion. *Physical review letters* **90**, 1–42 (2003)
- [83] Rosenfeld, A., Kaminka, G.A., Kraus, S.: A study of scalability properties in robotic teams. In: Scerri P., Vincent R., Mailler R. (eds) *Coordination of Large-Scale Multiagent Systems*, pp. 27–51. Springer (2006)
- [84] Rowe, M.D., Anderson, E.J., Wynne, T.T., Stumpf, R.P., Fanslow, D.L., Kijanka, K., Vanderploeg, H.A., Strickler, J.R., Davis, T.W.: Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *Journal of Geophysical Research: Oceans* **121**(7), 5296–5314 (2016)
- [85] Rubenstein, M., Ahler, C., Hoff, N., Cabrera, A., Nagpal, R.: Kilobot: A low

- cost robot with scalable operations designed for collective behaviors. *Robotics and Autonomous Systems* **62**(7), 966–975 (2014)
- [86] Rubenstein, M., Cabrera, A., Werfel, J., Habibi, G., McLurkin, J., Nagpal, R.: Collective transport of complex objects by simple robots. In: *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 47–54 (2013)
- [87] Sanderson, S.L., Roberts, E., Lineburg, J., Brooks, H.: Fish mouths as engineering structures for vortical cross-step filtration. *Nature communications* **7**, 11,092 (2016)
- [88] Scharf, I., Filin, I., Ovadia, O.: An experimental design and a statistical analysis separating interference from exploitation competition. *Population Ecology* **50**(3), 319–324 (2008)
- [89] Scheutz, M.: A Scalable, Robust, Ultra-Low Complexity Agent Swarm for Area Coverage and Interception Tasks. In: *Proceedings of the IEEE International Symposium on Intelligent Control, 2006*, pp. 1258–1263. IEEE (2006)
- [90] Schroeder, A., Subramanian, R., Kumar, M., Trease, B.: Efficient Spatial Coverage by a Robot Swarm based on an Ant Foraging Model and the Lévy Distribution. *Journal of Swarm Intelligence* pp. 1–31 (2017)
- [91] Schroeder, A., Trease, B., Arsie, A.: Balancing Robot Swarm Cost and Interference Effects by Varying Robot Quantity and Size. *Swarm Intelligence* (2018). DOI 10.1007/s11721-018-0161-1
- [92] Shaukat, M., Chitre, M.: Adaptive behaviors in multi-agent source localization using passive sensing. *Adaptive Behavior* **24**(6), 446–463 (2016)

- [93] Shkurti, F., Xu, A., Meghjani, M., Gamboa Higuera, J.C., Girdhar, Y., Giguere, P., Dey, B.B., Li, J., Kalmbach, A., Prahacs, C., Turgeon, K., Rekleitis, I., Dudek, G.: Multi-domain monitoring of marine environments using a heterogeneous robot team. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1747–1753 (2012)
- [94] Song, Z., Mohseni, K., Member, S.: Anisotropic Active Lagrangian Particle Swarm Control in a Meandering Jet. In: 2015 IEEE 54th Annual Conference on Decision and Control (CDC), pp. 240–245 (2015)
- [95] Soriano Marcolino, L., Tavares dos Passos, Y., Fonseca de Souza, Á.A., dos Santos Rodrigues, A., Chaimowicz, L.: Avoiding target congestion on the navigation of robotic swarms. *Autonomous Robots* **41**(6), 1297–1320 (2017)
- [96] Stelzer, R., Jafarmadar, K.: The robotic sailing boat asv roboat as a maritime research platform. In: Proceedings of the 22nd International HISWA Symposium on Yacht Design and Yacht Construction (2012)
- [97] Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., Burtner, A.: Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae* **54**, 160–173 (2016)
- [98] Sutanty, D.K., Kernbach, S., Nepomnyashchikh, V.A., Levi, P.: Multi-Robot Searching Algorithm using Levy Flight and Artificial Potential Field. In: IEEE International Workshop on Safety, Security, and Rescue Robotics. IEEE (2010)
- [99] Tarapore, D., Christensen, A.L., Timmis, J.: Generic, scalable and decentralized fault detection for robot swarms. *PLOS ONE* **12**(8), 1–30 (2017)
- [100] Tribe, M.A., Alpine, R.L.: Scale economies and the "0.6 rule". *Engineering Costs and Production Economics* **10**(1), 271–278 (1986)

- [101] Valada, A., Velagapudi, P., Kannan, B., Tomaszewski, C., Kantor, G., Scerri, P.: Development of a low cost multi-robot autonomous marine surface platform. *Springer Tracts in Advanced Robotics* **92**, 643–658 (2014)
- [102] Valdivia y Alvarado, P., Taher, T., Kurniawati, H., Weymouth, G., Khan, R.R., Leighton, J., Papadopoulos, G., Barbastathis, G., Patrikalakis, N.: A coastal distributed autonomous sensor network. *Oceans 2011* pp. 1–8 (2011)
- [103] Vanderploeg, H.A., Liebig, J.R., Carmichael, W.W., Agy, M.A., Johengen, T.H., Fahnenstiel, G.L., Nalepa, T.F.: Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. *Canadian Journal of Fisheries and Aquatic Sciences* **58**(6), 1208–1221 (2001)
- [104] World Health Organization: Guidelines for safe recreational water. Vol 1. Coastal and Fresh Waters. Tech. rep., World Health Organization (2003)
- [105] Wynne, T.T., Stumpf, R.P.: Spatial and Temporal Patterns in the Seasonal Distribution of Toxic Cyanobacteria in Western Lake Erie from 2002–2014. *Toxins* **7**, 1649–1663 (2015)
- [106] Wynne, T.T., Stumpf, R.P., Briggs, T.O.: Comparing MODIS and MERIS spectral shapes for cyanobacterial bloom detection. *International Journal of Remote Sensing* **34**(19), 6668–6678 (2013)
- [107] Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Dyle, J.: Characterizing a cyanobacterial bloom in Western Lake Erie using satellite imagery and meteorological data. *Limnology and Oceanography* **55**(5), 2025–2036 (2010)
- [108] Wynne, T.T., Stumpf, R.P., Tomlinson, M.C., Tester, P.A.: Relating spectral shape to cyanobacterial bloom in the Laurentian Great Lakes. *International Journal of Remote Sensing* **29**(12), 3665–3667 (2008)

- [109] Zhang, G., Zhang, P., Wang, B., Liu, H.: Ultrasonic frequency effects on the removal of *Microcystis aeruginosa*. *Ultrasonics Sonochemistry* **13**(5), 446–450 (2006)
- [110] Zhu, W., Li, M., Luo, Y., Dai, X., Guo, L., Xiao, M., Huang, J., Tan, X.: Vertical distribution of *Microcystis* colony size in Lake Taihu: Its role in algal blooms. *Journal of Great Lakes Research* **40**(4), 949–955 (2014)
- [111] Zoss, B.M., Mateo, D., Kuan, Y.K., Tokić, G., Chamanbaz, M., Goh, L., Vallegra, F., Bouffanais, R., Yue, D.K.P.: Distributed system of autonomous buoys for scalable deployment and monitoring of large waterbodies. *Autonomous Robots* (February) (2018)