DISSERTATION

FROM SAILORS TO SATELLITES:

INVESTIGATING THE MARITIME MYSTERY OF BIOLUMINESCENT MILKY SEAS

Submitted by

Justin Hudson

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2025

Doctoral Committee:

Advisor: Steven Miller

Steven Haddock
Eric Maloney
Kenneth Reardon
Peter Jan van Leeuwen

ABSTRACT

FROM SAILORS TO SATELLITES:

INVESTIGATING THE MARITIME MYSTERY OF BIOLUMINESCENT MILKY SEAS

Bioluminescence, the ability of living organisms to produce and emit light, has been a topic of human imagination and scientific study for millennia. Bioluminescence is observed in a myriad of forms in the ocean, among these bioluminescent displays milky seas stand out as perhaps one of the rarest, most poorly understood, and most awe-inspiring forms of bioluminescence. Milky seas are delineated from other more common forms of bioluminescence by their steady, non-flashing, eponymous white/green/gray glow which can cover 100,000 km^2 of the nocturnal ocean surface for possibly months at a time. Poetic descriptions of milky seas by eyewitnesses have compared this phenomenon to an episode of the 'Twilight Zone', the biblical apocalypse, and an ocean haunted by spirts.

Recent advances in spaceborne low-light imager technology, allowing milky seas to be identified remotely via satellite imagery, have greatly expanded our ability to study this phenomenon. Despite these technological advances and a modest compendium of published scientific literature on milky seas dating back to the 1700s, scientific understanding on milky seas has been historically limited by their remote, rare, and ephemeral nature. In addition, scientific research on milky seas has suffered the repeated loss of historical datasets.

This dissertation presents a collection of research that seeks to understand the global/macroscale properties (e.g. distribution and timing) of milky seas as well as more local and intrinsic properties that inform on their predictability. Combining centuries of eyewitness accounts

with recent satellite imagery, we reconstruct and build upon lost databases of milky sea observations. Leveraging this new and expanded database, we begin to address questions about milky sea occurrence, structure, and connection to the greater earth system.

The scientific analysis enabled by this database and the plethora of modern atmospheric and oceanic datasets allows new connections between milky seas, the South Asian and Indo-Australian monsoons, the Indian Ocean Dipole, and the El Niño Southern Oscillation to be drawn. These connections, which serve as sources of predictability, guided this research toward the first known prediction of a milky sea event, and offer the potential for proactive in-situ sampling of a milky sea event which is necessary to fully answer questions pertaining to their composition and formation mechanisms.

Furthermore, case study analysis of milky sea events near Java, Indonesia reveals insights into the physical processes that form, sustain, and eventually annihilate milky sea events. By way of this case study analysis, we test the natural flask hypothesis for milky seas, which postulates a physical mechanism for milky sea environments. Analysis of scatterometry data reveals the potential for coincident biological signals to be correlated with previously identified milky sea events, expanding the tools available to study and track the phenomenon from space across the lunar cycle and potentially overcome the limitations of current low-light visible observations.

DEDICATION

To my mom

TABLE OF CONTENTS

ABSTRACT ii
DEDICATION iv
Chapter 1 – Introduction 1
1.1 The Earth as a System 1
1.2 A History of Bioluminescence and Milky Seas 3
1.3 Milky Seas: A Remote Sensing Perspective 9
1.4 The Current State of Knowledge on Milky Seas 11
1.5 Current Gaps in Our Knowledge of Milky Seas 13
1.6 Science Questions 15
1.7 Dissertation Outline 17
Chapter 2 – A Curated Database of Milky Seas Since 1600 18
2.1 Introduction 18
2.2 Data Sources and Methodology 20
2.2.1 Atmospheric and Oceanic Data Sources 20
2.2.2 Eyewitness Accounts/Milky Sea Reports Sources 23
2.2.2.1 Archives of the Marine Observer Journal 23
2.2.2.2 Archives of Dr. Peter Herring 24
2.2.2.3 Archives of Dr. Tim Wyatt 26
2.2.2.4 Contemporary Accounts and Satellite Based Data 26
2.2.3 Milky Sea Database Construction Methodology 26
2.2.4 Potential Sources of Bias 27
2.2.4.1 Shipping Route Bias Between 1924 and 2003 28
2.2.4.2 Language Barriers 29
2.2.4.3 Low-Light Imager Limitations 29
2.2.4.4 Sampling Non-Uniformity in Time 32
2.2.5 Bootstrapping Methodology 32
2.3 Results 34
2.3.1 Spatial Distribution of Milky Seas 36
2.3.2 Intra-Year Distribution of Milky Seas 40
2.3.3 Milky Sea Environmental Conditions 42
2.3.4 Size and Duration of the ‘Typical’ Milky Sea Event 44
2.4 Identifying Potential Predictors for Milky Sea Events 47
2.4.1 Bootstrapping Results: NW Indian Ocean During Boreal Summer 48
2.4.2 Bootstrapping Results: NW Indian Ocean During Boreal Winter 52
2.4.3 Bootstrapping Results: Banda Sea During Austral Winter 55
2.4.4 Bootstrapping Results: South Java/Christmas Island During Austral Winter 57
2.5 Discussion and Conclusions 59
Chapter 3 – Exploring the Predictability of Milky Seas Near Java, Indonesia 61
3.1 Introduction 61
3.2 Data Sources 64

3.3 Verification and Construction of the Predictive Model 67
3.4 Model Insights 76
3.5 Forecasting Milky Seas Near Java 78
3.6 Discussion and Conclusions 80
Chapter 4 – The Natural Flask Hypothesis of Milky Sea Formation: A Case Study Analysis 82
4.1 Introduction 82
4.2 Data Sources 84
4.3 MS2017 Event Case Study Analysis 85
4.3.1 Depth Estimation of the MS2017 Event 85
4.3.2 Environmental Analysis of the MS2017 Event 88
4.3.3 Transect Analysis of the MS2017 Event 91
4.4 MS2019 Event Case Study Analysis 96
4.4.1 Environmental Analysis of Phase 1 of the MS2019 Event 96
4.4.2 Transect Analysis of Phase 1 of the MS2019 Event 98
4.4.3 Environmental Analysis of Phase 2 of the MS2019 Event 99
4.4.4 Transect Analysis of Phase 2 of the MS2019 Event 101
4.5 Conclusions 103
Chapter 5 – Scatterometers and Associated Biological Signals of Milky Seas 104
5.1 Introduction 104
5.2 Data Sources 106
5.3 Analysis 107
5.4 Conclusions 115
Chapter 6 – Sampling Guidance for an Active Milky Sea Event 117
6.1 Introduction 117
6.2 Anticipation of an Event 117
6.3 Sampling Logistics 118
6.4 Specific Measurands 119
Chapter 7 – Summary and Concluding Thoughts 122
7.1 Summary 122
7.2 Concluding Thoughts and Future Research 125
References 129

Chapter 1
Introduction

"But nature is always more subtle, more intricate,
more elegant than what we are able to imagine."
- Carl Sagan

1.1 The Earth as a System

Earth System science is, tautologically, the study of the Earth as a system. It is a holistic approach to understanding how the various 'spheres' that make up the Earth such as the atmosphere, biosphere, cryosphere, hydrosphere, and lithosphere interact with and influence one another. As an example, fluxes of moisture and energy between the atmosphere, ocean, and land surface play critical roles in driving weather both on large and small scales. This interconnected viewpoint is often central to truly understanding various facets of the Earth System that any single sub-discipline cannot describe adequately on its own.

The biosphere is tightly coupled to many other aspects of the earth system and vice versa. The rise of photosynthetic life and the proliferation of oxygen into Earth's previously anoxic primordial atmosphere, known as the great oxygenation event, has permanently altered the chemical and radiative properties of the atmosphere, the types of minerals present on the Earth's surface, and habitability of the Earth itself (Olejarz et al. 2021). More recent examples of the biosphere impacting the earth system include the mass emission of carbon from human industrial activities or changes in albedo due to land use change.

The carbon cycle, how carbon moves through the various aspects of the earth system, couples the atmosphere, ocean, land, and biosphere together at various timescales. On multi-million year timescales geological processes such as erosion and subduction dominate the global carbon cycle. On non-geological timescales the biosphere plays an important role in the

partitioning of carbon between various facets of the earth system via the growth and decay of organisms (Leblanc et al. 2018; Le Moigne 2019; Jiao et al. 2024).

In the ocean the partitioning of carbon by the marine biosphere is known as the biological carbon pump. The biological carbon pump moves carbon between the atmosphere, surface ocean, and deep ocean. The biological carbon pump is composed of many different processes and is relayed in a simplified form below and in Figure 1.1.

Autotrophic organisms which consume CO_2 for photosynthesis and heterotrophic organisms which produce CO_2 directly control the partial pressure of CO_2 in the upper ocean. This partial pressure mediates the CO_2 flux between the atmosphere and ocean at the air-sea interface. The death of organisms (Leblanc et al. 2018) and waste produced by organisms (Le Moigne 2019) forms the basis for marine snow, particulate matter rich in carbon, which sinks to greater depths in the ocean where the carbon therein can be sequestered for timespans ranging from thousands of years to geologic timescales. Particulate organic carbon that doesn't sink before being broken down by microbial organisms, and waste products from autotrophic and heterotrophic organisms forms the basis for Dissolved Organic Carbon (DOC) in the upper ocean (Le Moigne 2019).

The microbial carbon pump (Jiao et al. 2024), is the portion of the biological carbon pump focused on how microbial organisms (e.g. bacteria) process DOC. DOC is composed of organic carbon containing molecules dissolved into the ocean and can be divided into two categories: Labile Dissolved Organic Carbon (LDOC) and Refractory Dissolved Organic Carbon (RDOC; Jiao et al. 2024).

LDOC has a high bioavailability—it is readily processed by microbial organisms—and typically is quickly processed into RDOC and excess CO_2. RDOC has a low bioavailability—it is not readily processed by microbial organisms—and is composed of molecules which either quickly

decompose into CO_2 which is re-released back into the atmosphere, or stay dissolved in the ocean column. RDOC which remains dissolved in the ocean column can be sequestered for centuries or more until the various molecules decompose or an organism evolves which can make use of the RDOC which reclassifies it as LDOC (Jiao et al. 2024).

Marine bacteria play a key role in the Microbial Carbon Pump (MCP), and through it the fate of anthropogenic carbon emission. Understanding how the behavior of marine bacteria is organized and interacts with the MCP on large scales is critical to understanding the future of the earth system under anthropogenic climate change.

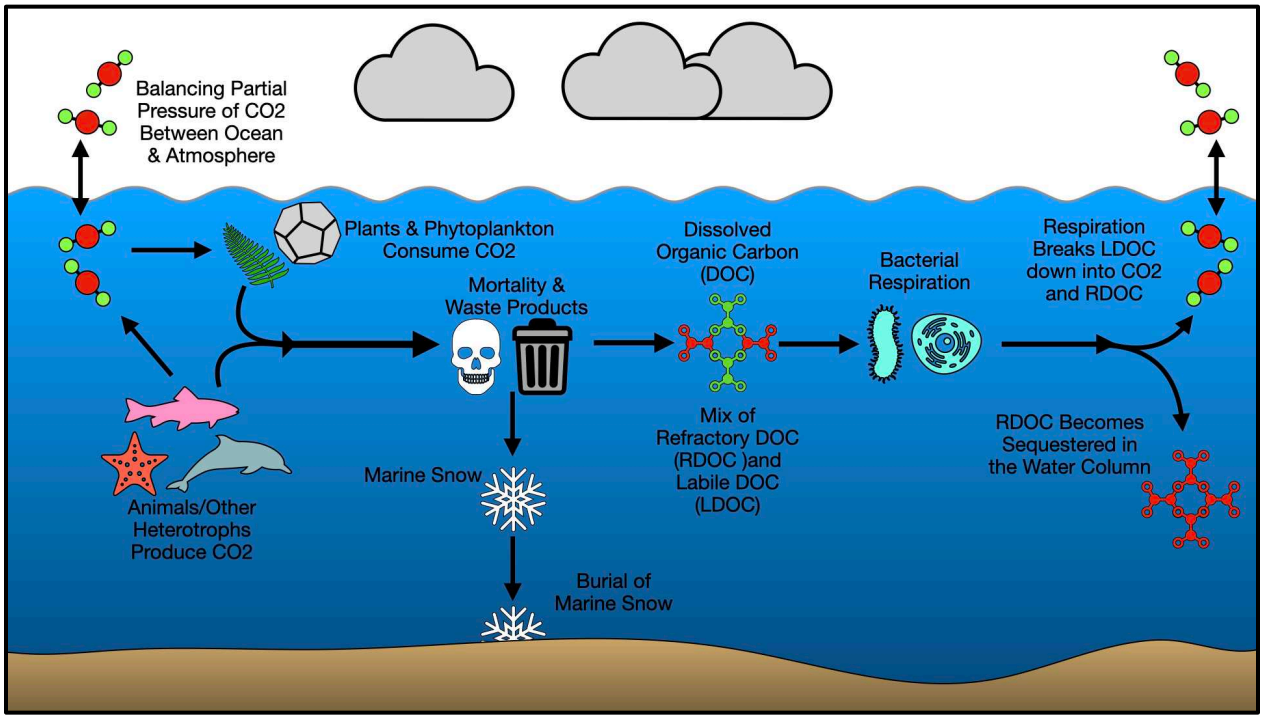


Figure 1.1: A cartoon schematic of the MCP showing where bacteria fit into the greater Biological Carbon Pump as well as the Solubility Carbon Pump which controls the balance of CO2 between the ocean and the atmosphere.

1.2 A History of Bioluminescence and Milky Seas

Bioluminescence, the ability for living beings to product and emit light, has been observed across a wide variety of organisms from bacteria, plants, protists, and animals (Figure 1.2 Panel

A; Haddock et al. 2010). Bioluminescence has been studied for millennia (Harvey 1920) and occurs due to a chemical reaction with a light producing molecule known as a luciferin (Haddock et al. 2010).

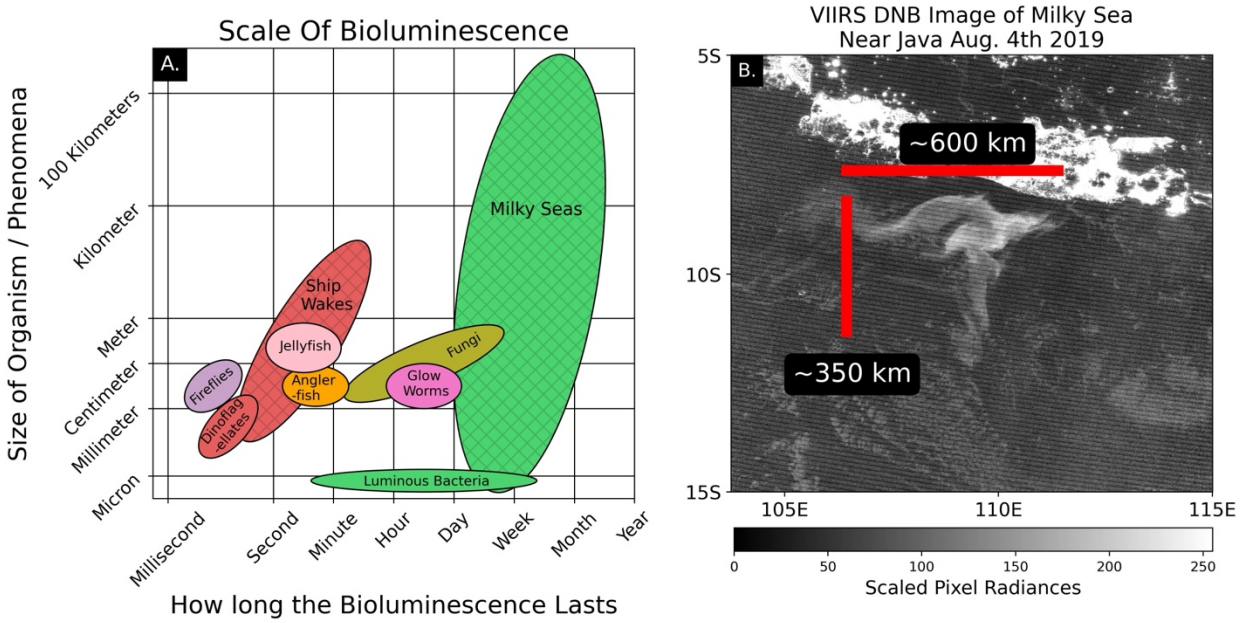


Figure 1.2: A comparison of the spatial and temporal scales of various forms of bioluminescence (Panel A.). A VIIRS DNB image of a milky sea event near Java, Indonesia in 2019 showing the spatial scales a milky sea event can reach (Panel B.). The hatched ellipses in Panel A are large scale displays of bioluminescence due to the organisms in the similarly colored ellipse.

Bioluminescent displays can be due to a lone organism or a group of organisms responding to a common stimulus (whether external or internal). Large numbers of dinoflagellates flashing due to agitation or turbulence in the water, as in crashing waves or a ship's wake, is a common example of externally stimulated marine bioluminescence on large scales (Figure 1.2 Panel A; Haddock et al. 2010).

Milky seas, a far-less frequently observed form of marine bioluminescence, are distinguished from other forms of marine bioluminescence by their immense size (observed to exceed 100,000 km^2) and their ability to last for potentially months at a time (Figure 1.2 Panel A; Miller et al. 2021). A typical milky sea involves large areas of the nocturnal ocean emitting a steady (non-flashing) white/green glow, oft likened to plastic glow in the dark stars by modern observers,

stretching from horizon to horizon. Photos taken from the yacht *Ganesha* during a chance transit of a milky sea in 2019 provide the only visual reference for what it is like to sail through a milky sea (Figure 1.3).

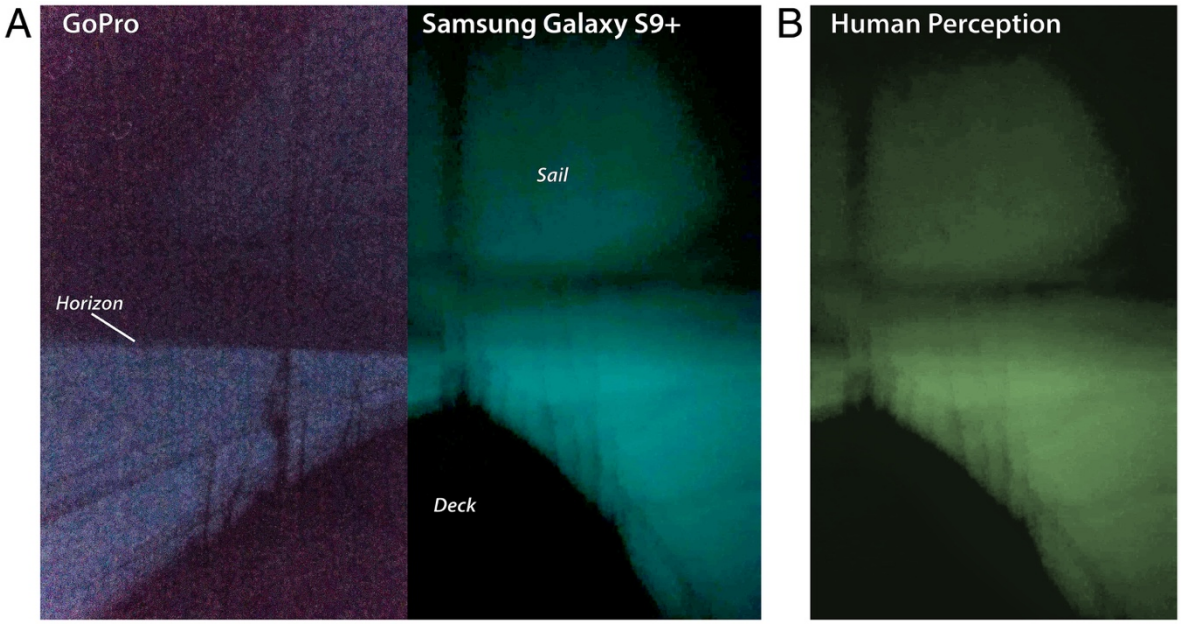


Figure 1.3: Photographs of a milky sea as captured by the crew of the yacht Ganesha during a transit of a milky sea event in August 2019 as seen through two different cameras a GoPro and a Samsung Galaxy S9+ (Panel A). The photo from the Samsung Galaxy S9+ in Panel A but with the color modified to better match what the eyewitnesses saw (Panel B). This figure was taken from Miller 2022 in the Proceedings of the National Academy of Sciences (PNAS).

While milky seas were likely already known to the peoples who inhabit the areas where they most commonly occur (the Northwest Indian Ocean and Maritime Continent), they begin to enter into the English language historical record in the early 1600s (Strachan & Penrose 1971). This early account of a milky sea, provided by Captain William Keeling and reproduced below, matches the descriptions of many later accounts:

> *"We fell in the night into a white water like an extreme shoald[1] but had no ground at 60 fathom line. The lattd about 10:30[2]"*

[1] Shoald is an archaic spelling of the word shoal which describes an area of shallow water

[2] Lattd is an abbreviation for latitude. This second sentence describes their position as 10° 30' N. No longitude information is provided as this predates the invention of accurate chronometers and the ability to easily determine longitude while at sea

Below are a selected sample of eyewitness descriptions of milky sea events:

"The whole appearance of the ocean was like a plain covered with snow. There was scarce a cloud in the heavens, yet the sky appeared black as if a storm was raging. The scene was one of awful grandeur, the sea having turned to phosphorous and the heavens being hung in blackness, and the stars going out seemed to indicate that all nature was preparing for that last grand conflagrations which we are taught to believe is to annihilate this material world."

- Capt. Kingman, Shooting Star, July 27th, 1854, near Java, Indonesia

"At about eight P.M., there being no moon, but the sky being clear, and the stars shining brightly, we suddenly passed from the deep blue water in which we had been sailing into a patch of water so white that it startled me; so much did it appear like a shoal. […] The patch was extensive. We were several hours running through it. Around the horizon there was a subdued glare, or flush, as though there were a distant illumination going on, whilst overhead there was a lurid, dark sky in which the stars paled. The whole face of nature seemed changed, and with but little stretch of the imagination, the Alabama might have been conceived to be a phantom ship, lighted up by the sickly and unearthly glare of a phantom sea, and gliding on under the pale stars one knew not whither."

- Capt. Raphael Semmes, CSS Alabama, January 30th, 1964, near Socotra

"December 29th, 1931, at 8:15 p.m. A.T.S., the ship entered, or was suddenly surrounded by an area of white water. It seemed as though shoal water stretched from horizon to horizon, being of a pale milky appearance, making the sky, already dark with Strato-Cumulus and Fracto-Nimbus clouds, look inky black. The most noticeable feature of the phenomenon was that, although there was a fresh northerly breeze blowing, the previously rough sea and moderate short swell were appreciably diminished. In point of fact, where there had been breaking seas all over, there were during the hour the white water was visible, very few crests on the sea, as if this luminescence was of an oleaginous quality. Floating objects appeared jet black and two dimensional. A faint haze of light seemed to be cast upward from some depth, causing a strain on the eye. At 1745 G.M.T. the phenomenon faded, and where the sky had seemed black and sea almost white the sea now became black and the sky lighter."

- Third Officer G. Drake, SS Clan Macphee, December 29th, 1931, near the Arabian Peninsula

"8th December, 1947, 0700 G.M.T. Vessel passed through a band of vivid phosphorescence which gave the water a milky-white appearance. It was approx. 100 ft. wide and had sharply defined limits. It ran in a WNW to ESE

direction and extended as far as was visible. Moderate SSW breeze and moderate sea."

- Third Officer H. J. Thompson, MV Port Macquarie, December 8th, 1947, eastern equatorial Pacific Ocean

"The initial observation was of 'speckles' running along the ship's side and in the bow waves, these appeared to be caused by the direct disturbance of the water by the vessel. This was observed for several minutes before the vessel entered an area of 'patchy', seemingly 'milky sea', where the water emitted a diffuse glow. It was noticed that the wind was blowing the luminescence in 'slicks' which could be seen to be quite distant from the ship. Although unable to gauge how extensive the phenomena was, it appeared to stretch half-way to the horizon, and certainly did not appear to be triggered by the ship's movement, as the area of the phenomena could be seen approaching and disappearing stern. The ship's sea track caused quite a distinctive glowing wake and 'speckles and patches' which were first observed continued through the encounter and for several minutes after. It was interested to note that the local speckled luminescence was quite sharp and distinct to the eye as opposed to the more remote 'milky/patchy' sea phenomena which produced an even diffuse glow in the surface water."

- Third Officer K. P. Mowat, MV Willowbank, December 10th, 1985, Tasman Sea

Milky seas have also been described in fiction, the second part of Jules Verne's 20,000 Leagues Under the Sea opens with their submersible, the Nautilus, entering the Indian Ocean and the Bay of Bengal, there the main characters encounter a milky sea:

"About seven o'clock in the evening, the Nautilus, half-immersed, was sailing in a sea of milk. At first sight the ocean seemed lactified. Was it the effect of lunar rays? No; for the moon, scarcely two days old, was still lying half hidden under the horizon in the rays of the sun. The whole sky, though lit by sidereal rays, seemed black by contrast with the whiteness of the waters."

- *Jules Verne, 20,000 Leagues Under the Sea, Chapter 1 Part 2*

The fictional character Pierre Aronnax, the narrator of *20,000 Leagues Under the Sea,* explains that milky seas were due to a vast number of "infusoria", a small worm like organism, which emit the light causing the milky sea. This hypothesis is in line with then contemporary ideas

for milky sea composition (Buist 1855) and those that predated Jules Verne's novel by a century or more (Newland 1772; Macartney 1810).

One of the earliest scientific publications in English on milky seas was read before the Royal Society of London in 1772 by a Captain Newland (first name unknown). Based on his observations of a milky sea that he sailed through, Captain Newland concluded that milky seas were due to vast quantities of an 'animalcule', an archaic word for microscopic organisms (Newland 1772). Scientific work in the 18th and 19th centuries agreed with Newland's hypothesis and the idea that the myriad forms of 'phosphorescence' (the word at the time to describe bioluminescence, chemiluminescence, and other forms of light production) observed were caused by living organisms (e.g. Franklin 1753; Macartney 1810). While there were competing explanations such as the accumulation of electric charge in the ocean (Franklin 1753), or the ocean absorbing sunlight and heat during the day and releasing it at night (Tachard 1688) many types of oceanic phosphorescence were eventually understood to be biological in nature.

In the 19th and early 20th century the theory behind what organism is causing milky seas was refined from unknown microscopic organisms to specifically bioluminescent bacteria (Fischer 1887; Katz 1888; Dahlgren 1915). This theory is, to this day, the leading explanation for the type of organism behind a milky sea (Haddock et al. 2010; Miller et al. 2021). This theory is bolstered by a chance encounter of a research vessel in the Arabian Sea with a milky sea in 1985. Water samples collected by this vessel within the milky sea it encountered found the presence of a bioluminescent bacteria known as *vibrio Harveyi* colonizing *Phaeocystis* algae (Lapota et al. 1988).

Per historical records, milky seas occur primarily in the northwest Indian Ocean and Maritime Continent regions (Smith 1931; Turner 1965; Herring & Watson 1993; Miller et al.

2021). Population background estimates of *v. Harveyi* in these regions is on the order of ~10 cells/ml (Ramaiah & Chandramohan 1987). While there are uncertainties on this population estimate (Takemura et al. 2014), they are not large enough to account for the ~10^8 cells/ml needed for bacterial bioluminescence to be visible to the human eye (Nealson & Hastings 1979). Exceptional biogeochemical conditions and mechanisms are needed to explain how the extraordinary conditions for milky sea events arise.

1.3 Milky Seas: A Remote Sensing Perspective

Research on milky seas has, historically, been limited to the examination of eyewitness accounts (Buist 1855; Smith 1926; Smith 1931; Turner 1965; Herring & Watson 1993). These analyses of when and where milky seas have been observed as well as common characteristics of a milky sea event have provided a wealth of information about this phenomenon. These collections of milky sea eyewitness accounts, including the most recent one of Herring & Watson 1993 (hereafter HW93), have been subsequently lost leaving only the publications describing these collections behind. As an example, the magnetic tape used to store the database of HW93 has been lost and any potential physical paper copy of the HW93 database that may still exist lies in an unknown location in the UK Navy Archives (Peter Herring, personal communication). The fate of the other historical databases is unknown aside from the fact that they are presently unavailable. The loss of these datasets has negatively impacted the ability to study milky seas.

In the 21st century, advances in remote sensing from satellites have greatly expanded the ability to study many earth system phenomena, and milky seas are no exception. Two low-light imagers, the Defense Meteorological Satellite Program's (DMSP) Operational Line Scanner (OLS), and the Visible Infrared Imaging Radiometer Suite's (VIIRS) Day/Night Band (DNB),

have allowed for remote verification and detection of milky sea events (Miller et al. 2005; Miller et al. 2021).

Miller et al. 2005 showed that given a specific time and location of a known milky sea event, based on an eyewitness account, it was possible for a satellite based low-light imager (i.e. DMSP OLS) to observe a milky sea event. The requirement of a priori information was due to the very noisy quality of OLS data at the extreme limits of low-light detection (where the signals from milky seas reside), making the task of identifying milky sea features nearly impossible without this advance knowledge of their location. The VIIRS DNB, with its higher sensitivity, was shown to be able to find milky sea events without a priori knowledge of their location (Miller et al. 2021)—a major step forward enabled by technological advancements, opening a promising new era for milky seas research.

The VIIRS DNB satellite observations of milky seas have greatly advanced the ability to study milky seas from random-chance eyewitness accounts towards direct observation by a battery of satellite-based sensors. Satellite observations of milky seas however do have one key limitation. The light emitted from milky seas is incredibly faint, approximately 10^3-10^4 times fainter than the light provided by the full moon, this means reflected moonlight off the ocean surface, clouds, aerosols, and the molecular atmosphere readily overpowers any milky sea signal in satellite imagery (Miller et al. 2005; Miller et al. 2021). This limitation relegates ‘milky sea hunting’ from these sensors to the windows of the lunar cycle when the moon is far below the horizon at the time of the satellite overpass (for the ~0130 local time descending node crossing of the DNB), this corresponds to an observational window beginning roughly two nights after the moon’s Third/Last Quarter phase until two nights after the First Quarter lunar phase, approximately two weeks.

The detection limitations of satellite sensors mirror the limitations of human vision to observe milky seas. Some milky sea eyewitness accounts specify that the milky sea seemingly disappears upon the moon rising above the horizon, consistent with satellite limitations.

In both cases, the satellite sensor and the human eye, the light detection is relegated to a broadband response (as opposed to color). The DNB's spectral response from 500-900 nanometers (nm) lacks color differentiation, and the rod cells of the human eye's retina under low-light conditions function in a similar, panchromatic way known as scotopic vision (Shin et al. 2004). It would in principle be possible to detect milky seas via color perception given a sensor of sufficient sensitivity across narrow-band red, green, and blue spectral channels, or in the case of human vision, for milky sea events of sufficient luminosity to trigger a response from color-partitioned cone cells of the retina. Some eyewitness accounts do describe a coloration to the glow, from greenish-yellow to a radiant cyan/turquoise.

1.4 The Current State of Knowledge on Milky Seas

Despite their limitations the current generation of low-light imagers have enabled the investigation of milky seas and greatly added to our knowledge of this phenomenon. This usefulness highlights how additional remote sensing technologies and techniques that can observe milky seas or related features would greatly expand the ability to study this phenomenon.

Satellite observations of milky sea events have revealed the true size and scale of milky seas. Based on the satellite events observed in Miller et al. 2021 (hereafter M21) the typical milky sea is on the order of 10,000 km^2 and lasts about two weeks. For the biogeochemical conditions that lead to and sustain milky seas to persist over such vast spatiotemporal scales an adequate theory is required.

The satellite observations analyzed in M21 found that milky seas are aligned with relatively cool upwelled waters with low surface current velocities bounded on one or more sides by intense ocean surface eddies/meanders near to, but not overlapping, local maxima in chlorophyll-a. These conditions were maintained over the lifespan of the milky sea events studied. Based on these observations M21 proposed the natural flask theory of milky sea formation.

In a laboratory setting a flask can be used to cultivate an isolated, ideal environment for bacteria. Under these conditions bacterial populations can thrive and reach the population density required for quorum sensing driven bioluminescence (Figure 1.4). The natural flask theory proposes that density driven or dynamical barriers in the ocean lead to isolated or semi-isolated parcels of water with ideal conditions for the bacteria behind milky seas.

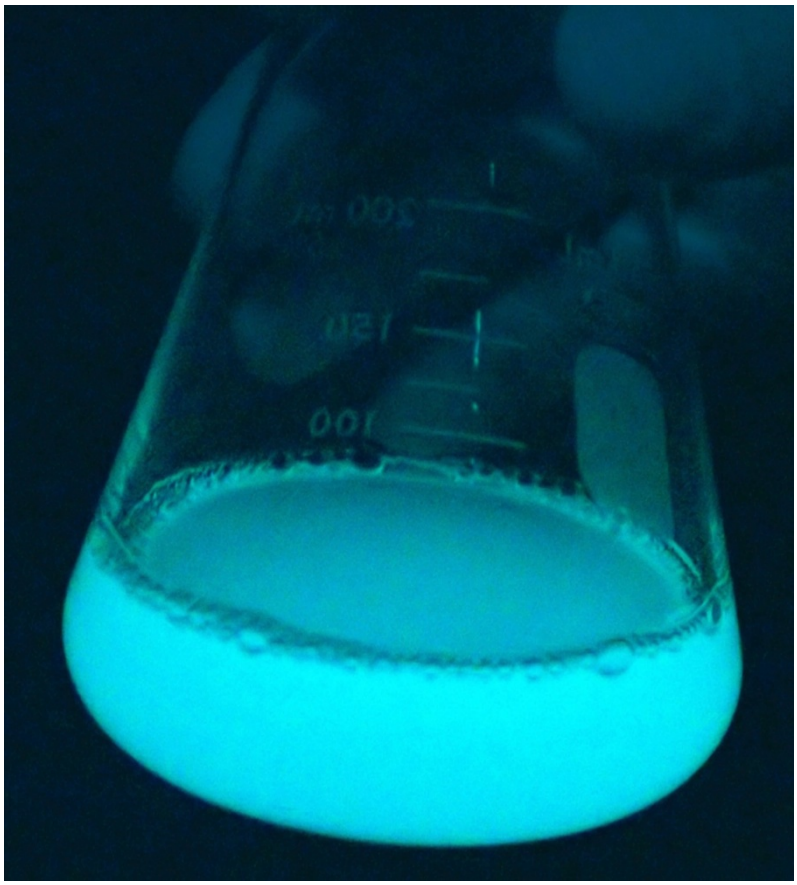


Figure 1.4: Bioluminescent bacteria cultured within a laboratory flask, the artificial analogue to the natural flask hypothesis. Photograph taken by Steve Haddock of the Monterey Bay Aquarium Research Institute (MBARI).

Within these natural flasks bacterial populations are able to reach the critical densities required for quorum sensing driven bioluminescence. The bioluminescence sustains itself as long as the flask is maintained and, upon either the conditions within the natural flask being disrupted or the bacteria's food source within the flask begins to diminish, the milky sea event dissipates. Understanding how natural flask conditions arise, their associated timescales (daily, weekly, subseasonal, or seasonal), and if these conditions can be predicted are key questions towards being able to sample a milky sea in-situ and answering the battery of biological and ecological questions about this phenomenon.

M21 also found that a milky sea near Java during the austral winter of 2019 was coincident with a positive Indian Ocean Dipole (IOD) event. It had long been speculated based on timing and location that milky seas were related to large-scale coupled atmosphere ocean phenomena such as the Indian Monsoon (Buist 1855; Smith 1931; Turner 1965; Herring & Watson 1993).

The Indian Monsoon and the IOD are connected to the El Niño Southern Oscillation (ENSO) with all three phenomena impacting the teleconnections and direct effects of each other (Saji et al. 1999, Ashok et al. 2001, Cai et al. 2009, Ummenhofer et al. 2011, Liu et al. 2015, Wirasatriya et al. 2020). Understanding how the interplay of these coupled earth system phenomena impact the frequency of and set the large-scale oceanic state for milky sea events is a critical first step towards being able to move from studying this phenomena remotely to in-situ.

1.5 Current Gaps in Our Knowledge of Milky Seas

Despite centuries of scientific inquiry into milky seas there is very little centralized data on this phenomenon. The databases generated by past authors have been lost to time. This greatly limits the ability to understand how milky seas relate to the larger coupled earth system.

Satellite observations, while they have provided enhanced insight into milky seas, are limited in the information they can provide. The chemical properties of the waters that make up and surround milky seas, the 3-dimensional structure of milky seas, the validity of the natural flask hypothesis, and the organism(s) responsible for milky seas are questions that can only be definitively answered via in-situ sampling of an active milky sea and the surrounding ocean.

In-situ sampling of a milky sea however is a difficult task. Despite low-light imagers giving us the ability to observe milky seas in near real-time the time needed to deploy a team with the proper equipment and vessel needed to a remote location before the milky sea event dissipates or the lunar illumination renders the event invisible is a tall task. Aside from anomalously long-lived events such as the 2019 July-September milky sea event near Java, Indonesia many milky sea events appear to be short-lived lasting approximately 1 to 2 weeks.

To properly address these key gaps in our knowledge regarding milky seas the ability to anticipate when and where a milky sea event will occur is critical. Historical databases such as that of HW93 have noted that in the northwest Indian Ocean, specifically near the coasts of Somalia, the island of Socotra, and the southern coast of the Arabian Peninsula, where many milky sea events occur, there is a bimodal intra-annual distribution of milky seas. This distribution peaks in January and August, coincidentally these are the peak time periods of the boreal winter and summer phases of the Indian Monsoon. Milky sea events near Java, Indonesia appear to have strong interannual variability and are anecdotally connected to both the Indo-Australian Monsoon and the IOD (Miller et al. 2021).

It is unknown if these anecdotal connections represent true interconnectedness between milky seas and the greater earth system or are the result of imperfect knowledge and records on milky seas occurrence. Understanding if these are true connections between milky seas and

coupled atmosphere-ocean behavior and whether these connections can be utilized to predict when and where a milky sea will occur is a critical knowledge gap regarding milky seas that limits our ability to study this phenomenon both remotely and in-situ.

Another gap in the knowledge surrounding milky seas is the typical depth of a milky sea in the water column. Milky seas are identified via a surface expression of bioluminescence over large areas. If a milky sea can occur at sufficient depth that any and all light emission is attenuated before reaching the ocean surface, then even the best attempts to catalog this phenomenon both from historical records and satellite imagery would fail to capture this fraction of events.

The thickness of milky sea events is also unknown. Whether milky seas are confined to thin layers or possess thickness on the order of centimeters or meters greatly changes the magnitude of a milky sea event in terms of carbon, nutrients, and bacterial population. For a milky sea event on the order of 10,000 km^2, the typical size of a milky sea from the satellite observations of M21, every cm of additional thickness to the event corresponds to 10^{22} additional bacteria (assuming 10^8 cells/ml as per Nealson & Hastings 1979). For the largest milky sea event which spanned 130,000 km^2 at its peak this would equate to approximately one mole of bacteria. If a typical bacterial cell has mass on the order of 100 femtograms or 10^{-16} kg (Lewis et al. 2014) this is an additional 10^7 kg of bacteria involved in the largest milky sea event for every cm of thickness.

The structure and thickness of a milky sea event greatly impact how this phenomenon could be studied in-situ but without in-situ sampling we cannot know the true structure and thickness of a milky sea event. When designing future strategies to sample a milky sea event this unknown physical structure as well as evolution of the event in time must be considered.

1.6 Science Questions

In this dissertation there are several key science questions surrounding milky seas that we seek to address. The first question is the spatiotemporal distribution of milky seas. The surviving records of previous databases roughly outline when and where milky seas have happened, but more granular information is needed to truly understand this phenomenon and potential connections to the broader earth system. If milky seas are as rare as the surviving historical record suggests, then only by combining decades or centuries of eyewitness accounts with the more recent satellite record can we begin to understand the patterns of inter-annual and intra-annual variation in milky seas.

The second science question we seek to address in this dissertation is whether or not milky seas can be predicted. Large-scale earth system phenomena are often coupled to or impacted by multiple other elements of the earth system. Assuming milky seas, as part of the biosphere, are also coupled to or impacted by multiple other elements of the earth system, then any patterns of inter-annual and intra-annual variability that impact when and where milky seas form can likely be leveraged to this end. With a long enough record, even one that is spotty and contains data quality issues, it should be possible to draw connections between milky seas and larger scale sources of variability and predict when they will happen next.

The validity of the natural flask hypothesis for explaining the behavior of milky sea events is the third science question we seek to address in this dissertation. As the leading theory for milky sea formation, understanding if this theory holds under a variety of conditions and can explain observed behavior of known milky sea events is critical for both determining its validity and better understanding milky seas. The satellite record of milky sea events allows us to know exactly when

and where the surface expression of a milky sea event is. While there is no satellite imagery of a milky sea forming, by combining the information the satellite record provides with atmospheric and oceanic models and reanalysis products we can understand how the large-scale state of the coupled atmosphere-ocean system plays into sustaining and eventually destroying milky sea events.

The final science question we seek to address in this dissertation is if there are additional remote sensing angles through which milky seas can be observed. There are a wide variety of sensors beyond low-light imagers that observe regions with active milky sea events. Milky seas and many of the commonly observed co-occurring oceanic features are of sufficient spatial scale to encompass multiple pixels for many modern remote sensing instruments. Exploring if the wealth of remote sensing resources currently available provides any additional perspectives on milky seas is critical to both overcoming current low-light imager limitations and better understanding the relationship between milky seas and other aspects of the coupled earth system.

1.7 Dissertation Outline

In this dissertation we create a centralized database of milky sea events (Chapter 2) and leverage this database to examine how predictable milky seas are (Chapter 3), the validity of the natural flask hypothesis (Chapter 4), and determine if there are any additional ways to detect and study milky seas remotely (Chapter 5). In Chapter 6 we discuss how the information presented in the previous four chapters informs future strategies and approaches to sampling milky seas in-situ, and in Chapter 7 we conclude this dissertation with a discussion of where future research into milky seas should focus.

Chapter 2
A Curated Database of Milky Seas Since 1600[3]

2.1 Introduction

As milky seas are a large-scale biological phenomenon (Figure 1.2 Panel A), understanding when and where they form is critical to understanding their role within and connections to the rest of the coupled earth system. There are many eyewitness-based catalogs of the global and regional distribution of milky seas (Buist 1855; Smith 1926; Smith 1931; Turner 1965; Herring & Watson 1993). These catalogs detail many of the known attributes of milky seas such as their prevalence in the Indian Ocean and their potential ties to the monsoons of South Asia and the Maritime Continent/Australia.

These catalogs also reveal what are the common features of a 'classical' milky sea:

- The eponymous, non-flashing, steady white/gray/green-blue glow (possibly impacted by depth of the phenomenon and human scotopic vision)
- A sometimes well-defined edge delineating glowing and non-glowing waters (possibly indicative of oceanic frontogenesis)
- A gradual fade-in or fade-out of brightness over multiple miles (possibly indicative of non-frontal oceanic conditions)
- Glowing waters extending from the observer's location to the horizon in one or more directions (large spatial scale)
- Little to no association with disturbed waters either by a ship's wake or other water-agitating sources such as dolphins or large schools of fish.
- The sea returning to a 'normal' appearance (ocean water darker than the sky and the horizon) with the rising of the sun or moon (eliminating reflective suspensions as the source of perceived luminosity)
- Visible over multiple nights (indicating both a large spatial extent and the ability to sustain internal conditions over extended periods of time)
- A perceived or actual calming of the ocean surface within a milky sea (owing possibly to surfactants/mucous from certain algal species potentially associated with milky seas; Lapota et al. 1988)

[3] The contents of this chapter are based on the contents of the publication by J. Hudson & S. D. Miller, "From Sailors to Satellites: A Curated Database of Bioluminescent Milky Seas Since 1600" published in the journal Earth & Space Science in April 2025. The version relayed here is modified from the publication to better match the overall dissertation and contains additional content that was not included in the publication.

Recent satellite-based studies leveraging low-light imagers such as the VIIRS DNB (Miller et al. 2005; Miller et al. 2021; Figure 1.2. Panel B) bolster historical attempts to understand milky seas. Combining the wealth of historical observations with the enhanced perspective of satellite observations would greatly expand the ability to understand the global footprint and physical underpinnings of milky seas.

Despite several databases of milky sea eyewitness accounts having been compiled by experts since the early 1900s (Smith 1926; Smith 1931; Turner 1965; Herring & Watson 1993) there has been no database of milky sea eyewitness accounts available in recent decades. The figures published alongside these databases, example eyewitness accounts, and recent satellite observations make up the available scientific information on milky seas. In this chapter we present a comprehensive and curated database of milky seas from both historical eyewitness accounts and modern satellite imagery. This effort serves as an important first step to producing a climatology of milky seas and a basis for any predictive algorithms or models to facilitate in-situ examination of an active milky sea event.

In this chapter we present details of database construction, data sources, and methodology in Section 2.2. In Section 2.3 we discuss how this database compares to previous efforts to catalog milky seas. Using this database, we examine the statistical relationship between milky seas and large-scale coupled earth system phenomena such as ENSO and the IOD (Section 2.4) to find potential sources of predictability. Section 2.5 considers how this work advances the objective study of milky seas.

2.2 Data Sources and Methodology

2.2.1 Atmospheric and Oceanic Data Sources

The VIIRS DNB (Miller et al. 2013) is an imaging radiometer designed to have high sensitivity to visibly/near-infrared light emission in the 500-900 nm spectral range, with sensitivity of 10^{-5} W m^{-2} sr^{-1} at a signal to noise ratio (SNR) of ~10. As a point of comparison, daytime reflected sunlight is ~10 million times brighter. The high sensitivity of the DNB is coupled with high spatial resolution, 0.742 km across a ~3000 km wide imagery swath allows the DNB to excel at providing direct measurement of milky sea bioluminescence.

The DNB is currently carried abord three satellites: Suomi-NPP, NOAA-20, and NOAA-21. These satellites fly in sun-synchronous orbits providing equatorial overpasses at ~1:30 AM and ~1:30 PM local time (descending and ascending nodes respectively). The combination of these three sensors allows milky seas to be distinguished from the environmental cloud field via temporal constancy and near-zero parallax displacement (in contrast to clouds which drift and produce parallax shift when viewed at different angles).

These tests of parallax and temporal constancy are coupled with other discriminators involving SST and cloud-cover analyses from infrared sensors present within the VIIRS sensor-suite. An in-depth discussion of how to distinguish milky seas from the environment is provided in M21 and we relay an abbreviated form of this discussion below.

On nights where the moon is below the horizon during the ~01:30 local-time overpass, DNB imagery is examined over multiple nights for features with high temporal constancy—the same shape—eliminating air glow as a potential illumination source (Miller et al. 2012; Miller et al. 2015). After a potential milky sea candidate is identified it is compared against the DNB cloud mask and infrared bands to confirm it is not a recurring cloud feature.

Using ocean currents taken from the U.S. Navy's Hybrid Coordinate Ocean Model (HYCOM) the milky sea candidate is examined to see if the observed motion, or lack thereof, aligns with what is expected. Day-time DNB imagery from the ascending node overpass (~13:30 local time overpass) is used to rule out a reflective suspension in the ocean.

The DNB imagery in this chapter was taken from the National Aeronautics and Space Administration's (NASA) Worldview application which is part of the NASA Earth Science Data and Information System (ESDIS). NASA Worldview is accessible online and allows for easy comparison between numerous co-located remote sensing products and VIIRS-DNB imagery.

Numerous phenomena impact the coupled earth system state in the Indian Ocean and Maritime Continent regions (Wyrtki 1961; Madden & Julian 1971; Madden & Julian 1972; Webster & Yang 1992; Saji et al. 1999; Kajikawa et al. 2010). In this section we focus on only a subset of such phenomena, namely the IOD, the boreal summer and boreal winter phases of the Indian/South-Asian Monsoon, ENSO, and the austral winter phase of the Indo-Australian Monsoon. To describe these phenomena we drew upon data from the following sources:

1. The Dipole Mode Index (DMI; Saji et al. 1999) for the Indian Ocean Dipole, calculated via the difference in spatial means between the western equatorial Indian Ocean [50°–70°E, 10°N–10°S] and the southeastern equatorial Indian Ocean [90°–110°E, EQ(0°N/S) –10°S]. We calculated the DMI from monthly Hadley Center Sea Ice and Sea Surface Temperature (HadISST; Rayner et al. 2003) SST anomalies at each grid point relative to the climate change corrected 1870-2023 monthly mean SST.
2. The Niño 3.4 index (Trenberth 1997) data was taken from the National Oceanic and Atmospheric Adinistration's (NOAA) Physical Sciences Laboratory (PSL) and was derived from the HadISST1 data of Rayner et al. 2003. This index was calculated via the mean SST anomaly over the 5°N–5°S, 120°–170°W region of the Pacific Ocean.
3. 1000 hPa, 850 hPa, and 200 hPa winds were taken from the European Center for Medium-Range Weather Forecasting's (ECMWF) Reanalysis version 5 dataset (ERA5; Hersbach et al. 2020; Hersbach et al. 2023) and used to construct both the Australian Monsoon Index (AUSMI; Kajikawa et al. 2010) for the Indo-Australian Monsoon as well as the Webster-Yang Index (WY; Webster & Yang 1992) for the Indian Monsoon. AUSMI was calculated via the

mean of 850 hPa zonal wind over 5° –15°S, 100°–130°E. The WY index was calculated via the mean difference between zonal winds at the 200 hPa and 850 hPa levels over EQ[0°N/S]–20°N, 40°–110°E.

This data is used to understand potential sources of predictability for milky seas as well as potential connections within the coupled earth system (Section 2.4). The timing and location of milky sea events are compared against these coupled atmosphere-ocean phenomena using a bootstrapping methodology detailed in Section 2.4.

Eyewitness accounts of milky seas come largely from ocean-going vessels and as such are biased by the location and timing of predominant shipping routes throughout history (see Section 2.2.4.1 for more detail on this bias). Data on global shipping density between January 2015 and February 2021 is taken from the International Monetary Fund's (IMF) World Seaborne Trade Monitoring System (WSTMS; Cerdeiro et al. 2020) and was derived from shipborne Automated Identification Systems (AIS). AIS collates Global Positioning System (GPS) data across the globe to improve the direction and flow of shipborne traffic. The data within WSTMS comes from a mixture of commercial, fishing, leisure, oil/gas, and passenger vessels.

The WSTMS data was originally downloaded as density of total GPS pings on a ~0.005° x ~0.005° grid over the globe before being coarsened onto a ~0.1° x ~0.1° grid. The ship track data used in this research are much later than almost all ship-based accounts in this database. While this does not truly represent the density and location of shipping during the mid-20th century, which makes up the majority of the data points within this database, it is the best available proxy to highlight ship traffic during this time period.

2.2.2 Eyewitness Accounts/Milky Sea Reports Sources

The eyewitness accounts comprising the database described in this chapter come primarily from two sources: the Marine Observer, and the bioluminescence experts recruited by the Marine Observer to explain bioluminescent phenomena observed at sea. Three individuals served as bioluminescence experts for the Marine Observer: E. W. Barlow who held this role from 1927 until his death in 1961, Prof. Roy Kay who held this position from 1963-1974, and finally Prof. Peter Herring who served in this role from 1974 until the Marine Observer ceased publication in 2003.

2.2.2.1 Archives of the Marine Observer Journal

The Marine Observer, published by the United Kingdom Meteorological Office (UK Met. Office) from 1924-2003, with publication paused from 1939-1946 due to World War II, was a journal focused on all aspects of marine meteorology for the members of the UK Met Office's volunteer observing fleet. The UK Met. Office's Digital Library and Archive provides a digital archive of Marine Observer issues.

In addition to publishing on marine meteorology the Marine Observer would also publish accounts of various oceanographic, atmospheric, astrophysical, and biological phenomena that readers would write-in about seeking answers for what they saw. Bioluminescence was one of the most written about topics over the journal's tenure.

The Marine Observer contains several first-hand accounts of rare bioluminescent phenomena. Every issue retained within the UK Met. Office's Digital Library and Archive was scoured for possible accounts of milky seas. Many accounts which were not published in the Marine Observer were retained in the records of the journal's bioluminescence experts.

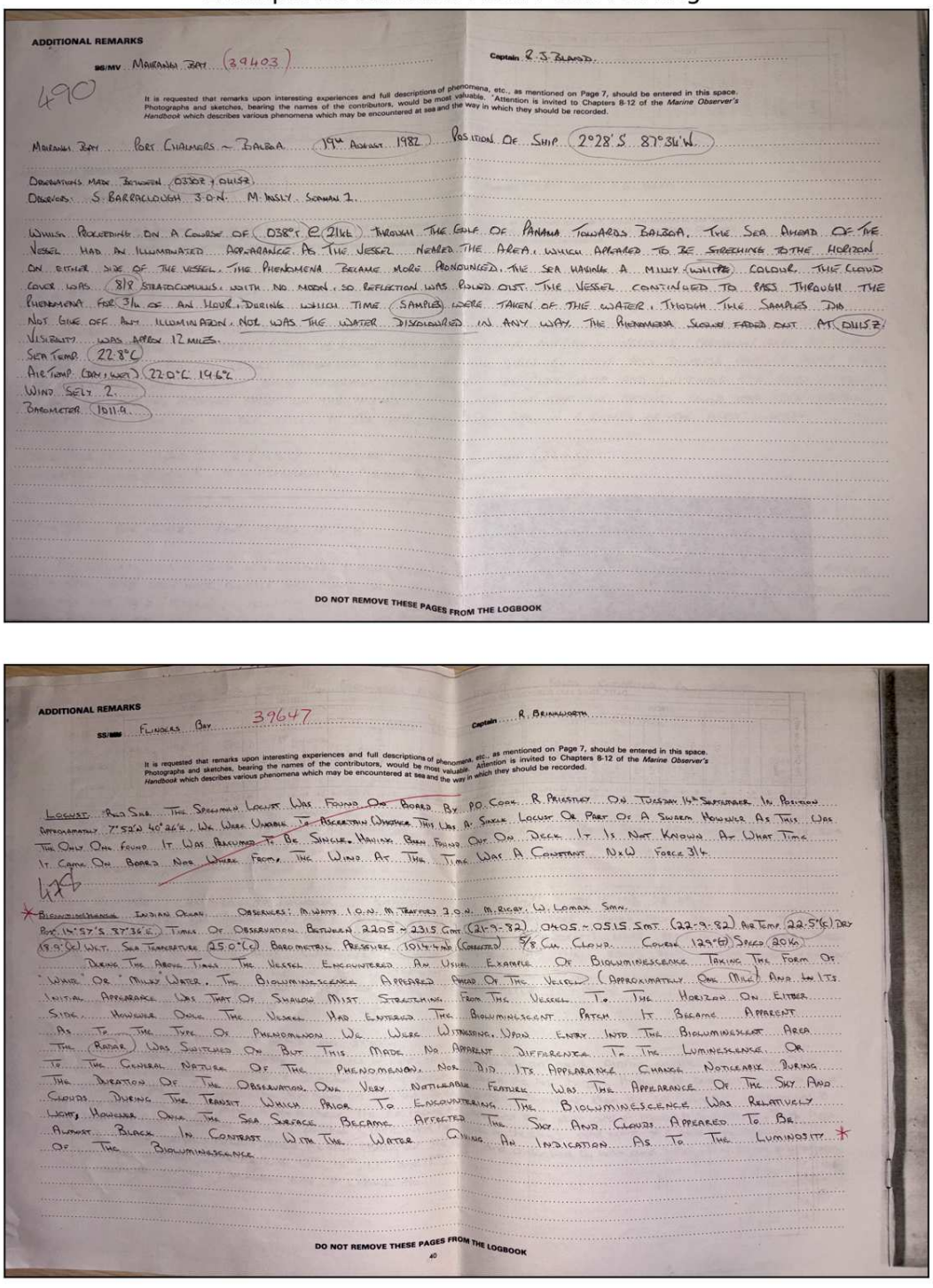
2.2.2.2 Archives of Dr. Peter Herring

Herring inherited the records of his predecessors, and in 1992 along with his then graduate student Maggie Watson published a seminal paper containing a database of general bioluminescent phenomena (Watson & Herring 1992). This database was succeeded by a database confined specifically to milky seas in 1993 (Herring & Watson 1993). While the original databases compiled in these papers have been lost, the publications describing these databases are still available and provide a wealth of insights into the global occurrence and qualitative properties of milky seas.

Herring generously donated his records and that of his predecessors to this research in hard copy form (Figure 2.1). These records, combined with those accounts published in the Marine Observer, comprise 281 of the 394 total eyewitness accounts within our updated database. The database of HW93 featured 232 eyewitness accounts, only 194 eyewitness accounts were identified within the records donated by Herring. There are also 87 eyewitness accounts taken from the Marine Observer which were not found in the records of Peter Herring. It is impossible to truly determine which of the missing 38 eyewitness accounts from HW93 are within these 87 accounts drawn from the Marine Observer, and subsequently, to truly recreate the database of HW93 for comparison.

In addition to missing accounts there were also disagreements in classifying events as milky seas between the database of HW93 and the current work. The subjective nature of classifying milky from primary sources such eyewitness accounts, given their oftentimes imprecise and vague/colloquial language, makes such disagreements inevitable. Fortunately, these instances were minimal and the updated database strives for consistency and uniformity in classification rules across the entire record.

Example Documents From Peter Herring



ADDITIONAL REMARKS

SS/MV MARIANA RAY (29403) Captain R. S. BLAND

490

It is requested that remarks upon interesting experiences and full descriptions of phenomena, etc., as mentioned on Page 7, should be entered in this space. Photographs and sketches, bearing the names of the contributors, would be most valuable. Attention is invited to Chapters 8-12 of the *Marine Observer's Handbook* which describes various phenomena which may be encountered at sea and the way in which they should be recorded.

MARIANA RAY PORT CHALMERS ~ BALBOA 19TH AUGUST 1982 POSITION OF SHIP 2°28'S 87°36'W

OBSERVATIONS MADE BETWEEN (0330 & 0415Z)
OBSERVERS S. BARRACLOUGH 3.O.N. M. INSLY SEAMAN 2.

WHILST PROCEEDING ON A COURSE OF (038°T @ 21kt) THROUGH THE GULF OF PANAMA TOWARDS BALBOA. THE SEA AHEAD OF THE VESSEL HAD AN ILLUMINATED APPEARANCE. AS THE VESSEL NEARED THE AREA, WHICH APPEARED TO BE STRETCHING TO THE HORIZON ON EITHER SIDE OF THE VESSEL. THE PHENOMENA BECAME MORE PRONOUNCED. THE SEA HAVING A MILKY (WHITE) COLOUR. THE CLOUD COVER WAS (8/8 STRATOCUMULUS, WITH NO MOON, SO REFLECTION WAS NOTED ONLY. THE VESSEL CONTINUED TO PASS THROUGH THE PHENOMENA FOR 3/4 OF AN HOUR, DURING WHICH TIME (SAMPLES) WERE TAKEN OF THE WATER, THOUGH THE SAMPLES DID NOT GIVE OFF ANY ILLUMINATION, NOR WAS THE WATER DISCOLOURED IN ANY WAY. THE PHENOMENA SLOWLY FADED OUT AT (0415Z)
VISIBILITY WAS APROX 12 MILES.
SEA TEMP. (22.8°C)
AIR TEMP. (DRY, WET) (22.0°C 19.6°C)
WIND SELY. 2.
BAROMETER (1011.9)

DO NOT REMOVE THESE PAGES FROM THE LOGBOOK

ADDITIONAL REMARKS

SS/MV FLINDERS BAY 39647 Captain R. BRINKWORTH

It is requested that remarks upon interesting experiences and full descriptions of phenomena, etc., as mentioned on Page 7, should be entered in this space. Photographs and sketches, bearing the names of the contributors, would be most valuable. Attention is invited to Chapters 8-12 of the *Marine Observer's Handbook* which describes various phenomena which may be encountered at sea and the way in which they should be recorded.

LOCUST. PALE SMR. THE SPECIMEN LOCUST WAS FOUND ON BOARD BY P.O. COOK R. PRIESTLEY ON TUESDAY 14TH SEPTEMBER. IN POSITION APPROXIMATELY 7°52'N 40°26'E. WE WERE UNABLE TO ASCERTAIN WHETHER THIS WAS A SINGLE LOCUST OR PART OF A SWARM HOWEVER AS THIS WAS THE ONLY ONE FOUND IT WAS ASSUMED TO BE SINGLE. HAVING BEEN FOUND OUT ON DECK. IT IS NOT KNOWN AT WHAT TIME IT CAME ON BOARD NOR WHERE FROM. THE WIND AT THE TIME WAS A CONSTANT NxW FORCE 3/4

BIOLUMINESCENCE INDIAN OCEAN OBSERVERS: M. WATTS 1.O.N. M. TRAFFORD 2.O.N. M. RIGBY. W. LOMAX SMN.
POS 14°57'S 87°36'E. TIMES OF OBSERVATION BETWEEN 2205 ~ 2315 GMT (21-9-82) 0405 ~ 0515 SMT (22-9-82) AIR TEMP 22.5°C (DRY 19.9°C) WET. SEA TEMPERATURE 25.0°C) BAROMETRIC PRESSURE 1014.4MB (CORRECTED) 5/8 CU. CLOUD. COVER 129°(T) SPEED 20Kts
DURING THE ABOVE TIMES THE VESSEL ENCOUNTERED AN USUAL EXAMPLE OF BIOLUMINESCENCE TAKING THE FORM OF 'WHITE' OR 'MILKY WATER'. THE BIOLUMINESCENCE APPEARED AHEAD OF THE VESSEL (APPROXIMATELY ONE MILE) AND IN ITS INITIAL APPEARANCE WAS THAT OF SHALLOW MIST STRETCHING FROM THE VESSEL TO THE HORIZON ON EITHER SIDE. HOWEVER ONCE THE VESSEL HAD ENTERED THE BIOLUMINESCENT PATCH IT BECAME APPARENT AS TO THE TYPE OF PHENOMENON WE WERE WITNESSING. UPON ENTRY INTO THE BIOLUMINESCENT AREA THE (RADAR) WAS SWITCHED ON BUT THIS MADE NO APPARENT DIFFERENCE TO THE LUMINESCENCE OR TO THE GENERAL NATURE OF THE PHENOMENON. NOR DID ITS APPEARANCE CHANGE NOTICEABLY DURING THE DURATION OF THE OBSERVATION. ONE VERY NOTICEABLE FEATURE WAS THE APPEARANCE OF THE SKY AND CLOUDS DURING THE TRANSIT WHICH PRIOR TO ENCOUNTERING THE BIOLUMINESCENCE WAS RELATIVELY LIGHT, HOWEVER ONCE THE SEA SURFACE BECAME AFFECTED THE SKY AND CLOUDS APPEARED TO BE ALMOST BLACK IN CONTRAST WITH THE WATER GIVING AN INDICATION AS TO THE LUMINOSITY OF THE BIOLUMINESCENCE

DO NOT REMOVE THESE PAGES FROM THE LOGBOOK
40

Figure 2.1: Examples of documents containing milky sea eyewitness accounts from the records provided by Dr. Peter Herring.

2.2.2.3 Archives of Dr. Tim Wyatt

Dr. Tim Wyatt, who compiled notes, research, and an unpublished database on milky seas, generously donated his records towards the creation of our updated database. The data provided by Wyatt features many accounts in Dutch which complement the largely English record of milky seas from the Marine Observer and Prof. Herring. Wyatt's notes and research also provided eyewitness accounts or locations/dates that complement the accounts from Herring. Among the 394 total eyewitness accounts in our updated database, 72 (~18%) come from the unpublished notes/database of Dr. Wyatt.

2.2.2.4 Contemporary Accounts and Satellite-Based Data

The remaining 41 eyewitness accounts of the updated database that did not come from the Marine Observer, Herring, or Wyatt originate from M21, scientific publications on milky seas, newspaper reports, ship logs, and travel diaries from individuals who sailed through a milky sea. At the time of this writing, there are also 24 satellite observations of milky seas. Together this gives the database presented here 417 observations of milky seas spanning from 1615 to 2025, or approximately 1 observation per year. Given the significant temporal gaps in this record, a more realistic number may come from the Marine Observer time period which yields ~2-3 milky sea observations a year as per HW93.

2.2.3 Milky Sea Database Construction Methodology

Eyewitness accounts of milky seas were judged based on how well the accounts conform to the features of a 'classical' milky sea account as described in Section 2.1. The language used in many eyewitness accounts is vague or imprecise, for the oldest accounts the language and

descriptions are also antiquated. An example of antiquated language is the use of the word phosphorescence. This term was not coined until the late 1700s and doesn't appear in the records to describe milky seas until 1832.

The eyewitness accounts examined during the creation of this database were judged in as unbiased and equal manner as possible. Given the unknowns around milky seas, accounts were considered primarily on their content versus their location and/or time of year. To add uncertainty to the database eyewitness accounts were classified based on our own assessment of confidence in truly representing a milky sea. The classification categories used were very low confidence, low confidence, and high confidence. Accounts classified as low and very low confidence may represent atypical milky sea events and were retained for completeness.

After eyewitness accounts were determined to be a milky sea and classified based on confidence, additional information about eyewitness accounts when present (e.g., location, date, time of day, observing ship, observers, source of the observations) and other details of the account itself were captured and paraphrased to accompany the entry within the database. This additional layer of qualitative/contextual information was not present in the strict date/location information of past databases and is a potentially valuable resource for future milky sea research.

2.2.4 Potential Sources of Bias

Based on the nature of the eyewitness accounts there are several potential sources of uncertainty and bias:

- Shipping lanes and UK economic activity between 1924 and 2003
- Language Barriers
- Sampling non-uniformity in time
- Possible multiple eyewitness observations for a single milky sea event

In addition, the low-light imagers used to remotely sense milky seas have their own limitations which must be accounted for. We explore each of these sources in the following sub-sections.

2.2.4.1 Shipping Route Bias Between 1924 and 2003

Many of the eyewitness accounts come from individuals traveling along established shipping/trade routes engaged in economic activity such as trade between the UK and India. The locations of these shipping lanes can bias the reported locations of milky sea sightings. Figure 2.5 shows how the density of milky sea sightings falls off with distance from major shipping lanes. This sampling bias means eyewitness accounts may under-report milky sea occurrence and lead to statistical anomalies. One such anomaly occurs in the South China Sea during February 1967. Of the nine milky sea observations in the South China Sea region, six come from a single well observed milky sea event in February 1967.

Because many of the ships encountering milky seas were engaged in economic activities, there may be biases in reporting frequency due to global economic swings, such as the Great Depression, and due to transformations in information sharing and the economy due to war such as World War II (see Figure 2.3 Panels A and B). Anecdotally, the Confederate States Ship (CSS) Alabama encounter in 1864 occurred during the American Civil War, and numerous reports in the northern Arabian Sea in the late 20th century come from the U.S. Naval fleet meaning the timing and location of military operations due to political events may also serve to bias when and where milky seas have been observed historically.

2.2.4.2 Language Barriers

The native tongue of most individuals who serve as sources for milky sea sightings is English. 346 of the eyewitness accounts within the database are written in English (~87%). The majority of non-English eyewitness accounts come from the unpublished database of Dr. Wyatt. There are likely accounts from peoples indigenous to the regions in and around the Indian Ocean but due to the language barrier for those who worked on this topic in English over the centuries these accounts have largely been inaccessible.

Only one account of a milky sea within this database comes from a language native to the regions where milky seas commonly occur. The Biosphere Foundation conducted an interview with a fisherman in Dungkean, Indonesia who had experienced a milky sea in 2017. Due to a lack of transcript in Bangaai, an English translation of the interview was used to generate the data-point corresponding to this interview.

2.2.4.3 Low-Light Imager Limitations

Two different low-light imagers have been utilized to study milky seas remotely. The VIIRS DNB and the Operation Line Scanner (OLS) that flew aboard Defense Meteorological Satellite Program (DMSP) satellites. Both sensors are sensitive to radiation in the visible to near-infrared band which allows them to see part of the bioluminescent emission spectrum thought to be responsible for milky seas. However, this sensitivity means that moonlight, which is several orders of magnitude brighter than the bioluminescence, completely dominates milky sea light emissions through atmospheric and surface backscatter.

Due to this limitation, milky sea events that are of a sufficient spatial scale and brightness to be observed by the VIIRS DNB and the DMSP OLS can only be seen when the moon is below

the local horizon at the time of the satellite overpass. The effect of this is that each sensor can, at best, detect a milky sea approximately 50% of the time. For the ~0130 local overpass time, the nights of moon-free imagery are roughly two nights after the Last Quarter lunar phase through two nights after the First Quarter lunar phase. This period is followed by approximately two weeks of moonlight-contaminated imagery.

We performed two simple statistical experiments to understand the extent of this moonlight-induced sampling bias. Milky seas primarily occur during the July-August-September time period. The number of moon-free nights for the VIIRS DNB to detect a milky sea during this 92-day period is dependent upon when the first new-moon in July occurs. As a spatially coherent signal over multiple nights is needed for a trained observer to distinguish a milky sea from ephemeral clouds and air glow, the detection of a milky sea is also dependent on how many nights the milky sea event persisted.

Using a uniform distribution over the 92-day time period of interest for one experiment and a normal distribution centered on August 16th with a standard deviation of 15-days for the other experiment to simulate the likelihood of a milky sea occurring during this time period, we can estimate how likely the milky sea is to be seen in a VIIRS DNB image before accounting for cloud cover (Figure 2.2). To estimate this probability of detection, we assumed that any milky sea of sufficient brightness and spatial scale to be detectable in DNB imagery would need to be observed across the span of at least 2 contiguous moon-free nights. Each distribution was sampled 100,000 times for every combination of milky sea event length and configuration of the lunar cycle during this time period based on when the first new moon occurred in July.

For milky sea events of sufficient brightness and spatial scale to be observed by the DNB that last 3 nights or less there is at most a 50% chance of satellite detection, and at most a 25%

chance of being observed under 50% cloud cover. In contrast, for events that last longer than 17 days there is a greater than 99% chance of being observed. Overall, these results show that like the shipboard accounts the VIIRS DNB record is likely to undercount the actual number of milky sea occurrences.

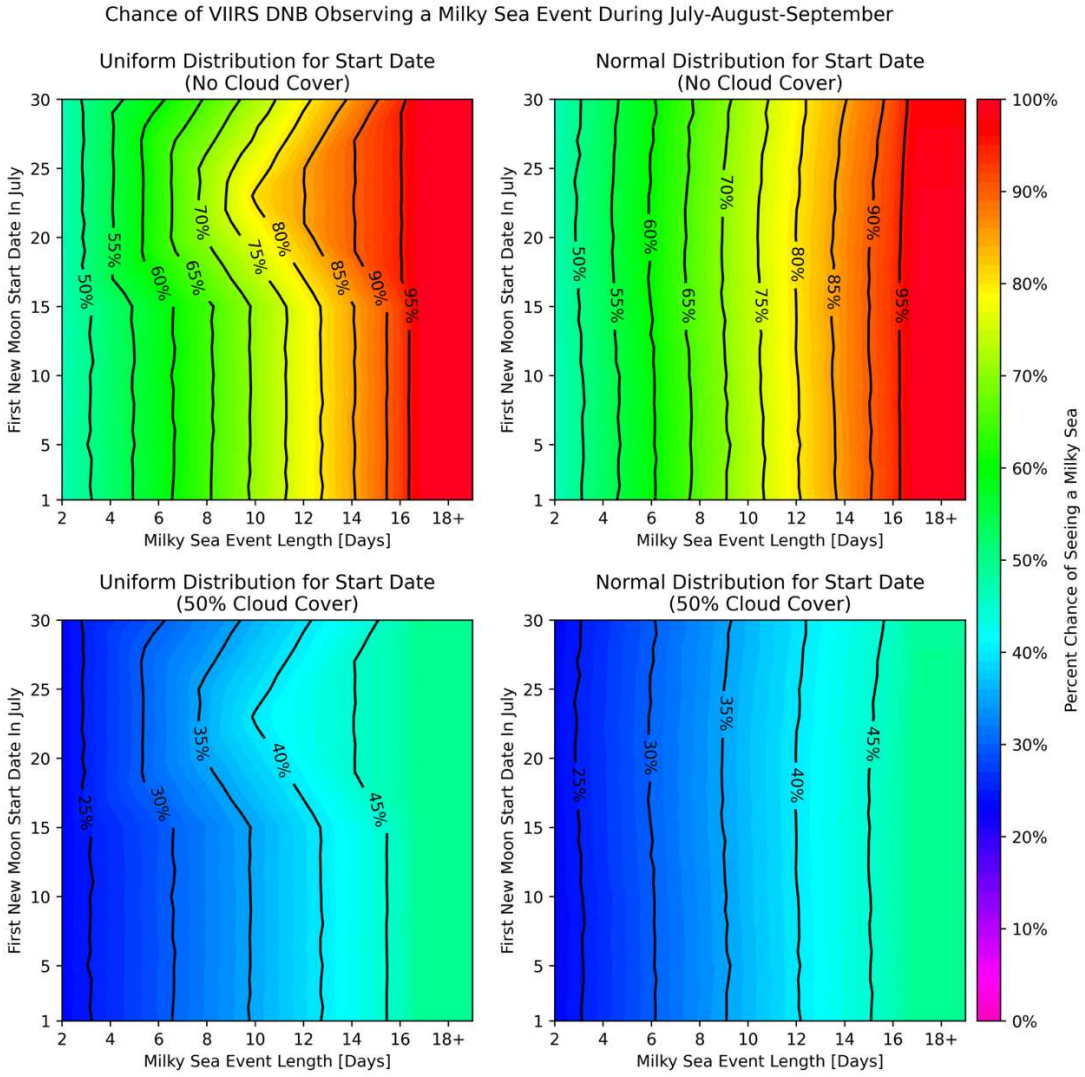


Figure 2.2: Result of simple statistical experiment to estimate the extent of the undersampling bias by the VIIRS DNB assuming a uniform distribution for the start date of a milky sea over the July-August-September (JAS) time period (left column). The results for a similar experiment assuming a normal distribution over the JAS time period centered on August 16th with a standard deviation of 15 days to more closely match the observed seasonal distribution of milky seas (right column).

2.2.4.4 Sampling Non-Uniformity in Time

The various biases described in Sections 2.2.4.1-2.2.4.3 lead to strong sampling non-uniformity across time that degrades the quality of the database. A strong example of this bias is seen near Java, Indonesia. This is one of the major hotspots for milky sea occurrences globally, despite this however, there are no reports in this region between August 1880 and August 1961. This 81-year gap is likely due to the paucity of surviving sources or changes in trade or politics rather than a physical reason. However, we cannot rule out some unknown multi-decadal change in the coupled earth system.

2.2.5 Bootstrapping Methodology

To understand statistical relationships between milky sea events and coupled atmosphere ocean phenomena, the eyewitness accounts must be grouped together into distinct and independent milky sea events. Doing so places well-observed milky sea events with multiple eyewitness observations and those with only a single eyewitness observation on equal footing for statistical analysis.

To collapse the observations into unique events, they must be grouped in space and time. A radius R in great-circle degrees and a maximum temporal extent N in days/nights were used to group together each milky sea sighting, S_M, in chronological order. For example, starting with the first sighting in chronological order, S_1, all subsequent sightings that occurred within N days of S_1 (i.e., $S_2, S_3, S_4, \ldots$) were examined to see if they occurred within the radius R drawn around S_1. All subsequent sightings S that could be grouped with S_1 based on our choice of R and N would be grouped together to create the event E_1, all sightings grouped into E_1 were removed from consideration for other events. The next sighting in chronological order after S_1, i.e. $S_{>1}$, would be

used as the basis for event E_2 and all sightings that occurred within N days of $S_{>1}$ not included in E_1 would be considered. This process would continue until all sightings had been grouped into distinct milky sea events.

Once this iterative process is completed and all milky sea sightings have been grouped into events the timing of each milky sea event relative to the state of the coupled atmosphere-ocean phenomena of interest can be examined. As an example, we can count the number of milky sea events that occurred in boreal winter during the peak of an El Niño event versus the number of milky sea events that occurred in boreal winter during the peak of La Niña.

Each region of interest for milky seas has a distribution, D_{OBS}, describing how many milky sea events in that region per year correspond to a particular coupled earth system phenomenon. By taking the statement that D_{OBS} is independent of a particular coupled earth system phenomenon as our null hypothesis we can randomly shuffle the years that make up D_{OBS} into a new randomized distribution D_{RD}. As an example, the number of observed milky sea events in 1999 is now the number for 1979, and the number for 1979 is now 1968 etc.

Bootstrapping is a statistical procedure used for estimating the distribution of a parameter by resampling (or shuffling) the data. For this research, the resampling to produce D_{OBS} was carried out 5,000 times to produce a bootstrapped distribution through which the observed value could be compared. We evaluated the observed value to see if it exceeded two-tailed 80%, 85%, 90%, and 95% significance levels for our bootstrapped distribution.

To carry out this bootstrapping procedure, a representative spatiotemporal scale for milky sea events is needed. These values can be estimated from the satellite observations and the eyewitness accounts in this database, but with a large degree of uncertainty due to the biases described in Section 2.2.4. Privileging a specific combination of R and N could skew the results if

the relationships are sensitive to the definitions used to group observations into events, vary between region, or even seasons. By carrying out the bootstrapping procedure for a range of values for *R* and *N* this uncertainty can be accounted for and the strength of the results increased.

N was varied from 2 days to 90 days. These values were chosen due to 2 days being the minimum number of days needed to identify a milky sea in satellite imagery, and 90 days being twice the length of the longest observed milky sea event in the satellite record. The values of N used were 2, 3, 4, 5, 7, 10, 15, 20, 30, 40, 50, 60, 70, 80, and 90 days. R was varied from 0.25^{o} to 8.0^{o}, near the equator 1^{o} ≈ 111 km, and this corresponds to the sizes of the smallest satellite observed milky sea events and slightly larger than the largest satellite observed milky sea event (Miller et al. 2021).

By conducting the bootstrapping experiment for each combination of *R* and *N* this allows for statistical relationships between milky seas and the phenomena of interest to be tested based on their dependence of *R* and *N*. The weaker the dependence on *R* and *N* for a relationship the more robust the relationship likely is. The results of this experiment are shown in Section 2.4.

2.3 Results

The final database contains 417 observations of milky seas, an increase of 182 compared to the 235 reported on in HW93. The eyewitness accounts provide information as to how often particular properties of milky seas were observed. 17% of accounts reported the milky sea having a well-defined edge (a demarcation between normal/dark and glowing near-surface waters), 21% of accounts reported a perceived or actual calming of the ocean surface within a milky sea. Nearly half of milky sea eyewitness accounts (46%) specify that the event spanned from the ship to the horizon.

Some less frequently reported features of milky seas that also emerged from the eyewitness accounts are that approximately 10% of the time other forms of bioluminescence were coincident, and approximately 10% of eyewitness accounts specify they observed the milky sea over multiple nights. 13 different eyewitness accounts specified that the light emission from the milky sea was of sufficient intensity to illuminate the underside of clouds.

Due to the matter-of-fact way that many of the reports were written (Figure 2.1), and instructions from the Marine Observer to focus particularly on the color of the bioluminescence and its reaction to stimuli versus the features used to distinguish bioluminescence, these values likely represent the lower bound with which they occur within milky seas. Overall, this shows that the database, in addition to capturing the classical features of a milky sea, could reveal other relationships and features of milky sea events not often reported in the literature.

Of the 417 entries within the database the majority (324, or 78%) occurred in the 20th century (Figure 2.3 Panel A), and 271 (65%) occurred in the second half of the 20th century. This disparity may be primarily due to the publishing timeline of the Marine Observer from which our primary sources are largely derived, and the data quality within the database likely peaks in the second half of the 20th century due to this bias as well. The importance of the Marine Observer for our database highlights the critical importance of a central repository for simple reporting of oceanic phenomena by seafarers to studying rare marine phenomena.

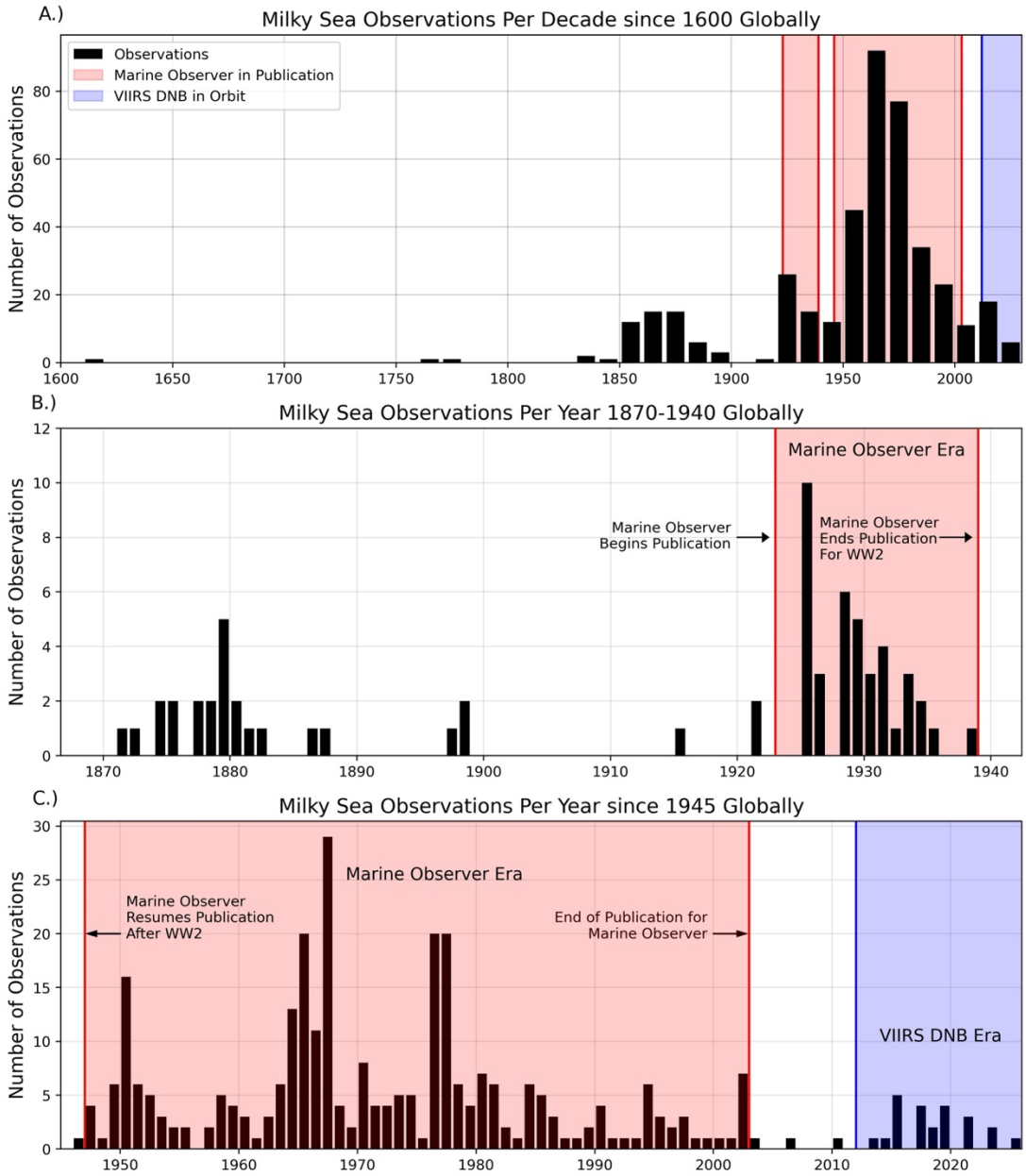


Figure 2.3: Temporal distribution of milky sea eyewitness accounts within the database. The publishing timeline of the marine observer (highlighted in red), and the time period where the VIIRS DNB is in orbit (highlighted in blue, panels A and C) bias the number of observations in the database.

2.3.1 Spatial Distribution of Milky Seas

The global distribution of milky seas in our database is primarily concentrated around the NW Indian Ocean with additional hotspots near the island of Java, Indonesia and in the Banda Sea (Figure 2.4). This distribution agrees with and builds upon the findings of previous efforts to

catalogue milky seas (Smith 1931, Turner 1965, Herring and Watson 1993, Miller et al. 2021). Over 60% of all observations within the database occur in the NW Indian Ocean, with 19% occurring in the Maritime Continent.

Sporadic observations of milky seas span the globe with accounts from the coasts of the United States, Canada, New Zealand, South Africa, and Spain. Understanding how these regions produce milky seas on rare occasions compared to the Indian Ocean where milky seas occur on a more regular basis is a topic of future study.

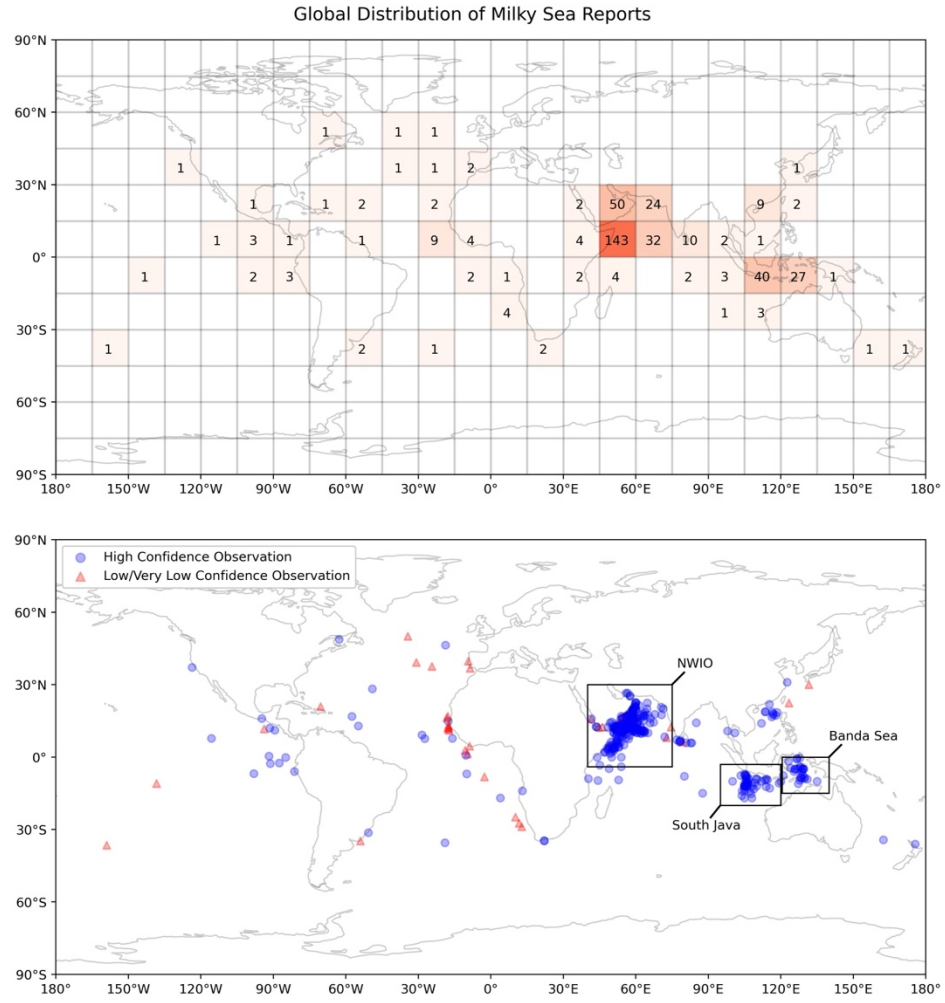


Figure 2.4: Global distribution of milky sea sightings per 15° x 15° box (upper panel) and as discrete data points coded by the author's confidence in the observation truly representing a milky sea (lower panel). The black boxes in the lower panel outline the main hotspots for milky seas globally.

Comparing milky sea observations in the NW Indian Ocean and near Java, Indonesia with modern shipping lanes from the WSTMS (Figure 2.5) shows the impact of sampling bias to our current database. Focusing on the two primary milky sea time periods (boreal winter and boreal summer) for the NW Indian Ocean (Figure 2.6) we see that this bias manifests in different ways. In the boreal summer, milky sea observations cluster along the shipping lane that runs to the north of Socotra, Yemen. During boreal winter, milky sea observations cluster along a shipping route that runs SW-NE from approximately 0° N, 50° E towards the Gulf of Oman [23° N, 60° E]. Such potential biases are pointed out originally by HW93.

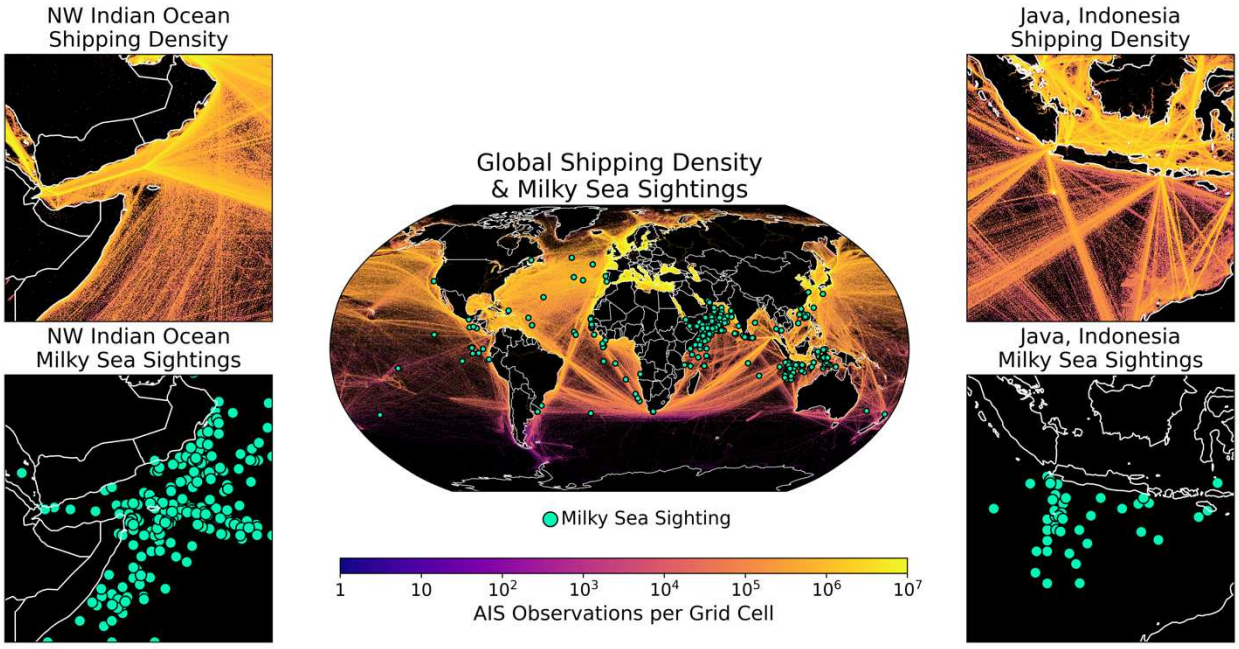


Figure 2.5: Comparison between milky sea observations and shipping density as observed by the IMF's WSTMS between 2015-2021. The location of milky sea sightings align closely with the locations of major shipping lanes in all regions.

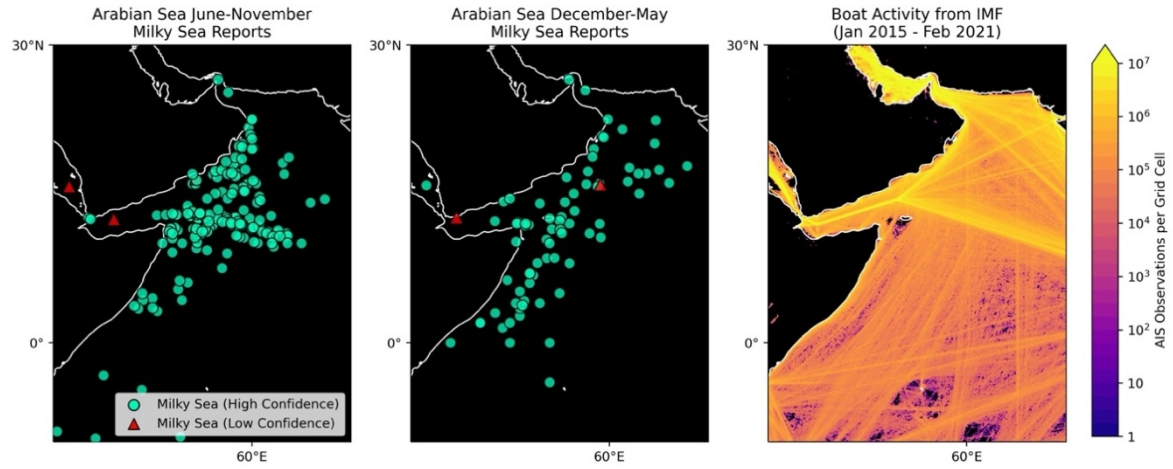


Figure 2.6: Comparison between boreal summer (June-November) and boreal winter (December-May) sightings of milky seas in the northwest Indian Ocean/Arabian Sea region compared to the IMF's WSTMS data. There is a possible seasonal pattern of milky sea sightings with more sightings south of Socotra, Yemen during boreal winter compared to boreal summer.

Our results are consistent with HW93, who suggest that there may be a seasonality to where milky seas occur in the NW Indian Ocean with boreal summer milky seas primarily occurring north of Socotra and boreal winter milky seas primarily occurring south of Socotra. This result may be tied to differing ocean currents and wind stress patterns between the different phases of the Indian Monsoon as well as the growth/decay phases of the IOD. Understanding how the dynamical forcings behind milky seas vary between boreal summer and boreal winter is a topic for future study.

Whereas the spatial distribution of milky seas reported this study reinforces the results of previous studies (Turner 1965, Herring & Watson 1993; Miller et al. 2021), they provide new insight on the global extent of the phenomenon, with several new data points in locations previously uncaptured. We anticipate that with increased awareness and steady curation, our new

milky sea database will continue to expand both in terms of future events and backfilling of historical ones.

2.3.2 Intra-Year Distribution of Milky Seas

Understanding when during the year milky seas occur provides insights into potential drivers within the coupled earth system as well as serves as a point of comparison between previous scholarship on milky seas. The global seasonality of milky seas within this database (Figure 2.7 Panel A) has a bimodal distribution with peaks in January and August. This finding agrees with the results of HW93 (Figure 2.7 Panel B).

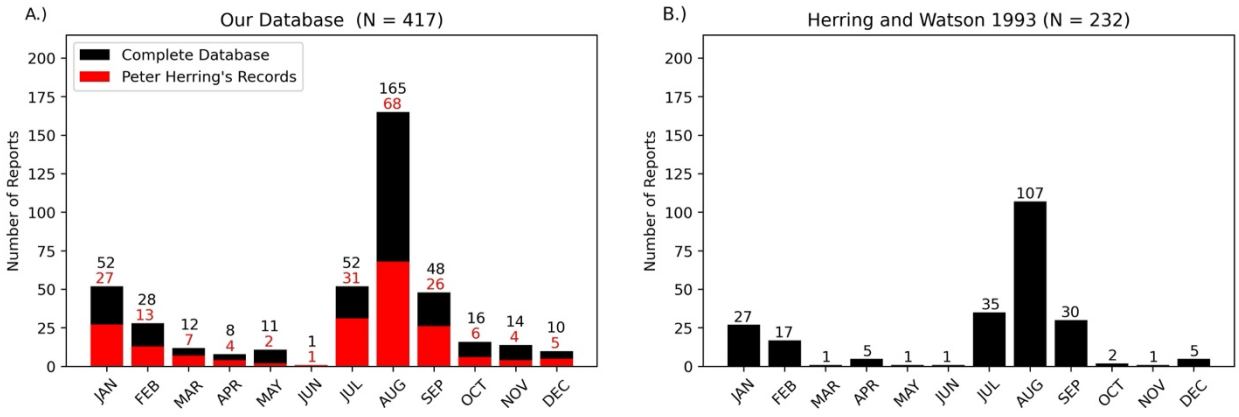


Figure 2.7: Comparison of the global seasonal distribution of milky seas between the database presented here (Panel A) and that of Herring and Watson 1993 (Panel B). The red numbers are vertical bars in Panel A represent the data taken directly from the records of Peter Herring, showing that there is likely both missing data and disagreement in classifying observations as milky seas between the two databases.

The red numbers and histogram bars in Figure 2.7 Panel A show the comparative distribution between our database as a whole and the records which came directly from Dr. Herring that went into HW93. This shows that fully recreating the database of HW93 is likely not possible

due to incomplete data and some disagreements in classifying eyewitness descriptions as milky seas between this work and HW93.

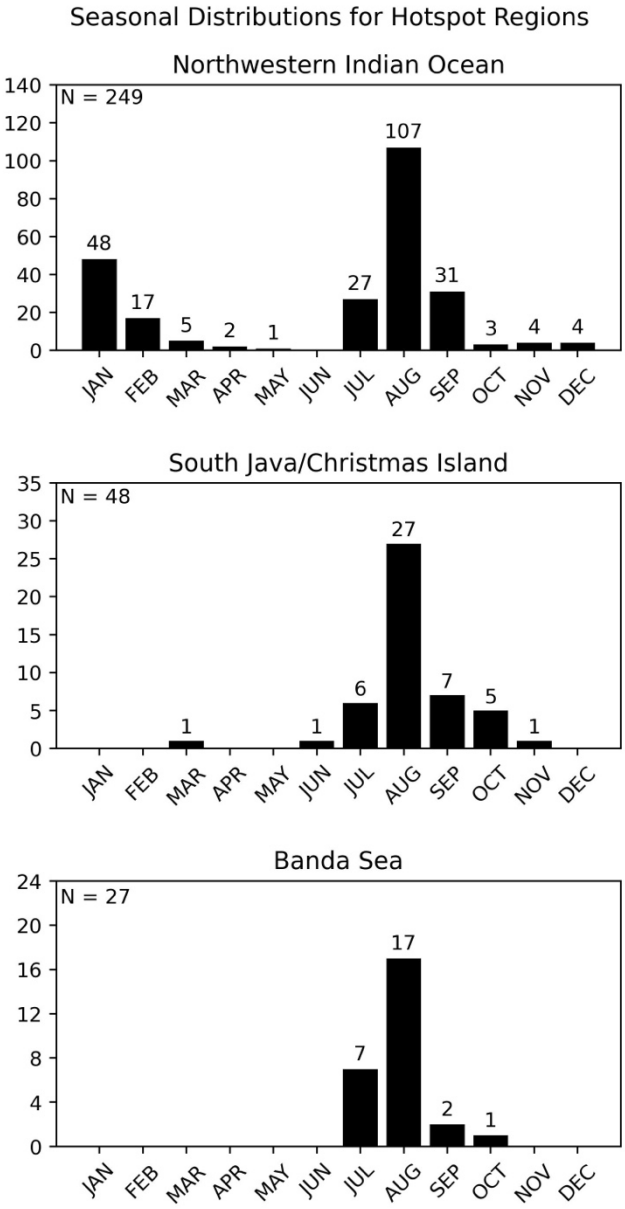


Figure 2.8: Seasonal distribution of milky sea sightings for the three main regions where milky seas occur. All three regions have the peak of milky sea occurrence in August, with the NW Indian Ocean having a secondary peak in January.

Breaking this intra-year distribution down by region (Figure 2.8) compares well with previous research. The NW Indian Ocean (accounting for more than half of the database's content) is the primary source of the bimodal distribution as seen in Figure 2.7 Panel A. The boreal winter and boreal summer peaks in this region correspond well with the timings of the Indian Monsoon. This relationship is investigated further in Section 4.

For the area south of Java, Indonesia aside from a single account from March the milky sea accounts are distributed tightly around August. This period corresponds with the austral winter phase of the Indo-Australian monsoon and the onset of the IOD (either positive, neutral, or negative). In the Banda Sea hotspot for milky sea reports and satellite observations, sightings are similarly distributed around August. How the timing of milky seas in these regions compares to large-scale atmospheric/ocean phenomena is investigated in Section 4.

2.3.3 Milky Sea Environmental Conditions

In their investigations of the environmental conditions surrounding satellite observed milky sea events M21 find that milky seas primarily occur in waters with relatively cool SSTs, calm ocean surface currents, and adjacent to high-values of chlorophyll-a. While exact ocean current and chlorophyll-a conditions cannot be drawn from the eyewitness accounts, 168 eyewitness accounts of our database reported observed windspeed conditions within or near the observed milky sea events, and 146 eyewitness accounts reported SSTs. By comparing these eyewitness accounts against decades of observed windspeed and SST data from the Arabian Sea/NW Indian

Ocean (the primary region where the windspeed and SST measurements originate) we can better understand the environmental conditions surrounding milky seas (Figure 2.9).

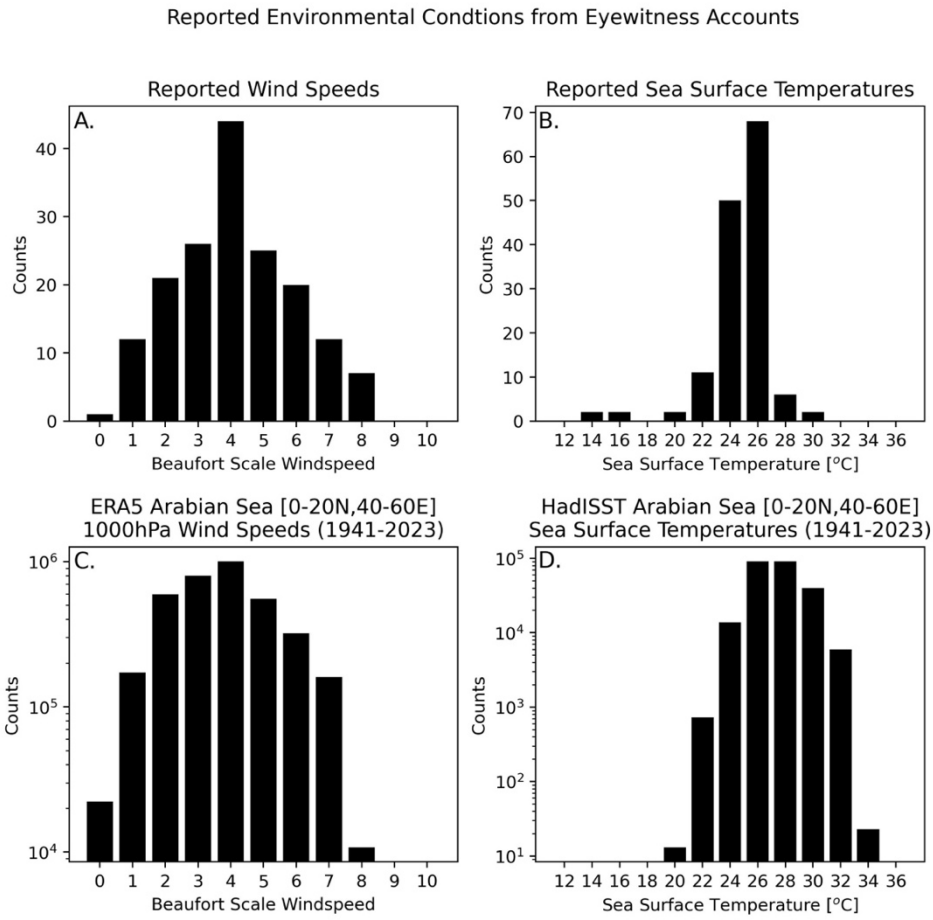


Figure 2.9: A comparison between reported wind speeds (Panel A) and SSTs (Panel B) from eyewitness accounts of milky seas and those taken from atmospheric and oceanic data sets. The winds in Panel C are 1000 hPa winds from ERA5, and the SSTs (Panel D) are taken from HadISST, both over the time period of 1941-2023.

The reported windspeeds from eyewitness accounts (Figure 2.9 Panel A) are approximately normally distributed about 4 on the Beaufort Scale for windspeed (approximately 6-8 m/s). The accounts at high windspeeds dispute the Lapota et al. (1988) surface slick hypothesis for milky seas morphology, something HW93 considered dubious as well, as surface slicks would be broken up under high windspeeds such as 6, 7, and 8 on the Beaufort Scale (Alpers & Hühnerfuss 1989).

Similarly, the observations at low windspeeds (Beaufort Scale 0-1) cast doubt on the hypothesis that milky seas are caused by widespread wind-driven disturbance of dinoflagellates, as at these low windspeeds there isn't enough wind stress available to agitate the ocean surface and produce dinoflagellate bioluminescence.

ERA5 1000 hPA windspeeds over the Arabian Sea/NW Indian Ocean between 1941-2023 (Figure 2.9 Panel C; converted to Beaufort Scale to better compare against eyewitness accounts) show a similar distribution as the eyewitness accounts. Comparing the two distributions shows the milky seas occur in every wind condition over the Arabian Sea/NW Indian Ocean further disputing both the surface slick and wind-agitated dinoflagellate hypotheses.

Comparing HadISST SSTs in the Arabian Sea/NW Indian Ocean (Figure 2.9 Panel D) against the SSTs drawn from eyewitness accounts (Figure 2.9 Panel B) aligns milky sea with the bottom half of the SST distribution for this region. Cooler waters are likely associated with upwelling and intensified biological activity from upwelled nutrients. This finding agrees with the results of M21 and the leading biological hypotheses for milky seas.

2.3.4 Size and Duration of the 'Typical' Milky Sea Event

Understanding the population distributions of size and how long a milky sea lasts is critical to assessing how likely we are to detect a milky sea with satellite-based instruments such as the VIIRS DNB (Section 2.2.4.3), and for any potential attempts to intercept and sample a milky sea event in-situ. These population level distributions are also critical for forecasting milky seas and for conducting statistical analyses.

The satellite results of M21, with all the caveats of satellite detection as described in Section 2.2.4.3 being considered, indicate that the typical milky sea lasts two weeks or less and

are on the order of 10,000 km^2 in size. The lifespan estimate from the satellite data however is limited by the approximately two-week lunar cycle time period where milky seas can be viewed via the VIIRS DNB. To estimate the distribution for how long a typical milky sea event lasts from our database, milky sea observations were grouped into events following the methodology outlined in Section 2.2.5, and using thresholds of $R = 8.0^o$, and $N = 90$ days. This grouping strategy resulted in 259 unique milky sea events.

The 259 unique milky sea events identified using this methodology is likely an underestimate of the true number of distinct events captured in our database. Satellite observations of milky sea events have shown that there can be multiple, spatially disjoint regions glowing due to milky sea bioluminescence in VIIRS DNB imagery coexisting (the 2015 Somalia event reported on in M21 is an example of this). While it is possible that the 3-D structure of milky sea events includes large portions of the milky sea event at sufficient depth to be invisible at the surface it is also likely that when the conditions necessary for genesis of a milky sea are present across a region multiple independent milky sea events can form simultaneously. A new milky sea event could also form 60-days after a previous milky sea event arose and dissipated and both events would be combined into a single long-lived event using this methodology.

When milky seas identified through this methodology consisted entirely of eyewitness accounts, we required at least two distinct eyewitness accounts to estimate how long the event occurred. Milky sea events with satellite information defaulted to the satellite record to estimate how long the event lasted (within the confines of lunar-window sampling periods). The distribution for how long the typical milky sea event lasts (Figure 2.10 Panel A) peaks around 1-2 weeks with a long tail extending out to ~90 days. The extent of this tail is sensitive to the threshold of N used

but the peak around 1-2 weeks holds until *N* is reduced below 1-2 weeks. This behavior agrees well this the satellite results of M21.

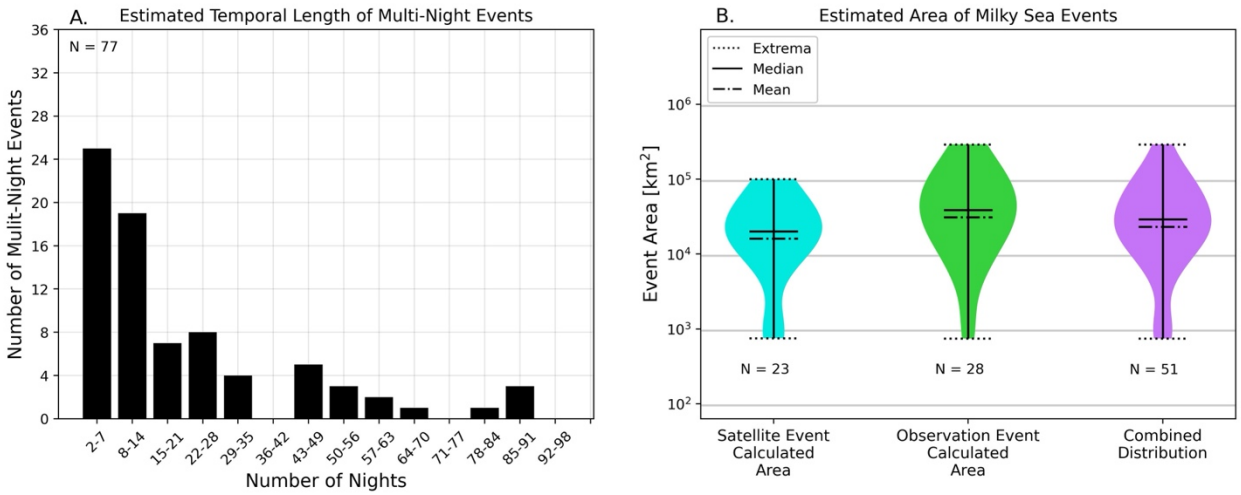


Figure 2.10: Estimated population distributions of milky seas in terms of how many nights they last (Panel A) and the size of milky sea events as estimated from eyewitness accounts and satellite observations (Panel B).

Milky sea events with at least three eyewitness accounts could be used to estimate the spatial extent of a milky sea. The eyewitness account events agree well with the satellite record (Figure 2.10 Panel B). The area of milky sea events identified through this methodology carries its own set of inherent biases. The first is that the method does not account for the drift of a particular milky sea event with the ocean currents. Satellite observed milky sea events reveal drifting of the glowing waters with ocean currents over distances comparable or greater to their horizontal extent during their lifetimes. Another possible bias with our method is that the eyewitness accounts, clustered along shipping lanes, may all encounter the milky sea event in roughly the same location, thus biasing the spatial extent towards smaller values. Despite these biases and those described in

Section 2.2 however, the observational record compares well against the satellite record in terms of the temporal extent and spatial scale of the typical milky sea event.

2.4 Identifying Potential Predictors for Milky Sea Events

Despite almost two centuries of association between milky seas and the monsoons of the Indian Ocean in the scientific literature (Buist 1855; Smith 1931; Turner 1965; Herring & Watson 1993), and recent work showing a plausible connection to the IOD (Miller et al. 2021) it is unknown if these phenomena actually play a role in milky sea formation or are merely coincidental. As large-scale biological phenomena the formation of milky seas is likely tied to wind-driven upwelling leading to enhanced primary production which, along with the time periods when milky seas tend to occur (Section 2.3.2), makes these connections plausible. Understanding what role, if any, large-scale atmospheric/oceanic phenomena play in milky sea occurrence is a critical first step towards a predictive algorithm for milky seas.

To understand a possible connection between milky seas and large scale coupled atmosphere/ocean phenomena the bootstrapping methodology outlined in Section 2.2.4.4 was utilized after dividing milky sea observations into the four spatiotemporal regions below:

1. The NW Indian Ocean [40^{o}-75^{o}E, 4^{o}S-30^{o}N] during boreal summer (June – October)
2. The NW Indian Ocean [40^{o}-75^{o}E, 4^{o}S-30^{o}N] during boreal winter (December – March)
3. South Java/Christmas Island [90^{o}-120^{o}E, 20^{o}-3^{o}S] during austral winter (June – October)
4. Banda Sea [120.5^{o}-140^{o}E, 15^{o}S – 0^{o}N/S] during austral winter (June – October)

These four spatiotemporal groups represent the major hotspots (Figure 2.4) and time periods (Figure 2.8) for milky seas globally, and together account for ~80% of all known milky sea observations from eyewitness accounts and 100% of satellite detected events.

Due to the nature of the sources used to generate the data on milky sea occurrences likely peaking in the second half of the 20th century this bootstrapping experiment was limited to milky sea events occurring between Dec 1st 1949 and Nov 30th 1999, inclusive. The results of the bootstrapping experiment for each of the four spatiotemporal groups above are outlined in Sections 2.4.1 – 2.4.4 respectively.

2.4.1 Bootstrapping Results: NW Indian Ocean During Boreal Summer

The NW Indian Ocean during Boreal Summer is, historically, the most prolific region for milky seas. For this region, three coupled atmosphere/ocean phenomena were considered to modify milky sea frequency: the Indian Monsoon, ENSO, and the IOD. These three phenomena were chosen both for their plausibility in influencing the formation of milky seas via their effects on the atmosphere-ocean state in the northwestern Indian Ocean, and their larger time scales. After the observations have been grouped into events there is uncertainty on exactly when the event began and ended. For example, many satellite observed events have uncertainty on the order of 2-4 weeks on their lifespan estimate due to the lunar cycle. This uncertainty limits the utility of sub-seasonal variation in explaining the frequency of milky sea events as a difference two to four weeks can relate to different local behavior from sub-seasonal phenomena. For example the Madden-Julian Oscillation which modulates the atmosphere-ocean state on sub-seasonal time scales can, based on its eastward propagation, switch from the generation of large scale westerly flow to large scale easterly flow in a location (or vice versa) over 14-28 days.

For the Indian Monsoon, milky sea events were grouped into three categories based on the mean strength of the Webster-Yang monsoon index (Webster & Yang 1992) for that boreal summer. Milky sea events occurring during the top 20% of SW monsoon years were classified as ‘Strong

SW Monsoon Events', milky sea events occurring during the bottom 20% of SW monsoon years were classified as 'Weak SW Monsoon Events', and milky sea events occurring during the remaining 60% of years were classified as 'Average SW Monsoon Events'.

ENSO, an interannual variation in equatorial Pacific sea surface temperatures and winds which impacts global weather patterns (Ummenhofer et al. 2011), is a seasonally phase-locked phenomenon (Battisti & Hirst 1989). During the boreal summer of year, *Y*, a particular phase of ENSO (e.g. La Niña) will begin to develop, peak during boreal winter of (*Y*, *Y*+1), and decay heading into the boreal summer of year *Y*+1. This trend means that boreal summer milky sea events may be associated with both decaying and developing ENSO events. Here, we classified milky sea events based on the decaying ENSO phase and the developing ENSO phase. These classifications were distinct, e.g. a milky sea event could be (e.g.,) a post-positive ENSO (i.e., El Niño) event and a pre-neutral ENSO event. To avoid over-splitting the data, mixed classifications such as post-positive ENSO/pre-neutral ENSO vs. post-positive ENSO/pre-negative ENSO (i.e., La Niña) were not considered.

The IOD is a basin-wide gradient in sea surface temperatures in the Indian Ocean which modifies the atmospheric and oceanic state as well as the biological state of the Indian Ocean (Shi & Wang 2021; Shi & Wang 2022; Shi & Wang 2024). Like ENSO the IOD is a seasonally phase-locked phenomena with the IOD initiating in boreal summer, peaking during boreal autumn, and decaying heading into boreal winter (Saji et al. 1999). Milky sea events were classified based on the developing IOD event for that boreal summer.

The bootstrapping results for this region (Figure 2.11) show, at best, moderate relationships between milky sea events in this region and the coupled atmosphere/ocean phenomena of interest.

Milky sea events in this region were more likely to occur following a La Niña event and less likely to occur following a neutral ENSO year.

The most noteworthy result for this region was the lack of signal for the SW monsoon. If the SW monsoon was the true driver of milky sea events, as speculated, a strong signal for an average SW monsoon would appear. The lack of signal for the SW monsoon suggests that the boreal summer Indian Monsoon likely sets a favorable background for milky seas independent of its intensity and other processes, perhaps at finer spatiotemporal scales, truly drive milky sea formation in this region.

Examples of smaller scale processes which were not examined in this work that may play a role are atmospheric transport and deposition of dust from nearby arid regions such as Somalia or the Arabian Peninsula, and the strength and location of the Great Whirl, an annually reoccurring oceanic eddy south of Socotra, Yemen.

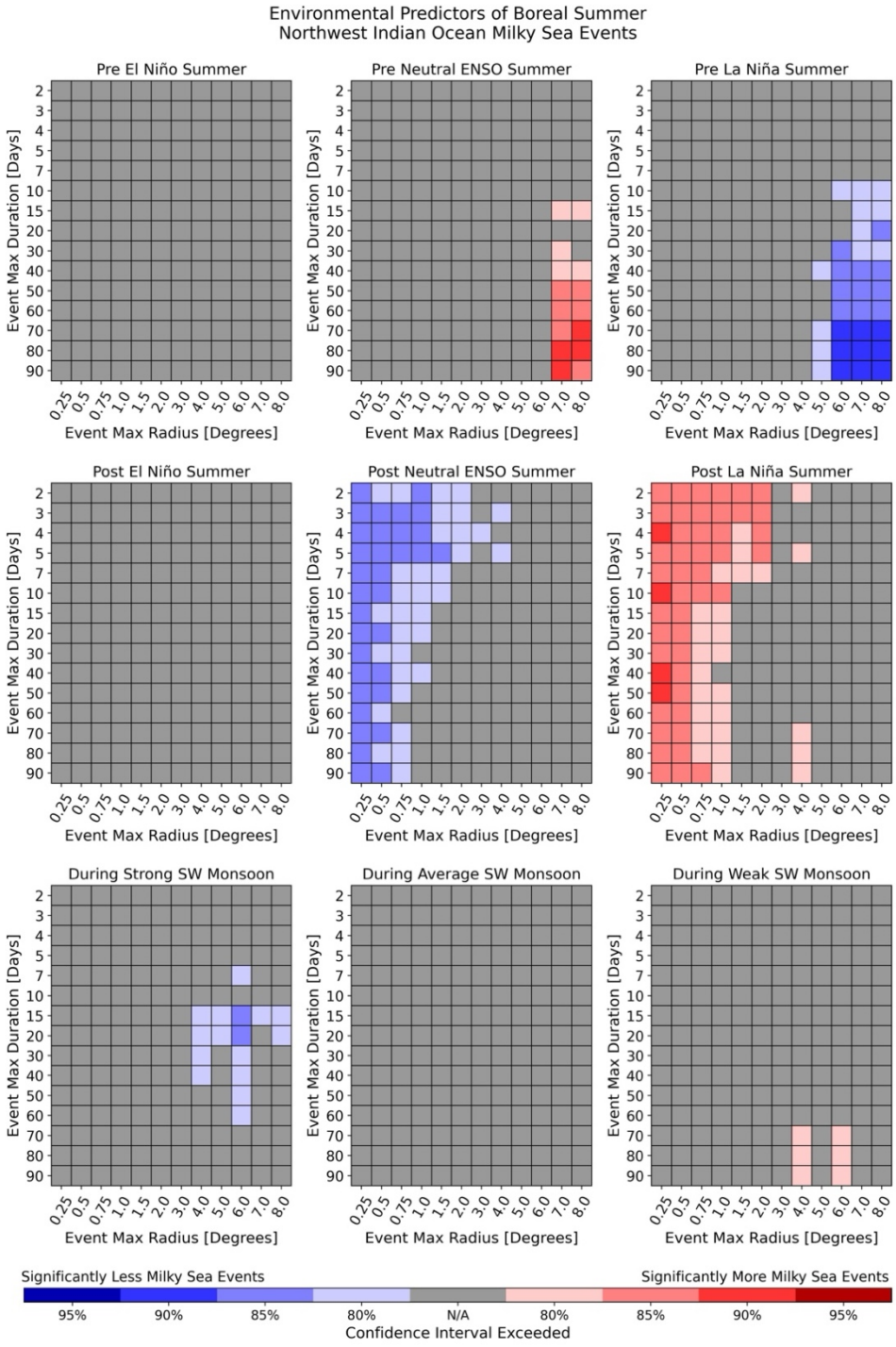


Figure 2.11: Results of the bootstrapping experiment for the Northwestern Indian Ocean during Boreal Summer (Section 2.4.1). Milky seas were more likely following a La Niña and less likely following a neutral ENSO event.

2.4.2 Bootstrapping Results: NW Indian Ocean During Boreal Winter

During the December-March time period of boreal winter the northeast (NE) phase of the Indian Monsoon initiates, and, for years where the IOD and ENSO are not neutral, the IOD decays, and ENSO ramps up to its peak strength. The relationship between milky sea events and these three phenomena were examined (Figure 2.12). Like for the SW monsoon milky sea events were classified based on the strength of the NE monsoon as defined by the Webster-Yang monsoon index. The Webster-Yang monsoon index is negative for the NE monsoon so the strongest NE monsoon events are the ones where the index is most negative (bottom 20% in terms of index value) and the weakest NE monsoon events are those where the index is at its most positive (top 20% in terms of index value), average monsoon events were defined as the remaining 60% of NE monsoon events.

When considering the role of the IOD, milky sea events were classified based on the phase of the decaying IOD event as defined by the DMI. When considering the role of ENSO, milky sea events were classified based on the concurrent behavior in the Pacific Ocean as defined by the Niño 3.4 index.

Our bootstrapping method found a strong signal of milky seas being more likely following a positive IOD for the majority of pairings of representative spatiotemporal scale for milky sea events. In the western equatorial Indian Ocean the positive IOD suppresses biological activity (as seen through chlorophyll-a) due to downwelling driving a deepening of the thermocline and a decrease in nutrient availability in the ocean's photic zone, this behavior can continue through boreal spring of the following year (Shi & Wang 2021). Understanding how the post-positive IOD

state of the NW Indian Ocean becomes more conducive to milky seas following an extended decrease in nutrient availability is a topic of future research.

Unlike the SW monsoon, for boreal summer milky sea events in this region the frequency of boreal winter milky sea events in the NW Indian Ocean is increased during a typical NE monsoon season and is decreased during a weak NE monsoon. This suggests that the NE monsoon plays a critical role in driving the formation of a milky sea beyond setting the stage for other phenomena. Understanding what differences between the NE and SW monsoons lead to this distinction as well as if this plays into the N/S seasonality of where milky seas occur in the NW Indian Ocean (Figure 2.6) is a topic of future research.

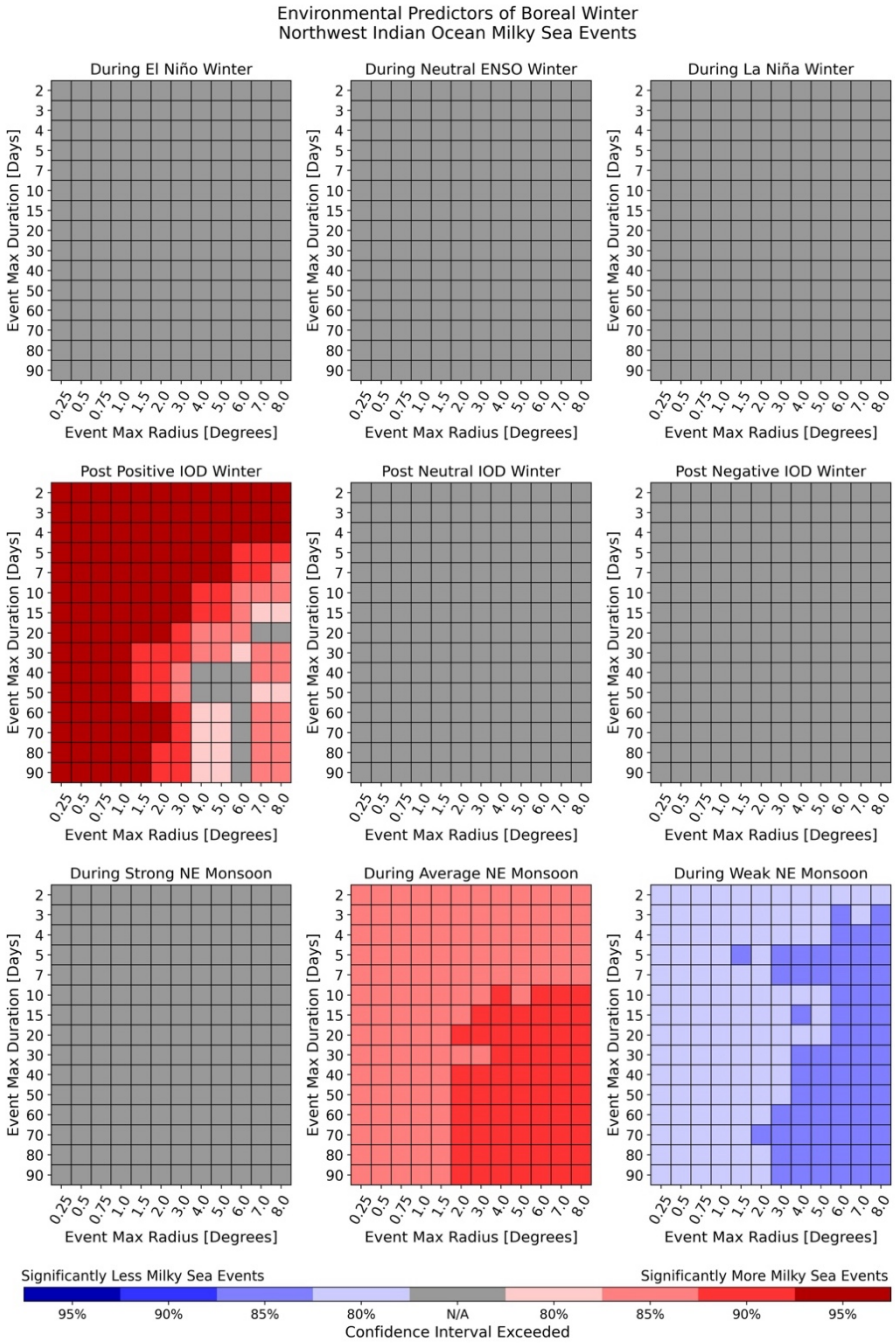


Figure 2.12: Results of the bootstrapping experiment for the Northwestern Indian Ocean during Boreal Winter (2.4.2). Milky seas are more likely following a positive IOD event and during a typical NE monsoon, a weak NE monsoon was found to decrease the likelihood of seeing a milky sea.

2.4.3 Bootstrapping Results: Banda Sea During Austral Winter

The frequency of austral winter milky sea events in the Banda Sea was examined based on their relationship to the IOD, ENSO, and the IAM. Similar to the NW Indian Ocean, in the Banda Sea region these phenomena impact the coupled atmosphere ocean state on interannual to seasonal time scales (Wirasatriya et al. 2021). The IAM is the dominant monsoonal mode over Australia and Indonesia. The southeasterly (SE) phase of the IAM occurs during Austral Winter wherein low-level southeasterly winds blow over northern Australia and the Maritime Continent (Kajikawa et al. 2010).

The IAM was described using the AUSMI index (Kajikawa et al. 2010). Like with the NE monsoon, the AUSMI index is negative for the austral winter phase of the IAM (the SE Monsoon phase). The top 20% of years of AUSMI Index (those where the mean value for the Austral Winter time period were most negative) were used to define strong SE monsoon years, the bottom 20% of years of AUSMI index (those where the mean value for the Austral Winter time period were most positive) were used to define the weak SE monsoon years, and the remaining 60% were used to define an average SE monsoon year.

The Banda Sea region saw very little signal or relationship to the various air-sea coupled phenomena considered, save for a potential weak relationship between increased milky sea frequency and a developing/on-going positive IOD event (Figure 2.13). This region has the least amount of data among the four spatiotemporal groups, and it is unknown if this lack of signal is due to data quality, milky seas being inherently less predictable in this region, or if milky seas in this region are controlled by factors that were not examined here.

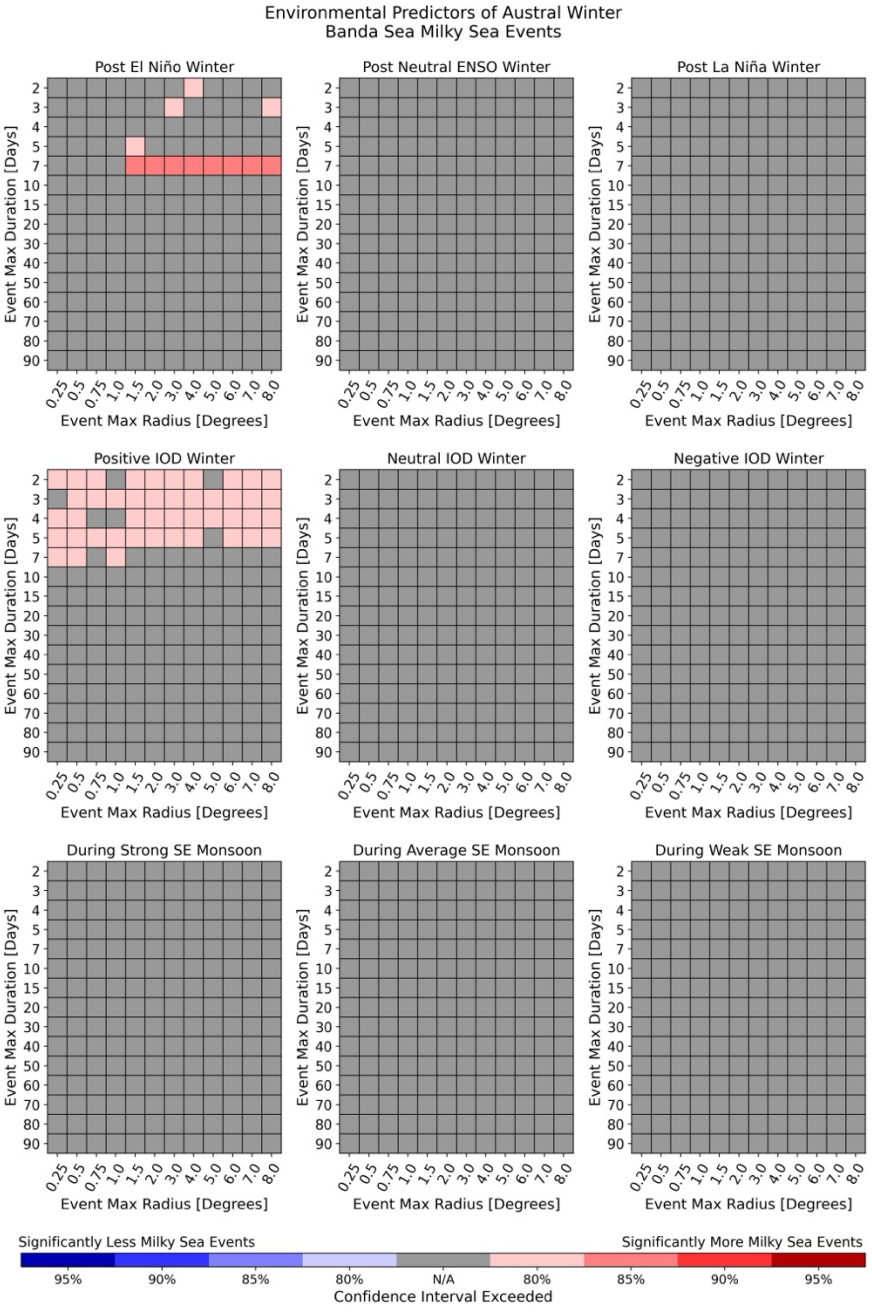


Figure 2.13: Results of the bootstrapping experiment for the Banda Sea region during Austral Winter (Section 2.4.3). There is little signal in this region aside from a possible increase in milky sea events during a positive IOD event.

2.4.4 Bootstrapping Results: South Java/Christmas Island During Austral Winter

For the area south of Java, Indonesia the same phenomena as for the Banda Sea were examined (Figure 2.14). This region saw potentially the strongest predictability among the four regions based on the methodology at hand.

The positive phase of the IOD is associated with a strong increase in the frequency of observed milky sea events and the negative phase of the IOD is associated with a strong decrease in the frequency of observed milky sea events at all pairings of *R* and *N*. This connection to the IOD for this region was surmised by M21 and agrees with recent work showing the connection between the IOD and biological activity in this region (Shi & Wang 2021, 2022, 2024).

There is also an uptick in milky sea events in this region during a typical SE monsoon year for all values of *R* and *N*. The SE monsoon drives upwelling along the southern coast of Java and modifies the behavior of the Indonesian Throughflow (ITF; Wirasatriya et al. 2020). Understanding how the IOD and the SE monsoon coincide and interact to foster milky seas is examined in Chapter 3.

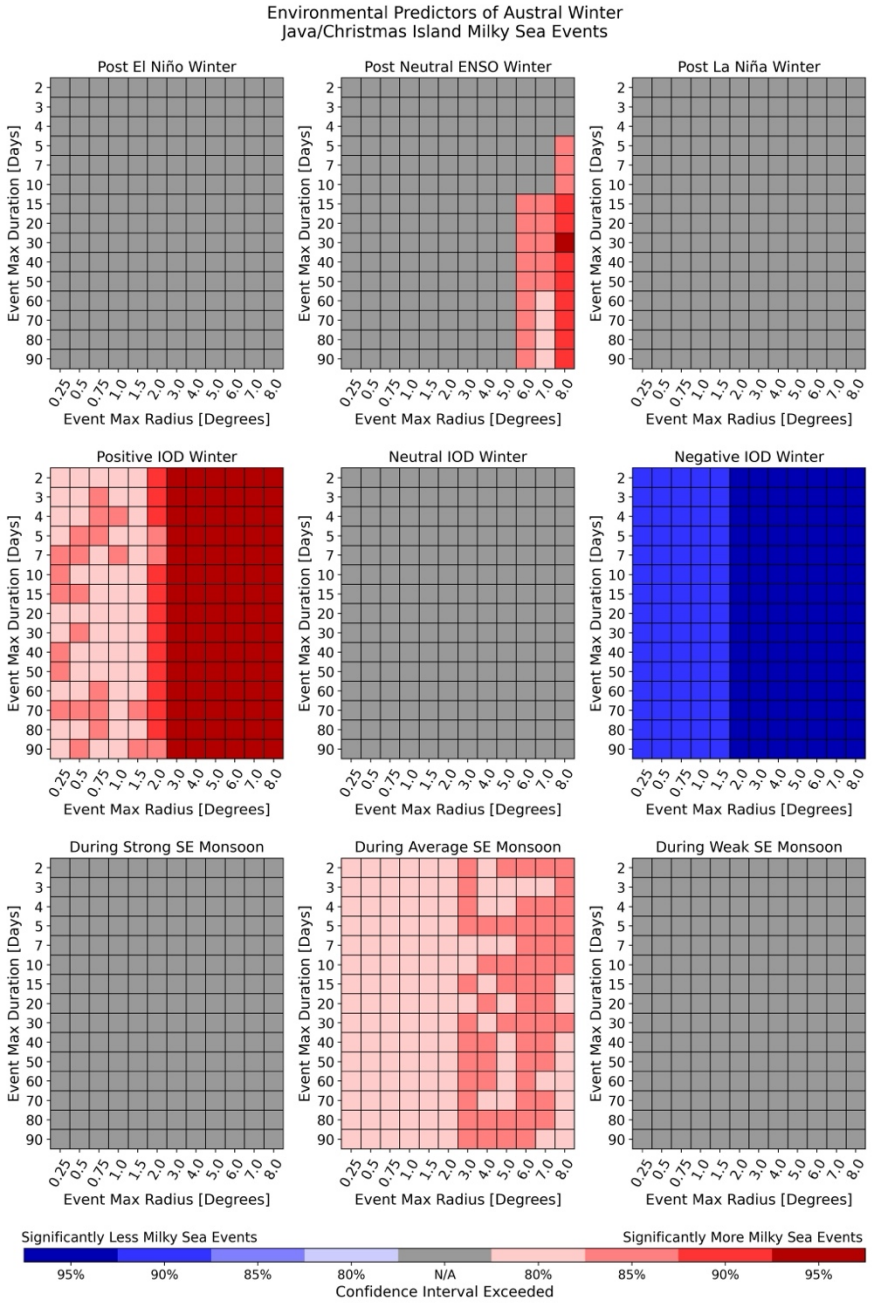


Figure 2.14: Results of the bootstrapping experiment for the region south of Java, Indonesia (Section 2.4.4). Milky sea events in this region are strongly controlled by the phase of the IOD, with the typical SE monsoon event also being associated with more milky sea events than expected. There is a possible increase in milky sea events following a neutral ENSO event.

2.5 Discussion and Conclusions

The largest database of milky sea sightings to date has been developed. This database recaptures and builds upon the results of previous works to reveal the global spatial and intra-year distributions of milky seas. This database serves as a launching point for future research to study this coupled earth system phenomena. The role of milky seas within the ecosystems in which they occur, how they fit within regional and global carbon cycles, and how their frequency may change in a future climate are outstanding science questions. Leveraging this database to address these questions may provide new insight into the biogeochemical state and evolution of the Indian Ocean.

Using this database, connections between milky seas and large-scale atmospheric/oceanic phenomena (e.g., the Indian Monson, ENSO and the IOD) were examined using a bootstrapping methodology. This examination revealed potential sources of large-scale predictability for milky seas which are leveraged in Chapter 3 toward prediction of milky seas near Java, Indonesia. Understanding how other phenomena not examined here such as oceanic eddies/meanders, the Madden-Julian Oscillation, and localized wind-driven upwelling impact milky sea occurrence is a topic of future research. Understanding how more regional dynamics interact with large-scale atmosphere/ocean dynamics to form and consequently sustain milky seas is critical to eventually being able to study a milky sea in-situ.

We note that this database was primarily limited to the English record. Recent advancements in machine learning have greatly expanded the ability to parse and extract information from documents in a variety of languages. There is likely a wealth of additional information on milky seas in records from/around the Indian Ocean that may greatly expand the

ability to study milky seas and elucidate potentially unknown connections to the coupled earth system.

Chapter 3
Exploring the Predictability of Milky Seas Near Java, Indonesia

3.1 Introduction

Due to their remote, ephemeral, and rare nature the in-situ sampling of a milky sea by a scientific vessel has been limited to a single chance encounter in 1985 (Lapota et al. 1988). The inability to know exactly when and where a milky sea will occur is one of the main barriers to sampling a milky sea in-situ and unraveling the biogeochemical properties of a milky sea.

The area south of Java, Indonesia (hereafter referred to as Java) is one of the main hotspots for milky sea eyewitness accounts historically (Figure 2.4; Figure 3.1 Panels A & D; Turner 1965; Herring & Watson 1993; Hudson & Miller 2025), and in the satellite record (Miller et al. 2021). Unlike the northwest Indian Ocean where milky seas tend to occur annually or biannually milky seas near Java tend to occur on a more irregular basis (Hudson & Miller 2025). This irregular basis hints that the conditions which bring about a milky sea near Java may be tied to interannual variability within the coupled earth system.

HW93 and Hudson & Miller 2025 (the publication corresponding to Chapter 2 and hereafter HM25) examined milky sea sightings near Java and found a strong seasonality focused around the month of August (Figure 3.1 Panel C) the peak of the austral winter or SE phase of the Indo-Australian Monsoon (IAM; Kajikawa et al. 2010). The satellite observations of M21 also found that a milky sea event during July-September 2019 was coincident with a positive Indian Ocean Dipole (IOD) event. This anecdotal association of milky seas with large scale coupled atmosphere ocean phenomena was examined statistically in Chapter 2 where we showed that the IAM and the IOD both influence the frequency of milky sea events near Java.

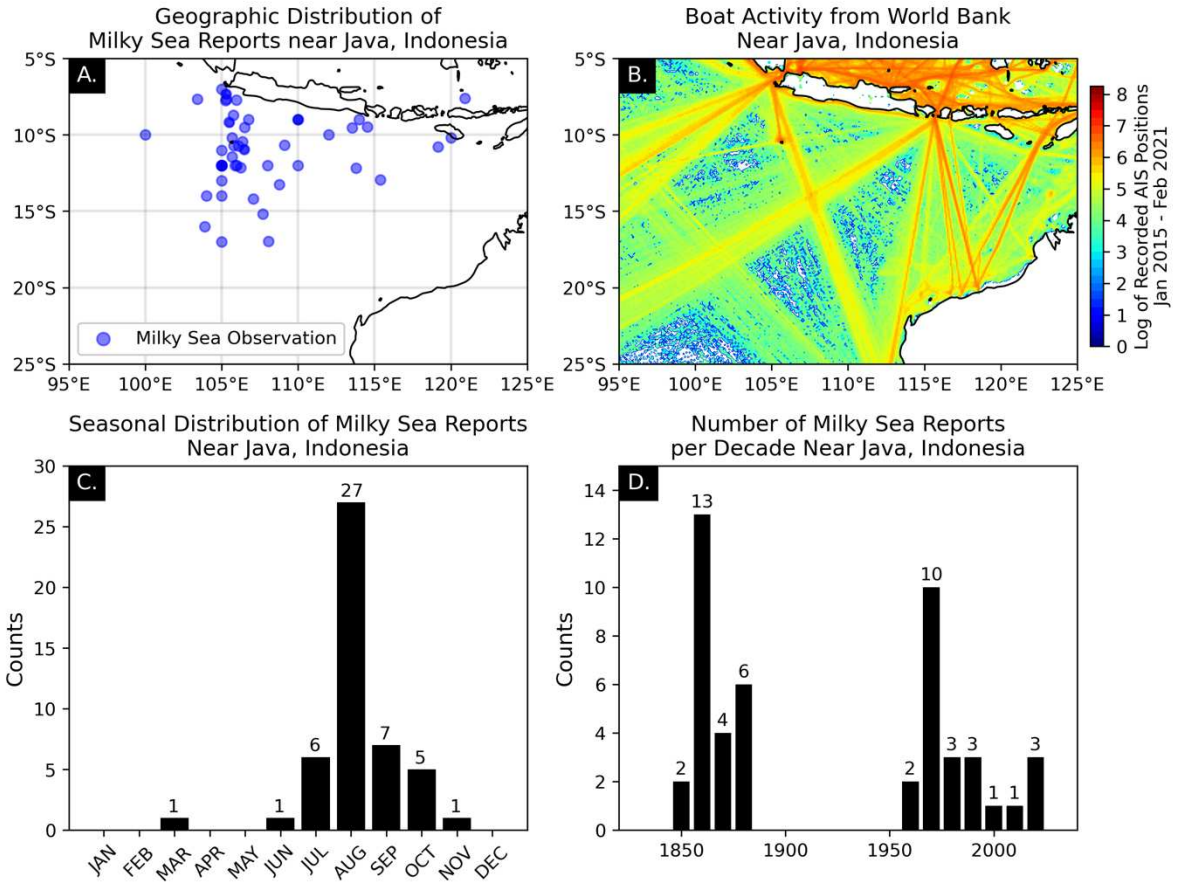


Figure 3.1: Geographic distribution of milky sea observations near Java (Panel A.), IMF WSTMS ship track data highlighting the impact of shipping lanes on milky sea data (Panel B.), the seasonal distribution of milky sea reports near Java (Panel C.), and the number of milky sea reports per decade near Java highlighting the data from the 1880s to the 1960s (Panel D.). The data in Panels A, C, and D comes from the database presented in Chapter 2. The data in Panel B was described in Chapter 2.

The austral winter phase of the IAM influences the sea state near Java via low-level southeasterly winds which drive coastal upwelling along Java (Figure 2.3; Wirasatriya et al. 2020). The positive (negative) phase of the IOD similarly drives coastal upwelling (downwelling) near Java via basin wide changes in circulation (Saji et al. 1999; Shi & Wang 2019; Shi & Wang 2022; Shi & Wang 2024).

The natural flask hypothesis of M21 theorizes that milky seas form within isolated or semi-isolated parcels of water bounded by dynamical or density driven horizontal barriers which contain

ideal biogeochemical conditions for the bacteria behind milky seas. This hypothesis is based on the wealth of satellite observations reviewed in M21 which saw many milky sea events forming within quiescent, relatively cool, nutrient rich waters, bounded horizontally by fast moving meanders or eddies in the local currents.

The IAM, IOD, and the state of the ocean near Java are also correlated to and impacted by the El Niño southern Oscillation (ENSO; Saji et al. 1999; Kajikawa et al. 2010). Despite the bootstrapping experiment in Chapter 2 finding no direct statistical connection between ENSO and the frequency of milky sea events near Java it is known that ENSO directly impacts the Indonesian Throughflow (ITF; Liu et al. 2015; Chen et al. 2016; Wirasatriya et al. 2020). The ITF may play a critical role in providing the dynamical barriers for the natural flask hypothesis of M21.

The IOD is also correlated with ENSO (Saji et al. 1999) with a positive (negative) IOD event often occurring alongside a positive (negative) ENSO event. The impacts of the IOD on the ITF can compete with and offset the impacts of ENSO (Liu et al. 2015; Chen et al. 2016; Wirasatriya et al. 2020; Wen et al. 2023). This complex interrelationship between the IOD, IAM, and ENSO on interannual timescales may play a role in determining when milky seas happen near Java.

The fact that Java lies at the nexus of many large-scale sources of coupled atmosphere ocean variability means there is potential to make use of these phenomena and their superposition of states to predict when milky seas will happen. In this chapter we examine the feasibility of constructing a predictive model for milky seas near Java as a first step towards being able to sample this phenomena in-situ.

In Section 3.2 the data used to construct the predictive model is presented. In Section 3.3 we construct and verify the accuracy of a hindcast model for milky seas near Java, in Section 3.4

the hindcast model is used to examine the natural flask hypothesis for milky seas. The feasibility of extending our model from hindcasting to forecasting is examined in Section 3.5, and finally in Section 3.6 we consider the constructed model's implications for a future expedition towards a milky sea and future work on predicting milky seas.

3.2 Data Sources

The database of Chapter 2 (also described in HM25) was used to ascertain the timing of milky sea events near Java. This dataset was limited to just those observations which occurred in the region of interest near Java [100^{o}-120^{o}E, 20^{o}-5^{o}S]. This subset originally yielded a total of 47 unique observations and 1 satellite observed event from 1854 to 2019 (this eventually became 2 satellite observed events, see section 3.3). The intra- and inter-year variations in milky sea sightings can be seen in Figure 3.1 Panels C and D respectively. The DNB imagery in this chapter was taken from NASA Worldview application which is part of the NASA ESDIS.

The intra-year variation in milky sea sightings near Java has a notable gap between 1880 and 1961 where no milky seas were observed. While this is likely the impact of data quality and a paucity of surviving eyewitness accounts it is also possible that this is due to some unknown multi-decadal change in the coupled earth system. To account for this gap in the data, both the dataset and subsequently the predictive model described in Section 3.3 were limited to the time period of 1960-2023.

The eyewitness accounts in this region are biased towards the timing and location of major shipping lanes (Figure 3.1 Panel B; Section 2.2.4.1). The ship traffic density data for this chapter

was taken from the IMF's WSTMS (Cerdeiro et al. 2020) as in Chapter 2. The WSTMS data in this chapter was coarsened from the original ~0.005° x ~0.005° grid to a ~0.1° x ~0.1° grid.

To assess the state of the IOD the Dipole Mode Index (DMI; Saji et al. 1999) was used. The DMI is calculated via the difference of spatial means for anomalous SST between the western equatorial Indian Ocean [50°-70°E, 10°N-10°S] and the southeastern equatorial Indian Ocean [90°-110°E, 10°S-0°N/S]. Climate change corrected monthly Hadley Center Ice and Sea Surface Temperature (HadISST) SST anomalies were used to calculate the DMI (Rayner et al. 2003).

The Niño 3.4 Index (N34; Trenberth 1997), taken from the NOAA PSL website was used to evaluate the state of ENSO for our period of interest. The N34 index is based on a 5-month running mean of HadISST1 (Rayner et al. 2003) SST anomalies in the equatorial Pacific [170°-120°W, 5°S-5°N].

The strength of the IAM was assessed using AUSMI (Kajikawa et al. 2010). AUSMI was calculated using monthly-mean 850 hPa zonal winds over 110°-130°E, 15°-5°S from the ERA5 for 1950-2023 (Hersbach et al. 2020). Monthly mean 10-meter zonal and meridional winds for the area near Java [100°-120°E, 20°-5°S] were also downloaded from ERA5 (Hersbach et al. 2020).

The state of the ocean was evaluated using Copernicus Marine Environmental Service (CMEMS) upper 20-meter zonal and meridional currents, mixed layer depth, and salinity from the Global Ocean Physics Reanalysis (doi: 10.46780/moi-0021) for 1993-2023 over 100°-120°E, 20°-5°S. Chlorophyll-a measurements from VIIRS (Wang & Son 2016) were used to relate milky sea activity to primary productivity and were taken from the NASA Ocean Color group via OB.DAAC for 2012-2023.

The seasonal cycle of milky seas varies between regions (Chapter 2; Hudson & Miller 2025). To capture the seasonal cycle of milky seas near Java the number of observations per month

(Figure 3.1 Panel C) was converted into a normalized time series spanning the range [0, 1]. This normalization was done by dividing the number of monthly observations by the maxima in August.

The DMI, AUSMI, and N34 indices have varying units. To account for this difference all three index time series were standardized by subtracting their respective means and dividing by their respective standard deviations. This places all 3 indices on similar scales, renders them unitless, and enables them to more easily be combined together in the predictive model.

Canadian Seasonal to Interannual Prediction System version 2 (CanSIPSv2; Lin et al. 2020) monthly forecasts of global SST and 850 hPa zonal winds at 1.5 to 10.5 month lead times over the December 2016 to November 2021 time period produced by Environment Canada were taken from Phase II of the North American Multi-Model Ensemble (NMME; Kirtman et al. 2014) project. Monthly climatology (version: sc9120) of SST from CanSIPSv2 at various forecast lead times of 1.5 to 10.5 months were taken from CanSIPSv2 data produced for Phase II of the NMME project. NMME Phase II data [https://www.earthsystemgrid.org/search.html?Project=NMME] was taken from the Columbia Climate School's International Research Institute for Climate and Society (IRI) and the Lamont-Doherty Earth Observatory's (LDEO) Climate Data Library. The CanSIPSv2 data was produced and made available by Environment Canada through the IRI/LDEO Climate Data Library.

3.3 Verification and Construction of the Predictive Model

The ocean state near Java is, on inter-annual and intra-annual timescales, modulated by phenomena such as the IOD, ENSO, and the IAM. These phenomena modulate both upwelling/SSTs in this region (Figure 3.2) as well as currents and prevailing wind patterns. To construct our hindcast model, indices for the IOD (DMI), ENSO (N34), and IAM (AUSMI) were

first limited to the time period of 1960-2023. To account for the varying units between these indices the indices were standardized by subtracting their respective means and dividing by their respective standard deviations. The seasonal cycle (SC) of milky sea events near Java (Figure 3.1 Panel C) was converted into a normalized time series on the range [0, 1] by dividing the number of monthly sightings by the maxima in August. These data were then combined via a weighted linear sum (Equation 3.1) to create a predictive monthly hindcast index for milky seas south of Java (South Java Milky Sea Index or "SJMI").

$$SJMI = W_1 * DMI + W_2 * AUSMI + W_3 * N34 + W_4 * SC \quad 3.1$$

The predictors for the SJMI model are not wholly independent from one another. For example, the IOD and ENSO and strongly correlated phenomena (Saji et al. 1999). The correlation between the DMI and Niño 3.4 indices used here is 0.33, and the correlation of both the DMI and Niño 3.4 indices with AUSMI is -0.11. This collinearity between predictors can potentially impact the interpretability of the model if the weights are determined via regression.

The SJMI is designed such that a month with a higher value of SJMI is more likely to have a milky sea than one with a lower value. To make use of the SJMI, however, the weights (W_1, W_2, W_3, and W_4) must first be determined. The limited number of milky sea events in this region over the 1960-2023 time period, the associated uncertainty over whether months/years with no reported milky sea actually contain a milky sea, and the potential collinearity between predictors precludes a regression based approach.

In order to determine the weights for the SJMI a coarse grid of parameters spanning the range [-3, 3] for every W_N with a resolution of 0.1 was searched. This parameters space was

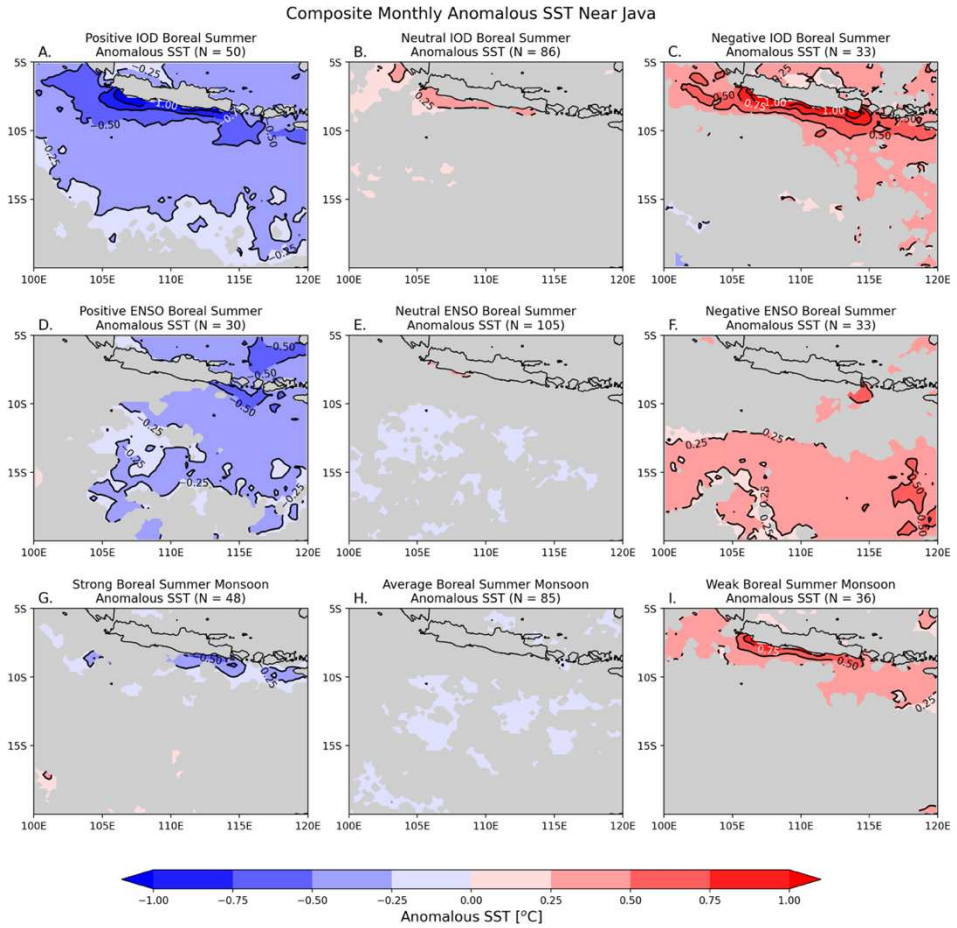


Figure 3.3: Composite monthly anomalies in SST by phase of the IOD (Top Row), ENSO phase (middle row), and strength of the Indo-Australian Monsoon (bottom row). All SSTs anomalies shown are significant at the 95% level using the scaled means technique of Brown and Hall (1999).

searched for the combination of weights which maximizes how many of the top 10 time periods identified by this model correspond to time periods with known milky sea events. A top 10 period was defined as the months with the 10 highest values of SJMI over the 1960-2023 time period. This is analogous to a hyperparameter grid search in machine learning.

This methodology often predicts multiple months from the same austral winter as being part of the top 10 time periods. To prevent a single year from dominating the top 10 time periods and skewing the results of the parameter fitting, months that were either temporally adjacent (e.g. June and July) or months separated by only one other month (e.g. June and August) were combined into a single continuous time period. This process condenses the top 10 time periods. Every time

a set of months were combined into a single top 10 time period the 11th ranked time period would rise into the top 10 time periods becoming the 10th ranked time period. If the new 10th ranked time period could be combined with another time period the two time periods would be combined and another time period would rise into the top 10 time periods. This process was repeated until the top 10 time periods identified by a particular parameter set became stable.

The top 10 time periods were then examined for whether or not there was a known milky sea event during those time periods. The parameter set being examined would be given a point for each of the top 10 time periods that contained a known milky sea event, this results in a minimum score of 0 and a maximum score of 10. The higher the score the better the parameter set.

Interactions between earth system phenomena can be expressed at multi-month lags. To account for this, we conducted our parameter search for a suite of potential models where the DMI, AUSMI, and N34 indices were allowed to lead/lag one another up to 1-2 months. This comparison was done for every combination of leads/lags for all three indices.

Table 3.1: Results of parameter fitting for various lead-lag relationships between input timeseries and the SJMI. Only those lead/lag combinations that correctly placed known milky sea observations within the top 10 time periods for SJMI are shown.

DMI Lead (Months)	AUSMI Lead (Months)	N34 Lead (Months)	Number Top 10 Verified
0	2	0	7
0	0	0	6
0	0	1	6
0	0	2	6
0	1	0	6
0	1	1	6
0	1	2	6
0	2	1	6
0	2	2	6
1	0	0	6
1	0	1	6

Table 3.1 shows the results of the parameter search described above for various combinations of lead/lag relationship between the predictors. Only those models which successfully put milky sea events within their top 10 time periods are shown.

The best performing model has AUSMI lead DMI and N34 indices by two months (table 3.1, Figures 3.3 Panels A and B). For this model there were 21 parameters sets which lead to 7 verified milky sea events within their top 10 time periods (Table 3.2). A more refined parameter set or search method may be able to produce a better result, for this analysis however we elected to forego such refinements to prevent potentially overfitting to a dataset which likely under-samples the true spatiotemporal distribution of milky seas near Java.

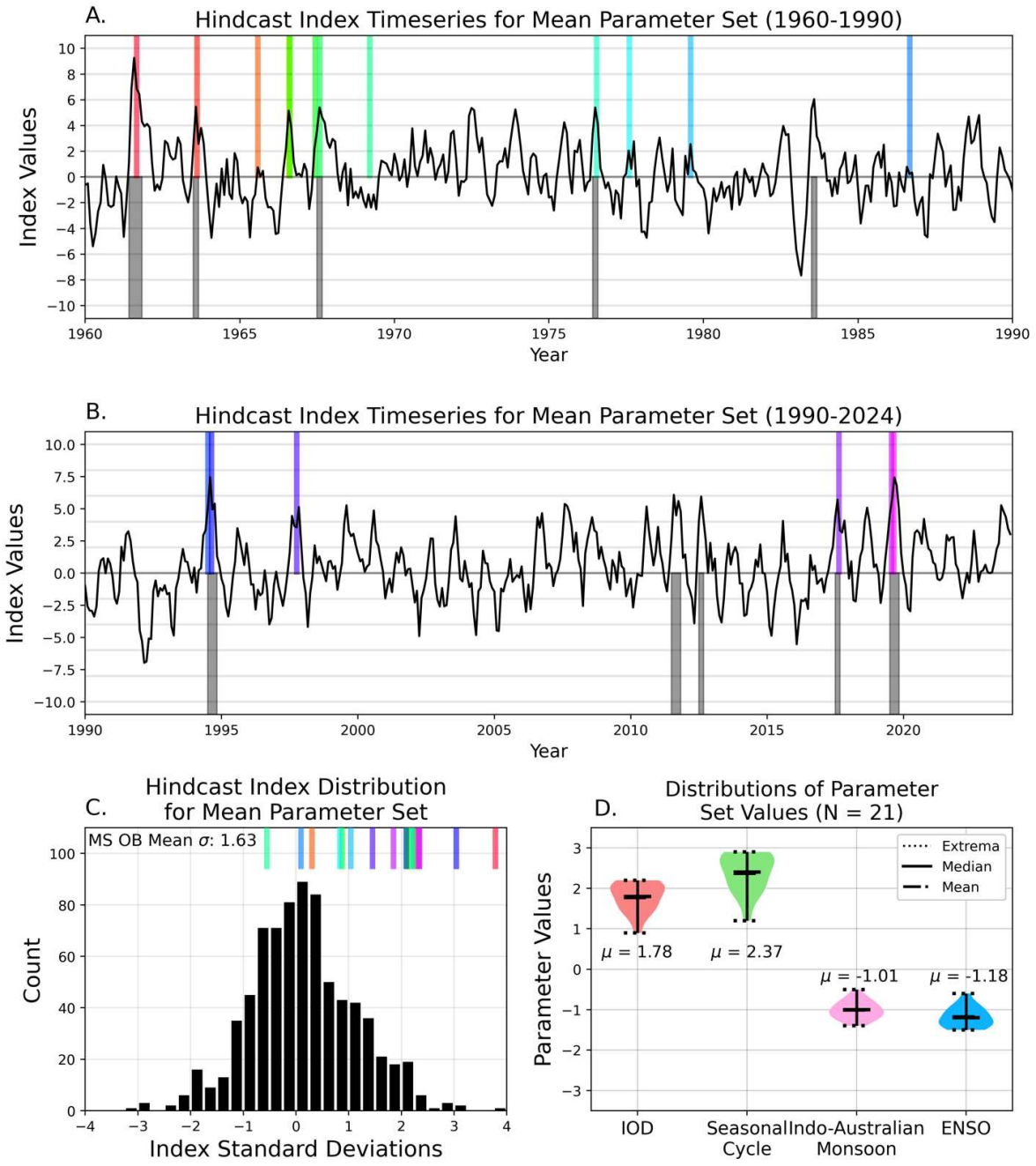


Figure 3.3: SJMI hindcast time series as generated by the mean of the 21 parameter sets for 1960-1990 (Panel A.), and 1990-2024 (Panel B.). The colored bars in the top half of panels A. and B. correspond to known milky sea observations near Java, the grey bars on the bottom half of panels A. and B. correspond to the top 10 time periods in Table 2. The distribution of values for the SJMI is plotted in Panel C. and the colored bars along the top of Panel C. correspond to the similar colored bars in Panels A. and B. A violin plot of the distributions of the parameters is shown in Panel D.

Table 3 2. The top 10 time periods for the SJMI as identified using the mean of the 21 parameter sets that produced the best predictive ability for milky seas near Java, Indonesia. 7 out of 10 time periods correspond to a milky sea event that was either previously known or identified due to the SJMI model.

Time Period Start	Time Period End	Milky Sea Observed
July 1961	October 1961	Yes
August 1994	October 1994	Yes
August 2019	October 2019	Yes
August 2011	October 2011	No
August 1983	August 1983	No
August 2012	August 2012	No
August 2017	August 2017	Yes
August 1963	August 1963	Yes
August 1967	August 1967	Yes
July 1976	July 1976	Yes

The 21 parameter sets (Table 3.3; Figure 3.3 Panel D) have a mean DMI weight (W_1) of 1.78, a mean AUSMI weight (W_2) of -1.01, a mean N34 weight (W_3) of -1.18, and a seasonal cycle weight (W_4) of 2.37. All four weights are unitless. The SJMI produced utilizing the means of these weights is approximately normally distributed with the mean SJMI value for a milky sea event lying 1.63 standard deviations above the mean (approximately the 95th percentile of the distribution).

Table 3.3 the individual weights for the 21 best performing parameter sets. These parameter sets are visualized in Figure 3.3

DMI (W_1); μ = 1.78	AUSMI (W_2); μ = -1.01	N34 (W_3); μ = -1.18	SC (W_4); μ = 2.37
0.9	0.5	0.6	1.2
1.2	0.7	0.8	1.6
1.4	0.8	0.9	1.9
1.4	0.7	0.9	1.9
1.5	0.9	1.0	2.0
1.5	0.8	1.0	2.0
1.6	0.9	1.1	2.1
1.7	1.0	1.1	2.3
1.7	0.9	1.1	2.3
1.8	1.1	1.2	2.4
1.8	1.0	1.2	2.4
1.9	1.1	1.3	2.5
2.0	1.2	1.3	2.7
2.0	1.1	1.3	2.7
2.0	1.0	1.3	2.7
2.1	1.3	1.4	2.8
2.1	1.2	1.4	2.8
2.1	1.1	1.4	2.8
2.2	1.4	1.5	2.9
2.2	1.3	1.5	2.9
2.2	1.2	1.5	2.9

The narrow ranges of the weights for the 21 parameter sets (Figure3.3 Panel D; Table 3.3) also show that the model is keying in on a particular region within the constructed parameter space. If the collinearity of predictors for the SJMI model were hindering the model's interpretability the 21 parameter sets would span a larger volume of the parameter space. This allows for the sign and relative value of these weights to be used to verify the SJMI model is making predictions based on physical relationships versus random chance.

A priori we expected W_1 (DMI) to be positive which matches the results here. This expectation is derived from the results in Chapter 2 of this dissertation which showed that a positive IOD event (noted by a positive value of DMI) increases the likelihood of seeing a milky sea event near Java and that a negative IOD event (corresponding to negative values of DMI) decreases the likelihood of seeing a milky sea event near Java.

AUSMI is constructed in such a way that it is negative during the austral winter phase of the IAM (Kajikawa et al. 2010). This means the more negative a value AUSMI attains the stronger the IAM during austral winter. From Chapter 2 we expect the austral winter IAM to positively contribute to the likelihood of a milky sea. This is why, a priori we expected the weight for AUSMI to be negative which matches the results of the parameter fitting experiment.

The seasonal cycle of milky seas is constructed from the observations presented in Chapter 2. By construction we expect this to positively contribute towards the likelihood of a milky sea near Java. This expectation is validated by the positive weight applied to the seasonal cycle term in the SJMI.

The contribution of ENSO towards the SJMI was unknown based on the results presented in Chapter 2. An a posteriori explanation for why the ENSO contribution is negative is based on the ITF. The impacts of ENSO on the ITF are opposed and, during the peak time period for milky

seas, largely overwritten by the impacts of the IOD (Liu et al. 2015; Chen et al. 2016; Wirasatriya et al. 2020; Wen et al. 2023). For example the positive (negative) phase of the IOD, which coincides with the positive (negative) phase of ENSO, strengthens (weakens) the ITF locally, the positive phase of ENSO however weakens (strengthens) the ITF (Liu et al. 2015). This competition between the IOD and ENSO, two strongly correlated phenomena (Saji et al. 1999), likely explains both the sign and scale of the weight for ENSO's contribution to the SJMI. The negative weight reflects the opposition of the IOD's positive influence on milky sea likelihood and the smaller absolute value reflects how the DMI's influence outstrips that of ENSO. This may also explain why there was no signal from ENSO in Chapter 2.

During the development of the SJMI hindcast model multiple versions of the model routinely highlighted the August 2017 time period as a top 10 time period for milky seas near Java. Prior to the development of these models there was no known milky sea event during this time period. Based on the results of these hindcast models the VIIRS DNB imagery for this time period was reexamined utilizing the methodology presented in M21. Doing so resulted in the identification of a previously unknown milky sea event from August 20th to August 26th 2017 (Figure 3.4). This represents the first time a milky sea has been discovered via a predictive method.

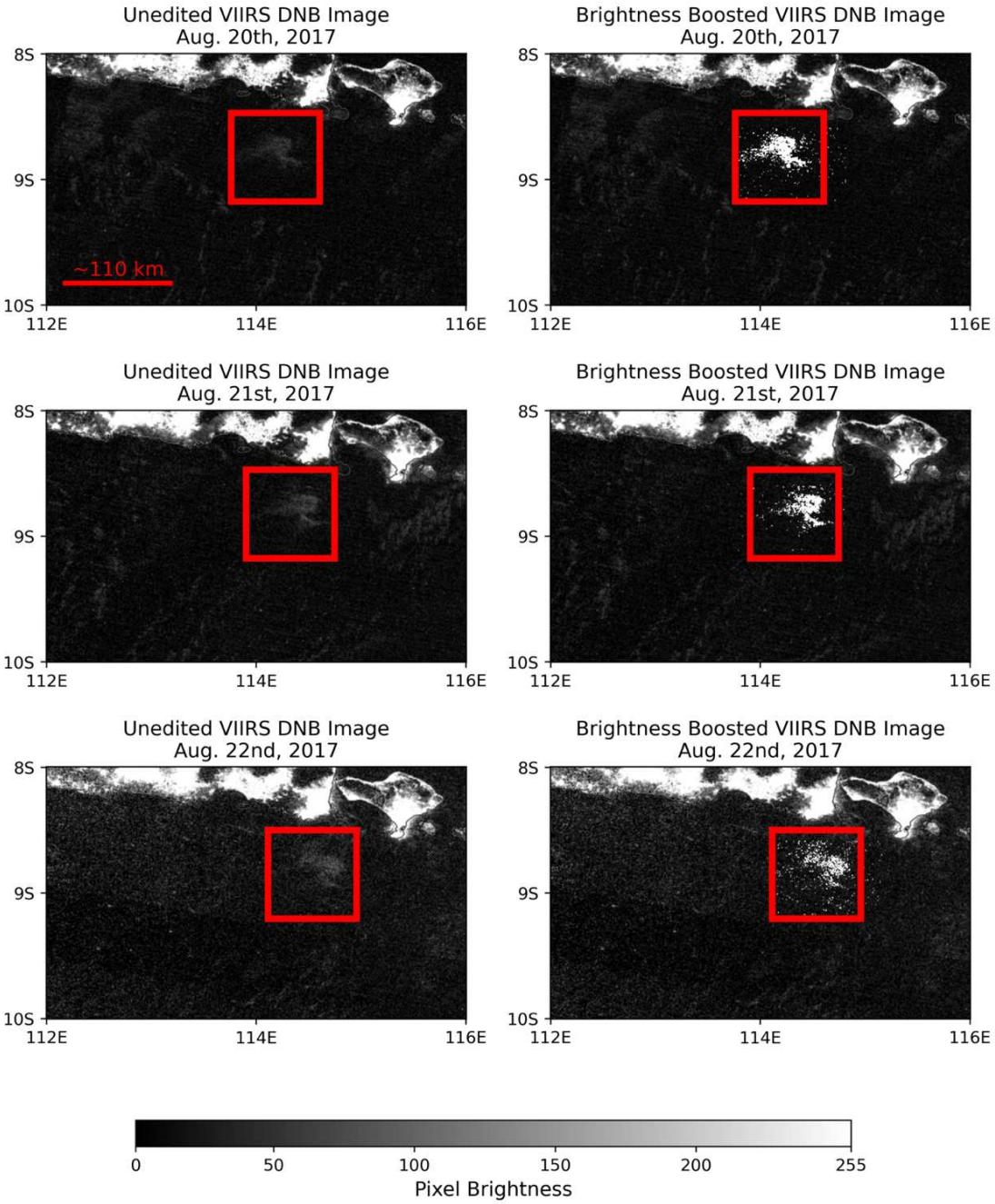


Figure 3.4: Unedited VIIRS DNB Near-Constant (NCC) imagery taken from NASA Worldview of the milky sea event near Java in August 2017 identified due to the SJMI model (left column), and the same imagery with pixels within the red square outlining the milky sea event modified to better highlight the milky sea (right column)

3.4 Model Insights

The natural flask hypothesis of milky sea formation (Miller et al. 2021) theorizes that milky seas form within isolated or semi-isolated parcels of water with ideal nutritional, chemical, substrate, and physical conditions for milky seas to form. These parcels are isolated by dynamical features in the ocean or gradients in density (Figure 3.5) which allow them to be sustained for the extended periods of time needed to match the observational characteristics of a milky sea. Understanding if the SJMI is keying in on the conditions that are believed to promote the formation of natural flasks is critical insights to both the validity of the SJMI model as well as the natural flask hypothesis.

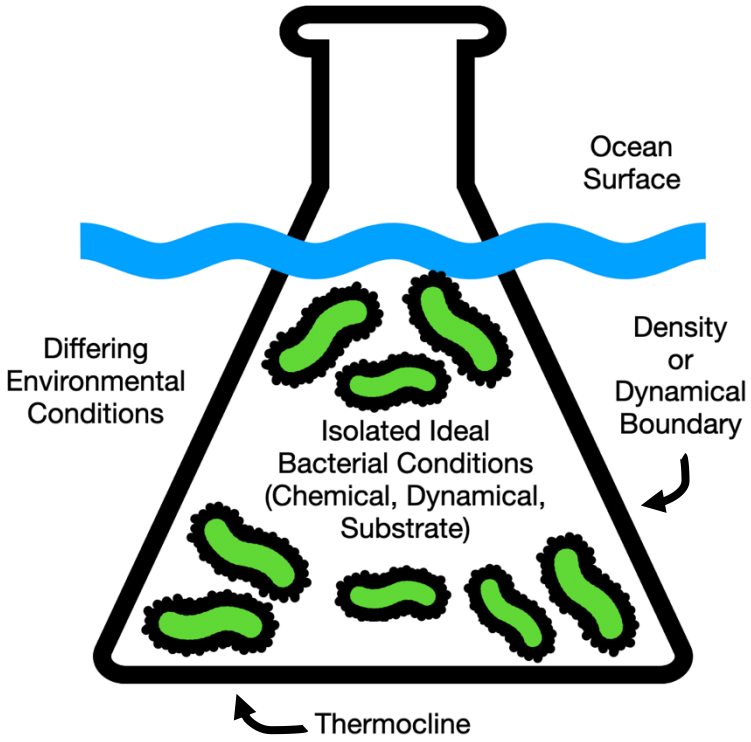


Figure 3.5 Cartoon schematic of the natural flask hypothesis of Miller et al. 2021 which theorizes that milky seas form within isolated or semi-isolated pockets of ideal bacterial conditions bounded laterally by density gradients or dynamical boundaries.

The SJMI hindcast in Section 3.3 was correlated with OISSTv2 SST, VIIRS Chlorophyll-a (Chl-a), CMEMS salinity and ocean current velocity, as well as ERA5 10-meter winds (Figure 3.6). All correlations plotted in Figure 3.6 are significant at the 95% level using a modified p-value to account for the false discovery rate as detailed in Wilks (2016).

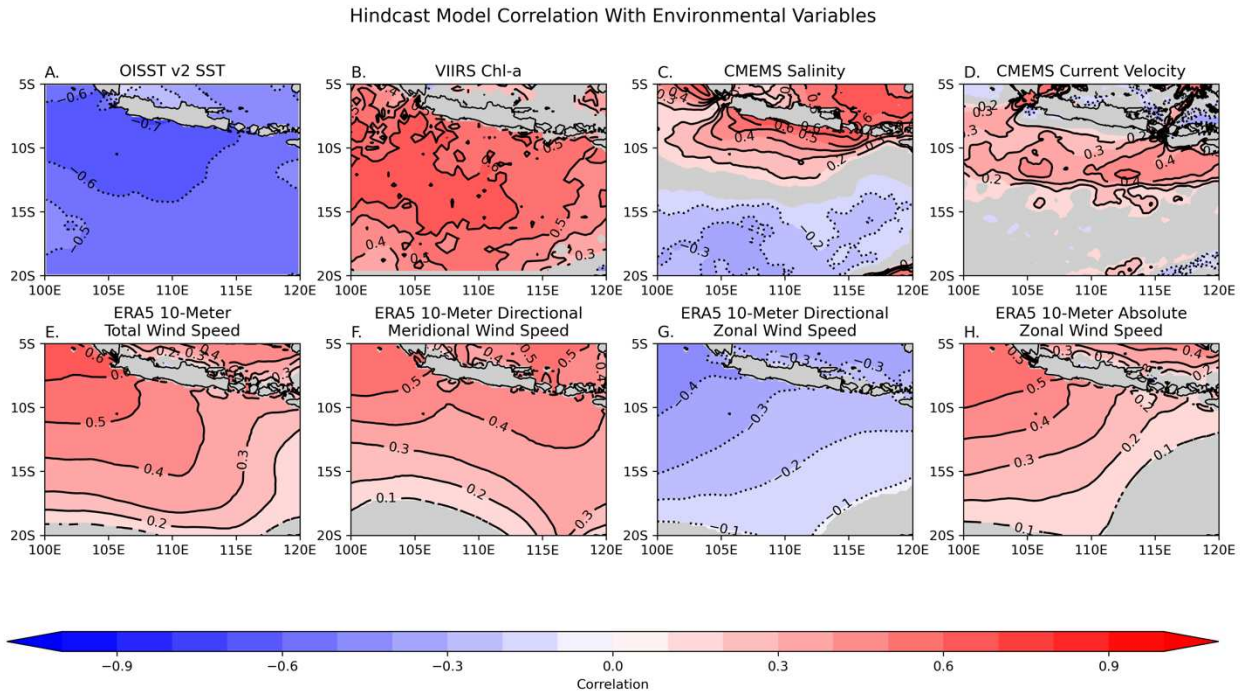


Figure 3.6: Correlation of the mean SJMI timeseries with oceanic and atmospheric variables. All plotted values are significant at the 95% level using a modified p-value to account for the false discovery rate as described in Wilks (2016).

The SJMI model is negatively correlated with SST across the region of interest with correlation decreasing with distance from the southern Java coast indicating that high values of SJMI are associated with coastal upwelling (Figure 3.6 Panel A). Given that the SJMI was constructed using phenomena known to modify upwelling in the region during austral winter this association serves as additional verification that the SJMI is highlighting the correct ocean state.

The SJMI is also correlated with enhanced values of Chl-a (Figure 3.6 Panel B) and increased salinity near Java (Figure 3.6 Panel C). Upwelled waters are nutrient rich and often denser due to both decreased temperature and increased salinity, this further aligns the SJMI with periods of upwelling and enhanced upper ocean biological activity near Java. The meridional

gradient in salinity could reflect a true meridional gradient in salinity for milky sea events serving as a potential wall for a natural flask type environment.

The relationship between the SJMI and the upper 20-meter mean ocean current velocity (Figure 3.6 Panel D) shows that the SJMI is associated with enhancement and widening of the ITF. Anecdotally the 2019 milky sea event near Java was bounded zonally by meanders in the ITF. This suggests that the meanders and eddies in the ITF may play a key role in milky sea formation. Investigating the connection between ITF behavior and the formation of natural flask environments may be critical to better understanding milky seas near Java.

The correlation of the SJMI with ERA5 10-meter winds (Figure 3.6 Panels E-H) shows that the SJMI is associated with enhanced winds blowing southeasterly. This wind configuration would drive upwelling through net Ekman transport to the southwest away from the southern shore of Java. Given that the IAM is baked into the SJMI via Equation 3.1 and blows southeasterly during the austral winter time period this may just be the signal of the IAM reasserting itself versus a particular enhancement or configuration of the IAM wind field leading up to milky sea events.

Overall, these results show that the SJMI is keying in on time periods with widespread upwelling, density gradients in the ocean, as well as sources of dynamical boundaries, key ingredients of the natural flask hypothesis. Further investigation of the biogeochemical conditions associated with milky seas and the natural flask hypothesis is carried out in Chapter 4.

3.5 Forecasting Milky Seas Near Java

The ability of the SJMI to find milky seas through hindcasting the highlight the conditions theorized by the natural flask hypothesis to create milky sea events shows that the SJMI or similar predictive metrics have the potential to forecast milky seas near Java. In this section we extend the

SJMI from a hindcasting tool to a forecasting tool to explore the feasibility of predicting milky seas and enabling the sampling of a future milky sea in-situ.

CanSIPSv2 data from Phase II of the North American Multi-Model Ensemble (NMME) includes monthly forecasts and climatology of SST and 850 hPa winds. This allows us to construct monthly forecasts of DMI, AUSMI, and N34 indices which are needed for a forecasting variant of the SJMI. The CanSIPSv2 model was chosen because, of the subset of data available for Phase II of the NMME, the CanSIPSv2 model was the only one with the data needed to produce an SJMI forecast.

Using the CanSIPSv2 data we constructed monthly forecast time series of DMI and N34 at forecast lead times of 3.5 months to 10.5 months and monthly forecast time series of AUSMI at forecast lead times of 1.5 months to 8.5 months. This matches the two-month lead of AUSMI relative to DMI and N34 which produced the best SJMI hindcast model in Section 3.3. The forecast initialization times for the CanSIPSv2 data spans December 2016 to November 2021. As the N34 index requires 5 months of data to be calculated forecasted SJMI begins with the April 2017 initialization. This allows us to determine if the SJMI would have been able to predict the 2017 and 2019 milky sea events near Java.

The constructed forecasts of DMI, AUSMI, and N34 were standardized by subtracting the respective means and dividing by the respective standard deviations as with the hindcast SJMI. These standardized indices and the normalized seasonal cycle data for milky seas were combined via Equation 3.1 using the mean parameter set of the best performing hindcast SJMI model (visualized in Figure 3.3 and listed in Table 3.3). This produces a forecasted SJMI spanning mid 2017 through 2022 at various forecast lead times (Figure 3.7). This was done for all 20 CanSIPSv2 ensemble members.

The forecasted ensemble of SJMI values correlated well with the hindcasted SJMI over the same time period, but the ensemble mean tends to underpredict extreme in SJMI which is where milky sea events lie (Figure 3.7). As milky seas are associated with local maxima in the SJMI we looked for a metric within the CanSIPSv2 ensemble data that would allow for successful prediction of both the 2017 and 2019 milky sea events near Java.

We found that time periods where the top 5 ensemble members (those with the largest monthly values of SJMI for a particular month) exceed a value of 4.3 (the mean observed hindcast SJMI value for milky sea events) would result in the CanSIPSv2 model successfully predicting the known milky events near Java with no false predictions at shortest forecast lead time (Figure 3.7 Panel A). This result shows that the SJMI, had it existed and been properly tuned for CanSIPSv2, could have potentially forecasted these milky sea events with a 1.5 month lead time.

As the forecast lead time is extended the underprediction of the hindcast SJMI lessens and the ensemble mean begins to better match the hindcast SJMI. This results in the chosen metric of the top 5 ensemble members exceeding an SJMI value of 4.3 producing false positives every austral winter starting at the 2.5 month/4.5 month lead time. Further fine tuning of a similar metric may produce a threshold which works well with CanSIPSv2 or a similar model at a wide variety of forecast lead times.

The ability of the SJMI to forecast milky sea events greatly expands the potential to accurately position resources to sample a milky sea in-situ. Future work will focus on refining the SJMI to smaller time scales and seeing if the SJMI or a similar predictive model can be paired with a dedicated Earth System Model forecast run to provide actionable forecast guidance on milky seas.

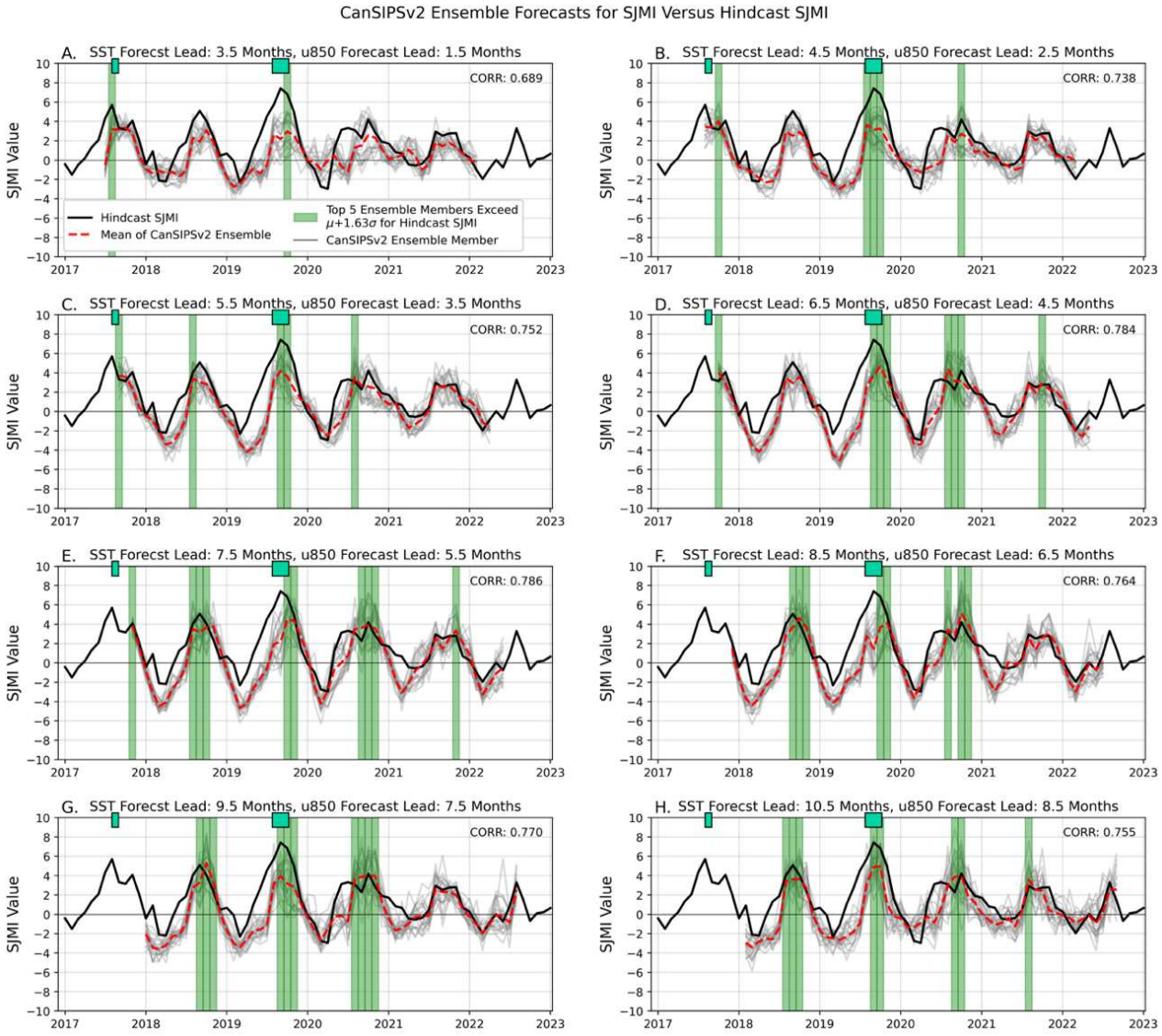


Figure 3.8: Forecasts of SJMI generated with the 20 CanSIPSv2 ensemble members using NMME Phase II data at various forecast lead times. The turquoise rectangles at the top of each plot represent time periods with an observed milky sea near Java. Time periods where the top 5 ensemble members are above an SJMI value of 4.3 (the mean SJMI value for observed milky sea events; highlighted in green) shows that the SJMI has potential as a forecasting tool to identify potential milky sea events within a year of their occurrence.

3.6 Discussion and Conclusions

Past research on bioluminescent milky seas emphasizes the importance of in-situ sampling of a milky sea to answer the myriad of biological and physical questions about the phenomena. In this chapter we presented the first ever predictive model for milky seas. This model serves as a critical first step to overcoming the practical barriers to sampling a milky sea in-situ.

The predictive model in this chapter led to the discovery of a new milky sea event in August 2017 near Java. This is the first time a milky sea has been discovered via a predictive methodology and serves as a verification for the potential of the model to identify future milky sea events. The potential of the SJMI model to forecast milky sea events was also examined using model output. At a forecast lead time of 1.5 months the SJMI was able to identify when milky seas will occur near Java. These results highlight the potential for predicting a milky sea before it happens to enable in-situ sampling.

Further work on this topic will focus on improving the temporal and spatial resolution of the SJMI via more localized sources of predictability. Crafting predictive models like the SJMI for the northwestern Indian Ocean is another key area of advancement for research into both milky sea predictability and understanding connections between milky seas and other facets of the coupled earth system.

Chapter 4
The Natural Flask Hypothesis of Milky Sea Formation: A Case Study Analysis

4.1 Introduction

The natural flask hypothesis of milky sea formation (Miller et al. 2021) proposes that the bacteria behind milky seas achieve the population densities necessary for quorum sensing driven bioluminescence via isolated to semi-isolated parcels of water which serve as natural analogues to a laboratory flask. In a laboratory setting the ideal conditions for a bacterial population such as light, temperature, salinity, nutrient availability, etc. can be maintained for extended periods of time. The natural flask hypothesis proposes that, given the right conditions, density gradients or dynamically driven processes can isolate parcels of nutrient and substrate rich water in the ocean where bacterial populations can thrive (Figure 3.5). Understanding if milky seas do form within natural flask type environments is important to advancing our understanding of this phenomenon.

In this chapter we perform a case study analysis on two milky sea events near Java, Indonesia. The first occurred in August 2017, hereafter referred to as MS2017, the second occurred over July – September 2019, hereafter referred to as MS2019. The MS2017 event, originally discovered via the hindcast model presented in Chapter 3, lasted from August 20th to August 26th, 2017. This milky sea event was initially visible on satellite imagery with well relatively well-defined boundaries (Figure 3.4) before rapidly becoming dimmer and more diffuse and disappearing entirely from DNB imagery.

The MS2017 event is one of the few milky sea events where its decay was visible to satellite imagery providing knowledge of exactly when and where the event was as it began to break down. This means the event can provide unique insight into the decay processes of milky

seas and provide information about how milky seas are maintained over wide areas for extended periods of time.

The MS2019 event, found via VIIRS DNB imagery and first reported on in M21, is the largest known milky sea in the satellite record and the longest lasting event in the satellite record (Figure 4.1). Understanding the conditions leading up to this event and across the lifetime of this event allows a view into how the natural flask hypothesis can explain behavior on large spatiotemporal scales compared to the relatively more modest MS2017 event (~1% the size of the MS2019 event).

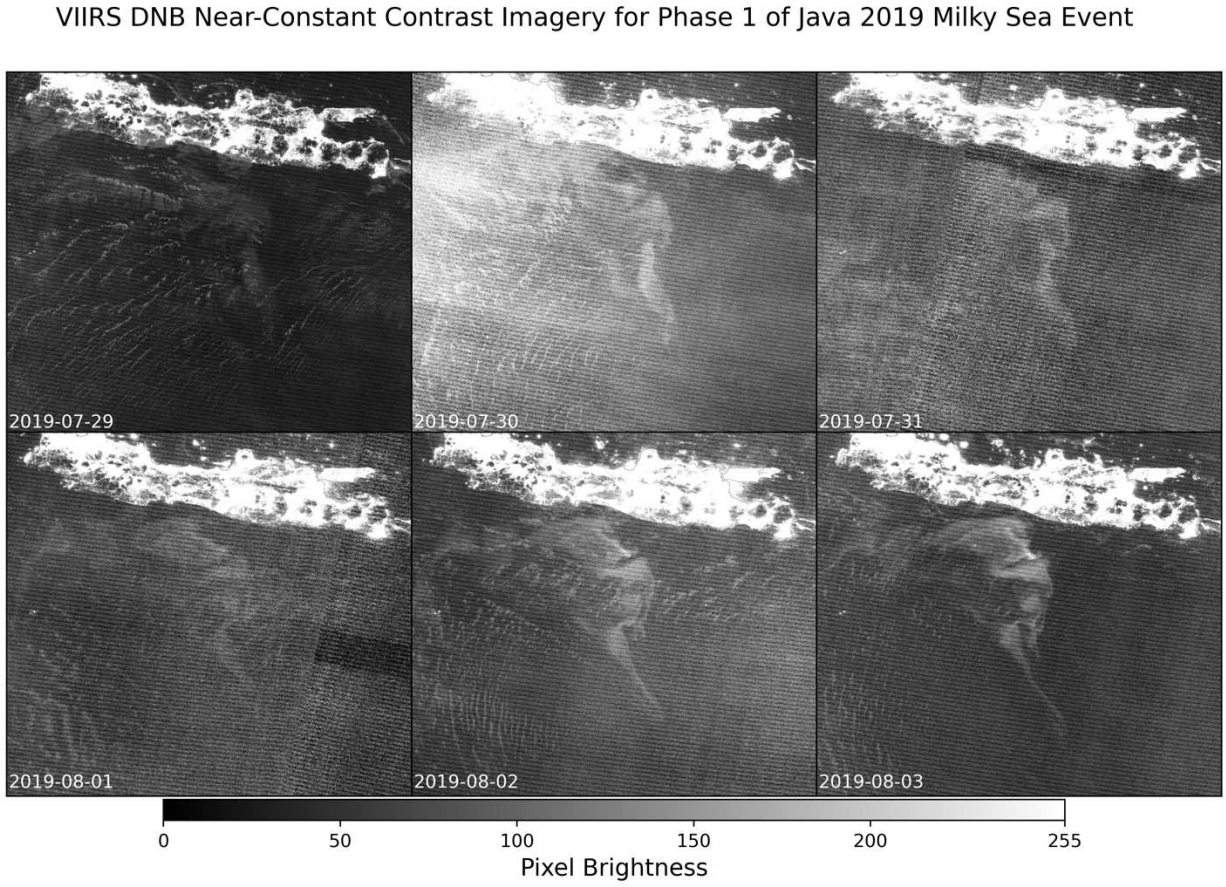


Figure 4.1: VIIRS DNB imagery of the MS2019 milky sea event near Java during the austral winter of 2019. This event lasted from July to September and exceeded 130,000 km^2 in size at its peak.

Section 4.2 of this chapter outlines the data sources used to carry out the case study analysis. Section 4.3 shows the results of the case study for MS2017, and Section 4.4 relays the results of the case study for MS2019. Section 4.5 concludes this chapter with a discussion of the insights this analysis provides for further research into milky seas.

4.2 Data Sources

VIIRS DNB imagery in this chapter was taken from the NASA Worldview application, part of NASA ESDIS. Polygons outlining the MS2017 and MS2019 events were manually extracted using the NASA Worldview application. These polygons match the visible outline of the milky sea event in VIIRS DNB imagery.

Zonal and meridional currents for the upper 100-meters of the ocean were taken from the US Navy's Hybrid Coordinate Ocean Model (HYCOM; hycom.org; Bleck 2002) for August 18th 2017 to August 29th 2017, and July 20th 2019 to September 20th 2019 (the time periods of the MS2017 and MS2019 events respectively). CMEMS chlorophyll for the upper 100-meters of the ocean were taken from the Global Ocean Biogeochemical Hindcast (doi: 10.46870/moi-00019) for the time periods of the MS2017 and MS2019 events. CMEMS salinity for the upper 100-meters of the ocean were taken from the Global Ocean Physics Reanalysis (doi: 10.48670/moi-00021) for the time periods of the MS2017 and MS2019 events respectively). Both HYCOM and CMEMS data were downloaded over the region near Java [100^{o}-120^{o}E, 15^{o}-5^{o}S].

4.3 MS2017 Event Case Study Analysis

4.3.1 Depth Estimation of the MS2017 Event

Eyewitness accounts of milky seas report the phenomenon as both being isolated to a surface slick or explicitly coming from some depth in the ocean. To accurately analyze the dynamics which shaped the breakdown of the MS2017 event an estimate of what depth in the ocean the event occurred at is necessary. To estimate the depth of the MS2017 event the polygon that outlines the MS2017 event on August 20th 2017, the first day the event was visible in DNB imagery, was overlaid with a 0.01^{o} x 0.01^{o} grid.

Every grid point on the 0.01^{o} x 0.01^{o} grid that fell within the bounds of the polygon was used as the starting position (in terms of latitude and longitude) to initialize a neutrally buoyant numerical particle. In total this produced 1444 numerical particles. The 1444 numerical particles were then advected in one-hour increments using the HYCOM currents for every layer in the ocean at or above 50 meters depth in the HYCOM model using a 4th order Runge-Kutta method. The full list of depths is 0, 2, 4, 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 45, and 50 meters depth. The particles were limited to the depth in which they were initiated with no impact of any vertical motion or mixing on their trajectory. The currents at the exact position of each particle were determined via a cubic interpolation in space from the HYCOM 1/12th degree grid and a linear interpolation in time from HYCOM's 3-hour timesteps.

The VIIRS DNB overpass from Suomi-NPP which captured the imagery of MS2017 and was used to generate the polygon outlining the MS2017 event occurred at ~01:30 local time on August 20th. As HYCOM data is provided in 3-hour increments and the MS2017 event is in the GMT+7 time zone the HYCOM 18Z time step from August 19th is the best time step through which

to initiate the numerical particles. Similarly, the HYCOM 18Z time step on August 25th is the best time step to end our advection experiment.

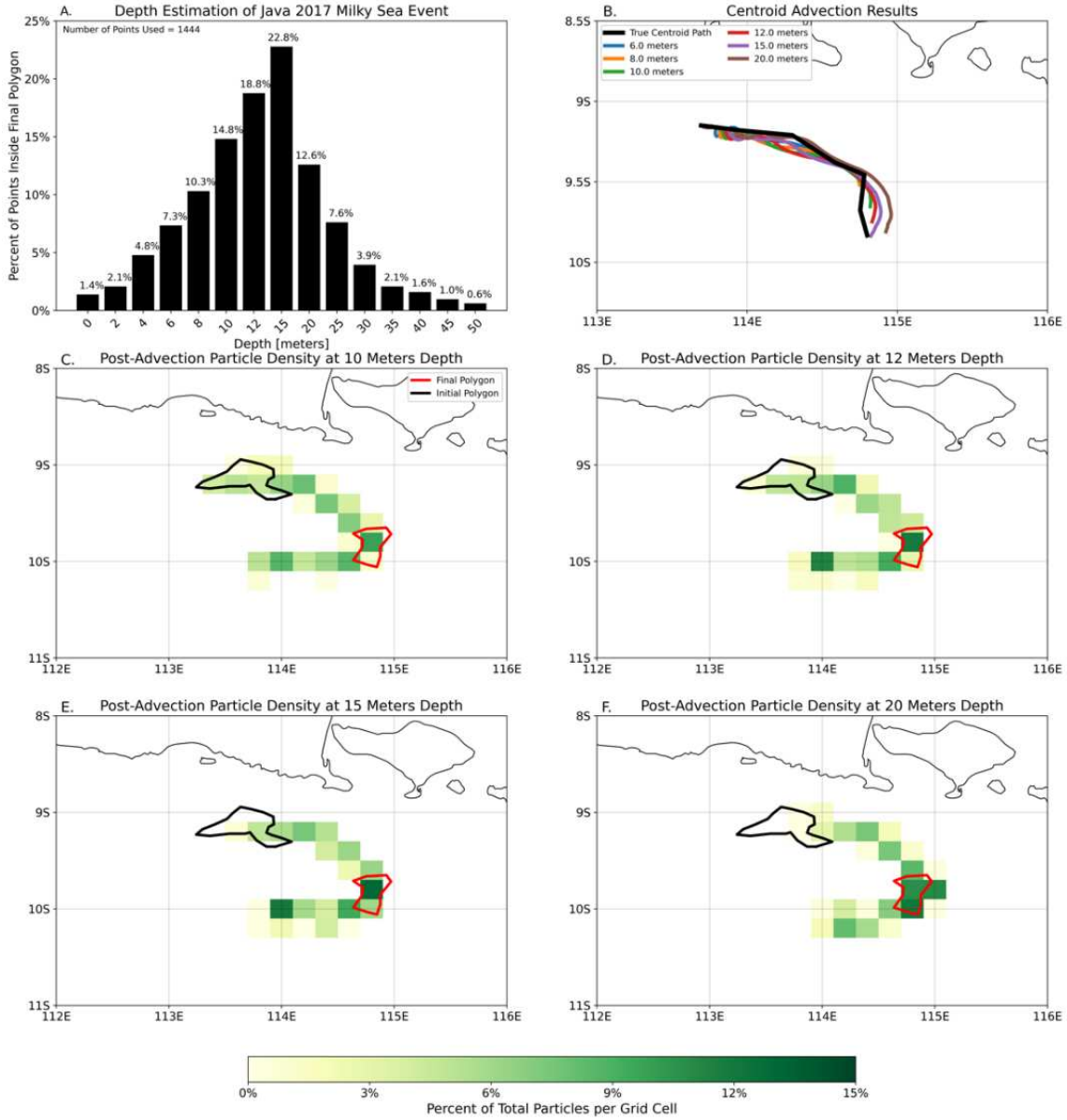


Figure 4.2: Depth estimation of the MS2017 event. Of the1444 numerical particles initiated within the polygon outlining the initial observation only 15 meters depth resulted in >20% of the particles being advected into the polygon corresponding to the final observation of the event (Panel A). Advecting a single numerical particle initialized at the geometric centroid of the MS2017 event for its initial observation the currents at 15-meters depth perform the best (Panel B). The relative density of particles after the advection experiment for the four most probable depths for the MS2017 event 10-meters (Panel C), 12 meters (Panel D), 15 meters (Panel E) and 20 meters (Panel F).

To estimate the depth of the MS2017 event we searched for the depth where the most particles initialized within the August 20th polygon outlining the event were inside the August 26th polygon outlining the event at the end of the advection experiment. The results of this experiment (Figure 4.2 Panels A and C-F) show that the MS2017 event likely occurred between 10-20 meters depth with 15 meters being the most likely depth.

A similar advection experiment using only a single neutrally buoyant particle initialized at the geometric centroid of the August 20th polygon was also carried out (Figure 4.2 Panel B). This experiment also shows 10-20 meters depth as the best estimate for the MS2017 event with 15 meters depth performing the best. The particle advected using the 15-meter depth current finished ~3 kilometers away from the true position of the geometric centroid of the MS2017 event.

Without in-situ sampling, determining the exact 3-dimensional structure and depth of the MS2017 event is impossible. Given the results however, for the rest of the analyses carried out in this chapter we assumed the MS2017 event was at 10-20 meters depth in the water column, with 15 meters being the most likely depth of the event.

4.3.2 Environmental Analysis of the MS2017 Event

With a likely depth for the MS2017 event determined we then analyzed the environment in the upper-20 meters of the ocean using the CMEMS Biogeochemical Hindcast chlorophyll, and HYCOM model output. To analyze the dynamical behavior, gradients in the meridional and zonal velocity fields were calculated and used to calculate the divergence (Equation 4.1), relative vorticity (Equation 4.2), horizontal shear (Equation 4.3), the horizontal strain (Equation 4.4) and the Okubo-Weiss Parameter (Equation 4.5).

$$\nabla \cdot V = \frac{dV_x}{dx} + \frac{dV_y}{dy} \quad 4.1$$

$$\omega = \nabla \times V = \frac{dV_y}{dx} - \frac{dV_x}{dy} \quad 4.2$$

$$s_n = \frac{dV_x}{dx} - \frac{dV_y}{dy} \quad 4.3$$

$$s_s = \frac{dV_y}{dx} + \frac{dV_x}{dy} \quad 4.4$$

$$W = s_n^2 + s_s^2 - \omega^2 \quad 4.5$$

Where V is the velocity field, V_x is the zonal component of the velocity field, and V_y is the meridional component of the velocity field. The Okubo-Weiss parameter is a measure of whether the local flow is dominated by horizontal shear and stretching (also known as strain) or by vorticity. Positive values correspond to a strain dominated flow, and negative values correspond to vorticity dominated flow.

The MS2017 event formed upstream and adjacent to a chlorophyll bloom, in calm, quiescent waters (Figure 4.3 top two panels). As the event was advected eastward by the current it eventually interacted with a parcel of vorticity dominated flow from the ITF where it exits the Banda Sea area east of Bali, Indonesia (Figure 4.3 bottom two panels). Upon interacting with the ITF outflow region, the MS2017 event became more diffuse in the VIIRS DNB imagery before disappearing entirely from imagery two nights later.

Analyzing this interaction through the lens of the natural flask hypothesis suggests that upon encountering this parcel with vorticity dominated flow, the density or dynamically driven barriers holding the natural flask together were broken. Once the walls of the flask were broken the milky sea quickly dissipated as its internal conditions mixed with its surroundings.

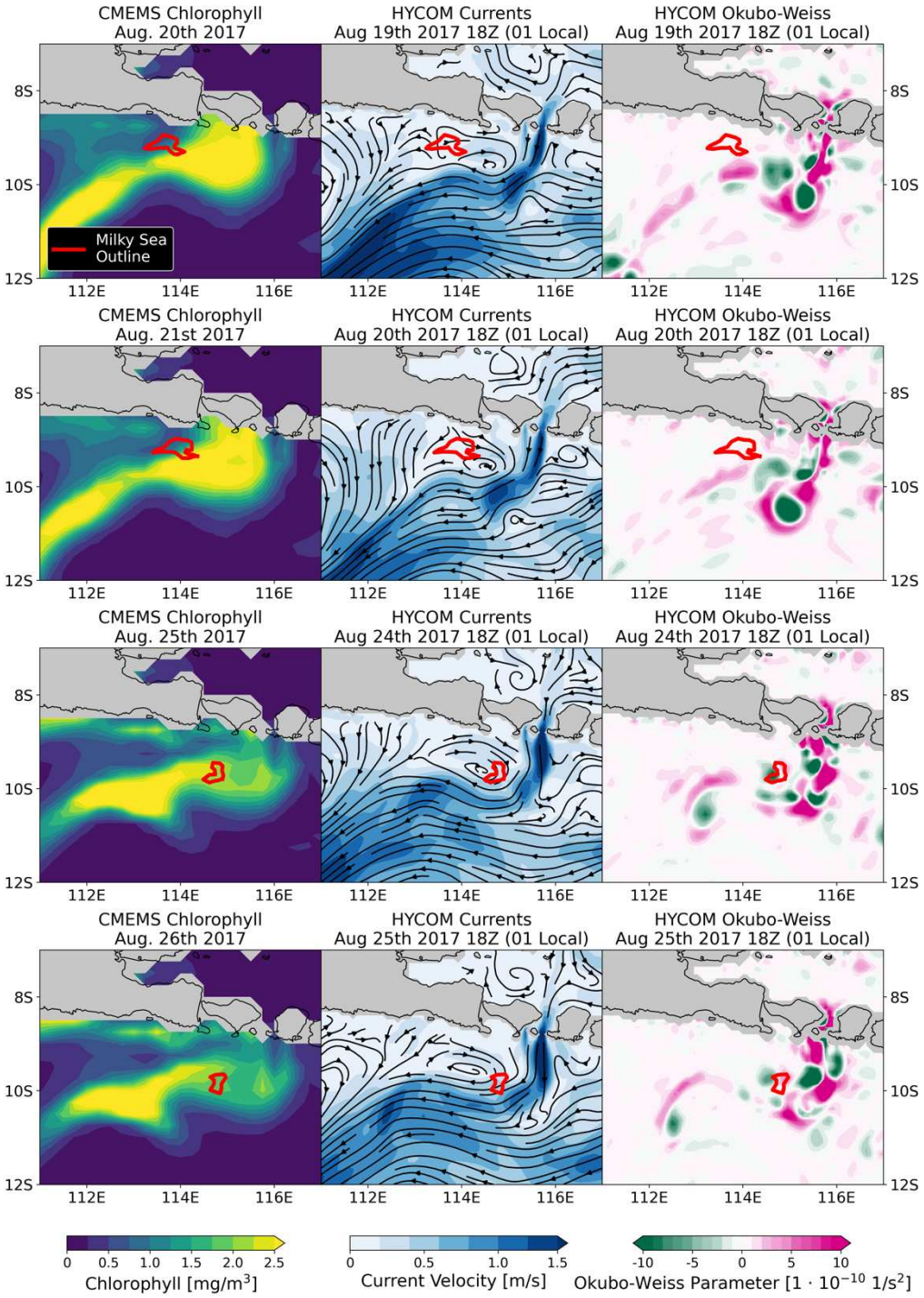


Figure 4.3: Environmental analysis of the MS2017 event for the first two days the event was observed (top two rows) and the final two days the event was observed (bottom two rows). The MS2017 event dissipated after interacting with a parcel of vorticity dominated water from the ITF outflow region near Bali, Indonesia

4.3.3 Transect Analysis of the MS2017 Event

To further investigate the impact of the vorticity dominated flow on breaking apart the MS2017 event we carried out a transect analysis of the MS2017 event to understand what occurred at depth. Given that the natural flask hypothesis theorizes that the interior waters are distinct from the surroundings, any analysis that doesn't compare how the interior of the event, as a whole, is distinct from the surroundings would not be able to prove to disprove the hypothesis.

For this reason, we elected to not perform linear transects but instead performed the transect analysis by taking the polygon outlining the MS2017 event for each day and identifying points within a distance $D + \delta D$ of the edge of the polygon. Once all points within a band bounded by D and $D + \delta D$ were identified, the mean of various quantities of interest for that collection of points was calculated at each depth down to 100-meters.

This process was then repeated for a band defined by $D + \delta D$ and $D + 2\delta D$, $D + 2\delta D$ and $D + 3\delta D$, and so on. This process allows for comparison between the interior of the MS2017 event and the surrounding environment. To create the transects distances in the range $[-0.1^o, 0.5^o]$ with a δD of 0.05^o were used. Negative distances lie in the interior of the polygon and represent internal conditions, positive distances are exterior to the polygon and represent external conditions.

The transects at various bands of $D + \delta D$ (hereafter called band transects) for HYCOM current velocity (Figure 4.4) show that for the first three days that the MS2017 event was visible (Figure 4.4 top row) the MS2017 event contained quiescent waters internally and was surrounded by similarly quiescent waters. For the last three days that the MS2017 event was visible (Figure 4.4 bottom row) the MS2017 event is embedded in higher velocity flow (due to its drift into the ITF outflow region), and internally contains higher velocity water due to mixing with the surrounding environment and entrainment into the ITF.

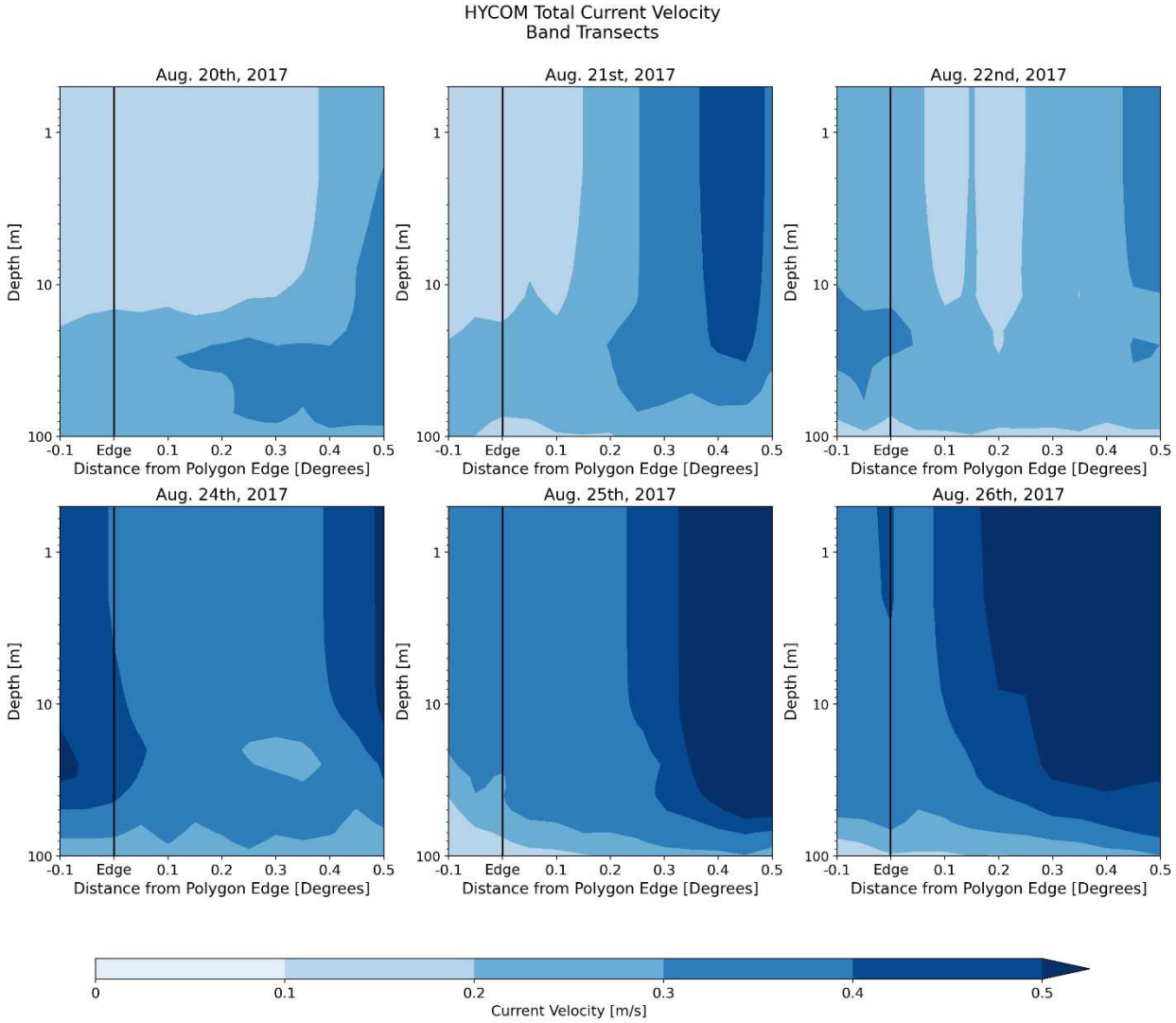


Figure 4.4: Band transects of current velocity from HYCOM for the first three days of the MS2017 event (top row) and the final three days of the MS2017 event (bottom row). Negative values along the X-axis represent values inside the MS2017 event, positive values along the X-axis represent the environment surrounding the MS2017 event. The Y-axis is log-scaled to emphasize values around 10-20 meters depth (the most probable depths for the MS2017 event).

Looking at band transects of relative vorticity (Figure 4.5) we see that initially the MS2017 event had positive relative vorticity internally versus the surrounding environment which had negative relative vorticity. After interacting with the vorticity dominated parcel from the ITF the relative vorticity of the MS2017 matches the surrounding environment. Looking at band transects of horizontal shear (Figure 4.6) we see a similar pattern where initially the MS2017 differed from the surrounding environment and, after interacting with the ITF outflow and, became more diffuse and began to decay as its internal conditions began to match those of the surrounding environment.

Taken together with the observations in section 4.3.2 we see that the MS2017 event initially formed within a dynamically distinct parcel of water inside a region of quiescent currents. This dynamically distinct parcel of water was maintained until interaction with a vorticity dominated region of flow in the ITF outflow west of Bali, Indonesia mixed the interior of the MS2017 event with the surroundings. This mixing broke the walls of the natural flask and lead to the eventual decay of the MS2017 event as seen in VIIRS DNB imagery.

This case study highlights the potential of the natural flask hypothesis in explaining the formation and decay of milky sea events. In the next section we examine how the natural flask hypothesis holds on large scales for the MS2019 event.

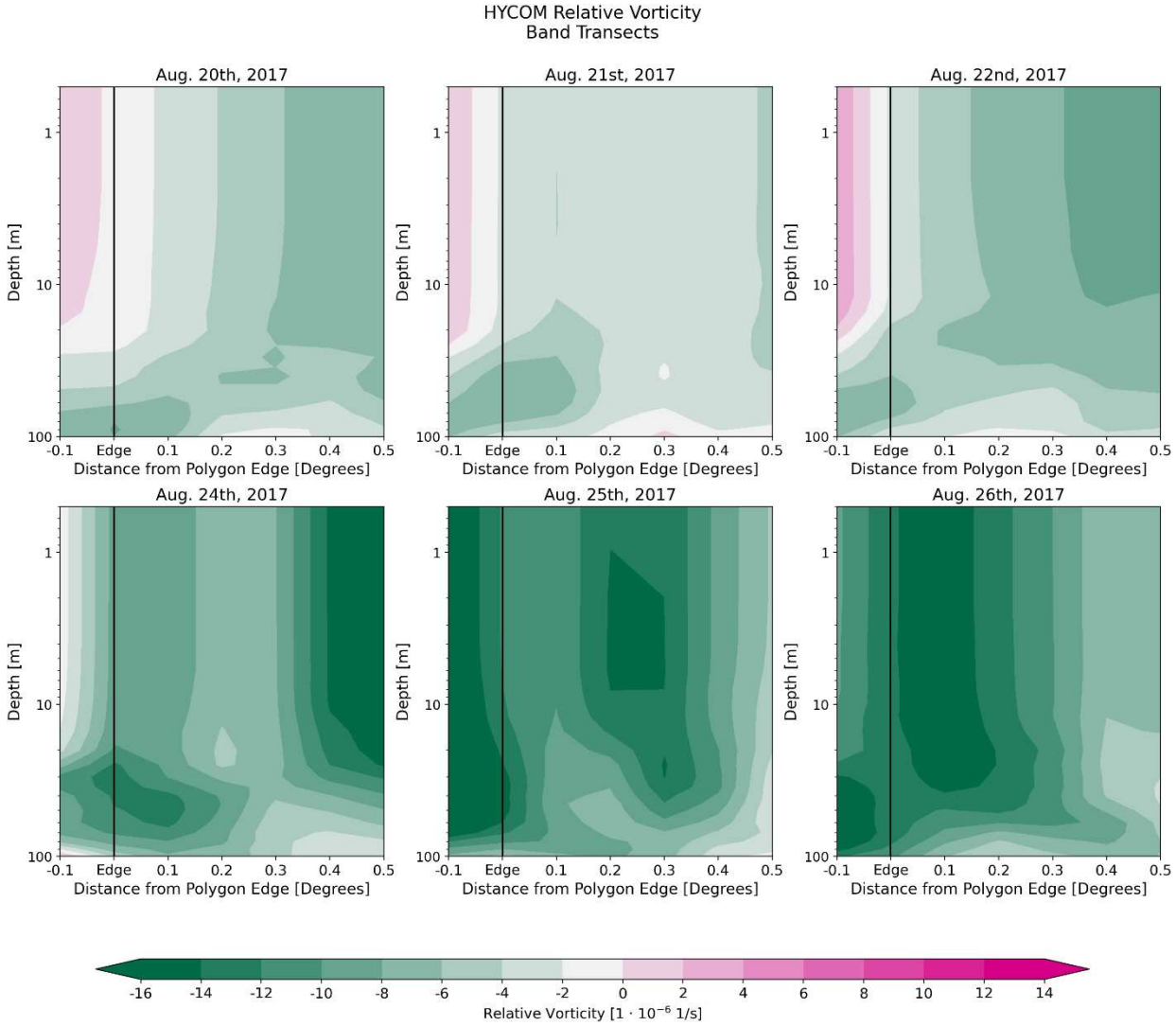


Figure 4.5: Band transects of relative vorticity calculated from HYCOM for the first three days of the MS2017 event (top row) and the final three days of the MS2017 event (bottom row). Negative values along the X-axis represent values inside the MS2017 event, positive values along the X-axis represent the environment surrounding the MS2017 event. The Y-axis is log-scaled to emphasize values around 10-20 meters depth (the most probable depths for the MS2017 event).

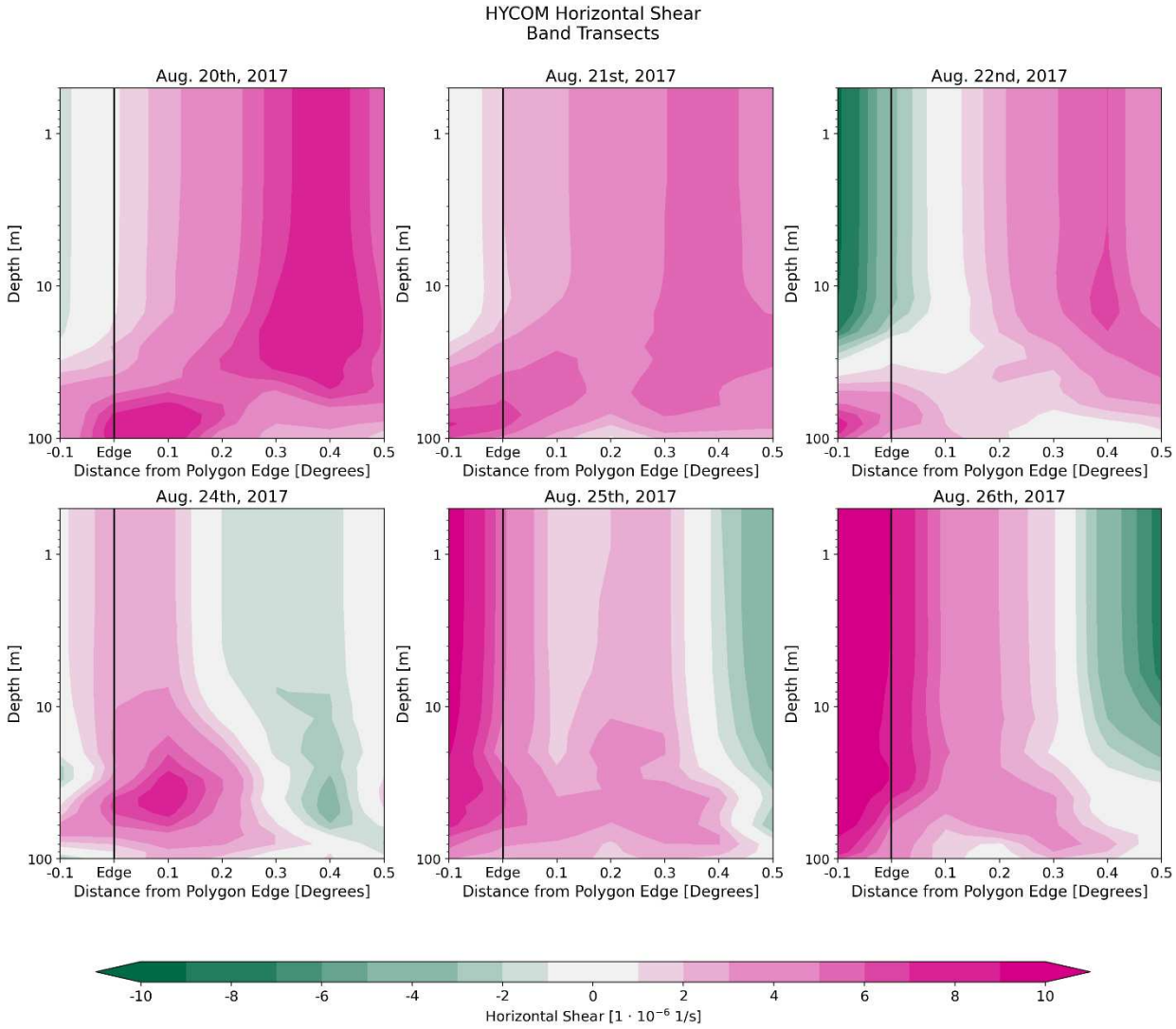


Figure 4.6: Band transects of horizontal shear calculated from HYCOM for the first three days of the MS2017 event (top row) and the final three days of the MS2017 event (bottom row). Negative values along the X-axis represent values inside the MS2017 event, positive values along the X-axis represent the environment surrounding the MS2017 event. The Y-axis is log-scaled to emphasize values around 10-20 meters depth (the most probable depths for the MS2017 event).

4.4 MS2019 Event Case Study Analysis

4.4.1 Environmental Analysis of Phase 1 of the MS2019 Event

The MS2019 event (Figure 4.1) exceeded 130,000 km^2 in size at its peak and potentially lasted three months subject to observational limitations of the VIIRS DNB (Section 2.2.4.3). This event represents the known upper extremes of milky sea events (Miller et al. 2021; Hudson & Miller 2025). Observation of the MS2019 event is split into two-moon free time periods, the first, called Phase 1, lasted from June 25th to August 9th, the second, called Phase 2, lasted from August 25th to September 9th. Understanding how well the natural flask hypothesis can explain the formation and subsequent maintenance of the MS2019 event at these scales serves as insight into the large-scale conditions that maintain milky seas.

Like with the MS2017 event, and many of the other events in the satellite record (Miller et al. 2021), the MS2019 event initially formed within quiescent waters (Figure 4.7). The MS2019 event was bounded to the southwest and southeast by two mesoscale eddies in the ITF. The quiescent waters that the MS2019 event formed within extend beyond 100 meters in depth according to both the HYCOM model and CMEMS reanalysis, and the quiescent region is largely moving coherently with the meanders in the ITF that push it westward. Due to this, the depth estimation analysis done for the MS2017 event does not yield any insight as to where in the water column the MS2019 event was located.

Looking at CMEMS salinity (Figure 4.8) we see that Phase 1 of the MS2019 event was bounded by salinity gradients on all sides. The upwelling near the coast brought high-salinity water from depth and the ITF outflow brought low salinity waters from the interior of the Maritime Continent (these low salinity waters can be seen on the northern coast of Java).

Phase 1 of the MS2019 event was bounded laterally by both dynamical barriers and gradients in salinity. These features worked together to prevent horizontal mixing of the MS2019 event with the surrounding environment and helped maintain the event over its lifetime.

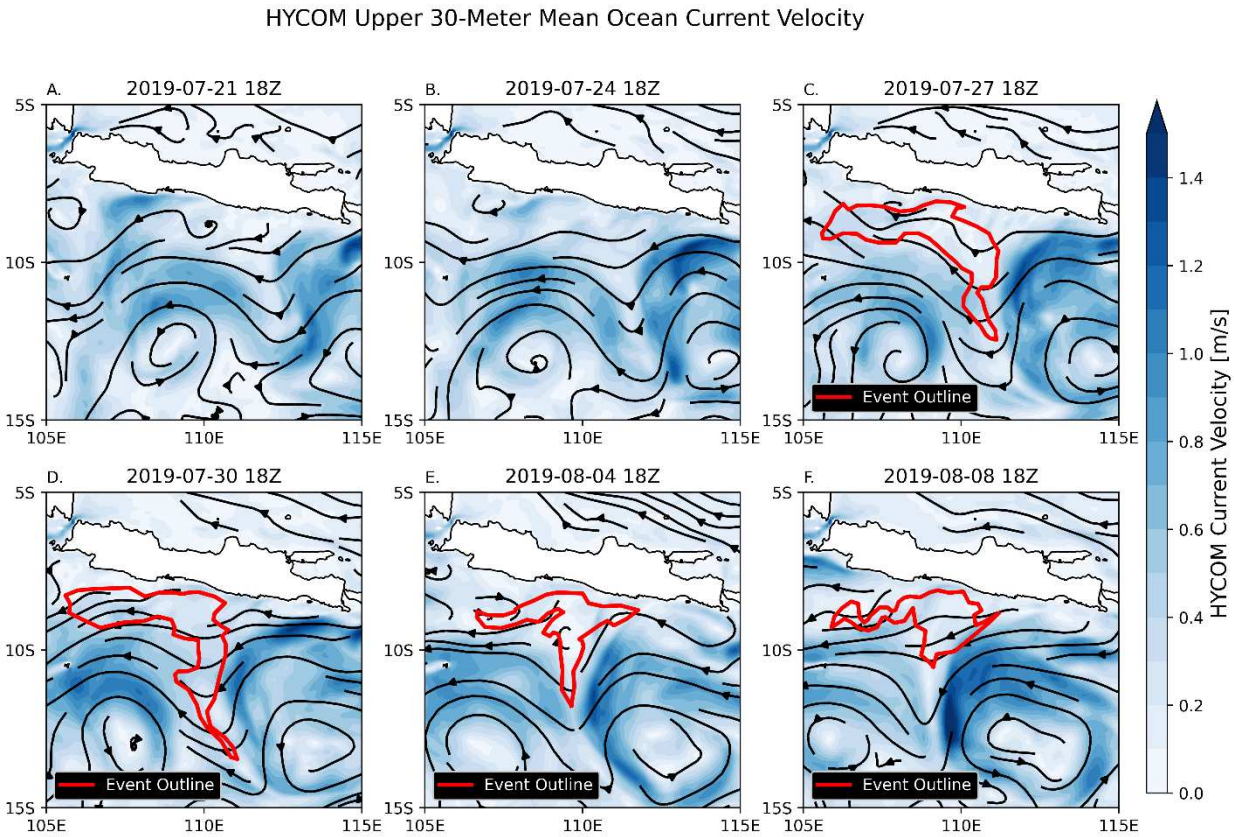


Figure 4.7: Plots of HYCOM mean current velocity for the upper 30-meters of the ocean near Java during Phase 1 of the MS2019 event. Red polygons representing the visible extent of the MS2019 event in VIIRS DNB imagery are overlaid to show where the MS2019 event was located relative to various features.

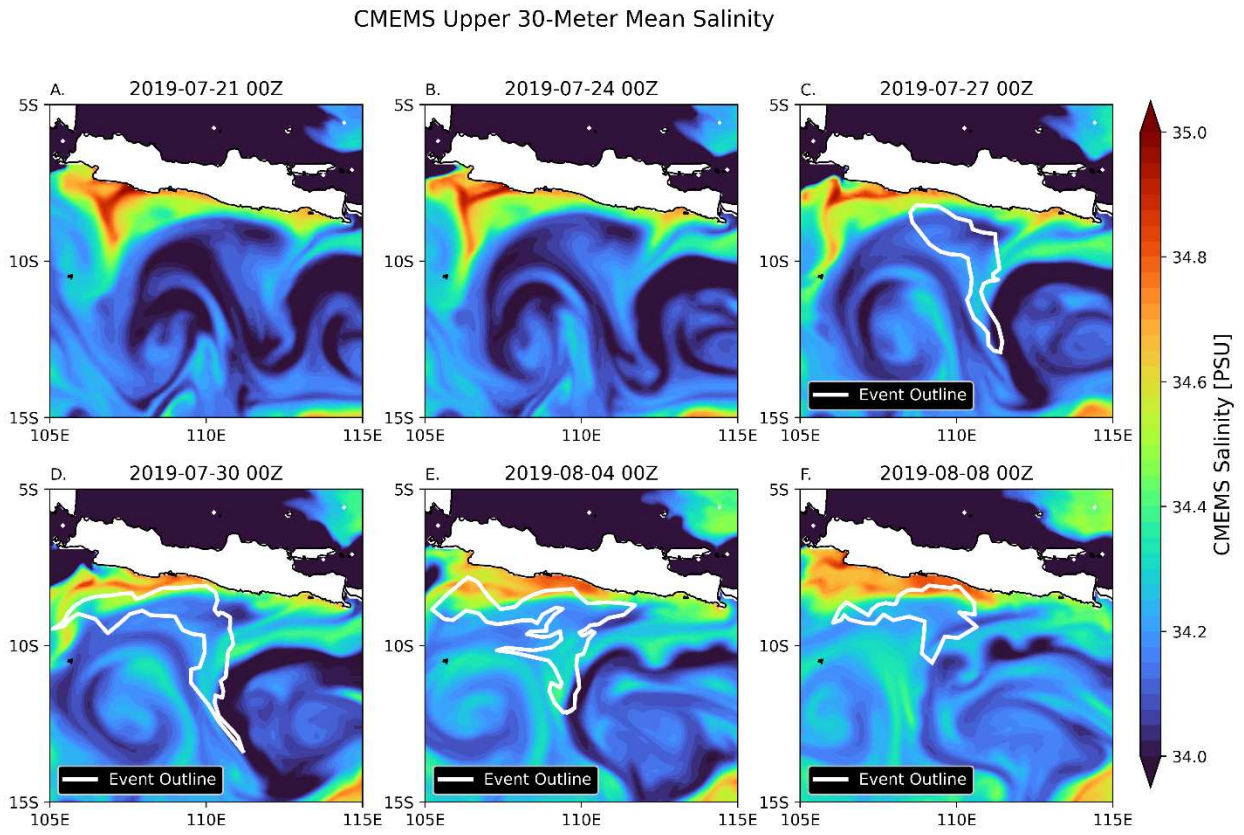


Figure 4.8 Plots of mean salinity in the upper 30-meters of the ocean from CMEMS data during Phase 1 of the MS2019 event. White polygons representing the visible extent of the MS2019 event in VIIRS DNB imagery are overlaid to show where the MS2019 event was relative to various features.

4.4.2 Transect Analysis of Phase 1 of the MS2019 Event

Performing a similar band transect analysis on the MS2019 event as with MS2017 over the range [-0.8°, 0.8°] with a δD of 0.05° we see that the dynamical and density barriers hold at depth (Figure 4.9). Band transects of CMEMS salinity show that the visible extent of Phase 1 of the MS2019 event lines up with gradients in salinity down to the halocline at approximately 40 meters depth. The band transect also shows that in a mean sense the halocline is shallower within the MS2019 event compared to the surrounding environment.

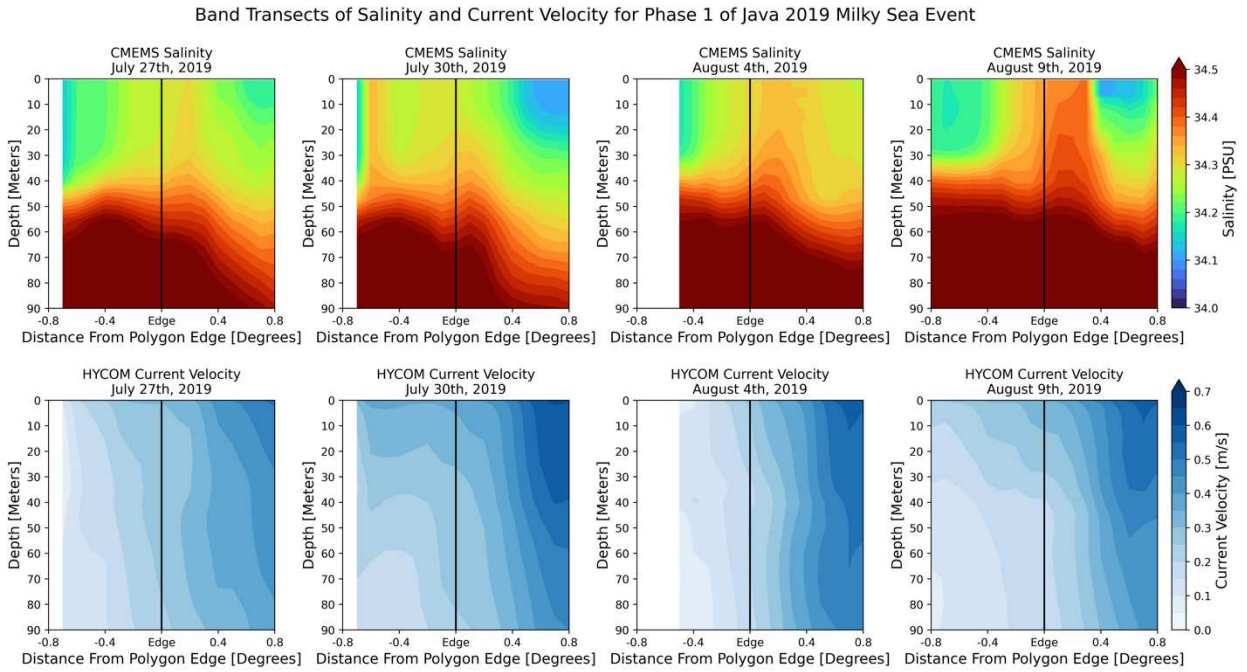


Figure 4.9 Band transects of Salinity from CMEMS and current velocity from HYCOM for phase 1 of the MS2019 event. Negative values along the x-axis represent points inside the MS2019 event, positive values along the x-axis represent points outside the MS2019 event.

Looking at the same band transects but for current velocity we see quiescent waters down to 90+ meters depth in the interior of the MS2019 event. These quiescent waters increase in velocity with distance away from the MS2019 event forming distinct horizontal gradients in the band transects. Like with the MS2017 event, Phase 1 of the MS2019 event is dynamically distinct from its surroundings.

4.4.3 Environmental Analysis of Phase 2 of the MS2019 Event

The MS2019 event is the only milky sea event in the satellite record that has been observed in multiple VIIRS DNB viewing windows. Phase 2 of the MS2019 event was elongated zonally with a long tail extending SW away from Java being stretched and advected by the ITF, and a triangular region near the SE corner of Java inside a patch of quiescent water between meanders in the ITF like with Phase 1 (Figure 4.10). The elongated tail portion of Phase 2 was observed in

DNB imagery to collide with Christmas Island and become divided into two pieces. This is a unique observation and both shows that the 'walls' of the natural flask may be more resilient than previously thought, and that a single milky sea event can fracture into multiple spatially disjoint regions which could be classified as separate events depending on when the event was confirmed in DNB imagery.

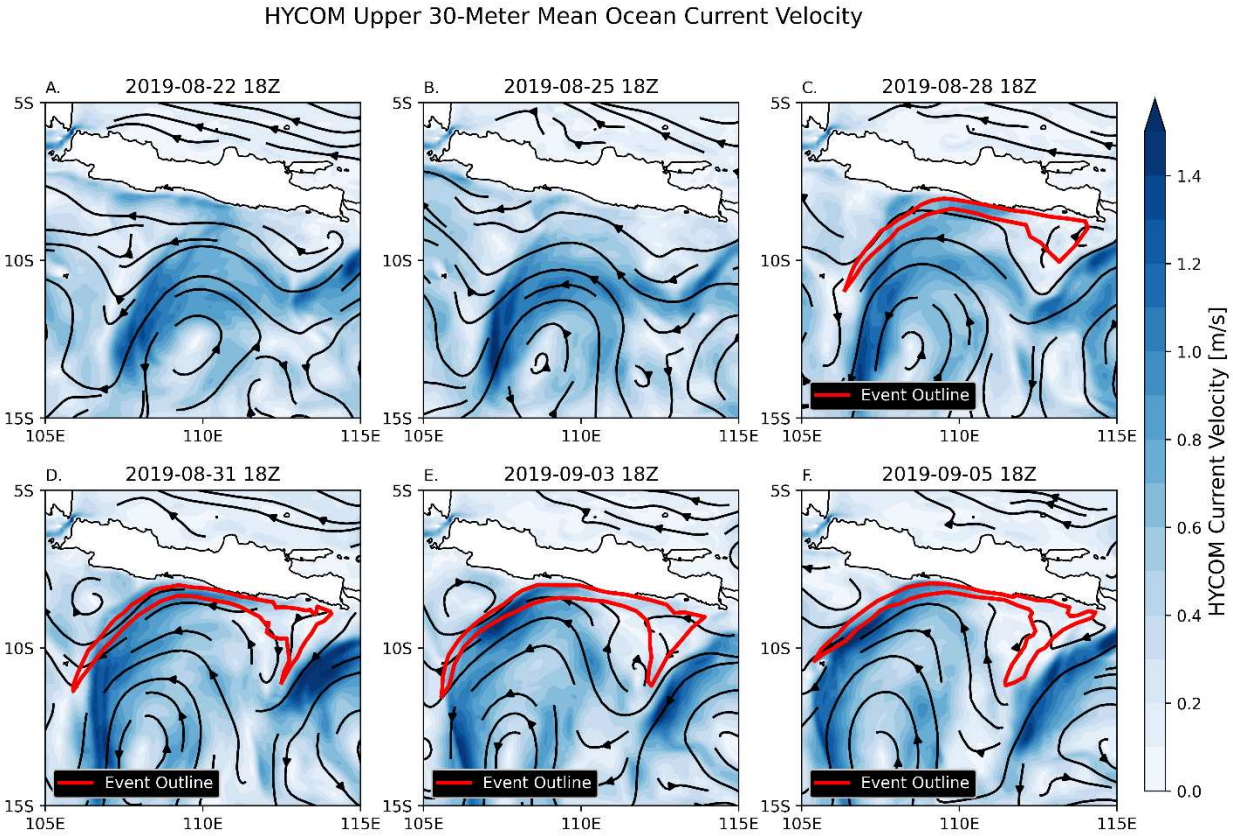


Figure 4.10: Plots of HYCOM mean current velocity for the upper 30-meters of the ocean near Java during Phase 2 of the MS2019 event. Red polygons representing the visible extent of the MS2019 event in VIIRS DNB imagery are overlaid to show where the MS2019 event was relative to various oceanic features.

Like phase 1 of the MS2019 event Phase 2 was bounded horizontally by gradients in salinity with high salinity to the north against the coast due to upwelled waters and lower salinity to the south due to the ITF carrying lower salinity waters from the interior of the Maritime Continent (Figure 4.11). For the duration of Phase 2 the entire region south of Java, was, as a whole, more saline than the same region during Phase 1 of the MS2019 event.

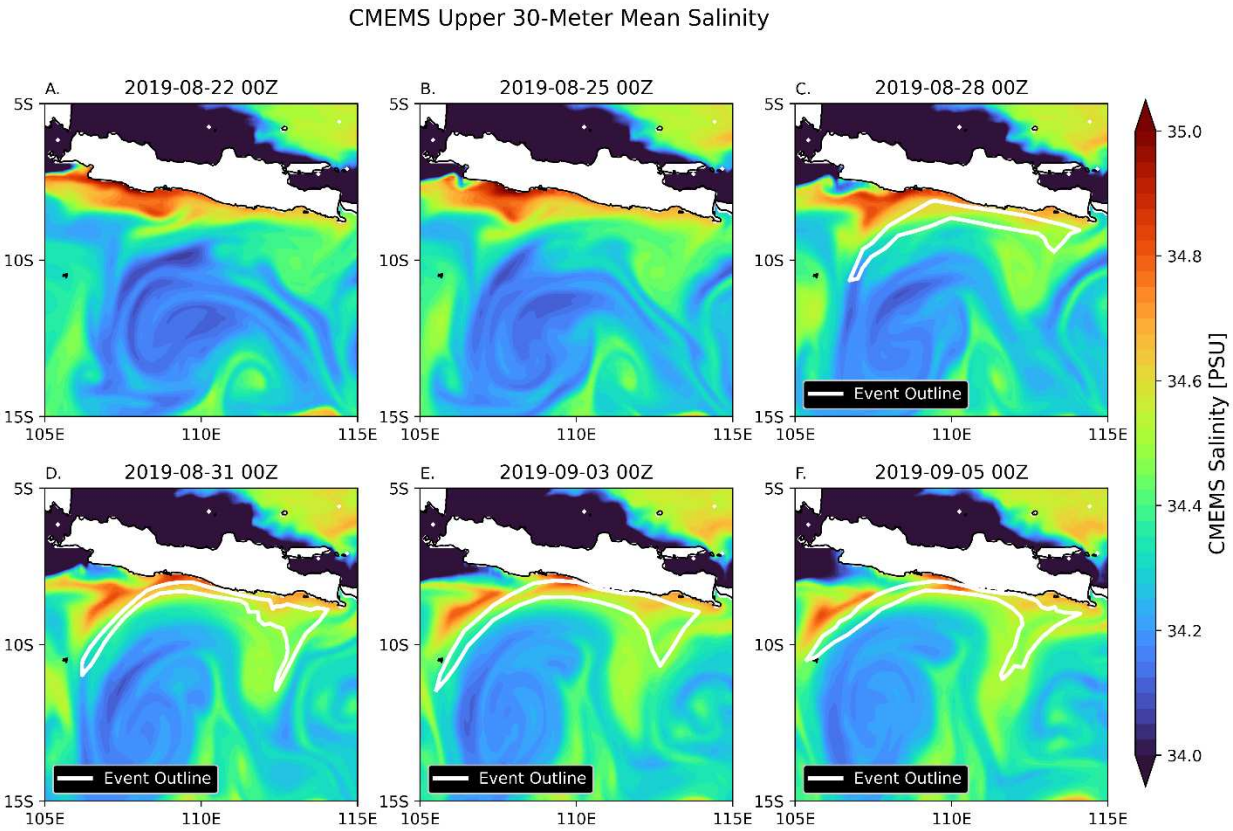


Figure 4.11 Plots of mean salinity in the upper 30-meters of the ocean from CMEMS data during Phase 1 of the MS2019 event. White polygons representing the visible extent of the MS2019 event in VIIRS DNB imagery are overlaid to show where the MS2019 event was relative to various features.

4.4.2 Transect Analysis of Phase 2 of the MS2019 Event

Like with the MS2017 event and Phase 1 of the MS2019 event we performed a band transect analysis over the range [-0.8°, 0.8°] with a δD of 0.05°. The band transects are strongly biased by the geometry of Phase 2 of the MS2019 event. The long thin tail of the event extending westward dominates the points in the band transect between -0.2° and 0° and the triangular feature embedded within quiescent waters is the entirety of the data for more interior points. This creates the sudden changes in salinity and current velocity in the band transects at -0.2° (Figure 4.12).

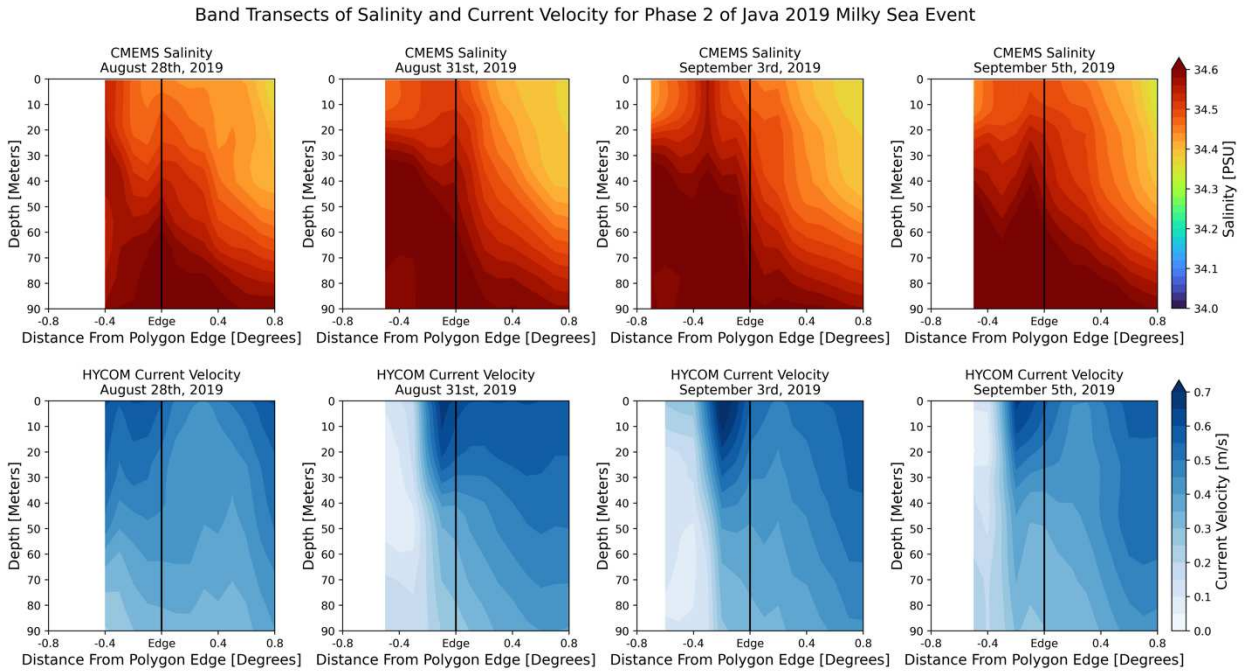


Figure 4.12: Band Transects of CMEMS Salinity and HYCOM current velocity for phase 2 of the MS2019 event. Negative values along the X-axis represent points inside the MS2019 event, positive values along the X-axis represent points in the environment surrounding the MS2019 event.

Turning again to the salinity gradients at depth (Figure 4.12 top row) like with Phase 1, the visible extent of Phase 2 of the MS2019 event aligns with gradients in salinity and the halocline within the MS2019 event is raised compared to the surrounding environment. The current velocity band transects (Figure 4.12 bottom row) are more visibly biased by the shape of the event as seen by the sudden change in current velocity at -0.2°.

Looking to the points exterior to the event there is no strong gradient in current velocity unlike with Phase 1 (Figure 4.9). While the different environments between the eastern and western extents of Phase 2 could be complicating the data in a mean sense this could also suggest that over the lifetime of a milky sea event the relative importance of dynamical barriers versus density gradients in sustaining the 'walls' of the natural flask can fluctuate both in space and time.

4.5 Conclusions

Two milky sea events were examined through the lens of the natural flask hypothesis. The first event, MS2017, provides insight into how well the natural flask hypothesis can explain the mechanisms governing decay of a milky sea event. Initially the MS2017 event was dynamically distinct from its environment. Upon interacting with the intense currents from the ITF outflow region near Bali, Indonesia the MS2017 event became visually diffuse in VIIRS DNB imagery as the interior was mixed with the surrounding environment before decaying entirely two nights later.

The MS2019 event, the largest event on record, was initially surrounded by a mixture of dynamical and density gradient features forming the walls of the natural flask. Upon re-appearing following the end of the lunar contamination window, the MS2019 event was largely bounded by salinity gradients versus dynamical features. This event shows that the natural flask hypothesis can be extended to the large spatial scales of the MS2019 event. The differing importance of dynamical features and density gradients in forming the 'walls' of the natural flask across the event and over time suggests that, as long as the 'walls' of the flask are maintained, the actual driver can vary.

This case study represents only a subsample of the milky sea events in the satellite record, examining how well these results extend to the 20+ other events in the satellite record would serve to further test the natural flask hypothesis. These results do show that, for in-situ sampling of a milky sea event samples will likely need to be taken at various depths across potentially large horizontal distances (100+ kilometers) to verify the natural flask hypothesis.

Chapter 5
Scatterometers and Associated Biological Signals of Milky Seas

5.1 Introduction

The ability of low-light imagers in recent years to remotely sense milky seas has greatly expanded the ability to study this phenomenon (Miller et al. 2005; Miller et al. 2021). The limitations of currently available satellite-based low-light imagers (discussed in Section 2.2.4.3) however, mean that natural obstacles such as cloud cover and lunar illumination can preclude the continuous and accurate detection of milky seas. A satellite-based method of tracking an active milky sea, if only approximately, during such time periods would greatly expand the ability to study milky seas and potentially deploy in-situ resources to sample them.

Scatterometers are a form of active microwave sensor that make use of changes in backscatter born from centimeter-scale variations in ocean surface roughness to detect the relative motion between the ocean and atmosphere (Kelly et al. 2001). Scatterometers, due to their operating wavelengths in the ~1-5 cm range, can peer through cloud cover producing high quality retrievals of the ocean surface. Measurements of ocean surface roughness via measured backscatter of the scatterometer pulse enables retrievals of wind speed under many atmospheric conditions (Garg et al. 2020; Ribal et al. 2021).

Surface slicks on the ocean surface modify the properties of the ocean-atmosphere boundary and can dampen waves at various scales (Franklin et al. 1774; Alpers & Hühnerfuss 1989; Lin et al. 2003; Hashizume & Liu 2004; Lindsley & long 2011). The impact of surface slicks on scatterometer data has been observed from both biogenic sources such as algal blooms (Lin et al. 2003; Hashizume & Liu 2004) and from human activity such as the Deepwater Horizon oil spill in 2010 (Lindsley & Long 2011).

Milky seas are, through the one known sample collected in 1985, associated with algae from the genus *Phaeocystis* (Lapota et al. 1988). Members of *Phaeocystis* play a key role in the global carbon cycle (Verity et al. 2007) and are known to produce a mucosal matrix which acts as a strong surfactant (Seuront et al. 2005). Members of *Phaeocystis* are also known for their ability to harbor and support bacterial populations (Mars Brisbin et al. 2022; Xu et al. 2022). These properties both make *Phaeocystis* a contender for part of the biological assemblage that make up milky seas and a possible source of surface films which may be visible to a space-borne scatterometer. It is important to note that other algal organisms besides members of *Phaeocystis* are associated with surfactant production on large scales (Lin et al. 2003; Hashizume & Liu 2004).

Surface film biases in scatterometer data are spatially and temporally coherent as it takes time for the film to be broken down (Alpers & Hühnerfuss 1989; Lin et al. 2003; Hashizume & Liu 2004; Lindsley & Long 2011). This coherence means that surface film biases in scatterometer data stand out over multiple overpasses versus other sources of shorter-term variability such as gradients in the wind field due to local meteorological conditions.

Satellite observations reveal that milky sea events often occur alongside large surface algal blooms, as seen through satellite-retrieved Chlorophyll-a (Chl-a; Miller et al. 2021); a metric related to the biomass of phytoplankton in near-surface waters. Ocean color measurements, the traditional way in which Chl-a is measured, use visible/near-visible sunlight reflectance bands, and thus, are daytime observations. This use of visible light means cloud cover and specular reflection of sunlight by the ocean surface (sun-glint) can lead to decreased data quality over the ocean. The ability of scatterometers to probe through cloud cover offers an advantage to ocean color measurements in tracking large scale algal blooms.

The large spatial scale of milky seas and associated/adjacent algal blooms allows for potential association of these signals in scatterometer data, providing a new lens through which to view and study milky sea events. In Section 5.2 of this chapter, we layout the data used in this investigation, and in Section 5.3 we use the MS2019 event as a case study to examine the potential utility of scatterometer data. Section 5.4 concludes this chapter with an examination of how scatterometer data could be utilized to provide guidance for a future sampling expedition to a milky sea event.

5.2 Data Sources

The VIIRS DNB imagery used in this chapter was taken from NASA's Worldview application, part of NASA ESDIS. The polygons outlining the MS2019 event were manually extracted using NASA Worldview. VIIRS Chl-a measurements (NASA Ocean Biology Processing Group) come from the VIIRS instrument aboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite, part of the Joint Polar Satellite System (JPSS) and was taken from NASA Ocean Color, part of the Ocean Biology Distributed Active Archive Center (OB.DAAC). The Chl-a data used spans 2012-2023 and is gathered during the daytime ascending node overpass of the satellite at approximately 13:30 local time. This is approximately 12 hours separated from the descending node VIIRS DNB measurements of night-time scenes which are used to identify milky sea events.

The scatterometer data used in this chapter comes from the Advanced Scatterometer (ASCAT) aboard the Meteorological Operation B (MetOp-B) polar orbiting satellite (EUMETSAT/OSI/SAF, KSNMI 2013) and consists of swaths (two 550 km wide swaths, separated by a gap of 360 km) of wind speed and direction at 25 km resolution across these swaths. The MetOp-B data span 2012-2023 and have spatial resolution of 25 km. In late 2022 MetOp-B

began to drift out of orbit in preparation for its eventual decommissioning. The time period for Chl-a and scatterometer data used in this chapter was chosen for the temporal overlap between S-NPP (the source of DNB imagery for milky sea events over this time period) and MetOp-B, and to limit any potential changes in quality due to ground track drift from MetOp-B.

10-meter winds from the ECMWF 5th Generation Reanalysis (ERA5) were taken once daily at 18Z from July 1st, 2019 to October 1st, 2019 in the region near Java, Indonesia [100^{o}-120^{o}E, 20^{o}-5^{o}S] (Hersbach et al. 2020; Hersbach et al. 2023). 1000 hPa winds from the Global Forecast System (GFS) were provided by the National Center for Environmental Prediction (NCEP)/National Weather Service (NWS)/NOAA/U.S. Department of Commerce and downloaded from the University Corporation for Atmospheric Research's (UCAR) NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive (doi: 10.5065/D65D8PWK) for August of 2019. These wind data are used to investigate biases in scatterometer data as well as their potential assimilation into other products.

5.3 Analysis

Milky seas are associated with large-scale surface Chl-a blooms (Miller et al. 2021; Figure 5.1 Panel A), some of which may produce dampened capillary waves on the ocean surface. To investigate whether scatterometer wind retrieval biases are present in large milky sea events, we compared VIIRS Chl-a data with ASCAT winds over the lifetime of the MS2019 event near Java, and for every Austral Winter (defined as June 1st to September 30th) between 2012 and 2023. As the ASCAT winds were initially at 25 km resolution across their swaths and the VIIRS Chl-a data is gridded at 4 km resolution the ASCAT winds were first re-gridded (oversampled) via nearest neighbor interpolation to match the VIIRS Chl-a data's resolution.

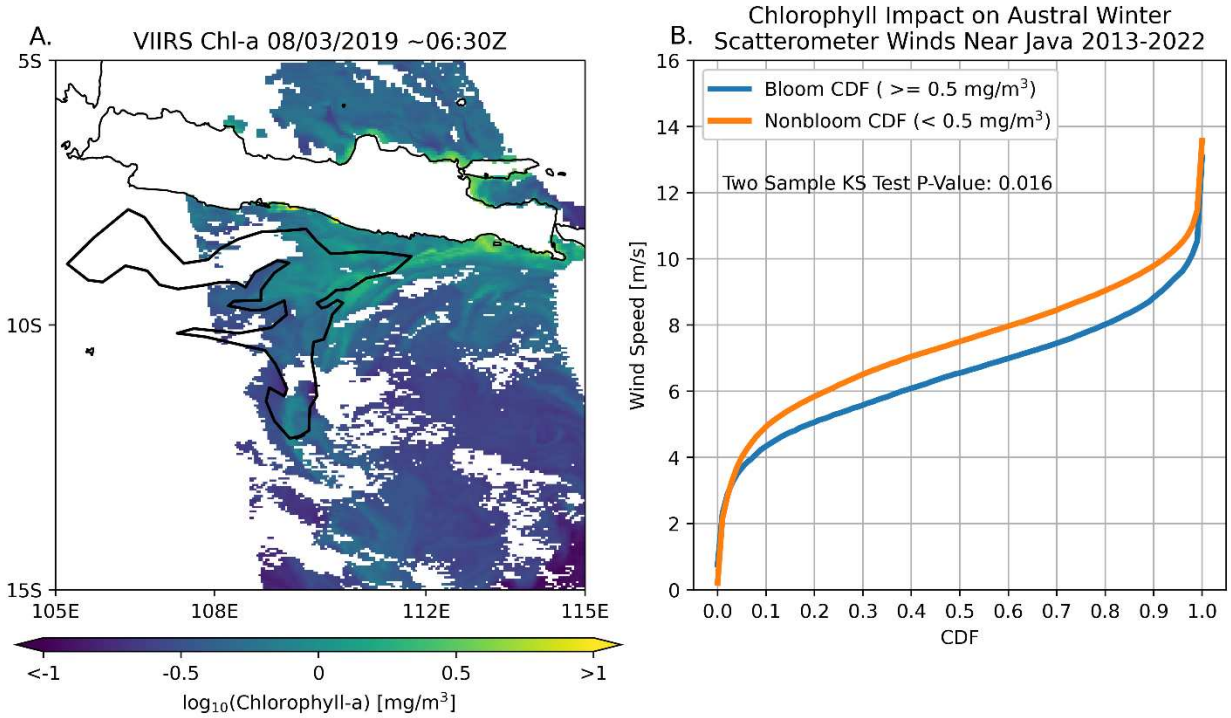


Figure 5.1: Chlorphyll-a (Chl-a) as measured from the VIIRS instrument aboard Suomi-NPP overlaid with a polygon outlining the location of the MS2019 event on August 3rd 2019 (Panel A). The Cumulative Distribution Functions (CDF) for observed MetOp-B ASCAT wind speeds for pixels that correspond to chlorophyll blooms (Chl-a >= 0.5 mg/m^3) and pixels that do not correspond to chlorophyll blooms (Chl-a < 0.5 mg/m^3; Panel B).

Like S-NPP, MetOp-B has multiple overpasses over Java during a single 24-hour time period at ~01Z and ~13Z (~08:00 local time, and ~21:00 local time respectively). These overpasses are reasonably equitemporal from daytime S-NPP overpasses (~13:30 local time). For the same calendar day on Java both MetOp-B overpasses were paired with the Chl-a data from S-NPP. ASCAT measurements from MetOp-B were then divided into two classes, pixels temporally collocated with a VIIRS Chl-a measurement greater than or equal to 0.5 mg/m^3 and pixels temporally collocated with a Chl-a measurement less than 0.5 mg/m^3.

Comparing the cumulative distribution function (CDF) of windspeed for our two classes near Java [105°-115°E, 12°-7°S] for every Austral Winter (June-August) between 2012 and 2023,

we found that scatterometer winds are, on average, ~1 m/s weaker for pixels belonging to the chlorophyll blooms class (Chl-a >= 0.5 mg/m^3) than pixels belonging to the non-bloom class (Chl-a < 0.5 mg/m^3; Figure 5.1 Panel B). Comparing the two distributions using a 2-Sample Kolmogorov-Smirnov Test (KS-Test) shows that the results are statistically distinct at the 95% confidence level (P-Value = 0.016).

The wind speed bias in scatterometer data near Java for the austral winter time period can be seen near the MS2019 event (Figure 5.2). The MS2019 event is the largest satellite observed milky sea event on record, and was associated with a large chlorophyll bloom along its eastern edge (Figure 5.1 Panel A) during Phase 1 of the event (June 25th to August 9th, 2019). This chlorophyll bloom expresses in ASCAT imagery as a region of suppressed (low-biased) wind speeds (Figure 5.2 Panel B). This area of chlorophyll associated low wind speed bias was of sufficient size that it was assimilated into ERA5 10-meter winds (Figure 5.2 Panel C). Forecasted GFS 1000 hPa winds for this period, which would not have had time to assimilate scatterometer data, do not show this bias (Figure 5.2 Panel D).

We then investigated how this scatterometer bias co-evolves with the MS2019 event. For Phase 1 of the MS2019 Event (Figure 5.3) there is a persistent SW-NE oriented low wind speed region approximately near 110^{o}E, 10^{o}S. This scatterometer feature lines up with the southern and eastern edge of the MS2019 event during this time period. In VIIRS Chl-a retrievals we can see elevated Chl-a on days with relatively cloud and sun-glint free conditions over this feature. This low windspeed feature is visible through the latter half of Phase 1 of the MS2019 event and continues to be observed after the event has disappeared from VIIRS DNB view.

Examining the time period between Phase 1 and Phase 2 of the MS2019 event, hereafter the Interphase time period (when moonlight precluded nightly tracking of the milky sea by the

DNB), this SE-NE low-biased wind speed feature remains visible in the scatterometer data (Figure 5.4). Likewise, the VIIRS Chl-a data continues to have limited visibility of this feature over the Interphase time period due to cloud cover and sun-glint. A low-bias wind feature aligning with high Chl-a is also visible on the southeastern coast of Java (~114°E, 9°S).

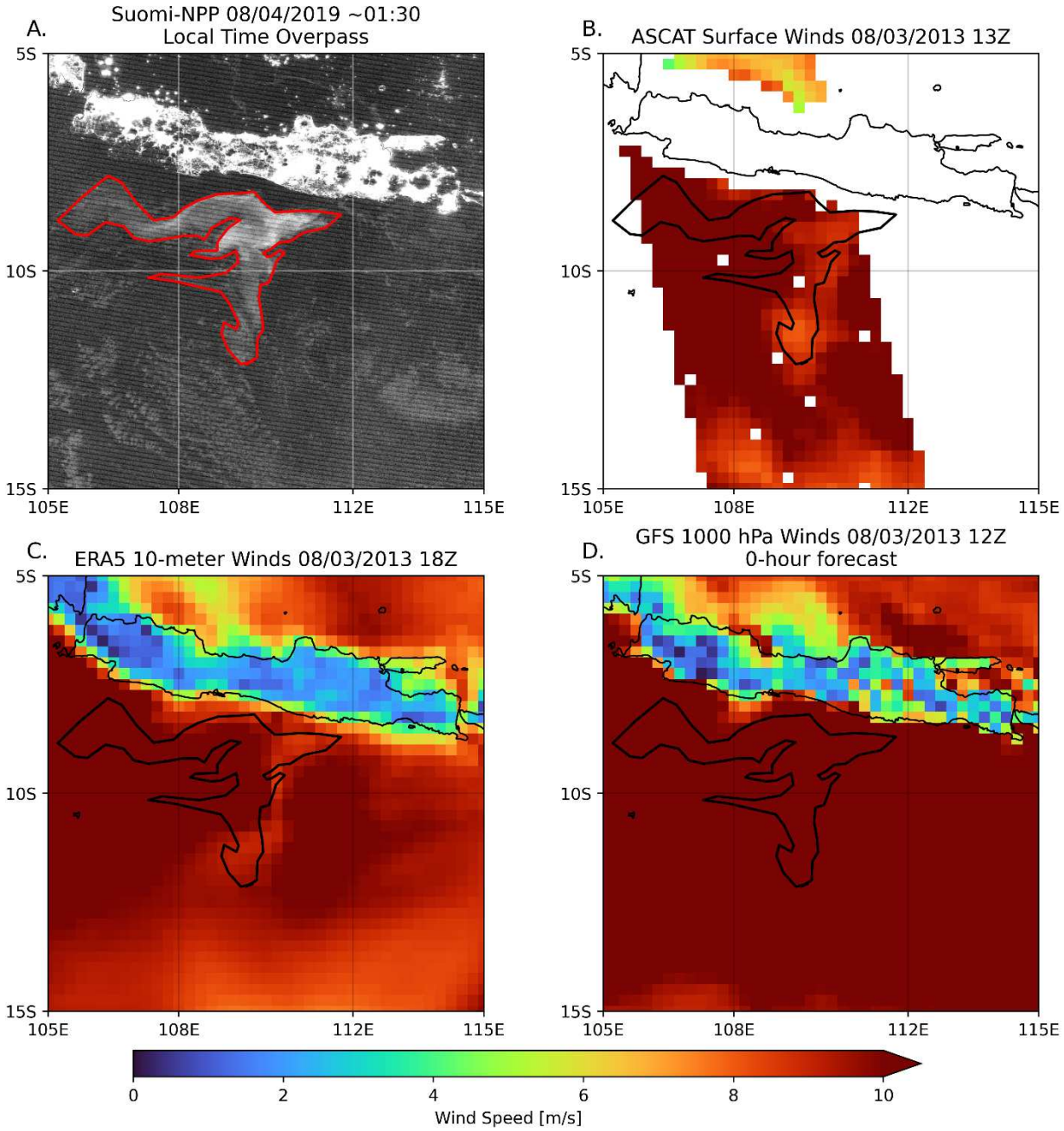


Figure 5.2: A VIIRS DNB image from Suomi-NPP overlaid with a red polygon highlighting the outline of the MS2019 milky sea event (Panel A). MetOp-B ASCAT surface winds for the closest time to the VIIRS DNB image in Panel A (Panel B). ERA5 10-meter winds from the closest time to the VIIRS DNB image in Panel A (Panel C). GFS Forecasted 1000 hPa winds for the closest time period to the other 3 panels (Panel D).

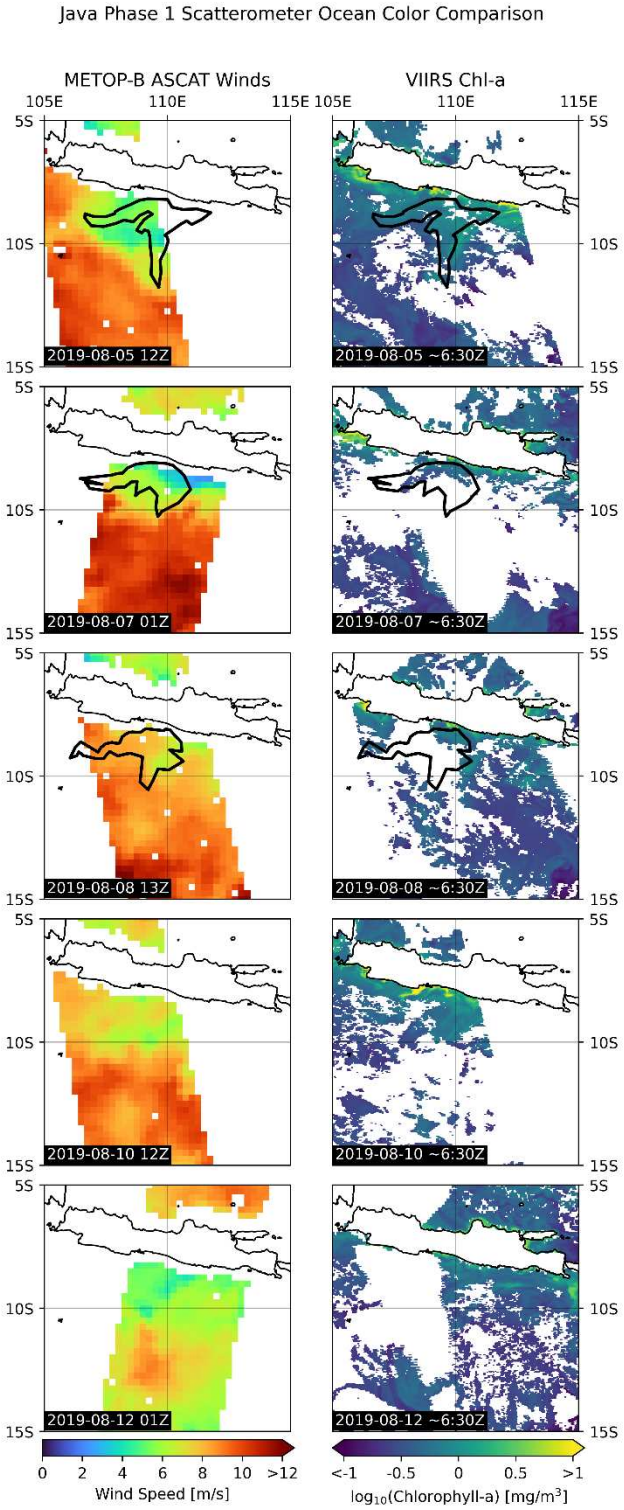


Figure 5.3: A Comparison of MetOp-B ASCAT scatterometer surface winds (left column) and VIIRS Chl-a from Suomi-NPP for Phase 1 of the MS2019 event (right column).

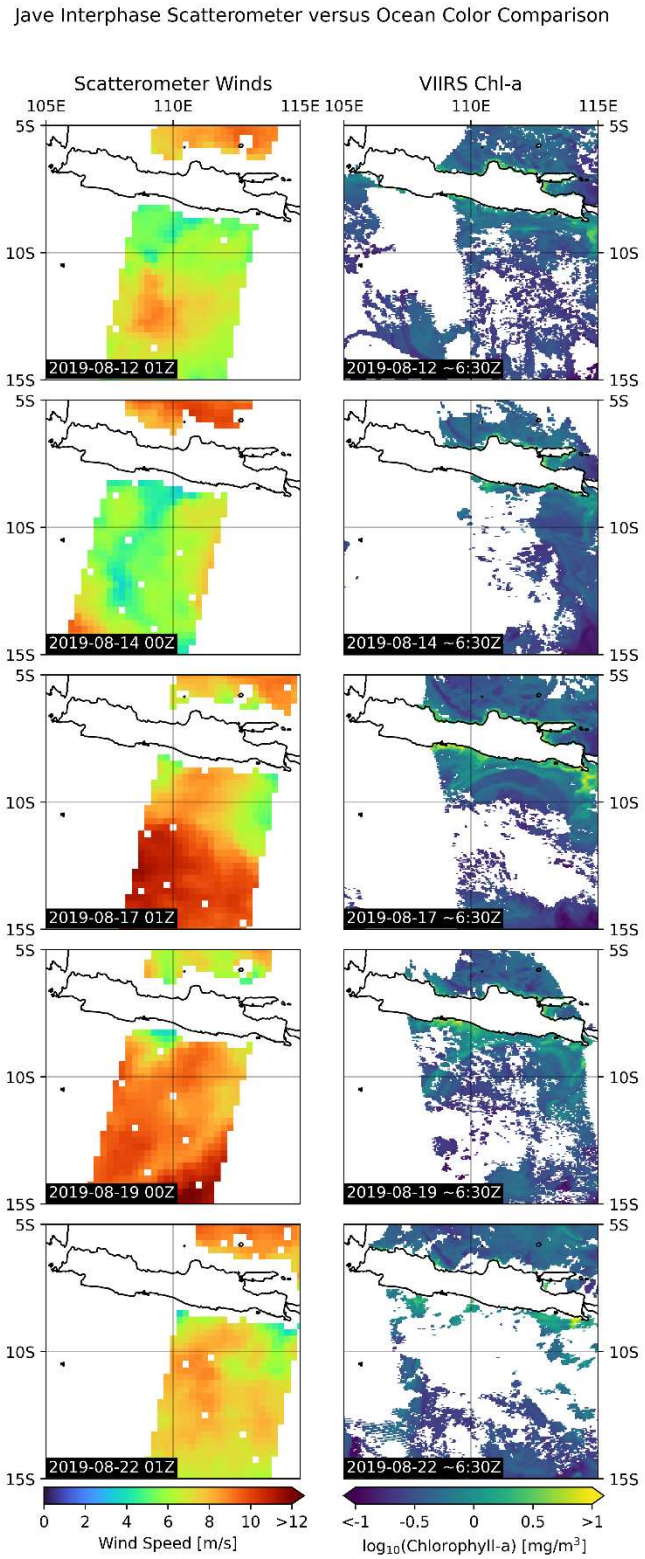


Figure 5.4: A comparison of MetOp-B ASCAT scatterometer surface winds (left column) and VIIRS Chl-a from Suomi-NPP (right Column) for the Interphase time period of the MS2019 event.

Transitioning to the Phase 2 time period for the MS2019 event (August 26th – September 6th) the low-wind speed bias features near the southeastern coast of Java and the SW-NE oriented feature that aligned with the southeastern edge of the MS2019 event continue to be visible in the scatterometer data. On August 26th when the MS2019 event became detectable again via VIIRS DNB imagery due to the moon once again being below the horizon during the ~01:30 local time of the satellite overpass, the event is now seen to be elongated to the west with a long tail, and a triangular shaped feature on its eastern flank. This elongated tail to the west aligns with the NW-SE oriented feature and the triangular shaped feature in the VIIRS DNB imagery aligns with the similarly shaped feature seen in the scatterometer data along the southeastern coast of Java.

By following the evolution of these spatiotemporally coherent reduced wind-speed features in the scatterometer data the location of the MS2019 event following the end of moonlight contamination of the VIIRS DNB imagery could be seen. These reduced wind-speed features align with regions of high Chl-a as seen through VIIRS, but cloud cover and sun-glint greatly limit the ability to follow these features in ocean color imagery. Given the relationship between Chl-a and scatterometer wind speed bias, these spatiotemporally coherent reduced wind speed features are likely a coincident biological signal to the MS2019 event.

The MS2019 event serves as an example for how with large milky sea events scatterometer data can be used to inform where the milky sea is located by tracking coincident biological signals versus the milky sea itself. VIIRS DNB imagery, like with Chl-a, is constrained by cloud cover. For time periods where an active milky sea event has been observed but cloud cover or lunar illumination limit the utility of nighttime VIIRS DNB imagery, tracking of coincident biological signals via scatterometer may serve as a form of backup guidance as to where a milky sea event is located once it has been identified in VIIRS DNB imagery.

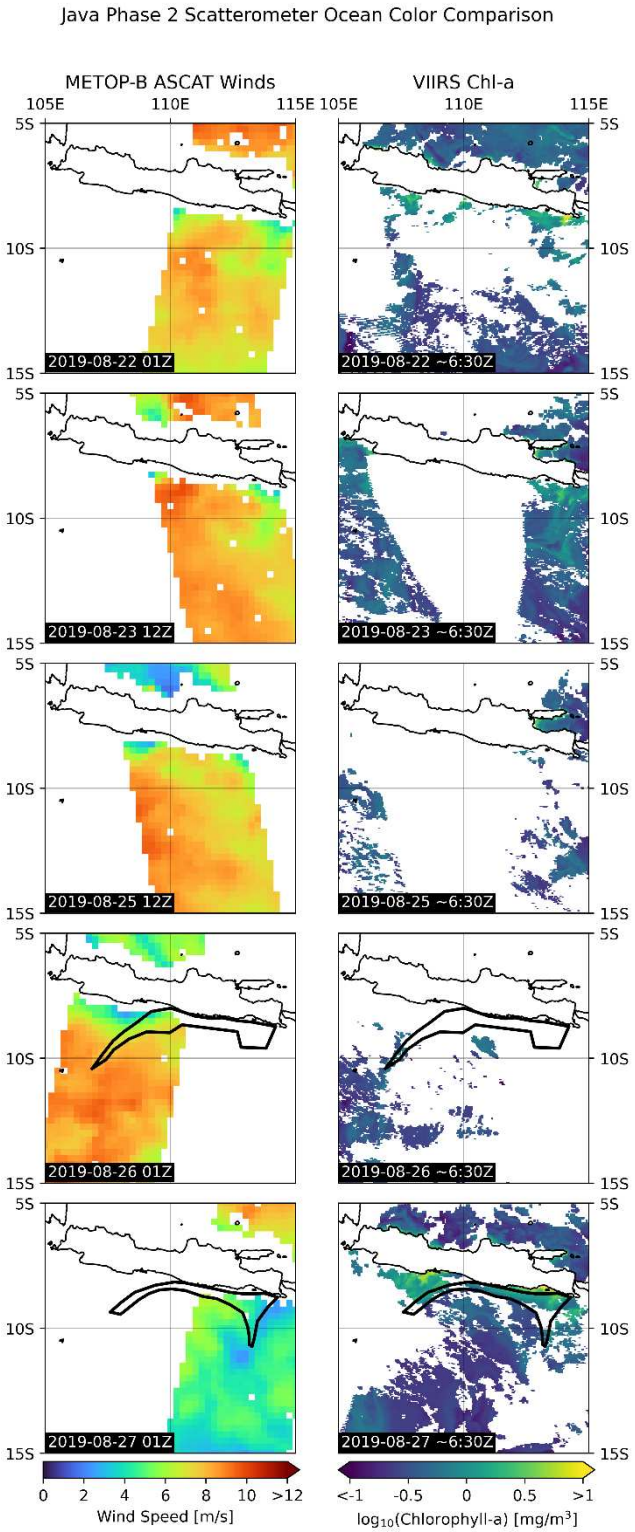


Figure 5.5: A comparison of MetOp-B ASCAT scatterometer surface winds (left column) and VIIRS Chl-a from Suomi-NPP (right column) for Phase 2 of the MS2019 event.

5.4 Conclusions

VIIRS DNB imagery is the primary means to observing milky sea events remotely and directly. Factors such as cloud cover and lunar illumination, however, limit the utility of the VIIRS DNB. In particular, lunar contamination limits the observation window to approximately half of each ~29 day lunar cycle. Biogenic surface slicks associated with algal blooms are a spatially and temporally coherent feature that is often coincident with milky sea events. The persistent wind speed suppression (low bias) in scatterometer retrievals due to biogenic surface slicks represent a spatially coherent feature in the ocean that coevolves with milky sea events that can be observed day and night as well as through cloud cover regardless of the lunar cycle.

Tracking these scatterometer biases between Phases 1 and 2 of the MS2019 event revealed the eventual location and form of the MS2019 event following the end of lunar contamination of the VIIRS DNB imagery. The ability to know where a drifting milky sea may reappear at the start of the VIIRS DNB window or where it may be located to an in-situ observer is critical guidance that, to this point, had not been possible.

The use of scatterometer data in this context is limited in that it requires a-priori knowledge of an active milky sea event, information largely only possible through VIIRS DNB imagery. Surface slicks can also break up under high wind speeds meaning large scale meteorological forcings associated with milky seas such as the IOD and the IAM may be able to provide the necessary wind speeds to break up biogenic surface slicks. Despite these limitations however, the utility of the scatterometer data to track coincident biological signals to a milky sea represents a new/complementary tool to both study remotely and aid the in-situ study of milky seas.

Future work would involve the creation of a blended scatterometer and Chl-a product from multiple instruments to better track scatterometer wind speed biases across time in milky sea hot

spot regions. Such a product could be able to provide the necessary ‘where’ information which with the ‘when’ information from products such as the SJMI (Chapter 3) to know when and where a milky sea will occur. Investigation into how well synthetic aperture radar can capture and track milky sea associated surface slicks is another topic of future work.

Chapter 6
Sampling Guidance for an Active Milky Sea Event

6.1 Introduction

The results of this dissertation show that milky seas are not only a fantastical visual phenomenon, but also a potentially important facet of the coupled earth system due to their relation to the global carbon cycle. Based on the analysis of eyewitness accounts and satellite observations, the typical milky sea is approximately 10,000 km^2 in size with a lifespan of approximately one to two weeks. This size and lifespan are the key critical factors to consider when deciding how to intercept and sample a milky sea event in-situ.

In this chapter, we will discuss how the results of Chapters 2-5 impact potential attempts to sample a milky sea in-situ. We will also discuss how these results serve to generate guidance both for how to sample and study a milky sea in-situ and what specific facets of a milky sea are of importance to be sampled and studied.

6.2 Anticipation of an Event

Due to the abbreviated (~2 week) moon-free windows for low-light imagers, and the need for two sequential nighttime observations to assess a potential VIIRS DNB feature as a milky sea with high confidence, there is a 2-15 day lag between a milky sea forming and its verification via satellite imagery. As discussed in Chapter 2, this time lag matches the estimated lifetime of a typical milky sea event. Given this timeframe, it is in principle possible to first observe a milky sea event in VIIRS DNB imagery and, with pre-staged equipment and rapid deployment, sample the event. It is also possible, however, that such a deployment could arrive too late to properly study the milky sea event in-situ.

The provisional success of the SJMI in Chapter 3 shows that, despite limited data and a relatively simple model, milky seas near Java are predictable phenomena. In this dissertation we further show that both seasonal and monthly guidance regarding the likelihood of milky sea formation in a given region can be generated with the proper forecasting tools. Based on the connection between milky seas near Java and the IOD, and the connections between the IOD and ENSO, knowledge of whether or not a particular austral winter is a 'good' milky sea season can be known several months in advance. The observed seasonal cycle of milky seas near Java also tells us that within a particular austral winter, August is by far the best time period to observe a milky sea.

Based on this information, field campaigns to study a milky sea in-situ near Java should be organized for, at minimum, the late July through early September time period during a positive IOD/ENSO year. If predictive tools like the SJMI can be further refined and tested with forecast models, more specific/refined guidance regarding when and where to stage resources to sample a milky sea event could be generated.

6.3 Sampling Logistics

Once a research vessel has reached a milky sea event the next question of how to properly sample a milky sea in general and the specific active milky sea event needs to be answered. Milky seas can span several hundred kilometers horizontally, exist at an unknown depth in the water column, and have an unknown thickness vertically. A sampling strategy to properly account for this combination of known unknowns is needed to properly take stock of a milky sea event.

The large spatial scales of a milky sea event are the first challenge to properly sampling a milky sea event. As seen in Chapter 4, the dynamical and biogeochemical properties of a milky

sea event vary through space and time. Sampling would need to be done in various parts of the event with the edges of the milky sea event being of particular interest for the natural flask hypothesis.

Eyewitness accounts report that some milky sea events have sharp demarcation between luminous and non-luminous water, others report a gradual fading in or out of the luminosity, and some individuals report both details for a single milky sea event. Sampling in horizontal transects across the boundaries of a milky sea event would serve to test the nature of the walls of the natural flask containing the milky sea.

Unmanned Underwater Vehicles (UUVs) allow for real-time examination of the water column. A photometer equipped UUV can be used to determine where in the water column the illumination from a milky sea is most concentrated and take water samples. This information coupled with a traditional research vessel simultaneously sampling multiple depths would allow for constructed vertical profiles to be related to the 'core' of the milky sea event. A UUV would also allow for examination of the bottom of the natural flask. The gradient in luminosity with depth would provide insight into whether a milky sea event's vertical extent ends discretely or gradually like the horizontal extent of the milky sea event.

6.4 Specific Measurands

Up to this point we have only spoken of sampling a milky sea in general terms. Little mention has been made of what exactly should be sampled. The main questions discussed in this dissertation so far are the 'when' and 'where' of milky seas, in sampling a milky sea we are interested in the questions of 'who' and how'.

The question of 'who' is behind milky seas, alternatively phrased as what organism or possibly array of organisms are causing this phenomenon is an open question that can best be answered in-situ. Consensus has largely fallen on bacterial bioluminescence and there is a prime suspect in *v. Harveyi* who has been caught, so to speak, red-handed at the scene of the crime (Lapota et al. 1988). Metagenomic analysis of water samples acquired at various depths both interior and exterior to a milky sea would allow for a thorough examination of 'who'. Such analysis would reveal how microbial communities vary in and around the milky sea event greatly refining the list of possible suspects if not outright identifying the culprit or culprits.

Ideally such an analysis would be done for multiple milky sea events. While it is entirely possible that the variations in light qualities such as brightness and coloration between milky sea events is due to facets of human scotopic vision, it is also possible that multiple organisms with different light emission spectra are behind the variety of descriptions for milky sea events. It is also possible that the light emitted by the bioluminescent organisms at some depth below the surface is being processed via spectral absorption by the intervening water or the organic composition within the column. By investigating the make-up of the microbial community for multiple milky sea events we can begin to understand if milky seas are due to one organism with specific environmental conditions for reproductive success or a class of organisms.

Based again on the findings of Lapota et al. (1988), we suspect that the bacteria within a milky sea are forming communities on particulate substrates. Algal organisms such as *Phaeocystis* being one example of a potential substrate. It is unknown if one specific organism acts as a preferred substrate for a milky sea as part of some as of yet unknown symbiotic, saprophytic, or parasitic relationship, or if there is no preferred algal substrate for a milky sea as well as if algae are entirely unneeded and the proliferation of microplastics and other human generated particulates

in the ocean may be an additional source of substrate. Genetic sequencing of filtered particulate matter (such as with 16S and 18S ribosomal RNA sequencing) or microscopic analysis of filtered samples to determine the algal organisms associated with a milky sea is a key measurand for any potential expedition to a milky sea. This identification combined with identification of the bioluminescent organism behind a milky sea would allow for examination of the exact relationship between the bacteria and the underlying substrates supporting their population boom. Such an examination would allow for better understanding as to why milky seas are located adjacent to but not overlapping other significant biological signals in the ocean (e.g. satellite retrieved Chl-a as seen in Chapter 5).

Finally, understanding how the conditions within a milky sea sustain the bioluminescent organism(s) behind the phenomena for extensive periods is another key question of interest. Relating vertical and horizontal profiles of key oceanic variables such as temperature, salinity, pH/pOH, and concentration of nutrients such as phosphate, nitrate, and silicate to vertical profiles of luminosity and information about ecological assemblages at various depths would highlight exact conditions which lead to the organism(s) responsible for a milky sea event proliferating to the population sizes needed.

Understanding the entire ecological assemblage associated with milky seas serves as a window into the role of milky seas in the global carbon cycle. Whether or not the sum respiratory and waste by-products of this ecological assemblage results in more or less carbon being sequestered in the ocean, and eventually entombed in the lithosphere or released back into the atmosphere to accelerate human driven climate change is currently unknown. Only by taking stock of the various pieces of the bioluminescent puzzle that is milky seas can we begin to answer this question.

Chapter 7
Summary and Concluding Thoughts

7.1 Summary

In a single sentence this dissertation can be summed up as an examination of when and where milky seas have occurred as well as when and where they may occur in the near future. In Chapter 2 we combined centuries of historical data with modern satellite imagery to understand when and where milky seas typically form and answer the first of the science questions we posed at the outset of this dissertation. The global distribution of milky seas we identified matches that of previous researchers and updates/augments previous databases—providing a new, curated and indexed repository of reports.

Milky seas occur primarily in the northwestern Indian Ocean and Maritime Continent regions. Combining these spatial locations with when milky seas occurred during the year we extended the anecdotal connections of previous researchers to statistical ones and recognized that beyond the monsoons of the Indian Ocean milky seas are in fact correlated with other coupled air-sea interactions within the earth system such as ENSO and the IOD.

Leveraging this novel milky sea database, in Chapter 3 we formulated the first predictive model for milky seas, the South Java Milky Sea Index (SJMI) addressing our second science question. The SJMI was verified as a skillful hindcasting tool via its ability to identify time periods where milky seas occurred near Java and the discovery of a previously unknown milky sea event near Java in August 2017, MS2017. This finding represents the first known prediction (in this case a hindcast prediction) of a bioluminescent milky sea. The retrospective discovery of MS2017 via the SJMI highlights that milky seas are likely predictable phenomena. Extending the SJMI from hindcasting to forecasting via forecast model output shows that, paired with both an appropriate

source of forecast data and an appropriate metric for the forecast model, milky sea events near Java could potentially be predicted months before their occurrence. This predictability has yet to be confirmed but future research will focus on applying the SJMI in a prognostic context and verifying the predictability of milky seas.

In Chapter 4 we turned our attention towards the question of where milky seas occur and why. Specifically, how might the natural flask hypothesis of milky sea formation explain where milky seas form, what conditions allow them to persist, and what leads to their eventual decay? In addressing this third science question, we analyzed two milky sea events via case study analysis. The first event that we examined, MS2017, formed within quiescent waters dynamically distinct from the surrounding ocean. This event persisted until interacting with the relatively intense currents from the Indonesian Throughflow which mixed the event with the surrounding waters leading to its eventual demise.

The next event that we examined, MS2019, formed within quiescent waters bounded to the south by intense eddies spawned by the Indonesian Throughflow and to the north by strong salinity gradients. Encapsulated in this environment, this event persisted for an anomalously long period of time until, at some unknown date, it dissipated. Over this time period the relative importance of dynamical barriers (e.g. strong currents) and density gradients in supporting the natural flask containing the MS2019 event vacillated. Suggesting that the 'walls' of the natural flask may vary in both space and time over the lifetime of a milky sea event.

Both the MS2017 and MS2019 events show the potential of the natural flask hypothesis to explain how milky seas are sustained and eventually dissipated. One missing piece of the puzzle regarding the natural flask hypothesis is the formation of a milky sea—how does the flask set up and what are the key ingredients that need to be inside of it? To date there is no VIIRS DNB

imagery of a milky sea forming: the ~20 DNB-detected events were already in-progress at the beginning of the moon-free detection period of the lunar cycle. Without the critical information regarding exactly when and where the milky sea event was when it formed, testing how well the natural flask hypothesis explains the formation of a milky sea remains, currently, out of reach.

In Chapter 5 we examined our fourth science question: can additional remote sensing tools beyond low-light imagery be utilized to understand where a milky sea is? Milky seas, like many biological phenomena in the ocean, are associated with algal blooms. Making use of spatially and temporally coherent biases in scatterometer data due to the effect of algal bloom generated surface films on capillary waves we showed that, given a priori knowledge of where a milky sea was, it is possible to track the feature to first-order without access to low-light imagery. Such a capability would allow, for example, a research team to deploy to a location for sampling despite not having current nighttime low-light imagery as a guide—thereby expanding the time window for intercepting an active milky sea. These low wind speed scatterometer biases thus serve as an additional tool in the remote sensing toolbox we use to study milky seas and a source of potential additional guidance.

Finally, in Chapter 6 we examined how the results presented in Chapters 2-5 may inform future attempts to sample a milky sea in-situ. Attempts to properly sample a milky sea must account both for the vast spatial scales of a milky sea and the unknown location in the water column relative to the ocean surface. Similarly, the biological assemblage which makes up a milky sea is largely a mystery and understanding exactly which organisms cause milky seas is both a top priority for a future expedition to study a milky sea and critical to understanding exactly where milky seas fit into the greater coupled earth system.

Collectively, these results increase our knowledge of when and where milky seas have occurred and how coupled air/sea patterns can inform when and where milky seas may form in the future. In the next section we discuss how the results presented in this dissertation inform future directions of research into milky seas.

7.2 Concluding Thoughts and Future Research

Since the 1800s, scientific knowledge and speculation on milky seas has been centered on the theory that they are bacterial in origin, and that milky seas are most likely connected to large scale weather systems such as the Indian Monsoon. Based on the results of this dissertation, we can finally assert that one of these foci of speculation has, for the first time in over a century, been advanced. Milky seas are shown to be a coupled facet of the earth system due to the state of various atmosphere-ocean phenomena such as the IOD, ENSO, and the monsoons of both India and Indonesia/Australia directly influencing when and where they occur. By way of this research to compile and curate the most comprehensive database of milky seas to date, we have been able to confirm that milky seas can be predicted by way of a simple forecast model composed of a weighted linear combination of these aforementioned coupled atmosphere-ocean interactions. However, specific details of milky sea formation which occur on finer spatial and temporal scales require further inquiry.

This dissertation's research has focused on how much of the physical underpinnings of milky seas and their relationship to other facets of the earth system can be ascertained from afar and leveraging these insights to predict milky seas. The traditional rationale for this approach has been that being in the right place at the right time to sample a milky sea is an impractical and arduous task that, up to this point in time, has only been accomplished through pure coincidence.

Sampling a milky sea, particularly in a comprehensive way as described in Chapter 6, would allow for greater insight into the physical underpinnings of milky seas and their relationship to other facets of the earth system expanding the ability to model and predict milky seas. Therein lies the core difficulty and inherent challenge of studying milky seas.

To truly understand milky sea composition, structure, and their role in nature, we need to be able to sample a milky sea. To reliably and practically sample a milky sea we need, in turn, deep insight and understanding of milky seas. This situation resembles the classic problem of the chicken and the egg and has been where much of the previous scholarship on milky seas has languished. New-generation satellite based low-light imagers such as the VIIRS DNB, coupled with both the vast array of remote sensing tools and the copious array of atmospheric and oceanic datasets presently available, have helped us 'sample' properties of a milky sea remotely and break this cycle.

In Chapter 5 we investigated how non-visible signals related to milky seas can be detected from space, extending the trackable horizon for known/active events. The study of milky seas beyond historical reports is still a very new scientific topic. With an ever-increasing variety of instruments, sensors, and experiments of opportunity occurring, it is possible that some important facet of milky seas remains undiscovered. Examining how the remote sensing toolbox can be leveraged to study milky seas is still a largely open topic with little examination by previous research.

This dissertation opened with a quote by Carl Sagan: *"But nature is always more subtle, more intricate, more elegant than what we are able to imagine."* This quote represents the direction future research into milky seas must focus. Investigating the subtle, intricate, and elegant connections between milky seas and the rest of the coupled earth system.

The conditions that we observe as being associated with many milky seas, quiescent waters rich in biological activity bounded by strong oceanic currents are common in the regions where milky seas occur most often. However, similar conditions exist in many other regions globally. The ITF and the seasonally reoccurring currents in the northwest Indian Ocean routinely trap waters that, ostensibly, should be primed for milky seas against the coasts of Somalia and Java. Despite this, milky seas are a rare phenomenon. Understanding the subtle difference between which environments do and do not produce a milky sea event should be the focus of future research into milky seas.

Understanding these distinctions, what causes them, and how to predict them is one of the main topics of future research raised by this dissertation. While the data on milky seas themselves are limited, data on ancillary things such as atmospheric and oceanic state are plentiful. Modern advances in data science have revealed a wealth of statistical and machine learning techniques for sifting through figurative mountains of data generated about the earth system.

Another topic for future research is how will milky seas respond to climate change. Depending on how the organisms which cause milky seas respond to an increasingly acidic and warmer ocean due to anthropogenic carbon emissions, we could be entering an era where milky seas become less common. Conversely, we could be entering an era where their frequency greatly increases as the oceans become a more hospitable environment.

Milky seas are a large, visible signal regarding the health of the ecosystems in which they occur. Whether they represent a healthy productive ecosystem or one that is stressed is presently unknown. Further study of milky seas both remotely and, if possible, in-situ is required to understand and decipher what exactly this glowing message means about the state of the ecosystems in which milky seas occur both now, and in the future.

In the 32 years since Herring and Watson published their paper “Milky Seas: a Bioluminescent Puzzle”, our understanding of the coupled earth system has greatly advanced. For example, the IOD has been recognized as a distinct phenomenon versus an extension of ENSO. Without this expanded understanding constructing the predictive model in Chapter 3 might not have been possible. Improving our understanding of the earth system to find the subtle, intricate, and elegant connections between milky seas and other facets of the earth system and leveraging our ever-expanding understanding of the earth system is where future research into milky seas and other rare biosphere phenomena lies.

References

Alpers, W., and H. Hühnerfuss, 1989: The damping of ocean waves by surface films: A new look at an old problem. *J. Geophys. Res.*, 94, 6251–6265, https://doi.org/10.1029/JC094iC05p06251.

Antunes, L. C. M., R. B. R. Ferreira, M. M. C. Buckner, and B. B. Finlay, 2010: Quorum sensing in bacterial virulence. *Microbiology*, 156, 2271–2282, https://doi.org/10.1099/mic.0.038794-0.

Ashok, K., Z. Guan, and T. Yamagata, 2001: Impact of the Indian Ocean dipole on the relationship between the Indian monsoon rainfall and ENSO. *Geophysical Research Letters*, 28, 4499–4502, https://doi.org/10.1029/2001GL013294.

Battisti, D. S., and A. C. Hirst, 1989: Interannual Variability in a Tropical Atmosphere–Ocean Model: Influence of the Basic State, Ocean Geometry and Nonlinearity. *J. Atmos. Sci.*, 46, 1687–1712, https://doi.org/10.1175/1520-0469(1989)046<1687:IVIATA>2.0.CO;2.

Bleck, R., 2002: An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Modelling*, 4, 55–88, https://doi.org/10.1016/S1463-5003(01)00012-9.

Buist, G., 1855: Notes on Certain Discoloured Appearances Met With on the Surface of the Sea in Warm Latitudes. *Proceedings of the Bombay Geographical Society*, 108–125.

Cai, W., A. Sullivan, and T. Cowan, 2009: Rainfall Teleconnections with Indo-Pacific Variability in the WCRP CMIP3 Models. *Journal of Climate*, 22, 5046–5071, https://doi.org/10.1175/2009JCLI2694.1.

Capt. Newland, 1772: XIII. Observations on the milky appearance of some spots of water in the sea; by the same. *Phil. Trans. R. Soc.*, 62, 93–94, https://doi.org/10.1098/rstl.1772.0014.

Cerdeiro, D., A. Komaromi, Y. Liu, and M. Saeed, 2020: World Seaborne Trade in Real Time: A Proof of Concept for Building AIS-based Nowcasts from Scratch. *IMF Working Papers*, 2020, 1, https://doi.org/10.5089/9781513544106.001.

Chen, G., W. Han, Y. Li, and D. Wang, 2016: Interannual Variability of Equatorial Eastern Indian Ocean Upwelling: Local versus Remote Forcing. *Journal of Physical Oceanography*, 46, 789–807, https://doi.org/10.1175/JPO-D-15-0117.1.

Dahlgren, U., 1915: The production of light by animals. *Journal of the Franklin Institute*, 180, 513–537, https://doi.org/10.1016/S0016-0032(15)90523-9.

Fan, Q., J. Zuo, H. Wang, D. Grenier, L. Yi, and Y. Wang, 2022: Contribution of quorum sensing to virulence and antibiotic resistance in zoonotic bacteria. *Biotechnology Advances*, 59, 107965, https://doi.org/10.1016/j.biotechadv.2022.107965.

Fischer, B., 1887: *Ueber Einen Lichtentwickelnden, Im Meerwasser Gefundenen Spaltpilz*.

Franklin, B., 1753: From Benjamin Franklin to Peter Collinson, September 1753.

——, W. Brownrigg, and Farish, 1774: XLIV. Of the stilling of waves by means of oil. Extracted from sundry letters between Benjamin Franklin, LL. D. F. R. S. William Brownrigg, M. D. F. R. S. and the Reverend Mr. Farish. *Phil. Trans. R. Soc.*, 64, 445–460, https://doi.org/10.1098/rstl.1774.0044.

Garg, P., S. W. Nesbitt, T. J. Lang, G. Priftis, T. Chronis, J. D. Thayer, and D. A. Hence, 2020: Identifying and Characterizing Tropical Oceanic Mesoscale Cold Pools using Spaceborne Scatterometer Winds. *JGR Atmospheres*, 125, e2019JD031812, https://doi.org/10.1029/2019JD031812.

Haddock, S. H. D., M. A. Moline, and J. F. Case, 2010: Bioluminescence in the Sea. *Annu. Rev. Mar. Sci.*, 2, 443–493, https://doi.org/10.1146/annurev-marine-120308-081028.

Harvey, E. N., 1920: *The Nature of Animal Light*. J. B. Lippincott Company.

Hashizume, H., and W. T. Liu, 2004: Systematic error of microwave scatterometer wind related to the basin-scale plankton bloom. *Geophysical Research Letters*, 31, https://doi.org/10.1029/2003GL018941.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Q.J.R. Meteorol. Soc.*, 146, 1999–2049, https://doi.org/10.1002/qj.3803.

——, and Coauthors, 2023: ERA5 Monthly Averaged Data on Pressure Levels from 1940 to Present. Copernicus Climate Change Service (C3S), https://doi.org/10.24381/cds.6860a573.

Huang, B., C. Liu, V. Banzon, E. Freeman, G. Graham, B. Hankins, T. Smith, and H.-M. Zhang, 2021: Improvements of the Daily Optimum Interpolation Sea Surface Temperature (DOISST) Version 2.1. *Journal of Climate*, 34, 2923–2939, https://doi.org/10.1175/JCLI-D-20-0166.1.

Hudson, J., and S. D. Miller, 2025: From Sailors to Satellites: A Curated Database of Bioluminescent Milky Seas Spanning 1600-Present. *Earth and Space Science*, 12, e2024EA004082, https://doi.org/10.1029/2024EA004082.

Jiao, N., and Coauthors, 2024: The microbial carbon pump and climate change. *Nat Rev Microbiol*, 22, 408–419, https://doi.org/10.1038/s41579-024-01018-0.

Kajikawa, Y., B. Wang, and J. Yang, 2010: A multi-time scale Australian monsoon index. *Intl Journal of Climatology*, 30, 1114–1120, https://doi.org/10.1002/joc.1955.

Katz, O., 1888: Preliminary Remarks on Phosphorescent Bacteria from Sea-Water. *Proceedings of the Linnean Society of New South Wales*, 12, 331–336.

Kelly, K. A., S. Dickinson, M. J. McPhaden, and G. C. Johnson, 2001: Ocean currents evident in satellite wind data. *Geophysical Research Letters*, 28, 2469–2472, https://doi.org/10.1029/2000GL012610.

Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction. *Bull. Amer. Meteor. Soc.*, 95, 585–601, https://doi.org/10.1175/BAMS-D-12-00050.1.

Kong, K.-F., C. Vuong, and M. Otto, 2006: Staphylococcus quorum sensing in biofilm formation and infection. *International Journal of Medical Microbiology*, 296, 133–139, https://doi.org/10.1016/j.ijmm.2006.01.042.

Lapota, D., C. Galt, J. R. Losee, H. D. Huddell, J. K. Orzech, and K. H. Nealson, 1988: Observations and measurements of planktonic bioluminescence in and around a milky sea. *Journal of Experimental Marine Biology and Ecology*, 119, 55–81, https://doi.org/10.1016/0022-0981(88)90152-9.

Le Moigne, F. A. C., 2019: Pathways of Organic Carbon Downward Transport by the Oceanic Biological Carbon Pump. *Front. Mar. Sci.*, 6, 634, https://doi.org/10.3389/fmars.2019.00634.

Leblanc, K., and Coauthors, 2018: Nanoplanktonic diatoms are globally overlooked but play a role in spring blooms and carbon export. *Nat Commun*, 9, 953, https://doi.org/10.1038/s41467-018-03376-9.

Lewis, C. L., C. C. Craig, and A. G. Senecal, 2014: Mass and Density Measurements of Live and Dead Gram-Negative and Gram-Positive Bacterial Populations. *Appl Environ Microbiol*, 80, 3622–3631, https://doi.org/10.1128/AEM.00117-14.

Lin, H., and Coauthors, 2020: The Canadian Seasonal to Interannual Prediction System Version 2 (CanSIPSv2). *Weather and Forecasting*, 35, 1317–1343, https://doi.org/10.1175/WAF-D-19-0259.1.

Lin, I.-I., W. Alpers, and W. T. Liu, 2003: First evidence for the detection of natural surface films by the QuikSCAT scatterometer. *Geophysical Research Letters*, 30, https://doi.org/10.1029/2003GL017415.

Lindsley, R. D., and D. G. Long, 2012: Mapping Surface Oil Extent From the Deepwater Horizon Oil Spill Using ASCAT Backscatter. *IEEE Trans. Geosci. Remote Sensing*, 50, 2534–2541, https://doi.org/10.1109/TGRS.2011.2174369.

Liu, Q., M. Feng, D. Wang, and S. Wijffels, 2015: Interannual variability of the Indonesian Throughflow transport: A revisit based on 30 year expendable bathythermograph data. *JGR Oceans*, 120, 8270–8282, https://doi.org/10.1002/2015JC011351.

Macartney, J., 1810: Observations upon Luminous Animals. *Philosophical Transactions of the Royal Society of London*, 100, 258–293.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 Day Oscillation in the Zonal Wind in the Tropical Pacific. *Journal of the Atmospheric Sciences*, 28, 702–708, https://doi.org/10.1175/1520-0469(1971)028<0702:DOADOI>2.0.CO;2.

——, and ——, 1972: Description of Global-Scale Circulation Cells in the Tropics with a 40–50 Day Period. *Journal of the Atmospheric Sciences*, 29, 1109–1123, https://doi.org/10.1175/1520-0469(1972)029<1109:DOGSCC>2.0.CO;2.

Mars Brisbin, M., S. Mitarai, M. A. Saito, and H. Alexander, 2022: Microbiomes of bloom-forming *Phaeocystis* algae are stable and consistently recruited, with both symbiotic and opportunistic modes. *The ISME Journal*, 16, 2255–2264, https://doi.org/10.1038/s41396-022-01263-2.

Miller, M. B., and B. L. Bassler, 2001: Quorum Sensing in Bacteria. *Annu. Rev. Microbiol.*, 55, 165–199, https://doi.org/10.1146/annurev.micro.55.1.165.

Miller, S. D., S. H. D. Haddock, C. D. Elvidge, and T. F. Lee, 2005: Detection of a bioluminescent milky sea from space. *Proc. Natl. Acad. Sci. U.S.A.*, 102, 14181–14184, https://doi.org/10.1073/pnas.0507253102.

——, S. P. Mills, C. D. Elvidge, D. T. Lindsey, T. F. Lee, and J. D. Hawkins, 2012: Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities. *Proc. Natl. Acad. Sci. U.S.A.*, 109, 15706–15711, https://doi.org/10.1073/pnas.1207034109.

——, and Coauthors, 2013: Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. *Remote Sensing*, 5, 6717–6766, https://doi.org/10.3390/rs5126717.

——, W. C. Straka, J. Yue, S. M. Smith, M. J. Alexander, L. Hoffmann, M. Setvák, and P. T. Partain, 2015: Upper atmospheric gravity wave details revealed in nightglow satellite imagery. *Proc. Natl. Acad. Sci. U.S.A.*, 112, https://doi.org/10.1073/pnas.1508084112.

——, S. H. D. Haddock, W. C. Straka, C. J. Seaman, C. L. Combs, M. Wang, W. Shi, and S. Nam, 2021: Honing in on bioluminescent milky seas from space. *Sci Rep*, 11, 15443, https://doi.org/10.1038/s41598-021-94823-z.

NASA Ocean Biology Processing Group, 2022: Suomi-NPP VIIRS Level 3 Mapped Chlorophyll Data, Version R2022.0. NASA Ocean Biology Distributed Active Archive Center, accessed 30 April 2025, https://doi.org/10.5067/SUOMI-NPP/VIIRS/L3M/CHL/2022.

Nealson, K. H., and J. W. Hastings, 1979: Bacterial bioluminescence: its control and ecological significance. *Microbiol Rev*, 43, 496–518, https://doi.org/10.1128/mr.43.4.496-518.1979.

Nealson, K. H., and J. W. Hastings, 2006: Quorum Sensing on a Global Scale: Massive Numbers of Bioluminescent Bacteria Make Milky Seas. *Appl Environ Microbiol*, 72, 2295–2297, https://doi.org/10.1128/AEM.72.4.2295-2297.2006.

Olejarz, J., Y. Iwasa, A. H. Knoll, and M. A. Nowak, 2021: The Great Oxygenation Event as a consequence of ecological dynamics modulated by planetary change. *Nat Commun*, 12, 3985, https://doi.org/10.1038/s41467-021-23286-7.

Ramaiah, N., and D. Chandramohan, 1987: Distribution & Species Composition of Planktonic Luminous Bacteria in the Arabian Sea. *Indian Journal of Marine Scienes*, 16, 139–142.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, 108, 2002JD002670, https://doi.org/10.1029/2002JD002670.

Ribal, A., A. Tamizi, and I. R. Young, 2021: Calibration of scatterometer wind speed under hurricane conditions. *Journal of Atmospheric and Oceanic Technology*, https://doi.org/10.1175/JTECH-D-21-0055.1.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature*, 401, 360–363, https://doi.org/10.1038/43854.

Seuront, L., D. Vincent, and J. G. Mitchell, 2006: Biologically induced modification of seawater viscosity in the Eastern English Channel during a Phaeocystis globosa spring bloom. *Journal of Marine Systems*, 61, 118–133, https://doi.org/10.1016/j.jmarsys.2005.04.010.

Shi, W., and M. Wang, 2021: A biological Indian Ocean Dipole event in 2019. *Sci Rep*, 11, 2452, https://doi.org/10.1038/s41598-021-81410-5.

——, and ——, 2022: Biological dipole mode indices: New parameters to characterize the physical and biological processes of the Indian Ocean Dipole event. *Progress in Oceanography*, 206, 102847, https://doi.org/10.1016/j.pocean.2022.102847.

——, and ——, 2024: A negative biological Indian Ocean dipole event in 2022. *Sci Rep*, 14, 1110, https://doi.org/10.1038/s41598-024-51347-6.

Shin, J. C., H. Yaguchi, and S. Shioiri, 2004: Change of Color Appearance in Photopic, Mesopic and Scotopic Vision. *OPT REV*, 11, 265–271, https://doi.org/10.1007/s10043-004-0265-2.

Smith, H. T., 1926: Phosphorescence of the Sea. *Marine Observer*, III, 193–196.
——, 1931: Phosphorescence Of The Sea. *Marine Observer*, VIII, 230–234.

Strachan, M., and B. Penrose, 1971: Keeling's Journal. *The East India Company Journals of Captain William Keeling and Master Thomas Bonner, 1615-1617*, University of Minnesota Press, p. 95.

Tachard, G., 1688: *A Relation of the Voyage to Siam Performed by Six Jesuits, Sent by the French King, to the Indies and China in the Year, 1685. With Their Astrological Observations, and their Remarks of Natural Philosophy, Geography, Hydrography, and History*.

Takemura, A. F., D. M. Chien, and M. F. Polz, 2014: Associations and dynamics of Vibrionaceae in the environment, from the genus to the population level. *Front. Microbiol.*, 5, https://doi.org/10.3389/fmicb.2014.00038.

Teeling, H., and Coauthors, 2012: Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. *Science*, 336, 608–611, https://doi.org/10.1126/science.1218344.

Turner, R., 1965: *Notes on the Nature and Occurrence of Marine Bioluminescent Phenomena*.

Ummenhofer, C. C., A. Sen Gupta, Y. Li, A. S. Taschetto, and M. H. England, 2011: Multi-decadal modulation of the El Niño–Indian monsoon relationship by Indian Ocean variability. *Environ. Res. Lett.*, 6, 034006, https://doi.org/10.1088/1748-9326/6/3/034006.

Verity, P. G., C. P. Brussaard, J. C. Nejstgaard, M. A. Van Leeuwe, C. Lancelot, and L. K. Medlin, 2007: Current understanding of Phaeocystis ecology and biogeochemistry, and perspectives for future research. *Biogeochemistry*, 83, 311–330, https://doi.org/10.1007/s10533-007-9090-6.

Watson, M., and P. J. Herring, 1992: A new database for observations of bioluminescence. *The Marine Observer*, 62, 182–183.

Webster, P. J., and S. Yang, 1992: Monsoon and Enso: Selectively Interactive Systems. *Q.J Royal Met. Soc.*, 118, 877–926, https://doi.org/10.1002/qj.49711850705.

Wen, C., Z. Wang, J. Wang, H. Li, X. Shi, W. Gao, and H. Huang, 2023: Variation of the coastal upwelling off South Java and their impact on local fishery resources. *J. Ocean. Limnol.*, 41, 1389–1404, https://doi.org/10.1007/s00343-022-2031-3.

Wirasatriya, A., J. D. Setiawan, D. N. Sugianto, I. A. Rosyadi, H. Haryadi, G. Winarso, R. Y. Setiawan, and R. D. Susanto, 2020: Ekman dynamics variability along the southern coast of Java revealed by satellite data. *International Journal of Remote Sensing*, 41, 8475–8496, https://doi.org/10.1080/01431161.2020.1797215.

——, R. D. Susanto, K. Kunarso, Abd. R. Jalil, F. Ramdani, and A. D. Puryajati, 2021: Northwest monsoon upwelling within the Indonesian seas. *International Journal of Remote Sensing*, 42, 5433–5454, https://doi.org/10.1080/01431161.2021.1918790.

Wyrtki, K., 1961: *Physical Oceanography of the Southeast Asian Waters*, https://escholarship.org/uc/item/49n9x3t4.

Xu, S., X. Wang, J. Liu, F. Zhou, K. Guo, S. Chen, Z. Wang, and Y. Wang, 2022: Bacteria Associated With Phaeocystis globosa and Their Influence on Colony Formation. *Front. Microbiol.*, 13, 826602, https://doi.org/10.3389/fmicb.2022.826602.