

University of Massachusetts Boston

ScholarWorks at UMass Boston

Graduate Doctoral Dissertations

Doctoral Dissertations and Masters Theses

12-2020

Novel Tools and Techniques to Investigate and Reduce the Impacts of Capture-and-Handling

Ryan J. Knotek

University of Massachusetts Boston

Follow this and additional works at: https://scholarworks.umb.edu/doctoral_dissertations



Part of the [Environmental Sciences Commons](#)

Recommended Citation

Knotek, Ryan J., "Novel Tools and Techniques to Investigate and Reduce the Impacts of Capture-and-Handling" (2020). *Graduate Doctoral Dissertations*. 630.

https://scholarworks.umb.edu/doctoral_dissertations/630

This Open Access Dissertation is brought to you for free and open access by the Doctoral Dissertations and Masters Theses at ScholarWorks at UMass Boston. It has been accepted for inclusion in Graduate Doctoral Dissertations by an authorized administrator of ScholarWorks at UMass Boston. For more information, please contact library.uasc@umb.edu.

NOVEL TOOLS AND TECHNIQUES TO INVESTIGATE AND REDUCE THE
IMPACTS OF CAPTURE-AND-HANDLING

A Dissertation Presented
by
RYAN J. KNOTEK

Submitted to the Office of Graduate Studies,
University of Massachusetts Boston,
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2020

Marine Science and Technology Program

© 2020 by Ryan J. Knotek
All rights reserved

NOVEL TOOLS AND TECHNIQUES TO INVESTIGATE AND REDUCE THE
IMPACTS OF CAPTURE-AND-HANDLING

A Dissertation Presented

by

RYAN J. KNOTEK

Approved as to style and content by:

John Mandelman, Adjunct Faculty
Chairperson of Committee

Robert F. Chen, Professor
Member

Jeff Kneebone, Research Scientist
Anderson Cabot Center for Ocean Life, New England Aquarium
Member

Tobey Curtis, Fishery Management Specialist
Atlantic Highly Migratory Species Management Division, NOAA NMFS
Member

Edd Brooks, Chief Executive Officer
Cape Eleuthera Island School
Member

Juanita Urban-Rich, Program Director
Environmental Sciences, Marine Science & Technology Program

Robert F. Chen, Interim Dean
School for the Environment

ABSTRACT

NOVEL TOOLS AND TECHNIQUES TO INVESTIGATE AND REDUCE THE IMPACTS OF CAPTURE-AND-HANDLING

December 2020

Ryan J. Knotek, B.S., University of New England
M.S., University of New England
M.S., University of Massachusetts Boston
Ph.D., University of Massachusetts Boston

Directed by Professor John Mandelman

Fish have been an important food resource for humans throughout history, but with growing populations increasing the demand for this resource, many fish stocks have become overfished. In response, management has traditionally addressed overfishing by establishing regulations that reduce the directed fishing mortality associated with harvesting. However, these practices often do not take into account the mortality associated with fish that are incidentally captured and discarded (as bycatch). This discard mortality (DM) can represent a potentially large source of removals for species that are particularly susceptible to the stressors of capture-and-handling and/or discarded at high rates. It is therefore vital to understand overall DM in a fishery and the factors associated with capture-and-handling

driving mortality. But this information can be very challenging to collect, with many species/fisheries having yet to be addressed. In light of these difficulties, my dissertation aimed to develop novel tools and techniques that improve our ability to evaluate and/or reduce the impacts associated with fisheries-interactions. In particular, this included (1) the development of a technique that addresses capture behavior as an underlying driver of mortality, (2) a novel application of electronic tags to better deduce fate in a family of frequently discarded fish, and (3) a new tool for forecasting where fisheries-interactions (i.e., incidental capture) are most likely to occur in both time and space. The information provided by these tools and techniques will help inform stock assessments and bycatch mitigation strategies, which in turn will help to promote the sustainability of bycatch species. Moreover, two species of elasmobranch [thorny skate (*Amblyraja radiata*) and oceanic whitetip shark (*Carcharhinus longimanus*)] were used as case study species in this dissertation. Both have been heavily overfished, but the impact of capture-and-handling remains poorly understood for each species. Thus, my dissertation contributes critical capture-related information for thorny skate and oceanic whitetip shark, which could aid in the development of future mitigation strategies to promote their recovery.

ACKNOWLEDGEMENTS

I would first like to thank, John Mandelman, for his continued guidance from serving on my thesis committee (University of England) to being my advisor here at the University of Massachusetts Boston. I would not be where I am today without your mentorship and I cannot thank you enough for putting up with me for more than eight years now! I also want to thank my committee members for helping me navigate through this degree. To Jeff Kneebone, thank you for many, many things, from showing me the ropes in the field on various projects, to your patience and help with my writing. To Edd Brooks, thank you for the huge opportunity to work with you and your team on the longstanding oceanic whitetip shark research effort, and your continued guidance with those research projects. To Tobey Curtis, thank you for your help with all things thorny skate and management-related, but I do not thank you for your poor choice in sports jersey on the day of my defense! And to Bob Chen, thank you for challenging me to think outside of my fishery-specific world and look at research from a much more transdisciplinary vantage point. All of you have played a huge role in shaping me into the scientist I am today and I thank you over and over again.

I also want to thank the research team from the Anderson Cabot Center for Ocean Life (ACCOL) at the New England Aquarium. In particular, Connor Capizzano, who has been a great friend and fellow student throughout both my Master's and Doctoral programs. It's been quite the road getting here and I'm glad we got to traverse it all together. I also want to thank Emily Jones, who was always there to lend a helping hand, keep me on track with

paperwork, or maybe most importantly, provide the best baked goods. Thank you to Nick Whitney as well, for his guidance on a separate research project, exceptional sarcasm, and help in developing grant proposals for future work that I hope we get to conduct. In addition, thank you to all of the other folks from ACCOL that I have had the pleasure of working with at the New England Aquarium. Thank you as well to the many people involved with the Oceanic Whitetip Consortium that I have also had the pleasure of working with over the past four years. I look forward to continuing to work with you all and collaborating on future research efforts.

I also want to give a special thanks to my wife and family for always supporting me throughout my pursuit of higher education and putting up with my incessant love for the ocean and fishing. More specifically, to my parents, thank you for always encouraging me to do what I love and follow my passions. To my brother, thank you for sending me daily shark-related news articles and updates, and always being down to talk sharks. To my sister, thank you for briefly being a part of my world of research, it led to some great days out on the water. To my wife, there is no way I could ever thank you enough for dealing with my never ending status of being a student and being my rock through it all. And last, but certainly not least, I would like to give a shout out to my newborn son, Jack. I cannot wait to share with you my passion for the ocean and I dedicate this dissertation to you. Love you, little man.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	vi
LIST OF TABLES	x
LIST OF FIGURES	xi
CHAPTER	Page
INTRODUCTION	1
1. UTILIZATION OF POP-UP SATELLITE ARCHIVAL TRANSMITTING TAGS TO EVALUATE THORNY SKATE (<i>AMBLYRAJA RADIATA</i>) DISCARD MORTALITY IN THE GULF OF MAINE GROUND FISH BOTTOM TRAWL FISHERY	14
Abstract	14
Introduction	15
Methods	18
Results	27
Discussion	30
Acknowledgements	36
Literature cited	38
2. DESCRIPTION OF A BYCATCH “HOTSPOT” AVOIDANCE TOOL AND ITS APPLICATION TO THORNY SKATE, <i>AMBLYRAJA RADIATA</i> , IN THE U.S. NORTHWEST ATLANTIC OCEAN	51
Abstract	51
Introduction	52
Methods	54
Results	62
Discussion	63
Acknowledgements	71
Literature cited	72
3. MERGING TECHNOLOGIES AND SUPERVISED CLASSIFICATION METHODS TO QUANTIFY CAPTURE BEHAVIOR ON HOOK-AND-LINE	85
Abstract	85
Introduction	86
Methods	89
Results	96

CHAPTER	Page
Discussion.....	97
Acknowledgements.....	103
Literature cited.....	104
SYNTHESIS AND FUTURE DIRECTIONS	120
APPENDIX	
A. CHAPTER 1 SUPPLEMENTARY MATERIAL	126
B. CHAPTER 3 SUPPLEMENTARY MATERIAL	134
BIOGRAPHICAL SKETCH OF AUTHOR	148

LIST OF TABLES

Table	Page
1.1. Injury code classification	43
1.2. Survival mixture model results	44
1.3. Thorny skate sample sizes	45
1.4. Fishing conditions, practices, and biological characteristics	46
2.1. Vessel-speed thresholds	77
2.2. Hotspot characteristics by evaluation period for each fishery examined in the U.S. Northwest Atlantic Ocean.....	78
3.1. Final ensemble weights assigned to predictions from each supervised classification method	115
3.2. F1-measure performances	116

LIST OF FIGURES

Figure	Page
1.1. Tow locations.....	47
1.2. Pop-up satellite archival transmitting tag attachment method.....	48
1.3. Depth time-series of two mortality events.....	49
1.4. Probability of survival over time for thorny skate with respect to size-category.....	50
2.1. Diagram of the bycatch hotspot avoidance tool workflow.....	79
2.2. Relevant spatial extent of potential fisheries-interactions.....	80
2.3. Thorny skate fishery-interaction potential hotspots in the Northeast multispecies fishing industry.....	81
2.4. Thorny skate (51 to 70 cm total length) fishery-interaction potential hotspots in the Northeast multispecies fishing industry.....	82
2.5. Thorny skate fishery-interaction potential hotspots in the monkfish fishing industry.....	83
2.6. Thorny skate fishery-interaction potential hotspots in the Atlantic sea scallop fishing industry.....	84
3.1. Diagram of the modified handline.....	117
3.3. Framework of the phased-workflow in the capture behavior analysis ...	118
3.4. Raw tri-axial acceleration data and capture behaviors.....	119
4.1. Key components in the “toolbox” researchers have when addressing the impact of capture-and-handling on bycatch species.....	125

INTRODUCTION

Elasmobranchs (sharks, skates, and rays) have functioned in marine ecosystems for at least 420 million years, playing an integral role in the trophic ecology and stability of an ecosystem (Camhi et al., 2008; Huelgel et al., 2014). These fishes are k-selected strategists that exhibit late maturity, long lifespans, low fecundity, and extended gestation periods (Dulvy et al., 2008). But while these life-history characteristics have contributed to the evolutionary success of this group, low reproductive potential and intrinsic rate of population growth make elasmobranchs particularly vulnerable to anthropogenic stressors (e.g., climate change and fishing pressure; Worm and Branch, 2013). In particular, sharks, skates, and rays face fishing pressure from recreational and commercial industries as targeted species and/or incidental bycatch. Landings of elasmobranchs are primarily tied to commercial “meat” markets or the global shark fin trade (Dent and Clarke, 2015), but more often they are discarded due to regulatory measures that prohibit landings or low economic value relative to other species (Molina and Cooke, 2012). Collectively, fishing pressures have led to overfishing and large-scale population declines for this group of fishes, with nearly one-quarter of elasmobranch species facing some level of extinction risk (Dulvy et al., 2008; Ferretti et al., 2010).

Management of elasmobranchs and other fish is largely predicated upon an identification of all sources of input (e.g., recruitment and immigration) and output (e.g. emigration, natural and fishing mortality) influencing stock dynamics (Hoggarth et al., 2006). This includes direct fishing mortality associated with landings, but also the mortality of discarded fish (Crowder and Murawski, 1998). However, discard mortality (DM) is much more challenging to address and is often unknown for many species (Crowder and Murawski, 1998). This presents a major problem for stock assessors and management bodies because DM can potentially represent a large source of removals, particularly if a species is frequently discarded and subject to high DM rates (Hoggarth et al., 2006). It is therefore vital to gain a better understanding of the mortality of discarded fish and the factors associated with capture-and-handling that impact animal health/fate. This information will yield more accurate stock assessments and provide management and other end-users with the necessary information to develop bycatch mitigation strategies that reduce fishing mortality (Crowder and Murawski, 1998; Hall and Mainprize, 2005).

The capture-and-handling process is an inherently stressful event that can lead to physical trauma and blood chemistry alterations that disrupt homeostasis, inciting a cascade of adaptive physiological responses to return an organism back to its original state (Chrousos, 1998; Wendelaar-Bonga, 1998; Barton, 2002; Davis, 2002; Skomal and Mandelman, 2012). However, if physical trauma and/or physiological perturbations are too severe or unresolved, organisms can experience longer term health consequences (e.g., impaired growth, reproduction, and immune system) or even mortality (Barton, 2002; Skomal and Mandelman, 2012). The factors associated with capture events that can lead to health consequences or

mortality includes fishing conditions and practices and biological characteristics (Davis, 2002; Broadhurst et al., 2006; Veldhuizen et al., 2018). In particular, extended capture duration is often cited with links to increased physical trauma and physiological stress because of prolonged bouts of struggling and impaired respiration (i.e., metabolic and respiratory acidosis; Wood et al., 1993; Dapp et al., 2016; Ellis et al., 2017). Catch composition (e.g., rocks or fish with spines) and larger catch biomasses can also exacerbate physical damage and lead to hypoxic conditions in mobile fishing gear (Davis, 2002). In addition, extreme temperature gradients (between seawater and air) can incite thermal stress due to elevated reaction rates (Gale et al., 2013), while prolonged air exposure could collapse gill structures and impede oxygen uptake (Arlinghaus, 2007). Biological characteristics also dictate the impact of capture-and-handling on animal health/fate. For instance, smaller animals can be more susceptible to physical trauma in mobile gear given their less robust morphology (relative to larger conspecifics; Davis, 2002; Ellis et al., 2017).

The aforementioned capture-related factors were presented in terms of how they generally impact fish, but ultimately, impacts can be highly variable between species and fisheries (e.g., physiological outcomes and mortality rates; Mandelman and Skomal, 2009; Veldhuizen et al., 2018; Musyl and Gilman, 2019). It is speculated that these differences attribute to underlying mechanisms that influence how an animal reacts and responds to the stressors of capture (e.g., metabolic scope, ventilation strategy, and capture behavior; Skomal and Bernal, 2010; Dapp et al., 2016; Gallagher et al., 2017; Bouyoucos et al., 2018). However, these mechanisms have received minimal direct research attention and there is little to no consensus on the true role of each underlying mechanism.

Though our understanding of fisheries impact on elasmobranchs has evolved over the years, many data gaps still exist that require additional research efforts. To investigate capture-and-handling, researchers can match various tools and techniques (i.e., the “toolbox”) to study objectives on a case-by-case basis in order to provide the most relevant information for a given species/fishery. This toolbox can be categorized into tools and techniques that measure factors associated with capture-and-handling, and those that can quantify animal health and mortality (Davis, 2002; Pollock and Pine, 2007). Most fishing conditions and practices and biological characteristics are simple to record (e.g., visual quantification of catch composition, handling method, weather conditions, and sea state), but other factors require additional instrumentation. Hook timers are needed to address capture duration (hook-and-line fisheries) and gear-borne temperature loggers are required to measure bottom seawater temperature in demersal fisheries. In addition, while maturation state is straightforward for male elasmobranchs (i.e., calcified claspers), ultrasonography or steroid hormone analyses are often needed to assess maturation/gravidity of females (Gelsleichter et al., 2002; Daly et al., 2007). Moreover, some factors remain incredibly challenging to evaluate with the tools and techniques available to date.

To evaluate the impact of capture-and-handling on the immediate health of an animal, researchers can choose from various health indicators that focus on either the degree of overt physical trauma, activity level, or a combination of the two (i.e., a vitality index; Davis, 2002; Ellis et al., 2017). In addition, phlebotomy and blood chemistry analyses can be incorporated to directly evaluate the degree of physiological stress (Skomal and Mandelman, 2012; Ellis et al., 2017). Researchers could also test reflex impairment as a proxy for a

compromised physiological state (Raby et al., 2012; Davis, 2010), but little work has been done with elasmobranchs to validate the connection between reflexes and actual physiological status (Davis, 2010).

Evaluating animal fate can also be accomplished with a variety of tools and techniques. The simplest form of fate monitoring relies on confinement periods (e.g., on-board holding tank or sea cages; Portz et al., 2006) post-capture, with mortalities visually identified either at the end of confinement or intervals throughout (Benoît et al., 2012; Mandelman et al., 2013; Sulikowski et al., 2020). However, confinement studies can produce misleading results if fish exhibit confinement stress and/or shielding from predation precludes accounting for a major source of mortality (Davis, 2002; Laptikhovsky, 2004; Portz et al., 2006). Tagging studies (conventional or electronic) offer an alternative form of monitoring that allows researchers to evaluate fate under more realistic discarding conditions where animals can swim freely and interact with their environment (Davis, 2002; Ellis et al., 2017).

Conventional tags are fisheries-dependent and reliant upon high recapture and reporting rates (Hueter et al., 2006). These tags are a more economical monitoring tool due to their low-cost, but several factors that influence reporting rates, such as tag shedding, participation of fishers, and emigration, make it very difficult to come to any definitive conclusions about fate with this technique (Kohler and Turner, 2001). Mortality is also not directly observed with these tags, but rather inferred from reporting rates (Ellis et al., 2017). In contrast, electronic tags record movement and provide researchers with a more continuous monitoring approach where mortality is directly identified at discrete times (via movement or

lack thereof; Davis, 2002; Hammerschlag et al., 2011; Ellis et al., 2017). Pop-up satellite archival transmitting (PSAT) tags are a type of electronic tag that has become one of the more frequently used tools (e.g., Campana et al., 2015; Sulikowski et al., 2020; Mohan et al., 2020). These tags record pressure (depth), water temperature, and light-level data over a pre-determined time-period, upon which they detach from the animal and float to the surface where archived data are then transmitted via the Argos satellite array (Hammerschlag et al., 2011). In addition to the appeal of fisheries-independent data recovery, PSAT tags monitor animals with little to no spatial restrictions outside of extreme depths that can damage sensors (Hammerschlag et al., 2011). High-cost of PSAT tags does make it logistically difficult to collect large sample sizes, but with an appropriate subsample of animals across the range capture-and-handling factors, relevant estimates of DM can be generated via these tags (Hammerschlag et al., 2011).

Information provided by these tools and techniques can be incorporated into stock assessments and used by fisheries managers and other end-users (e.g., industry stakeholders) when developing bycatch mitigation strategies (Benaka et al., 2016). In particular, DM estimates benefit stock assessments because they permit a comprehensive approximation of fishing mortality that translates into more relevant biological reference points, which are used to determine that status of a stock (i.e., overfished/overfishing; Crowder and Murawski, 1998; Hoggarth et al., 2006; Benaka et al., 2016). This status is then used by fisheries managers when determining the level of regulation needed to promote a sustainable stock, which will also require knowledge of DM rates and the factors driving mortality. In general, input management controls are used to regulate the amount of allowable fishing effort (e.g.,

days-at-sea or gear-limitations), while output controls are used to control the amount of legal catch (e.g., total allowable catch or individual transferable quotas; Crowder and Murawski, 1998; Hoggarth et al., 2006; Benaka et al., 2016). Best fishing practices and gear-modifications can also be required to reduce the impact of capture and number of fisheries-interactions, while maintaining fishing opportunities for the industry (Hoggarth et al., 2006; Benaka et al., 2016). In addition, time-area closures can be afforded to defined areas of concern (e.g., spawning or nursery grounds or areas with high catch rates) to more directly reduce interactions and fishing mortality (Hall and Mainprize, 2005; Broadhurst et al., 2006; Benaka et al., 2016). But while many bycatch mitigation strategies are available, it can often be difficult for managers and other stakeholders to identify the most practical and effective strategies for a given species/fishery, especially when the species being managed are data-poor and/or interact with multiple fisheries.

In light of the potential impact of capture-and-handling on animal health/fate and the need for capture-related information from a management and mitigation standpoint, the primary objective of this dissertation was to contribute new tools and techniques to our toolbox for addressing fisheries-interactions. This included a novel application of electronic tagging to deduce fate in an elasmobranch family (Chapter I), a tool that can identify spatiotemporal bycatch hotspots where fisheries-interactions are most likely to occur (Chapter II), and a technique for evaluating capture behavior (Chapter III). These tools and techniques complement the pre-existing toolbox and are applicable to a wide range of species and fisheries, beyond those described in this dissertation. In addition, two overfished species of elasmobranch (thorny skate, *Amblyraja radiata*; and oceanic whitetip shark, *Carcharhinus*

longimanus) were used as case study species in this dissertation. Both are discarded at high rates and it is not clearly understood how capture-and-handling affects the population dynamics of these species. Therefore, the secondary objective of this dissertation was to provide information for these species that will improve our understanding of their fisheries-interactions and aid in the development of future bycatch mitigation strategies that will promote their recovery.

Literature cited

Arlinghaus, R., & Hallermann, J. (2007). Effects of air exposure on mortality and growth of undersized pikeperch, *Sander lucioperca*, at low water temperatures with implications for catch and-release fishing. *Fisheries Management and Ecology*, *14*(2), 155-160.

Barton, B. A. (2002). Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. *Integrative and comparative biology*, *42*(3), 517-525.

Benaka, L. R., Sharpe, L., Abrams, K. H., Campbell, M., Cope, J. M., Darby, F., & Rioux, D. D. (2016). Action plan for fish release mortality science.

Bouyoucos, I. A., Talwar, B. S., Brooks, E. J., Brownscombe, J. W., Cooke, S. J., Suski, C. D., & Mandelman, J. W. (2018). Exercise intensity while hooked is associated with physiological status of longline-captured sharks. *Conservation Physiology*, *6*(1), coy074.

Camhi, M. D. (2008). Conservation status of pelagic elasmobranchs. *Sharks of the Open Ocean: Biology, Fisheries and Conservation*, 397-417.

Chrousos, G. P. (1998). Stressors, stress, and neuroendocrine integration of the adaptive response: the 1997 Hans Selye Memorial Lecture. *Annals of the New York Academy of Sciences*, *851*(1), 311-335.

Crowder, L. B., & Murawski, S. A. (1998). Fisheries bycatch: implications for management. *Fisheries*, *23*(6), 8-17.

Daly, J., Gunn, I., Kirby, N., Jones, R., & Galloway, D. (2007). Ultrasound examination and behavior scoring of captive broadnose sevengill sharks, *Notorynchus cepedianus* (Peron, 1807). *Zoo Biology: Published in affiliation with the American Zoo and Aquarium Association*, *26*(5), 383-395.

Danylchuk, A. J., Suski, C. D., Mandelman, J. W., Murchie, K. J., Haak, C. R., Brooks, A. M., & Cooke, S. J. (2014). Hooking injury, physiological status and short-term mortality of juvenile lemon sharks (*Negaprion brevirostris*) following catch-and-release recreational angling. *Conservation physiology*, *2*(1), cot036.

Dapp, D. R., Walker, T. I., Huvneers, C., & Reina, R. D. (2016). Respiratory mode and gear type are important determinants of elasmobranch immediate and post-release mortality. *Fish and Fisheries*, *17*(2), 507-524.

Davis, M. W. (2002). Key principles for understanding fish bycatch discard mortality. *Canadian Journal of Fisheries and Aquatic Sciences*, *59*(11), 1834-1843.

- Davis, M. W. (2010). Fish stress and mortality can be predicted using reflex impairment. *Fish and Fisheries*, *11*(1), 1-11.
- Dent, F., & Clarke, S. (2015). State of the global market for shark products. *FAO Fisheries and Aquaculture technical paper*, (590), I.
- Dulvy, N. K., Baum, J. K., Clarke, S., Compagno, L. J., Cortés, E., Domingo, A., & Martínez, J. (2008). You can swim but you can't hide: the global status and conservation of oceanic pelagic sharks and rays. *Aquatic Conservation Marine and Freshwater Ecosystems*, *18*(5), 459-482.
- Ellis, J. R., McCully Phillips, S. R., & Poisson, F. (2017). A review of capture and post-release mortality of elasmobranchs. *Journal of Fish Biology*, *90*(3), 653-722.
- Ferretti, F., Worm, B., Britten, G. L., Heithaus, M. R., & Lotze, H. K. (2010). Patterns and ecosystem consequences of shark declines in the ocean. *Ecology Letters*, *13*(8), 1055-1071.
- Frick, L. H., Reina, R. D., & Walker, T. I. (2009). The physiological response of Port Jackson sharks and Australian swellsharks to sedation, gill-net capture, and repeated sampling in captivity. *North American Journal of Fisheries Management*, *29*(1), 127-139.
- Frick, L. H., Reina, R. D., & Walker, T. I. (2010). Stress related physiological changes and post-release survival of Port Jackson sharks (*Heterodontus portusjacksoni*) and gummy sharks (*Mustelus antarcticus*) following gill-net and longline capture in captivity. *Journal of Experimental Marine Biology and Ecology*, *385*(1-2), 29-37.
- Gallagher, A. J., Serafy, J. E., Cooke, S. J., & Hammerschlag, N. (2014). Physiological stress response, reflex impairment, and survival of five sympatric shark species following experimental capture and release. *Marine Ecology Progress Series*, *496*, 207-218.
- Gallagher, A. J., Staatterman, E. R., Cooke, S. J., & Hammerschlag, N. (2017). Behavioural responses to fisheries capture among sharks caught using experimental fishery gear. *Canadian Journal of Fisheries and Aquatic Sciences*, *74*(1), 1-7.
- Gale, M. K., Hinch, S. G., & Donaldson, M. R. (2013). The role of temperature in the capture and release of fish. *Fish and Fisheries*, *14*(1), 1-33.
- Gelsleichter, J., Rasmussen, L. E. L., Manire, C. A., Tyminski, J., Chang, B., & Lombardi-Carlson, L. (2002). Serum steroid concentrations and development of reproductive organs during puberty in male bonnethead sharks, *Sphyrna tiburo*. *Fish Physiology and Biochemistry*, *26*(4), 389-401.

- Gurshin, C. W. D., & Szedlmayer, S. T. (2004). Short-term survival and movements of Atlantic sharpnose sharks captured by hook-and-line in the north-east Gulf of Mexico. *Journal of Fish Biology*, 65(4), 973-986.
- Hall, S. J., & Mainprize, B. M. (2005). Managing by-catch and discards: how much progress are we making and how can we do better?. *Fish and Fisheries*, 6(2), 134-155.
- Hammerschlag, N., Gallagher, A. J., & Lazarre, D. M. (2011). A review of shark satellite tagging studies. *Journal of Experimental Marine Biology and Ecology*, 398(1-2), 1-8.
- Heupel, M. R., Knip, D. M., Simpfendorfer, C. A., & Dulvy, N. K. (2014). Sizing up the ecological role of sharks as predators. *Marine Ecology Progress Series*, 495, 291-298.
- Hoggarth, D. D. (2006). *Stock assessment for fishery management: A framework guide to the stock assessment tools of the fisheries management and science programme* (No. 487). Food & Agriculture Organization.
- Hueter, R. E., Manire, C. A., Tyminski, J. P., Hoenig, J. M., & Hepworth, D. A. (2006). Assessing mortality of released or discarded fish using a logistic model of relative survival derived from tagging data. *Transactions of the American Fisheries Society*, 135(2), 500-508.
- Kohler, N. E., & Turner, P. A. (2001). Shark tagging: a review of conventional methods and studies. In *The behavior and sensory biology of elasmobranch fishes: an anthology in memory of Donald Richard Nelson* (pp. 191-224). Springer, Dordrecht.
- Laptikhovskiy, V. V. (2004). Survival rates of rays discarded by the bottom trawl squid fishery off the Falkland Islands. *Fisheries Bulletin*, 102(4), 757-759.
- Mandelman, J. W., & Skomal, G. B. (2009). Differential sensitivity to capture stress assessed by blood acid–base status in five carcharhinid sharks. *Journal of Comparative Physiology B*, 179(3), 267.
- Mandelman, J. W., Cicia, A. M., Ingram Jr., G. W., Driggers III, W. B., Coutre, K. M., & Sulikowski, J. A. (2013). Short-term post-release mortality of skates (family Rajidae) discarded in a western North Atlantic commercial otter trawl fishery. *Fisheries Research*, 139, 76-84.
- Molina, J. M. & Cooke, S. J. (2012). Trends in shark bycatch research: current status and research needs. *Reviews in Fish Biology and Fisheries*, 22: 719-737.
- Pollock, K. H., & Pine III, W. E. (2007). The design and analysis of field studies to estimate catch-and-release mortality. *Fisheries Management and Ecology*, 14(2), 123-130.

Raby, G. D., Donaldson, M. R., Hinch, S. G., Patterson, D. A., Lotto, A. G., Robichaud, D., English, K. K., Willmore, W. G., Farrell, A. P., Davis, M. W., & Cooke, S. J. (2012). Validation of reflex indicators for measuring vitality and predicting the delayed mortality of wild coho salmon bycatch released from fishing gears. *Journal of Applied Ecology*, 49(1), 90-98.

Portz, D. E., Woodley, C. M., & Cech, J. J. (2006). Stress-associated impacts of short-term holding on fishes. *Reviews in Fish Biology and Fisheries*, 16, 125-170.

Skomal, G., & Bernal, D. (2010). Physiological responses to stress in sharks. *Sharks and their relatives II: biodiversity, adaptive physiology, and conservation*. CRC Press, Boca Raton, 459-490.

Skomal, G. B., & Mandelman, J. W. (2012). The physiological response to anthropogenic stressors in marine elasmobranch fishes: a review with a focus on the secondary response. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 162(2), 146-155.

Talwar, B. S., Bouyoucos, I. A., Brooks, E. J., Brownscombe, J. W., Suski, C. D., Cooke, S. J., Grubbs, R. D., & Mandelman, J. W. (2020). Variation in behavioural responses of sub-tropical marine fishes to experimental longline capture. *ICES Journal of Marine Science*.

Veldhuizen, L. J. L., Berentsen, P. B. M., de Boer, I. J. M., van de Vis, J. W., & Bokkers, E. A. M. (2018). Fish welfare in capture fisheries: a review of injuries and mortality. *Fisheries Research*, 204, 41-48.

Wendelaar Bonga, S. E. (1997). The stress response in fish. *Physiological Reviews*, 77(3), 591-625.

Whitney N. M., White C. F., Gleiss A. C., Schwieterman G. D., Anderson P., Hueter R. E., & Skomal, G.B. (2016). A novel method for determining post-release mortality, behavior, and recovery period using acceleration data loggers. *Fisheries Research*. 183: 210-221.

Whitney, N. M., White, C. F., Anderson, P. A., Hueter, R. E., & Skomal, G. B. (2017). The physiological stress response, post-release behavior, and mortality of blacktip sharks (*Carcharhinus limbatus*) caught on circle and J-hooks in the Florida recreational fishery. *Fishery Bulletin*, 115(4), 532-544.

Wood, C. M., Turner, J. D., & Graham, M. S. (1983). Why do fish die after severe exercise?. *Journal of Fish Biology*, 22(2), 189-201.

Worm, B., Davis, B., Ketteimer, L., Ward-Paige, C. A., Chapman, D., Heithaus, M. R., Kessel, S. T., & Gruber, S. H. (2013). Global catches, exploitation rates, and rebuilding options for sharks. *Marine Policy*, 40, 194-204.

CHAPTER 1

UTILIZATION OF POP-UP SATELLITE ARCHIVAL TRANSMITTING TAGS TO EVALUATE THORNY SKATE (*AMBLYRAJA RADIATA*) DISCARD MORTALITY IN THE GULF OF MAINE GROUND FISHIN BOTTOM TRALWL FISHERY

Abstract

Thorny skate (*Amblyraja radiata*) remain one of the most overfished species in the Gulf of Maine (GOM) despite being designated as a prohibited (zero-possession, mandatory release) species by the New England Fishery Management Council in 2003. To better understand the extent to which discard mortality (DM) occurring after incidental capture in the GOM groundfish bottom trawl fishery may be impeding recovery, 75 individuals (55 – 94 cm total length, TL) were tagged with pop-up satellite archival transmitting (PSAT) tags and monitored for up to 28 days following capture under representative commercial trawl fishing practices. Data recovered from 61 PSAT-tagged skate were analyzed with a longitudinal survival analysis to estimate DM and identify influential capture-related variables. DM rate was found to be a function of TL, with larger skates (> 70 cm; DM = 16.5%) experiencing lower mortality than smaller conspecifics (55 – 70 cm; DM = 24.5%).

From our results, we estimate annual thorny skate DM in the GOM groundfish bottom trawl fishery to be 79.2 ± 0.2 mt, which accounts for less than one percent of the existing stock biomass in the GOM (8,400 mt). This study confirms that thorny skate are relatively resilient to bottom trawl fishing practices in the GOM, and suggests that other sources of mortality may be impeding population recovery.

Introduction

In the U.S. Gulf of Maine (GOM), skate species managed under the Northeast Skate Complex Fishery Management Plan (FMP) are both harvested and taken as bycatch by fisheries operating in the region. Under the FMP, fishing mortality for seven species is managed through a single complex-wide quota on total allowable landings (NEFMC, 2003). However, highly restrictive species-specific regulations such as retention limits and complete prohibitions (i.e., zero-possession) on commercial landings have been enacted to reduce mortality on overfished populations (NEFMC, 2003; Curtis and Sosebee, 2015). While these measures may mitigate the impact of direct harvest (Waring, 1984; Hoenig and Gruber, 1990; Sulikowski *et al.*, 2003), some species still experience high rates of incidental capture (as bycatch), and discard mortality (DM) can continue to have significant impacts on skate populations due to their life history characteristics (e.g., slow growth, late maturation, and low fecundity). In fact, high DM rates may even perpetuate population declines by impeding species recovery (Davis, 2002).

Thorny skate (*Amblyraja radiata*) are one of the most overfished species in the GOM, with National Oceanic and Atmospheric Administration (NOAA) Northeast Fisheries

Science Center (NEFSC) bottom trawl survey biomass indices currently near 55 year lows (i.e., 0.17 kg/tow, less than 5% of the biomass target; NMFS, 2017). In response to this chronically poor stock status, the Northeast Skate Complex FMP prohibited this species in 2003 to reduce fishing mortality associated with commercial landings. However, thorny skate biomass in the GOM has declined further since the prohibition (Sosebee *et al.*, 2016), thereby suggesting that other sources of mortality, and/or environmentally-driven distributional shifts (e.g., Nye *et al.*, 2009), are impeding stock recovery. For example, DM associated with the groundfish bottom trawl fishery, which was recently estimated to account for 72.6% (or 322.8 metric tons, mt) of total annual thorny skate live-discards in the GOM (Sosebee *et al.*, 2016), may represent a significant source of mortality.

DM in the GOM groundfish bottom trawl fishery was previously addressed by Mandelman *et al.* (2013), wherein short-term mortality was found to be 23% for 351 skates that were caught under representative trawl conditions and monitored in submerged enclosures for 72-hours. However, the authors urged caution when considering the adoption of this estimate for stock assessments because of uncertainties associated with the 72-hour cross-sectional results, particularly the compromised state (i.e., listless or moribund) of surviving skates at the completion of the 72-hour trials, and the higher mortality (54%) evident in a subsample ($n = 35$) of thorny skate that were transported to shore-based tanks and monitored for seven days. Long-term DM beyond 72-hours has been reported for several other fish species (Kaimmer and Trumble, 1998; Knotek *et al.*, 2018; Schram and Molenaar, 2018), and shown to extend upwards of 27-days post-capture in a simulated trawl-capture with Pacific halibut, *Hippoglossus stenolepis* (Davis and Olla, 2001). However, the extent to

which confinement stress throughout monitoring and/or transport or shielding from post-release predation, two potential limitations of confinement studies (Portz et al., 2006), affected the 23% DM estimate of Mandelman et al. (2013) is unknown. Nonetheless, the NEFMC incorporated the 23% DM rate into stock assessments under Framework Adjustment 2 to the Northeast Skate Complex FMP (NEFMC, 2014). Given the impact of assumed DM rates on annual catch limits and quotas set for the Northeast Skate Complex, a further investigation of trawl-caught thorny skate DM with monitoring periods exceeding 72-hours is warranted to address the aforementioned uncertainties and confirm the accuracy of current DM estimates.

Pop-up satellite archival transmitting (PSAT) tags are powerful fisheries-independent tools for monitoring post-release fate over extended durations (Campana et al., 2009). These tags can be rapidly applied (externally) to an animal prior to release, and allow the animal to swim freely and interact with its environment for a user-defined period prior to reporting their archived data via the Argos satellite system. These data can then be used to infer the post-release fate of an animal, determine the timing of mortality (if evident), and evaluate the effects of the capture and handling process on survival. Although PSAT tags have not previously been used to investigate skate DM, they have been used to examine movement patterns of several skate species including the common skate (*Dipturus batis*; Wearmouth and Sims, 2009), Arctic skate (*Amblyraja hyperborea*; Peklova et al., 2014) and big skate (*Beringraja binocularata*; Farrugia et al., 2016) over long-term periods (weeks to months), thereby demonstrating their utility for DM estimation. For these reasons, the objective of this study was to derive DM estimates for trawl-caught thorny skate over extended monitoring

periods via PSAT tags. In addition to estimating DM rates, this study aimed to identify the most-influential factors (e.g., fishing conditions and practices, individual biological traits, and degree of capture-related physical trauma) that affect thorny skate mortality, which could then be incorporated into a best-practice framework to reduce mortality in the GOM groundfish bottom trawl fishery.

Methods

Study site and fishing protocol

At-sea trials (n = 9 single-day trips) to collect information on thorny skate condition and DM rate following capture with typical commercial fishing practices in the GOM groundfish bottom trawl fishery were conducted during August and September 2016 (n = 7 days) and September 2017 (n = 2 days) in the coastal waters off northern Massachusetts (Fig. 1.1). All trials occurred aboard a mid-sized (~ 13 m) commercial fishing trawl vessel, the F/V *Mystique Lady* (Gloucester, MA, USA), and tows (n = 27) were conducted using a standard trawl net (16.51 cm codend mesh; 20.12 m head rope length; 21.95 m footrope length; 8.90 cm cookie sweep) with 127 cm trawl doors. Fishing was conducted in depths that ranged from 55 to 161 m and occurred on predominately soft-bottom substrate in calm seas (< 1 m wave height) for durations of 30 – 240 minutes and at speeds of 2 to 4 kts. Bottom seawater and air temperatures were measured during all tows and on-deck sampling with HOBO temperature loggers (Onset Computer Corp., Bourne, MA, USA) affixed to the head rope of the trawl net and to the topside of the vessel, respectively. The temperature logger on the topside of the vessel was housed in an RS1 Solar Radiation Shield (Onset Computer Corp.)

to more accurately measure on deck air temperatures. From these measurements, the temperature gradient evident during each sampling event was calculated as the difference between bottom seawater and air temperatures. Temperature gradients during this study reflect “worst-case scenario” summer temperature regimes in the GOM (elevated bottom to surface seawater gradients and highest air temperatures), which are known to exacerbate capture-related stress in congeneric species of skate from this region (e.g., $> 9^{\circ}\text{C}$ temperature gradients; Cicia et al., 2012).

Sampling procedure

Following each tow, the catch was deposited on-deck and the total catch weight (in kg) was estimated by the vessel captain. Fishermen then followed standard commercial practices in the GOM trawl fishery when culling/sorting catch. This included sorting of the catch by species into fish totes (71 cm \times 41 cm \times 28 cm) either by hand or by inserting a fish pick (a hand-held tool with a metal hook or nail for lifting and maneuvering catch) into the wing musculature. To examine the effect of air exposure on DM, individuals were sampled continuously over a period of ~ 75 minutes from the initial air exposure of the net codend, which represents the maximum elapsed on-deck time prior to discard in the fishery.

For each skate sampled, the sex and total length (TL; cm) were recorded. Maturity status was also recorded, with male skates considered mature when claspers were long and fully calcified (Ruocco et al., 2006), and female skates recorded as mature based on palpitation for developing egg cases or having a distended/vascularized cloaca (i.e., suggesting recent parturition; Sosebee, 2005). Each skate was also assigned a nominal injury

code based upon the level of overt physical trauma as originally described in Mandelman et al. (2013) (Table 1.1; Appendix Fig. A.1). Injury assessments for all skate were conducted by a single researcher and therefore were not subject to interpretational bias. Skate were then released, with the exception of a subsample described in the following section.

Monitoring the fate of discarded skate

To examine the fate of discarded thorny skate, a subset of 75 individuals (55 – 94 cm TL) were tagged with PSAT tags (Model: PSATLIFE, Lotek Wireless Inc., St. John's, Newfoundland, Canada). These tags measured pressure (i.e., depth) and ambient water temperature at 10-second intervals over 14- or 28-day deployment periods. Following this period, the tags were programmed to detach, float to the surface, and transmit the archived data compiled into 5-minute bins. Higher resolution data (i.e., 10-second intervals) were only accessible in cases where the tag was physically recovered and directly downloaded using the “Tag Talk” program (Lotek Wireless Inc.). Tags were affixed near the center of one of the skate's pectoral wings using a method modified from Wearmouth and Sims (2009), wherein a monofilament line and two pairs of 1” Petersen discs and baffles were crimped on either side of the skate's wing musculature (dorsal and ventral; Fig. 1.2). To minimize the potential impact of sampling and tagging on post-release fate, individual skates were processed as quickly as possible (1 – 6 minutes). Prior to at-sea sampling, tag attachment/retention trials were performed at the University of New England Marine Science Center (Biddeford, ME USA) with captive thorny skate (55 – 65 cm TL) to confirm that our attachment technique resulted in no tag shedding and/or tag-induced mortality up to 28 days post-tagging. Tags

were not applied to skate < 55 cm TL due to concerns that the PSAT tags would compromise the mobility of smaller individuals and therefore potentially result in tag-induced mortality.

PSAT tags were strategically deployed on individuals across a range of tow and air exposure durations, and for each injury code and handling method (i.e., hand sorting versus fish pick), in order to assess the impact of several variables experienced during typical commercial fishing conditions. Furthermore, PSAT tags (14-day deployment) were also deployed on dead skates (n = 2) to establish “negative control” movement signatures, which would be used in subsequent analyses to infer mortality events.

Survival analysis

We utilized a longitudinal survival analysis to model the probability of survival as a function of time (e.g., Cox and Oakes, 1984; Benoît et al., 2012, 2015). In this analysis there were four primary objectives. First, PSAT tag depth time-series data were examined to determine the fate of our subsampled skates. Second, the fishing conditions and biological covariates associated with each subsampled skate were evaluated to identify which covariates were best able to predict survival following discard. Third, using this subset of covariates, a suite of survival mixture models (SMMs; developed by Benoît et al., 2012, 2015) were fitted to explain survival and ultimately provide DM estimates. Finally, best-practice frameworks were formulated based upon the subset of covariates to mitigate thorny skate mortality. All analyses were performed using R 3.4.5 (R-Core Team, 2017), and when applicable, statistical significance was accepted at $p < 0.05$.

Depth-variance survival test

To determine the fate of individual discarded animals, we first converted pressure data collected by the PSAT tags to depth using the package ‘rtide’ (Thorley et al., 2017) and removed any tidal-noise associated with the tidal cycle of the area with the ‘oce’ package (Kelly and Richards, 2018). Depth time-series for each skate were then subjected to a modified depth-variance survival test adapted from Capizzano et al. (2016), with an additional clause to ensure the test was not falsely classifying extended periods of on-bottom (live) behavior as mortality events (details in Appendix A). If skate survived throughout the monitoring period they were treated as right-censored data because their fate is unknown following tag detachment (i.e., time of death occurs after the monitoring period). If mortality was confirmed (i.e., a censored event), time of death was estimated as the time bin with the first non-significant result in the sequence leading up to the end of the trial. One individual was recaptured alive ~ eight days after release, but died shortly after being released again. Since this mortality event could not be attributed to the original capture event, its time-series was truncated to the recapture event and was treated as a right-censored (i.e., alive) observation at the point of recapture.

Assessment of fishing conditions and practices, individual biological traits, and injury

For this step of the analysis we considered all covariates (i.e., fishing conditions and practices, individual biological traits, and injury scores) that could influence survivorship of trawl-caught thorny skate. Initial examination of correlation between covariates and relevant interactions (i.e., TL and catch weight, TL and air exposure, and air temperature and air

exposure) using a Pearson correlation revealed that interactions terms were all highly collinear with respect to catch weight ($r = 0.97$) or air exposure ($r = 0.97$). Therefore, interaction terms were not considered in the survival analysis.

The empirical Kaplan-Meier estimator (Cox and Oakes, 1984), a non-parametric analysis that provides an estimate of the survival function by following the proportion of live skate throughout the time-at-large and in the absence of censored values, was utilized to generate preliminary graphical depictions of the influence of each covariate on survivorship, and to provide an empirical estimate of survival that can be used to assess the fit of SMMs developed in later steps.

The multivariate (semi-parametric) survival analysis known as the Cox proportional-hazards model (CPHM; Cox, 1972; Therneau and Grambsch, 2000) was then used to identify which covariates (i.e., tow duration, fishing depth, temperature gradient, catch weight, handling method, air duration, TL, sex, maturity, and injury score) predicted the survival of discarded thorny skate following the methodologies outlined in Knotek et al. (2018). This model is defined as:

$$\hat{h}(t) = h_0(t)^{(X'\beta + Z'b)} \quad (1)$$

where $\hat{h}(t)$ is the hazard function at time t (i.e., the risk of incidental mortality occurring at time t) determined by a non-parametric baseline hazard function $h_0(t)$, a suite of relevant covariates X' , and a Gaussian random effect Z' (with tow as the subject) to account for any within-tow correlations (Benoît et al., 2010; Knotek et al., 2018). To estimate parameters in

this model we used partial maximum likelihood (Cox, 1972; Ripatti and Palmgren, 2000). Model building was two-fold: (1) identify whether or not the random effect was appropriate, and (2) find the most parsimonious set of covariates to predict survival (details in Appendix A; Benoît et al., 2010; Knotek et al., 2018).

Model building revealed a fixed-effects modeling approach was appropriate (i.e., random effect of tow was not significant; $p = 0.97$) and TL was the only covariate that predicted survival. TL was then converted into a categorical predictor (binned every 10 cm) to reflect size classes used in NEFSC bottom trawl surveys, and to allow us to identify which of the TL-specific DM rates (generated in subsequent steps) best reflected the size-distribution of thorny skate discards in the GOM. Log-rank tests were then performed to test for differences in the underlying survival function between size classes and if the resulting p -value was not significant ($p > 0.05$), these classes were combined. This led to the consolidation of bins greater than 70 cm TL (χ^2 : 0.48 to 0.82) and bins less than or equal to 70 cm TL ($\chi^2 = 0.70$). TL was therefore used as a two-level categorical covariate in subsequent analyses.

Predicting post-release survival based upon relevant covariates

To estimate post-release survivorship rates, we chose to use a parametric survival analysis modeling approach (i.e., SMMs) developed by Benoît et al. (2012, 2015). This approach is favorable due to its ability to explain the survivorship of thorny skate over time with longitudinal data, and estimate the time at which survivorship asymptotes (i.e., survival rate). The underlying model of the SMMs is defined as:

$$\hat{S}(t) = \pi \cdot \exp[-(\alpha \cdot t)^\gamma] + (1 - \pi) \quad (2)$$

where $\hat{S}(t)$ is the probability of survival at time t . The probability of a skate being adversely affected by capture-and-handling is denoted by π and the survival function for these skate is explained by $\exp[-(\alpha \cdot t)^\gamma]$, which is assumed to follow a Weibull distribution that includes scale and shape parameters, α and γ , respectively. The skates not adversely affected by capture-and-handling are assumed to have a survival probability of 1, as natural mortality was assumed to be negligible during the relatively short-duration of the tag monitoring period (14 – 28 days). In Eq. (2) the survival rate is defined as $\hat{S}(t) = 1 - \pi$ because as $t \rightarrow \infty$, $\exp[-(\alpha \cdot t)^\gamma]$ eventually leads to nil survival for affected individuals, leaving only the non-affected skate alive. To incorporate the influence of TL on survivorship, α and π parameters can be manipulated to address the effect of the covariates on the survival rate over time and/or the probability of a skate being adversely affected, respectively. This approach produces several competing models according to the different assumptions of how TL impacts survivorship (Table 1.2). All model variants were fit using maximum likelihood methodology (additional details provided in Benoît et al., 2012, 2015) and convergence and fit were evaluated by overlaying the 95% confidence bands from Kaplan-Meier estimates with the predicted survival functions. Model fits were also compared using Akaike's Information Criterion corrected for small samples sizes (AICc; Burnham and Anderson, 2002), wherein ΔAICc values (relative to lowest AICc value between models) were used to identify the model variant(s) most suited for predicting the survival function (details in

Appendix A; Burnham and Anderson, 2002). Based upon a comparison of ΔAICc (Table 1.2) there was similar support for several of the competing models and therefore no single-model could be considered a best-fit. Therefore, model averaging (of SMMs) based on Akaike weights (e.g., Lukacs et al., 2010; Benoît et al., 2010; details in Appendix A) was used to generate a single survival rate for each size class that reflected the suite of model variants according to their ΔAICc values (e.g., Knotek et al., 2018).

Estimation of DM on a fishery-scale

To estimate fishery-wide DM for thorny skate in the GOM groundfish bottom trawl fishery, we used a modified modeling approach described by Knotek et al. (2018). Here, model averaged survival rates (see above) were converted into DM rates (i.e., $1 - \text{survival rate}$), and then applied to the most recently available estimate of total annual live-discards (by weight) in the fishery to derive an estimate of DM (in mt). However, since recent NEFSC bottom trawl data suggests that $\sim 95\%$ of all thorny skate occupying the GOM are < 74 cm (Sosebee et al., 2016), only the DM rate associated with the smaller-size category (55 – 70 cm TL) was used to estimate DM. Finally, to account for model variant-specific parameter uncertainty and variability in fishery-scale estimates of DM, Monte Carlo simulations based on bootstrapping (Efron and Tibshirani, 1993; details in Appendix A) were used with this approach to derive the final fishery-scale estimate (mean and standard deviation) of DM.

Results

Fieldwork and capture characteristics

A total of 612 thorny skate (24 – 94 cm TL; 305 male and 307 female) were captured and sampled after being handled by hand (n = 323) or with a fish pick (n = 289; Table 1.3). The majority of skates were categorized as having minor (i.e., injury code 1; 38.4%) to moderate (i.e., injury code 2; 44.0%) overt physical trauma with the fewest observations made for severely injured skate (i.e., injury code 3; 17.6%; Table 1.3). Tows that captured thorny skate were primarily composed of common commercial groundfish species (e.g., grey sole, *Glyptocephalus cynoglossus*; yellowtail flounder, *Pleuronectes ferruginea*; winter flounder, *Pseudopleuronectes americanus*; haddock, *Melanogrammus aeglefinus*; and Atlantic cod, *Gadus morhua*), monkfish (*Lophius americanus*), spiny dogfish (*Squalus acanthias*), skate (e.g., little skate, *Leucoraja erinacea*; winter skate, *Leucoraja ocellata*; and barndoor skate, *Dipturus laevis*), and to a lesser extent, invertebrates (e.g., American lobster, *Homarus americanus*; and rock crab, *Cancer borealis*). Total estimated catch weight per tow ranged from 29 to 907 kg. Bottom seawater and air temperature ranged from 6.3 to 8.3°C and 16.5 to 26.8°C, respectively, while the temperature gradient between the two locations varied from 9.3 to 19.6°C (Table 1.4).

Monitoring the fate of discards and negative controls

Of the 75 PSAT tags deployed, 59 transmitted Argos data and 35 were physically recovered and/or returned and directly downloaded, including two tags that did not transmit. In total, we retrieved data from 61 of the 75 PSAT tags (81.3%) via transmission (n = 30)

and/or physical recovery and direct download ($n = 31$). Tags that only provided Argos transmissions, including one of the two negative controls (i.e., known dead skate), yielded an average of 88.7% (out of 4,032 expected observations) and 72.3% (out of 7,888 expected observations) of the time-series data in 5-minute resolution for 14- and 28-day deployment periods, respectively. In contrast, physically downloaded tags provided access to the entire time series in 10-second resolution for the 14- and 28-day deployment periods (i.e., 120,960 and 236,640 observations, respectively).

Depth-variance survival test

The depth-variance survival test identified seven mortality events, with the majority ($n = 5$ individuals) of capture-related DM having occurred within 18-hours of release. However, two mortalities did occur at 241 and 263-hours post-release (Fig. 1.3). After mortality occurred, PSAT tags generally remained attached to dead-skates for 121 to 489 hours before dislodging and floating to the surface. Surviving skates ($n = 53$) occupied depths ranging from 22.9 to 184.0 m (mean \pm standard deviation: 89.8 ± 19.0 m) and displayed vertical movement behavior that can be broadly categorized as periods remaining at-depth (maximum span of 102 hours) with brief off-bottom forays (associated with diel-cycles), and on/off shelf movement (Appendix Fig. A.2). Ultimately, these behaviors made it possible to characterize the status of skate (live or dead) throughout a time-series and identify mortality events.

Discard survival rates and estimates of fishery-scale DM

Model variants (i.e., SMMs) fit the 95% confidence bands relatively well (i.e., within the center of the bands) with the exception of the timeframe preceding the first mortality in the smaller-sized category (55 to 70 cm TL). However, model fit largely improved over time for this category (Fig. 1.4). Individual model variants generated a range of size-specific DM rates according to the various model assumptions (Table 1.2). For example, Weibull 2 assumed that all skate were adversely affected (nil survival), whereas Mixture 4 allowed the survival function and proportion of adversely affected skate to vary by size category (6.7 and 22.2% mortality for large and smaller-sized skate, respectively; Table 1.2). However, because each variant carried different Akaike weights (Table 1.2) with no clear best-fitting model (according to Benoît et al., 2012), model averaging was used to generate a single overall DM rate for each size category that incorporated model variant DM rates based on their relative contribution to the survival function (i.e., Akaike weight). This approach estimated that larger-sized skate have a lower DM rate (16.5 ± 0.1 %) than smaller thorny skate (24.5 ± 0.1 %; Table 1.2). Based upon the mortality rate for the smaller-sized category and average annual estimates of bottom trawl thorny skate live discards (i.e., 322.8 mt in 2014; Sosebee et al., 2016), annual DM for the GOM groundfish bottom trawl fishery is approximately 79.2 ± 0.2 mt.

Discussion

Thorny skate DM in the GOM groundfish bottom trawl fishery

This study demonstrated the utility of PSAT tags for assessing the long-term survivorship of a benthic skate species following discard from one of the GOM's largest commercial fisheries. By modeling PSAT tagging data in concert with detailed observations collected during the capture event, a size-dependent DM rate was identified that suggested larger thorny skate (DM rate = 16.5%) experience lower mortality than smaller conspecifics (DM rate = 24.5%) when caught in bottom otter trawls under "worst-case scenario" summer temperature regimes (Cicia et al., 2012). Furthermore, given the average size of thorny skate in the GOM, the DM rate estimate for the smaller-sized category (55 – 70 cm TL) was determined to best reflect the size-distribution of discards and the overall DM rate in the bottom trawl fishery. This overall DM rate is consistent with the previous estimate (23%) reported by Mandelman et al. (2013) and is markedly lower than estimates for other trawl-caught species of skate (Laptikhovsky, 2004; Enever et al., 2009; Mandelman et al., 2013; Saygu and Deval, 2014; Tsagarakis et al., 2015), which suggests thorny skate are fairly resilient to capture-and-handling in this fishery.

Our results indicate that the majority of thorny skate DM following capture in bottom trawls occurred within 18-hours of discarding; however, longer-term manifestations of DM were also evident up to 10 days post-release. Previous studies with trawl-caught skate in the Bristol Channel and Gulf of St. Lawrence (Enever et al., 2009; Benoît *et al*, 2012) found that DM occurred entirely within 24-hours post-release when using monitoring periods of up to 64 and 110 hours, respectively. In contrast, extended monitoring periods (up to 28 days) in

this study revealed delayed DM events, which corroborates recent observations of little and winter skate captured in the New England scallop dredge fishery (Knotek et al., 2018). Of note, most previous studies of skate DM have not been able to quantify delayed DM because of abbreviated monitoring periods (≤ 5 days) associated with traditional confinement methods (i.e., on-board holding tanks and submerged cages; Kaiser and Spencer, 1995; Endicott and Agnew, 2004; Laptikhovsky, 2004; Enever et al., 2009; Benoît et al., 2012; Mandelman et al., 2013; Depestele et al., 2014; Saygu and Deval, 2014; Tsagarakis et al., 2015; Sulikowski et al., 2018), thereby making our PSAT tag-derived results more reliable and robust to both short and delayed DM events. Nonetheless, the occurrence of delayed DM for thorny skate in this study was relatively low, which confirms that the 54% (7-day) DM rates reported by Mandelman et al. (2013) were more likely an artifact of confinement stress than a result of capture-and-handling (Portz et al., 2006).

The impact of capture-related and biological factors on DM rates

All of the mortalities documented in this study can be attributed to the effects of capture-and-handling in the GOM groundfish bottom trawl fishery, and are not considered a reflection of natural mortality. Natural mortality in these types of studies (e.g., Benoît et al., 2012; Mandelman et al., 2013; Depestele et al., 2014; Knotek et al., 2018) is considered negligible because the monitoring period is relatively short compared to the life-span of elasmobranchs (e.g., ~ 16 years for thorny skate in the GOM; Sulikowski et al., 2005), and any predation within this timeframe would have likely been due to decreased predator avoidance abilities from the effects of the capture event (Raby et al., 2014). Therefore, DM

rates for trawl-caught thorny skate were driven by capture-related factors that included the size of the animal, the degree of incurred physical trauma, thermal stress and stressors associated with air exposure during the capture-and-handling process.

TL was identified with our modeling approach as the most influential factor affecting trawl-caught thorny skate DM. In this size-dependent relationship, smaller thorny skate displayed a higher degree of DM than larger conspecifics, which has similarly been reported for skate discarded from beam trawls in the North Sea (Depestele et al., 2014), and speculated for thornback skate, *Raja clavata*, and brown skate, *Raja miraletus*, captured in bottom trawls in the eastern Mediterranean (Saygu and Deval, 2014). In addition, size has also been shown to influence at-vessel mortality (i.e., mortality occurring at or prior to landing) for ray species (family Dasyatidae and Myliobatidae) captured in the prawn trawl fishery off of Northern Australia (Stobutzki et al., 2002). In all of these scenarios, increased survivorship of larger-sized individuals may be the result of a thicker integument providing added protection (Mandelman et al., 2013; Depestele et al., 2014), having a larger body mass and core body temperature that is less sensitive to ambient temperature changes (Spigarelli et al., 1977), and/or having more energetic reserves to cope and recover from the capture-and-handling process. This size-dependent relationship may have also been an artifact of tag-induced mortality wherein individuals in the smaller-sized category (55 – 70 cm TL) were more susceptible to the adverse effects associated with carrying PSAT tags. However, there was no relationship between smaller size and increased mortality within this size category, and no mortality was observed during the initial 28-day retention trials that were conducted on skates measuring 55 – 65 cm TL (i.e., the smallest tagged in the field study). As such,

there is no evidence to suggest that our results were greatly confounded by tag-induced mortality.

Thorny skate DM rates were also influenced by the degree of physical trauma incurred during the capture event in this study. For example, the DM rate (based on live-dead proportion) increased by 25% across injury scores (code 1 = 6%; code 2 = 9%; code 3 = 31%). Injury is a well-documented factor impacting DM in skates (e.g., Depestele et al., 2014; Knotek et al., 2018; Sulikowski et al., 2018) and has been previously shown to influence DM for trawl-caught thorny skate in the Gulf of St. Lawrence (Benoît et al., 2012). Furthermore, higher degrees of physical trauma (injury codes 2 and 3) were likely the underlying driver for longer-term (delayed) manifestations of thorny skate mortality; wherein sublethal injuries (e.g., puncture wounds and lacerations) could have compromised the skate's integument and epidermal mucus layer, and led to osmoregulatory dysfunction, secondary infections, and/or susceptibility to disease (Luer, 2014; Cook et al., 2018).

Though there was evidence that injury influenced DM, higher incidental PSAT tag non-reporting rates for skate that incurred more severe physical trauma (injury codes 2 and 3) limited our ability to comprehensively evaluate the effect of injury on mortality. For example, while all tags reported from animals with no to minor injuries (code 1), non-reporting rates rose to 14.3% (n = 4) and 34.6% (n = 9) for skate that suffered moderate (code 2) to more severe injuries (code 3), respectively. This trend becomes particularly relevant to the study if we consider the scenario of non-reporting rates as an artifact of DM events; wherein tags were either damaged during a scavenging event or inhibited from transmitting following premature detachment from the dead skate (via

decomposition/scavenging) and the tag washing ashore and/or being covered by debris before the pre-programmed transmission at 14 or 28 days after deployment. Had more tags reported for the higher injury codes, it is likely that additional mortality events would have been documented and the degree of physical trauma would have had a more pronounced role in establishing DM rates. Therefore, the estimated mortality rates for skates with injury codes 2 and 3 should be considered minimums.

Despite the effect of TL and injury on thorny skate DM, this species seems capable of surviving extended tow durations of up to 3.9 hours, prolonged bouts of air exposure up to 70 min, and large-scale temperature gradients up to 19.6°C, to which many skate species have previously shown susceptibility. For example, mortality increased with tow duration for trawl-caught cuckoo skate, *Leucoraja naevus*, small-eyed skate, *Raja microocellata*, blonde skate, *Raja brachyura*, and thornback skate in the Bristol Channel (Enever et al., 2009) and smooth skate, *Malacoraja senta*, in the GOM (Mandelman et al., 2013). Though there are inherent limitations in evaluating tow duration (i.e., inability to pinpoint the exact time skate entered the codend; Neilson et al., 1989), thorny skate are likely more resilient to longer periods of compaction in the codend because of their larger size and more rigid morphology that provides added physical protection (Sulikowski et al., 2005; Natanson et al., 2007). To this end, Mandelman et al. (2013) reported that trawl-caught smooth skate were 10.9% more likely to incur severe injuries than thorny skate in the GOM, which is further substantiated by the nearly 82% of trawl-caught thorny skate sustaining only minor to moderate injuries in the present study.

Individual and cumulative effects of prolonged bouts of air exposure and abrupt changes in bottom seawater to air temperature did not appear to result in high levels of thorny skate DM. Previous studies have demonstrated that both the individual (e.g., 30-minute air exposure or 18.1°C temperature gradient; Knotek et al., 2018) and cumulative (e.g., interaction of 50-minute air exposure and 9.1°C temperature gradient; Cicia et al., 2012) impacts of these factors have resulted in increased DM for little and winter skate. However, in this study trawl-caught thorny skate were capable of surviving both the individual and cumulative effects of more pronounced exposure events (up to 20 minutes longer) and abrupt temperature gradients (up to a 10.5°C larger gradient) than was reported as being lethal in little and winter skates (Cicia et al., 2012; Knotek et al., 2018). The reasons for this resiliency are unknown, but nonetheless, it suggests that thorny skate may be tolerant of acute hypoxia and thermal stress.

Management implications and conclusions

This study confirms that thorny skate are relatively resilient to bottom trawl fishing practices in the GOM, with annual DM accounting for less than one percent of the existing stock biomass in the region (Sosebee et al., 2016; NMFS, 2017). This suggests that factors other than DM (e.g., environmentally-driven change; Di Santo, 2015) may be contributing to the impaired recovery of this species that has been evident in the 15 years since its prohibition. However, given that thorny skate biomass is currently at only 2.3% of its highest historical level (i.e., 364 million mt in 1966; Sosebee et al., 2016) the population is likely highly sensitive to even small-scale sources of mortality, such as DM. Furthermore, the

impact of DM on thorny skate population recovery is further elevated if extensive fishing effort occurs within areas of high thorny skate biomass (Knotek et al., in prep.). Therefore, best-practices aimed at reducing mortality in this fishery should still be taken into consideration in fisheries management, especially when providing collateral benefits to other non-target species/size-classes. This could include best-practice handling methods such as prioritizing the immediate release of skates to reduce time on-deck and air exposure and encouraging fishermen to avoid picking skate in areas of vital organs (i.e., cranial, pericardial, and abdominal cavities). Other factors to consider include mitigation measures that address animal size; however, traditional best-practices such as size limitations are not applicable for this species given its prohibited status. Bycatch avoidance strategies are therefore a likely more appropriate option, particularly the avoidance of areas where smaller-sized skate are found in abundance. Lastly, given the gear-specific nature of DM, future research efforts should be directed towards evaluating mortality in the other major gear-types responsible for thorny skate discarding, including scallop dredge, sink gillnet, and bottom longline gears.

Acknowledgements

I would like to thank all of the co-authors on the manuscript published from this work (J. Mandelman, J. Kneebone, T. Curtis, J. Sulikowski, J. Jurek). Data collection in this study was made possible by research scientists from the Anderson Cabot Center for Ocean Life at the New England aquarium, who assisted me throughout this study during fieldwork. This study was approved by the New England Aquarium Institutional Animal Care and Use

Committee (Protocol 2016-02). Funding for this research was provided by the 2015 NOAA NMFS Bycatch Reduction Engineering Program (Award # NA15NMF4720373).

Literature cited

Benoît, H.P., Hurlbut, T., & Chassé, J. (2010). Assessing the covariates influencing discard mortality of demersal fishes in four fisheries using a semi-quantitative indicator of survival potential. *Fish. Res.*, 106, 436-447.

Benoît, H.P., Hurlbut, T., Chasse, J., & Jonsen, I.D. (2012). Estimating fishery-scale rates of discard mortality using conditional reasoning. *Fish. Res.*, 125-126, 318-330.

Benoît, H. P., Capizzano, C. W., Knotek, R. J., Rudders, D. B., Sulikowski, J. A., Dean, M. J., ...

& Mandelman, J. W. (2015). A generalized model for longitudinal short-and long-term mortality data for commercial fishery discards and recreational fishery catch-and-releases. *ICES J. Mar. Sci.*, 72(6), 1834-1847.

Burnham, K.P., & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*, 2nd ed. Springer-Verlag, New York

Campana, S.E., Joyce, W., & Manning, M.J. (2009). Bycatch and discard mortality in commercially caught blue sharks *Prionace glauca* assessed using archival satellite pop-up tags. *Mar. Ecol. Prog. Ser.*, 387, 241-253.

Capizzano, C.W., Mandelman, J.W., Hoffman, W.S., Dean, M.J., Zemeckis, D.R., Benoît, H.P., & Langan, J.A. (2016). Estimating and mitigating the discard mortality of Atlantic cod (*Gadus morhua*) in the Gulf of Maine recreational rod-and-reel fishery. *ICES J. Mar. Sci.: Journal du Conseil*, fsw058.

Cicia, A. M., Schlenker, L. S., Sulikowski, J. A., & Mandelman, J. W. (2012). Seasonal variations in the physiological stress response to discrete bouts of aerial exposure in the little skate, *Leucoraja erinacea*. *Comp. Biochem. Physiol., Part A Mol. Integr. Physiol.*, 162(2), 130-138.

Cook, K. V., Hinch, S. G., Drenner, S. M., Raby, G. G., Patterson, D. A., & Cooke, S. J. (2018). Dermal injuries caused by purse seine capture result in lasting physiological disturbances in coho salmon. *Comp. Biochem. Physiol., Part A Mol. Integr. Physiol.*, 227, 75-83.

Cox, D.R. (1972). Regression models and life tables. *J. R. Stat. Soc.*, 34, 187-200.

Cox, D.R., & Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall Ltd, London.

- Curtis, T. H., & Sosebee, K. A. (2015). Landings composition of the Northeast US skate, Rajidae, wing fishery and the effectiveness of prohibited species regulations. *Mar. Fish. Rev.*, 77(4), 1-8.
- Davis, M.W. (2002). Key principles for understanding fish bycatch discard mortality. *Can. J. Fish. Aquat. Sci.*, 59, 1834–1843.
- Davis, M.W. & Olla, B.L. (2001). Stress and delayed mortality induced in Pacific halibut *Hippoglossus stenolepis* by exposure to hooking, net towing, elevated sea water temperature and air: implications for management of bycatch. *N. Am. J. Fish. Manag.*, 21, 725-732.
- Depestele, J., Desender, M., Benoît, H. P., Polet, H., & Vincx, M. (2014). Short-term survival of discarded target fish and non-target invertebrate species in the “eurocutter” beam trawl fishery of the southern North Sea. *Fish. Res.*, 154, 82-92.
- Di Santo, V. (2015). Ocean acidification exacerbates the impacts of global warming on embryonic little skate, *Leucoraja erinacea* (Mitchill). *J. Exp. Mar. Biol. Ecol.*, 463, 72-78.
- Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, NY.
- Endicott, M., & Agnew, D. J. (2004). The survivorship of rays discarded from the South Georgia longline fishery. *CCAMLR Sci.*, 11, 155-164.
- Enever, R., Revill, A.S., & Grant, A. (2009). Discarding in the North Sea and on the historical efficacy of gear-based technical measures in reducing discards. *Fish. Res.*, 95, 40-46.
- Farrugia, T. J., Goldman, K. J., Tribuzio, C., & Seitz, A. C. (2016). First use of satellite tags to examine movement and habitat use of big skates *Beringraja binoculata* in the Gulf of Alaska. *Mar. Ecol. Prog. Ser.*, 556, 209-221.
- Hoenig, J.M., & Gruber, S.H. (1990). Life-history patterns in the elasmobranchs: implications for fisheries management. NOAA Tech. Rep. NMFS, 90, 1-16.
- Kaiser, M. J., & Spencer, B. E. (1995). Survival of by-catch from a beam trawl. *Mar. Ecol. Prog. Ser.*, 126, 31-38.
- Kaimmer, S. M., & Trumble, R. J. (1998). Injury, condition, and mortality of Pacific halibut bycatch following careful release by Pacific cod and sablefish longline fisheries. *Fish. Res.*, 38(2), 131-144.

Kelley, D. & Richards, C. (2018). oce: Analysis of Oceanographic Data. R package version 0.9-23. <https://CRAN.R-project.org/package=oce>

Knotek, R. J., Rudders, D. B., Mandelman, J. W., Benoît, H. P., & Sulikowski, J. A. (2018). The survival of rajids discarded in the New England scallop dredge fisheries. *Fish. Res.*, 198, 50-62.

Laptikhovsky, V. V. (2004). Survival rates of rays discarded by the bottom trawl squid fishery off the Falkland Islands. *Fish. Bull.*, 102(4), 757-759.

Luer, C. A. (2014). Novel compounds from shark and stingray epidermal mucus with antimicrobial activity against wound infection pathogens. Mote Marine Laboratory, Sarasota, FL.

Lukacs, P.M., Burnham, K.P., & Anderson, D.R. (2010). Model selection bias and Freedman's paradox. *Ann. Inst. Stat. Math.*, 62(1), 117-125.

Mandelman, J.W., Cicia, A.M., Ingram Jr., G.W., Driggers III, W.B., Coutre, K.M., & Sulikowski, J.A. (2013). Short-term post-release mortality of skates (family Rajidae) discarded in a western North Atlantic commercial otter trawl fishery. *Fish. Res.*, 139, 76-84.

Natanson, L. J., Sulikowski, J. A., Kneebone, J. R., & Tsang, P. C. (2007). Age and growth estimates for the smooth skate, *Malacoraja senta*, in the Gulf of Maine. In *Biology of Skates* (pp. 187-202). Springer, Dordrecht.

National Marine Fisheries Service (NMFS). (2017). Status review report: Thorny skate (*Amblyraja radiata*). Final report to National Marine Fisheries Service, Office of Protected Resources. February 2017. 60 pp.

Neilson, J. D., Waiwood, K. G., & Smith, S. J. (1989). Survival of Atlantic halibut (*Hippoglossus hippoglossus*) caught by longline and otter trawl gear. *Can. J. Fish. Aquat. Sci.*, 46(5), 887-897.

New England Fishery Management Council (NEFMC) (2003). Northeast Skate Complex Fishery Management Plan. NEFMC, Newburyport, MA, USA.

New England Fishery Management Council (NEFMC). (2014). Framework Adjustment 2 to the Northeast Skate Complex Fishery Management Plan. NEFMC, Newburyport, MA, USA.

Nye, J. A., Link, J. S., Hare, J. A., & Overholtz, W. J. (2009). Changing spatial distribution of fish stocks in relation to climate and population size on the Northeast United States continental shelf. *Mar. Ecol. Prog. Ser.*, 393, 111-129.

- Peklova, I., Hussey, N. E., Hedges, K. J., Treble, M. A., & Fisk, A. T. (2014). Movement, depth and temperature preferences of an important bycatch species, Arctic skate *Amblyraja hyperborea*, in Cumberland Sound, Canadian Arctic. *End. Sp. Res.*, 23(3), 229.
- Portz, D.E., Woodley, C.M., & Cech, J.J. (2006). Stress-associated impacts of short-term holding on fishes. *Rev. Fish. Biol. Fisheries.*, 16, 125-170.
- Raby, G. D., Packer, J. R., Danylchuk, A. J., & Cooke, S. J. (2014). The understudied and underappreciated role of predation in the mortality of fish released from fishing gears. *Fish. Fish.*, 15(3), 489-505.
- R Core Team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Ripatti, S., & Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, 56, 1016-1022.
- Ruocco, N. L., Lucifora, L. O., Díaz de Astarloa, J. M., & Wöhler, O. (2006). Reproductive biology and abundance of the white-dotted skate, *Bathyraja albomaculata*, in the Southwest Atlantic. *ICES J. Mar. Sci.*, 63(1), 105-116.
- Saygu, I., & Deval, M.C. (2014). The post-release survival of two skate species discarded by bottom trawl fisheries in Antalya Bay, Eastern Mediterranean. *Turkish J. Fish. Aqu. Sci.*, 14, 1-2.
- Schram, E., & Molenaar, P. (2018). Discards survival probabilities of flatfish and rays in North Sea pulse-trawl fisheries (No. C037/18). Wageningen Marine Research.
- Sosebee, K. A. (2005). Maturity of skates in northeast United States waters. *J. Northwest Atl. Fish. Sci.*, 35, 141-153.
- Sosebee, K., Miller, A., O'Brien, L., McElroy, D., & Sherman S. (2016). Update of Thorny Skate (*Amblyraja radiata*) Commercial and Survey Data. *US Dept. Commer., Northeast Fish. Sci. Cent. Ref. Doc.*, 16-08; 145 p.
- Spigarelli, S. A., Thommes, M. M., & Beitinger, T. L. (1977). The influence of body weight on heating and cooling of selected Lake Michigan fishes. *Comp. Biochem. Phys. A*, 56(1), 51-57.
- Stobutzki, I. C., Miller, M. J., Heales, D. S., & Brewer, D. T. (2002). Sustainability of elasmobranchs caught as bycatch in a tropical prawn (shrimp) trawl fishery. *Fish. Bull.*, 100(4), 800-821.

Sulikowski, J.A., Morin, M.D., Suk, S.H., & Howell, W.H. (2003). Age and growth estimates of the winter skate (*Leucoraja ocellata*) in the western Gulf of Maine. *Fish. Bull.*, 101, 405-413.

Sulikowski, J. A., Kneebone, J., Elzey, S., Jurek, J., Danley, P. D., Howell, W. H., & Tsang, P. C. (2005). Age and growth estimates of the thorny skate (*Amblyraja radiata*) in the western Gulf of Maine. *Fish. Bull.*

Sulikowski, J. A., Kneebone, J., Elzey, S., Jurek, J., Howell, W. H., & Tsang, P. C. W. (2006). Using the composite variables of reproductive morphology, histology and steroid hormones to determine age and size at sexual maturity for the thorny skate *Amblyraja radiata* in the western Gulf of Maine. *J. Fish Bio.*, 69(5), 1449-1465.

Sulikowski, J. A., Benoît, H. P., Capizzano, C. W., Knotek, R. J., Mandelman, J. W., Platz, T., & Rudders, D. B. (2018). Evaluating the condition and discard mortality of winter skate, *Leucoraja ocellata*, following capture and handling in the Atlantic monkfish (*Lophius americanus*) sink gillnet fishery. *Fish. Res.*, 198, 159-164.

Therneau, T.M., & Grambsch, T.M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer, New York.

Thorley, J., Fleishman, A., & Miller, L. (2017). rtide: Tide Heights. R package version 0.0.4. <https://CRAN.R-project.org/package=rtide>

Tsagarakis, K., Nikolioudakis, N., Papandroulakis, N., Vassilopoulou, V., & Machias, A. (2015). Survival of discards in a Mediterranean bottom trawl fishery. *Panhellenic Symposium on Oceanography and Fisheries*

Waring, G.T. (1984). Age, growth and mortality of the little skate off the northwest coast of the United States. *Trans. Am. Fish. Soc.*, 113, 314-321.

Wearmouth, V.J. & Sims, D.W. (2009). Movement and behaviour patterns of the critically endangered common skate *Dipturus batis* revealed by electronic tagging. *J. Exp. Mar. Bio. Ecol.*, 380, 77087.

Table 1.1. Injury code classification. Individual classes based on overt physical trauma adapted from Mandelman et al. (2013) that were used to evaluate the condition of trawl-caught thorny skate. All skates were evaluated by a single researcher to avoid any subjectivity in scoring.

Code	Description
1	None to minor physical trauma (< 10 mm lacerations, no hemorrhaging or internal bleeding) and normal coloration
2	Moderate physical trauma (11-20 mm lacerations, slight to moderate hemorrhaging or internal bleeding) and may or may not show signs of discoloration
3	Severe physical trauma (> 20 mm lacerations, extensive hemorrhaging or internal bleeding) and body discoloration

Table 1.2. Survival mixture model results. Model variants for survival mixture models (SMMs) modeled thorny skate survival as a function of total length (TL) that was treated as a two-level categorical covariate. TL’s effect on survivorship was modeled under different assumptions of model parameters α and π , which were treated either as fixed (equal to 1), constant (estimated) or as a function of TL categories, $f(\text{size})$. Model variant fit is represented by the change in Akaike’s Information Criterion corrected for small sample sizes (ΔAICc) relative to the lowest AICc value (indicated with a dash “-”). Model averaged discard mortality rates (in %; derived from survival rates generated by SMMs) are presented as the mean [standard error] from Monte Carlo simulations ($n = 5,000$) with parametric bootstrapping to simulate model uncertainty. Further description of model variant interpretation is provided in table footnotes.

Model variant	Parameters		Model fitting	Discard mortality rates	
	α	π	ΔAICc	55 - 70 cm TL	> 70 cm TL
Weibull 2 ^a	$f(\text{size})$	1	5.195	100	100
Mixture 2 ^b	$f(\text{size})$	Constant	-	21.63 [0.11]	21.63 [0.11]
Mixture 3 ^c	Constant	$f(\text{size})$	2.540	21.98 [0.11]	5.13 [0.08]
Mixture 4 ^d	$f(\text{size})$	$f(\text{size})$	0.370	22.16 [0.11]	6.67 [0.13]
Model average				24.53 [0.06]	16.50 [0.07]

^a Common survival function for each total length category

^b Common survival function within each total length category for a fixed proportion of affected individuals

^c Common survival function for affected individuals, with the proportion affected dependent on total length category

^d Common survival function within each total length category, where the proportion of affected individuals also depends on total length category

Table 1.3. Thorny skate sample sizes. Number of trawl-caught thorny skate sampled as either observational samples or tagged with pop-up satellite archival transmitting (PSAT) tags broken down for each category of handling method, sex, and injury code. Negative control PSAT-tagged skate (n = 2) are not included in this table.

Sample	Handling method		Sex		Injury code		
	Hand	Fish pick	Male	Female	Code 1	Code 2	Code 3
Observation	323	289	305	307	235	269	108
PSAT	39	34	40	33	19	28	26

Table 1.4. Fishing conditions, practices, and biological characteristics. Mean [minimum, maximum] values are provided for thorny skate released as either observational samples or with pop-off satellite archival tags (PSAT).

Variable	Observational		PSAT	
Tow duration (min)	141	[37, 236]	137	[40, 236]
Depth (m)	101	[55, 161]	99	[75, 161]
Temperature (°C)				
Bottom seawater	7.2	[6.3, 8.3]	7.1	[6.4, 8.3]
Air (on-deck)	20.8	[16.5, 26.8]	19.6	[16.5, 26.9]
Gradient	13.6	[9.3, 19.6]	12.5	[9.9, 19.6]
Catch weight (kg)	354	[29, 907]	468	[29, 907]
Air exposure (min)	30.7	[3.7, 76.7]	30.2	[5.8, 69.5]
Total length (cm)	46.9	[24.0, 94.0]	70.6	[55.5, 94.0]

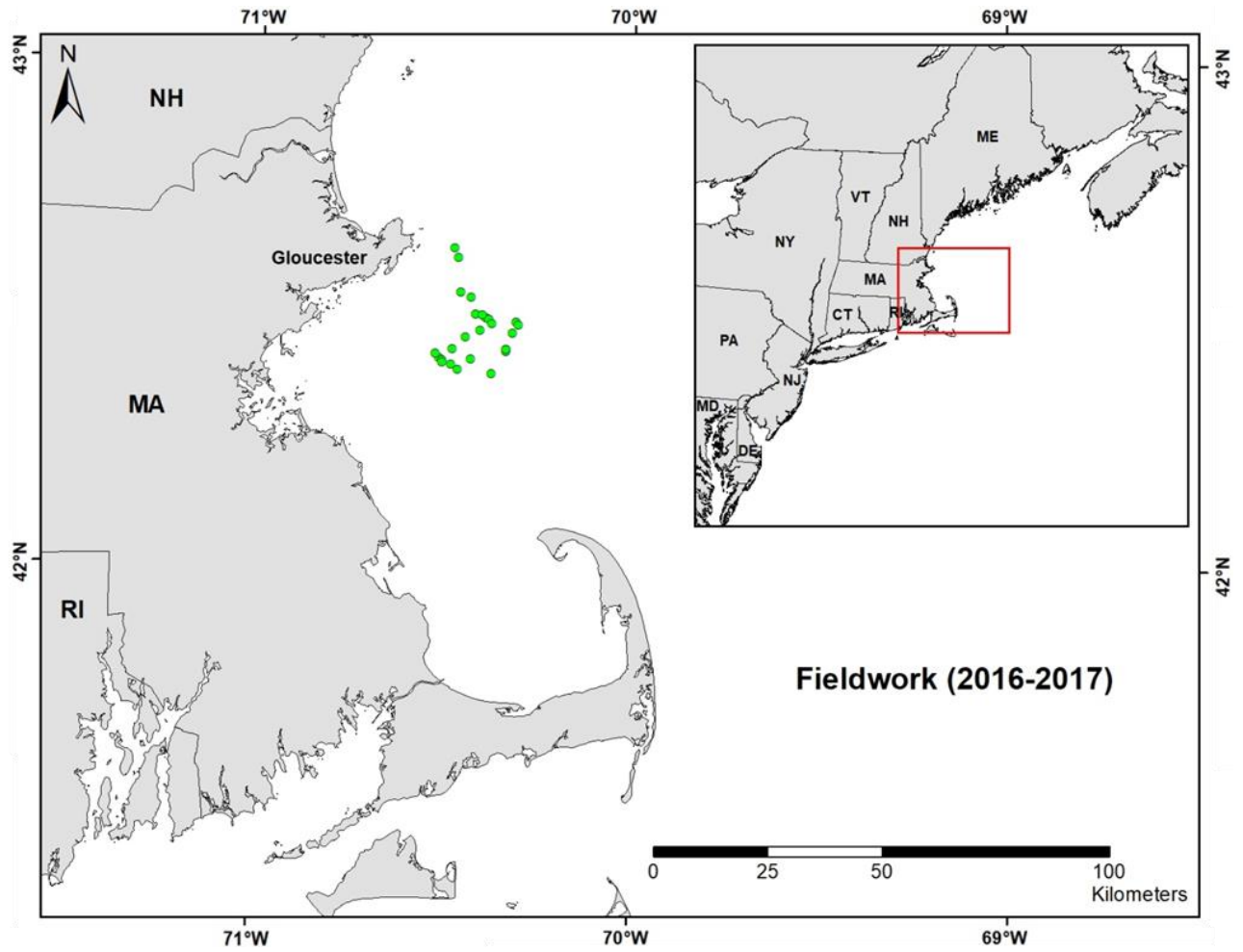


Figure 1.1. Tow locations. Individual locations ($n = 27$; depicted as light green points) are shown from the fieldwork conducted in August and September (2016-2017) aboard the *F/V Mystique Lady* in the Gulf of Maine.



Figure 1.2. Pop-up satellite archival transmitting tag attachment method. Tags (Lotek LIFE LAT3500) were attached using monofilament line and two pairs of 1" Petersen discs and baffles secured on either side of the skate (dorsal and ventral). Image in the lower-left corner displays the Petersen disc/baffle/crimp pair on the ventral side of the skate. Image on the top-left corner displays the Floy® dart tag that was inserted at the base of the tail for all skates ≥ 30 cm total length.

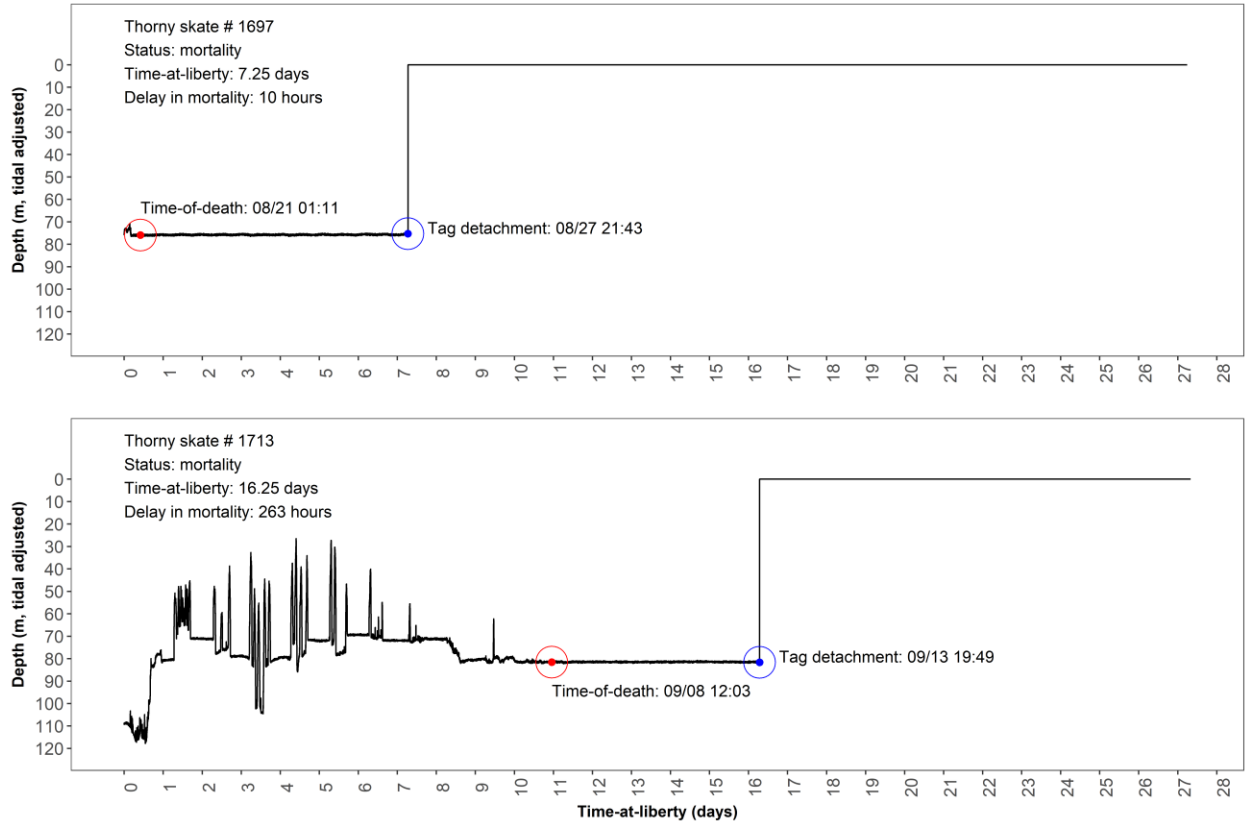


Figure 1.3. Depth time-series of two mortality events. Depth time-series (black line) were tide-adjusted and obtained from pop-up satellite archival transmitting tags (10-second resolution). Mortalities were either immediate (top panel) or delayed (~ 10 days; bottom panel). Transmission characteristics are provided in the upper left corners of the figures. Time of death derived from the depth-variance survival test is indicated with a red dot/circle and annotated with text. Time of tag detachment (i.e., pop-up) is shown with a blue dot/circle respectively.

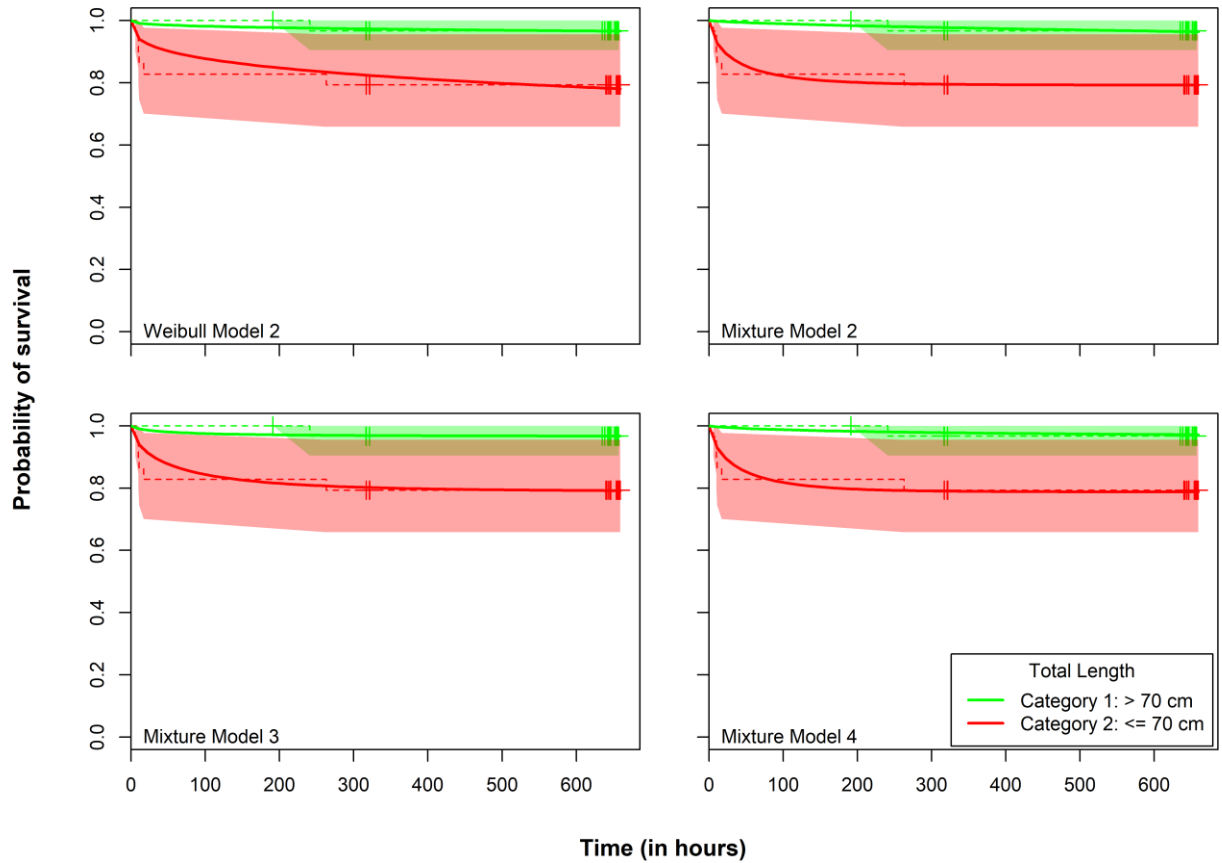


Figure 1.4. Probability of survival over time for thorny skate with respect to size-category. Size-categories are shown in varying colors for each model variant (indicated in the bottom-left corner of each panel). Dashed lines and shaded areas represent Kaplan-Meier estimates and 95% confidence bands up until the last observation for each size-category. Censored observations are indicated with “+” along the Kaplan-Meier estimate (i.e., dashed line). Solid lines show the model variant prediction.

CHAPTER 2

A SPATIOTEMPORAL BYCATCH “HOTSPOT” AVOIDANCE TOOL AND ITS APPLICATION TO THORNY SKATE, *AMBLYRAJA RADIATA*, IN THE U.S. NORTHWEST ATLANTIC OCEAN

Abstract

Fishing mortality associated with incidentally-captured and discarded fish can have population-level impacts that require attention from managers and members of the fishing industry. Attention can include avoiding areas where capture is most likely to occur (i.e., hotspots) as a direct mechanism for mitigating fishing mortality. Here, we developed a spatiotemporal bycatch avoidance tool that utilized fishing effort and species distribution/abundance data to provide a comprehensive assessment of hotspots throughout a fishery. Key components of the tool included (1) an adaptable framework to incorporate various sources of distribution or abundance data, (2) the estimation of fishery-interaction potential (FIP) based on the degree of overlap between fishing effort and species abundance, (3) the use of spatial statistics to identify areas of highest FIP as hotspots, and (4) the flexibility to further refine hotspots to account for animals that are particularly susceptible to

the stressors of capture in a given fishery. To demonstrate our method, multiple forms of data on fishing effort and species distribution were used to identify spatiotemporal bycatch hotspots for thorny skate, *Amblyraja radiata*, in three fisheries in the U.S. Northwest Atlantic Ocean. Reported size-specific discard mortality was also integrated to illustrate how hotspots can be refined with external information. Hotspots identified by this tool can assist with the development of bycatch reduction/mitigation measures.

Introduction

Fish have been an important food resource for humans throughout history (Bardey, 2019), but with increasing demand (FAO, 2018), many fish stocks have become overfished (Hall and Mainprize, 2005). Fisheries management has traditionally addressed overfishing by establishing regulations to reduce the directed fishing mortality associated with harvesting (Ward-Paige et al., 2012). However, such measures often do not mitigate fishing mortality of non-target species that are captured incidentally and discarded (as bycatch) due to regulatory measures or economic incentives (e.g., catch quotas or high-grading; Alverson et al., 1994). In many instances, incidental mortality can have a large impact on bycatch species, particularly those that have high discard rates and/or are susceptible to the stressors of capture (reviewed by Ellis et al., 2017; Veldhuizen et al., 2018), and impede the recovery of overfished stocks. Because of these negative impacts, additional management measures focused on reducing fisheries-interactions with bycatch species are often warranted to comprehensively mitigate the various sources of fishing mortality that can influence a population or stock.

There are several strategies managers can use to reduce fisheries-interactions with bycatch. For example, output controls (e.g., annual catch limits) can be set with accountability measures that restrict fishing effort after a certain level of bycatch is exceeded (e.g., windowpane flounder, *Scophthalmus aquosus*, in the Atlantic Sea Scallop fishery; 50 C.F.R. §648.64). Gear-modifications can also be required to improve selectivity towards target species (e.g., square mesh panels and sorting grids with trawl gear; Beutel et al., 2016; Lomeli et al., 2017). Or, fisheries-interactions can be reduced or avoided altogether with spatiotemporal management measures that regulate fishing in defined areas.

In particular, avoiding areas with high fisheries-interactions (i.e., bycatch hotspots) has continued to gain traction with researchers, management bodies, and members of the fishing industry because it provides a clear and direct mechanism for reducing incidental mortality while maintaining fishing opportunity in other areas where bycatch is less of a concern. Hotspots can be identified in near real-time via communication with fishers (e.g., 50 C.F.R. §635.21; Bethoney et al., 2017), or, forecasted off past spatiotemporal trends in fishing effort and/or species abundance (e.g., Cortes and Waessel, 2017; Hurley et al., 2018). Real-time hotspots have been used to effectively mitigate bycatch interactions (Bethoney et al., 2017), but tend to only perform well in restricted access fisheries with complete industry support and engagement (Little et al., 2014). Alternatively, forecasted hotspots can be used to proactively establish avoidance measures in instances where real-time communication would be logistically challenging, such as in larger, open access fisheries (Little et al., 2014). However, in order for forecasted hotspots to be effective, multiple sources of data are needed

to assess overlap in fishing effort and species abundance and comprehensively characterize hotspots throughout the spatial and temporal footprint of a fishery.

The objective of this study was to develop a spatiotemporal bycatch avoidance tool to identify practical hotspots on a case-by-case basis using the most relevant species abundance data and fishing effort within a region. To build and demonstrate the functionality of our tool, we used thorny skate, *Amblyraja radiata*, as our case study species. Thorny skate are overfished (NMFS, 2017) in the U.S. Northwest Atlantic Ocean (NWAO) and highly-susceptible to capture by demersal fishing gear (e.g., bottom otter trawl, sink gillnet, and scallop dredge; Sosebee et al., 2016). The New England Fishery Management Council (NEFMC) has prohibited their retention (i.e. mandated discard) for the past 17 years to reduce fishing mortality and promote stock recovery (NEFMC, 2003). However, the stock remains at only 4.3% of its target biomass and there is limited evidence of it rebuilding under the current prohibited measures (NEFMC, 2020). Therefore, our secondary objective was to investigate how spatiotemporal bycatch avoidance strategies may be used to complement existing measures and further reduce incidental mortality for this species in the U.S. NWAO.

Methods

The function of our spatiotemporal bycatch avoidance tool is to compare the distributions of fishing effort and species abundance within a region and determine areas of overlap where incidental capture would be most likely to occur (i.e., “hotspots”). To describe the analytics of this tool and demonstrate its implementation, we focused on identifying areas where thorny skate overlapped with the Northeast multispecies (NMS), monkfish (MNK),

and Atlantic sea scallop (SES) fisheries, which account for nearly 99% of annual thorny skate discards in the U.S. NWAO (Sosebee et al., 2016). Each fishery was analyzed separately to provide hotspots specific to each. The avoidance tool itself follows a sequence of steps that generally involved (1) joining fishing effort with species abundance data on a fixed-resolution grid, (2) evaluating cells where data overlap as fishery-interaction potential (FIP), and (3) identifying cells where FIP was the highest, as hotspots, in both space and time (Fig. 2.1). All analyses were performed using ArcGIS 10.7.1 (ESRI, 2019).

Data compilation

Commercial fishing effort

NOAA Fisheries Vessel Monitoring System (VMS) reports were used to quantify the distribution of commercial fishing effort for NMS, MNK, and SES fisheries. Fishing vessels operating in these industries are required to report once (i.e., NMS and MNK) or twice (i.e., SES) per hour when fishing seaward of the VMS Demarcation Line (50 CFR § 648.10). Each report consists of a Declaration Code that provides information about the vessel's trip (e.g., the fishery plan they are operating under), the vessel speed (in knots, kts), geographic coordinates, and date/time of the report. The VMS reports acquired spanned from 2014 to 2016 (NMS = 4,037,199; MNK = 525,008; SES = 12,712,041 reports) and did not include gear-type or permit category information. Reports were pooled across years and aggregated into three month bins beginning at the start of each fishery's fishing year to incorporate inter- and intra-annual changes to regulations (e.g., Gulf of Maine cod protection closures in the NMS fishery) and target species distributions, both of which dictate fishing effort. In the

subsequent analyses, hotspots were identified for four, three-month periods that reflect interactions in each fishery.

Because VMS reports do not specify vessel activity (e.g., on-anchor, fishing, or transitting), additional steps were needed to identify fishing effort. First, VMS reports were referenced to fishery-specific speed thresholds (Table 2.1) associated with relevant gear-types used in the fishery (Davis et al., 2015; Palmer, 2017). Because reports did not specify gear-type, lower and upper speed thresholds reflected all potential gear (Table 2.1). If the reported vessel speed was below or above the fishery threshold, the report was classified as on-anchor or transit, respectively. Location-based criteria were then used to identify instances when the vessel may have been transitting at reduced speeds (within the threshold), but was not engaged in fishing activities. Reports were reclassified as transit if they fell within three nautical miles of a major port (O’Shea et al., 2006) or in areas where fishing was prohibited throughout the timeframe of the study (e.g., Western Gulf of Maine Groundfish Closure). Without information on gear-type and/or permit category, we could not address areas where fishing was more-selectively prohibited.

Reports were also reclassified if located within “transit corridors” where transitting was the clear mode of activity and fishing was unlikely (i.e., between ports and/or fishing grounds; Watson and Haynie, 2016). However, additional steps were required to identify the footprint of fishery-specific transit corridors. Identifying these areas began with joining fishing effort and transit activity observations on a fixed 5 x 5 km grid. In each cell, activity overlap was calculated as the product of the total number of both activities (i.e., transit vs. fishing activity) within the cell’s extent. The Hot Spot analysis tool (Getis Ord G_i^* statistic;

further explained in a later section; Getis and Ord, 1992) was then used to identify and remove cells where significant overlap occurred between activities, such as the fishing grounds of Massachusetts Bay. Even if transiting occurred more frequently in these cells, they were not identified as corridors because of the significant overlap with fishing effort that suggested both activities existed. In the remaining cells where overlap was not a concern, total transit activity was evaluated (with the Getis Ord G_i^* statistic) to identify areas where transiting most frequently occurred (i.e., transit corridors), and any fishing effort observations within these cells were reclassified as transit at lower speeds.

Species distribution and relative abundance

To quantify the distribution and relative abundance of thorny skate throughout the study region, NOAA Northeast Fisheries Science Center bottom trawl (BTS) and bottom longline (LLS) survey data were compiled. The geographic extent of surveys overlapped in the Gulf of Maine (GOM; McElroy et al., 2019); however, BTS data had a larger extent that ranged from Cape Hatteras to the Canadian border (Politis et al., 2014; Fig. 2.2). Detailed descriptions of survey design and protocol can be found in Politis et al. (2014; BTS) and McElroy et al. (2019; LLS). Surveys provided the number and size (total length, TL) of thorny skate sampled at each station and a pair of geographic coordinates associated with either the deployment/haul-back of trawl gear (BTS), or terminal positions of a longline set (LLS). Midpoints were calculated between paired coordinates as the representative location of each station. Both surveys also provided a measure of effort, which included the minimum straight line distance for BTS (in km) or soak time (in hours) for LLS.

BTS and LLS data were acquired from 2009 to 2015 (n = 4,731 trawls) and 2014 to 2017 (n = 449 longline sets), respectively. Data were then pooled across years within each survey to achieve a more complete spatiotemporal representation of thorny skate throughout their U.S. geographic range. This approach was supported by the limited (< ~50 km) horizontal movement of thorny skate, both in the short- (monthly) and long-term (annual), and the lack of extensive seasonal movements in the GOM (Kneebone et al., 2020). Due to inherent differences in gear selectivity and catchability in surveys (e.g., less complex, softer bottom habitats, such as sand and mud, are sampled effectively by BTS gear, whereas LLS gear is more suited for complex, rock-bottom habitats; McElroy et al., 2019), data could not be combined into a single index of relative abundance over all substrates on which thorny skate occur. Therefore, BTS and LLS data were used separately to identify two sets of hotspots in each fishery/period that would reflect the distribution/abundance of thorny skate from the respective survey.

Calculating fishery-interaction potential (FIP)

FIP was estimated as the product of thorny skate per-unit effort (TPUE) and overall fishing activity (OFE). This value represented the relative likelihood of thorny skate interacting with a fishery inside the extent of a cell, with higher FIP values (FIP \rightarrow 1) representing an increased interaction potential. Prior to calculating FIP, fishing effort and thorny skate relative abundance were initially joined onto a fixed-resolution grid (10 x 10 km cells) that extended from North Carolina to the Canadian border (bounded by the U.S. Economic Exclusive Zone) to cover for the purported distribution of thorny skate (Fig. 2.2;

Sosebee et al., 2016). Because we were only concerned with areas where interaction potential existed, the grid was then refined to only include cells that contained fishing effort and thorny skate in survey data. TPUE was then calculated for each cell as the total number of skate divided by the total amount of sampling effort, while OFE was quantified as the total number of fishing effort observations. Both metrics were then normalized to avoid biasing FIP towards larger OFE values, which would have had a greater influence on our estimate of interaction potential relative to TPUE, using the following equation:

$$X_{ij}^{norm} = \frac{X_{ij} - X_{min_j}}{X_{max_j} - X_{min_j}} \quad (1)$$

where X_{ij}^{norm} is the normalized value of X (i.e. TPUE or OFE) from 0 to 1 of the i th observation that was calculated using the minimum and maximum values of the entire j dataset. Normalized TPUE and OFE were then used to calculate FIP.

Hotspot analysis

To identify cells where thorny skate would be expected to have the highest risk of fishery-specific interaction, FIP was evaluated using with the Hot Spot Analysis tool (Spatial Statistics, ArcGIS 10.7.1), which calculates the Getis-Ord G_i^* statistic for each cell (Getis and Ord, 1992). This calculation proportionally compares the FIP of a given cell, within the context of its neighborhood (i.e., all cells that fall within a user-defined distance band), to the global sum of FIP for the respective period. The G_i^* statistic is calculated as:

$$Gi^* = \frac{\sum_{j=1}^n w_{i,j} x_j - \bar{X} \sum_{j=1}^n w_{i,j}}{S \sqrt{\frac{n \sum_{j=1}^n w_{i,j}^2 - (\sum_{j=1}^n w_{i,j})^2}{n-1}}} \quad (2)$$

where x_j is the FIP value for grid cell j , $w_{i,j}$ is the spatial weight between grid cells i and j , n is the total number of grid cells, and \bar{X} and S are the mean and standard deviation of all FIP values within the grid, respectively. The resulting Gi^* statistic is a z-score (standard deviation) that is based on the Randomization Null Hypothesis computation. Z-scores were then converted into p-values using a one-tailed test to identify statistically significant areas of FIP at varying α -levels (0.01, 0.05, and 0.1). Hotspots were defined as cells with high z-scores and significant p-values and were displayed in figures at the highest confidence level they were detected (i.e., α -level = confidence level; 0.01 = 99%, 0.05 = 95%, 0.1 = 90%). For example, hotspots detected at all α -levels were assigned a 99% confidence level when displayed.

In calculating the Gi^* statistic, the spatial extent of the neighborhood should bear relevance to the context of the data (Getis and Ord 1992), which here should represent the potential interaction area for a single fishing event (e.g., distance trawled or length of gillnet). Therefore we chose to use neighborhood distance bands that reflected maximum distances covered by the relevant gear-types used in each respective fishery. In NMS and MNK fisheries, where bottom otter trawls and sink gillnets (primarily MNK) are used, the distance was set based on a single four-hour trawl at four kts (maximum duration and speed; 29.6 km; Palmer, 2017). This distance subsumed the smaller footprint of gillnet gear, which has

regulatory constraints on the number of nets that can be fished at once [i.e., maximum of 160 nets (91.4 m each); 50 § 648.92]. In the SES fishery, the distance band reflected the primary gear-type, a scallop dredge, towed for a maximum of 90 minutes at six kts (16.7 km; Palmer, 2017; Knotek et al., 2018).

Lastly, total hotspot area (in km²) at the each confidence level was calculated within each period/fishery to qualitatively compare the risk of fisheries-interaction both within and between fisheries. If a hotspot transected areas where fishing was prohibited (e.g., Cashes Ledge Closure Area), that portion of the grid cell was not included in the total area.

Informed bycatch hotspot analysis

Hotspots identified by this tool can also be refined to account for relevant information about bycatch, such as considerations towards animals that are more susceptible to the stressors of capture. To illustrate this functionality, thorny skate distribution/abundance data were modified to reflect a defined size range (55 – 70 cm TL) that Knotek et al. (2019) identified as has having higher discard mortality rates than larger conspecifics in the NMS fishery (bottom otter-trawl). Note that size resolution in BTS data (10 cm bins) required slightly smaller animals to also be included (i.e., 51 – 54 cm TL). These data were then used in another iteration of the tool to calculate FIP and identify hotspots solely within the NMS fishery. The resulting hotspots therefore reflected areas where the most susceptible-sized thorny skate (Knotek et al., 2019) were most likely to interact with this fishery.

Results

Spatiotemporal bycatch avoidance tool

Fishing effort was identified in 21.7 (NMS), 15.7 (MNK), and 16.9% (SES) of the original VMS reports. Incorporating BTS and LLS surveys increased the spatial extent of our evaluation to 175,104 km² (1,764 cells), including 14,596 km² (147 cells) where surveys coincided and 154,937 km² (1,560 cells) and 5,570 km² (57 cells) that were covered solely by the BTS or LLS, respectively (Fig. 2.2). FIP was calculated for a total of 2,722 cells [NMS = 1,099; NMS (size-specific) = 652; MNK = 209; SES = 762; Table 2.2]. Of these, 127 hotspots, 52 identified from LLS and 75 from BTS data (9 concurrent hotspots), were identified at the 90, 95, and 99% confidence levels (Table 2.2).

Thorny skate hotspots

In the NMS fishery, the most hotspots were identified between the months of May and July (hotspot area = 1,427 km²) and least from February to April (hotspot area = 848 km²; Table 2.2). The location of hotspots varied between period and survey; however, consistently occurred within Massachusetts Bay and off of the Southwest edge of the Western Gulf of Maine Closure Area (Fig. 2.3). Hotspots associated with LLS data were also consistently identified throughout the year off the Southeast edge of the Cashes Ledge Closure Area, whereas hotspots derived with BTS were clustered off Cape Cod between the months of November and January (Fig. 2.3). Informed NMS hotspots for thorny skate from 51 to 70 cm TL had a more concentrated spatial extent than what was identified for thorny skate of all sizes (i.e., 21.2% reduction in total hotspot area throughout the year; Table 2.2).

However, hotspots displayed similar spatiotemporal trends, with an additional hotspot revealed further offshore on Georges Bank between February and April (Fig. 2.4).

The MNK fishery had the fewest number of hotspots throughout the year (hotspot area = 100 to 341 km²; Fig. 2.3), relative to NMS and SES, and data from both surveys pointed to Platts Bank and the Southeast edge of the Cashes Ledge Closure Area as being the locations with the most potential for interaction (Fig. 2.3). The LLS data also suggested that Southern Jeffrey's Ledge may present a risk to thorny skate interacting with this fishery throughout most of the year (Fig. 2.5).

The highest number of hotspots in the SES fishery was identified between the period of March to May (hotspot area = 703 km²) and December to February (hotspot area = 770 km²), with the fewest hotspots from June to August (hotspot area = 281 km²; Table 2.2). Hotspots were regularly identified throughout the year off the Southeast edge of the Western Gulf of Maine Closure Area (Fig. 2.6). Hotspots associated with BTS data also occurred off Cape Cod throughout the year, with the winter months of December to February presenting the greatest interaction risk and occasional hotspots in the Great South Channel (Fig. 2.6).

Discussion

The spatiotemporal bycatch avoidance tool described uses overlap between fishing effort and species abundance to estimate the relative likelihood of interaction and identifies hotspots where interaction potential is the highest in both space and time. This method also allows users to choose input data that best-reflects the focal fishery and species, with adjustable parameters to control the spatiotemporal resolution of the analysis. The outputs of

this tool can be used to evaluate bycatch risk of a species (or group of species) in multiple fisheries and used to develop practical bycatch avoidance strategies for reducing fisheries-interactions and incidental mortality in a given fishery or region. Identified hotspots can be considered by fishery managers when developing bycatch mitigation strategies tailored at rebuilding the severely depleted thorny skate population in the U.S. NWAO.

Bycatch avoidance tool advantages, successes, and technical considerations

The easily adjustable framework of this tool allows users to modify the analysis to meet the specific needs of a species and/or fishery. Of potentially the greatest utility is the ability to specify the spatiotemporal resolution of hotspots to best describe the resolution of the (abundance) data and the nature of a fishery and its regulations. For example, in our case study we chose to generate thorny skate hotspots using three-month periods to incorporate seasonal shifts in fishing effort and changes to regulations that would produce the most applicable results for fishery managers in New England. However, temporal resolution could be easily increased to a month-to-month basis to cater towards more mobile species that display more rapid changes to their distribution (e.g., spiny dogfish, *Squalus acanthias*; Sulikowski et al., 2010). Alternatively, temporal resolution could be adjusted to reflect specific changes in regulations that prohibit fishing effort and alter FIP (e.g., Sea Scallop Rotational Areas). In addition, while 10 x 10 km grid cells were appropriate for capturing small-scale seasonal horizontal movements of thorny skate (Kneebone et al., 2020), increased spatial resolution would benefit species with high site-fidelity to complex habitat (e.g., cusk or wolfish, *Anarhichas lupus*; Davies and Jonsen, 2011, Simpson et al., 2015). Conversely, a

coarser resolution would be suitable for species exhibiting more pronounced movement behavior (e.g., spiny dogfish; Sulikowski et al., 2010).

Nested into the tool is also a simple, step-wise method for differentiating fishing and non-fishing activity in NOAA Fisheries VMS data. Previous studies have identified fishing effort in VMS data using fishery-specific speed-based thresholds (Lee et al., 2010; Gerritsen and Lordan, 2011) and sophisticated statistical modeling that incorporates information not available in U.S. VMS data (e.g., vessel heading; Bez et al., 2001; Guillot et al., 2017). The sole use of speed thresholds likely resulted in the overestimation of fishing effort due to scenarios of reduced transiting speeds by vessels to, among other reasons, avoid pot buoys and/or other boats while transiting near-shore and in/out of ports (Mills et al., 2007). Therefore, to minimize the risk of falsely classifying VMS reports and consequently bias in FIP, we chose to integrate our fishery-specific fishing speed thresholds (Davies et al., 2015; Palmer, 2017) with location-based criteria that focused on observations associated with areas of reduced transiting speeds. Location-based criteria included port-distance and prohibited area restrictions, as well as an overlay approach that reclassified fishing effort as transiting when located within transit corridors. Ultimately, we believe our integrated approach will provide users with a more realistic depiction of fishing effort based on VMS data.

Hotspots are also provided at varying levels of confidence (i.e., 90, 95 and 99%) to allow managers and other end-users the flexibility to select the level of confidence that matches the management or conservation need. The 90% confidence level is the most conservative approach for identifying hotspots and results in the largest area for mitigation consideration (Table 2.2), whereas the 99% confidence level is a more focused approach

where hotspots are only identified in areas with highest FIP. Bycatch mitigation measures afforded under the conservative approach (90% confidence level) provide the most comprehensive strategy for addressing fisheries-interactions and overall fishing mortality, which may be required for overfished species that have heightened management concerns (e.g., thorny skate or Atlantic cod, *Gadus morhua*; NMFS 2017; NEFSC 2017), high discard rates, moderate-to-high discard mortality, and/or limited mechanisms for reducing the adverse impacts of incidental capture-and-handling. Taking a conservative approach may also be warranted for ecologically important areas where aggregations of key life-history stages are vulnerable to fishing effort, such as nursery grounds or migration pathways (Dunn et al., 2011; Suuronen and Gilman, 2019). In contrast, more resilient species that are less of a management concern might only require mitigation measures focused on areas with the highest likelihood of fisheries-interactions (99% confidence level). Hotspots identified at varying confidence levels can also be addressed in concert with each other, wherein regulations become more stringent as the likelihood of fisheries-interaction increases. For example, managers could prohibit fishing in hotspots detected at the 99% confidence level, and require best fishing practices and/or gear-modifications in the remaining hotspots to promote survivability and lower interaction rates (Beutel et al., 2016; Lomeli et al., 2017; reviewed by Cook et al., 2019).

This tool can also be informed by other capture-related information (e.g., factors driving discard mortality rates) to further increase the potential impact of avoiding hotspots. In the NMS fishery, hotspots were isolated for thorny skate within a given size range shown to be more susceptible to the impacts of capture (Knotek et al., 2019). Therefore, avoiding

these particular locations represents a targeted means for reducing incidental fishing mortality for this species. As such, we strongly suggest that relevant biological factors driving discard mortality (e.g., fish size, sex, and life-stage) are taken into consideration when addressing prohibited bycatch species. For example, fish size has been shown to dictate the impacts of capture-and-handling in several studies (e.g., Benoît, 2006; Mandelman et al., 2013; Depestele et al., 2014; Ellis et al., 2018) and is typically available in survey datasets. The inclusion of sex-specific information may also be particularly useful for elasmobranchs with sex-specific discard mortality rates (e.g., Coehlo et al., 2012; Campbell et al., 2018; Sulikowski et al., 2018), because of easily identifiable male reproductive organs (i.e., claspers) that can be recorded during surveys.

Hotspot applicability (i.e., relevance to actual fisheries-interactions) hinges upon how well fishing effort and species abundance data reflects the true distribution and overlap of these metrics. In this study, NOAA Fisheries Vessel Monitoring Systems thoroughly characterized fishing effort of the focal industries due to mandated reporting requirements (50 CFR § 648.10); however, not all fisheries are monitored with this level of detail. Other sources that could be used to address fishing effort include logbooks, observer programs, or Automatic Identification Systems (AIS) that operate similarly to VMS data (Kroodsma et al., 2018). Recent studies have demonstrated the utility of AIS data as a measure of fishing effort, particularly on a global-scale where it is the most widely-implemented form of vessel monitoring (Kroodsma et al., 2018; White et al., 2019). Fishery-independent surveys can be used as a reliable source of species distribution and relative abundance (Politis et al., 2014).

But, they may be less applicable or effective for species that inhabit areas not accessible to the gear and/or whose distributions extend outside of the surveying extent.

Here, we demonstrated how multiple surveys can be combined to comprehensively evaluate the relative abundance of thorny skate in the U.S. NWAO (i.e., NEFSC bottom trawl and bottom longline surveys; Sosebee et al., 2016; McElroy et al., 2019). Multiple sources of data with different structures (e.g., different metrics of sampling effort) can be incorporated without altering the functionality or end-product of this tool because of the normalization step that transforms data onto the same scale (i.e., 0 to 1; Eq. 1). As such, probability density functions (e.g., kernel density estimator; NEFSC 2019) or species distribution models (e.g., flapper skate, *Dipturus cf. intermedia*; Pinto et al., 2016) can be used to describe the distribution and abundance of a species. Distribution models in particular may also provide a mechanism for forecasting shifts in fisheries-interactions if climate change alters future species' distributions (e.g., Kleisner et al., 2017).

The hotspots identified by this tool are reliable predictions of FIP when the distribution of species abundance and fishing effort remain relatively constant from year-to-year. In the thorny skate case study, hotspot predictions are reliable because of the restricted movement ecology of thorny skate (Kneebone et al., 2020) and regulatory prohibitions that are reflected in the data and expected to perpetuate (e.g., Sea Scallop Rotational Areas and Open Access Areas). However, careful attention must be paid towards future changes to regulations that prohibit fishing in locations not accounted for in the current data, which can ultimately displace fishing effort and change the dynamics of FIP. Predictions may also be less relevant when considering more active schooling species whose distributions can change

rapidly (e.g., river herring and American shad; *Alosa sapidissima*) and alter FIP outside of what this tool is capable of accounting for (Little et al., 2014). In those scenarios, real-time bycatch avoidance programs are likely the more suitable approach given their ability to adapt to dynamic fluctuations in FIP (Little et al., 2014; Bethoney et al., 2017).

Thorny skate hotspots in U.S. NWAQ

Hotspots from our case study suggest the NMS had the highest likelihood of interacting with thorny skate throughout timeframe of the study (58.9% of all hotspots), followed by SES (28.9% of all hotspots) and MNK (12.2% of all hotspots; Table 2.2). These findings were consistent with thorny skate discards by gear type reported from the same timeframe, which indicate otter-trawls account for nearly three-quarters (73.0%) of all discards (Sosebee et al., 2016). Fisheries managers (i.e., NEFMC) can pair these practical hotspots with bycatch avoidance strategies to complement pre-existing prohibited measures and more comprehensively address fishing mortality for this overfished species. For example, managers could consider avoiding/reducing interactions in hotspots within Massachusetts Bay for the NMS fishery (Figs. 2.3 and 2.4), Platts Bank in the MNK fishery (Fig. 2.5), and off of Cape Cod in the SES fishery (Fig. 2.6), all of which present the highest risk of incidental capture to thorny skate within each fishery. Or, mitigation efforts could be targeted toward the NMS fishery that was most likely to interact with thorny skate throughout the year (Table 2.2), with an additional emphasis placed on avoiding/reducing capture of the size-class found to be more susceptible to the stressors of capture (Knotek et al., 2019; Fig. 2.3). Lastly, providing fishing industries with hotspot locations can serve as a

voluntary avoidance method for reducing interactions; however, the efficacy of this method would likely require additional output controls for thorny skate (i.e., bycatch limits; Reid et al., 2019) to further incentivize fishers to avoid their capture.

Future applications

The bycatch avoidance hotspot tool described herein can be broadly used to address any fisheries-interaction wherein both fishing effort and species distribution/abundance data are available. In particular, data sources collected here can be directly applied to other prohibited and/or overfished species incidentally captured by demersal fisheries in the U.S. NWA0 (e.g., Atlantic cod; yellowtail flounder, *Limanda ferruginea*; winter flounder, *Pseudopleuronectes americanus*; witch flounder, *Glyptocephalus cynoglossus*; windowpane flounder, *Scophthalmus aquosus*; halibut, *Hippoglossus stenolepis*; and white hake, *Urophycis tenuis*; NMFS 2017; NEFSC 2019). It may also be possible to merge hotspots from species within the same region (i.e., U.S. NWA0) to identify areas where mitigation efforts could target fisheries-interactions for multiple species. Furthermore, this tool may prove useful for pelagic species that are overfished (e.g., porbeagle shark, *Lamna nasus*, and shortfin mako shark, *Isurus oxyrinchus*; Curtis et al., 2016; ICCAT 2017); however, additional modifications may be required to incorporate data (e.g., utilization distributions derived from satellite tag data; Vaudo et al., 2017) from more migratory species that are not adequately addressed by traditional surveys.

Acknowledgements

I would like to thank all of the co-authors on this forthcoming manuscript (J. Mandelman, J. Kneebone, T. Curtis, and J. Sulikowski) and also K. Sosebee and D. McElroy for providing the NEFSC bottom trawl and bottom longline survey data, respectively. Funding for this research was provided by the 2015 NOAA NMFS Bycatch Reduction Engineering Program (Award # NA15NMF4720373).

Literature cited

Alverson, D. L., Freeberg, M. H., Murawski, S. A., & Pope, J. G. (1994). *A global assessment of fisheries bycatch and discards* (Vol. 339). Food & Agriculture Org..

Bardey, D. J. (2019). Overfishing: pressure on our oceans. *Research in Agriculture Livestock and Fisheries*, 6(3), 397-404.

Benoît, H. P. (2006). Estimated discards of winter skate (*Leucoraja ocellata*) in the southern Gulf of St. Lawrence, 1971-2004. *Fisheries and Oceans Canada*, Science.

Bethoney, N. D., Schondelmeier, B. P., Kneebone, J., & Hoffman, W. S. (2017). Bridges to best management: effects of a voluntary bycatch avoidance program in a mid-water trawl fishery. *Marine Policy*, 83, 172-178.

Beutel, D., Skrobe, L., Castro, K., Ruhle Sr, P., Ruhle Jr, P., O'Grady, J., & Knight, J. (2008). Bycatch reduction in the Northeast USA directed haddock bottom trawl fishery. *Fisheries Research*, 94(2), 190-198.

Bez, N., Walker, E., Gaertner, D., Rivoirard, J., & Gaspar, P. (2011). Fishing activity of tuna purse seiners estimated from vessel monitoring system (VMS) data. *Canadian Journal of Fisheries and Aquatic Sciences*, 68(11), 1998-2010.

Campbell, M. J., McLennan, M. F., Courtney, A. J., & Simpfendorfer, C. A. (2018). Post-release survival of two elasmobranchs, the eastern shovelnose ray (*Aptychotrema rostrata*) and the common stingaree (*Trygonoptera testacea*), discarded from a prawn trawl fishery in southern Queensland, Australia. *Marine and Freshwater Research*, 69(4), 551-561.

Coelho, R., Fernandez-Carvalho, J., Lino, P. G., & Santos, M. N. (2012). An overview of the hooking mortality of elasmobranchs caught in a swordfish pelagic longline fishery in the Atlantic Ocean. *Aquatic Living Resources*, 25(4), 311-319.

Cook, K. V., Reid, A. J., Patterson, D. A., Robinson, K. A., Chapman, J. M., Hinch, S. G., & Cooke, S. J. (2019). A synthesis to understand responses to capture stressors among fish discarded from commercial fisheries and options for mitigating their severity. *Fish and Fisheries*, 20(1), 25-43.

Cortés, F., & Waessle, J. A. (2016). Hotspots for porbeagle shark (*Lamna nasus*) bycatch in the southwestern Atlantic (51° S–57° S). *Canadian Journal of Fisheries and Aquatic Sciences* 74(7):1100-1110.

Curtis, T., Cortes, E., DuBeck, G., & McCandless, C. T. (2016). Statues review report: porbeagle shark (*Lamna nasus*).

Davies, T. D., & Jonsen, I. D. (2011). Identifying nonproportionality of fishery-independent survey data to estimate population trends and assess recovery potential for cusk (*Brosme brosme*). *Canadian Journal of Fisheries and Aquatic Sciences*, 68(3), 413-425.

Depestele, J., Desender, M., Benoît, H. P., Polet, H., & Vincx, M. (2014). Short-term survival of discarded target fish and non-target invertebrate species in the “eurocutter” beam trawl fishery of the southern North Sea. *Fisheries Research*, 154, 82-92.

Dunn, D. C., Boustany, A. M., & Halpin, P. N. (2011). Spatio-temporal management of fisheries to reduce by-catch and increase fishing selectivity. *Fish and Fisheries*, 12(1), 110-119.

Ellis, J. R., Burt, G. J., Grilli, G., McCully Phillips, S. R., Catchpole, T. L., & Maxwell, D. L. (2018). At-vessel mortality of skates (Rajidae) taken in coastal fisheries and evidence of longer-term survival. *Journal of Fish Biology*, 92(6), 1702-1719.

FAO. (2018). The State of World Fisheries and Aquaculture 2018 - Meeting the sustainable development goals. Rome. Licence: CC BY-NC-SA 3.0 IGO.

Gerritsen, H., & Lordan, C. (2011). Integrating vessel monitoring systems (VMS) data with daily catch data from logbooks to explore the spatial distribution of catch and effort at high resolution. *ICES Journal of Marine Science*, 68(1), 245-252.

Getis, A., and Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographic Analysis* 24(3):189-206

Guillot, G., Benoit, P., Kinalis, S., Bastardie, F., & Bartolino, V. (2017). Enhancing and comparing methods for the detection of fishing activity from Vessel Monitoring System data. *arXiv preprint arXiv:1708.09663*.

Hall, S. J., & Mainprize, B. M. (2005). Managing by-catch and discards: how much progress are we making and how can we do better? *Fish and Fisheries*, 6(2), 134-155.

ICCAT, S. (2017). Report of the Standing Committee on Research and Statistics (SCRS), Madrid, Spain, October 2–6, 2017. ICCAT, Madrid, Spain.

Kleisner, K. M., Fogarty, M. J., McGee, S., Hare, J. A., Moret, S., Perretti, C. T., & Saba, V. S. (2017). Marine species distribution shifts on the US Northeast Continental Shelf under continued ocean warming. *Progress in Oceanography*, 153, 24-36.

Kneebone, J., Sulikowski, J., Knotek, R., McElroy, W. D., Gervelis, B., Curtis, T., Jurek, J., & Mandelman, J. Using conventional and pop-up satellite transmitting tags to assess the horizontal movements and habitat use of thorny skate (*Amblyraja radiata*) in the Gulf of Maine. *ICES Journal of Marine Science*.

Knotek, R., Kneebone, J., Sulikowski, J., Curtis, T., Jurek, J., & Mandelman, J. (2020). Utilization of pop-up satellite archival transmitting tags to evaluate thorny skate (*Amblyraja radiata*) discard mortality in the Gulf of Maine groundfish bottom trawl fishery. *ICES Journal of Marine Science*, 77(1), 256-266.

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T.D., Bock, B.A., & Woods, P. (2018). Tracking the global footprint of fisheries. *Science*, 359(6378), 904-908.

Lee, J., South, A. B., & Jennings, S. (2010). Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data. *ICES Journal of Marine Science*, 67(6), 1260-1271.

Little, A. S., Needle, C. L., Hilborn, R., Holland, D. S., & Marshall, C. T. (2015). Real-time spatial management approaches to reduce bycatch and discards: experiences from Europe and the United States. *Fish and Fisheries*, 16(4), 576-602.

Lomeli, M. J., Wakefield, W. W., & Herrmann, B. (2017). Testing of two selective flatfish Sorting-Grid bycatch reduction devices in the US West Coast groundfish bottom trawl fishery. *Marine and Coastal Fisheries*, 9(1), 597-611.

Mandelman, J. W., Cicia, A. M., Ingram Jr, G. W., Driggers III, W. B., Coutre, K. M., & Sulikowski, J. A. (2013). Short-term post-release mortality of skates (family Rajidae) discarded in a western North Atlantic commercial otter trawl fishery. *Fisheries Research*, 139, 76-84.

McElroy, W. D., O'Brien, L., Blaylock, J., Martin, M. H., Rago, P. J., Hoey, J. J., Seremet, V. A. (2019) Design, implementation, and results of a cooperative research Gulf of Maine longline survey, 2014-2017. NOAA Tech Memo NMFS NE. 249; 155 p. Online at: <https://doi.org/10.25923/2sgn-mx62>

Mills, C. M., Townsend, S. E., Jennings, S., Eastwood, P. D., & Houghton, C. A. (2007). Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. *ICES Journal of Marine Science*, 64(2), 248-255.

New England Fishery Management Council (NEFMC). (2003). Northeast Skate Complex Fishery Management Plan. NEFMC, Newburyport, MA, USA.

New England Fishery Management Council (NEFMC). (2020). Northeast Skate Complex Fishery Management Plan. Annual monitoring report for fishing year 2019. NEFMC, Newburyport, MA, USA.

Northeast Fisheries Science Center (NEFSC). (2017). Operational Assessment of 19 Northeast Groundfish Stocks, Updated Through 2016. US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center.

Northeast Fisheries Science Center (NEFSC). (2019). Operational Assessment of 14 Northeast Groundfish Stocks, Updated Through 2018. US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center.

National Marine Fisheries Service (NMFS). (2017). Status review report: Thorny skate (*Amblyraja radiata*). Final report to National Marine Fisheries Service, Office of Protected Resources. February 2017. 60 pp.

Palmer, M. C. (2017). Vessel trip reports catch-area reporting errors: potential impacts on the monitoring and management of the northeast United States groundfish resource.

Pinto, C., Thorburn, J. A., Neat, F., Wright, P. J., Wright, S., Scott, B. E., Cornulier, T., & Travis, J. M. (2016). Using individual tracking data to validate the predictions of species distribution models. *Diversity and Distributions*, 22(6), 682-693.

Politis, P. J., Galbraith, J. K., Kostovick, P., Brown, R. W. (2014). Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow. US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-06; 138 p.

Reid, D. G., Calderwood, J., Afonso, P., Bourdaud, P., Fauconnet, L., González-Irusta, J. M., Mortensen, L.O., Ordines, F., Lehuta, S., Pawlowski, L., & Plet-Hansen, K. S. (2019). The best way to reduce discards is by not catching them!. *The European Landing Obligation*, 257.

Simpson, M. (2015). A preliminary analysis of habitat use and movement patterns of Wolffish (*Anarhichas* spp.) in coastal Newfoundland waters. Canadian Science Advisory Secretariat (CSAS).

Sosebee, K. A., O'Brien, L., McElroy, D., & Sherman, S. (2016). Update of thorny skate (*Amblyraja radiata*) commercial and survey data.

Sulikowski, J. A., Galuardi, B., Buble, W., Furey, N. B., Driggers III, W. B., Ingram Jr, G. W., & Tsang, P. C. (2010). Use of satellite tags to reveal the movements of spiny dogfish

Squalus acanthias in the western North Atlantic Ocean. *Marine Ecology Progress Series*, 418, 249-254.

Sulikowski, J. A., Benoît, H. P., Capizzano, C. W., Knotek, R. J., Mandelman, J. W., Platz, T., & Rudders, D. B. (2018). Evaluating the condition and discard mortality of winter skate, *Leucoraja ocellata*, following capture and handling in the Atlantic monkfish (*Lophius americanus*) sink gillnet fishery. *Fisheries Research*, 198, 159-164.

Suuronen, P., & Gilman, E. (2019). Monitoring and managing fisheries discards: New technologies and approaches. *Marine Policy*, 116 103554.

Ward-Paige, C. A., Keith, D. M., Worm, B., & Lotze, H. K. (2012). Recovery potential and conservation options for elasmobranchs. *Journal of Fish Biology* 80(5), 1844-1869.

Watson, J. T., & Haynie, A. C. (2016). Using Vessel Monitoring System data to identify and characterize trips made by fishing vessels in the United States North Pacific. *PLoS one*, 11(10), e0165173.

White, T. D., Ferretti, F., Kroodsma, D. A., Hazen, E. L., Carlisle, A. B., Scales, K. L., Bograd, S.J., & Block, B. A. (2019). Predicted hotspots of overlap between highly migratory fishes and industrial fishing fleets in the northeast Pacific. *Science advances*, 5(3), eaau3761.

Vaudo, J. J., Byrne, M. E., Wetherbee, B. M., Harvey, G. M., & Shivji, M. S. (2017). Long-term satellite tracking reveals region-specific movements of a large pelagic predator, the shortfin mako shark, in the western North Atlantic Ocean. *Journal of Applied Ecology*, 54(6), 1765-1775.

Veldhuizen, L. J. L., Berentsen, P. B. M., De Boer, I. J. M., Van De Vis, J. W., & Bokkers, E. A. M. (2018). Fish welfare in capture fisheries: A review of injuries and mortality. *Fisheries Research*, 204, 41-48.

Table 2.1. Vessel-speed thresholds. Speeds (knots; kts) were used to distinguish fishery-specific fishing effort from on-anchor or transit activity in NOAA Vessel Monitoring Systems reports.

Fishing industry	On-anchor (kts)	Fishing (kts)^a	Transit (kts)
Northeast multispecies	< 0.1	0.1 – 4.0	4.0 – 20.0
Monkfish	< 0.1	0.1 – 4.0	4.0 – 20.0
Sea scallop	< 0.1	0.1 – 6.0	6.0 – 20.0

^a Fishing speeds reflected standard commercial practices of relevant gear types according to Davis *et al.* (2015) and Palmer (2017).

Table 2.2. Hotspot characteristics by evaluation period for each fishery examined in the U.S. Northwest Atlantic Ocean. The number of grid cells identified as hotspots at varying confidence levels (90% | 95% | 99%) and the number of cells identified with interaction potential (*italicized*) based on Northeast Fisheries Science Center bottom trawl and bottom longline surveys are presented. Total hotspot area (km²) at varying confidence levels (90% | 95% | 99%) is provided based on cells where hotspots were identified. The total number of grids cells and hotspot area over all periods for each fishery are shown in bold. Note that the months included in each period differ by fishery.

Fishery	Trawl survey			Longline survey			Hotspot (km ²)				
Northeast multispecies (NMS)											
May-July	10	7	5	<i>176</i>	7	7	5	<i>105</i>	1427	1127	741
Aug-Oct	9	8	6	<i>180</i>	4	4	3	<i>104</i>	1219	1119	838
Nov-Jan	7	7	4	<i>172</i>	5	5	3	<i>99</i>	982	982	688
Feb-Apr	4	3	3	<i>164</i>	7	6	4	<i>99</i>	848	658	493
	30	25	18	<i>692</i>	24	23	15	<i>407</i>	4476	3886	2761
Northeast multispecies (NMS) (51-70 cm TL)^a											
May-July	4	4	3	<i>85</i>	7	5	2	<i>80</i>	1021	867	482
Aug-Oct	6	6	2	<i>91</i>	6	5	2	<i>79</i>	1067	867	367
Nov-Jan	2	2	2	<i>88</i>	5	3	2	<i>74</i>	586	400	300
Feb-Apr	7	5	2	<i>80</i>	5	4	3	<i>75</i>	855	665	486
	19	17	9	<i>344</i>	23	17	9	<i>308</i>	3529	2800	1634
Monkfish (MNK)											
May-July	1	1	1	<i>31</i>	3	2	2	<i>35</i>	341	241	241
Aug-Oct	1	1	1	<i>31</i>	2	2	0	<i>19</i>	300	100	100
Nov-Jan	1	1	1	<i>31</i>	1	1	1	<i>24</i>	200	200	200
Feb-Apr	1	1	1	<i>17</i>	1	1	1	<i>21</i>	100	100	100
	4	4	4	<i>110</i>	7	6	4	<i>99</i>	941	641	541
Sea Scallop (SES)											
Mar-May	6	6	5	<i>126</i>	2	2	2	<i>89</i>	703	703	615
June-Aug	1	1	1	<i>97</i>	2	2	1	<i>56</i>	281	281	181
Sept-Nov	4	4	4	<i>137</i>	3	3	1	<i>68</i>	581	581	500
Dec-Feb	7	5	5	<i>123</i>	1	1	1	<i>66</i>	770	588	588
	18	16	15	<i>483</i>	8	8	5	<i>279</i>	2335	2154	1884

^a Northeast Fisheries Science Center bottom trawl and bottom longline survey data subset for thorny skate ranging from 51 to 70 cm total length (TL).

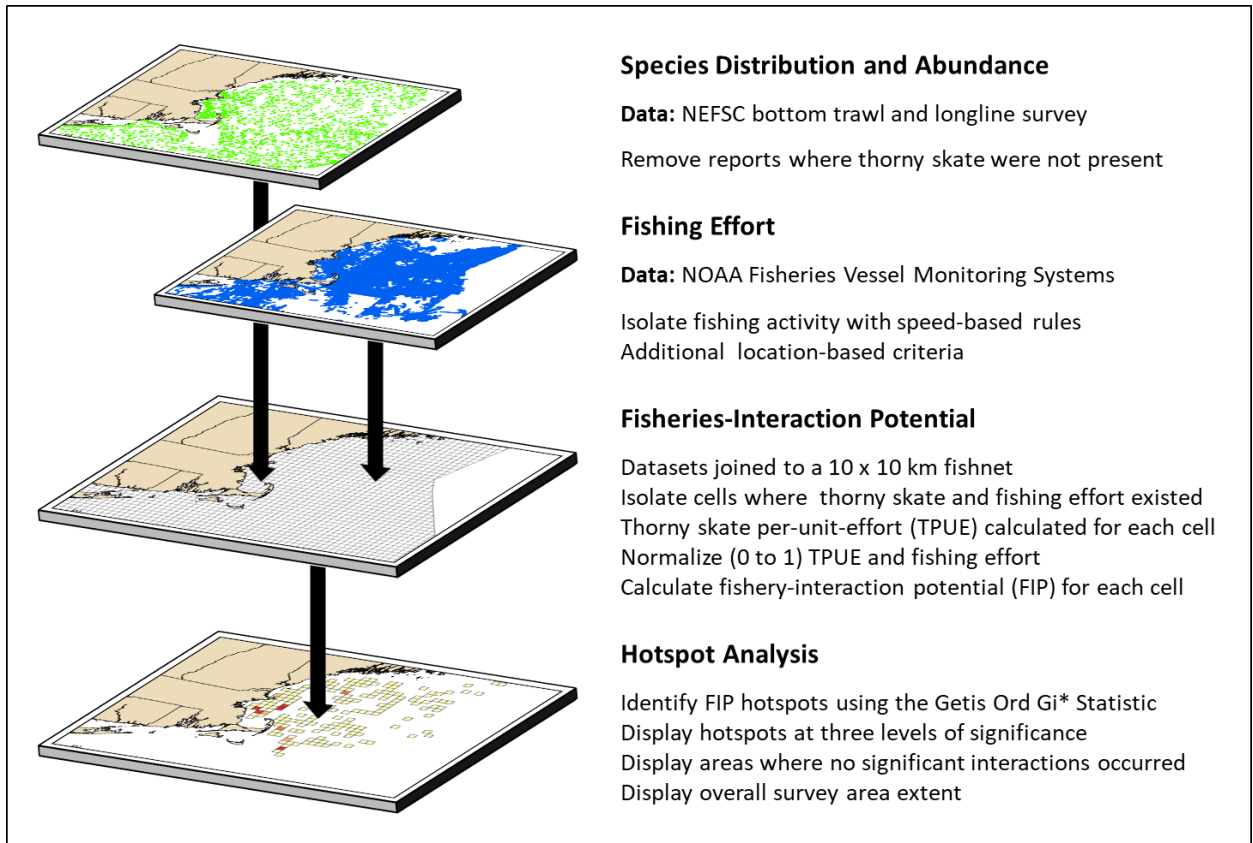


Figure 2.1. Diagram of the bycatch hotspot avoidance tool workflow. Key steps are provided along with the data sources used to identify thorny skate fisheries-interaction potential hotspots.

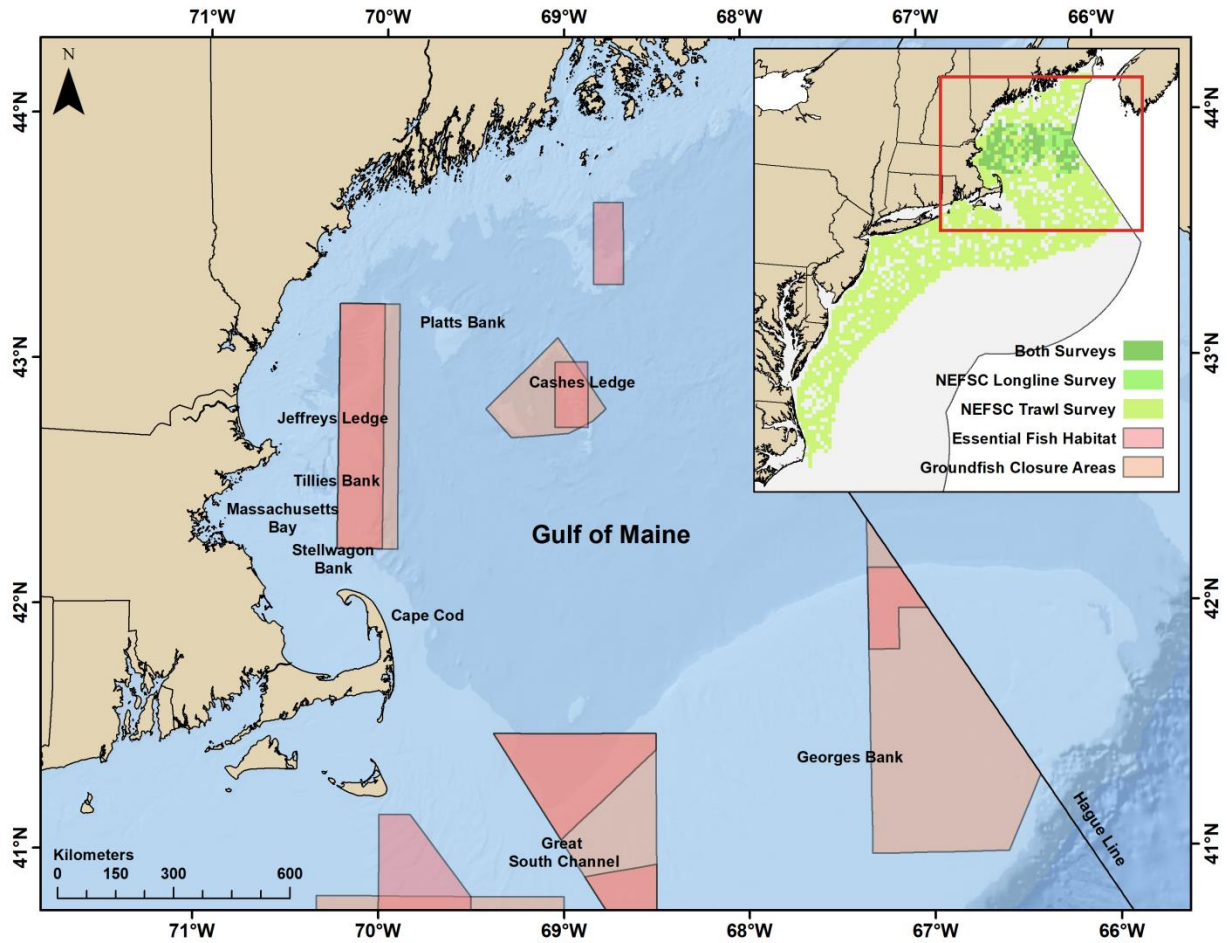


Figure 2.2. Relevant spatial extent of potential fisheries-interactions. Fisheries-interactions were between thorny skate and the Northeast multispecies, monkfish, and Atlantic sea scallop fisheries. Management areas where all fishing effort was prohibited throughout the duration of the study are illustrated with transparent red (i.e., Essential Fish Habitat areas) and pink (i.e., Groundfish Closures Areas) polygons. Major bathymetric features are denoted with text and depth contours shaded in blue. Inset map (upper-right corner) displays the overall spatial extent of 10km x 10km grid (shown with grey fill), with cells that contained Northeast Fisheries Science Center (NEFSC) bottom trawl and bottom longline surveys highlighted in shades of green. Bathymetry sourced from Esri, GEBCO, NOAA, National Geographic, DeLorme, HERE, Geonames.org, and other contributors.

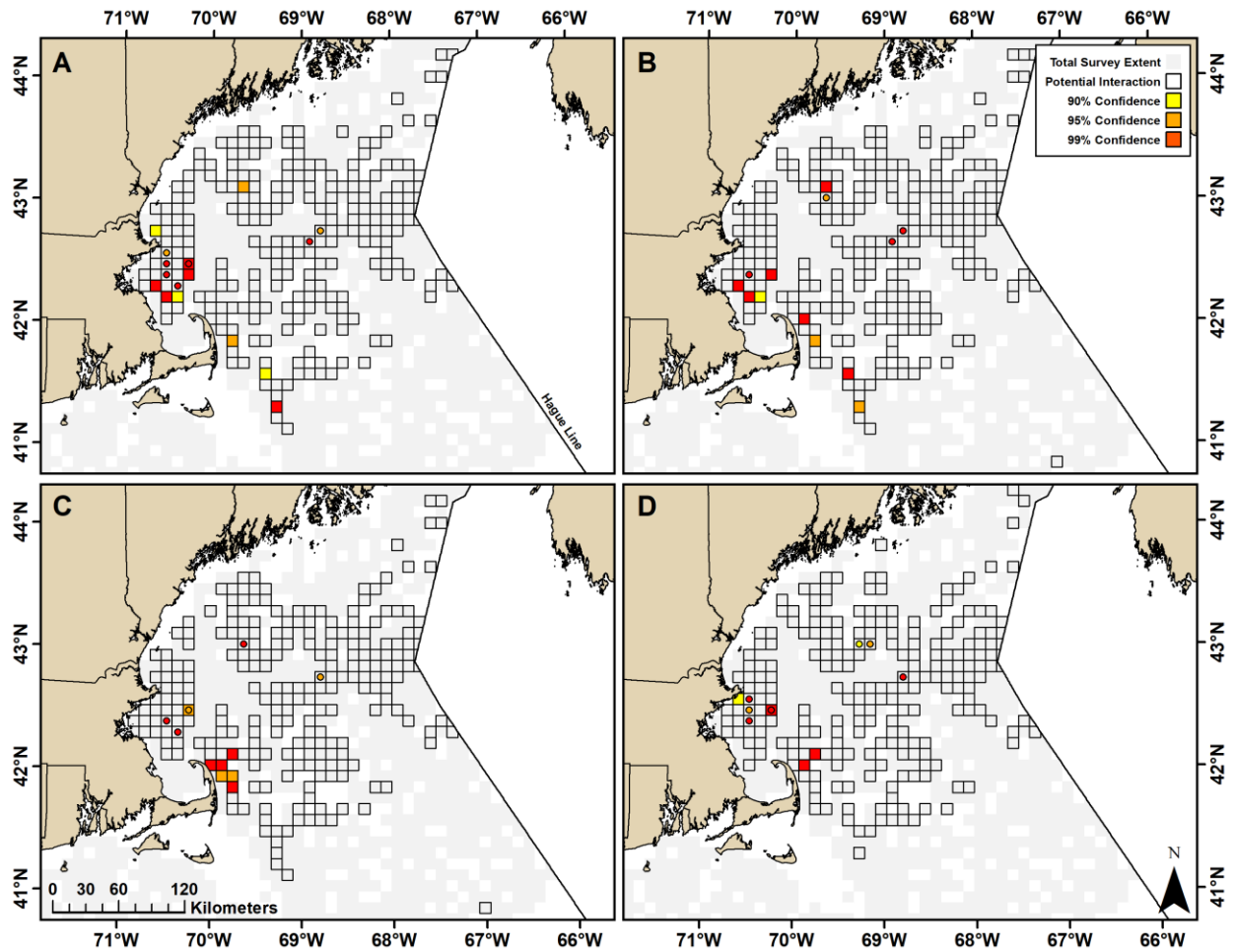


Figure 2.3. Thorny skate fishery-interaction potential hotspots in the Northeast multispecies fishing industry. Hotspots are provided in three-month windows (beginning at the start of the fishing year in May; A-D) and in 10 x 10 km resolution. Hotspots are shown at varying confidence levels (90, 95, and 99%) for the bottom trawl (in squares; i.e., grid cells) and bottom longline (circles within cells) surveys. In addition, grid cells where surveys occurred are provided (bottom trawl and/or bottom longline), along with cells where fishery-interaction potential existed when considering either bottom trawl and/or bottom longline survey data.

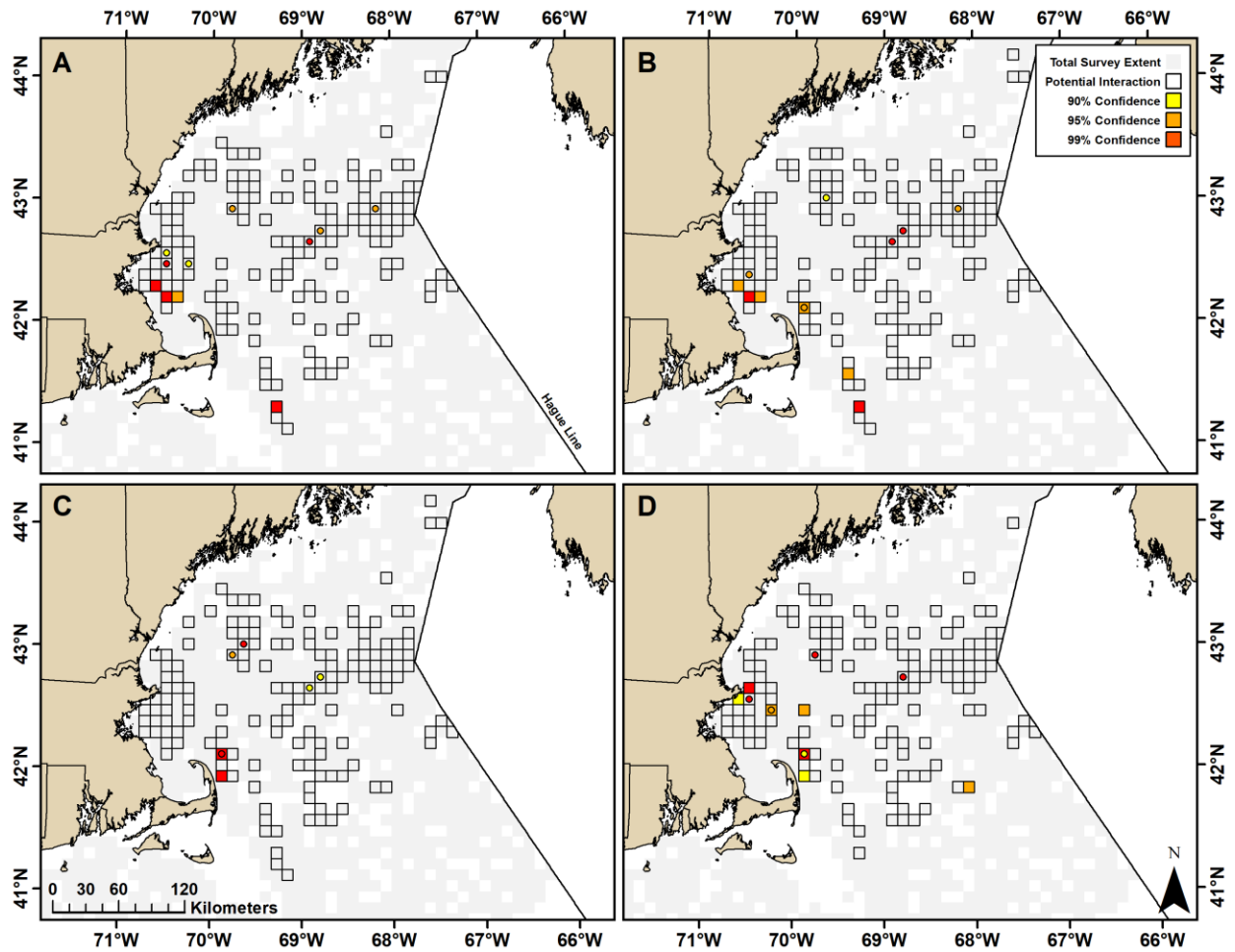


Figure 2.4. Thorny skate (51 to 70 cm total length) fishery-interaction potential hotspots in the Northeast multispecies fishing industry. Hotspots are provided three-month windows (beginning at the start of the fishing year in May; A-D) and in 10 x 10 km resolution. Hotspots are shown at varying confidence levels (90, 95, and 99%) for the bottom trawl (in squares; i.e., grid cells) and bottom longline (circles within cells) surveys. In addition, grid cells where surveys occurred are provided (bottom trawl and/or bottom longline), along with cells where fishery-interaction potential existed when considering either bottom trawl and/or bottom longline survey data.

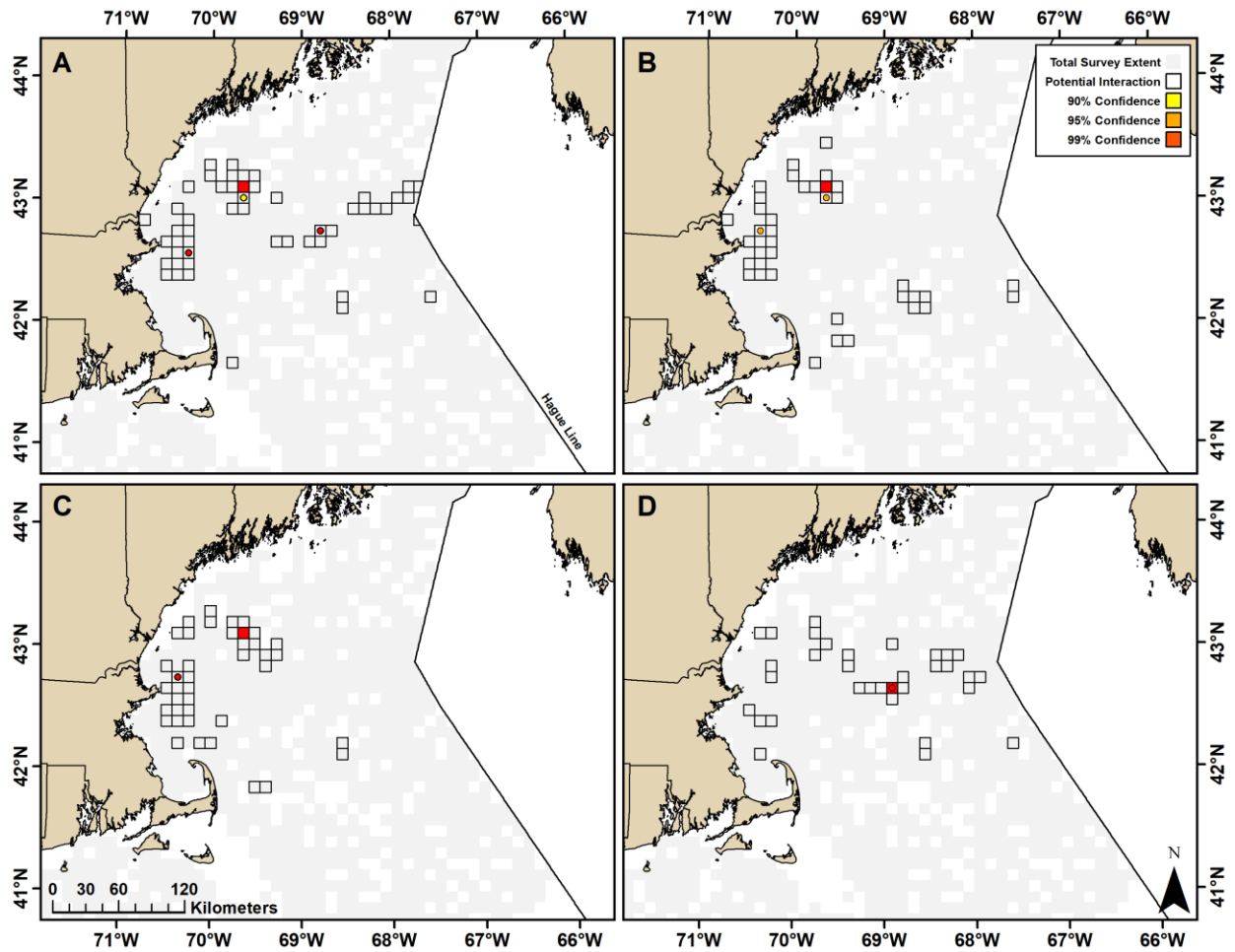


Figure 2.5. Thorny skate fishery-interaction potential hotspots in the monkfish fishing industry. Hotspots are provided in three-month windows (beginning at the start of the fishing year in May; A-D) and in 10 x 10 km resolution. Hotspots are shown at varying confidence levels (90, 95, and 99%) for the bottom trawl (in squares; i.e., grid cells) and bottom longline (circles within cells) surveys. In addition, grid cells where surveys occurred are provided (bottom trawl and/or bottom longline), along with cells where fishery-interaction potential existed when considering either bottom trawl and/or bottom longline survey data.

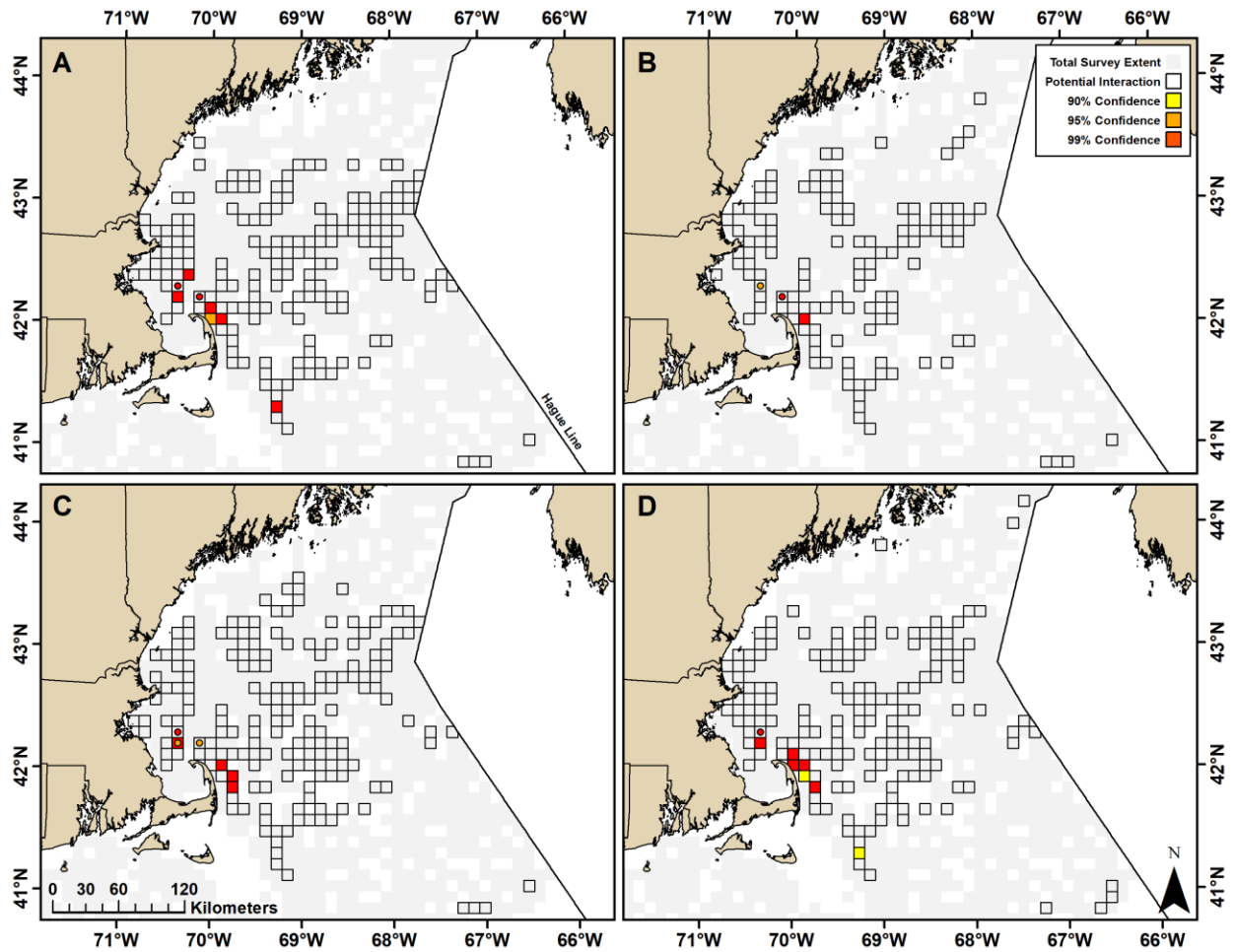


Figure 2.6. Thorny skate fishery-interaction potential hotspots in the Atlantic sea scallop fishing industry. Hotspots are provided in three-month windows (beginning at the start of the fishing year in March; A-D) and in 10 x 10 km resolution. Hotspots are shown at varying confidence levels (90, 95, and 99%) for the bottom trawl (in squares; i.e., grid cells) and bottom longline (circles within cells) surveys. In addition, grid cells where surveys occurred are provided (bottom trawl and/or bottom longline), along with cells where fishery-interaction potential existed when considering either bottom trawl and/or bottom longline survey data.

CHAPTER 3

MERGING TECHNOLOGIES AND SUPERVISED CLASSIFICATION METHODS TO QUANTIFY CAPTURE BEHAVIOR ON HOOK-AND-LINE

Abstract

Behavior in response to capture varies between species/fisheries and may be contributing to incidental mortality given the physiological ramifications tied its expression. However, the relationship between capture behavior, animal health, and discard mortality is poorly understood, and it remains a logistical challenge to address behavior over all conditions evident in a fishery. Here, we describe a technique that uses practical experimental gear to characterize capture behavior in hook-and-line fisheries and provides an avenue for predicting behaviors from acceleration data, which are often used to quantify animal movement and energy expenditure. The main components of the technique included (1) simultaneously monitoring the behavioral response to capture with accelerometers and cameras, (2) characterizing behavior from video footage and linking discrete behaviors to acceleration data, and (3) predicting behavior based solely on acceleration data using an ensemble of supervised classification methods. To develop and demonstrate our technique,

we collected observations from oceanic whitetip sharks, *Carcharhinus longimanus*, caught with experimental (hook-and-line) gear. The output of our technique provides not only a better understanding of the behavioral response to capture on hook-and-line fishing gear, but can be paired with physical or physiological assessments of animal health and/or fate to better understand the extent to which capture behavior influences mortality. Such information will assist in both the understanding of how species respond to capture on hook-and-line gear and in the formulation of species and fishery-specific strategies for mitigating mortality.

Introduction

The physiological and physical stressors incurred throughout capture can have lethal repercussions for incidentally caught and discarded fish, and population-level effects when discard mortality (DM) rates are high (Davis, 2002). DM has subsequently been a focal point of research (reviewed by Ellis et al., 2017; Veldhuizen et al., 2018), with estimated rates being used to inform stock assessments and management frameworks (e.g., establishing total allowable catch; Benaka et al., 2016). Studies have identified clear trends in DM (i.e., mortality connected to elevated stress and physical trauma; reviewed by Cook et al., 2019), but also that there is a great deal of variability between species and fisheries. High-variability in DM rates between species/fisheries has created challenges for managing individual stocks and stock complexes (e.g., shallow-water grouper or large coastal sharks; SAFMC, 1983; NMFS, 2006) that can require different management frameworks to control fishing mortality. This variability has also made it challenging to accurately approximate DM for data-poor species/fisheries (Benaka et al., 2014).

An animal's behavioral response to capture on fishing gear may lead to mortality if certain behaviors associated with the "fight or flight" response result in severe physiological disruptions tied to their expression (Skomal and Bernal, 2010). The relationship between behavior and an animal's welfare has been well studied for free-swimming fish (reviewed by Wood et al., 1991), but only a handful of studies have addressed this relationship with respect to capture, most of which having dealt with elasmobranchs. Studies have shown that behavioral expression is species-specific (Gallagher et al., 2017; Bouyoucos et al., 2018; Talwar et al., 2020) and that higher-intensity capture behaviors (e.g., burst-swimming) are connected to elevated physiological stress. In contrast, tempered behaviors (e.g., steady swimming) have been linked to a muted physiological response and/or recovery (Frick et al., 2009, 2010; Brooks et al., 2012; Gallagher et al., 2017; Bouyoucos et al., 2018). Based on these findings it appears that species-specific behaviors and their physiological consequences can be (in part) responsible for the variability in DM rates. Indeed, it has been speculated that certain species with notoriously pronounced fighting behaviors are subject to high DM rates as a result of these behaviors (Gallagher et al., 2014). However, the connection between capture behavior, physiology, and mortality has yet to be confirmed, partially due to an incomplete understanding of how individuals (or species) respond to capture on different fishing gears.

Capture behavior can be visually assessed by monitoring animals throughout a capture event either in-person or with cameras (reviewed by Graham et al., 2004), but visual monitoring can be challenging under certain conditions (e.g., low-visibility) and is not always feasible, particularly in large, high-seas fisheries (Guida et al., 2017). Instead, many

studies have turned to accelerometers that measure movement (tri-axial acceleration) throughout capture to gain insight to an animal's activity levels and energy expenditure (Wilson et al., 2020) in response to struggle on fishing gear. Accelerometers archive high-resolution data and are easier to deploy and operate under a wide-range of fishery conditions. However, they cannot identify the actual behavior that an animal is exhibiting at any point during the capture event. Discrete behaviors (e.g., burst swimming or lack of swimming) are often correlated with the physiological stress response, energy expenditure, and fate (Guida et al., 2016; Gallagher et al., 2014, 2017; Bouyoucos et al., 2018). Thus, a comprehensive understanding of capture behavior requires the characterization of specific behaviors and empirical data on animal physiology or movement kinematics associated with those behaviors. To characterize behavior from acceleration data it must be linked with visual observations of behavior (Nathan et al., 2012). Supervised classification methods (SCMs) can establish this linkage and also be used to predict behavior from movement alone (reviewed by Valetta et al., 2018).

Here we develop a technique to monitor capture behavior in hook-and-line fisheries (e.g., longlines, handlines, and rod-and-reel) and link acceleration data to visual observations of behavior. Using observations from oceanic whitetip sharks, *Carcharhinus longimanus*, caught with experimental hook-and-line gear, we applied an ensemble of SCMs to predict behavior from acceleration data, and demonstrate a mechanism for addressing behavior in more applied fishery settings where only accelerometers are needed. Experimental gear has been widely used as a proxy for characterizing behavior in response to capture in a given fishery, because it provides a semi-controlled platform for monitoring animals, and is often

more practical than working in concert with large-scale fishing operations (Guida et al., 2017). However, it does not obviate the need for fishery-scale observations of behavior where expression is expected vary to under the range authentic conditions and practices (reviewed by He 2006; Guida et al., 2017). Because behaviors are tied to a range of sublethal and lethal consequences, understanding behavioral expression on a fishery-scale will help in addressing the cause of species/fishery-specific DM rates.

Methods

Experimental fishing gear and shark capture

To observe the capture behavior of oceanic whitetip sharks on experimental hook-and-line gear, we captured sharks over a over a three year span off of Cat Island (2017 – 2019) and Mayaguana Island (2019), The Bahamas. Fishing was conducted with a modified handline constructed of two large polyballs (37 mm maximum diameter) that were attached to 8 – 9 m (6 mm diameter) of nylon line, which connected to four 1.2 m strands of 181.4 kg monofilament leader with an 18/0 non-offset carbon circle hook (Fig. 2.1). This gear served as a proxy for commercial pelagic longline gear because it similarly restricted the movement of sharks to a hemispherical area around a point that is relatively fixed (i.e., polyballs or mainline with handline or longline, respectively). To account for some of the range in capture duration expected in longline fisheries, we varied the amount of time sharks were left on-hook up to 60 minutes. Extended capture durations beyond 60 minutes could not be observed due to the limitations of concurrent research objectives.

Monitoring capture behavior

Capture behavior of oceanic whitetip sharks was monitored during 38 capture events using single accelerometers attached to the handline roughly 1 m from the hook. Accelerometers measured tri-axial acceleration (in g's) of the gear, which by extension registered the movement of sharks during capture (Gallagher et al., 2017; Bouyoucos et al., 2018; Talwar et al., 2020). HOBO Pendant G accelerometers were used from 2017 to 2018 (Onset Computer Corp., Bourne, MA, USA) and MDL-x16+p14 accelerometers in 2019 (Gulf Coast Data Concept, LLC, Waveland, MS, USA). Because MDLs sampled at a higher resolution and range (100 Hz and ± 16 g's), these data were down-sampled to match the configuration of Pendant Gs (1 Hz and ± 3 g's) using methods described by Hounslow et al. (2019). For a subsample of 10 capture events, GoFish cameras (GoFish Cam, Inc., Austin, Texas, USA) were attached to handlines to collect video at 1080p and 60 frames-per-second (FPS), or 720p and 30 FPS when observing events longer than one hour (i.e., due to battery consumption restrictions; Fig. 3.1). Initial exploratory trials revealed that a single camera did not provide enough viewing angle to keep the shark within frame throughout capture, regardless of where the camera was positioned on the line (Supplemental Video S1). Two individual cameras were therefore used and placed at ~ 3 and 6 m from the hook, with both placements determined as being the best combination of viewing angles to ensure the shark remained in frame during capture (Supplemental Video S1). This sampling design resulted in 10 capture events that were simultaneously-monitored by line-borne accelerometers and video cameras, and 28 events that were solely monitored by accelerometers.

Data analysis

To establish a link between acceleration data and videoed observations of behavior, an ensemble of SCMs (referred to collectively as the “ensemble”) was optimized to predict behaviors according to the spectral characteristics of acceleration data. Our analytical approach progressed in three phases: (I) an ethographic analysis to characterize behavior in the subsample of individuals monitored with accelerometers and cameras, (II) feature construction from raw acceleration data for all animals, and (III) the prediction of capture behavior in animals solely monitored with acceleration data based on features identified in (II) (Fig. 3.2). All analyses were performed using R 3.6.2 (R Core Team 2019) unless otherwise stated.

Phase I: Ethographic analysis

GoFish camera video footage of the 10 dually-monitored capture events was visually-reviewed to characterize discrete behaviors with a semi-qualitative approach. A total of 6.4 hours of footage was reviewed, representing capture events spanning from 7.2 to 36.6 min (mean \pm standard deviation: 19.1 ± 9.0 minutes). Four behaviors were identified including steady swimming, thrashing, burst-swimming, and motionless behavior. The most frequently observed behavior was sustained, “steady swimming”, which resembled kinematics of free-swimming oceanic whitetip sharks and included occasional gliding and no visual indications of struggling (Fig. 3.3b; Supplemental Video S2). Brief, erratic thrashing and burst-swimming were the next most common behaviors displayed. Thrashing was characterized by sequences of head-shaking/body contortions and burst-swimming by an increase in tailbeat

frequency and amplitude (Supplemental Video S3). Both behaviors were also associated with several instances of regurgitation (n = 15 events from five individuals; Supplemental Video S4). *A priori* visual comparisons of acceleration data between instances of thrashing and burst swimming revealed that our relatively coarse sampling resolution (1 Hz) was unable to discern acceleration signatures between the two. Thus, both behaviors were reclassified as “high-energy” response to capture (Fig. 3.3d; Talwar et al., 2020). The most infrequent capture behavior was hanging from the hook vertically in the water column with no apparent signs of movement, which was only documented in one individual (referred to herein as “motionless behavior”; Fig. 3.3f; Supplemental Video S3).

The expressions of behaviors during a capture event were tracked using the Behavioral Observation Research Interactive Software (BORIS Version 7.9.7; Friard and Gamba, 2016; Andrzejaczek et al., 2018, 2019), which creates an output time-series file with behavioral observations. Observed behavior was then matched to the corresponding acceleration data using timestamps, which will be referred to herein as “labeled” data. “Unlabeled” data will refer to capture events that were monitored solely with accelerometers and are described further in subsequent sections.

Phase II: Acceleration features

Features were derived from raw-acceleration data to further describe movement in a variety of ways (e.g., instantaneous or over-time), adding more context to behavior-specific acceleration signals, and improving behavioral recognition (Kleanthous et al., 2019). Frequentist features (n = 102) were calculated including static and dynamic acceleration

(Yoda et al., 1999; Shepard et al., 2008), vectorial and overall dynamic body acceleration (Gleiss et al., 2011; Qasem et al., 2012), and time-domain statistics (e.g., mean, median, standard deviation, skewness, etc.) using a three-second moving window (Appendix Table B.1). The ReliefF algorithm (Kononenko, 1994) feature selection method in the “CORElearn” package (Robnik-Sikonja and Savicky, 2020) was then used to identify irrelevant and redundant features that lead to overfitting and reduced classification performance, if included (Tang et al., 2014). Irrelevant features ($n = 13$) were removed from our labeled and unlabeled datasets, leaving 89 features of acceleration by which to describe oceanic whitetip capture behavior.

Phase III: Behavior classification and prediction

To predict capture behavior, our ensemble was trained using labeled data with acceleration features. SCMs used in the ensemble have been frequently used predict animal behaviors and included multinomial logistic regressions (MLR), K-nearest neighbors (KNN), support vector machines (SVM), multilayer perceptron neural networks (MLP), random forests (RF), and extreme gradient boosted decision trees (XGBoost; Nathan et al., 2012; Rescheff et al., 2014; Ladds et al., 2016, 2017; Brewster et al., 2018; Benaissa et al., 2019; Hounslow et al., 2019; additional details provided in Appendix B). Relevant hyperparameters for each SCM were optimized during training to maximize prediction performance to the structure of the data (Appendix Table B.2 and Fig. B.1). Each SCM was used to independently predict the probability of an observation belonging to each of the three behaviors (i.e., steady swimming, high-energy, and motionless behaviors). This array of

behavioral probabilities was subject to an ensembling approach that calculated the weighted-average of each behavior's probabilities for a given observation, with the highest probability assigned as the final behavior prediction. Ensembling was incorporated because it is known to improve overall classification performance relative to any of the constituent SCMs (Schapire et al., 1998; Catal et al., 2015; Brewster et al., 2018). SCM behavioral prediction weights were determined using a brute-force grid search that varied the weights of constituent SCMs in 0.05 increments (summing to one between all SCMs; $n = 49,854$ combinations). After the ensemble was optimized and informed by labeled data and acceleration features, it was used to make behavioral predictions for unlabeled data.

Behavior predictions made for unlabeled data were also subjected to a series of additional criteria designed *a posteriori* to address instances where our ensemble was not able to predict behavior with a high degree of certainty. Most notably, additional noise observed in acceleration data (e.g., rougher sea conditions) led to several bouts of steady swimming that contained individual observations of high-energy behavior that were only marginally more probable. To account for noise potentially conflating high-energy predictions, we reclassified any observation of this behavior if the ensemble did not provide a definitive prediction (i.e., 90% or higher probability). In addition, because accelerometers sampled at a relatively coarse sampling frequency (1 Hz) it was possible that the acceleration signature of more rapid high-energy responses (e.g., thrashing) went undetected. To account for this potential shortcoming, any single observation between bouts of high-energy behavior was reclassified as a continuous bout. Lastly, sequences of motionless behaviors were revised as steady swimming if the bouts were shorter than the minimum duration observed in the

labeled data (4 seconds), which serves as an additional buffer against any mischaracterizations of movement signatures due to sampling frequencies.

SCM performance evaluation

Prediction performance of behaviors by the ensemble was evaluated using a stratified 5-fold cross-validation approach that partitioned labeled data into 80% training and 20% testing within each fold, while retaining the proportion of observations across behaviors. Steady-swimming observations (92.2% of total observations) were then down-sampled by 10-fold to avoid biasing predictions towards the most frequent behavior (Japkowicz and Stephan, 2002). In each fold, SCMs were supplied with training data and used to predict behavior in testing data, construct confusion matrices (i.e., predicted versus known behaviors), and evaluate prediction performance. True positive (TP) and true negative (TN) observations represented instances where behaviors were predicted correctly, whereas false positives (FP) and false negatives (FN) reflected incorrect predictions. The following performance metrics were then calculated:

$$\mathbf{Recall} = \frac{TP}{(TP+FN)} \quad \text{And} \quad \mathbf{Precision} = \frac{TP}{(TP+FP)} \quad (1) \text{ and } (2)$$

Recall measures the proportion of observations for a given capture behavior that were correctly identified, whereas precision reflects the proportion of predictions for a given

capture behavior that were relevant (i.e., correct). Both metrics were calculated for each capture behavior and incorporated into these additional $F1$ -measure metrics:

$$\mathbf{F1} = \frac{2*Precision*Recall}{(Precision+Recall)} \quad \text{And} \quad \mathbf{macro - F1} = \frac{\sum F_1}{M} \quad (3) \text{ and } (4)$$

$F1$ represents the harmonic mean of precision and recall and was calculated for each capture behavior, while $macro - F1$ was the average $F1$ measure across all capture behaviors (M) when each behavior was weighted equally. As $F1$ measures $\rightarrow 1$ the prediction performance based on a balance between precision and recall improved.

Results

Hyperparameter optimization was successful and resulted in marginal (1.7 – 4.2%; KNN, MLP, RF, and XGBoost) to major (256.9%; SVM) improvements in the overall predictive performance (0.0138 – 0.6170 increase in $macro - F1$) of SCMs (Appendix Fig. B.1). SCMs predicted the three capture behaviors with varying degrees of success (Table 3.2). Collectively, steady swimming was the most reliably-predicted capture behavior ($F1_{swim} = 0.906 \pm 0.029$) followed by high-energy ($F1_{high-energy} = 0.863 \pm 0.028$), while motionless behavior was consistently predicted with the least accuracy ($F1_{motionless} = 0.809 \pm 0.180$; Table 3.2). The overall predictive performance of SCMs varied when taking into account all capture behaviors ($macro - F1 = 0.877 \pm 0.040$), with MLR identified as the

poorest performing and XGBoost the best ($macro - F1 = 0.819$ and 0.926 , respectively; Table 3.2; Fig. 3.4).

The optimal ensemble omitted behavior predictions (i.e., assigned a weight of zero) for MLR, while XGBoost contributed the majority of predictive information (70%) and KNN the second most information (19%; Table 3.1). The remaining SCMs individually contributed to no more than 8% of the final prediction probabilities (8%). The ensemble achieved the highest performance ($macro - F1 = 0.927$; Table 3.2) and was marginally better (increase of 0.17% in $macro - F1$) than the best performing SCM (XGBoost; Table 3.2). This improvement was mostly due to prediction performance of the motionless behavior ($F1_{motionless} = 0.951$), though performance was boosted in the other behaviors as well (Table 3.2). Individual behaviors were correctly identified by the ensemble in $94.2 \pm 0.3\%$ of steady swimming observations, $88.2 \pm 0.3\%$ of high-energy observations, and $95.2 \pm 0.6\%$ of observations for the motionless behavior.

Discussion

Behavior in response to capture may have repercussions on an animal's health and fate, but this connection is poorly understood and remains difficult to address in applied settings. Our technique provides researchers a more practical approach to collecting basic behavioral information without the logistical challenges of monitoring capture in a fishery. Behavioral information collected by our technique can be paired with health/fate assessments to better understand the mechanistic relationship between behavior and mortality (Guida et al., 2016; Gallagher et al., 2017a; Bouyoucos et al., 2018), which up until now has been

largely speculative. If the link between behavior and discard mortality becomes more clearly defined, behavioral expressions throughout a fishery will help to explain the variability being reported in mortality. Behavior on a fishery-scale can be addressed using our ensemble that is capable of predicting behavior from acceleration data, which can be more feasibly collected across the range of fishing conditions and practices of a fishery. Having a more comprehensive understanding of what is driving variability in DM will better inform managers when formulating species and fishery-specific strategies for mitigating mortality.

Technique performance and considerations

The ensemble used by our technique identified discrete behaviors exhibited during a capture event with a level high level of accuracy and balance between recall and precision (Table 3.2; Appendix Fig. B.2). In addition to outperforming any of the SCMs (Table 3.2), the level of performance achieved by our ensemble matched or exceeded what has been reported in the larger body of free-ranging animal behavior studies (McClune et al., 2014; Dutta et al., 2015; Hammond et al., 2016; Ladds et al., 2017; Barker et al., 2018; Benaissa et al., 2019; Studd et al., 2019; Jeantet et al., 2018, 2020). However, because performance is also subject to inherent differences in data structuring, comparing our ensemble to other animal behavior studies should only be treated as context to what level of performance is considered acceptable in this field. That said, ensemble frameworks can account for various data structures (i.e., acceleration signals from different species/fisheries) while maintaining a high level of performance (Dutta et al., 2015; Brewster et al., 2018). Because SCMs operated independently in our ensemble when evaluating acceleration data, performance can be

optimized on a case-by-case basis by weighting constituents differently to achieve the best-overall performance, relative to the structure of acceleration data (Dutta et al., 2015; Brewster et al., 2018). For instance, while XGBoost and KNN contributed the majority of information to oceanic whitetip shark behavior predictions (Table 3.1), other studies will require different combinations of constituent SCMs that our ensemble can easily adapt to.

The sampling frequency used to record acceleration will dictate your ability to identify discrete behaviors (Hounslow et al., 2019). Relatively coarse acceleration data (1 Hz) led to signal aliasing of higher kinematic burst-swimming and thrashing of oceanic whitetip sharks that would have hindered the ability of our ensemble to confidently predict each behavior. But by merging both behaviors we were able to maintain a high level of prediction performance and still address these two behaviors (grouped together) that are expected to have comparable physiological implications tied to higher-energy expenditure (Skomal and Bernal, 2010; Wilson et al., 2020). This simpler approach of lower resolution acceleration data and grouping like behaviors, with similar impacts on animal welfare, will be a useful avenue for addressing capture behavior moving forward. For example, in addition to bolstering prediction performance, lower resolution data will extend the monitoring period of accelerometers (i.e., memory space not as quickly exhausted) to better match authentic fishing practices needed to address behavior on a fishery-scale. However, if your goal is to identify all discrete behaviors, consider sampling at a frequency that is twice the resolution of the most rapid behavior you expect to observe (i.e., the Nyquist criterion; Weik, 2001). Hounslow et al. (2019) also recommended a 5 Hz acceleration threshold for behavioral classification of free-swimming sharks (based on juvenile lemon sharks, *Negaprion*

brevirostris) and several other studies have used 20 Hz or more (Nakamura et al., 2011; Lear et al., 2017, 2019; Whitney et al., 2017; Brewster et al., 2018; Karan et al., 2019; Coffey et al., 2020). But ultimately sampling frequency should be tailored to specific study objectives on a case-by-case basis, taking into account the level of behavioral detail required, while also keeping in mind that higher resolutions will limit monitoring durations.

Essential to our technique and the construction of SCMs was the collection of behavioral information from experimental gear that served as a proxy for capture in a fishery. In this study, experimental hook-and-line gear was considered an acceptable proxy for pelagic longlines based on similar behaviors observed between oceanic whitetip sharks and other congeneric species caught with modified longlines (Bouyoucos et al., 2018; Talwar et al., 2020). By using experimental gear, monitoring capture events with cameras and accelerometers becomes more feasible (relative to an applied setting) and makes it possible to characterize discrete behaviors and establish the definitive link between behavior and acceleration for a focal species. This linkage informs our ensemble to predict behavior based on acceleration data, which can be more easily collected on a larger, fishery-scale, across a wide-range of conditions and practices that are not addressed with experimental gear. For instance, accelerometers are more equipped to monitor pelagic longlines (relative to cameras) because of having greater depth-ratings and longer battery-lives (days-weeks) that are better suited to fishing operations (e.g., Axy5depth; TechnoSmArt, Rome, Italy). In addition, experimental gear in a semi-controlled setting can be matched with post-capture physiological assessments and/or post-release monitoring (e.g., pop-up satellite archival transmitting tags) to examine the relationship between behavior and animal welfare (Guida et

al., 2016; Gallagher et al., 2017; Bouyoucos et al., 2018). Having a firmer understanding of the relationship between behavior and animal welfare is needed to interpret the results of behavioral expression on a fishery-scale and determine whether or not behaviors are contributing to variable DM rates.

Lastly, performance of our technique hinges heavily on the ability of cameras to collect visual observations of behavior. While cameras perform well in clear, well-lit waters, inhabited by oceanic whitetip shark, they can become less effective in low-visibility conditions associated with turbid environments or fishing at night and/or depth (Guida et al., 2017). Battery life can also restrict the length of the capture event you are able to film and therefore limit the amount of behavioral information you are able to collect. Pressure at depth can also damage cameras (i.e., specified as a depth rating) and vertically confine monitoring to portions of the water column above where fishing occurs. For GoFish cameras, operating at the highest resolution (i.e., 1080p and 60 FPS) constrained capture events to less than an hour, but by reducing video quality to 720p and 30 FPS we were able to extend monitoring periods while still maintaining enough resolution to observe behaviors clearly. Other cameras may better address the monitoring challenges mentioned above (O'Shea et al., 2015; Bouyoucos et al., 2018; Talwar et al., 2020), but we opted for GoFish cameras because of their unique design specific to in-line attachment (Fig. 3.1) and ability to capture quality footage at an affordable price (\$150 USD). Regardless of the cameras being used, it is important to optimize their placement on the gear to ensure the animal's movements remain in frame throughout capture (Supplemental Video S1)

Future applications and conclusion

If the relationship between behavior and animal welfare becomes more clearly defined, our technique can be used to help (in part) identify or address the efficacy of certain best-practices and gear-modifications. For example, reducing soak times has often been touted as an approach for mitigating mortality (Diaz and Serafy, 2005; Campana et al., 2009; Morgan and Carlson, 2010; Gallagher et al., 2014). However, this best-practice may be less effective if the focal species completely exhausts itself shortly after being hooked and acute physiological stress leads to DM (Gallagher et al., 2014). Our technique can provide behavioral information to confirm this type of response to capture and can be used to interpret the efficacy of the practice, if behaviors are also linked to animal health. In addition, our technique can be used to address the efficacy of extending gangions to increase motility, by quantifying the expression of steady swimming behavior at various gangion lengths. If paired with confirmation that steady swimming promotes survivability, we could identify an appropriate gangion length as an effective mitigation strategy. (Gulak et al., 2015; Musyl and Gilman, 2019)

In addition, the application of our technique to other teleost (e.g., Atlantic cod, *Gadus morhua*, Pacific halibut, *Hippoglossus stenolepis*; swordfish, *Xiphias gladius*; and several *Thunnus* spp.) and elasmobranch (e.g., blacknose shark, *C. acronotus*; blue shark, *Prionace glauca*; porbeagle shark, *Lamna nasus*; sandbar shark, *C. plumbeus*; shortfin mako shark; *Isurus oxyrinchus*; silky shark, *C. falciformis*) species that commonly interact with hook-and-line fisheries is warranted, especially for those with highly variable discard mortality rates that remain unexplained. Handlines were used as a proxy for pelagic longline gear; however,

our technique could easily be adapted to other hook-and-line fisheries, such as such as bottom longlines (e.g., using drumlines as a proxy; Gallagher et al., 2014, 2017a) or rod-and-reel. It may be possible to use a similar monitoring approach with gillnet gear as well, using appropriately placed accelerometers and cameras (e.g., attached to floatlines). But additional validation studies would be required to confirm this approach.

Acknowledgements

I would like to thank all of the co-authors involved with a forthcoming manuscript (E. Brooks, L. Howey, J. Gelsleichter, M. Bond, B. Talwar, S. Williams, O. Shipley, J. Mandelman, and J. Kneebone). I would also like to thank D. Grubbs, L. Jordan, M. Winchester, D. Abercrombie, A. Brooks, M. Henriksen, S. LaMay, D. Brake, D. Kasana, K. Palmrose, S. Pinder, and J. Aquino for their help with data collection. I would also like to acknowledge the Capt. S. Genereux, J. Salamone, and the crew of the M/V Alucia for helping to make this research possible. Support for this project was provided by Moore Charitable Foundation, Bloomberg Philanthropies, Discovery Channel, and OceanX. Lastly, thank you the Oceanic Whitetip Consortium for their continued research efforts with oceanic whitetip sharks.

Literature cited

Alvarenga, F. A. P., Borges, I., Palkovič, L., Rodina, J., Oddy, V. H., & Dobos, R. C. (2016). Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. *Applied Animal Behaviour Science*, *181*, 91-99.

Andrzejaczek, S., Gleiss, A. C., Pattiaratchi, C. B., & Meekan, M. G. (2018). First insights into the fine-scale movements of the sandbar shark, *Carcharhinus plumbeus*. *Frontiers in Marine Science*, *5*, 483.

Andrzejaczek, S., Gleiss, A. C., Lear, K. O., Pattiaratchi, C. B., Chapple, T. K., & Meekan, M. G. (2019). Biologging tags reveal links between fine-scale horizontal and vertical movement behaviors in tiger sharks (*Galeocerdo cuvier*). *Frontiers in Marine Science*, *6*, 229.

Barker, Z. E., Diosdado, J. V., Codling, E. A., Bell, N. J., Hodges, H. R., Croft, D. P., & Amory, J. R. (2018). Use of novel sensors combining local positioning and acceleration to measure feeding behavior differences associated with lameness in dairy cattle. *Journal of Dairy Science*, *101*(7), 6310-6321.

Benaissa, S., Tuytens, F. A., Plets, D., De Pessemier, T., Trogh, J., Tanghe, E., Martens, L., Vandaele, L., Van Nuffel, A., Joseph, W., & Sonck, B. (2019). On the use of on-cow accelerometers for the classification of behaviours in dairy barns. *Research in Veterinary Science*, *125*, 425-433.

Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In *Neural networks: Tricks of the Trade* (pp. 437-478). Springer, Berlin, Heidelberg.

Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., Gao, L., Quintana, F., & Wilson, R. P. (2014). Love thy neighbour: automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PloS One*, *9*(2), e88609.

Bom, R. A. (2010). *Can speed and tri-axial acceleration measured by biologgers be used to classify oystercatcher behaviour* (Doctoral dissertation, Master's thesis, University of Amsterdam, The Netherlands).

Bom, R. A., Bouten, W., Piersma, T., Oosterbeek, K., & van Gils, J. A. (2014). Optimizing acceleration-based ethograms: the use of variable-time versus fixed-time segmentation. *Movement Ecology*, *2*(1), 6.

Bouyoucos, I. A., Suski, C. D., Mandelman, J. W., & Brooks, E. J. (2017). The energetic, physiological, and behavioral response of lemon sharks (*Negaprion brevirostris*) to simulated longline capture. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 207, 65-72.

Bouyoucos, I. A., Talwar, B. S., Brooks, E. J., Brownscombe, J. W., Cooke, S. J., Suski, C. D., & Mandelman, J. W. (2018). Exercise intensity while hooked is associated with physiological status of longline-captured sharks. *Conservation Physiology*, 6(1), coy074.

Bradley, P. S., & Mangasarian, O. L. (1998, July). Feature selection via concave minimization and support vector machines. In *ICML* (Vol. 98, pp. 82-90).

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Brewster, L. R., Dale, J. J., Guttridge, T. L., Gruber, S. H., Hansell, A. C., Elliott, M., Cowx, I. G., Whitney, N. M., & Gleiss, A. C. (2018). Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data. *Marine Biology*, 165(4), 62.

Brooks, E. J., Mandelman, J. W., Sloman, K. A., Liss, S., Danylchuk, A. J., Cooke, S. J., Skomal, G. B., Philipp, D. P., Sims, D. W., & Suski, C. D. (2012). The physiological response of the Caribbean reef shark (*Carcharhinus perezii*) to longline capture. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 162(2), 94-100.

Brownscombe, J. W., Marchand, K., Tisshaw, K., Fewster, V., Groff, O., Pichette, M., Seed, M., Gutowski, L. F., Wilson, A. D., & Cooke, S. J. (2014). The influence of water temperature and accelerometer-determined fight intensity on physiological stress and reflex impairment of angled largemouth bass. *Conservation Physiology*, 2(1).

Bridle, J. S. (1990). Probabilistic interpretation of feed forward classification network outputs, with relationships to statistical pattern recognition. In *Neurocomputing* (pp. 227-236). Springer, Berlin, Heidelberg.

Butcher, P. A., Peddemors, V. M., Mandelman, J. W., McGrath, S. P., & Cullis, B. R. (2015). At-vessel mortality and blood biochemical status of elasmobranchs caught in an Australian commercial longline fishery. *Global Ecology and Conservation*, 3, 878-889.

Campana, S. E., Marks, L., Joyce, W., & Kohler, N. E. (2006). Effects of recreational and commercial fishing on blue sharks (*Prionace glauca*) in Atlantic Canada, with inferences on the North Atlantic population. *Canadian Journal of Fisheries and Aquatic Sciences*, 63(3), 670-682.

- Campana S.E., W. Joyce, M.J. Manning. 2009. Bycatch and discard mortality on commercially caught blue sharks *Prionace glauca* assessed using archival satellite pop-up tags. *Marine Ecology Progress Series*. 387, 241-253.
- Campbell, H. A., Gao, L., Bidder, O. R., Hunter, J., & Franklin, C. E. (2013). Creating a behavioural classification module for acceleration data: using a captive surrogate for difficult to observe species. *Journal of Experimental Biology*, 216(24), 4501-4506.
- Carroll, G., Slip, D., Jonsen, I., & Harcourt, R. (2014). Supervised accelerometry analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology*, 217(24), 4295-4302.
- Carruthers, E. H., Neilson, J. D., & Smith, S. C. (2011). Overlooked bycatch mitigation opportunities in pelagic longline fisheries: Soak time and temperature effects on swordfish (*Xiphias gladius*) and blue shark (*Prionace glauca*) catch. *Fisheries Research*, 108(1), 112-120.
- Catal, C., Tufekci, S., Pirmit, E., & Kocabag, G. (2015). On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing*, 37, 1018-1022.
- Cerezuela-Escudero, E., Rios-Navarro, A., Dominguez-Morales, J. P., Tapiador-Morales, R., Gutierrez-Galan, D., Martín-Cañal, C., & Linares-Barranco, A. (2016). Performance evaluation of neural networks for animal behaviors classification: Horse gaits case study. In *Distributed Computing and Artificial Intelligence, 13th International Conference* (pp. 377-385). Springer, Cham.
- Chang, C. C., & Lin, C. J. (2016). *A practical guide to support vector classification*. Department of Computer Science, National Taiwan University, Taipei, 106.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Coffey, D. M., Royer, M. A., Meyer, C. G., & Holland, K. N. (2020). Diel patterns in swimming behavior of a vertically migrating deepwater shark, the bluntnose sixgill (*Hexanchus griseus*). *PloS one*, 15(1), e0228253.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

- Dapp, D. R., Walker, T. I., Huveneers, C., & Reina, R. D. (2016). Respiratory mode and gear type are important determinants of elasmobranch immediate and post-release mortality. *Fish and Fisheries*, 17(2), 507-524.
- Dasarathy, B. V. (1991). Nearest neighbor (NN) norms: NN pattern classification techniques. *IEEE Computer Society Tutorial*,
- Davis, M. W. (2002). Key principles for understanding fish bycatch discard mortality. *Canadian Journal of Fisheries and Aquatic Sciences*, 59(11), 1834-1843.
- Diaz, G. A., & Serafy, J. E. (2005). Longline-caught blue shark (*Prionace glauca*): factors affecting the numbers available for live release. *Fishery Bulletin*, 103(4), 720.
- Dulvy, N. K., Fowler, S. L., Musick, J. A., Cavanagh, R. D., Kyne, P. M., Harrison, L. R., & Pollock, C. M. (2014). Extinction risk and conservation of the world's sharks and rays. *Elife*, 3, e00590.
- Dulvy, N. K., Simpfendorfer, C. A., Davidson, L. N., Fordham, S. V., Bräutigam, A., Sant, G., & Welch, D. J. (2017). Challenges and priorities in shark and ray conservation. *Current Biology*, 27(11), R565-R572.
- Dutta, R., Smith, D., Rawnsley, R., Bishop-Hurley, G., Hills, J., Timms, G., & Henry, D. (2015). Dynamic cattle behavioural classification using supervised ensemble classifiers. *Computers and Electronics in Agriculture*, 111, 18-28.
- Ellis, J. R., McCully Phillips, S. R., & Poisson, F. (2017). A review of capture and post-release mortality of elasmobranchs. *Journal of Fish Biology*, 90(3), 653-722.
- Escalante, H. J., Rodriguez, S. V., Cordero, J., Kristensen, A. R., & Cornou, C. (2013). Sow-activity classification from acceleration patterns: a machine learning approach. *Computers and Electronics in Agriculture*, 93, 17-26.
- Fehlmann, G., O'Riain, M. J., Hopkins, P. W., O'Sullivan, J., Holton, M. D., Shepard, E. L., & King, A. J. (2017). Identification of behaviours from accelerometer data in a wild social primate. *Animal Biotelemetry*, 5(1), 6.
- Friard, O., & Gamba, M. (2016). BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11), 1325-1330.
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232.

- Gallagher, A. J., Staaterman, E. R., Cooke, S. J., & Hammerschlag, N. (2016). Behavioural responses to fisheries capture among sharks caught using experimental fishery gear. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(1), 1-7.
- Gerencsér, L., Vásárhelyi, G., Nagy, M., Vicsek, T., & Miklósi, A. (2013). Identification of behaviour in freely moving dogs (*Canis familiaris*) using inertial sensors. *PloS One*, 8(10).
- Gleiss, A. C., Wilson, R. P., & Shepard, E. L. (2011). Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, 2(1), 23-33.
- Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 315-323).
- Graf, P. M., Wilson, R. P., Qasem, L., Hackländer, K., & Rosell, F. (2015). The use of acceleration to code for animal behaviours; a case study in free-ranging Eurasian beavers *Castor fiber*. *PloS One*, 10(8).
- Guida, L., Walker, T. I., & Reina, R. D. (2016). Temperature insensitivity and behavioural reduction of the physiological stress response to longline capture by the gummy shark, *Mustelus antarcticus*. *PLoS One*, 11(2), e0148829.
- Guida, L., Dapp, D. R., Huveneers, C. P., Walker, T. I., & Reina, R. D. (2017). Evaluating time-depth recorders as a tool to measure the behaviour of sharks captured on longlines. *Journal of Experimental Marine Biology and Ecology*, 497, 120-126.
- Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, 219(11), 1618-1624.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.
- Hoang, Q. T., Phi Khanh, P. C., Trung Ninh, B., Phuong Dung, C. T., & Tran, T. (2018). Cow Behavior Monitoring Using a Multidimensional Acceleration Sensor and Multiclass SVM. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(03), 110-118.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed forward networks are universal approximators. *Neural Networks*, 2(5), 359-366.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Hounslow, J. L., Brewster, L. R., Lear, K. O., Guttridge, T. L., Daly, R., Whitney, N. M., & Gleiss, A. C. (2019). Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*, 512, 22-30.

Howey-Jordan, L. A., Brooks, E. J., Abercrombie, D. L., Jordan, L. K., Brooks, A., Williams, S., Gospodarczyk, E., & Chapman, D. D. (2013). Complex movements, philopatry and expanded depth range of a severely threatened pelagic shark, the oceanic whitetip (*Carcharhinus longimanus*) in the western North Atlantic. *PloS One*, 8(2), e56588.

Howey, L. A., Tolentino, E. R., Papastamatiou, Y. P., Brooks, E. J., Abercrombie, D. L., Watanabe, Y. Y., Williams, S., Brooks, A., Chapman, D. D., & Jordan, L. K. (2016). Into the deep: the functionality of mesopelagic excursions by an oceanic apex predator. *Ecology and Evolution*, 6(15), 5290-5304.

Japkowicz, N. (2000, July). Learning from imbalanced data sets: a comparison of various strategies. In *AAAI workshop on learning from imbalanced data sets* (Vol. 68, pp. 10-15).

Japkowicz, N., & Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429-449.

Jeantet, L., Dell'Amico, F., Forin-Wiart, M. A., Coutant, M., Bonola, M., Etienne, D., Gresser, J., Regis, S., Lecerf, N., Lefebvre, F., & de Thoisy, B. (2018). Combined use of two supervised learning algorithms to model sea turtle behaviours from tri-axial acceleration data. *Journal of Experimental Biology*, 221(10).

Jeantet, L., Planas-Bielsa, V., Benhamou, S., Geiger, S., Martin, J., Siegwalt, F., Lelong, P., Gresser, J., Etienne, D., Hiélard, G., & Arque, A. (2020). Behavioural inference from signal processing using animal-borne multi-sensor loggers: a novel solution to extend the knowledge of sea turtle ecology. *Royal Society Open Science*, 7(5), 200139.

Karan, S., Meese, E. N., Yang, Y., Yeh, H. G., Lowe, C. G., & Zhang, W. (2019, November). Classification of Shark Behaviors using K-Nearest Neighbors. In *2019 IEEE Green Energy and Smart Systems Conference (IGESSC)* (pp. 1-6). IEEE.

Kleanthous, N., Hussain, A., Mason, A., & Sneddon, J. (2019, August). Data Science Approaches for the Analysis of Animal Behaviours. In *International Conference on Intelligent Computing* (pp. 411-422). Springer, Cham.

- Kneebone, J., Chisholm, J., Bernal, D., & Skomal, G. (2013). The physiological effects of capture stress, recovery, and post-release survivorship of juvenile sand tigers (*Carcharias taurus*) caught on rod and reel. *Fisheries Research*, *147*, 103-114.
- Kononenko, I. (1994, April). Estimating attributes: analysis and extensions of RELIEF. In *European Conference on Machine Learning* (pp. 171-182). Springer, Berlin, Heidelberg.
- Ladds, M. A., Thompson, A. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2016). Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PLoS One*, *11*(12).
- Ladds, M. A., Thompson, A. P., Kadar, J. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2017). Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry. *Animal Biotelemetry*, *5*(1), 8.
- Lear, K. O., Whitney, N. M., Brewster, L. R., Morris, J. J., Hueter, R. E., & Gleiss, A. C. (2017). Correlations of metabolic rate and body acceleration in three species of coastal sharks under contrasting temperature regimes. *Journal of Experimental Biology*, *220*(3), 397-407.
- Lear, K. O., Poulakis, G. R., Scharer, R. M., Gleiss, A. C., & Whitney, N. M. (2019). Fine-scale behavior and habitat use of the endangered smalltooth sawfish (*Pristis pectinata*): insights from accelerometry. *Fishery Bulletin*, *117*(4), 348-359.
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science*, *119*(1-2), 32-38.
- McClune, D. W., Marks, N. J., Wilson, R. P., Houghton, J. D., Montgomery, I. W., McGowan, N. E., Gormley, E., & Scantlebury, M. (2014). Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelemetry*, *2*(1), 5.
- Morgan, A., Carlson, J., Ford, T., Siceloff, L., Hale, L., Allen, M. S., & Burgess, G. (2010). Temporal and spatial distribution of finfish bycatch in the US Atlantic bottom longline shark fishery. *Marine Fisheries Review*, *72*(2), 34-38.
- Musyl, M. K., & Gilman, E. L. (2019). Meta-analysis of post-release fishing mortality in apex predatory pelagic sharks and white marlin. *Fish and Fisheries*, *20*(3), 466-500.
- Nakamura, I., Watanabe, Y. Y., Papastamatiou, Y. P., Sato, K., & Meyer, C. G. (2011). Yo-yo vertical movements suggest a foraging strategy for tiger sharks *Galeocerdo cuvier*. *Marine Ecology Progress Series*, *424*, 237-246.

- Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology*, 215(6), 986-996.
- Nguyen, H. D., Tran, K. P., Zeng, X., Koehl, L., & Tartare, G. (2020). An Improved Ensemble Machine Learning Algorithm for Wearable Sensor Data Based Human Activity Recognition. In *Reliability and Statistical Computing* (pp. 207-228). Springer, Cham.
- Oliver, S., Braccini, M., Newman, S. J., & Harvey, E. S. (2015). Global patterns in the bycatch of sharks and rays. *Marine Policy*, 54, 86-97.
- Pagano, A. M., Rode, K. D., Cutting, A., Owen, M. A., Jensen, S., Ware, J. V., Robbins, C. T., Durner, G. M., Atwood, T. C., Obbard, M. E., & Middel, K. R. (2017). Using tri-axial accelerometers to identify wild polar bear behaviors. *Endangered Species Research*, 32, 19-33.
- Papastamatiou, Y. P., Iosilevskii, G., Leos-Barajas, V., Brooks, E. J., Howey, L. A., Chapman, D. D., & Watanabe, Y. Y. (2018). Optimal swimming strategies and behavioral plasticity of oceanic whitetip sharks. *Scientific Reports*, 8(1), 1-12.
- Patterson, A., Gilchrist, H. G., Chivers, L., Hatch, S., & Elliott, K. (2019). A comparison of techniques for classifying behavior from accelerometers for two species of seabird. *Ecology and Evolution*, 9(6), 3030-3045.
- Poisson, F., Crespo, F. A., Ellis, J. R., Chavance, P., Bach, P., Santos, M. N., Séret, B., Korta, M., Coelho, R., Ariz, J., & Murua, H. (2016). Technical mitigation measures for sharks and rays in fisheries for tuna and tuna-like species: turning possibility into reality. *Aquatic Living Resources*, 29(4), 402.
- Probst, P., & Boulesteix, A. L. (2017). To tune or not to tune the number of trees in random forest. *The Journal of Machine Learning Research*, 18(1), 6673-6690.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), e1301.
- Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L. G., Shepard, E. L., Gleiss, A. C., & Wilson, R. (2012). Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we be summing values or calculating the vector?. *PloS One*, 7(2), e31187.

Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology*, 2(1), 27.

Rifkin, R., & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of Machine Learning Research*, 5(Jan), 101-141.

Rigby, C.L., Barreto, R., Carlson, J., Fernando, D., Fordham, S., Francis, M.P., Herman, K., Jabado, R.W., Liu, K.M., Marshall, A., Pacoureau, N., Romanov, E., Sherley, R.B., & Winker, H. (2019). *Carcharhinus longimanus*. *The IUCN Red List of Threatened Species 2019*: e.T39374A2911619. <https://dx.doi.org/10.2305/IUCN.UK.2019-3.RLTS.T39374A2911619.en>. Downloaded on 10 August 2020.

Chauvin, Y., & Rumelhart, D. E. (Eds.). (1995). *Backpropagation: theory, architectures, and applications*. Psychology press.

Ryan, L. A., Chapuis, L., Hemmi, J. M., Collin, S. P., McCauley, R. D., Yopak, K. E., Gennari, E., Huveneers, C., Kempster, R. M., Kerr, C. C., & Schmidt, C. (2018). Effects of auditory and visual stimuli on shark feeding behaviour: the disco effect. *Marine Biology*, 165(1), 11.

Sakai, K., Oishi, K., Miwa, M., Kumagai, H., & Hirooka, H. (2019). Behavior classification of goats using 9-axis multi sensors: The effect of imbalanced datasets on classification performance. *Computers and Electronics in Agriculture*, 166, 105027.

Sakamoto, K. Q., Sato, K., Ishizuka, M., Watanuki, Y., Takahashi, A., Daunt, F., & Wanless, S. (2009). Can ethograms be automatically generated using body acceleration data from free-ranging birds?. *PloS One*, 4(4), e5379.

Schapire, R. E., Freund, Y., Bartlett, P., & Lee, W. S. (1998). Boosting the margin: A new explanation for the effectiveness of voting methods. *The Annals of Statistics*, 26(5), 1651-1686.

Schölkopf, B., Smola, A. J., & Bach, F. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.

Shepard, E. L., Wilson, R. P., Quintana, F., Laich, A. G., Liebsch, N., Albareda, D. A., Halsey, L. G., Gleiss, A., Morgan, D. T., Myers, A. E., & Newman, C. (2008). Identification of animal movement patterns using tri-axial accelerometry. *Endangered Species Research*, 10, 47-60.

- Skomal, G. B., & Mandelman, J. W. (2012). The physiological response to anthropogenic stressors in marine elasmobranch fishes: a review with a focus on the secondary response. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 162(2), 146-155.
- Studd, E. K., Landry-Cuerrier, M., Menzies, A. K., Boutin, S., McAdam, A. G., Lane, J. E., & Humphries, M. M. (2019a). Behavioral classification of low-frequency acceleration and temperature data from a free-ranging small mammal. *Ecology and Evolution*, 9(1), 619-630.
- Studd, E. K., Boudreau, M., Majchrzak, Y., Menzies, A., Peers, M., Seguin, J., Lavergne, S. G., Boonstra, R., Murray, D L., Boutin, S., & Humphries, M. (2019b). Use of acceleration and acoustics to classify behavior, generate time budgets, and evaluate responses to moonlight in free-ranging snowshoe hares. *Frontiers in Ecology and Evolution*, 7, 154.
- Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. *Data Classification: Algorithms and Applications*, 37.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.
- Tolotti, M. T., Filmlalter, J. D., Bach, P., Travassos, P., Seret, B., & Dagorn, L. (2015). Banning is not enough: The complexities of oceanic shark management by tuna regional fisheries management organizations. *Global Ecology and Conservation*, 4, 1-7.
- Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics*, 85, 189-203.
- Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124, 203-220.
- Veldhuizen, L. J. L., Berentsen, P. B. M., de Boer, I. J. M., van de Vis, J. W., & Bokkers, E. A. M. (2018). Fish welfare in capture fisheries: a review of injuries and mortality. *Fisheries Research*, 204, 41-48.
- Wang, Y., Nickel, B., Rutishauser, M., Bryce, C. M., Williams, T. M., Elkaim, G., & Wilmers, C. C. (2015). Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. *Movement Ecology*, 3(1), 2.
- Whitney, N. M., Papastamatiou, Y. P., Holland, K. N., & Lowe, C. G. (2007). Use of an acceleration data logger to measure diel activity patterns in captive whitetip reef sharks, *Triaenodon obesus*. *Aquatic Living Resources*, 20(4), 299-305.

Whitney, N. M., White, C. F., Anderson, P. A., Hueter, R. E., & Skomal, G. B. (2017). The physiological stress response, postrelease behavior, and mortality of blacktip sharks (*Carcharhinus limbatus*) caught on circle and J-hooks in the Florida recreational fishery. *Fishery Bulletin*, 115(4), 532-544.

Wilson, R. P., Börger, L., Holton, M. D., Scantlebury, D. M., Gómez-Laich, A., Quintana, F., Rosell, F., Graf, P. M., Williams, H., Gunner, R., & Hopkins, L. (2020). Estimates for energy expenditure in free-living animals using acceleration proxies: A reappraisal. *Journal of Animal Ecology*, 89(1), 161-172.

Young, C. N., & Carlson, J. K. (2020). The biology and conservation status of the oceanic whitetip shark (*Carcharhinus longimanus*) and future directions for recovery. *Reviews in Fish Biology and Fisheries*, 1-20.

Zollett, E. A., & Swimmer, Y. (2019). Safe handling practices to increase post-capture survival of cetaceans, sea turtles, seabirds, sharks, and billfish in tuna fisheries. *Endangered Species Research*, 38, 115-125.

Table 3.1. Final ensemble weights assigned to predictions from each supervised classification method. Supervised classification methods included multinomial logistic regressions (MLR), k-nearest neighbors (KNN), support vector machines (SVM), multilayer perceptron neural network (MLP), random forests (RF), and extreme gradient boosted decision tress (XGBoost). Weights were identified using a brute-force grid search (resolution of 0.05) and cross-validation to find the optimal combination of weights for maximizing predictive performance without overfitting. Weights are presented as the mean \pm standard deviation, with the mean value used as the ensemble weight.

MLR	KNN	SVM	MLP	RF	XGBoost
0.00	0.19 \pm 0.07	0.08 \pm 0.08	0.02 \pm 0.04	0.01 \pm 0.02	0.70 \pm 0.09

Table 3.2. F1-measure performances. Performances (behavior-specific and macro; (mean \pm standard deviation) are from cross-validation of optimized supervised classification methods (multiple logistic regression, MLR; k-nearest neighbors, KNN; support vector machines, SVM; multilayer perceptron neural network, MLP; random forests, RF; extreme gradient boosted decision trees; XGBoost) and the final ensemble. Best and worst performing classification methods are identified with bold and italicized lettering, respectively.

Method	$F1_{\text{high-energy}}$	$F1_{\text{motionless}}$	$F1_{\text{steady swimming}}$	Macro-F1
MLR	<i>0.819 \pm 0.004</i>	<i>0.458 \pm 0.009</i>	<i>0.856 \pm 0.002</i>	<i>0.819 \pm 0.004</i>
KNN	0.844 \pm 0.002	0.799 \pm 0.003	0.896 \pm 0.002	0.846 \pm 0.002
SVM	0.863 \pm 0.003	0.839 \pm 0.007	0.905 \pm 0.002	0.869 \pm 0.003
MLP	0.872 \pm 0.002	0.890 \pm 0.005	0.916 \pm 0.001	0.893 \pm 0.001
RF	0.886 \pm 0.003	0.921 \pm 0.006	0.927 \pm 0.002	0.911 \pm 0.003
XGBoost ^a	0.895 \pm 0.002	0.947 \pm 0.003	0.935 \pm 0.001	0.926 \pm 0.002
Ensemble ^b	0.895 \pm 0.003	0.951 \pm 0.004	0.936 \pm 0.002	0.927 \pm 0.002

^a Extreme gradient boosting decision trees were the best performing supervised classification method

^b The ensemble displayed the best overall performance and was the final method used for predictions

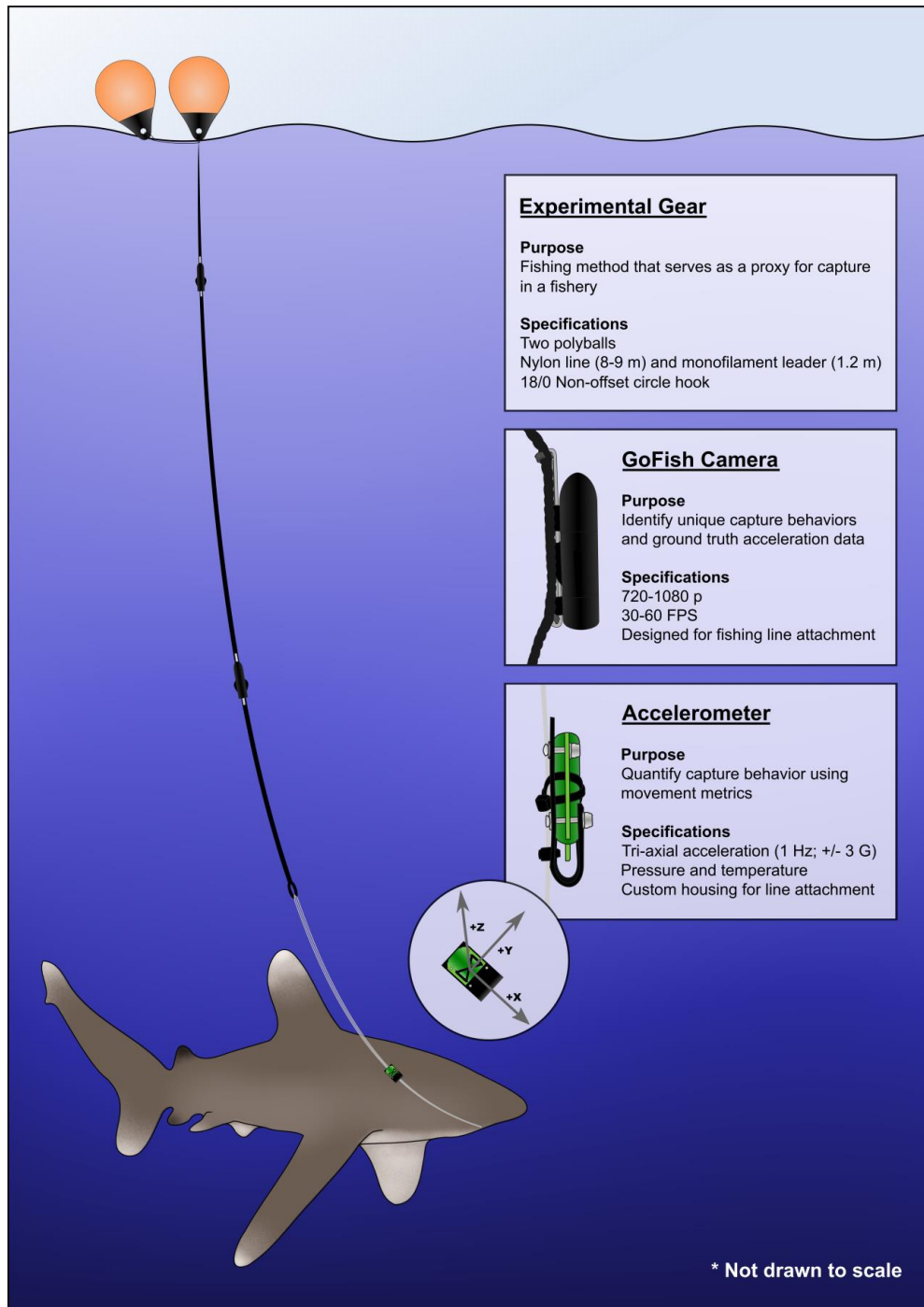


Figure 3.1. Diagram of the modified handline. Handlines were equipped with line-borne GoFish cameras and an accelerometer, used to capture oceanic whitetip shark and monitor and quantify their behavior.

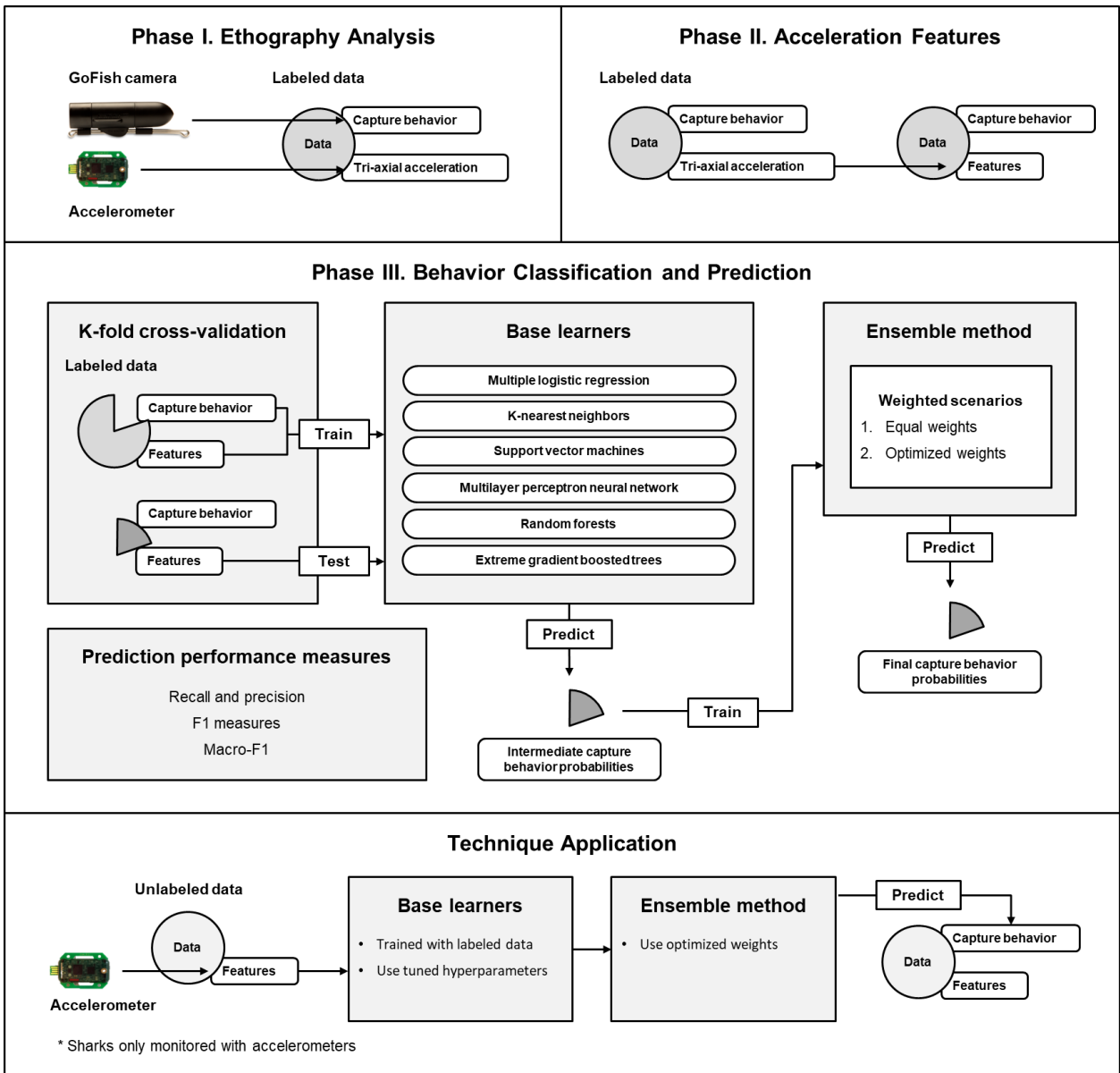


Figure 3.2. Framework of the phased-workflow in the capture behavior analysis.

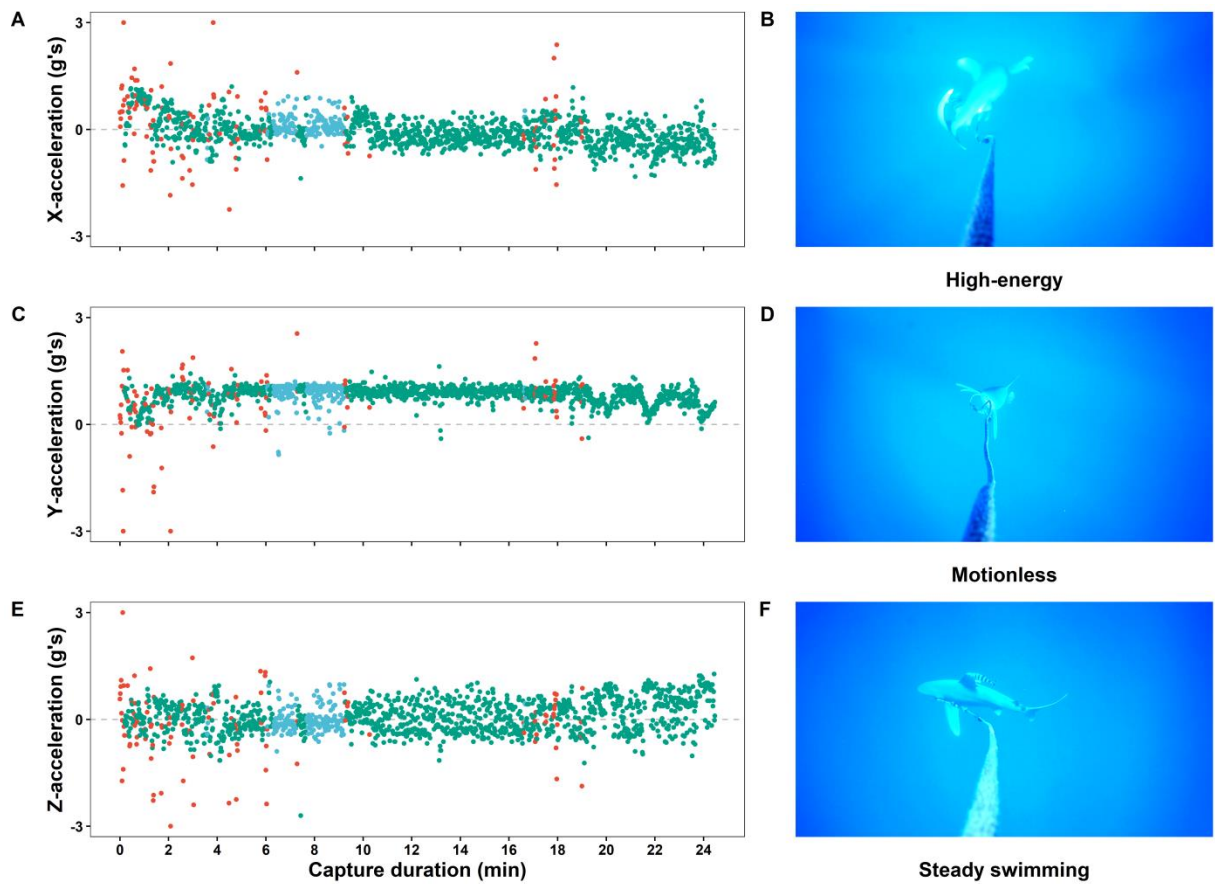


Figure 3.3. Raw tri-axial acceleration data and capture behaviors. Acceleration data (A, C, and E) are shown throughout a 24-minute capture event of an individual oceanic whitetip shark (185 cm fork length; female) with observations categorized by behaviors (red = high-energy; blue = motionless; and green = steady swimming) that were identified from simultaneously collected GoFish video footage. Screenshots of each behavior are provided in right-hand panels (B, D, and F).

SYNTHESIS AND FUTURE DIRECTIONS

The novel tools and techniques presented in this dissertation contribute to the “toolbox” researchers have when investigating capture-and-handling and its impact on incidentally caught and discarded (as bycatch) fish. Integral to this toolbox are tools for monitoring the fate of discarded animals, with pop-up satellite archival transmitting (PSAT) tags being one of the more popular monitoring tools (Fig. 4.1). Chapter I demonstrated how PSAT tags can be used to deduce fate in rajids for the first time, improving our ability to address discard mortality (DM) for these species that are often discarded at high rates. But these tags are not without shortcomings that need to be addressed moving forward. Namely, tag size precludes monitoring of smaller animals, and reduced resolution in transmitted data can make it difficult to identify mortalities via movement alone. The latter is particularly relevant for more sedentary demersal species (e.g., families *Dasyatidae*, *Rajidae*, *Pleuronectidae*, and *Lophiidae*) that exhibit intermittent, brief off-bottom forays that may go undetected at lower resolutions.

Pairing fate monitoring with tools and techniques that quantify factors related to capture-and-handling can also shed light on what is causing mortality (e.g., temperature and

air exposure; Fig. 4.1). However, underlying mechanisms that are more challenging to examine are often left to speculation, leaving a major gap in our understanding of DM. The technique developed in Chapter III improves our ability to evaluate one of these underlying mechanisms, capture behavior, in hook-and-line fisheries. But other gear-types where behavior has not yet been addressed and remains difficult to quantify (e.g., gillnet) require future directed research efforts. In addition, we need to continue exploring ways of better addressing other underlying mechanisms that are still poorly understood, but likely play a role in mortality. For example, chronic stressors of climate change and repeated capture events in areas of high fishing effort, may exacerbate acute stressors of capture-and-handling.

Moreover, when monitoring animals with electronic tags we should pay closer attention to whether the tag (application/bearing) leads to health consequences that contribute to mortality. Typically, minimally stressed captive animals tagged in laboratory trials are used to assess the impact of tags on health. But future studies might consider integrating laboratory trials with simulated capture to account for animals whose health has already been compromised due to capture, which may make them more susceptible to health consequences associated with these tags. In addition, we need to more clearly understand how electronic tags affect the cryptic abilities of species that rely upon being hidden for feeding and/or predator avoidance, both of which could contribute or result in mortality.

Tools and techniques from the toolbox provide information for fishery managers and other end-users that can be used mitigate the impact of capture-and-handling on bycatch species using a variety of strategies (Fig. 4.1). If addressing overfished bycatch species, more stringent strategies that reduce fisheries-interactions may be required to mitigate fishing

mortality. This can be accomplished with the help of the tool developed in Chapter II, which forecasts where fisheries-interactions are most likely to occur (i.e. hotspots). Time-area closures or other mitigation strategies (e.g., bycatch reduction devices) can be afforded to these hotspots as an effective strategy for reducing DM in larger, open access fisheries. Moving forward, resources should be put into developing a more dynamic forecasting approach and/or near real-time hotspot identification systems for these open-access fisheries, to better account for changes to fishing regulations and species distribution that displace fisheries-interactions.

The other objective of this dissertation was to provide capture-related information for two elasmobranch species that are often discarded by commercial fisheries. Thorny skate in the U.S. Northwest Atlantic Ocean (NWAOC) have been overfished for nearly 20 years despite their prohibited status, with little to no signs of recovery up until this point. In Chapters I & II, DM was addressed as a mechanism precluding this species recovery, with a focus on the Northeast (NE) multispecies fishery that is primarily responsible for discarding of thorny skate. But because DM in this fishery only accounted for ~ 1% of existing stock biomass, other factors are likely contributing to the status of this stock. In particular, recruitment may be a concern for this oviparous species whose embryos require extended incubation periods within egg cases deposited on the seafloor. Throughout this incubation period, egg cases are susceptible to damage from mobile demersal gear and subject to pronounced effects of climate change in the region (e.g., elevated seawater temperature and ocean acidification) that can impair embryonic development and reduce hatching success. However, it is not yet understood how this may be impacting recruitment for thorny skate and other oviparous

species (e.g., families Heterodontidae, Rajidae, Scyliorhinidae) and will require further research.

In addition, reduced recruitment could exacerbate DM's impact on a population and warrant additional bycatch mitigation measures even when mortality is low. The New England Fishery Management Council should therefore consider reducing fisheries-interactions for thorny skate to promote recovery by lowering DM and potentially, mitigating damage to egg cases (i.e., promoting recovery). Interactions can be reduced by using hotspots (Chapter II) as time-area closures or establishing bycatch limits and accountability measures that incentivize voluntary avoidance of these hotspots. Recent 100% monitoring requirements for the NE multispecies fishery may also permit near real-time hotspot identification as another strategy.

Oceanic whitetip shark populations have been severely depleted on a global-scale due to commercial fishing, which has resulted in widely-implemented prohibited measures on an international stage (e.g., Regional Fishery Management Organizations). However, without a clear understanding of how capture-and-handling impacts the welfare of this species, it is difficult to assess retention ban efficacy and formulate bycatch mitigation strategies. This dissertation (Chapter III) contributes to our knowledge of how this species interacts with pelagic longline fisheries, but ultimately, estimates of DM are required for this and the other fisheries that discard oceanic whitetip sharks. Moreover, additional bycatch mitigation strategies should be implemented by relevant management bodies (e.g., RFMOs) to promote the recovery of this species. In particular, time-area closures could be considered for ecologically important areas, while real-time communication between fishers (e.g.,

Amendment 5b to the Consolidated Atlantic Highly Migratory Species Fishery Management Plan) may provide an avenue for dynamic identification of hotspots for this highly migratory species. However, because of the transient nature of oceanic whitetip sharks, implementation and enforcement of any mitigation measures will be a challenge and ultimately require cooperation across multiple jurisdictions.

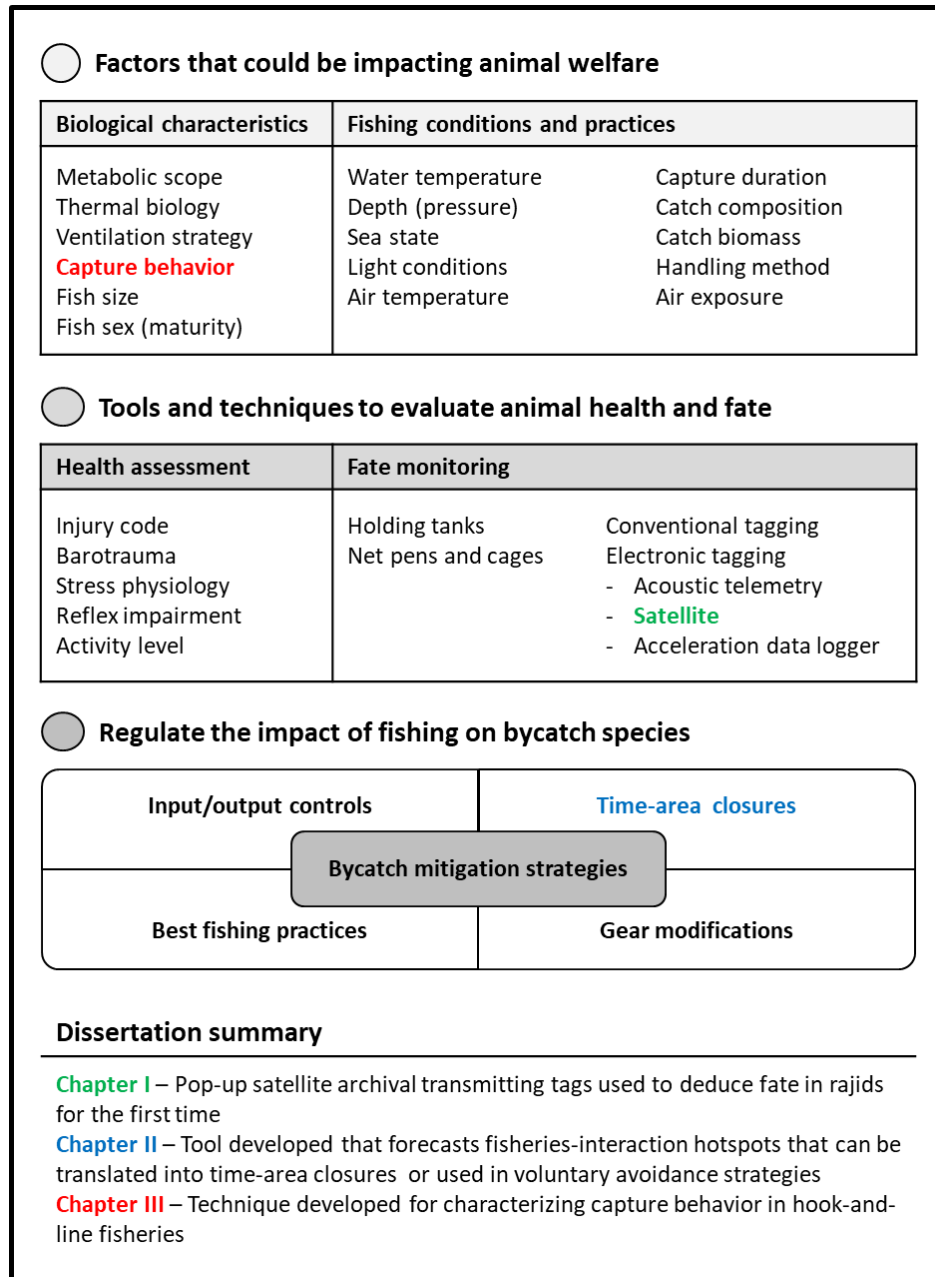


Figure 4.1. Key components in the “toolbox” researchers have when addressing the impact of capture-and-handling on bycatch species. Biological characteristics and fishing conditions/practices (relative to gear-types mentioned in this dissertation) that could impact animal welfare are listed, as well as the various tools and techniques that can quantify animal health/fate. The primary forms of bycatch mitigation strategies are also provided. Lastly, areas of the toolbox that this dissertation contributed to are highlighted with colors coded to each chapter, with a summary of the work provided at the bottom of the diagram.

APPENDIX A

CHAPTER 1 SUPPLEMENTARY MATERIAL

Description of depth-variance survival test

The following steps of the depth-variance survival test were adapted from Capizzano et al. (2016) and modified to reflect thorny skate behavior.

- (i) Each candidate skates' time-series were parsed into bins of equal size with respect to the resolution of the data (10-second = one hour bins; 5-minute = 6 hour bins) to ensure equal sensitivity of the test and provide a temporal component.
- (ii) Depth variance from each bin of a candidate skate was sequentially compared to the depth variance of the negative control (i.e., known dead skate) using a one-tailed t-test of the absolute difference from the median (i.e., modified Browne-Forsythe-Levene test for homogeneity of population variance; Ott, 2010).
- (iii) Mortality was inferred from this test if there were no significant differences ($p \geq 0.05$) in depth variance between the candidate skate and negative control for a period of 102 hours from the final observation. This timeframe reflected the maximum duration between off-seafloor movements in live skate (characterized a priori using higher resolution data from recovered tags, see Results) and was used to identify cases where the animal was exhibiting on-bottom (live) behavior at the end of deployment, which would otherwise be falsely characterized as a mortality event.

Description of model building for the Cox proportional-hazards model

Part I

Determine if a random effect (i.e., mixed-effect CPHM) is appropriate for the data using the following steps:

- (i) Fit a fully-saturated (i.e., all fishing conditions and practices, biological traits, and injury scores) mixed-effects CPHM.
- (ii) Compute the likelihood ratio-statistic against a Chi-squared distribution and correct the resulting p-value for boundary testing (Knotek et al., 2018).
- (iii) If the p-value was significant ($p \leq 0.05$) = retain the random effect and proceed with a mixed-effect modeling approach. Or, if the p-value was not significant ($p > 0.05$) = discard the random effect and utilize a fixed-effects modeling approach for the remainder of the analysis.

Part II

Find the most parsimonious set of covariates to explain survivorship using a forward selection process outlined in the following steps:

- (i) Incrementally add covariates to the intercept-only model and compare these candidate models with Akaike's Information Criterion corrected for small sample sizes (AICc; Burnham and Anderson, 2002).

- (ii) If the candidate model (i.e., additional covariate) resulted in a reduction of three or more units in AICc (relative to intercept-only model), the covariate is retained.
- (iii) Repeat these steps until additional covariates no longer reduce the AICc by more than three units. If models are produced with an equal number of covariates and have AICc values within three units of each other, these models were considered equally plausible and retained.

Model fit comparison of Survival Mixture Models

Model fits were compared via Akaike's Information Criterion corrected for small sample sizes (AICc) with support for each model evaluated under the following Δ AICc (relative to the lowest AICc value) guidelines:

- (i) If Δ AIC < 2 units, the competing model shows comparable support to the best-fit model
- (ii) If Δ AIC is between 3 to 7 units, the competing models shows less support to the best-fit model
- (iii) If Δ AIC > 10 units, the competing model is not supported

Model averaging of Survival Mixture Models

Model averaging is appropriate when no single model-variant explains the data (i.e., survivorship) better than others. In this case a single estimate of survivorship is estimated from all model-variants using the following equation:

$$\tilde{\beta} = \sum_{i=1}^R w_i \hat{\beta}_i \quad (\text{A1})$$

where $\tilde{\beta}$ is the model-averaged post-release survival for all model-variants, $\hat{\beta}_i$ is the model variant-specific survival estimate for model i , and w_i is the Akaike weight (Akaike, 1978, 1979) for model i (i.e., the probability of model i providing the best fit amongst the several competing models R ; Tables 1.2). The Akaike weight is defined as:

$$w_i = \frac{\exp(-\frac{1}{2}\Delta AICc_i)}{\sum_{r=1}^R \exp(-\frac{1}{2}\Delta AICc_r)} \quad (\text{A2})$$

Monte Carlo simulations based on bootstrapping

Monte Carlo simulations were performed to estimate the variability of fishery-scale discard DM using the following the steps (per iteration):

- (i) Random parameter values for each model variant (i.e., SMMs; Table 1.2) were drawn from multivariate normal distributions that were conditional on model variant-specific maximum-likelihood values and their associated covariance matrix (i.e., size categories; Benoît et al., 2012).
- (ii) These parameter values were then used to calculate the size-specific probability of survival (i.e., survival rate) for each model variant using Eq. (2).

- (iii) Model averaging based on Akaike weights (Eq. A1 & A2) was then performed to assimilate model variant size-specific survival rates into an overall estimate of survival for each size category.
- (iv) Size-specific survival rates were converted into a mortality rate (i.e., DM rate = 1 – survival rate).
- (v) The DM rate from the smaller-sized category (i.e., 55.5 – 70.0 cm TL) was applied to the estimate of the annual number of mortalities for discarded thorny skate.

This simulation was repeated for 5,000 iterations to derive a fishery-scale estimate (mean and standard deviation) of DM for thorny skate captured in the GOM groundfish bottom trawl fishery.

Literature cited

Akaike, H. (1978). On newer statistical approaches to parameter estimation and structure determination. *Int. Fed. Auto. Ctrl.*, 3, 1877-1884.

Akaike, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika*, 66, 237-242.

Benoît, H.P., Hurlbut, T., Chasse, J., & Jonsen, I.D. (2012). Estimating fishery-scale rates of discard mortality using conditional reasoning. *Fish. Res.*, 125-126, 318-330.

Burnham, K.P., & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*, 2nd ed. Springer-Verlag, New York

Capizzano, C.W., Mandelman, J.W., Hoffman, W.S., Dean, M.J., Zemeckis, D.R., Benoît, H.P., & Langan, J.A. (2016). Estimating and mitigating the discard mortality of Atlantic cod (*Gadus morhua*) in the Gulf of Maine recreational rod-and-reel fishery. *ICES J. Mar. Sci.: Journal du Conseil*, fsw058.

Knotek, R. J., Rudders, D. B., Mandelman, J. W., Benoît, H. P., & Sulikowski, J. A. (2018). The survival of rajids discarded in the New England scallop dredge fisheries. *Fish. Res.*, 198, 50-62.

Ott, R. (2010). Lyman and Michael Longnecker. *An introduction to statistical methods and data analysis*.

1| None to minor

2| Moderate

3| Severe



Figure A.1. Examples of thorny skate assigned to each injury code. This injury code was adapted from Mandelman et al. (2013) and Knotek et al. (2018) and is based upon increasing overt physical trauma.

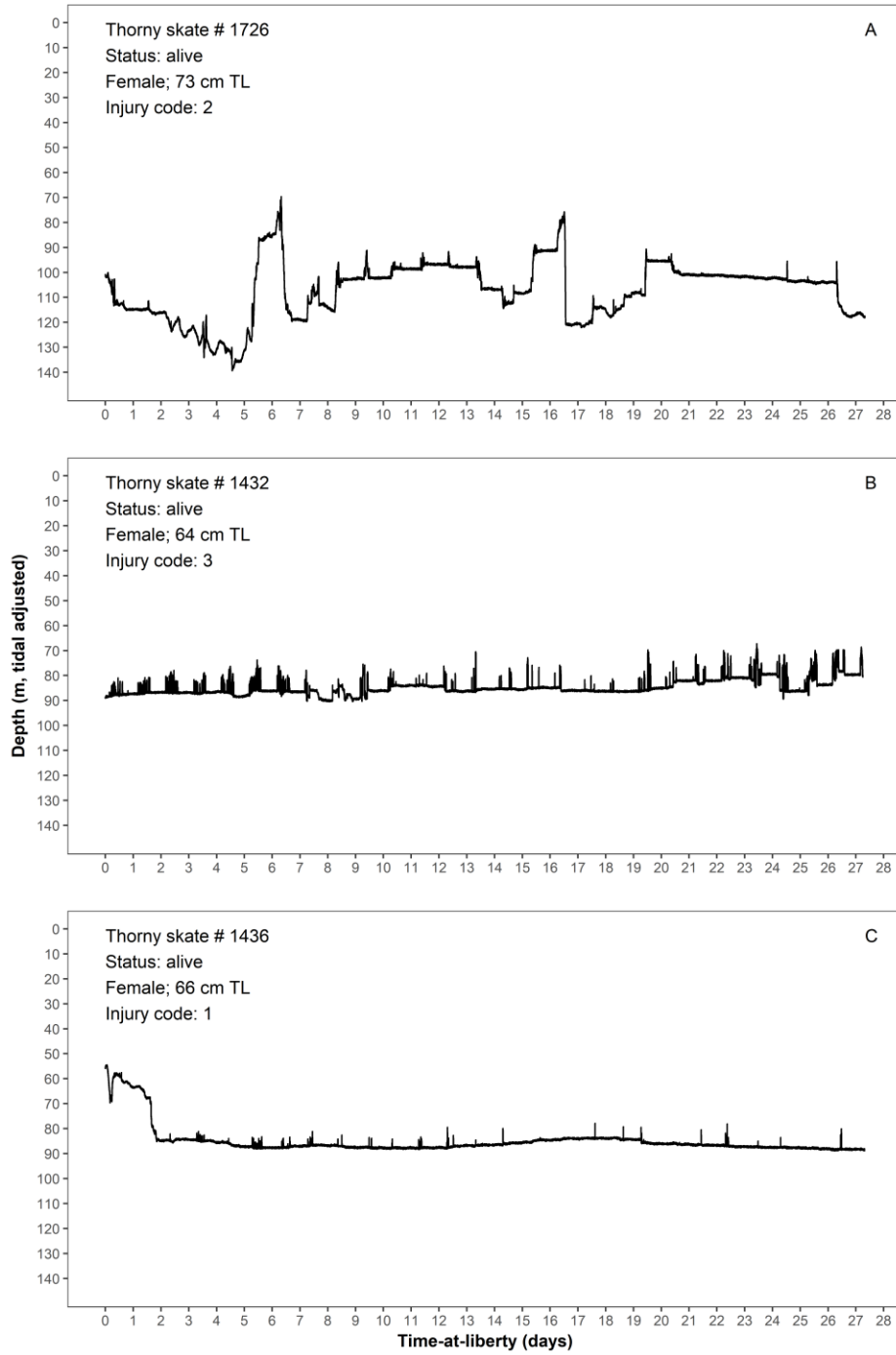


Figure A.2. Tide-adjusted depth time-series (black line) obtained from pop-up satellite archival transmitting tags (10-second resolution) deployed on three individual skate that show different vertical movement patterns including (A) on/off shelf movements and diel vertical migrations that occurred at both (B) high and (C) low activity levels.

APPENDIX B

CHAPTER 3 SUPPLEMENTARY MATERIAL

ReliefF algorithm

The ReliefF algorithm is a popular filtering selection method that estimates the independence (i.e. weight) of the information provided by a feature, relative to its k nearest neighbors, to a classification problem (for more details refer to Kononenko 1994). Here, we chose to consider the 10-nearest neighbors based on recommendations in Urbanowicz et al., (2018). The advantage to this method is that the selected features are not biased to any induction algorithm (e.g. wrapper or embedded methods; Tibshirani et al., 1996; Bradley et al., 1998), which means the subset of selected features can then be used across any classification techniques (Urbanowicz et al., 2018). Feature weights generated from the ReliefF algorithm identified 13 features with markedly low weights (< 0.005) that were omitted from the remaining analyses.

Description of supervised classification methods

Multinomial logistic regression is one of the most fundamental and widely-used classification techniques in acceleration-based behavioral studies (Catal et al., 2015; Ladds et al., 2016, 2017; Brewster et al., 2018; Nguyen et al., 2020). Herein, MLRs were constructed using a “one-versus-all” scheme comprised of three binary logistic regressions (Rifkin and Klautau 2004). In each, the model predicts the probability of new data belonging to a given

class (as opposed to the remaining) using a logistic (sigmoidal) function that can be mapped to two discrete classes using a decision threshold. The highest class probability is then assigned as the final classification for a given observation (Hosmer et al., 2013). MLRs in this study were constructed with normalized features (zero mean, unit variance), and no hyperparameters were considered as this served as our most simplistic, baseline learner (Table B.2).

K-nearest neighbor is an instance-based classification technique that operates much differently than other learners used in this study (Bidder et al., 2014; McClune et al., 2014; Rescheff et al., 2014; Patterson et al., 2018; Benaissa et al., 2019). This technique uses all of the input features directly for classification without building a model that requires learning (Dasarthy 1991). Classification is then simply assigned based on the majority vote of classes from the surrounding *k* neighbors, with a distance measure (e.g. Euclidean distance) weighting each neighbors' vote (Dasarthy 1991). KNNs in this study were applied to normalized features, and the hyperparameter of *k*, the number of nearest neighbors considered when voting, was tuned to maximize predictive performance and ensure that our modeling was not overfitting (Table B.2).

Support vector machines are one of the most popular classification techniques (e.g. Martiskainen et al., 2009; Nathan et al., 2012; Campbell et al., 2013; Escalante et al., 2013; Gerenscer et al., 2013; Carroll et al., 2014; Rescheff et al., 2014; Hammond et al., 2016; Ladds et al., 2016; Hoang et al., 2018; Benaissa et al., 2019). The objective of the SVM algorithm is to find the best hyperplane in *N*-dimensional space (*N* = the number of input features) that maximizes the margin between classes, which it then uses as a separation plane

for classification of new data (Cristianini and Taylor 2000). Hyperplanes can be made non-linear in SVMs using a kernelling function, such as the radial basis function selected in this study (Schölkopf and Smola 2002). SVMs were constructed herein using normalized data, while the hyperparameters gamma and C were tuned, which control the spread of kernel and cost of misclassification, respectively (Table B.2; Chang and Lin 2016).

Multilayer perceptron neural networks are one of the most widely used types of neural networks in classification tasks (e.g., Nathan et al., 2012; Escalante et al., 2013; Cerezuela-Escudero et al., 2016; Brewster et al., 2018; Patterson et al., 2019). The architecture of the three-layer MLP (the frequentist approach) used in this study starts with an input layer that consists of all candidate features that are then fully-connected, with weights assigned to each connection, to nodes within the hidden layer (Hornik et al., 1989). Each node within the hidden layer maps the sum of the connected weighted features with a bias term, using a non-linear activation function (i.e. rectified linear units; Glorot et al., 2011). This produces an output value that is fed-forward to the final output layer with another set of weighted connections and bias term (Hornik et al., 1989). Here, multi-classification predictions of capture behavior are made using the softmax function (Bridle 1990). MLP weights and bias terms are optimized during training via the back-propagation of prediction errors and gradient descent (Hornik et al., 1989). MLPs were built in this study using normalized data and the hyperparameters of learning rate and number of hidden nodes were tuned to optimize its predictive performance (Bengio 2012; Table B.2). MLP training was terminated early if the cost-function for the training data showed no reduction within 100

iterations (with a maximum of 1,000 iterations), to avoid overfitting, and mini-batch gradient descent ($n = 32$; Bengio 2012) was used to reduce computational costs.

Random forests are one of the most commonly used ensembling classification techniques in field of acceleration-based behavioral studies, particularly in recent years (e.g. Bom et al., 2010, 2014; Graf et al., 2015; Wang et al., 2015; Alvarenga et al., 2016; Ladds et al., 2016, 2017; Fehlman et al., 2017; Brewster et al., 2018; Jeantet et al., 2018; Hounslow et al., 2019; Studd et al., 2019a, 2019b). RFs assign classification based on majority class vote from an ensemble of decision trees that are independently grown via recursively splitting the data, at nodes, according to the features that lead to the greatest reduction in Gini impurity (i.e. a measure of a features ability to classify data; Breiman 2001). Individual decision trees can be prone to overfitting and high variance; however, because RFs rely upon the average of many trees constructed with bootstrapped samples of the data, their predictive performance is largely safeguarded against these pitfalls (Breiman 2001). In addition, individual nodes of decision trees are only supplied with a random subset of features from which to split (based on Gini impurity), which is an additional RF technique to reduce the overall variance (Breiman 2001). RFs were constructed in this study using the selected features without normalization, a large number of decision trees to ensure stabilization of results ($n = 1,000$; Probst and Boulesteix 2017), and tuned over a varying number of features considered at each node (i.e. *mtry*; Probst et al., 2019; Table B.2).

Gradient boosted decision **trees** are another popular ensemble classification technique in the field of acceleration-based behavioral studies (e.g. Ladds et al., 2016, 2017; Brewster et al., 2018). In this technique, classification is assigned as the weighted prediction from an

ensemble of shallow decision trees that are iteratively built using pseudo-residual errors from the previous tree (Friedman 2001). This sequential approach allows the model to learn from misclassified predictions and reduce error in subsequent trees via gradient descent (Friedman 2001). In this study, we utilized a specific implementation of this technique called XGBoost (extreme gradient boosting), characterized by a more effective approach to minimizing the loss function via calculating its second-order gradient (Chen and Guestrin 2016). XGBoost was applied to non-normalized data and the two most commonly tuned hyperparameters were optimized to control for overfitting and improve modeling performance. This included building a large number of trees ($n = 1,000$, similar to RFs) while varying the maximum tree depth (i.e. number of nodes) and learning rate to control the complexity of feature relationships and weight of each tree's prediction, respectively (Table B.2; Friedman 2001, Chen and Guestrin, 2016). Similar to MLPs, XGBoost training was terminated early if the cost-function was not minimized within 100 iterations.

Literature cited

Alvarenga, F. A. P., Borges, I., Palkovič, L., Rodina, J., Oddy, V. H., & Dobos, R. C. (2016). Using a three-axis accelerometer to identify and classify sheep behaviour at pasture. *Applied Animal Behaviour Science*, *181*, 91-99.

Benaissa, S., Tuytens, F. A., Plets, D., De Pessemier, T., Trogh, J., Tanghe, E., Martens, L., Vandaele, L., Van Nuffel, A., Joseph, W., & Sonck, B. (2019). On the use of on-cow accelerometers for the classification of behaviours in dairy barns. *Research in Veterinary Science*, *125*, 425-433.

Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. In *Neural networks: Tricks of the Trade* (pp. 437-478). Springer, Berlin, Heidelberg.

Bidder, O. R., Campbell, H. A., Gómez-Laich, A., Urgé, P., Walker, J., Cai, Y., Gao, L., Quintana, F., & Wilson, R. P. (2014). Love thy neighbour: automatic animal behavioural classification of acceleration data using the k-nearest neighbour algorithm. *PloS One*, *9*(2), e88609.

Bom, R. A. (2010). *Can speed and tri-axial acceleration measured by biologgers be used to classify oystercatcher behaviour* (Doctoral dissertation, Master's thesis, University of Amsterdam, The Netherlands).

Bom, R. A., Bouten, W., Piersma, T., Oosterbeek, K., & van Gils, J. A. (2014). Optimizing acceleration-based ethograms: the use of variable-time versus fixed-time segmentation. *Movement Ecology*, *2*(1), 6.

Bradley, P. S., & Mangasarian, O. L. (1998, July). Feature selection via concave minimization and support vector machines. In *ICML* (Vol. 98, pp. 82-90).

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Brewster, L. R., Dale, J. J., Guttridge, T. L., Gruber, S. H., Hansell, A. C., Elliott, M., Cowx, I. G., Whitney, N. M., & Gleiss, A. C. (2018). Development and application of a machine learning algorithm for classification of elasmobranch behaviour from accelerometry data. *Marine Biology*, *165*(4), 62.

Bridle, J. S. (1990). Probabilistic interpretation of feed forward classification network outputs, with relationships to statistical pattern recognition. In *Neurocomputing* (pp. 227-236). Springer, Berlin, Heidelberg.

Campbell, H. A., Gao, L., Bidder, O. R., Hunter, J., & Franklin, C. E. (2013). Creating a behavioural classification module for acceleration data: using a captive surrogate for difficult to observe species. *Journal of Experimental Biology*, 216(24), 4501-4506.

Carroll, G., Slip, D., Jonsen, I., & Harcourt, R. (2014). Supervised accelerometry analysis can identify prey capture by penguins at sea. *Journal of Experimental Biology*, 217(24), 4295-4302.

Catal, C., Tufekci, S., Pirmit, E., & Kocabag, G. (2015). On the use of ensemble of classifiers for accelerometer-based activity recognition. *Applied Soft Computing*, 37, 1018-1022.

Cerezuela-Escudero, E., Rios-Navarro, A., Dominguez-Morales, J. P., Tapiador-Morales, R., Gutierrez-Galan, D., Martín-Cañal, C., & Linares-Barranco, A. (2016). Performance evaluation of neural networks for animal behaviors classification: Horse gaits case study. In *Distributed Computing and Artificial Intelligence, 13th International Conference* (pp. 377-385). Springer, Cham.

Chang, C. C., & Lin, C. J. (2016). *A practical guide to support vector classification*. Department of Computer Science, National Taiwan University, Taipei, 106.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.

Dasarathy, B. V. (1991). Nearest neighbor (NN) norms: NN pattern classification techniques. *IEEE Computer Society Tutorial*.

Escalante, H. J., Rodriguez, S. V., Cordero, J., Kristensen, A. R., & Cornou, C. (2013). Sow-activity classification from acceleration patterns: a machine learning approach. *Computers and Electronics in Agriculture*, 93, 17-26.

Fehlmann, G., O'Riain, M. J., Hopkins, P. W., O'Sullivan, J., Holton, M. D., Shepard, E. L., & King, A. J. (2017). Identification of behaviours from accelerometer data in a wild social primate. *Animal Biotelemetry*, 5(1), 6.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189-1232.

Gerencsér, L., Vásárhelyi, G., Nagy, M., Vicsek, T., & Miklósi, A. (2013). Identification of behaviour in freely moving dogs (*Canis familiaris*) using inertial sensors. *PloS One*, 8(10).

Gleiss, A. C., Wilson, R. P., & Shepard, E. L. (2011). Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, 2(1), 23-33.

Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (pp. 315-323).

Graf, P. M., Wilson, R. P., Qasem, L., Hackländer, K., & Rosell, F. (2015). The use of acceleration to code for animal behaviours; a case study in free-ranging Eurasian beavers *Castor fiber*. *PloS One*, 10(8).

Hammond, T. T., Springthorpe, D., Walsh, R. E., & Berg-Kirkpatrick, T. (2016). Using accelerometers to remotely and automatically characterize behavior in small animals. *Journal of Experimental Biology*, 219(11), 1618-1624.

Hoang, Q. T., Phi Khanh, P. C., Trung Ninh, B., Phuong Dung, C. T., & Tran, T. (2018). Cow Behavior Monitoring Using a Multidimensional Acceleration Sensor and Multiclass SVM. *International Journal of Machine Learning and Networked Collaborative Engineering*, 2(03), 110-118.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feed forward networks are universal approximators. *Neural Networks*, 2(5), 359-366.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Hounslow, J. L., Brewster, L. R., Lear, K. O., Guttridge, T. L., Daly, R., Whitney, N. M., & Gleiss, A. C. (2019). Assessing the effects of sampling frequency on behavioural classification of accelerometer data. *Journal of Experimental Marine Biology and Ecology*, 512, 22-30.

Jeantet, L., Dell'Amico, F., Forin-Wiart, M. A., Coutant, M., Bonola, M., Etienne, D., Gresser, J., Regis, S., Lecerf, N., Lefebvre, F., & de Thoisy, B. (2018). Combined use of two supervised learning algorithms to model sea turtle behaviours from tri-axial acceleration data. *Journal of Experimental Biology*, 221(10).

Kononenko, I. (1994, April). Estimating attributes: analysis and extensions of RELIEF. In *European Conference on Machine Learning* (pp. 171-182). Springer, Berlin, Heidelberg.

Ladds, M. A., Thompson, A. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2016). Seeing it all: evaluating supervised machine learning methods for the classification of diverse otariid behaviours. *PLoS One*, 11(12).

- Ladds, M. A., Thompson, A. P., Kadar, J. P., Slip, D. J., Hocking, D. P., & Harcourt, R. G. (2017). Super machine learning: improving accuracy and reducing variance of behaviour classification from accelerometry. *Animal Biotelemetry*, 5(1), 8.
- Martiskainen, P., Järvinen, M., Skön, J. P., Tiirikainen, J., Kolehmainen, M., & Mononen, J. (2009). Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Applied Animal Behaviour Science*, 119(1-2), 32-38.
- McClune, D. W., Marks, N. J., Wilson, R. P., Houghton, J. D., Montgomery, I. W., McGowan, N. E., Gormley, E., & Scantlebury, M. (2014). Tri-axial accelerometers quantify behaviour in the Eurasian badger (*Meles meles*): towards an automated interpretation of field data. *Animal Biotelemetry*, 2(1), 5.
- Nathan, R., Spiegel, O., Fortmann-Roe, S., Harel, R., Wikelski, M., & Getz, W. M. (2012). Using tri-axial acceleration data to identify behavioral modes of free-ranging animals: general concepts and tools illustrated for griffon vultures. *Journal of Experimental Biology*, 215(6), 986-996.
- Nguyen, H. D., Tran, K. P., Zeng, X., Koehl, L., & Tartare, G. (2020). An Improved Ensemble Machine Learning Algorithm for Wearable Sensor Data Based Human Activity Recognition. In *Reliability and Statistical Computing* (pp. 207-228). Springer, Cham.
- Patterson, A., Gilchrist, H. G., Chivers, L., Hatch, S., & Elliott, K. (2019). A comparison of techniques for classifying behavior from accelerometers for two species of seabird. *Ecology and Evolution*, 9(6), 3030-3045.
- Probst, P., & Boulesteix, A. L. (2017). To tune or not to tune the number of trees in random forest. *The Journal of Machine Learning Research*, 18(1), 6673-6690.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), e1301.
- Qasem, L., Cardew, A., Wilson, A., Griffiths, I., Halsey, L. G., Shepard, E. L., Gleiss, A. C., & Wilson, R. (2012). Tri-axial dynamic acceleration as a proxy for animal energy expenditure; should we be summing values or calculating the vector?. *PloS One*, 7(2), e31187.
- Resheff, Y. S., Rotics, S., Harel, R., Spiegel, O., & Nathan, R. (2014). AcceleRater: a web application for supervised learning of behavioral modes from acceleration measurements. *Movement Ecology*, 2(1), 27.

Rifkin, R., & Klautau, A. (2004). In defense of one-vs-all classification. *Journal of Machine Learning Research*, 5(Jan), 101-141.

Schölkopf, B., Smola, A. J., & Bach, F. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.

Studd, E. K., Landry-Cuerrier, M., Menzies, A. K., Boutin, S., McAdam, A. G., Lane, J. E., & Humphries, M. M. (2019a). Behavioral classification of low-frequency acceleration and temperature data from a free-ranging small mammal., *Ecology and Evolution*, 9(1), 619-630.

Studd, E. K., Boudreau, M., Majchrzak, Y., Menzies, A., Peers, M., Seguin, J., Lavergne, S. G., Boonstra, R., Murray, D L., Boutin, S., & Humphries, M. (2019b). Use of acceleration and acoustics to classify behavior, generate time budgets, and evaluate responses to moonlight in free-ranging snowshoe hares. *Frontiers in Ecology and Evolution*, 7, 154.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics*, 85, 189-203.

Wang, Y., Nickel, B., Rutishauser, M., Bryce, C. M., Williams, T. M., Elkaim, G., & Wilmers, C. C. (2015). Movement, resting, and attack behaviors of wild pumas are revealed by tri-axial accelerometer measurements. *Movement Ecology*, 3(1), 2.

Table B.1. Descriptions of the features generated from tri-axial acceleration data that were used to predict capture behavior in hook-and-line caught oceanic whitetip shark.

Features	Description
Static acceleration (SA; g)	Calculated for each axis with a 3-second box-smoothing window (Shephard et al. 2008) of the raw acceleration data
Line pitch (°)	Calculated using the following equation adapted from HOBO Pendant G Data Logger manual (Onset Computer Corporation): $\text{Pitch} = 180 - \arccos(SA_x) * 180/\pi$
Dynamic acceleration (DA; g)	Calculated for each axis as the difference in static and raw acceleration data (Gleiss et al. 2011)
Zero-crossings	Calculated for each axis as the number of times dynamic acceleration changes signs (+/-) in a 3-second moving window
Vectorial dynamic acceleration (g)	Calculated using the following equation (Qasem et al. 2012): $\sqrt{(DA_x^2 + DA_y^2 + DA_z^2)}$
Overall dynamic acceleration (ODA; g)	Calculated using the following equation (Wilson et al. 2006): $ DA_x + DA_y + DA_z $
Partial dynamic acceleration (PDA; g)	Calculated for each axis as the absolute value of dynamic acceleration
Dynamic acceleration ratio	Calculated for using the following equation: $\frac{PDA_x \text{ or } PDA_y \text{ or } PDA_z}{ODA}$

Table B.2. Base learners used to predict capture behavior of hook-and-line caught oceanic whitetip sharks. Relevant fixed modeling parameters (i.e. selected without tuning) and hyperparameters are provided along with the R statistical program function and package (italicized) used to implement these base learners. Hyperparameters were tuned using a grid-search across the provided range of values, respective to the hyperparameter, with a 5-fold stratified cross-validation approach that identified the values (shown in italics) that maximized overall classification performance with the macro-F1 measure.

Base learner	R statistical program	Fixed parameters	Hyperparameters
MLR	<i>multinom</i> <i>nnet</i>	-	-
KNN	<i>knn3</i> <i>caret</i>	Distance: Euclidean	K: 1 – 15 <i>5</i>
SVM	<i>svm</i> <i>e1071</i>	Kernel: RBF ^a	Cost: $1e^{-6} - 1e^6$ <i>$1e^4$</i> Gamma: $1e^{-6} - 1e^6$ <i>$1e^1$</i>
MLP	<i>keras_model_sequential</i> <i>fit</i> <i>keras</i>	Activation: RELU ^b Softmax ^c Batch size: 32	Hidden nodes: 16 – 178 <i>92</i> Learning rate: $1e^{-3} - 1e^{-1}$ <i>$1e^{-1}$</i> Epochs: 1,000 <i>Stopping^d</i>
RF	<i>ranger</i> <i>ranger</i>	Trees: 1,000	Mtry: 1 – 30 <i>19</i>
XGBoost	<i>xgb.train</i> <i>xgboost</i>	-	Learning rate: 0.01 – 0.1 <i>0.05</i> Tree depth: 1 – 20 <i>10</i> Trees: 1,000 <i>Stopping^d</i>

a RBF = radial basis function

b RELU = rectified linear units; selected as the hidden layer activation function

c Softmax was selected as the output layer activation function

d Number of learning iterations (i.e. epochs) or trees were reduced in cases where model construction was stopped early to avoid overfitting (i.e. instances where multiclass log-loss showed no decline in the testing data after 10 iterations or 100 trees, respectively).

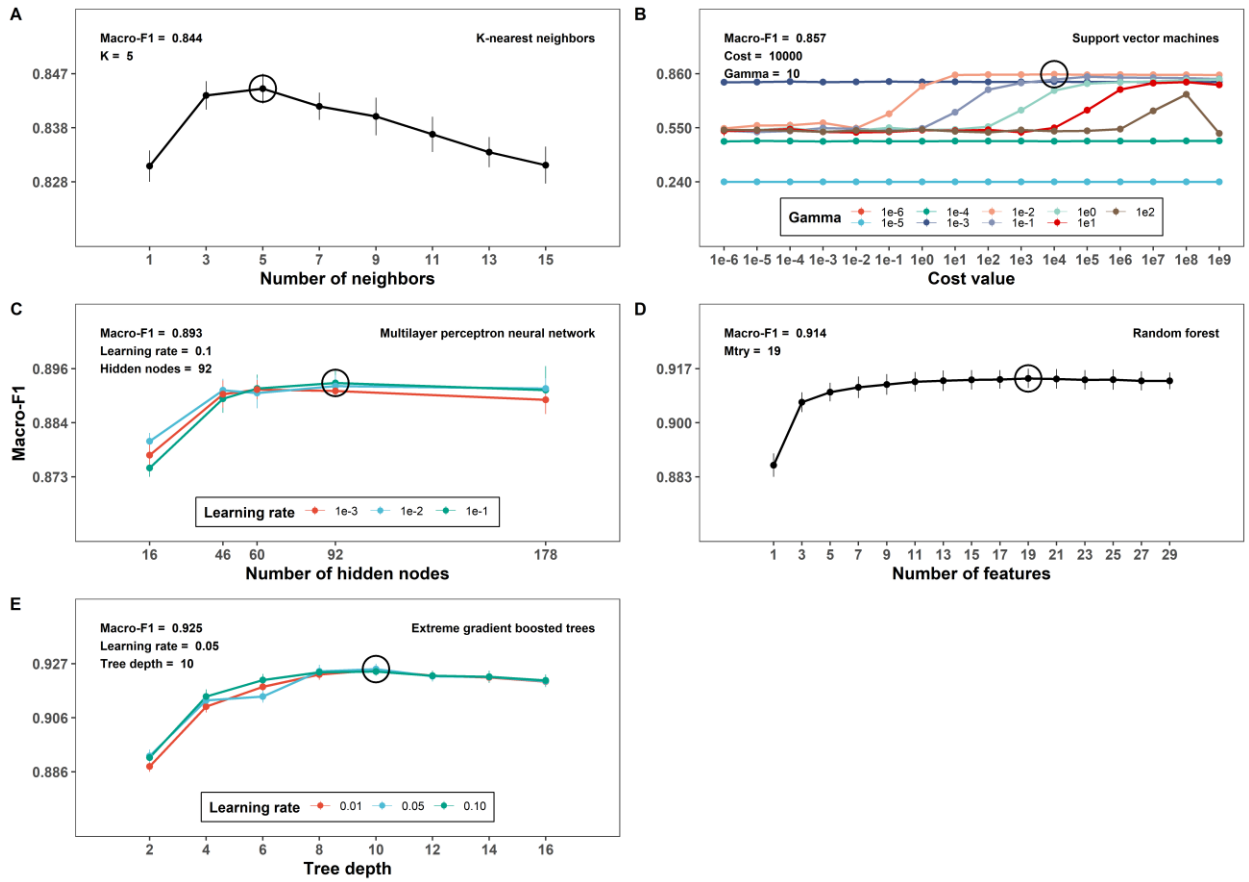


Figure B.1. Overall performance measures (macro-F1) for hyperparameter values considered during the tuning process of each base learner. Macro-F1 measures are presented as the mean \pm standard deviation from five-fold cross-validation stratified across behavioral categories and majority down sampling of the steady swimming behavior. Hyperparameter values that best-optimized performance are provided in the top-left corner of each panel and indicated within the plotted data using a black circle.

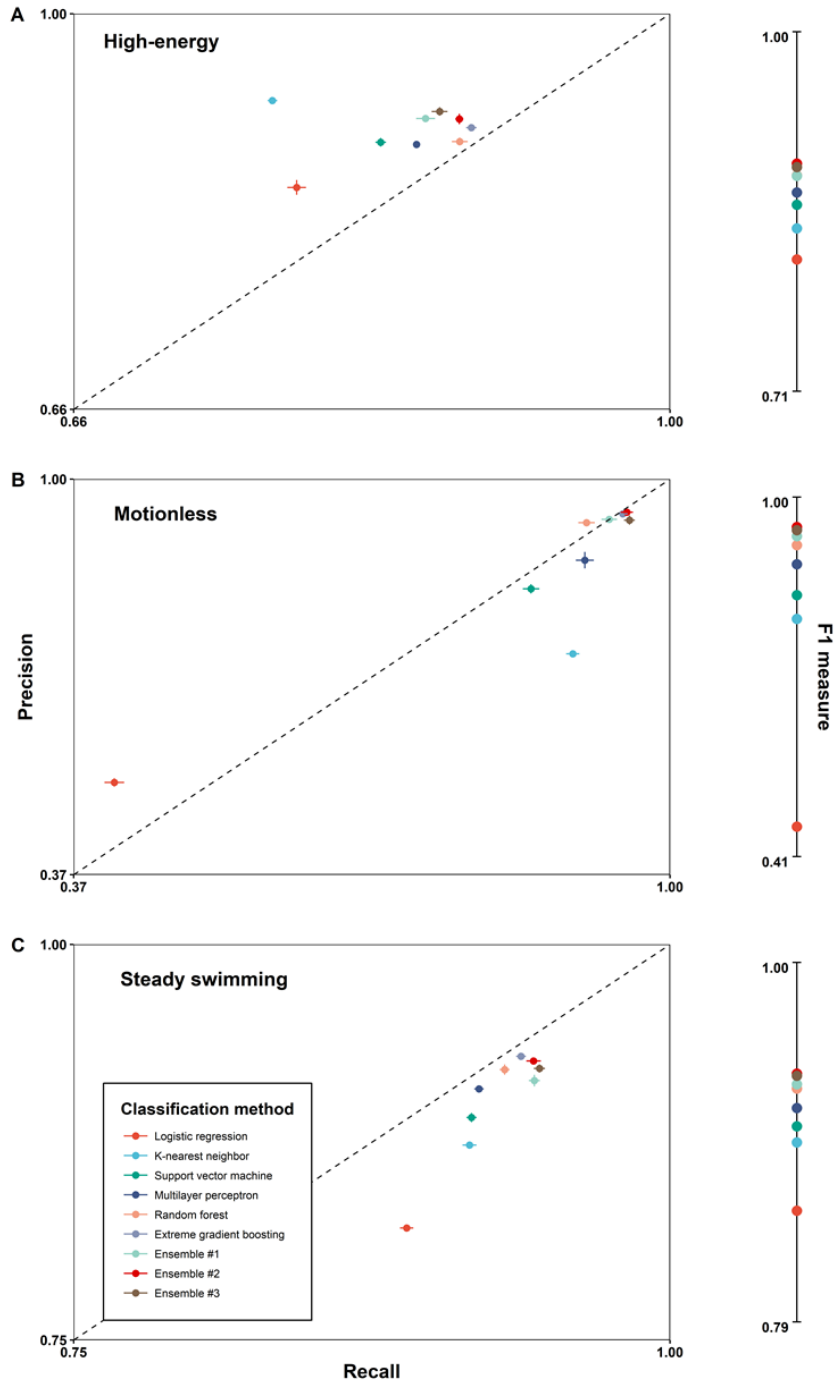


Figure B.2. Final performance measures (recall, precision, and F1 measures) within each behavior for optimized base learners and ensembling techniques. Measures are presented as the mean \pm standard deviation from five-fold cross-validation stratified across behavioral categories and majority down sampling of the steady swimming behavior.

BIOGRAPHICAL SKETCH

Ryan J. Knotek earned his Bachelors of Science degree in Marine Biology from the University of New England in May 2012. He then continued his education at the University of New England and received a Master's of Science degree in Marine Science in 2015. He subsequently enrolled in the Marine Science and Technology Doctorate program at the University of Massachusetts Boston in 2015. Throughout this program, he has received a Master's of Science degree in Marine Science and Technology (2019) and worked as a research technician at the Anderson Cabot Center for Ocean Life at the New England Aquarium.

Throughout Ryan's career as a marine scientist, his research has focused on merging cutting-edge technologies and techniques with modern analytics, to advance our understanding of how fishing impacts our world's sharks, skates, and ray populations. This work has had direct implications for management measures promoting species conservation, while also considering the continued success of the fishing industries that support the livelihoods of so many. Beyond the manuscripts and numbers, Ryan is a strong advocate for breaking down the communication barrier between the science and public, and is constantly exploring new ways of sharing his research across popular social media platforms. To this end, he finds education hugely important and part of his mission, as we mentor the next generation of stewards for our world's oceans.