

## ARTICLE

# Developing a predictive model to identify Sea Lamprey parasitism on Lake Trout using biologgers

Connor Reeve<sup>1</sup>  | Jean V. Adams<sup>2</sup> | Scott M. Miehls<sup>3</sup>  | Michael R. Lowe<sup>3</sup> | Steven J. Cooke<sup>1,4</sup>  | Mary Moser<sup>5</sup> | Jacob W. Brownscombe<sup>1,6</sup> 

<sup>1</sup>Department of Biology, Carleton University, Ottawa, Ontario, Canada

<sup>2</sup>U.S. Geological Survey, Great Lakes Science Center, Ann Arbor, Michigan, USA

<sup>3</sup>U.S. Geological Survey, Hammond Bay Biological Station, Millersburg, Michigan, USA

<sup>4</sup>Institute of Environmental and Interdisciplinary Science, Carleton University, Ottawa, Ontario, Canada

<sup>5</sup>National Oceanic and Atmospheric Administration, Northwest Fisheries Science Center, Seattle, Washington, USA

<sup>6</sup>Fisheries and Oceans Canada, Great Lakes Laboratory for Fisheries and Aquatic Sciences, Burlington, Ontario, Canada

## Correspondence

Connor Reeve

Email: [connorreve@cmail.carleton.ca](mailto:connorreve@cmail.carleton.ca)

## Abstract

**Objective:** Sea Lamprey *Petromyzon marinus* remain problematic for Lake Trout *Salvelinus namaycush* restoration in the Laurentian Great Lakes. Fisheries assessments would benefit from knowledge of spatial–temporal patterns of Sea Lamprey parasitism on Lake Trout; however, such patterns are challenging to estimate from wounding rates on caught Lake Trout. Electronic tags have been used to identify distinct fish behaviors (e.g., foraging or spawning) using measurements of acceleration or heart rate. We hypothesized that Sea Lamprey attachment would elicit changes in the heart rate and swimming behavior of Lake Trout. Here, we determined whether tagging devices could record these changes and whether we could accurately predict lamprey attachment on Lake Trout using these recordings.

**Methods:** Adult Lake Trout ( $n = 34$ ) were implanted with acceleration and heart rate tags and then were subjected to Sea Lamprey parasitism within a laboratory setting. Approximately 70 different acceleration and heart rate metrics were collected and tried as predictors of lamprey attachment. The top variables were used to train random forest models and then tried on test data sets. The accuracy of these models was then validated using a jackknife approach.

**Result:** Metrics related to body orientation and heart rate were identified as the best predictors of Sea Lamprey attachment. The best models predicted lamprey attachments with high accuracy; however, individual-level jackknife tests resulted in less accurate cross-individual prediction and regularly predicted false negatives. These findings may be related to individual variance in the Lake Trout response to attachment, but there was evidence that the shifting of tags after implantation impacted predictive performance, which could be remedied with adjustments during implantation.

**Conclusions:** Our study highlights the potential to use tagging devices for quantifying Sea Lamprey attachments on Lake Trout in the wild. Further development appears necessary; however, once improved, these predictive models have the potential to generate field-based estimates of Sea Lamprey attack rates on Lake Trout.

## KEYWORDS

biogger, machine learning, Lake Trout, Sea Lamprey

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Transactions of the American Fisheries Society* published by Wiley Periodicals LLC on behalf of American Fisheries Society.

## INTRODUCTION

Sea Lamprey *Petromyzon marinus* pose a threat to native fish restoration in the Laurentian Great Lakes. Their diets are wide ranging but primarily consist of coregonids, cactostomids, and salmonids in the Great Lakes (Harvey et al. 2008; Johnson et al. 2021); however, several studies have shown that Sea Lamprey will disproportionately parasitize Lake Trout *Salvelinus namaycush* (Pycha and King 1975; Swink 2003; Harvey et al. 2008). In fact, the Great Lakes once supported one of the largest Lake Trout fisheries, but this fishery collapsed in the 1950s due to overharvest and Sea Lamprey parasitism (Coble et al. 1990; Brant 2019). Various Sea Lamprey control measures were implemented shortly after the fishery collapsed and have been relatively effective at suppressing Sea Lamprey abundance across the Great Lakes (Smith and Tibbles 1980; Great Lakes Fishery Commission 2023). Given the lamprey control measures combined with native fish restoration programs, Lake Trout populations are recovering (He et al. 2012; Great Lakes Fishery Commission 2023). Today, the Great Lakes Sea Lamprey population is estimated to be roughly one-tenth of its peak abundance, which was observed in 1961 (Robinson et al. 2021), and both sport and commercial fisheries for Lake Trout exist in areas of the Great Lakes (Great Lakes Fishery Commission 2023). Despite these successes, Sea Lamprey continue to contribute significantly to Lake Trout mortality, hindering restoration efforts (Kornis et al. 2019). Managing Sea Lamprey-induced mortality remains among the most important factors for Lake Trout rehabilitation in the Great Lakes (Hansen et al. 1995; Ebener 1998; Kornis et al. 2019; He et al. 2020). Other fishes also benefit from Sea Lamprey control, and over time, this program has diversified to more broadly restore native fish populations.

To guide Lake Trout restoration efforts, fisheries managers rely on estimates of Sea Lamprey-induced mortality, which are derived most often from Sea Lamprey marking rates on collected Lake Trout (i.e., the prevalence of marks or wounds on fish resulting from Sea Lamprey attachment; Pycha 1980; Youngs 1980; Koonce 1987; Weeks 1997; Sitar et al. 1999; He et al. 2020). However, such estimates may be impacted by survivor detection bias, potentially underestimating both the lethality and number of Sea Lamprey attacks per fish (Adams et al. 2021). This bias could be corrected if the proportion of Lake Trout that survive attacks could be determined (Adams et al. 2021). When calculating Sea Lamprey-induced mortality rates, the probability of survival is typically assumed to be constant over time (Sitar et al. 1999; Ebener et al. 2005). If constant, marking rates can be a good index of Sea Lamprey-induced mortality (Adams et al. 2021); however, attack lethality

### Impact statement

There is a need for direct field observations of Sea Lamprey attack rates on Lake Trout to determine spatial-temporal patterns of parasitism and to help guide Sea Lamprey control and Lake Trout restoration efforts. Here, we determine whether measurements of heart rate and acceleration recorded from implantable tagging devices (biologgers) can be used to detect Sea Lamprey attachment on Lake Trout.

varies and is influenced by factors such as Sea Lamprey and Lake Trout sizes (Swink 2003), water temperature (McKee et al. 2004), and Lake Trout strain (Schneider et al. 1996). Although there is evidence to support the assumption of relatively constant lethality over longer periods of time (based on mortality rates predicted from Sea Lamprey marking rates in survey-collected Lake Trout; e.g., Pycha 1980; Schneider et al. 1996), there is little information on the annual variation in attack lethality, which could help to guide more immediate Sea Lamprey control decisions. Further, relying on marking rates is greatly hindered by the logistical constraints of contemporary sampling efforts and therefore offers limited insights into parasitism rates over space and time at the scale of the Great Lakes.

Direct measurement of Sea Lamprey parasitism and mortality on wild fishes is inherently challenging but may be aided by recent advances in biologging and biotelemetry technology that involve the use of various sensors (reviewed by Cooke et al. 2016). For example, triaxial accelerometer sensor tags have been used to examine fine-scale locomotor behaviors in free-swimming fish, thereby enabling the detection of foraging, burst swimming, or spawning events (e.g., Brownscombe et al. 2014; Brewster et al. 2018; Clarke et al. 2021). Electrocardiogram (ECG) sensors have been used to log the heart rates of fish in an effort to quantify physiological states, such as exhaustion (e.g., from exercise) and stress (e.g., Svendsen et al. 2021; Yousaf et al. 2022). In theory, these technologies may be useful for remote measurement of behavioral and/or physiological changes that occur in response to Sea Lamprey parasitism given the considerable physiological burden to host individuals. To detect these responses, calibrations must be conducted that collect behavioral observations in tandem with tag-borne measurements. Machine learning algorithms can then aid in identifying free-swimming individuals with and without parasitic lamprey attached. Additionally, one could determine the associated mortality from Sea Lamprey attacks in acoustic-tagged Lake

Trout, as it is relatively simple to infer mortality based on the absence of movement and/or flatline triaxial acceleration values (Villegas-Rios et al. 2020).

There is immense theoretical potential for electronic tags equipped with sensors to remotely identify fish parasitism by Sea Lamprey, greatly expanding our capacity to empirically quantify this phenomenon at larger spatial-temporal scales. Doing so would provide fisheries managers and Sea Lamprey control agents with important data to inform models and decisions. To this end, we conducted controlled, laboratory-based calibrations to generate predictive models for identifying behavioral (accelerometry) and physiological (heart rate) signatures of Sea Lamprey parasitism on Lake Trout. We hypothesized that given the physiological burden caused by Sea Lamprey attachment on Lake Trout (i.e., via injury, stress, and increased drag), Lake Trout would exhibit elevated heart rates and irregular swimming behavior during Sea Lamprey attachment, which could be subsequently recorded by using implantable biologging tags. We predicted that if such changes were recorded, machine learning classification trees could discriminate periods of Sea Lamprey attachment from periods when Sea Lamprey were not attached to tagged Lake Trout. The overall purpose of this research was to develop functional predictive models that would support future field deployments. Although biologgers were used in this study to seed models with abundant data, field deployments will ultimately require acoustic telemetry transmitter tags; thus, we consider those future requirements in our quantitative methods. Successful development of these predictive models would improve our ability to remotely measure Sea Lamprey parasitism in free-swimming Lake Trout.

## METHODS

### Species collection and tag implantation

Lake Trout were collected from Lake Huron during the fall of 2021 and 2022. The 2021 sample was collected by angling from spawning reefs near Hammond Bay, Michigan. The 2022 sample was collected from a Michigan Department of Natural Resources gill-net assessment survey (overnight sets; 1.8-m-high, 366-m-long, 10.2-cm-bar-mesh monofilament nets) on spawning reefs in Thunder Bay, Michigan. Parasitic-phase Sea Lamprey were collected each fall by commercial fishers on Lake Huron as incidental catch during fishing efforts for Lake Whitefish *Coregonus clupeaformis* and Lake Trout. Collected fish were transported in aerated coolers to the Hammond Bay Biological Station and were housed in flow-through tanks supplied with ambient-temperature

Lake Huron water. Lake Trout were fed 6.0-mm trout pellets (Skretting) at rates based on guidelines from BioMar (see [www.biomar.com/en/require/products-and-species](http://www.biomar.com/en/require/products-and-species) for details). To increase the probability of parasitic attachment, Sea Lamprey were not fed during holding.

Prior to experimentation, weight and length were recorded and each Lake Trout ( $3000 \pm 261$  g;  $716 \pm 19$  mm; all subsequent data are reported as mean  $\pm$  standard error, unless otherwise stated) was implanted with two different biologging tags: the Star-Oddi milli-HRT ACT heart rate and acceleration biologger (hereafter, heart rate and acceleration [HRA] tag) and the Technosmart Axy-5 XS accelerometer (hereafter, acceleration-only [AO] tag). For more details and tag recording specifics, see the respective manufacturer websites ([www.star-oddi.com](http://www.star-oddi.com) or [www.technosmart.eu/axy-5-xs/](http://www.technosmart.eu/axy-5-xs/)). Weight and length were also recorded for each Sea Lamprey ( $160 \pm 10$  g;  $459 \pm 9$  mm), and a 12-mm, half-duplex PIT tag was implanted to identify individuals. The AO tags were programmed to continuously record acceleration in three axes at 25 Hz. Because the Star-Oddi tags cannot be recharged and their batteries cannot be changed, a more conservative sampling approach was needed to avoid depleting the battery life too rapidly. In the first three trials ( $n = 19$ ), the HRA loggers were programmed to record the heart rate (at 125 Hz for 12 s), its associated ECG, and acceleration metrics (at 0.5 Hz for 60 s) every hour for the first week and then every 20 min until the trial finished. In the later two trials ( $n = 15$ ), the HRA tags were programmed to collect one recording of heart rate (at 125 Hz for 12 s) and acceleration metrics (at 5 Hz for 60 s) every 30 min. Because the quality of the heart rate recordings was good during the first three trials, ECGs were only saved for every fourth measurement (instead of every measurement) to corroborate heart rate values derived by the HRA tags in the later trials (i.e., to save tag battery and memory because HRA tags cannot be recharged).

For tag implantation, Lake Trout were electrocuted using a transcutaneous electrical nerve stimulation (TENS) unit (pulse width = 180  $\mu$ s; pulse rate = 120 Hz; 3–5 V; Model 7000 TENS unit, Roscoe Medical Inc.) while a constant flow of water was maintained over the gills for respiration during surgery (see Reid et al. 2019). Two incisions (both  $\sim 2.5$  cm in length) were made on the ventral side of the fish. Incisions for the HRA tags started on the midline directly between the pectoral fins and extended posteriorly. Incisions for the AO tags started about 5 cm anterior of the pelvic fins and were positioned just left of the midline. The HRA tag was inserted blunt side forward and was pushed anteriorly until the tag made contact with the pectoral girdle. Suturing thread (Oasis Medical PDO II monofilament sutures, size 2) was passed through the built-in anchor point on the HRA tag and then was used

to close the anterior end of the incision, thereby anchoring the tag to the body wall. A second suture closed the posterior end of the incision. Similarly, suturing thread was routed through an attached anchor point (a nylon washer fixed to the bottom of the AO tag as an attachment point) and was used to close the posterior end of the incision, anchoring the AO tag to the body wall. A second suture closed the anterior end of the incision. The same orientation was maintained for all tags. The combined weight of the HRA and AO tags was 15 g (HRA tag: 12 g, 39.5 × 13.0 mm [length × diameter]; AO tag: 3 g, 20 × 10 × 6 mm [length × width × height]), so the tag burden was minimal ( $0.49 \pm 0.04\%$  of body mass; maximum burden = 1.08%). The surgeons had previous training prior to the implantation of these devices. Future users should be aware of the impacts of tag implantation and surgeon experience on surgical recovery (for more details, see Cooke et al. 2003; Wagner et al. 2011; Smircich and Kelly 2014).

Fish collections were approved by the Michigan Department of Natural Resources (Fisheries Division Scientific Collector's Permit FSCP10312022203230), and experimental protocols involving the handling of fishes were carried out in accordance with U.S. federal guidelines for the care and use of animals and in accordance with American Fisheries Society guidelines (Use of Fishes in Research Committee 2014).

## Experiment

After surgery, Lake Trout were transferred into individual holding cages inside an oval-shaped artificial stream (internal channel width = 1.2 m; water depth = 1 m; linear distance following the channel centerline = 17.4 m; volume = 20,800 L; Frigid Units Inc.). The date and time of transfer were noted for later analyses. Experiments were conducted from December 2021 to March 2022 and from November to December 2022. The artificial stream was supplied with Lake Huron water at a rate of 25 L/min, which passed through a heat exchanger to maintain water temperatures at approximately 9°C (mean ± standard deviation [SD] =  $9.09 \pm 1.29^\circ\text{C}$ ). Water circulation was maintained using two 12-V, 13.6-kg-thrust trolling motors (Model R3; MotorGuide) positioned in the curved sections at either end of the stream. Baffles were placed in the channel upstream of the cages to produce a flow of water that was as linear as possible through the cages. Eight holding cages (1.1 m long; 1.2 m wide) were constructed by inserting blocker gates into the 4.5-m straight sections on either side of the stream. Blocker gates consisted of 1.27-cm square, vinyl-coated, wire-mesh panels that were attached to frames constructed of extruded aluminum

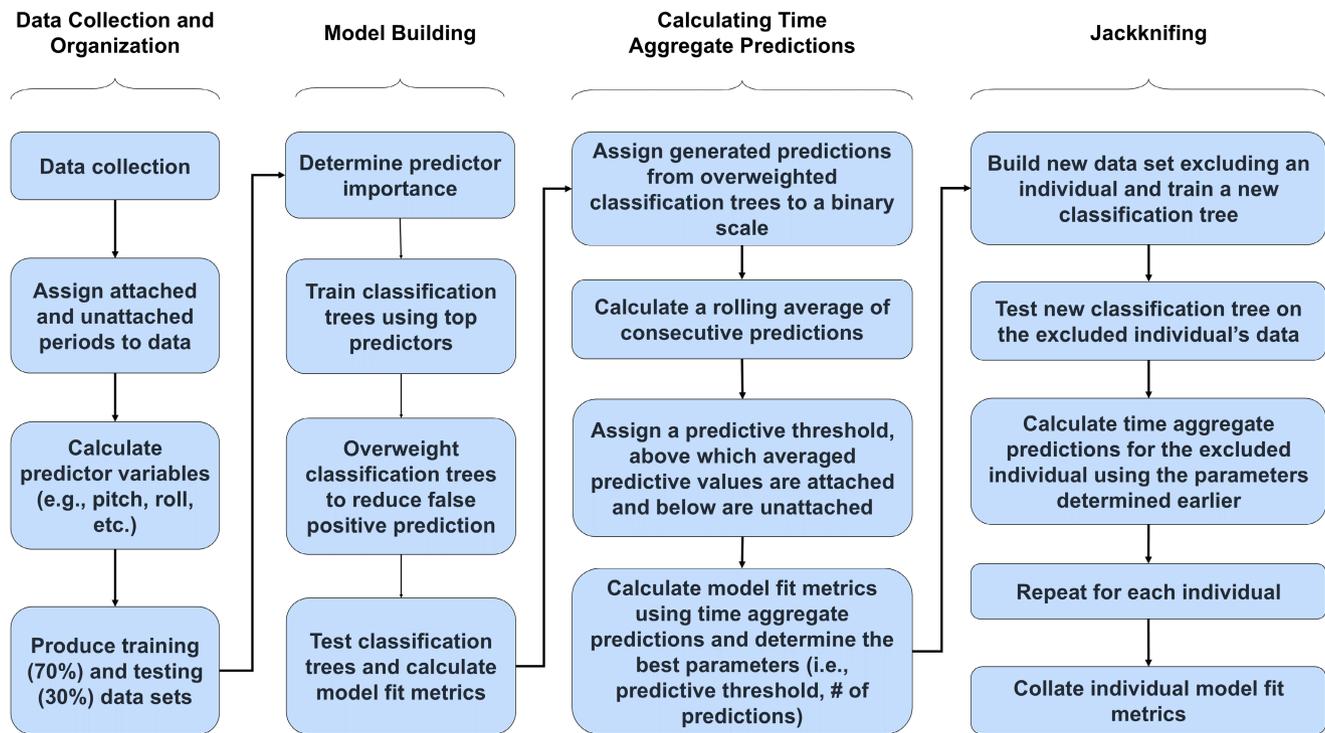
tube (2.54 × 2.54-cm, 80/20 T-slot). Frames were wrapped with garage door weather stripping to make a seal with the stream walls and were secured in place with wedges. Mesh was placed overhead to prevent fish from jumping out.

Once the Lake Trout were placed in their holding cages, they had 7–10 days to recover from surgery prior to the addition of Sea Lamprey. Heart rates require a minimum of 4 days to stabilize after surgical implantation in salmonids (Hvas et al. 2020; Zrini and Gamperl 2021). This time also served as a period to collect baseline swimming behaviors. Additionally, in the later two trials ( $n = 15$ ), a 1-min chase was performed using a blunt pole to stimulate burst swimming behavior. These Lake Trout had a minimum of 2 days to recover from the chase event prior to the addition of Sea Lamprey. Sea Lamprey were then added to each individual holding cage. Typically, only one lamprey was added to each cage, but in a few circumstances lamprey were added in an attempt to reduce the time to attachment ( $n = 7$ ). If the Sea Lamprey did not naturally attach to a Lake Trout after a given period of time (~2–3 days; time varied somewhat depending on time constraints), the Sea Lamprey were manually attached to the Lake Trout. This involved grasping the Sea Lamprey with a long grabbing pole and initiating contact between the lamprey and the Lake Trout ( $n = 16$  were manually attached). Similarly, if the Sea Lamprey did not naturally detach from the Lake Trout after a given period of time (~4 days; this period varied somewhat to ensure that there was still ample time to monitor the Lake Trout after detachment), the Sea Lamprey was manually removed from the Lake Trout ( $n = 12$  were manually removed). Additionally, active parasitism that persists for more than 5 days increases mortality in Lake Trout (Smith et al. 2016). After detachment, Sea Lamprey were removed from the cage and weighed, and Lake Trout were monitored for a minimum of 3 days unless mortality occurred. After experimentation, Lake Trout were euthanized and the biologgers were removed.

Video of the Lake Trout was recorded continuously during the experiment using overhead, wide-angle infrared cameras (Axis Model Q1604; one camera for each holding cage) mounted 0.76 m above the water surface. Red-light strips mounted behind the cameras permitted 24-h video collection. The stream was covered with a large tarp to reduce overall light levels, but several small gaps at regular spacing permitted a small amount of ambient light to enter during daytime hours.

## Data collection and organization

Figure 1 highlights the main steps involved in data analysis. Some tags were corrupted during experimentation



**FIGURE 1** Simplified flow diagram of the steps between data collection and the model fit metrics calculated for each individual Lake Trout's time aggregate predictions.

(7 AO tags and 2 HRA tags) and several Lake Trout died ( $n=9$ , presumably related to injury sustained from Sea Lamprey attacks) or were removed prematurely ( $n=3$ , removed due to fungal infection), resulting in a loss of data or proportionally less data collected from these individuals. However, any data collected from these fish prior to data corruption or mortality/removal were still included and used for model building. Two AO tags moved within the body cavity throughout the experiment to a degree that could not be corrected for (see correction methods below) and were subsequently removed from analyses. Data from the first day were excluded for all fish to mitigate the effects of surgery on their behavior. In the final trial, eight AO tags were not correctly initiated prior to implantation; therefore, no AO measurements were collected from that trial. In total, data were collected from 34 Lake Trout, 28 of which experienced Sea Lamprey attachment. The HRA tags collected data from 33 fish and specifically recorded Sea Lamprey attachment data in 28 fish, whereas the AO tags collected data from 23 fish and specifically recorded Sea Lamprey attachment data in 14 fish.

Video recordings were observed to note the timing of Sea Lamprey attachment and detachment for each fish. These observations were used to assign HRA recordings to periods of either Sea Lamprey attachment or unattachment. Other important observations included when a Lake Trout lost equilibrium and died, when the simulated

chase occurred (in the later trials), and when Lake Trout were removed from the artificial stream.

Prior to data analysis and model building, the quality of the HRA tag heart rate data was assessed using Star-Oddi's quality index (QI) metric and through visual observation of saved ECGs. The QI is determined by a built-in tag algorithm and represents the quality of the ECG signal. High-quality signals are denoted by a QI of 0, whereas QI values of 1, 2, and 3 indicate decreasing quality. Most recordings were of high quality (76% of recordings corresponded to QI=0). If a given heart rate was assigned a QI value greater than 1 or if the heart rate appeared irregular, the associated ECG was observed in the Star-Oddi HRT Analyzer (see <https://www.star-oddi.com/support/software> for details). In these cases, ECGs typically suffered from noise, but the R-R interval (i.e., the interval between two successive R-peaks of the QRS waveform, the characteristic waveform that is produced from the depolarization of ventricles) was often observable and the heart rate could be manually calculated to reduce data loss. Heart rates that could not be corroborated by or manually calculated from saved ECGs were removed from the data set (<1%).

In addition to the heart rate recordings, metrics of acceleration and orientation were also collected or calculated to be tried in predictive models. Metrics of acceleration collected from the HRA tags included static acceleration

in the X-, Y-, and Z-axes, total acceleration (the sum of absolute static acceleration from all three axes), external acceleration (an on-board tag measurement of acceleration; average, maximum, and minimum values were calculated over the 60-s recording interval), variation in external acceleration, and tilt data (HRA tags recorded X-tilt, Y-tilt, and Z-tilt in degrees). Metrics of acceleration collected from the AO tags included static acceleration in the X-, Y-, and Z-axes; dynamic acceleration in the X-, Y-, and Z-axes (static acceleration and dynamic acceleration were separated using weighted smoothing at an interval of 2s; see Brownscombe et al. 2013); and overall dynamic body acceleration (the absolute sum of the dynamic acceleration from all three axes; Gleiss et al. 2011). Static acceleration and tilt values were corrected for differences in tag orientation among individuals using two different methods. The first method used average static X, Y, and Z acceleration values over the first 3 days, and the second method used 24-h rolling averages of static X, Y, and Z acceleration to correct for differences in tag orientation. This second method was tried to correct for tag movement that occurred over time within the body cavity. Pitch and roll were also calculated from normalized static acceleration for both HRA and AO tags using the following formulas (Pedley 2013):

$$\text{Pitch} = \text{atan2}(Y_{\text{static}}/Z_{\text{static}}) \times 180 \div \pi,$$

$$\text{Roll} = \text{atan2}\left[-X_{\text{static}}/\sqrt{(Y_{\text{static}}^2/Z_{\text{static}}^2)}\right] \times 180 \div \pi.$$

Since future field deployments will require the use of acoustic transmitter tags (i.e., to retrieve data without tag recollection), the recorded and calculated variables from both the HRA and AO data sets were averaged to the minute to simulate the optimal data resolution collected from R-code acoustic receivers, which comprise the most widely distributed receiver infrastructure in the Great Lakes (see <https://glatos.glos.us/>; Heupel et al. 2006). Additionally, variance was calculated for tilt, total acceleration, and static acceleration in all three axes for the HRA data set, and variance was calculated for all AO variables. Absolute values were calculated for all acceleration- and orientation-specific metrics for both tag types. These recordings and subsequent calculations resulted in the formulation of 71 variables in the HRA data set and 66 variables in the AO data set that could be used in predictive model building.

## Model building

Several models were produced to predict Sea Lamprey attachment for the two tag types used; “HRA models”

refers to those models informed by HRA data, and “AO models” refers to those informed by AO data. Because acoustic transmitters can currently only record and transmit a limited number of variables, we fitted models with a reduced set of predictor variables (InnovaSea acoustic transmitters can transmit a maximum of four separate variables; InnovaSea, personal communication). Each data set was randomly split into training and test data sets (split using individual data; 70% training and 30% testing for each individual fish). Training data were then fitted with random forest algorithms. Random forests are an extension of classification and regression trees, which create a hierarchical series of binary partitions in the data using the predictors to optimize prediction of the response (Breiman 2001). Random forests fit a series of trees with random subsets of data and predictors and then aggregate the predictions, often improving prediction accuracy and reducing model overfitting. The dependent variable in all models was Sea Lamprey attachment (yes or no; i.e., the Sea Lamprey was or was not attached to the Lake Trout). Using the complete data sets, top predictors were selected by calculating the mean decrease in accuracy (MDA; the percent decrease in model accuracy in trees when the variable was not included) after fitting classification trees with 1000 trees and trying all variables at each split at the square root of the number of predictors. If any of the top-four predictors were derivative of each other (e.g., Z static and absolute Z static), the variable with the next highest MDA was selected.

In both the HRA and AO models, static X, Y, and Z acceleration as well as roll and pitch were among the top predictors. Although there is a limited number of variables that can be transmitted by modern acoustic transmitters, if the static X, Y, and Z acceleration variables (among the top predictors) are transmitted, the roll and pitch can be calculated secondarily (see formulas above). This would allow for calculation of additional predictors while limiting the number of variables that would need to be transmitted. With this in mind, we trained and tested several different random forest classification trees: (1) the selected model, which used the top-four predictors selected from calculation of the MDA; (2) the triaxial model, which used static triaxial acceleration (and heart rate values in the HRA models) and secondarily calculated the pitch and roll; and (3) the absolute model, which used absolute static triaxial acceleration values (and heart rate values in the HRA models) and secondarily calculated the absolute pitch and roll. Absolute values were calculated and tested with the idea that absolute values would consider right (positive X acceleration) and left (negative X acceleration) or up (negative Y acceleration) and down (positive Y acceleration) movement alike within a model, potentially reducing the influence of variance in Sea Lamprey attachment location.

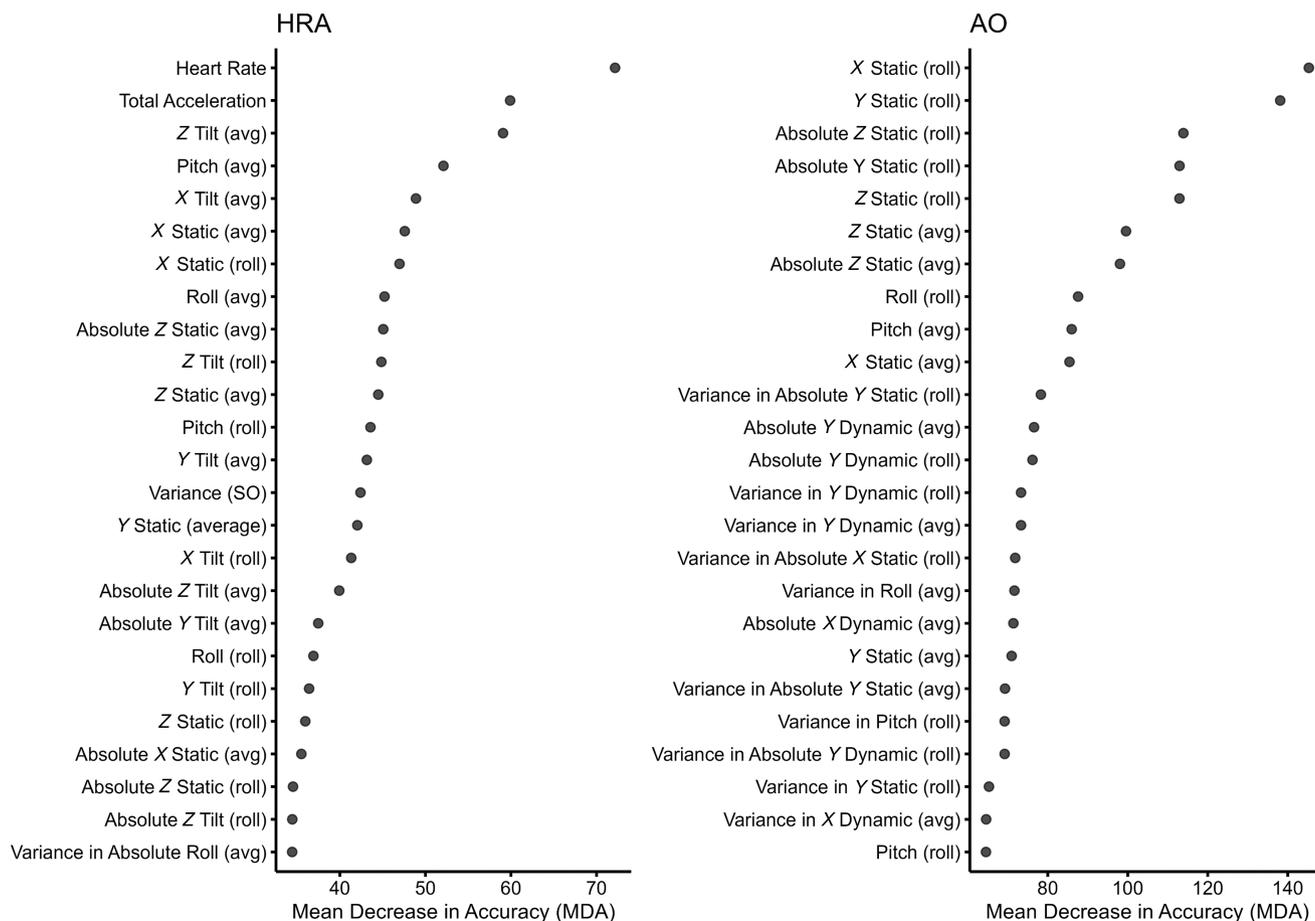
All predictors used were continuous. It should be noted that based on the MDA, the first method of correction (i.e., using the average values over the first 3 days to correct orientation) appeared to function better for HRA data and the second method of correction (i.e., using a 24-h rolling average to correct orientation) was better suited for the AO data (see Figure 2); therefore, these corrected variables were used in the calculation of triaxial and absolute models, respectively.

All tree models (i.e., classification trees) were fitted with 1000 trees, and the default number of variables was tried at each split at the square root of the number of predictors. Models were overweighted (weighted 1:100 for attached : unattached; see Brownscombe et al. 2021) to penalize misclassifications of unattached values (i.e., data that corresponded to unattached periods). Overweighting is not common practice; however, this was done to reduce the likelihood of a type I error (false positive) when predicting attachment and was accomplished by iteratively testing the weighting combinations

in training data. Additionally, we must consider what the class proportions might be if our models were applied in a field setting, where Sea Lamprey attachments would be much less frequent. In general, overweighting models marginally improved model sensitivity (~1–2% increase) and reduced model specificity (~1% decrease; see Table S1). Model fit metrics were then calculated by predicting onto test data (see Table S2; Brownscombe et al. 2021).

### Calculating time aggregate predictions

A secondary method was used to improve the predictive performance of HRA and AO models; this method involved smoothing consecutive predictions generated from their random forest classification trees. Predicted attached and unattached values were assigned to a binary scale, with 1 representing predicted attached values and 0 representing predicted unattached values. Binary predictions



**FIGURE 2** Predictor variables used in producing the heart rate and acceleration (HRA; left) and acceleration-only (AO; right) models and their mean decrease in accuracy (MDA) calculated after fitting random forest classification trees with 1000 trees and trying all variables. The 25 variables with the greatest values of MDA are shown. Within parentheses, “avg” refers to values corrected using their respective average of the first 3 days of recording, “roll” refers to values corrected using their respective 24-h rolling average, and “SO” refers to values that were internally calculated within Star-Oddi tags.

were then smoothed by using a rolling average of consecutive predictions. In theory, averaged predictions closer to 1 should indicate higher support for true-positive predictions. Therefore, a predictive threshold could then be set, above which the smoothed predicted values could more accurately predict attachment (Figure 3 provides a visual demonstration of this method). Hereafter, these smoothed predicted values will be referred to as time aggregate predictions (TAPs).

We assessed the relationships between the predictive threshold (between 0.00 and 0.99) and the number of consecutive predictions included (between 5 and 100), and secondarily calculated model fit metrics (i.e., levels of sensitivity and specificity and positive and negative predictive values) from TAPs. In doing so, we could choose appropriate predictive parameters to use in the calculation of TAPs during model validation (see below). Several data sets were produced in which the number of consecutive predictions considered in the rolling average differed. In total, 20 different data sets were produced using rolling averages of consecutive predictions in intervals of 5 up to a maximum of 100 consecutive predictions (e.g., one data set was produced using rolling averages for every five consecutive predictions, another was produced using rolling averages for every 10 consecutive predictions, and so on). Levels of sensitivity and specificity and positive and negative predictive values were calculated using the TAPs associated with each predictive threshold between 0 and 1 for every interval of 0.01 (i.e., 100 calculations per data set) and for each of the 20 data sets produced from the rolling averages of consecutive predictions. All data sets were then combined. To determine the optimal predictive threshold and the number of averaged consecutive predictions that should be used for best predictive performance, we calculated a balanced predictive value by multiplying all four secondarily calculated model fit metrics (Table S2). These metrics were multiplied to penalize low fit values more heavily. Higher balanced predictive values should indicate an overall greater ability to accurately predict periods of Sea Lamprey attachment and unattachment with high precision and low rates of false positives.

## Model validation

To validate the models produced (i.e., the HRA and AO selected, triaxial, and absolute models), a jackknife approach was used. This involved excluding an individual fish's data from the total data set and producing a new classification tree (again fitted with 1000 trees and over-weighted as described earlier) using the predictive variables corresponding to the model being tested, which was

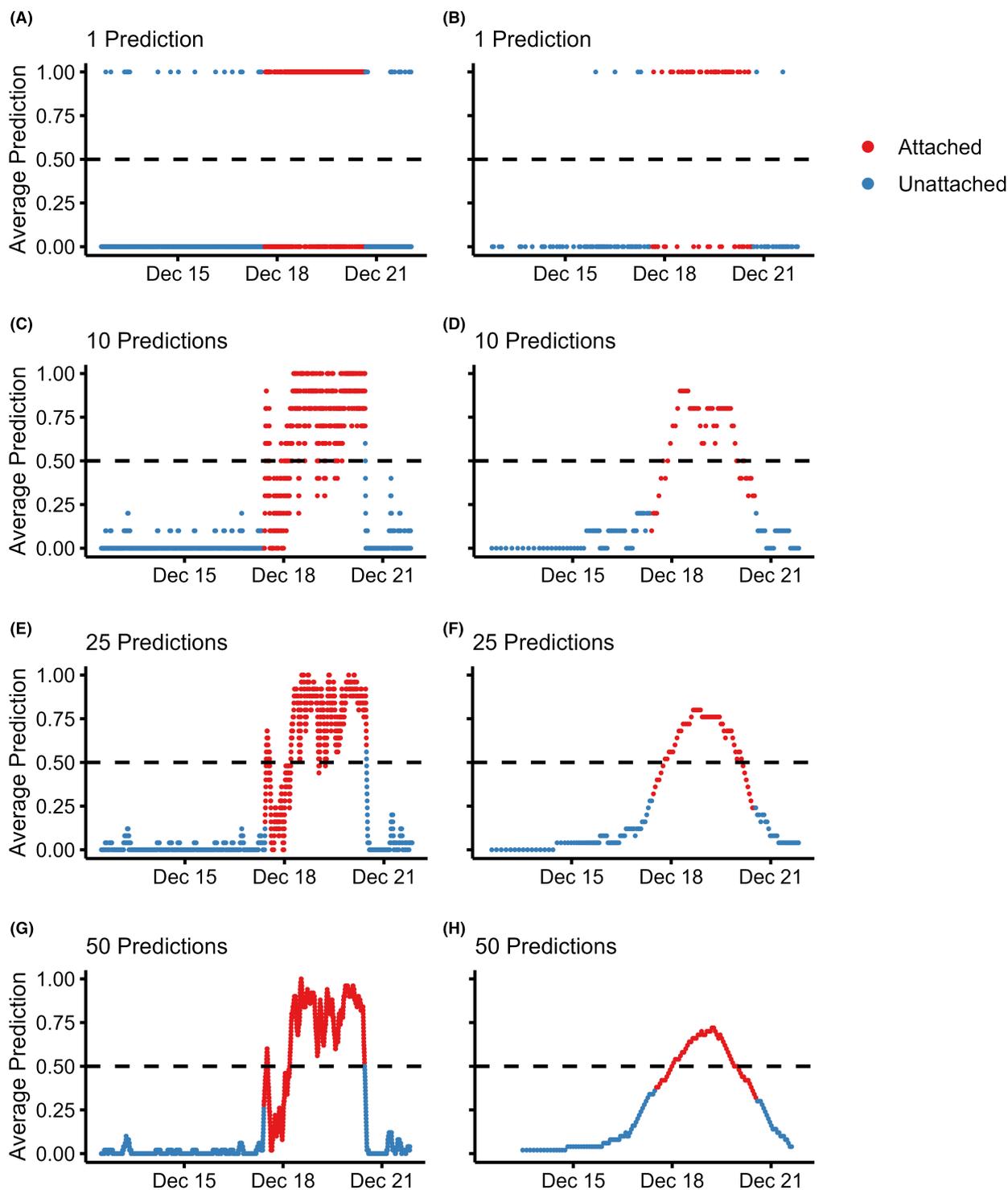
then tested on the individual's data. The TAPs were calculated from their classification tree predictive outputs. The balanced predictive values calculated earlier were used to identify a conservative number of consecutive predictions and a predictive threshold to be used for individual TAP calculations. A predictive threshold of 0.25 and a rolling average of 40 consecutive predictions were used to calculate TAPs, as this approach appeared to be somewhat conservative for all models and limited the amount of data required for detection (see Supplemental Information for a more detailed rationale; Figure S1; Tables S2 and S3 [all available in the online version of this article]). Model fit metrics were then calculated by comparing TAPs to observed values. This was repeated for each individual and for all models. Model fit metrics from individuals were then collated, and the average specificity and sensitivity, average positive and negative predictive values, average accuracy, and total accuracy ([the sum of correct predictions from all individuals]/[the total number of predictions from all individuals]) were calculated.

These collated, post-jackknife model fit metrics were then compared between the different predictive models (i.e., selected, triaxial, and absolute) to determine the best model. Partial dependencies were calculated for all predictors used in the best HRA and AO models to examine relationships between predictors and Sea Lamprey attachment. To investigate sources of error on positive detection accuracy (i.e., positive predictive value) in the final models, we performed simple linear regression analyses. This was simply exploratory, as we lacked power to appropriately assess these relationships. Sources of error that were investigated included the tag orientation (determined by averaging the uncorrected static X, Y, and Z acceleration values prior to Sea Lamprey attachment), Lake Trout weight, Sea Lamprey weight, the interaction between Lake Trout weight and Sea Lamprey weight, whether or not the Lake Trout died, whether or not the Sea Lamprey was manually attached, and the location of the Sea Lamprey attachment (head, side, ventral, or dorsal).

All analyses were conducted in RStudio (R version 4.2.1; R Core Team 2019). Random forest classification trees were produced using the randomForests package (Liaw 2022). Model fit metrics and partial dependencies were calculated using the caret (Kuhn 2023) and pdp (Greenwell 2022) packages, respectively.

## RESULTS

Over the five trials, 39,576 min of data were recorded by HRA tags ( $n = 33$ ), 8236 min of which included lamprey



**FIGURE 3** This figure highlights how time aggregate predictions (TAPs) were calculated using different predictive thresholds and consecutive predictions. Panels A, C, E, and G include predictions that were generated using the absolute acceleration-only (AO) model, and panels B, D, F, and H include predictions that were generated using the triaxial heart rate and acceleration (HRA) model. The y-axis (average prediction) refers to the average binary predictive value calculated using the rolling average of consecutive predictions. The number of consecutive predictions used in this calculation is indicated in the title of each panel. Red points represent periods when the Sea Lamprey was attached to the Lake Trout, whereas blue points represent periods when the Sea Lamprey was not attached. Here, we show how a predictive threshold of 0.50 (dashed black line) influences the accuracy in TAPs; the TAPs above this threshold would be considered to indicate attachment, and the TAPs below this threshold would be considered to indicate nonattachment. One can observe how increasing the number of consecutive predictions reduces the likelihood for false positives (i.e., fewer blue points above the threshold) and generally increases model accuracy. Additionally, one could envision how shifting this threshold could alter model precision and rates of false positives. Using a threshold of 0.50, all or most values greater than this threshold represent true attachments; however, some true attachments may have been missed.

attachment; 467,144 min of data were recorded by AO tags ( $n=25$ ), 67,692 min of which included lamprey attachment. Sea Lamprey remained attached to Lake Trout for an average of  $2.6 \pm 0.3$  days. Most Sea Lamprey actively fed while attached, but four individuals were nonfeeding. Sea Lamprey that actively fed gained  $29.6 \pm 4.0$  g during the trials, and those that did not feed lost  $4.9 \pm 1.8$  g during the trials.

Calculations of the MDA indicated that the heart rate, total acceleration, Z-tilt, and pitch were the top predictors of Sea Lamprey attachment when using HRA data and that static X, Y, and Z acceleration and roll (after several derivatives of static Y and Z acceleration; see Figure 2) were the top predictors of Sea Lamprey

attachment when using AO data. Classification trees that used MDA-selected variables (i.e., the selected HRA and AO models) had the highest model fit metrics prior to jackknife testing and TAP calculation but had the lowest fit metrics after jackknife testing, which suggested that overfitting may have inflated their performance (see Tables 1–3; see Figure S2 for an example classification tree).

Calculation of TAPs for all three models demonstrated similar trends (see Figures 6 and 7; data shown for HRA triaxial and AO absolute models). Because predicted attachments occurred in greater densities during periods of true Sea Lamprey attachment (e.g., see Figure 3), TAPs improved model precision and specificity (provided that

**TABLE 1** Model fit metrics from random forest classification trees in the test data (prior to jackknife testing and producing time aggregate predictions [TAPs]). Model fit metrics shown in bold correspond to the best models after jackknife testing and TAP calculation. AO, acceleration only; HRA, heart rate and acceleration.

Metric	HRA selected	HRA triaxial	HRA absolute	AO selected	AO triaxial	AO absolute
Number of variables	4	<b>6</b>	6	4	5	<b>5</b>
Accuracy	0.872	<b>0.858</b>	0.847	0.895	0.830	<b>0.823</b>
No information rate	0.790	<b>0.787</b>	0.787	0.817	0.817	<b>0.816</b>
Kappa	0.576	<b>0.518</b>	0.461	0.605	0.254	<b>0.195</b>
Sensitivity	0.575	<b>0.503</b>	0.432	0.568	0.229	<b>0.173</b>
Specificity	0.951	<b>0.954</b>	0.959	0.968	0.964	<b>0.969</b>
Positive predictive value	0.756	<b>0.746</b>	0.741	0.801	0.591	<b>0.563</b>
Negative predictive value	0.893	<b>0.877</b>	0.862	0.909	0.848	<b>0.838</b>
Prevalence	0.211	<b>0.213</b>	0.213	0.183	0.183	<b>0.183</b>
Detection rate	0.121	<b>0.107</b>	0.092	0.104	0.042	<b>0.032</b>
Detection prevalence	0.160	<b>0.143</b>	0.125	0.130	0.071	<b>0.057</b>
Balanced accuracy	0.763	<b>0.728</b>	0.696	0.768	0.597	<b>0.572</b>

**TABLE 2** Model fit metrics after jackknife testing without time aggregate predictions (TAPs). Model fit metrics shown in bold correspond to the best models following jackknife and TAP calculation. Values are presented as mean  $\pm$  standard error. AO, acceleration only; HRA, heart rate and acceleration.

Metric	HRA selected	HRA triaxial	HRA absolute	AO selected	AO triaxial	AO absolute
Number of variables	4	<b>6</b>	6	4	5	<b>5</b>
Total accuracy	0.691	<b>0.718</b>	0.722	0.751	0.794	<b>0.804</b>
Average accuracy	$0.695 \pm 0.032$	<b><math>0.725 \pm 0.032</math></b>	$0.723 \pm 0.027$	$0.805 \pm 0.024$	$0.831 \pm 0.022$	<b><math>0.845 \pm 0.022</math></b>
Average sensitivity	$0.238 \pm 0.040$	<b><math>0.291 \pm 0.052</math></b>	$0.262 \pm 0.043$	$0.336 \pm 0.052$	$0.186 \pm 0.049$	<b><math>0.181 \pm 0.051</math></b>
Average specificity	$0.797 \pm 0.042$	<b><math>0.806 \pm 0.047</math></b>	$0.834 \pm 0.033$	$0.877 \pm 0.024$	$0.933 \pm 0.010$	<b><math>0.952 \pm 0.006</math></b>
Average positive predictive value	$0.366 \pm 0.046$	<b><math>0.378 \pm 0.053</math></b>	$0.358 \pm 0.055$	$0.404 \pm 0.057$	$0.387 \pm 0.054$	<b><math>0.404 \pm 0.054</math></b>
Average negative predictive value	$0.820 \pm 0.042$	<b><math>0.837 \pm 0.027</math></b>	$0.835 \pm 0.026$	$0.887 \pm 0.024$	$0.880 \pm 0.023$	<b><math>0.879 \pm 0.023</math></b>

**TABLE 3** Model fit metrics from time aggregate predictions (TAPs) after jackknife testing. Model fit metrics shown in bold correspond to the best models after jackknife and TAP calculation. Values are presented as mean  $\pm$  standard error. AO, acceleration only; HRA, heart rate and acceleration.

Metric	HRA selected	HRA triaxial	HRA absolute	AO selected	AO triaxial	AO absolute
Number of variables	4	<b>6</b>	6	4	5	<b>5</b>
Total accuracy	0.664	<b>0.712</b>	0.706	0.706	0.810	<b>0.821</b>
Average accuracy	0.657 $\pm$ 0.058	<b>0.728 <math>\pm</math> 0.043</b>	0.710 $\pm$ 0.041	0.771 $\pm$ 0.039	0.845 $\pm$ 0.024	<b>0.864 <math>\pm</math> 0.022</b>
Average sensitivity	0.326 $\pm$ 0.067	<b>0.417 <math>\pm</math> 0.080</b>	0.394 $\pm$ 0.071	0.411 $\pm$ 0.062	0.201 $\pm$ 0.059	<b>0.206 <math>\pm</math> 0.059</b>
Average specificity	0.733 $\pm$ 0.066	<b>0.779 <math>\pm</math> 0.060</b>	0.784 $\pm$ 0.053	0.819 $\pm$ 0.043	0.949 $\pm$ 0.015	<b>0.971 <math>\pm</math> 0.008</b>
Average positive predictive value	0.340 $\pm$ 0.064	<b>0.470 <math>\pm</math> 0.070</b>	0.452 $\pm$ 0.081	0.456 $\pm$ 0.064	0.501 $\pm$ 0.059	<b>0.588 <math>\pm</math> 0.062</b>
Average negative predictive value	0.812 $\pm$ 0.038	<b>0.855 <math>\pm</math> 0.038</b>	0.856 $\pm$ 0.025	0.894 $\pm$ 0.022	0.884 $\pm$ 0.023	<b>0.884 <math>\pm</math> 0.022</b>

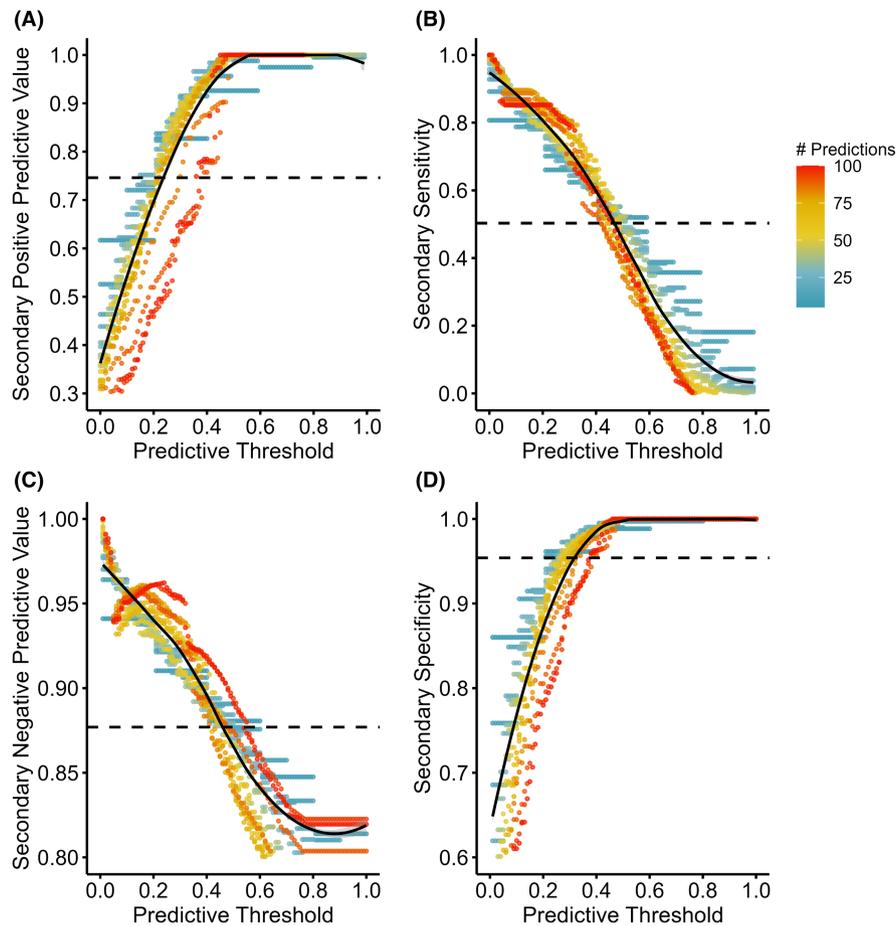
an appropriate predictive threshold was used; i.e., between 0.1 and 0.4) even when using few consecutive predictions (more details in the [Supplemental Information](#)). Generally, increasing the predictive threshold improved both the positive predictive value and the specificity, which were further improved by increasing the number of consecutive predictions included in TAPs. The TAPs improved both models' negative predictive value and sensitivity at low to moderate predictive thresholds; however, at higher predictive thresholds, these metrics were impaired (see [Figures 4B,D](#) and [5B,D](#)). This suggests that a greater number of true attachments may be missed at higher predictive thresholds.

Although these trends still existed after jackknife testing, the predictive models exhibited relatively poor performance when predicting periods of Sea Lamprey attachment in individuals even after TAPs were calculated (e.g., see [Tables 2](#) and [3](#)). Triaxial and absolute models outperformed the selected models after jackknife and TAP calculations. The triaxial model produced the best postjackknife fit metrics among the HRA models tested, whereas the absolute model produced the best fit metrics among the AO models tested. However, differences in model fit metrics between triaxial and absolute models were relatively small ([Table 3](#)). In general, both the triaxial HRA model and the absolute AO model were relatively accurate (triaxial HRA model: average = 0.728  $\pm$  0.043, total accuracy = 0.712; absolute AO model: average = 0.864  $\pm$  0.022, total accuracy = 0.821; see [Table 3](#)) and would regularly predict unattached periods correctly (triaxial HRA model: average negative predictive value = 0.855  $\pm$  0.038; absolute AO model: average negative predictive value = 0.884  $\pm$  0.022). However, these models frequently predicted false negatives (triaxial HRA model: average sensitivity = 0.417  $\pm$  0.080; absolute AO model: average sensitivity = 0.206  $\pm$  0.059) and lacked some precision in predicting periods of Sea Lamprey

attachment (triaxial HRA model: average positive predictive value = 0.470  $\pm$  0.070; absolute AO model: average positive predictive value = 0.588  $\pm$  0.062).

Partial dependencies were calculated for all predictors used in the triaxial HRA and absolute AO models and showed how Lake Trout body orientation and heart rate responded to Sea Lamprey attachment ([Figure 6](#)). Sea Lamprey attachments were most often associated with elevated heart rates, negative values of static *X* and *Y* acceleration, low (i.e., <1 g) to negative values of static *Z* acceleration, and relatively high positive or negative values of pitch and roll (roughly  $-50^\circ$  or  $+50^\circ$  for pitch and roughly  $-100^\circ$  or  $+100^\circ$  for roll, although the negative values had a stronger effect) when using the triaxial HRA model and were associated with absolute values of *Z* static acceleration less than 1 g (and, to a lesser degree, <0.6 g) and high absolute values of static *X* and *Y* acceleration as well as pitch and roll (see [Figure 7](#); [Figure S4](#)).

There was considerable variation in individual responses—particularly the acceleration responses—to Sea Lamprey attachment (see [Figure 8](#); [Figure S5](#)). Changes in the heart rate were more consistent across individuals, with most individuals demonstrating elevated heart rates during periods of attachment. Heart rates on average increased by 6.54  $\pm$  1.30 beats per minute (bpm) during Sea Lamprey attachment (i.e., from 36.76  $\pm$  2.30 bpm to 43.29  $\pm$  1.92 bpm). Interestingly, Lake Trout that experienced nonfeeding attachments did not exhibit elevated heart rates. The average heart rates during periods of nonfeeding Sea Lamprey attachments and when Sea Lamprey were not attached were 48.80  $\pm$  8.45 bpm and 52.11  $\pm$  5.75 bpm, respectively; the average heart rates during periods of feeding Sea Lamprey attachments and when Sea Lamprey were not attached were 42.63  $\pm$  1.94 bpm and 34.91  $\pm$  2.24 bpm, respectively. However, heart rate responses to nonfeeding attachments were only recorded in three Lake Trout (one



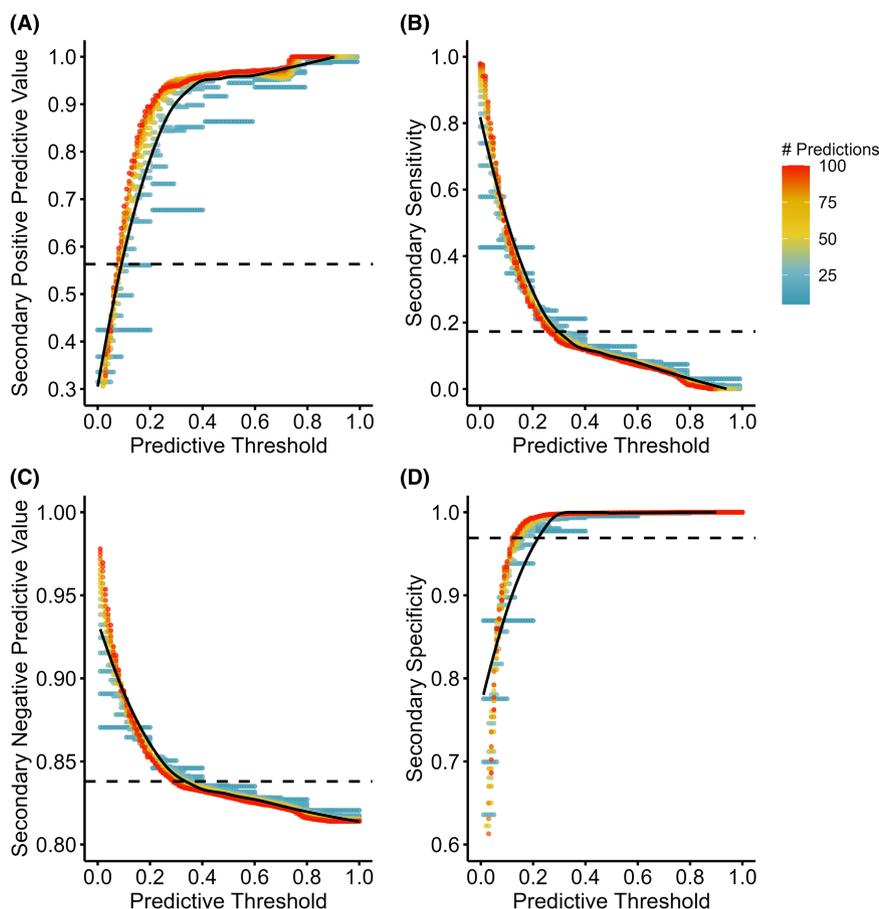
**FIGURE 4** Panels highlight how changes to the predictive threshold and the number of included consecutive predictions influence different model fit metrics using time aggregate predictions (TAPs) generated from the triaxial heart rate and acceleration model. Secondarily calculated model fit metrics include (A) positive predictive values, (B) levels of sensitivity, (C) negative predictive values, and (D) levels of specificity. The dashed black lines represent the original model fit metric (i.e., prior to calculating the TAPs). “# Predictions” refers to the number of consecutive predictions that were included in TAPs for each model fit metric calculation. The number of consecutive predictions averaged ranges from 5 to 100 (in intervals of 5), represented by colors from blue to red, respectively. Solid black lines show the data in each panel fitted with a locally estimated scatterplot smoother. For details on the model fit metric calculations, see [Table 1](#) and *Data Analysis and Model Building*.

nonfeeding attachment only lasted about 15 min and occurred in between heart rate and acceleration sampling intervals).

Although we lacked power, linear regressions suggested that there was a relationship between tag orientation and positive detection accuracy. There was a negative relationship between uncorrected static  $Y$  acceleration and positive detection accuracy in the triaxial HRA model ( $R^2=0.22$ ), and there was a negative relationship between uncorrected static  $X$  acceleration and positive detection accuracy ( $R^2=0.45$ ) in the absolute AO model, suggesting that tags that shifted or were implanted at an angle reduced positive detection accuracy. There was also a relationship between Sea Lamprey weight and positive prediction accuracy in the triaxial HRA model ( $R^2=0.19$ ), suggesting that Sea Lamprey weight may have impacted the model's ability to accurately detect Sea Lamprey attachment.

## DISCUSSION

The capacity to remotely measure Sea Lamprey attack rates on fishes would dramatically improve our ability to manage their populations and fisheries over space and time. Our findings here suggest that Lake Trout exhibit behavioral and physiological signatures that can be measured and used to identify Sea Lamprey attachments by using predictive models. Predictor selection methods found that variables related to body orientation (i.e., static  $X$ ,  $Y$ , and  $Z$  acceleration, as well as pitch and roll) were among the most important predictors of Sea Lamprey attachment in both models, indicating that Sea Lamprey attachment impacts Lake Trout body orientation during swimming. Heart rate was also a top predictor in the triaxial HRA model, with Lake Trout typically displaying elevated heart rates during periods of Sea Lamprey



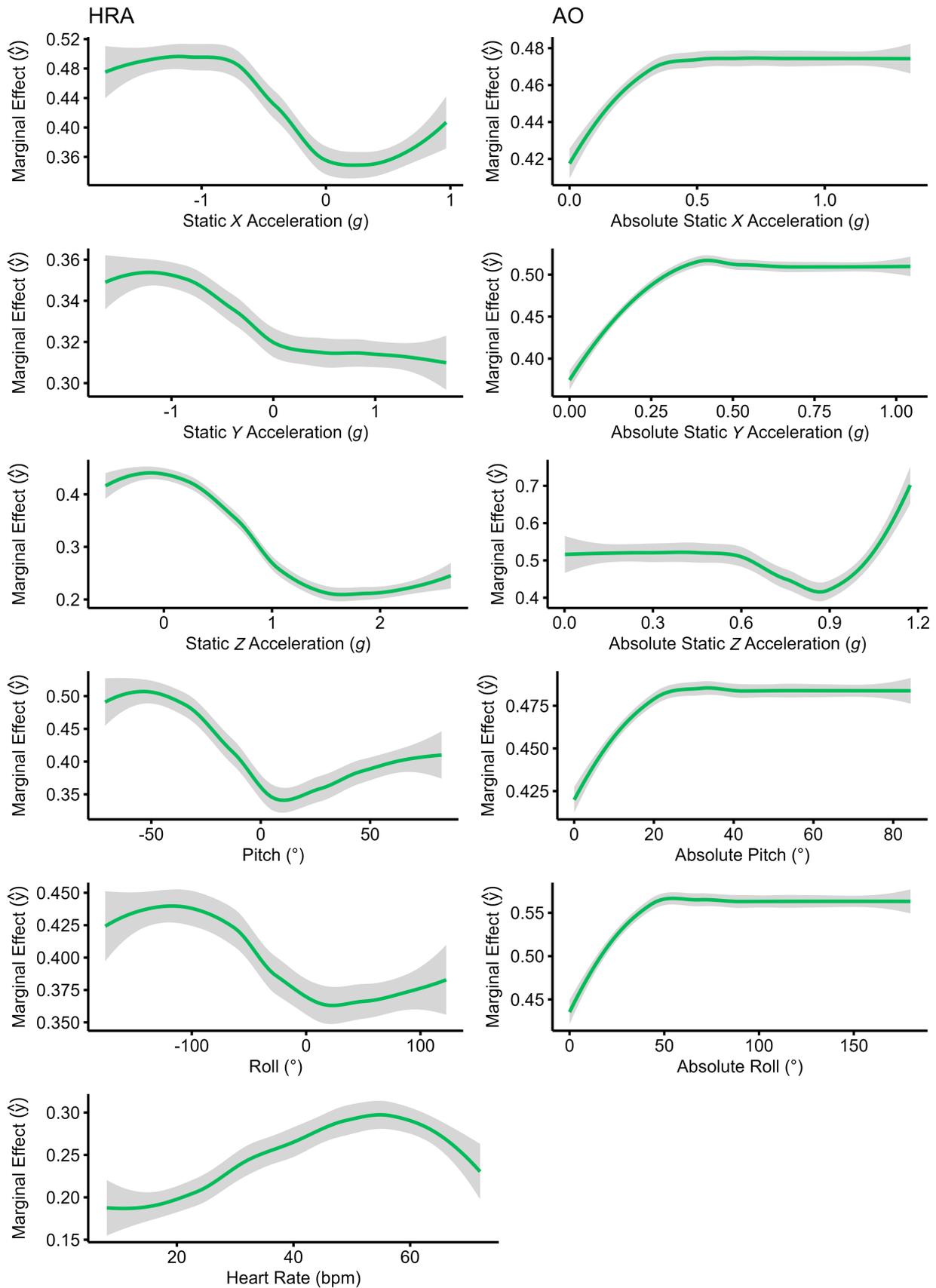
**FIGURE 5** Data shown correspond to the absolute acceleration-only (AO) model. Panels highlight how changes to the predictive threshold and the number of included consecutive predictions influence different model fit metrics using time aggregate predictions generated from the absolute AO model. See the [Figure 4](#) caption for further details.

attachment—presumably due to the combined effects of stress, injury, and increased drag forces resulting from the attached Sea Lamprey. Using these unique behavioral and heart rate responses, both the triaxial HRA model and the absolute AO model could predict Sea Lamprey attachments on Lake Trout with reasonable accuracy, but they lacked some precision and frequently predicted false negatives. Although these models are currently imperfect, the present results are promising and minor modifications could improve their performance, thereby permitting their use within the field.

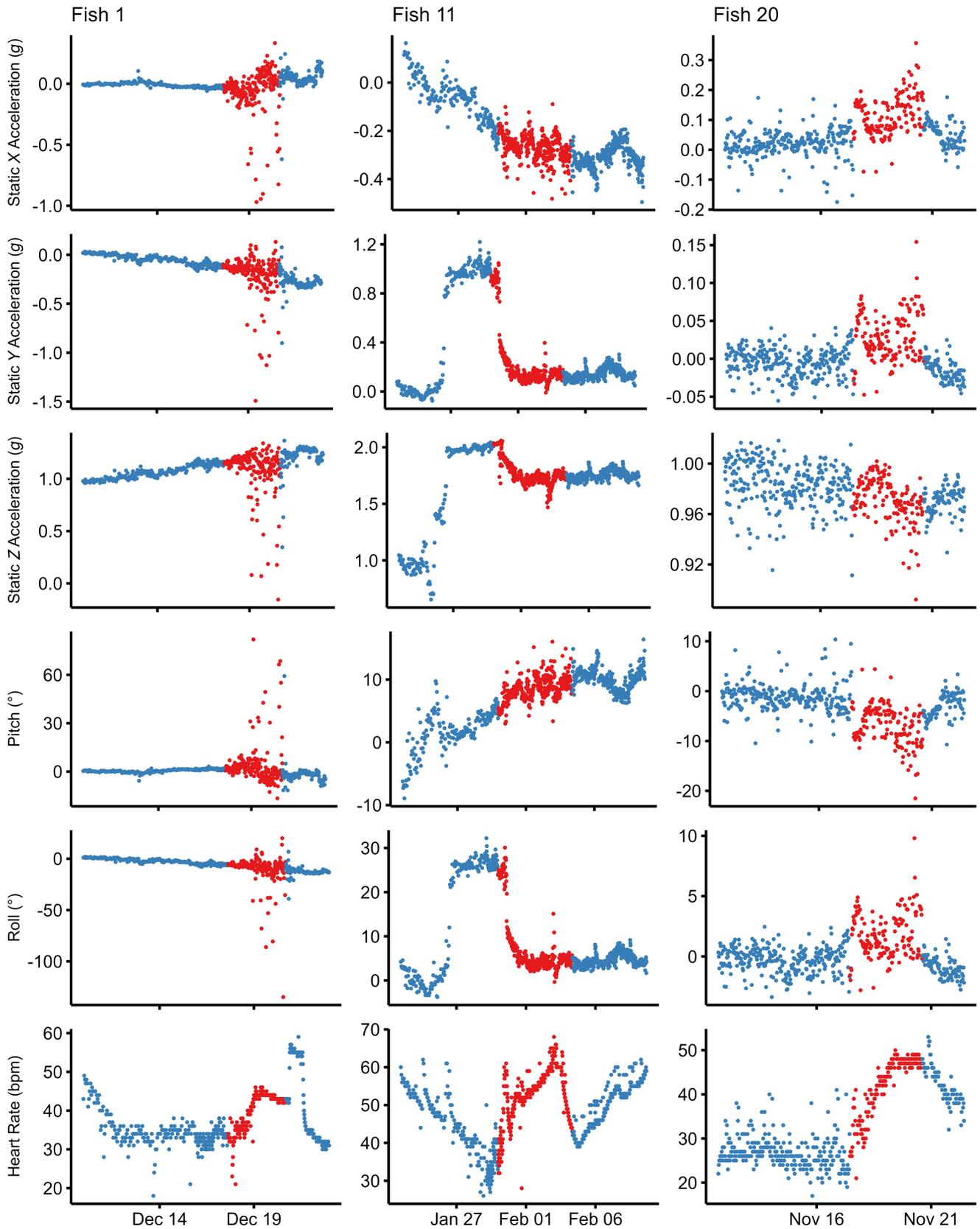
## Model considerations

Several issues presented themselves during data analysis and model building. Notably, most tags (in particular, AO tags) were inserted at an angle or rotated shortly after implantation, despite our use of anchoring sutures. Although corrections were used, if the tag rotated beyond a certain angle, axial measurements of acceleration no

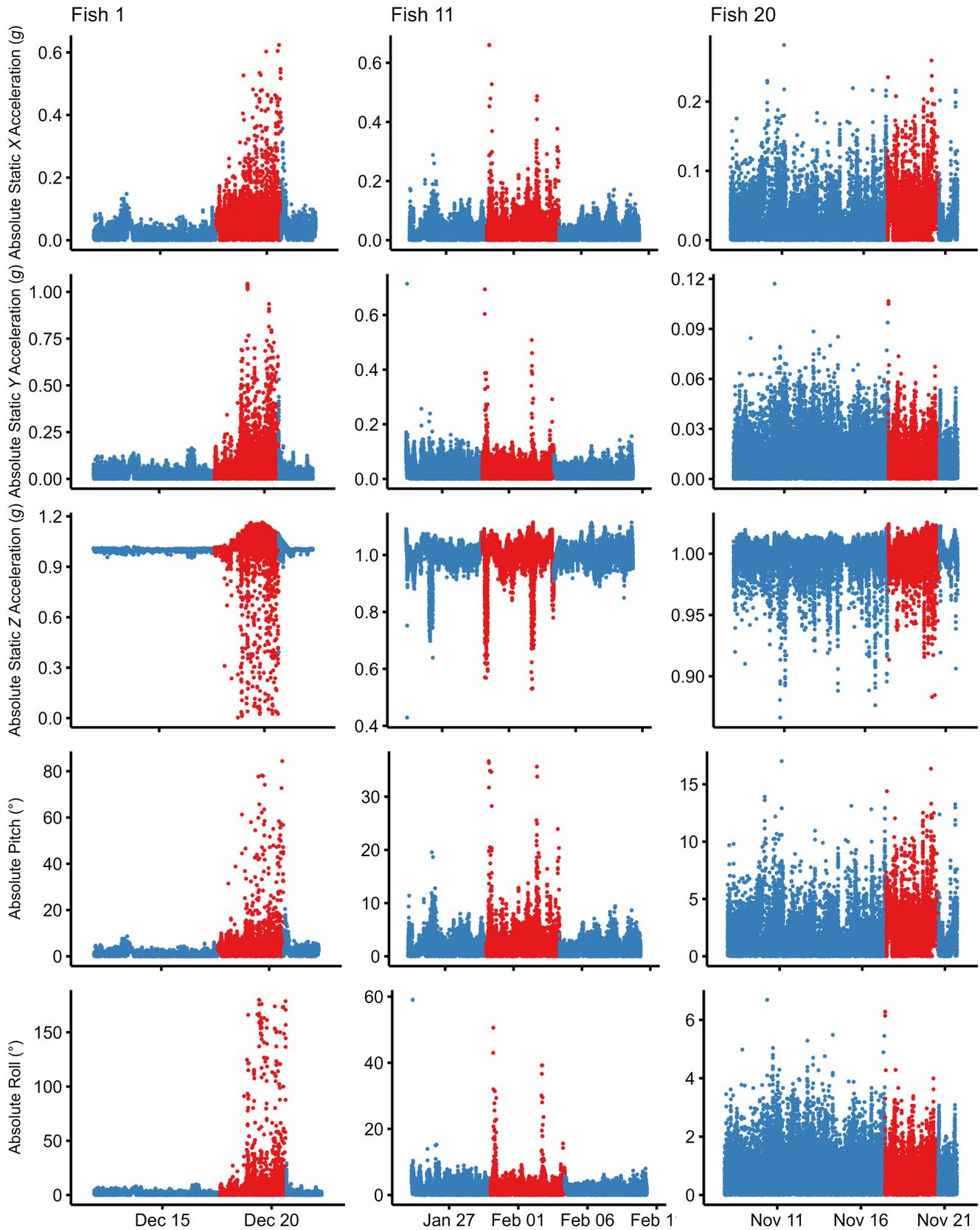
longer approximated their real values. For example, if a tag rotates to the side by 45°, the measurements of  $X$  acceleration would now include a component of gravity ( $Z$  acceleration) and measurements of  $Z$  acceleration would now include a component of  $X$  acceleration. It may be possible to calibrate the tags if the angle were known; however, this is not possible when the tag is inside of the fish. Moreover, a number of tags appeared to rotate or move throughout the experiment (this can be observed as a sloping line in acceleration values or unnaturally large variation in acceleration values). To correct these movements, we used a 24-h rolling average correction, which yielded some success in correcting AO tag movements (highlighted by MDA calculations; see [Figure 2](#)), but this likely dampened the overall responses during Sea Lamprey attachment, potentially impacting the predictive performance of the final model. Unsurprisingly, linear regression analyses indicated that poor tag orientation may have impacted both models' ability to accurately detect Sea Lamprey attachments; however, we lacked the power to verify this. Therefore, future studies seeking to predict



**FIGURE 6** Partial dependency plots of marginal predicted values ( $\hat{y}$ ) of the predictor variables used in the triaxial heart rate and acceleration (HRA; left) and absolute acceleration-only (AO; right) models. Higher  $\hat{y}$ -values indicate better positive prediction accuracy. Data are fitted with a locally estimated scatterplot smoother (solid green line), and 95% confidence intervals (gray bands) are presented. bpm, beats per minute.



**FIGURE 7** Examples of how three Lake Trout responded to Sea Lamprey attachment, visualized by changes in the top predictors included in the triaxial heart rate and acceleration model. Red points represent periods of Sea Lamprey attachment, whereas blue points represent periods when the Sea Lamprey was not attached. To see the associated plots for all Lake Trout, refer to [Figure S4](#).



**FIGURE 8** Examples of how three Lake Trout responded to Sea Lamprey attachment, visualized by changes in the top predictors included in the absolute acceleration-only model. Red points represent periods of Sea Lamprey attachment, whereas blue points represent periods when the Sea Lamprey was not attached. To see the associated plots for all Lake Trout, refer to [Figure S5](#).

behavioral responses from acceleration values or body orientation should take precautions to ensure that the tag is implanted at the correct angle and is heavily secured to the body wall.

Behavior in the laboratory rarely equates exactly to that observed in the field. Thus, certain behaviors that were unaccounted for in the laboratory could trigger false-positive detections if applied to wild fish. Efforts were made to collect a suite of behaviors to test our models—which, in theory, should reduce false positives when used in wild fish. For example, burst swimming was recorded using a chase protocol in the later trials. Additionally, irregular behaviors that were displayed near death and while recovering from surgery were also recorded. Chases performed in the later trials did not appear to cause changes in the final models' performance, and linear regressions indicated that the positive prediction accuracy was not different between Lake Trout that died and those that did not die. However, we lacked power in our analyses, and more testing is needed to ascertain whether burst swimming or irregular premortality behaviors impair prediction accuracy. Similarly, we could not determine how the Sea Lamprey attachment location impacted predictive performance due to our small sample size and the large range in attachment points.

Other untested field behaviors (e.g., spawning behavior) may be problematic; however, TAPs may reduce the likelihood of false-positive predictions in field applications that use longer time frames (e.g., if 40 consecutive predictions are used in the calculation of TAPs, this would require ~13.5 h of HRA data and 40 min of AO data). Most behaviors that could feasibly trigger a false-positive prediction occur over shorter periods of time (e.g., foraging bouts or spawning behaviors), whereas Sea Lamprey typically remain attached to Lake Trout for many days or weeks during an attack (Swink 2003). It is possible that irregularly quick attachments (e.g., minutes to hours in length) could be missed using this method, but such short attachments are presumably rare and impose lesser physiological costs on Lake Trout.

Factors such as the size of the Lake Trout and Sea Lamprey used in our trials could affect model performance in the field; to address this, we used a range of sizes for both species in our trials. Lake Trout ranged from 1394 to 6080 g (total length = 576–885 mm), and Sea Lamprey ranged from 83 to 261 g (total length = 375–544 mm). The size of the Lake Trout did not seem to affect the performance of either model; however, there appeared to be a positive relationship between Sea Lamprey size and positive prediction accuracy in the triaxial HRA model, suggesting that attachments from smaller Sea Lamprey may have been underestimated. This could be particularly problematic for

small, postmetamorphosed Sea Lamprey. The average size at metamorphosis is about 140 mm, which is less than half the size of the smallest Sea Lamprey that was tested in our models (Purvis 1980; Youson et al. 1993). Host mortality is known to be influenced by Sea Lamprey size, so underestimations of smaller, less lethal Sea Lamprey attachments may be relatively less important from a management perspective and attachments by medium to large Sea Lamprey are less likely to be underestimated (Swink 2003). There is also evidence that Sea Lamprey host selection is influenced by their size, with smaller lamprey selecting Lake Trout less frequently when compared to larger lamprey (Harvey et al. 2008). In a few instances ( $n=3$ ), multiple Sea Lamprey were attached to a single Lake Trout for a period of time during our experiment; however, this did not seem to alter model performance, and simultaneous Sea Lamprey attachments rarely occur in the Great Lakes (Swink 2003).

Our models could not delineate between feeding and nonfeeding Sea Lamprey attachments. Although there was some indication that the heart rate response differed between nonfeeding and feeding attachments, there was an insufficient nonfeeding sample to produce a model that could delineate attack types. Nonfeeding attachments are somewhat common and are hypothesized to occur when Sea Lamprey are satiated, allowing them to reduce energetic costs associated with swimming (Farmer 1980). Although our model could not distinguish between attack types, it could determine attack rates and one may be able to assess associated lethality rates (i.e., by inferring mortality from the absence of movement and/or flatline acceleration or heart rate values in tagged individuals after a Sea Lamprey attachment; Villegas-Ríos et al. 2020), which are ultimately the sought-after end points for fisheries managers. However, improvements in model precision would be necessary prior to field deployments.

## Model application

The use of TAPs requires numerous measurements over a relatively short period of time for optimal performance. We tested the final models using 40 consecutive predictions, which equated to approximately 13.5 h of HRA data or 40 min of AO data (due to differences in the tag sampling rate), although even more consecutive predictions would likely be needed for optimal predictive performance (see Tables S3 and S4). Because numerous measurements are required, a relatively high transmission frequency will be necessary. Furthermore, because Sea Lamprey attacks are a somewhat rare occurrence, a longer battery life would be favorable if included in acoustic transmitter tags. If incorporated into an acoustic transmitter tag, successful

detections would also depend on the density of acoustic receivers within a given area and the likelihood of Sea Lamprey attack for the Lake Trout population of interest. The density and configuration of receivers (e.g., grids versus lines) will also influence one's ability to measure the associated mortality rate (Peterson et al. 2021). Simulations may help researchers to determine the optimal locations for first deployments and the necessary number of tags and receivers that are needed to adequately sample a population. Depending on the amount of data available, researchers could also tweak the TAP parameters to fit their needs. For instance, if ample data are collected, one could increase the predictive threshold, which should increase positive prediction accuracy and reduce false positives but may also result in more missed detections (e.g., see Figures 4 and 5). One should also employ ecological common sense when reviewing positive TAPs. For example, a dense cluster of positive predictions over numerous hours to days would be more likely to be a true attachment than a singular positive prediction among negative predictions.

The triaxial HRA model demonstrated model fit metrics that were comparable to those of the absolute AO model and performed similarly after TAPs were calculated; however, the triaxial HRA model is unlikely to be currently applicable in the field. Notably, heart rate sensor-equipped acoustic tags remain a “niche” technology that has yet to be widely produced or embraced (detailed by Cooke et al. 2016). Moreover, an acoustic transmitter that measures both triaxial acceleration and heart rate, which would be needed to use the HRA model, has not yet been produced. Additionally, numerous factors are known to influence heart rates in fish, such as temperature, digestion (i.e., specific dynamic action [SDA]), stress (e.g., presence of predators; noise from human infrastructure), and activity level (Kalinin et al. 2009; Chabot et al. 2016; Svendsen et al. 2021). Temperature has a particularly strong influence on the heart rate (Clarke 2017). Temperature was controlled in our experimental trials to a relatively narrow range (average  $\pm$ SD =  $9.09 \pm 1.29^\circ\text{C}$ ) to exclude potential thermal confounds. Thus, is it probable that the triaxial HRA model would not perform adequately across the range of temperatures that Lake Trout are likely to experience in the wild. Further model calibration would likely be required using a range of temperatures prior to application of the HRA weighted model in the field. Additionally, because Lake Trout were not fed during the experiment, the SDA effects on heart rate were not included in the training or testing of our models; thus, the resulting increase in heart rate after a meal could trigger false-positive prediction. For those reasons, we recommend starting with field deployments of telemetered versions that focus on the locomotor activity-derived metrics from acceleration sensors.

## Summary

Lake Trout display unique behavioral and heart rate responses to Sea Lamprey attachment that can be subsequently recorded using implantable tags. Capitalizing on these recorded changes, we produced models to predict Sea Lamprey attachment on Lake Trout. Our models showed reasonable accuracy, but they lacked some precision and frequently predicted false negatives using test data sets from Lake Trout in the laboratory. Further development is required to improve these models prior to field deployments. However, we believe that these models were rather accurate considering the tag orientation issues, which could be easily improved with some minor corrections in surgical methodology. Pilot field testing would be required to determine the likelihood of false positives in the wild; however, we believe that the temporal aggregation of predictions would be robust to false positives, as it considers longer time frames. Although both models performed similarly, the absolute AO model appears to be the most viable for field application because its required measurements (i.e., using acceleration sensors rather than ECG sensors) are simpler to record and transmit. We believe that this model could be a useful tool for investigating the impacts of Sea Lamprey on wild populations of Lake Trout without survivor detection bias and is deserving of further research. The ability to record direct observations of Sea Lamprey attack rates in the field will be valuable for fisheries managers and Sea Lamprey control agents to inform fisheries management and control decisions.

## ACKNOWLEDGMENTS

This research was funded by the Great Lakes Fishery Commission. S. J. Cooke and C. Reeve are supported by the Natural Sciences and Engineering Research Council of Canada and GenFish (Genome Canada).

## CONFLICT OF INTEREST STATEMENT

The authors report no conflicts of interest.

## DATA AVAILABILITY STATEMENT

Data are available from the corresponding author (C. Reeve) upon reasonable request.

## ETHICS STATEMENT

This research was performed in accordance with U.S. federal guidelines for the care and use of animals and in accordance with American Fisheries Society guidelines.

## ORCID

Connor Reeve  <https://orcid.org/0000-0002-0238-0575>

Scott M. Miehl  <https://orcid.org/0000-0002-5546-1854>

Steven J. Cooke  <https://orcid.org/0000-0002-5407-0659>  
 Jacob W. Brownscombe  <https://orcid.org/0000-0003-2496-8764>

## REFERENCES

- Adams, J. V., Jones, M. L., & Bence, J. R. (2021). Using simulation to understand annual Sea Lamprey marking rates on Lake Trout. *Journal of Great Lakes Research*, 47(Supplement 1), S628–S638. <https://doi.org/10.1016/j.jglr.2020.08.008>
- Brant, C. (2019). *Great Lakes Sea Lamprey: The 70 year war on a biological invader*. University of Michigan Press. <https://doi.org/10.3998/mpub.10209106>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- 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, Article 62. <https://doi.org/10.1007/s00227-018-3318-y>
- Brownscombe, J. W., Gutowsky, L. F. G., Danylchuk, A. J., & Cooke, S. J. (2014). Foraging behaviour and activity of a marine benthivorous fish estimated using tri-axial accelerometer biologgers. *Marine Ecology Progress Series*, 505, 241–251. <https://doi.org/10.3354/meps10786>
- Brownscombe, J. W., Midwood, J. D., & Cooke, S. J. (2021). Modeling fish habitat: Model tuning, fit metrics, and applications. *Aquatic Sciences*, 83, Article 44. <https://doi.org/10.1007/s00027-021-00797-5>
- Brownscombe, J. W., Thiem, J. D., Hatry, C., Cull, F., Haak, C. R., Danylchuk, A. J., & Cooke, S. J. (2013). Recovery bags reduce post-release impairments in locomotory activity and behavior of bonefish (*Albula* spp.) following exposure to angling-related stressors. *Journal of Experimental Marine Biology and Ecology*, 440, 207–215. <https://doi.org/10.1016/j.jembe.2012.12.004>
- Chabot, D., Koenker, R., & Farrell, A. P. (2016). The measurement of specific dynamic action in fishes. *Journal of Fish Biology*, 88, 152–172. <https://doi.org/10.1111/jfb.12836>
- Clarke, A. (2017). *Principles of thermal ecology: Temperature, energy, and life*. Oxford University Press. <https://doi.org/10.1093/oso/9780199551668.001.0001>
- Clarke, T. M., Whitmarsh, S. K., Hounslow, J. L., Gleiss, A. C., Payne, N. L., & Huvener, C. (2021). Using tri-axial accelerometer loggers to identify spawning behaviours of large pelagic fish. *Movement Ecology*, 9, Article 26. <https://doi.org/10.1186/s40462-021-00248-8>
- Coble, D. W., Bruesewitz, R. E., Fratt, T. W., & Scheirer, J. W. (1990). Lake Trout, Sea Lampreys, and overfishing in the upper Great Lakes: A review and reanalysis. *Transactions of the American Fisheries Society*, 119(6), 985–995. [https://doi.org/10.1577/1548-8659\(1990\)119<0985:LTSLAO>2.3.CO;2](https://doi.org/10.1577/1548-8659(1990)119<0985:LTSLAO>2.3.CO;2)
- Cooke, S. J., Brownscombe, J. W., Raby, G. D., Broell, F., Hinch, S. G., Clark, T. D., & Semmens, J. M. (2016). Remote bioenergetics measurements in wild fish: Opportunities and challenges. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 202, 23–37. <https://doi.org/10.1016/j.cbpa.2016.03.022>
- Cooke, S. J., Graeb, D. S., Suski, C. D., & Ostrand, K. G. (2003). Effects of suture material on incision healing, growth, and survival of juvenile Largemouth Bass implanted with miniature radio transmitters: Case study of a novice and experienced fish surgeon. *Journal of Fish Biology*, 62, 1366–1380. <https://doi.org/10.1046/j.1095-8649.2003.00119.x>
- Ebener, M. P. (1998). *A Lake Trout rehabilitation guide for Lake Huron*. Great Lakes Fishery Commission.
- Ebener, M. P., Bence, J. R., Newman, K. R., & Schneeberger, P. J. (2005). Application of statistical catch-at-age models to assess Lake Whitefish stocks in the 1836 Treaty-ceded waters of the upper Great Lakes. In L. C. Mohr & T. F. Nalepa (Eds.), *Proceedings of a workshop on the dynamics of Lake Whitefish (Coregonus clupeaformis) and the amphipod Diporeia spp. in the Great Lakes* (Technical Report 66, pp. 271–309). Great Lakes Fishery Commission.
- Farmer, G. J. (1980). Biology and physiology of feeding in adult lampreys. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(11), 1751–1761. <https://doi.org/10.1139/f80-220>
- Gleiss, A. C., Wilson, R. P., & Shepard, E. L. C. (2011). Making over-all dynamic body acceleration work: On the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution*, 2, 23–33. <https://doi.org/10.1111/j.2041-210X.2010.00057.x>
- Great Lakes Fishery Commission. (2023). Status of Sea Lamprey. <http://www.glfsc.org/status.php>
- Greenwell, B. M. (2022). pdp: Partial dependence plots. Version 0.8.1. <https://cran.r-project.org/web/packages/pdp/index.html>
- Hansen, M. J., Peck, J. W., Schorfhaar, R. G., Selgeby, J. H., Schreiner, D. R., Schram, S. T., Swanson, B. L., MacCallum, W. R., Burnham-Curtis, M. K., Curtis, G. L., Heinrich, J. W., & Young, R. J. (1995). Lake Trout (*Salvelinus namaycush*) populations in Lake Superior and their restoration in 1959–1993. *Journal of Great Lakes Research*, 21, 152–175. [https://doi.org/10.1016/S0380-1330\(95\)71088-7](https://doi.org/10.1016/S0380-1330(95)71088-7)
- Harvey, C. J., Ebener, M. P., & White, C. K. (2008). Spatial and ontogenetic variability of Sea Lamprey diets in Lake Superior. *Journal of Great Lakes Research*, 34, 434–449. [https://doi.org/10.3394/0380-1330\(2008\)34\[434:SAOVOS\]2.0.CO;2](https://doi.org/10.3394/0380-1330(2008)34[434:SAOVOS]2.0.CO;2)
- He, J. X., Bence, J. R., Madenjian, C. P., & Claramunt, R. M. (2020). Dynamics of Lake Trout production in the main basin of Lake Huron. *ICES Journal of Marine Science*, 77, 975–987. <https://doi.org/10.1093/icesjms/fsaa030>
- He, J. X., Ebener, M. P., Riley, S. C., Cottrill, A., Kowalski, A., Koproski, S., & Mohr, L. (2012). Lake Trout status in the main basin of Lake Huron, 1973–2010. *North American Journal of Fisheries Management*, 32, 402–412. <https://doi.org/10.1080/02755947.2012.675947>
- Heupel, M. R., Semmens, J. M., & Hobday, A. J. (2006). Automated acoustic tracking of aquatic animals: Scales, design and deployment of listening station arrays. *Marine and Freshwater Research*, 57(1), 1–13. <https://doi.org/10.1071/MF05091>
- Hvas, M., Folkedal, O., & Oppedal, F. (2020). Heart rate bio-loggers as welfare indicators in Atlantic Salmon (*Salmo salar*). *Aquaculture*, 529, Article 735630. <https://doi.org/10.1016/j.aquaculture.2020.735630>
- Johnson, N. S., Lewandoski, S. A., & Merkes, C. (2021). Assessment of Sea Lamprey (*Petromyzon marinus*) diet using DNA metabarcoding of feces. *Ecological Indicators*, 125, Article 107605. <https://doi.org/10.1016/j.ecolind.2021.107605>
- Kalinin, A. L., Costa, M., Rantin, F. T., & Glass, M. L. (2009). Effects of temperature on cardiac function in fish. In M. L. Glass &

- S. C. Wood. (Eds.), *Cardio-respiratory control in vertebrates: Comparative and evolutionary aspects*. Springer-Verlag. [https://doi.org/10.1007/978-3-540-93985-6\\_6](https://doi.org/10.1007/978-3-540-93985-6_6).
- Koonce, J. F. (1987). *Application of models of Lake Trout/Sea Lamprey interaction to the implementation of integrated pest management of Sea Lamprey in Lake Ontario*. Great Lakes Fishery Commission.
- Kornis, M. S., Bronte, C. R., Holey, M. E., Hanson, S. D., Treska, T. J., Jonas, M., C., Claramunt, R. M., Robillard, S. R., Breidert, S. R., Donner, K. C., Lenart, S. J., Martell, A. W., McKee, P. C., & Olsen, E. J. (2019). Factors affecting post-release survival of coded-wire-tagged Lake Trout in Lake Michigan at four historical spawning locations. *North American Journal of Fisheries Management*, 39, 868–895. <https://doi.org/10.1002/nafm.10338>
- Kuhn, M. (2023). caret: Classification and regression training. Version 6.94. <https://cran.r-project.org/web/packages/caret/caret.pdf>
- Liaw, A. (2022). randomForest: Breiman and Cutler's random forests for classification and regression. Version 4.7-1.1. <https://cran.r-project.org/web/packages/randomForest/randomForest.pdf>
- McKee, P. C., Toney, M. L., Hansen, M. J., & Holey, M. E. (2004). Performance of two strains of Lake Trout stocked in the mid-lake refuge of Lake Michigan. *North American Journal of Fisheries Management*, 24, 1101–1111. <https://doi.org/10.1577/M03-142.1>
- Pedley, M. (2013). Tilt sensing using a three-axis accelerometer (Freescale Semiconductor Application Note AN3461). Freescale.
- Peterson, L. K., Jones, M. L., Brenden, T. O., Vandergoot, C. S., & Krueger, C. C. (2021). Evaluating methods for estimating mortality from acoustic telemetry data. *Canadian Journal of Fisheries and Aquatic Sciences*, 78(10), 1444–1454. <https://doi.org/10.1139/cjfas-2020-0417>
- Purvis, H. A. (1980). Effects of temperature on metamorphosis and the age and length at metamorphosis in Sea Lamprey (*Petromyzon marinus*) in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 1827–1834. <https://doi.org/10.1139/f80-225>
- Pycha, R. L. (1980). Changes in mortality of Lake Trout (*Salvelinus namaycush*) in Michigan waters of Lake Superior in relation to Sea Lamprey (*Petromyzon marinus*) predation, 1968–78. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 2063–2073. <https://doi.org/10.1139/f80-247>
- Pycha, R. L., & King, G. R. (1975). *Changes in the Lake Trout population of southern Lake Superior in relation to the fishery, the Sea Lamprey, and stocking, 1950-70* (Technical Report 28). Great Lakes Fishery Commission.
- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reid, C. H., Vandergoot, C. S., Midwood, J. D., Stevens, E. D., Bowker, J., & Cooke, S. J. (2019). On the electroimmobilization of fishes for research and practice: Opportunities, challenges, and research needs. *Fisheries*, 44, 576–585. <https://doi.org/10.1002/fsh.10307>
- Robinson, K. F., Miehl, S. M., & Siefkes, M. J. (2021). Understanding Sea Lamprey abundances in the Great Lakes prior to broad implementation of Sea Lamprey control. *Journal of Great Lakes Research*, 47(1), S328–S334. <https://doi.org/10.1016/j.jglr.2021.04.002>
- Schneider, C. P., Owens, R. W., Bergstedt, R. A., & O'Gorman, R. (1996). Predation by Sea Lamprey (*Petromyzon marinus*) on Lake Trout (*Salvelinus namaycush*) in southern Lake Ontario, 1982–1992. *Canadian Journal of Fisheries and Aquatic Sciences*, 53, 1921–1932. <https://doi.org/10.1139/cjfas-53-9-1921>
- Sitar, S. P., Bence, J. R., Johnson, J. E., Ebener, M. P., & Taylor, W. W. (1999). Lake Trout mortality and abundance in southern Lake Huron. *North American Journal of Fisheries Management*, 19, 881–900. [https://doi.org/10.1577/1548-8675\(1999\)019](https://doi.org/10.1577/1548-8675(1999)019)
- Smircich, M. G., & Kelly, J. (2014). Extending the 2% rule: The effects of heavy internal tags on stress physiology, swimming performance, and growth in Brook Trout. *Animal Biotelemetry*, 2, Article 16. <https://doi.org/10.1186/2050-3385-2-16>
- Smith, B. R., & Tibbles, J. J. (1980). Sea Lamprey (*Petromyzon marinus*) in Lakes Huron, Michigan, and Superior: History of invasion and control, 1936–78. *Canadian Journal of Fisheries and Aquatic Sciences*, 37(11), 1708–1801. <https://doi.org/10.1139/f80-222>
- Smith, S. E., Sitar, S. P., Goetz, F. W., Mar Huertas, G., Armstrong, B. M., & Murphy, C. A. (2016). Differential physiological response to Sea Lamprey parasitism between Lake Trout (*Salvelinus namaycush*) morphotypes from Lake Superior. *Canadian Journal of Fisheries and Aquatic Sciences*, 73, 1815–1829. <https://doi.org/10.1139/cjfas-2015-0448>
- Svendsen, E., Føre, M., Økland, F., Gråns, A., Hedger, R. D., Alfredsen, J. A., Uglem, I., Rosten, C. M., Frank, K., Erikson, U., & Finstad, B. (2021). Heart rate and swimming activity as stress indicators for Atlantic Salmon (*Salmo salar*). *Aquaculture*, 531, Article 735804. <https://doi.org/10.1016/j.aquaculture.2020.735804>
- Swink, W. D. (2003). Host selection and lethality of attacks by Sea Lampreys (*Petromyzon marinus*) in laboratory studies. *Journal of Great Lakes Research*, 29, 307–319. [https://doi.org/10.1016/S0380-1330\(03\)70496-1](https://doi.org/10.1016/S0380-1330(03)70496-1)
- Use of Fishes in Research Committee (joint committee of the American Fisheries Society, the American Institute of Fishery Research Biologists, and the American Society of Ichthyologists and Herpetologists). (2014). *Guidelines for the use of fishes in research*. American Fisheries Society. <https://doi.org/10.47886/9781934874394>
- Villegas-Ríos, D., Freitas, C., Moland, E., Thorbjørnsen, S. H., & Olsen, E. M. (2020). Inferring individual fate from aquatic acoustic telemetry data. *Methods in Ecology and Evolution*, 11, 1186–1198. <https://doi.org/10.1111/2041-210X.13446>
- Wagner, G. N., Cooke, S. J., Brown, R. S., & Deters, K. A. (2011). Surgical implantation techniques for electronic tags in fish. *Reviews in Fish Biology and Fisheries*, 21, 71–81. <https://doi.org/10.1007/s11160-010-9191-5>
- Weeks, C. T. (1997). *Dynamics of Lake Trout (Salvelinus namaycush) size and age structure in Michigan waters of Lake Superior 1971–1995* [Master's thesis, Michigan State University]. MSU Libraries Digital Collection. <https://doi.org/10.25335/hxx4-ef80>
- Youngs, W. D. (1980). Estimate of lamprey-induced mortality in a Lake Trout (*Salvelinus namaycush*) population. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 2151–2158. <https://doi.org/10.1139/f80-257>
- Yousaf, M. N., Røn, O., Hagen, P. P., & McGurk, C. (2022). Monitoring fish welfare using heart rate bio-loggers in farmed Atlantic Salmon (*Salmo salar* L.): An insight into the surgical recovery. *Aquaculture*, 555, Article 738211. <https://doi.org/10.1016/j.aquaculture.2022.738211>

- Youson, J. H., Holmes, J. A., Guchardi, J. A., Seelye, J. G., Beaver, R. E., Gersmehl, J. E., Sower, S. A., & Beamish, F. W. H. (1993). Importance of condition factor and the influence of water temperature and photoperiod on metamorphosis of Sea Lamprey *Petromyzon marinus*. *Canadian Journal of Fisheries and Aquatic Sciences*, 50(11), 2448–2456. <https://doi.org/10.1139/f93-269>
- Zrini, Z. A., & Gamperl, A. K. (2021). Validating Star-Oddi heart rate and acceleration data storage tags for use in Atlantic Salmon

(*Salmo salar*). *Animal Biotelemetry*, 9, Article 12. <https://doi.org/10.1186/s40317-021-00235-1>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.