Advance View
https://doi.org/10.3354/meps14274

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

July 6, 2023

Contribution to the Theme Section 'Small pelagic fish: new research frontiers'

Evaluating factors affecting the distribution and timing of Pacific Herring *Clupea pallasii* spawn in British Columbia

Christopher N. Rooper[1,*], Jennifer L. Boldt[1], Jaclyn Cleary[1], M. Angelica Peña[2], Matthew Thompson[1], Matthew Grinnell[1]

[1]**Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia V9T 6N7, Canada**

[2]**Institute of Ocean Sciences, Fisheries and Oceans Canada, 9860 West Saanich Road, Sidney, British Columbia V8L 4B2, Canada**

ABSTRACT: Pacific herring *Clupea pallasii* spawn in nearshore areas in late winter to early spring, but factors influencing the timing and spatial distribution of spawning are not well known. We modeled the temporal and spatial distribution of spawning for 5 herring stocks in British Columbia from egg deposition surveys conducted from 1988–2018 using different sets of environmental predictors and modeling methods. Random forest modeling showed that the timing of spawning in each year was mainly influenced by the number of daylight hours being >10.5, cumulative degree days >100 and salinity at 30.5. The spatial distribution of spawning tended to occur at consistent locations over time. Results showed that the probability of spawning occurring at a transect in a given year was largely determined by the biomass of herring and location of the transect relative to the center of spawning. Environmental factors at individual transects played a much smaller role in determining spawn distribution. There was mixed evidence for spatial expansion of spawning distribution in years of high biomass, with some stocks, such as Haida Gwaii, showing no expansion of the spawning area in years of higher biomass. Since Pacific herring recruitment has been linked to their ability to time larval hatch to spring bloom timing, future warming temperatures may result in earlier herring spawning relative to the spring bloom. This will increase the probability of mismatch with prey, impacting larval herring starvation, growth and perhaps mortality, leading to reductions in recruitment to these important stocks.

KEY WORDS: Spawn timing · Spatial distribution · Pacific herring · British Columbia · Species distribution modeling · Random forest

1. INTRODUCTION

Small pelagic fish populations can undergo large fluctuations, and the mechanisms determining survival and recruitment are not fully understood (Peck et al. 2021, Boldt et al. 2022). Linkages between small pelagic fish population fluctuations and oceanographic and biological conditions have generally been attributed to processes occurring in early life history stages of fish that impact recruitment to adult populations (Hjort 1914, Cushing 1969, Houde 1987). Spawning phenology relative to ocean conditions and productivity seems to be especially important (Schweigert et al. 2013, Arula et al. 2015, Polte et al. 2021, Dias et al. 2022) in both Pacific herring *Clupea pallasii* Valenciennes, 1847 and Atlantic herring *C. harengus* Linnaeus, 1758, as growth through the initial months may determine survival to first feeding and through the first winter (Norcross et al. 2001, Hufnagl & Peck 2011, Polte et al. 2014, Sewall et al. 2018).

Pacific herring are an important forage base for birds, mammals and piscivorous fishes in coastal

*Corresponding author: chris.rooper@dfo-mpo.gc.ca

Publisher: Inter-Research · www.int-res.com

waters of the North Pacific Ocean (Olesiuk et al. 1990, Rooper & Haldorson 2000, Sullivan et al. 2002, Pearsall & Fargo 2007, Trites et al. 2007, Schweigert et al. 2010). They are also highly valued culturally, recreationally and commercially for human consumption, supporting roe fisheries and Spawn on Kelp fisheries during spawning events as well as food and bait fisheries throughout the year. As a small pelagic species, their abundance can be highly variable due to changing ocean conditions (Stocker et al. 1985, Stocker & Noakes 1988, Williams & Quinn 2000), mortality from predation (Ware & McFarlane 1986, Purcell et al. 1987, Moran et al. 2018) and changes in prey abundance (Foy & Paul 1999, Schweigert et al. 2010, Boldt et al. 2019).

Because of their importance, Pacific herring life history is reasonably well studied. Herring spawn in late winter to early spring on vegetation in intertidal and subtidal areas, most within 10 m of the mean low tide level (Haegele et al. 1981, Haegele & Schweigert 1985). Their eggs incubate for up to 3 wk (depending on temperatures) and then larvae hatch and move into nearshore areas for rearing (Outram 1955, Taylor 1971, Alderdice & Hourston 1985). Spawning in British Columbia (BC) has historically occurred during early March in the Strait of Georgia (SOG), with generally later spawning dates in more northerly latitudes, although there are exceptions (Hay 1985, Haegele & Schweigert 1985). BC herring are managed as 5 main migratory stocks: east coast of Haida Gwaii (HG), Prince Rupert District (PRD), Central Coast (CC), SOG and West Coast of Vancouver Island (WCVI) (DFO 2021). Pacific herring spawning occurs in the same general areas every year but does not occur at all locations within each general area each year, leading to some uncertainty about the factors controlling spawn distribution (Hay & McCarter 1999, Hay et al. 2009).

Herring spawn amounts, distribution and timing have changed over time, and the factors causing these changes are not fully understood (Hay & Kronlund 1987, Hay & McCarter 1999). Spawning grounds are characterized as tidally active habitats with marine vegetation, where silt cannot suffocate eggs but that are sheltered enough to minimize egg loss by wave action (Haegele & Schweigert 1985). Eggs are deposited in deeper depths, where the slope of the sea floor is steeper (Haegele et al. 1981). Most spawning occurs over a narrow range of temperatures and salinities (6.5–9.8°C and 22.4–28.7; Outram 1975, Hay et al. 1984) in both day and night (Hay 1985). Variations in spawn intensity and deposition have been attributed to vegetation types (Haegele et al. 1981), herring biomass (Haegele & Schweigert 1985, Hay & Kronlund 1987), pre-spawning school size and density, temperature and tides (current and cycle) (Hay & Kronlund 1987). Some studies have indicated that most spawn timing is related to temperature (e.g. Haegele & Schweigert 1985), while other studies suggest herring spawn just prior to the spring bloom, temporally aligning first-feeding larvae with their prey (Schweigert et al. 2013, Boldt et al. 2019). With the latter, the alignment of first-feeding larvae with the spring bloom would then result in higher prey availability, higher survival and higher recruitment (Schweigert et al. 2013, Boldt et al. 2019).

Spatial and temporal changes in Pacific herring spawning affect their availability to First Nations, roe fisheries and predator populations, and may drive changes in the marine ecosystem. In BC, there has been a change in herring egg density over time across a wide range of spawning biomass which has been attributed to changes in herring size-at-age and temperature (Hay et al. 2019). Determining how herring will respond to future environmental and biological changes requires an evaluation of how spawn distributions have changed across space and time and an understanding of the drivers that influence these distributions. Therefore, the objectives of this study were to (1) model the relationships between oceanographic variables and the timing of spawn deposition over 31 yr and (2) model the relationship between the observed spawn deposition in each year and the physical characteristics of the spawning habitat, while accounting for spawn biomass and location. Finally, we comment on the potential changes in spawn timing and distribution relative to ocean warming in BC to elucidate potential changes in Pacific herring stock productivity in future years.

2. METHODS

This study was carried out using data collected during annual spawning ground surveys for Pacific herring in BC from 1988–2018 (Fig. 1). There are 5 major stocks of herring identified by their general spawning location (DFO 2021). The HG major stock generally spawns on the eastern side of Haida Gwaii; the PRD stock spawns along the north coast of the BC mainland (primarily in Kitkatla and Big Bay); the CC stock spawns in the region from Kitasu Bay to Rivers Inlet on the BC mainland; the SOG stock spawns in inside waters between the southern BC mainland and the east coast of Vancouver Island in the Salish Sea; and the WCVI stock spawns predominantly

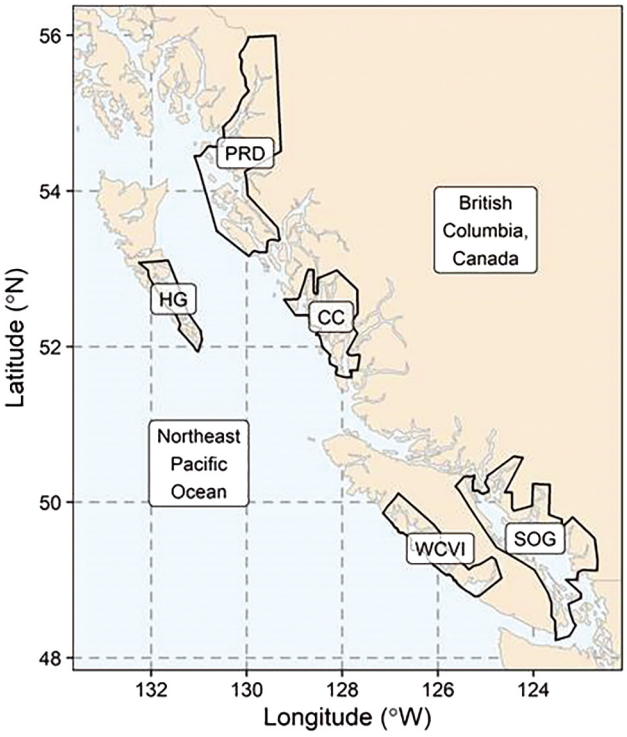


Fig. 1. British Columbia, Canada, showing the main Pacific herring spawning grounds and the 5 major stock assessment regions: HG: Haida Gwaii; PRD: Prince Rupert District; CC: Central Coast; SOG: Strait of Georgia; WCVI: West Coast of Vancouver Island

around Barkley and Clayoquot sounds on the west coast of Vancouver Island (Fig. 1). All of these stocks have been fished both for food, social, and ceremonial purposes and commercially, although in recent years commercial fishing has occurred primarily on the SOG and PRD stocks (DFO 2020). Commercial fisheries include food and bait, special use, spawn-on-kelp products and roe herring.

2.1. Spawn deposition survey data

The spawning grounds for each of the 5 major stocks of herring (HG, PRD, CC, SOG, WCVI) are surveyed using an aerial survey-fixed diver transect methodology (Fig. 1). The details of the herring spawning ground surveys can be found in Fort et al. (2013) and Grinnell et al. (in press). In brief, aerial surveys of the coastline of BC where spawning has historically occurred are observed daily beginning in early March. Spawning events are determined by observations of milt in the water, which causes a distinctive turquoise colour in the water near the shoreline. The position and length of coastline covered by spawning is confirmed by grapple, snorkel or short dive, and the confirmed position and length of the spawning bed is recorded. At a later date (usually within 5 d), a team of divers is deployed at fixed transects along the coastline where spawning occurred. The transects are assigned to a 'location' designation that is also a fixed section of shoreline. Locations are subsections of the spawning ground, each covering an ~1000–1500 m section of continuous shoreline. Within each location, fixed transects running perpendicular to the shore are spaced approximately 350 m apart. New transects have been added throughout the time series where new spawning locations have been identified (Grinnell et al. in press). For each location where spawning occurs, a starting date for each spawning event is recorded from the aerial survey or on-ground observations (Fort et al. 2013). For each transect, the diver first swims the transect to determine the width of the spawn (e.g. from the highest point on the beach where spawn is observed to the deepest point offshore where the diver observes spawn). Next, the diver records the number of layers of deposited spawn on either the seafloor or vegetation as well as the substrate type and vegetation type for 5–6 quadrats (0.5 m^2) evenly spaced along the transect within the spawn. The number of layers of spawn is recorded to the nearest 0.25 (e.g. complete coverage of the quadrat by 1 layer of eggs and ¼ coverage of the quadrat by 2 layers of eggs would be a layer of 1.25). The number of egg layers in the quadrat are later converted to density of eggs using standard conversions (Grinnell et al. in press) and expanded to abundance using the observed transect widths and shoreline lengths of spawn in a location. For this analysis, we did not use the abundance or substrate data, but only the presence or absence of spawn at individual transects. On average, historical transect lengths where spawn was observed were 112.7 m in length (median: 53.8 m; range: 1–2713 m) and minimum and maximum depths of observed spawn were 6 m above the water line and 32 m below the surface.

2.2. Modeling spawn timing for Pacific herring

2.2.1. Dependent variables

To model the timing of spawning events for Pacific herring in response to environmental conditions, the dependent variable was the presence or absence of the first spawning at a location in a given year on a given day of the year as indicated by the estimated

date of spawn (Grinnell et al. in press). Only year–location combinations with observed spawn were used in this analysis. There were 3115 combinations of location and year where spawn occurred from 1988–2018 (HG: 296; PRD: 537; CC: 671; SOG: 1173; WCVI: 438). For each of these presence observations, inferred (or assumed) absences were generated for those locations in 7 d intervals up to 5 wk prior to the spawning event. For example, a location with an observed first spawning (presence observation) on 15 March of a given year (day of the year 74) would have 5 associated absences at that location generated on Days 39, 46, 53, 60 and 67. Egg development begins in late summer and early fall when the herring may still be in summer feeding grounds and then continues as the fish migrate to overwintering areas (Hay 1985). Five weeks was chosen as the cut-off because this is when late stages of development occur, when the number of eggs is determined and near the time when the maturing adults are thought to be moving to the spawning grounds (Hay 1985). This procedure resulted in n= 15 575 absence observations for locations where spawn was eventually deposited.

2.2.2. Independent variables

The oceanographic variables used to explain the annual timing of Pacific herring spawning were obtained primarily from a hindcast simulation of the British Columbia continental margin (BCCM) model (Peña et al. 2019), which is an extension of Masson & Fine's (2012) implementation of the regional ocean modeling system (Haidvogel et al. 2008). The model has a horizontal resolution of 3 km and 42 terrain-following vertical layers of depth for all of BC. The model outputs used in this study are from a hindcast simulation for the 1988–2018 period derived from 3 d averages, so there were predictions of temperature, salinity and bottom currents near the seafloor for each grid point every 3 d from 1988–2018.

The oceanographic conditions near the ocean bottom used as predictor variables in these analyses were current speed, temperature and salinity, the change (measured as the trend of the variable over time) of temperature and salinity in the week prior to the presence or inferred absence observation and the change (measured by the slope of the variable over time) of temperature and salinity in the 4 wk prior to the observation (Table 1).

Oceanographic conditions were generated from the BCCM model output for each spawn deposition transect for each time step (date of spawning plus weekly time intervals up to 5 wk prior to the spawning event). The value for each variable from the closest BCCM model grid point in space and time was extracted to each spawn deposition transect and time step. The oceanographic data generated for each time step and transect were then averaged across the transects in a given location, giving a total number of independent observations of 18 690 (n = 3115 observations of locations of spawn presence and n = 15 575 inferred absence observations at locations where spawn was eventually deposited that year).

In addition, the number of degree days (days since 1 January of each year when the water temperature was above 5°C) was also included as an explanatory variable (Table 1). We chose 5°C as the reference temperature based on laboratory work that has found reduced egg hatching and viability below 5°C in Pacific herring (Alderdice & Velsen 1971) and Atlantic herring (Peck et al. 2012). The date of 1 January was chosen arbitrarily to reflect a time when herring are thought to have moved to inshore overwintering grounds and are in the last months of egg development (Hay 1985). The day length (photoperiod) was calculated for the date and latitude of each of the presence or inferred absence observations, using the 'geosphere' package in R (Hijmans 2019). In previous studies, lunar and tidal cycles have been found to impact spawn timing in Pacific herring (Hay 1990), so the lunar phase calculated from the 'lunar' package in R (Lazaridis 2022) was included as an explanatory variable in the analysis. The final variable used in the analysis was local wind speed, representing the potential occurrence of storms, which have been found to influence egg loss and mortality (Rooper et al. 1999, Moll et al. 2018). Hourly wind speeds were compiled from observations at 5 airports in BC (Sandspit, Prince Rupert, Bella Coola, Comox and Tofino) for the period 1988–2018, representing HG, PRD, CC, SOG and WCVI, respectively. Hourly data were obtained from Environment Canada (https://services.pacificclimate.org/met-data-portal-pcds/app/) and averaged for each day. In some cases (n = 127), there were gaps in the wind speed records; these gaps were filled in order of preference using data from the closest Environment Canada station within 50 km, the nearest BC Ministry of Forests, Lands, and Natural Resource Operations weather station or the average (March–May of the same year). The average was used in only 9 instances to fill gaps. Lunar phase and wind speed were calculated for each of the presence or inferred absence observations.

Table 1. Explanatory variables used to predict Pacific herring spawn timing and distribution in British Columbia from 1988–2018. Variables in **bold** were used for both the temporal and spatial analyses, while *italicized* variables were used only in the analysis of spatial distribution of spawn deposition. BCCM: British Columbia continental margin; ROMS: regional ocean modeling system

Variable	Units	Definition	Source
Temperature	°C	Near-bottom water temperature at the time of spawning	BCCM model (Peña et al. 2019)
Salinity	–	Near-bottom water salinity at the time of spawning	BCCM model (Peña et al. 2019)
Current speed	cm s^{-1}	Near-bottom current speed at the time of spawning	BCCM model (Peña et al. 2019)
Change in salinity and temperature (1 wk prior)	–	The slope of a linear regression of water temperature and salinity against date for the week prior to a spawning event. Temperature and salinity estimated from BCCM model output	BCCM model (Peña et al. 2019)
Change in salinity and temperature (5 wk prior)	–	The slope of a linear regression of water temperature and salinity against date for the 5 wk prior to a spawning event. Temperature and salinity estimated from BCCM model output	BCCM model (Peña et al. 2019)
Degree days	Number	Number of days per year with temperature above 5°C at the spawn location since 1 January	BCCM model (Peña et al. 2019)
Day length	Hours	Number of hours of daylight at the latitude of the spawn location	Hijmans (2019)
Lunar phase	Radians	Phase of the moon calculated at each date of presence or inferred absence of spawning activity	Lazaridis (2022)
Wind speed	km h^{-1}	Average daily wind speed at weather stations in each of the stock areas calculated from hourly data	Environment Canada and BC Ministry of Forests, Lands, and Natural Resource Operations [a]
Slope of bathymetry	%	Percent change of seafloor depth at each spawn deposition transect	Derived from Davies et al. (2020) using Hijmans (2021)
Vector ruggedness measure (*VRM*)	–	Value of the change in depth (or rugosity) at each spawn deposition transect	Derived from Davies et al. (2020) using Sappington et al. (2007)
Topographic position index (*TPI*)	–	Position of the midpoint of a spawn deposition transect relative to the surrounding bathymetry (positive values denote higher than the surrounding seafloor, while negative values are lower than the surrounding seafloor)	Derived from Davies et al. (2020) using Hijmans (2021)
Fetch	m	Cumulative distance to land surrounding the midpoint of the spawn deposition transects (a proxy for exposure to wave action)	Marchand & Gill (2018)
Alignment	degrees	Direction of the seafloor slope relative to the average current direction predicted by a ROMS for the 3 d prior to a spawning event	Derived from Davies et al. (2020) and Peña et al. (2019)
Current speed (*total*)	cm s^{-1}	Sum of the tidal current estimated using a tidal inversion program and the bottom current estimated using a ROMS averaged for the 3 d prior to a spawning event	Egbert & Erofeeva (2002), Peña et al. (2019)
Distance to spawn center	Degrees latitude and longitude	The distance of each transect from the median (most occupied) spawn deposition transect location for each stock across the time series (1988–2018)	Estimated from the spawn deposition survey data
Biomass anomaly	–	Annual biomass estimate for each stock normalized to mean = 0 and SD = 1	DFO (2020)

[a]https://services.pacificclimate.org/met-data-portal-pcds/app/

2.2.3. Modeling method

A random forest model was used to determine the effect of environmental variables on spawn timing (Breiman 2001, Cutler et al. 2007, Strobl et al. 2009). Random forest models are a classification-tree-based method where the data are bootstrapped, and many individual trees are fit to the bootstrapped data. A random selection of explanatory variables is chosen to consider at each split (branch) of the tree. Then, the predictions of the multiple resulting trees were combined into a single prediction by ensemble methods (Breiman 2001). The random forest modeling and evaluation were conducted using the 'randomForest' package in R (Liaw & Wiener 2002). This package predicts the response variable using non-parametric

ensemble methods to first construct multiple regression trees using recursive partitioning. For each node of a tree, a random subset (n = 3) of explanatory variables are used as candidates for splitting the data. A random bootstrap of the data is used to construct the models and held-back observations are used to cross-validate model performance. A minimum number of 1000 trees were computed in model fitting, and these were integrated into a final ensemble model.

The equation function included the full suite of explanatory variables so that:

Spawn presence or absence ~ temperature + salinity + current speed + change in temperature and salinity (1 wk and 5 wk prior) + degree days + day length + lunar phase + wind speed

Variable importance was used to sequentially reduce the number of predictor variables (Evans et al. 2011). The model was fit including the full set of explanatory variables, and variable importance was assessed using the permuted variable importance measure (the error increase when the variable is removed from the model). The least important variable was removed, the reduced model was refit and the prediction error was compared between the full and reduced models. This procedure was repeated until variable removal resulted in an increase in prediction error (i.e. a reduction in model performance).

Response plots from the random forest model were generated for each predictor variable by holding all variables (except the variable of interest) at their median values. A lowess smoother was fit to the expected values across the range of each predictor variable to visualize the relationship between explanatory variables in the model and the probability of spawning activity. The importance of variables that remained in the reduced model was assessed using the permuted variable importance measure (the error increase when the variable is removed from the model) as in Baker (2021); a decrease in model accuracy of >10% when a variable was removed indicates a substantial contribution of the variable to the model.

2.2.4. Model validation

Validation of the Pacific herring spawning model was approached in 2 ways. First, the model was fit as described above on a randomly selected 80% of data (training data) and then tested against the remaining 20% of the data in a standard model validation exercise. For this model validation, metrics of the model fit: root-mean-square error (RMSE), true-skill statistic, area under the receiver operating curve (AUC) and the difference between AUC for the training data and AUC for the testing data (AUC_{diff}) were all used to measure model performance. In addition to the randomly selected model validation set, the best-fitting overall model was also tested against the test data set for the individual Pacific herring stocks (HG, PRD, CC, SOG and WCVI). The same set of performance metrics were used in this model validation step, and the exercise was designed to ensure that the model was universally applicable to the 5 major Pacific herring stock areas in BC.

2.3. Modeling the spatial distribution of spawn for Pacific herring

2.3.1. Dependent variables

To model the distribution of spawning events for Pacific herring, all unique diver survey transects (n = 3851) were used. Spawning was recorded to have occurred at each of these transects at some point between 1988 and 2018 (hence the reason for their establishment); however, in a given year, Pacific herring may (presence) or may not (absence) have spawned at an individual transect. Over the history of the spawn surveys, transects have been added when spawning in new areas was observed (Grinnell et al. in press), with 14–16 transects added each year on average in HG, PRD and WCVI since 1988, 41 transects in CC and 35 transects in SOG. These transects have been designated as spawn absences in the years prior to their establishment for this analysis. In addition, there were some gaps in the survey data for some years and transects. For example, in 1988–1993, aerial and surface surveys of the first spawning of the WCVI stock were all completed, but only 77–92% of that spawn was surveyed using divers. For this analysis, years were removed for any of the stocks where the proportion of spawn not surveyed by divers was >1%, yet a spawning event was detected. This resulted in a sample size of 74 627 year–transect combinations, where presence was indicated by the observation of spawn at the transect during diver surveys and absence was indicated by transects that did not receive spawn (either through absence of spawn during aerial surveys of the location or observations of no spawn present on the transects during diver surveys). It was assumed that throughout the time-series, the aerial survey did not

miss any notable spawning events or locations, and if spawn events were missed they were missed at random and without a spatial or temporal bias.

The start and end longitude and latitude data for each spawn deposition transect (and all other geographical data, including the raster layers described in Section 2.3.2 below) were projected into Albers Equal Area Conic projection (center latitude: 45° N; center longitude: 126° W) for analysis. Mean underlying environmental data were extracted from the extent of the transects or from the closest point to the transect in the case of BCCM-derived variables.

2.3.2. Independent variables

The oceanographic conditions used as predictor variables in this analysis were the near-bottom water temperature and salinity at the time of the spawning event (Table 1). The value for each variable from the closest BCCM grid point in space and time was extracted for each spawn deposition transect and time of spawning.

Three variables derived from bathymetry data were also used as explanatory variables in the model of Pacific herring spawning. The depth (bathymetry) of nearshore waters of BC is available on a 20 × 20 m grid raster (Davies et al. 2019). The variables derived from the bathymetry layer were slope, the vector ruggedness measure (VRM) and the topographic position index (TPI; Table 1). Slope for each raster grid cell was computed as the maximum difference in angle (range: 0–90°) between the depth at a cell and its surrounding cells. VRM was calculated according to the methods of Sappington et al. (2007) and measures the roughness of the seafloor while accounting for differences in local slope of the seafloor. The TPI is the difference in bottom depth between the value of a cell and the mean value of the 8 surrounding cells, which indicates whether the cell was on a 'peak' or in a 'valley' relative to its surroundings. All terrain indices were calculated using the 'raster' package on the 20 × 20 m bathymetry layer in R (R Core Development Team 2019, Hijmans 2021).

Three additional variables were utilized as independent variables in the analyses: fetch, aspect relative to near-bottom current direction (referred to as alignment) and total current speed (Table 1). The fetch of each transect measured its potential exposure to wave action and was calculated for each transect location using the 'waver' package in R (Marchand & Gill 2018). The 'waver' package takes the bearings from a point (in this case, the mid-point of each spawn deposition transect) and estimates the distance to land in a specified set of directions. To implement this calculation, a shoreline was constructed from the bathymetry layer contour at depth = 0, and the distance to shore was calculated at 30° increments around a circle (i.e. 0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330°) and summed for the site as an index of exposure. The alignment variable captured the intersection of the direction the slope faced and the near-bottom mean current direction estimated at the mid-point of each transect. The aspect of the seafloor (i.e. angle the seafloor faces) in degrees relative to north (0°) was computed using the 'raster' package in R. The current direction used was the output from the BCCM model described above for February–May (since this period captures the expected period of incubation to post-hatch for the deposited eggs). The absolute value of the difference between the current direction and the aspect of the seafloor at each spawn deposition transect was calculated as a value ranging from 0° (where the mean current was flowing in the same direction the seafloor was facing) to 180° (where the mean current was flowing directly opposite the aspect of the seafloor). Finally, the total current at each spawn deposition survey location represented by the mid-point of the transect was calculated for the period from the time of the spawning event to 3 d following spawning. The mean of the hourly tidal current speeds was estimated for each site using a tidal inversion program (Egbert & Erofeeva 2002), and near-bottom current speeds were obtained from the BCCM model. These 2 components of the current speed were summed to create a total current speed variable.

Prior to initiating the spatial modeling, variables were examined for evidence of collinearity > 0.7, at which point correlation was likely to distort model estimation (Dormann et al. 2013) (see Fig. S1 in the Supplement at www.int-res.com/articles/suppl/m14274_supp.pdf) and variance inflation factors (VIFs) were calculated (Zuur et al. 2010). The statistics indicated that depth and slope were highly correlated (r = 0.75), and including both resulted in a maximum VIF of 3.5; thus, depth was not used further in the modeling (reducing the maximum VIF to 1.6).

2.3.3. Modeling method

A generalized additive model (GAM) was used to model the distribution of Pacific herring spawning (Wood 2006). The 8 independent variables described

above were included in the model representing mechanistic links to Pacific herring biology. In addition, a biomass anomaly variable with smoothing by each stock was used to account for differences in spawning biomass across years. The anomaly was calculated by subtracting the mean and dividing by the standard deviation of the biomass estimates from 1988–2018 (DFO 2020). A variable that measured the distance of each transect from the central position of spawning for the stock was also used to account for spatial autocorrelation within each stock. The central position of spawning was estimated as the median transect position (latitude and longitude) for each stock. It should be noted that this represented the geographical center of spawning, and there were often discontinuities (blank regions) between the center of spawning and the farthest point of spawning for any given year. Also of note is that the center of spawning may not have represented the region with the most abundant spawn, as the positions were not weighted by egg abundance; instead, this variable was used to capture the potential geographic spread of spawning events in a given year. Finally, an interaction term between biomass anomaly and distance from the center of spawning was included in the model (also smoothed within each stock). This interaction term was included to account for any basin effects (MacCall 1990), where at larger stock sizes the spawning locations might naturally expand from the central spawn location. The full model was:

Presence ~ temperature + salinity + slope + VRM + TPI + fetch + alignment + current speed + biomass anomaly (smoothed within each stock) + distance from center of spawning (smoothed within each stock) and biomass anomaly * distance from center of spawning (smoothed within each stock) + ε

The biomass anomaly and distance from the central spawning location variables represented extrinsic factors related to the size of the spawning stock, rather than the environment at spawning locations. The model assumed a binomial error structure (ε) and used a 'cloglog' link.

The GAM predicted the presence or absence of spawn at a given transect in a given year using a thin plate regression spline smoothing function to parameterize the relationship between predictor variables and the response (Wood 2006). The degrees of freedom used in the smoothing function were limited to ≤4 (except for the year and distance from central spawning location variables where the degrees of freedom were ≤10). Backward stepwise elimination was used to remove non-significant variables. Initially, a full model containing all independent variables was fit to the data. The least significant variable was then removed from the model, provided it had $p > 0.05$ and the unbiased risk estimator score was lower with the elimination of the variable. The resulting reduced model was re-fit to the data. The likelihood ratio test with a chi-squared distribution was used to compare the full and reduced models and determine if they were significantly different. Stepwise variable removal was continued until all variables were significant in the model or removal of additional variables resulted in significantly different models according to the likelihood ratio test. The remaining variables in the best model were determined to explain a significant portion of the variability in the location of Pacific herring spawn deposition.

2.3.4. Model validation

As with the temporal model validation, the Pacific herring spawn distribution model validation was approached 2 ways. First, the model was fit on a randomly selected subset (80%) of data (training data) and then tested against the remaining 20% using a standard model validation exercise. For this model validation, metrics of the model fit (RMSE, true-skill statistic, AUC, and AUC_{diff}) were all used to measure the model performance. In addition to the randomly selected model validation set, the best-fitting overall model was also tested against the individual Pacific herring stocks (HG, PRD, CC, SOG and WCVI) using the same performance metrics.

3. RESULTS

3.1. Modeling spawn timing for Pacific herring

3.1.1. Model results

The random forest model predicting the probability of spawn timing at locations in BC explained the maximum amount of variability when 5 variables were incorporated: day length, degree days (the number of days above 5°C), near-bottom salinity (at time of spawning), wind speed and lunar phase. The random forest model explained about 64% of the variability in spawn timing (Table 2). The AUC was 0.98 for the training data and the RMSE was 0.20. The true-skill statistic (0.86) indicated that the model was very useful in predicting spawn timing, based on

Table 2. Model fit statistics for a random forest model predicting the spawn timing (presence or absence of first spawning) for Pacific herring spawning at a location in a given year on a given day of the year in British Columbia. Goodness-of-fit statistics are provided for the model training data and the randomly selected testing data. Test statistics are also generated for individual herring stocks used in the modeling. AUC: area under the receiver operator curve; AUC_{diff}: difference in AUC between training and test data; RMSE: root-mean-square error; TSS: true-skill statistic

	Training data (out-of-bag sample)	Test data	Haida Gwaii	Prince Rupert district	Central Coast	Strait of Georgia	West coast Vancouver Island
n	14952	3738	366	609	778	1452	533
AUC	0.98	0.98	0.94	0.98	0.98	0.99	0.95
AUC_{diff}	–	0.00	0.04	0.00	0.00	–0.01	0.03
RMSE	0.202	0.195	0.211	0.204	0.182	0.180	0.230
TSS (threshold)	0.86 (0.26)	0.86 (0.23)	0.81 (0.28)	0.90 (0.38)	0.90 (0.24)	0.88 (0.23)	0.79 (0.19)
R^2 (Spearman)	0.64	0.60	0.53	0.61	0.61	0.62	0.56

a probability threshold (0.26) that maximized sensitivity and specificity.

The probability of a spawning event occurring on a given date increased sharply with day length >10.5 h to an asymptote at ~12 h (Fig. 2). There was also an asymptotic response of the probability of a spawning event to the number of degree days, at 100 d. A peak probability of spawning occurred at near-bottom salinity of 30.5 (at the time of spawning), but the probability of spawning decreased at salinities between 28 and 29.5 and then decreased with increasing salinities over 31 (Fig. 2). There were few

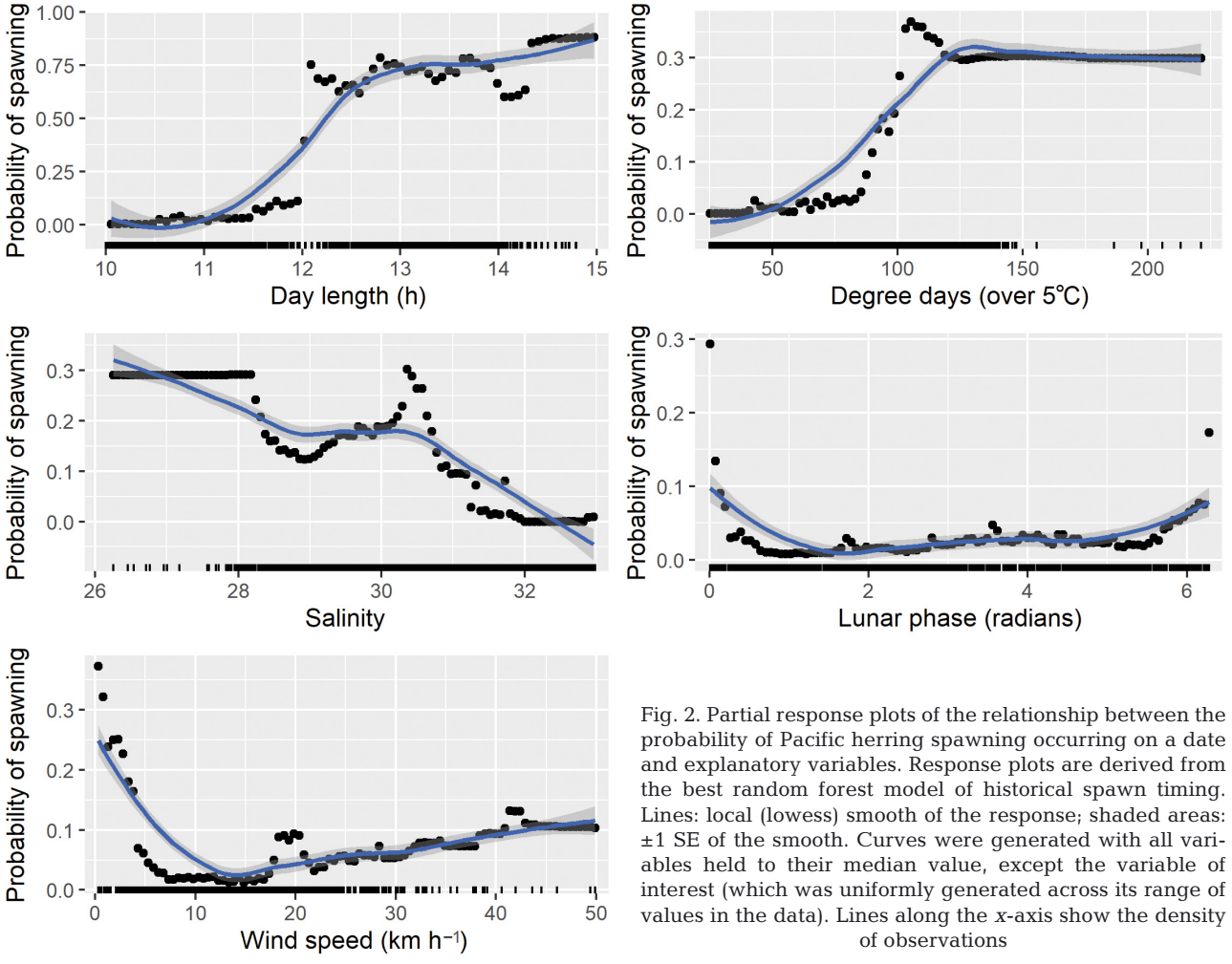


Fig. 2. Partial response plots of the relationship between the probability of Pacific herring spawning occurring on a date and explanatory variables. Response plots are derived from the best random forest model of historical spawn timing. Lines: local (lowess) smooth of the response; shaded areas: ±1 SE of the smooth. Curves were generated with all variables held to their median value, except the variable of interest (which was uniformly generated across its range of values in the data). Lines along the *x*-axis show the density of observations

observations at salinities less than 28. The probability of spawning was high during calm weather (indicated by average wind speeds near 0 km h^{-1}) and decreased as wind speeds increased to 12 km h^{-1} (Fig. 2). The probability of spawning then increased with increasing wind speeds above 12 km h^{-1}, although there were very few observations at wind speeds above 30 km h^{-1}. A clearer pattern was observed for the effect of lunar phase (Fig. 2), where peak probability of spawning was observed around the last quarter of the moon (lunar phase values >4.7) and the new moon (lunar phase: 0). Variable importance measures indicated that the spawn timing was explained primarily by a single variable, day length, which resulted in a decrease in model accuracy of 17.2% when it was removed. Removing degree days (the number of days above 5°C) and salinity at time of spawning resulted in decreases in model accuracy of 8.3 and 6.9%, respectively. The variable importance measures showed that the decrease in accuracy when wind speed and lunar phase were removed were 4.6 and 4.1%, indicating they were less helpful in explaining patterns of spawn timing.

3.1.2. Model validation

When the random forest model was tested against the remaining 20% of data held back from the model fitting, the results were similar to the training data (Table 2). There was a slight decrease in the correlation between predictions and observations, but the AUC for the test data was the same. There was a slight improvement in the RMSE, and the true-skill statistic was the same. When the model was tested against the different stocks, there were some notable changes; for example, the model did not perform as well for the HG and WCVI stocks, but the cross-validation statistics were all very high for individual stocks (Table 2).

3.2. Modeling the spatial distribution of spawn for Pacific herring

3.2.1. Model results

Model reduction for the GAM predicting the spatial distribution of Pacific herring spawn resulted in the retention of 5 environmental variables in the best-fitting model (Table 3). The relationship between the distribution of spawn and biomass anomaly, the distribution of spawn and distance to the center of spawning and the distribution of spawn and the interaction between biomass anomaly and distance to the center of spawning were also highly significant (except the WCVI and HG distance to the central spawning location). The GAM explained 20.1% of the deviation in the data set, and the AUC of the

Table 3. Summary of best-fitting generalized additive model explaining the spatial distribution of Pacific herring spawning (presence or absence of spawn deposition at dive survey transects) from 1988–2018. Distance: the distance of the transect to the central location of spawning for that stock; edf: estimated df for the covariate in the model

Species	edf	χ^2	p
Slope	2.93	117.91	<0.0001
Salinity	2.07	174.781	<0.0001
Temperature	2.85	11.933	0.010
Current speed	2.26	63.379	<0.0001
Fetch	2.74	10.275	0.024
Central Coast: distance	8.73	128.418	<0.0001
Haida Gwaii: distance	1.00	2.539	0.111
Prince Rupert District: distance	1.00	20.668	<0.0001
Strait of Georgia: distance	8.73	104.386	<0.0001
West Coast Vancouver Island: distance	1.00	1.232	0.267
Central Coast: biomass anomaly	9.00	277.442	<0.0001
Haida Gwaii: biomass anomaly	2.52	99.199	<0.0001
Prince Rupert District: biomass anomaly	4.04	28.519	<0.0001
Strait of Georgia: biomass anomaly	8.94	180.359	<0.0001
West Coast Vancouver Island: biomass anomaly	5.27	22.596	0.0004
Central Coast: biomass anomaly × distance	26.62	848.959	<0.0001
Haida Gwaii: biomass anomaly × distance	25.93	303.131	<0.0001
Prince Rupert District: biomass anomaly × distance	25.89	936.191	<0.0001
Strait of Georgia: biomass anomaly × distance	27.00	477.471	<0.0001
West Coast Vancouver Island: biomass anomaly × distance	26.93	406.466	<0.0001

model fit to the trained data set was 0.79 (SD = 0.002). The 3 near-bottom oceanographic variables (temperature, salinity and current speed) were significant in the model (Table 3). The probability of spawn occurrence increased linearly with salinity (Fig. 3). The effect was nonlinear for temperature, with a peak probability of spawn occurrence at ~8–10°C (Fig. 3). Transects with higher current speeds at the time of spawning also had a higher probability of occurrence of spawn. The probability of spawn occurrence generally decreased with increasing slope and increased with increasing fetch (a proxy for exposure to wave action; Fig. 3). The other effects related to the physical characteristics of the spawning location (VRM, alignment and TPI) did not add explanatory power and were removed from the final model.

The variables describing biomass anomaly (smoothed by stock), distance from the center of spawning (smoothed by stock) and the interaction of these 2 terms were also highly significant in most cases (Table 3). As the distance from the center of spawning increased, there was a significant decline in the probability of spawn presence across all levels of biomass for the SOG stock (see Fig. S2). However, the pattern was not as consistent for the other stocks. For the PRD, HG and WCVI stocks, there were multiple modes in the probability of spawning, where the probability of presence of spawn decreased farther from the center of spawning but then peaked again at multiple distances away from the center of spawning (see Fig. S2 in the Supplement). This effect was fairly consistent across all levels of biomass anomaly (Fig. 4). In HG, the probability of presence of spawn peaked ~15 and 65 km from the center of spawning. At higher biomasses, the spawn probability increased around 15 km from the center of spawning. The WCVI stock also exhibited a multimodal effect of probability of spawning, with peaks at ~30, 60 and 80 km away from the center of spawning (Fig. 4). Interestingly, for PRD and CC stocks, the probability of spawn at distances farther from the center of spawning was higher in years of relatively low biomass, exactly the opposite trend as would be expected if Pacific herring were expanding their spawning area in years when they were more abundant. The multiple modes of probability of spawning as the distance from the geographical center of spawning increased for PRD, HG and WCVI stocks are consistent with the

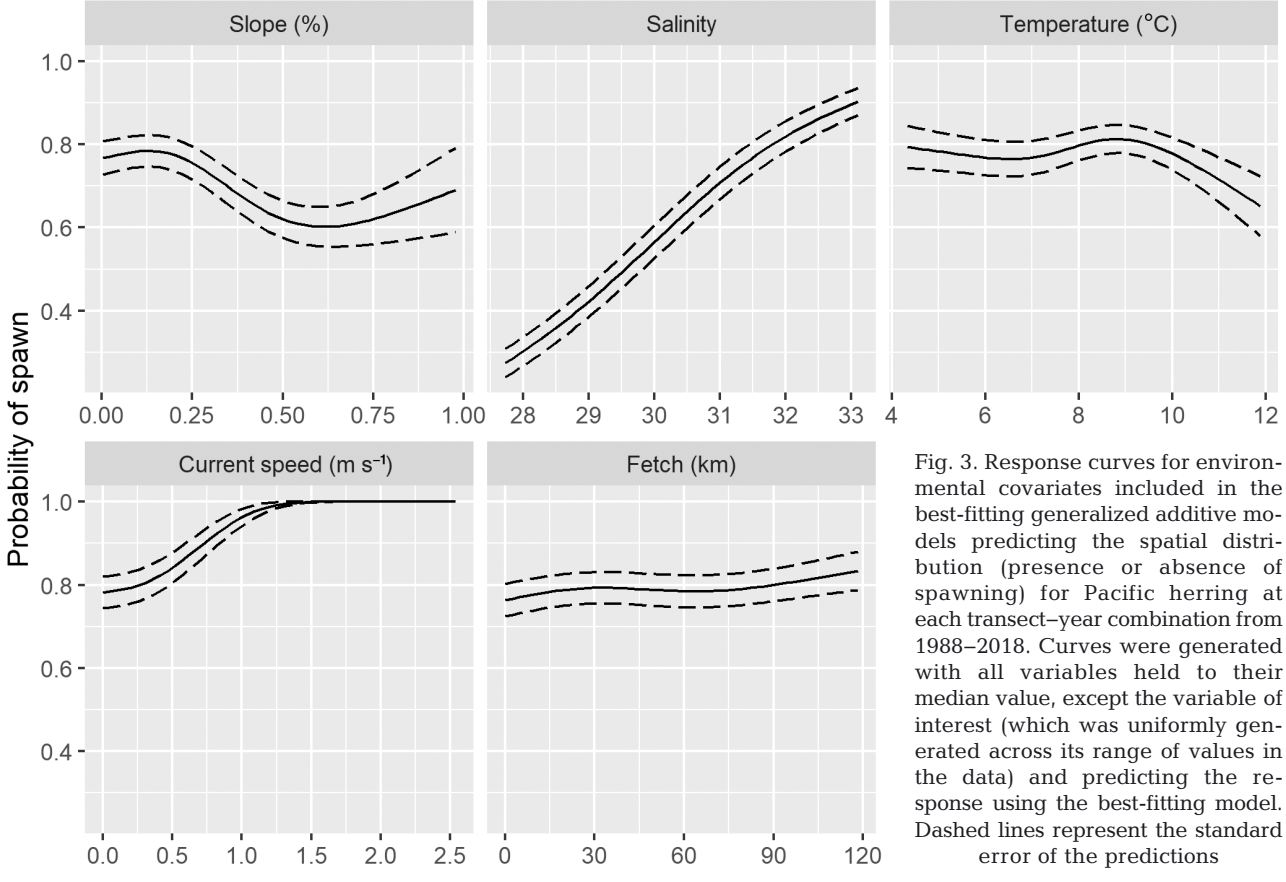
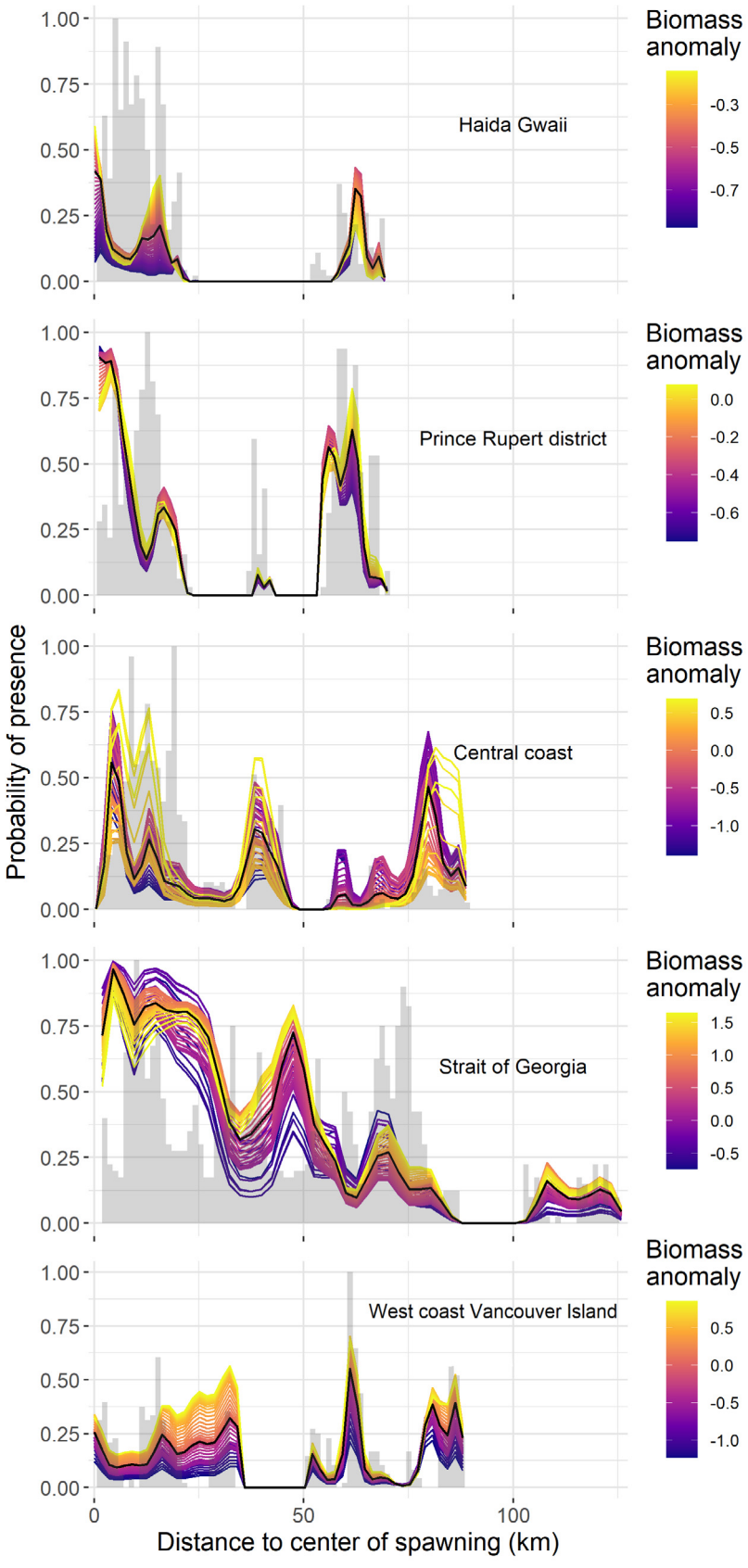


Fig. 3. Response curves for environmental covariates included in the best-fitting generalized additive models predicting the spatial distribution (presence or absence of spawning) for Pacific herring at each transect–year combination from 1988–2018. Curves were generated with all variables held to their median value, except the variable of interest (which was uniformly generated across its range of values in the data) and predicting the response using the best-fitting model. Dashed lines represent the standard error of the predictions



known multiple key spawning areas for these stocks (Grinnell et al. in press).

The biomass anomaly and distance to the center of spawning terms were the most important variables, measured by the drop in deviance explained by the model when these variables were removed (see Fig. S3). The removal of the interaction term resulted in the largest drop in deviance accounted for by the model but was still only ~4% of the total deviance. Removal of all environmental variables resulted in an ~1% drop in the deviance explained by the model.

3.2.2. Model validation

The model performed well at predicting the testing data, with AUC values as high for the testing data (AUC = 0.79) as the training data (Table 4). When compared to the individual stock components, the model performed reasonably well for the PRD, HG, SOG and CC test data sets (AUC_{diff} < 0.04), but the model performed marginally worse in predicting the presence or absence of spawning in WCVI (AUC_{diff} = 0.05). All model validation tests had AUCs > 0.70, indicating the model was useful in describing the distribution of spawn (Hosmer & Lemeshow 2013) and all exhibited true-skill statistics > 0.21, indicating model performance was better than random assignment. Contingency plots showed that the HG stock had a large proportion of false negatives predicted by the model (transects where spawning was observed but the model predicted absence) (Fig. 5).

Fig. 4. Response plots for the probability of Pacific herring spawn presence predicted for the interaction of the distance of each transect from the center spawning location and the biomass anomaly for each Pacific herring stock. Solid black line: median response; grey bars: frequency (scaled to 1) of the occurrence of transects at each distance from the center of spawning. Note that the range of the biomass anomaly (represented by the color scale) varies among Pacific herring stocks

Table 4. Model fit statistics for generalized additive model predicting the spawn spatial distribution (presence or absence of spawn at dive survey transects) for Pacific herring in British Columbia. Goodness-of-fit statistics are provided for the model training and randomly selected testing data. Test statistics are also generated for individual herring stocks used in the modeling. See Table 3 for acronym definitions

	Training data	Test data	Haida Gwaii	Prince Rupert district	Central Coast	Strait of Georgia	West coast Vancouver Island
n	59702	14925	2141	1978	4135	4579	2092
AUC	0.79	0.79	0.75	0.78	0.75	0.81	0.74
AUCdiff	–	0.00	0.04	0.01	0.04	−0.02	0.05
RMSE	0.37	0.37	0.31	0.39	0.38	0.40	0.34
TSS (threshold)	0.44 (0.21)	0.43	0.21	0.36	0.37	0.45	0.31
R^2 (Spearman)	0.43	0.43	0.28	0.43	0.35	0.49	0.30

4. DISCUSSION

From 1988–2018, there was a decreasing trend in the date of spawning for Pacific herring in 3 of the 5 stocks ranging from 2–6 d earlier (PRD, HG and WCVI). There has been no change in the average date of SOG spawning, and in the CC, the trend has been to a later spawn date (9 d later) over the time series. In this study, we modeled the relationships between the timing of Pacific herring spawning and

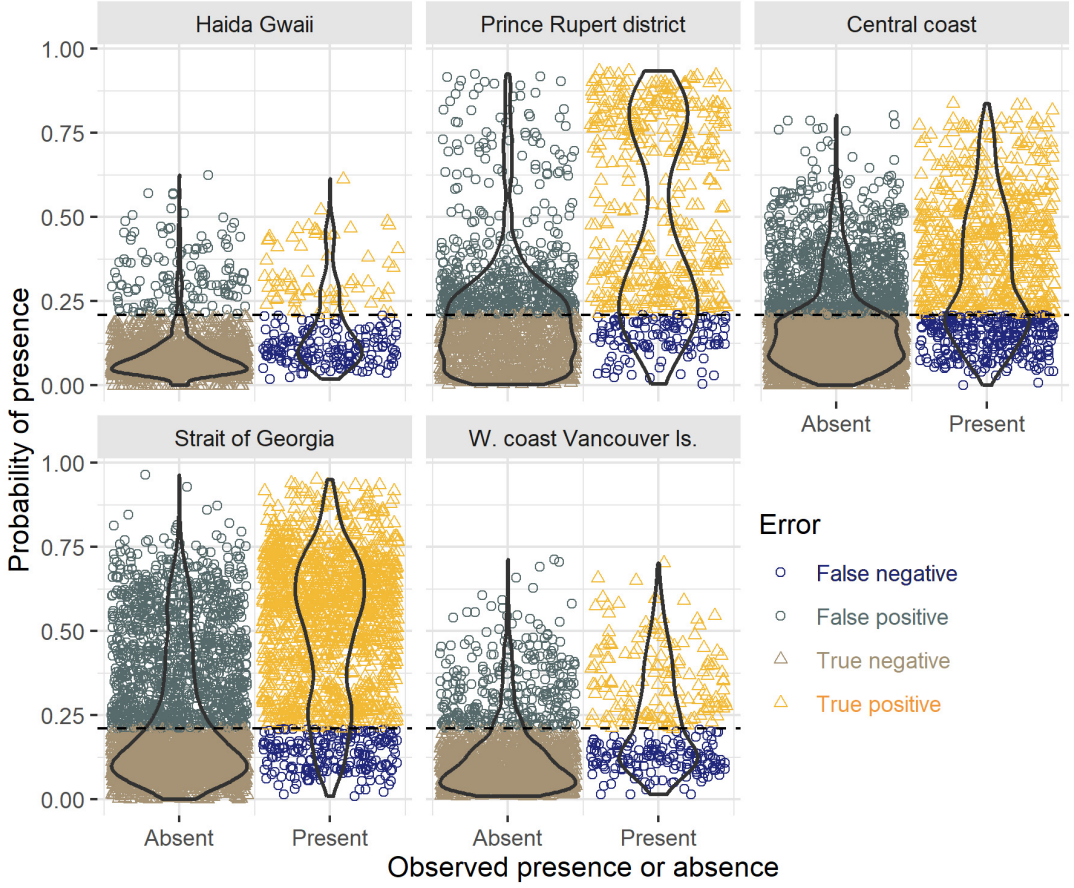


Fig. 5. Contingency matrix of predictions and observations of spawn presence or absence for the best-fitting generalized additive model applied to the test data set for Pacific herring spawn distribution (1988–2018). The predicted density of true positives (model predicted presence and presence was observed), true negatives (model predicted absence and absence was observed), false positives (model predicted presence, but absence was observed), and false negatives (model predicted absence but presence was observed) is shown relative to the threshold probability (dashed horizontal line) where presence is indicated (p = 0.21). Vertical dark lines (violin plot): continuous distribution of probability values

oceanographic variables and between the spatial pattern in deposition of spawn and the physical characteristics of the habitat. These novel analyses of 31 yr of spawn deposition data indicate that Pacific herring spawn timing is triggered by day length, degree days and near-bottom salinity. To the extent that these environmental cues coincide with the onset of the spring bloom and the most productive time of the year, this would tend to enhance year-class strength for Pacific herring. The relationship between spawn timing and the onset of the spring bloom forms the basis of the 'match–mismatch' hypothesis that links recruitment to the timing of larval hatch relative to the presence of suitable prey species noted for Norwegian spring spawning herring (Cushing 1969). Research has highlighted the importance of food availability at hatch with larval growth and survival across a variety of stocks (McGurk et al. 1992, Hufnagl & Peck 2011, Schweigert et al. 2013) There is some evidence that in BC, herring spawn just prior to the spring bloom, temporally aligning first-feeding larvae with their prey and resulting in higher prey availability, survival and recruitment (Schweigert et al. 2013, Boldt et al. 2019).

Recent studies have shown that there is a genetic basis for the relationship observed between day length and spawn timing (Petrou et al. 2021), helping to explain observations of genetic stock structure in Pacific herring. Links between spawn timing and genetic stock structure have also been observed for Atlantic herring (Lamichhaney et al. 2017). That study found that the genetic differences in spring or autumn spawning for Atlantic herring stocks showed consistent associations with the length of the day, which may explain the evolution of geographically distinct herring populations. Recent laboratory studies have also shown that by artificially manipulating day length, Atlantic herring can be induced to develop eggs out of phase with their natural cycle, indicating the strength of the photoperiod in determining spawning phenology (dos Santos Schmidt et al. 2022). However, Lamichhaney et al. (2017) also indicated that herring spawn timing shows some plasticity and that they could adjust spawning time according to water temperature. Our model results indicate that a day length of ~10.5 h is an important threshold beyond which the probability of spawning occurring is high. The good fit of the overall model to test data sets from individual areas suggested that this threshold in daylight hours was consistent across the 5 stocks of herring in BC. In part, then, the difference in spawn timing and stock structure from southern (e.g. SOG) to northern (e.g. PRD) BC is likely explained by the length of daylight hours in the 2 locations. However, both northern and southern BC reach day lengths of ~10.5 h within ~5 d of each other in late February, so attaining the threshold day length does not fully explain the timing of first spawning for herring, which tended to occur in March.

Many previous studies have also linked herring spawning activity to the effects of temperature and salinity, with later spawning occurring in cold years compared to warm years (e.g. Haegele & Schweigert 1985). This appears to also be true for spring spawning Atlantic herring (Messieh 1987, Polte et al. 2021, Ory et al. preprint doi:10.2139/ssrn.4164412). Laboratory and field studies have shown that although there can be a fairly wide range of temperatures and salinities over which successful Pacific herring spawning can occur (6.5–10°C and 22.4–28.7; Outram 1975, Hay et al. 1984), at temperatures below 4°C, eggs are unable to survive, and the maximum viable hatch occurs at 8.3°C and between 12 and 26 salinity (Alderdice & Velsen 1971). There have also been laboratory studies which assigned a developmental mechanism to the relationship with temperature, such as Brett & Solmie (1982), who found a relationship between gonadosomatic index and spawn timing that was mediated by temperature. This finding is consistent with the modeling results from our study, where the number of degree days was a stronger determinant of spawn timing than temperature on the spawning grounds. The modeling here indicates that when the number of degree days reaches 100, the probability of a spawning event occurring is greater. The relationships with individual variables except salinity were not important and did not have an impact on spawn timing. This is consistent with previous laboratory and field studies on Pacific herring that indicate there is plasticity in the actual oceanographic conditions where spawning can occur (Hay et al. 1984, Haegele & Schweigert 1985). In Atlantic herring, there is also considerable plasticity and adaptation to ambient temperature and salinity as can be seen in the different life history and maturity strategies employed by spring and autumn spawning stocks (van Damme et al. 2009, dos Santos-Schmidt et al. 2021). The importance of salinity in predicting spawn timing was interesting and may reflect not only the physiological impact of the variable on successful larval hatch but also may reflect environmental cues associated with homing behaviour (Bekkevold et al. 2005). The salinity differences among spawning grounds in BC reflect not only seasonal changes (through inputs of freshwater from winter runoff and precipitation) but also differ-

ences between relatively inshore areas (like the SOG) and areas closer to offshore waters (like the WCVI). Salinity patterns may thus play an important role in both the timing and the spatial distribution of spawning.

Model results were less clear about potential mechanisms influencing interannual variability in the spatial distribution of spawn. There was a large effect of spawner biomass anomalies and distance from the central spawning location as well as a significant but small effect of environmental conditions at the potential spawning locations. Removing all environmental variables from the model resulted in only a minimal (<2%) loss of predictive power. This is consistent with the findings that spawning can occur at a range of temperatures and salinities (Outram 1975, Hay et al. 1984, Haegele & Schweigert 1985). In past studies, variations in spawn intensity and deposition have been attributed to vegetation types (Haegele et al. 1981), herring biomass (Haegele & Schweigert 1985, Hay & Kronlund 1987), temperature and tides (Hay & Kronlund 1987). We were not able to address interannual differences in vegetation types at the different sites (because no data were collected in years where an individual transect was not surveyed due to the absence of herring spawn); however, the other variables were examined in the context of the modeling. The results from this study did not consistently show that years with larger herring biomass resulted in an increase in the overall probability of spawning at transects. This is shown by the trends for the SOG, PRD and CC (Fig. 4). This result was slightly surprising, as in years of higher herring biomass the overall probability of an individual transect receiving spawn would be expected to be higher.

The modeling results showed little evidence of a basin effect (MacCall 1990) in the spatial distribution of herring spawning. For the majority of stocks, there was no clear pattern that as the herring biomass increased the spawning expanded to surrounding locations (Fig. 4). For example, at higher biomasses, the probability of spawning generally decreased as the distance from the center of spawning increased for HG (Fig. 4). However, in the SOG, higher biomasses did result in a higher probability of spawning at greater distances from the center of spawning (Fig. 4). The absence of a consistent basin effect can also be seen in the kernel density of the distribution of transects relative to the center of spawning (see Fig. S4), where the distribution of transects with and without spawn did not change with distance from the center of spawning at different levels of biomass. The significance of the main effect of distance from the center of spawning in determining the spatial patterns of spawning activity was consistent for all stocks, indicating that some portions of the coast in each of the stock areas had an overall higher probability of receiving spawn, regardless of the size of the Pacific herring biomass, and that the capacity of these areas to receive spawn was not exceeded across the observed range of biomass (see Fig. S3). This accounts for the different modes in distance from the center of spawning in Fig. 4, where there are peaks in the probability of spawning across all levels of biomass. Historically, there have been longer-term shifts in spatial patterns of Pacific herring spawning distribution in BC and other regions. For example, Pacific herring in Prince William Sound showed spatial shifts in spawning, although the reasons for these shifts are not clear (McGowan et al. 2021). There does not appear to be strong evidence that Pacific herring spawning habitat is limited in its spatial extent (Shelton et al. 2014), which may in part explain some of the plasticity in spawn location over longer time periods. In the SOG, some of the areas where spawning activity occurred in the 1970s and early 1980s have not been utilized in more recent years (Hay & McCarter 1999, Hay & McCarter 2006). These changes were speculated to have been the result of the recovering population abundance at the time, although it is not clear what the exact mechanisms of this change were. Interestingly, the concentration of spawning activity in the SOG tended to occur in areas where the herring roe fishery was active and as the stock increased to higher biomasses (Hay & McCarter 1999, Hay et al. 2008). These shifts in spatial distribution occurred prior to the data examined in the current study (i.e. before 1988), so are not addressed by the modeling.

The modeling presented here indicated some potential caveats both in the methodology and the variables examined. The aerial surveys used for locating spawning have the potential to miss small spawning events that occur during nighttime hours or when poor weather prevents the surveys from occurring. However, spawning activity is also reported by fisheries officers, field biologists, commercial fishermen and the general public, when observed. This triggers a response by the dive survey team to investigate reported spawning activity. Implicit in our analysis here is the assumption that spawning events that did occur were observed and included in the data collection. There were undoubtedly some variables that were not included in the analysis because they were unavailable, such as veg-

etation type mentioned above (this section). Interannual variability in the distribution of vegetation could have accounted for some of the unexplained spatial patterns in the spawning data, such as the patterns in spawn occurrence with distance from the center of spawning. However, the spawn deposition data could not capture interannual changes in the distribution and abundance of vegetation data. In addition, the oceanographic variables were not measured at the site of spawning but were themselves model outputs from a regional ocean model (Peña et al. 2019). The regional ocean model used here had a 3 km resolution, which may not have been adequate to capture some of the transect-to-transect variability but likely was adequate for the analysis of spawn timing which used the location data that had a larger spatial scale. Both Masson & Fine (2012) and Peña et al. (2019) validated the BCCM against observed data collected in all areas of the continental shelf and coastal waters and found that the model reasonably captured water characteristics, in particular during winter months (Peña et al. 2019). The model resolution (3 km) is likely unable to fully resolve smaller features of the coastline but has been shown to capture spatial and temporal trends in the data well (Masson & Fine 2012, Peña et al. 2019). However, the separation of individual transects (~350 m) was much smaller than the grid size used by the BCCM, likely reducing the ability of the GAM to resolve a signal in the relationship. BCCM outputs have been previously used as covariates for modeling species distributions (Robinson et al. 2021, Thompson et al. 2023), ocean dispersal of larvae (Snauffer et al. 2014) and marine spatial planning activities (Blackford et al. 2021, Friesen et al. 2021) in marine waters of BC.

The use of the date of first spawning and the use of only the 5 major BC herring stocks meant that not all variability in both spawn timing and spawn distribution could be utilized. There are a number of locations that exhibit different spawning activity both in time and space than the major stocks examined here (Hay 1985). Some of these spawns occur as late as June and July and, given the potential for genetic differences based on differences in spawn timing, it is likely that they are responding to different cues in space and time. Some of the unexplained variability in the modeling might also have been due to differences in age structure in the fish populations; older fish, in both Atlantic and Pacific herring, have been observed to spawn earlier than their younger counterparts (Hay 1985, Lambert 1987, Arula et al. 2019, Dias et al. 2022). This phenomenon has been observed in other species of small pelagic fishes as well (e.g. Rogers & Dougherty 2019). These demographic differences may impact the pre-spawning behaviour of herring with similar mechanisms to those seen in Norwegian spring spawning herring (Huse et al. 2010), leading to changes in the spatial patterns of spawn distribution. Interestingly, the number of new spawn transects established in a year for a stock in this analysis was not related to the ratio of young (age 3 recruits) to older (ages 4+) fish ($p = 0.69$, $df = 154$) overall, or within any of the individual stocks, suggesting that younger, naïve fish were not likely to colonize new spawning areas. However, the pre-spawning migration patterns of herring have not been well studied relative to their age demographics, and it is important to note that individual spawning events are not sampled to determine the age structure of the spawners. Thus, it may be that the demographics of the spawning fish had an effect on our results that was undetectable with our data. There was unexplained variability, particularly in the analysis of the spatial patterns of spawn deposition; however, the cross-validation exercises showed that the models of spawn timing were able to predict the held-back data, both in tests against random subsets and tests against individual areas, indicating that the models were capturing the spatial and temporal patterns in herring spawning well across the region.

As nearshore waters warm in BC, the model presented here would predict that spawn timing would proceed earlier in the year, as the threshold of 100 degree days could occur earlier in the year. However, day length is constant across years, so given the genetic basis for the day length effect on spawn timing, it is unlikely that this threshold would shift. This will likely mean that the shift in spawn timing will be limited to no earlier than late February for most stocks in BC. In the SOG, the spring bloom usually occurs between day of the year 60 and 100 (Allen & Wolfe 2013), and recruitment of herring in the SOG is higher when spawning activity begins 1–3 wk prior to the spring bloom (Schweigert et al. 2013, Boldt et al. 2020). In recent years (since 2011), the SOG spring bloom has consistently occurred on about Day 85 (26 March), with little interannual variability (Allen et al. 2020). Temperature data from Entrance Island, BC (https://pacgis01.dfo-mpo.gc.ca/FGPPublic/BClightstations/BC_Lightstations_and_Other_Sample_Sites_V2.csv) shows that the average number of degree days on 6 March of each year (~3 wk before the spring bloom) has increased by about 5% during this time period. The earlier development of environmental conditions conducive to the onset of Pacific herring spawning may result in

spawning activity occurring earlier relative to the spring bloom, creating the potential to mismatch larval food supplies and potentially impact recruitment (Boldt et al. 2020). Future climate changes are likely to amplify these mismatches for BC herring stocks, with the potential to impact the productivity and sustainability of this important component of the marine ecosystem.

Acknowledgements. This paper was produced thanks to the support of many people, including those who have contributed to Pacific herring research and collection of data over the 40 plus year history of the DFO Spawn Surveys. We thank Kristen Daniel, Sarah Davies, Cole Fields, Jessica Nephin, Dana Haggarty and Jessica Finney for explaining and sharing their data. Isaac Fain prepared the forcing fields to update the regional ocean model. Jim Thorson, Sarah Power and Matt Baker provided helpful technical advice around the data analyses. Three anonymous reviewers provided very helpful and insightful comments and suggestions to improve the analysis.

LITERATURE CITED

Alderdice DF, Hourston AS (1985) Factors influencing development and survival of Pacific herring (*Clupea harengus pallasi*) eggs and larvae to beginning of exogenous feeding. Can J Fish Aquat Sci 42(Suppl 1):56–68

Alderdice DF, Velsen FPJ (1971) Some effects of salinity and temperature on early development of Pacific herring (*Clupea pallasi*). J Fish Res Board Can 28:1545–1562

Allen SE, Wolfe MA (2013) Hindcast of the timing of the spring phytoplankton bloom in the Strait of Georgia, 1968–2010. Prog Oceanogr 115:6–13

Allen S, Latornell D, Olson E (2020) Spring phytoplankton bloom timing, interannual summer productivity in the Strait of Georgia. Can Tech Rep Fish Aquat Sci 3377: 164–168

Arula T, Laur K, Simm M, Ojeveer H (2015) Dual impact of temperature on growth and mortality of marine fish larvae in a shallow estuarine habitat. Estuar Coast Shelf Sci 167:326–335

Arula T, Shpilev H, Raid T, Sepp E (2019) Thermal conditions and age structure determine the spawning regularities and condition of Baltic herring (*Clupea harengus membras*) in the NE of the Baltic Sea. PeerJ 7:e7345

Baker MR (2021) Contrast of warm and cold phases in the Bering Sea to understand spatial distributions of Arctic and sub-Arctic gadids. Polar Biol 44:1083–1105

Bekkevold D, André C, Dahlgren TG, Clausen LAW and others (2005) Environmental correlates of population differentiation in Atlantic herring. Evolution 59:2656–2668

Blackford C, Krkošek M, Fortin MJ (2021) A data-limited modeling approach for conserving connectivity in marine protected area networks. Mar Biol 168:86

Boldt JL, Thompson M, Rooper CN, Hay DE and others (2019) Bottom-up and top-down control of small pelagic forage fish: factors affecting juvenile herring in the Strait of Georgia, British Columbia. Mar Ecol Prog Ser 617-618: 53–66

Boldt JL, Javorski A, Chandler PC (eds) (2020) State of the physical, biological and selected fishery resources of Pacific Canadian marine ecosystems in 2019. Can Tech Rep Fish Aquat Sci 3377:1–288

Boldt JL, Murphy HM, Chamberland JM, Debertin A and others (2022) Canada's forage fish: an important but poorly understood component of marine ecosystems. Can J Fish Aquat Sci 79:1911–1933

Breiman L (2001) Random forests. Mach Learn 45:5–32

Brett JR, Solmie A (1982) Roe herring impoundment research—report on the 1980/81 studies. Can Tech Rep Fish Aquat Sci 1061:1–51

Cushing DH (1969) The regularity of the spawning season of some fishes. ICES J Mar Sci 33:81–92

Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler JJ (2007) Random forests for classification in ecology. Ecology 88:2783–2792

Davies SC, Gregr EJ, Lessard J, Bartier P, Wills P (2019) Coastal digital elevation models integrating ocean bathymetry and land topography for marine ecological analyses in Pacific Canadian waters. Can Tech Rep Fish Aquat Sci 3321:1–38

DFO (Fisheries and Oceans Canada) (2020) Stock status update with application of management procedures for Pacific herring (*Clupea pallasii*) in British Columbia: status in 2019 and forecast for 2020. Can Sci Advis Sec Sci Resp 2020/004

DFO (2021) Stock status update with application of management procedures for Pacific herring (*Clupea pallasii*) in British Columbia: status in 2020 and forecast for 2021. DFO Can Sci Advis Sec Sci Resp 2021/001

Dias BS, McGowan DW, Campbell R, Branch TA (2022) Influence of environmental and population factors on Prince William Sound herring spawning phenology. Mar Ecol Prog Ser 696:103–117

Dormann CF, Elith J, Bacher S, Buchmann C and others (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46

dos Santos Schmidt TC, Hay DE, Sundby S, Devine JA and others (2021) Adult body growth and reproductive investment vary markedly within and across Atlantic and Pacific herring: a meta-analysis and review of 26 stocks. Rev Fish Biol Fish 31:685–708

dos Santos Schmidt TC, Berg F, Folkvord A, Pires AMA, Komyakova V, Tiedemann M, Kjesbu OS (2022) Is it possible to photoperiod manipulate spawning time in planktivorous fish? A long-term experiment on Atlantic herring. J Exp Mar Biol Ecol 552:151737

Egbert GD, Erofeeva SY (2002) Efficient inverse modeling of barotropic ocean tides. J Atmos Ocean Technol 19: 183–204

Evans JS, Murphy MA, Holden ZA, Cushman SA (2011) Modeling species distribution and change using random forest. In: Drew AC, Wiersma Y, Huettmann F (eds) Predictive species and habitat modeling in landscape ecology. Springer, New York, NY, p 139–159

Fort C, Daniel K, Thompson M (2013) Herring spawn survey manual. Fisheries and Oceans Canada. https://www.pac.dfo-mpo.gc.ca/science/species-especes/pelagic-pelagique/herring-hareng/hertags/pdf/SurveyManual.pdf

Foy RJ, Paul AJ (1999) Winter feeding and changes in somatic energy content of age-0 Pacific herring in Prince William Sound, Alaska. Trans Am Fish Soc 128: 1193–1200

Friesen SK, Rubidge E, Martone R, Hunter KL, Peña MA, Ban NC (2021) Effects of changing ocean temperatures

on ecological connectivity among marine protected areas in northern British Columbia. Ocean Coast Manag 211: 105776

Grinnell MH, Schweigert JF, Thompson, M, Cleary JS (in press) Calculating the spawn index for Pacific herring (*Clupea pallasii*) in British Columbia, Canada. Can Tech Rep Fish Aquat Sci

Haegele CW, Schweigert JF (1985) Distribution and characteristics of herring spawning grounds and description of spawning behavior. Can J Fish Aquat Sci 42(Suppl 1): 39–55

Haegele CW, Humphreys RD, Hourston AS (1981) Distribution of eggs by depth and vegetation type in Pacific herring (*Clupea harengus pallasi*) spawnings in southern British Columbia. Can J Fish Aquat Sci 38:381–386

Haidvogel DB, Arango H, Budgell WP, Cornuelle BD and others (2008) Ocean forecasting in terrain-following coordinates: formulation and skill assessment of the Regional Ocean Modeling System. J Comput Phys 227:3595–3624

Hay DE (1985) Reproductive biology of Pacific herring (*Clupea harengus pallasi*). Can J Fish Aquat Sci 42(Suppl 1): 111–126

Hay DE (1990) Tidal influence on spawning time of Pacific herring (*Clupea harengus pallasi*). Can J Fish Aquat Sci 47:2390–2401

Hay DE, Kronlund AR (1987) Factors affecting the distribution, abundance, and measurement of Pacific herring (*Clupea harengus pallasi*) spawn. Can J Fish Aquat Sci 44:1181–1194

Hay DE, McCarter PB (1999) Distribution and timing of herring spawning in British Columbia. Can Stock Assess Sec Res Doc 99/14

Hay DE, McCarter PB (2006) Herring spawning areas of British Columbia. A review, geographic analysis and classification, rev. edn. Can Manuscr Rep Fish Aquat Sci 2019 rev edn. https://publications.gc.ca/collections/collection_2014/mpo-dfo/Fs97-4-2019-1-eng.pdf

Hay DE, Levings CD, Haney MJ (1984) Distribution of a herring fishery relative to submerged vegetation, herring spawn distribution and oceanographic factors. Can Manuscr Rep Fish Aquat Sci 1760:1–53

Hay DE, McCarter PB, Daniel KS (2008) Potential impacts of the British Columbia herring roe fishery on the spatial and temporal distribution of herring spawn: examination of the serial depletion hypothesis. Can Stock Assess Sec 2007/086

Hay DE, McCarter PB, Daniel KS, Schweigert JF (2009) Spatial diversity of Pacific herring (*Clupea pallasi*) spawning areas. ICES J Mar Sci 66:1662–1666

Hay DE, Schweigert J, Boldt JL, Thompson M (2019) Temporal changes in size-at-age: impacts and implications for reproductive biology, egg density and management of Pacific herring in British Columbia. Deep Sea Res II 159:42–51

Hijmans RJ (2019) geosphere: spherical trigonometry. R package version 1.5-10. https://CRAN.R-project.org/package=geosphere

Hijmans RJ (2021) raster: geographic data analysis and modeling. R package version 3.4-13. https://CRAN.R-project.org/package=raster

Hjort J (1914) Fluctuation in the great fisheries of northern Europe reviewed in the light of biological research. Rapp P-V Reùn Cons Int Explor Mer 20:1–228

Hosmer DW, Lemeshow S (2013) Applied logistic regression. Wiley, New York, NY

Houde ED (1987) Fish early life dynamics and recruitment variability. Am Fish Soc Symp 2:17–29

Hufnagl M, Peck MA (2011) Physiological individual-based modelling of larval Atlantic herring (*Clupea harengus*) foraging and growth: insights on climate-driven life-history scheduling. ICES J Mar Sci 68:1170–1188

Huse G, Fernö A, Holst JC (2010) Establishment of new wintering areas in herring co-occurs with peaks in the 'first time/repeat spawner' ratio. Mar Ecol Prog Ser 409: 189–198

Lambert TC (1987) Duration and intensity of spawning in herring *Clupea harengus* as related to the age structure of the mature population. Mar Ecol Prog Ser 39:209–220

Lamichhaney S, Fuentes-Pardo AP, Rafati N, Ryman N and others (2017) Parallel adaptive evolution of geographically distant herring populations on both sides of the North Atlantic Ocean. Proc Natl Acad Sci USA 114: E3452–E3461

Lazaridis E (2022) lunar: lunar phase & distance, seasons and other environmental factors. R package version 0.2-01. https://CRAN.R-project.org/package=lunar

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2:18–22

MacCall AD (1990) Dynamic geography of marine fish populations. University of Washington Press, Seattle, WA

Marchand P, Gill D (2018) waver: calculate fetch and wave energy. R package version 0.2.1. https://CRAN.R-project.org/package=waver

Masson D, Fine I (2012) Modeling seasonal to interannual ocean variability of coastal British Columbia. J Geophys Res 117:C10019

McGowan DW, Branch TA, Haught S, Scheuerell MD (2021) Multi-decadal shifts in the distribution and timing of Pacific herring (*Clupea pallasii*) spawning in Prince William Sound, Alaska. Can J Fish Aquat Sci 78: 1611–1627

McGurk MD, Warburton HD, Galbraith M, Kusser WC (1992) RNA–DNA ratio of herring and sand lance larvae from Port Moller, Alaska: comparison with prey concentration and temperature. Fish Oceanogr 1:193–207

Messieh SN (1987) Some characteristics of Atlantic herring (*Clupea harengus*) spawning in the southern Gulf of St. Lawrence. Northwest Atl Fish Organ Sci Counc Stud 11: 53–61

Moll D, Kotterba P, von Nordheim L, Polte P (2018) Storm-induced Atlantic herring (*Clupea harengus*) egg mortality in Baltic Sea inshore spawning areas. Estuaries Coasts 41:1–12

Moran JR, Heintz RA, Straley JM, Vollenweider JJ (2018) Regional variation in the intensity of humpback whale predation on Pacific herring in the Gulf of Alaska. Deep Sea Res II 147:187–195

Norcross BL, Brown ED, Foy RJ, Frandsen M and others (2001) A synthesis of the life history and ecology of juvenile Pacific herring in Prince William Sound, Alaska. Fish Oceanogr 10:42–57

Olesiuk PF, Bigg MA, Ellis GM, Crockford SJ, Wigen RJ (1990) An assessment of the feeding habits of harbour seals (*Phoca vitulina*) in the Strait of Georgia, British Columbia, based on scat analysis. Can Tech Rep Fish Aquat Sci 1730:1–135

Outram DN (1955) The development of the Pacific herring egg and its use in estimating age of spawn. Pacific Biological Station Circular No. 40, Fisheries Research Board of Canada, Nanaimo

Outram D (1975) Temperature and salinity levels associated with natural herring spawning in southern British Columbia. Can Manuscr Rep Fish Aquat Sci 1342:61

Pearsall IA, Fargo JJ (2007) Diet composition and habitat fidelity for groundfish assemblages in Hecate Strait, British Columbia. Can Tech Rep Fish Aquat Sci 2692: 1–149

Peck MA, Kanstinger P, Holste L, Martin M (2012) Thermal windows supporting survival of the earliest life stages of Baltic herring (*Clupea harengus*). ICES J Mar Sci 69: 529–536

Peck MA, Alheit J, Bertrand A, Catalán IA and others (2021) Small pelagic fish in the new millennium: a bottom-up view of global research effort. Prog Oceanogr 191: 102494

Peña MA, Fine I, Callendar W (2019) Interannual variability in primary production and shelf–offshore transport of nutrients along the northeast Pacific Ocean margin. Deep Sea Res II 169–170:104637

Petrou EL, Fuentes-Pardo AP, Rogers LA, Orobko M and others (2021) Functional genetic diversity in an exploited marine species and its relevance to fisheries management. Proc R Soc B 288:20202398

Polte P, Kotterba P, Hammer C, Gröhsler T (2014) Survival bottlenecks in the early ontogenesis of Atlantic herring (*Clupea harengus*, L.) in coastal lagoon spawning areas of the western Baltic Sea. ICES J Mar Sci 71:982–990

Polte P, Gröhsler T, Kotterba P, von Nordheim L and others (2021) Reduced reproductive success of western Baltic herring (*Clupea harengus*) as a response to warming winters. Front Mar Sci 8:589242

Purcell JE, Siferd TD, Marliave JB (1987) Vulnerability of larval herring (*Clupea harengus pallasi*) to capture by the jellyfish *Aequorea victoria*. Mar Biol 94:157–162

R Core Development Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Robinson CLK, Proudfoot B, Rooper CN, Bertram DF (2021) Comparison of spatial distribution models to predict subtidal burying habitat of the forage fish *Ammodytes personatus* in the Strait of Georgia, British Columbia, Canada. Aquat Conserv 31:2855–2869

Rogers LA, Dougherty AB (2019) Effects of climate and demography on reproductive phenology of a harvested marine fish population. Glob Change Biol 25:708–720

Rooper CN, Haldorson LJ (2000) Consumption of Pacific herring (*Clupea pallasi*) eggs by greenling (Hexagrammidae) in Prince William Sound, Alaska. Fish Bull 98: 655–659

Rooper CN, Haldorson LJ, Quinn TJ (1999) Habitat factors controlling Pacific herring (*Clupea pallasi*) egg loss in Prince William Sound, Alaska. Can J Fish Aquat Sci 56: 1133–1142

Sappington J, Longshore K, Thompson D (2007) Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. J Wildl Manag 71:1419–1426

Schweigert JF, Boldt JL, Flostrand L, Cleary JS (2010) A review of factors limiting recovery of Pacific herring stocks in Canada. ICES J Mar Sci 67:1903–1913

Schweigert JF, Thompson M, Fort C, Hay DE, Therriault TW, Brown LN (2013) Factors linking Pacific herring (*Clupea pallasi*) productivity and the spring plankton bloom in the Strait of Georgia, British Columbia, Canada. Prog Oceanogr 115:103–110

Sewall F, Norcross B, Mueter F, Heintz R (2018) Empirically based models of oceanographic and biological influences on Pacific herring recruitment in Prince William Sound. Deep Sea Res II 147:127–137

Shelton AO, Francis TB, Williams GD, Feist B, Stick K, Levin PS (2014) Habitat limitation and spatial variation in Pacific herring egg survival. Mar Ecol Prog Ser 514: 231–245

Snauffer EL, Masson D, Allen SE (2014) Modelling the dispersal of herring and hake larvae in the Strait of Georgia for the period 2007–2009. Fish Oceanogr 23:375–388

Stocker M, Noakes DJ (1988) Evaluating forecasting procedures for predicting Pacific herring (*Clupea harengus pallasi*) recruitment in British Columbia. Can J Fish Aquat Sci 45:928–935

Stocker M, Haist V, Fournier D (1985) Environmental variation and recruitment of Pacific herring (*Clupea harengus pallasi*) in the Strait of Georgia. Can J Fish Aquat Sci 42(Suppl 1):174–180

Strobl C, Malley J, Tutz G (2009) An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 14:323–348

Sullivan TM, Butler RW, Boyd WS (2002) Seasonal distribution of waterbirds in relation to spawning Pacific herring, *Clupea pallasi*, in the Strait of Georgia, British Columbia. Can Field Nat 116:366–370

Taylor FCH (1971) Variation in hatching success in Pacific herring (*Clupea pallasii*) eggs with water depth, temperature, salinity and egg mass thickness. Rapp P-V Reùn Cons Int Explor Mer 100:34–41

Thompson PL, Nephin J, Davies SC, Park AE and others (2023) Groundfish biodiversity change in northeast Pacific waters under projected warming and deoxygenation. Phil Trans R Soc B 378:20220191

Trites AW, Calkins DG, Winship AJ (2007) Diets of Steller sea lions (*Eumatopias jubatus*) in southeast Alaska, 1993–1999. Fish Bull 105:234–248

van Damme CJG, Dickey-Collas M, Rijnsdorp AD, Kjesbu OS (2009) Fecundity, atresia, and spawning strategies of Atlantic herring (*Clupea harengus*). Can J Fish Aquat Sci 66:2130–2141

Ware DM, McFarlane GA (1986) Relative impact of Pacific hake, sablefish and Pacific cod on west coast of Vancouver Island herring stocks. Bull Int N Pac Fish Comm 47: 67–77

Williams EH, Quinn TJ II (2000) Pacific herring, *Clupea pallasi*, recruitment in the Bering Sea and north-east Pacific Ocean. II: Relationships to environmental variables and implications for forecasting. Fish Oceanogr 9:300–315

Wood SN (2006) Generalized additive models: an introduction with R. CRC Press, Boca Raton, FL

Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. Methods Ecol Evol 1:3–14

Editorial responsibility: Myron Peck, Den Burg, The Netherlands
Reviewed by: P. Polte and 2 anonymous referees

Submitted: July 23, 2022
Accepted: February 21, 2023
Proofs received from author(s): April 21, 2023