Random forest classification to determine environmental drivers and forecast paralytic shellfish toxins in Southeast Alaska with high temporal resolution

John R. Harley[1*], Kari Lanphier[2], Esther Kennedy[2], Chris Whitehead[2], Allison Bidlack[1]

[1] *University of Alaska Southeast, Alaska Coastal Rainforest Center, 11066 Auke Lake Way, Juneau, AK, 99801*

[2] *Sitka Tribe of Alaska, Environmental Research Laboratory, 456 Katlian St, Sitka, AK, 99835*

**corresponding author*

Abstract

Paralytic shellfish poison toxins (PSTs) produced by the dinoflagellate in the genus *Alexandrium* are a threat to human health and subsistence lifestyles in Southeast Alaska. It is important to understand the drivers of *Alexandrium* blooms to inform shellfish management and aquaculture, as well as to predict trends of PST in a changing climate. In this study, we aggregate environmental data sets from multiple agencies and tribal partners to model and predict concentrations of PSTs in Southeast Alaska from 2016 to 2019. We used daily PST concentrations interpolated from regularly sampled blue mussels (*Mytilus trossulus*) analyzed for total PSTs using a receptor binding assay. We then created random forest models to classify shellfish above and below a threshold of toxicity (80µg $100g^{-1}$) and used two methods to determine variable importance. We obtained a multivariate model with key variables being sea surface temperature, salinity, freshwater discharge, and air temperature. We then used a similar model trained using lagged environmental variables to hindcast out-of-sample (OOS) shellfish toxicities during April-October in 2017, 2018, and 2019. Hindcast OOS accuracies were low (37-50%); however, we found forecasting using environmental variables may be useful in predicting the timing of early summer blooms. This study reinforces the efficacy of machine learning to determine important drivers of harmful algal blooms, although more complex models incorporating other parameters such as toxicokinetics are likely needed for accurate regional forecasts.

1. Introduction

Harmful algal blooms (HABs) are a significant and increasing threat to coastal communities and those dependent on marine resources (Glibert et al., 2018). Proliferation of single celled phytoplankton and cyanobacteria can lead to marine fouling, physical irritation, or the production of biotoxins that can bioaccumulate and cause mortality in aquatic organisms and humans (James et al., 2010; Moore et al., 2008). Of particular concern with respect to human and wildlife health are blooms that produce paralytic shellfish poisoning toxins (PSTs) which include saxitoxin (STX) and a number of analogues including gonyautoxins and neosaxitoxins that can cause paralytic shellfish poisoning (PSP) (Anderson et al., 2012). In the northeastern Pacific, the most common producer of PSTs are dinoflagellates in the genus *Alexandrium* (Anderson et al., 2012). Deaths and illness from shellfish toxins have been documented in the Northeast Pacific since 1793 when four members of Captain George Vancouver's expedition became sick after consuming mussels in an area now known as Poison Cove in British Columbia (Hallegraeff, 1993).

In Alaska there have been more than 240 reported cases of PSP since 1973, with the majority occurring in Southeast Alaska (Castrodale, 2015; Gessner and Middaugh, 1995). Shellfish are an important subsistence diet item for many Native Alaskans and tribes in the region have been cultivating shellfish gardens for at least 10,000 years (Jackley et al., 2016; Newton and Moss, 1984). Even as the incidence of PSP in Alaska is high compared to other regions (Gessner and Middaugh, 1995), the reported rates likely represent an underestimation of the true incidence of PSP due to the remoteness of communities in Southeast Alaska as well as variability in the severity of symptoms (Knaack et al., 2016).

Shellfish monitoring programs in areas with high recreational and subsistence harvests of shellfish have increased awareness and reduced risks of illness and fatalities due to PST (Etheridge, 2010). Programs in the states of Washington and Maine have been established for several decades and routinely provide beach closures and openings for recreational harvesters (Bean et al., 2005; Trainer et al., 2003). Despite growing concern in Alaska over marine biotoxins there is no state-sponsored monitoring

program for shellfish toxins or harmful algal blooms. As a result the Southeast Alaska Tribal Ocean Research (SEATOR) network was created to monitor HABs and minimize risks associated with shellfish harvests in Southeast Alaska (Harley et al., 2020). Since 2016, this program has analyzed over 1,700 shellfish samples, and currently monitors phytoplankton and shellfish PST concentrations in more than 17 communities. The initiation of a tribally-driven shellfish monitoring program allows subsistence and recreational harvesters to make more informed choices about when and where to harvest, and provides critical data for elucidating environmental drivers of HABs in Southeast Alaska.

One of the main challenges of HAB research is determining physical and chemical drivers of blooms which may be ephemeral and region specific (Wells et al., 2020). Due to the complex coastline and remoteness of Southeast Alaska, several aspects of physical, chemical, and biological oceanography are poorly characterized (Weingartner et al., 2009). Physiographic features (i.e., glaciers and fjords), large tidal ranges (6-8 meters), and highly seasonal meteorology and freshwater input drive complex oceanographic patterns (Weingartner et al., 2009). Accordingly, HABs in Southeast Alaska might be triggered by multiple environmental factors which are not present (e.g., glacial discharge) or less important (e.g., tidal flux) for HAB development in other coastal regions such as Oregon or Washington.

Several efforts have been made to model and predict spatiotemporal dynamics of *Alexandrium* and PSTs using environmental variables (Finnis et al., 2017; Laabir et al., 2013; Moore et al., 2009; Valbi et al., 2019). For instance, Moore et al. (2009) examined correlations between PST concentrations in blue mussels (*Mytilus edulis*) and sea surface temperature (SST), salinity, air temperature, precipitation, freshwater discharge, tidal flux, upwelling indices, and wind speed in Puget Sound, Washington. They found evidence that SST, freshwater discharge, air temperature, and upwelling contributed significantly to PST dynamics. More recently, Finnis et al. (2017) used machine learning algorithms to model monthly averaged PST dynamics from *Mytilus* spp. around Vancouver Island, British Columbia and found similar influence of SST, air temperature, salinity, discharge, upwelling, and photosynthetically active radiation (PAR). Both Finnis et al. (2017) and Moore et al. (2009) found evidence that warmer, drier conditions in

summer were associated with increased PST concentrations, but both studies found some high PST concentrations during colder winter months (i.e., December and January).

Most previous attempts at modelling environmental PST dynamics examined toxin concentrations on monthly or longer time scales. There are several challenges associated with modelling daily PST concentrations in the environment, foremost of which is that most state sponsored shellfish monitoring programs sample sites weekly or biweekly at most. Modelling using monthly or annual averages might give insight into regional long-term trends of PST (Moore et al., 2009); however, since blooms of *Alexandrium* can respond rapidly to physical and chemical conditions, these models might miss more ephemeral aspects of environmental forcing. For instance, *Alexandrium* blooms often occur when a brief period of wind-driven mixing (which can result in the resuspension of cysts resting in sediment) is followed by a short period (i.e., 1 week) of relatively calm winds, which increases water column stratification and allows for the germination of the cysts (Laanaia et al., 2013; Moore et al., 2009; Tobin et al., 2019). Models examining environmental drivers based on smaller time scales are therefore necessary and can aid in our understanding of HAB dynamics as well as be used in short-term predictive forecasts based on meteorological and oceanic conditions (Laanaia et al., 2013).

Forecasting HABs has progressed rapidly in the past decade, with operational forecasts now available for several species and areas (recently summarized by Franks, 2018). Some predictive models rely on HAB species cell counts (i.e., Gulf of Mexico HAB-OFS) or other metrics requiring human observers or technicians. However a lingering question is whether machine learning techniques have the capability of modelling and forecasting HABs using environmental data collected via automated processes (Finnis et al., 2017; Valbi et al., 2019). Environmental forecasts contain their own uncertainties (e.g., weather forecasts) and are unavailable for some variables (e.g., salinity); however, since the endpoint of interest for human and wildlife health (PST concentrations) lags the environmental conditions that lead to a bloom of *Alexandrium*, we can potentially leverage current observations of environmental conditions to forecast PST concentrations rather than relying on forecasted environmental data.

104 In this study we evaluate the efficacy of machine learning algorithms to determine environmental
105 drivers of PST dynamics in Southeast Alaska and hindcast PST concentrations in shellfish. We used daily
106 PST concentrations interpolated from regular sampling of blue mussels (*Mytilus trossulus*, Reese et al.,
107 2012) from 2016-2019 by the SEATOR network. We evaluate the ability of the model to hindcast out-of-
108 sample (OOS) observations using lagged environmental variables, a model which could be implemented
109 using current environmental observations to provide short-term forecasts. The value of these models will
110 not only increase our understanding of HAB dynamics in a complex oceanographic system, but also
111 potentially allow for shellfish safety advisories for communities without regular shellfish testing.

112 **2. Methods**

113 *2.1 Blue mussel samples*

114 The SEATOR network consists of 17 tribal partners that conduct weekly phytoplankton net tows
115 and biweekly shellfish sampling (see Harley et al., 2020). During each shellfish sampling effort
116 approximately 20-100 blue mussel samples are collected and shipped to the Sitka Tribe of Alaska
117 Environmental Research Lab (STAERL) for analysis. The process of sample processing and analysis is
118 described in more depth in Harley et. al (2020). Briefly, whole blue mussels are shucked and whole body
119 contents are collected until 100 grams of tissue is obtained. Samples are homogenized using an immersion
120 blender and extracted using the HCl method as described in (Van Dolah et al., 2009). Supernatant from
121 the acid extraction is then analyzed using the receptor binding assay (RBA) method (Van Dolah et al.,
122 2009, 2012). Concentrations of PSTs are reported in µg STX equivalent per 100g tissue (henceforth µg
123 $100g^{-1}$).

124 *2.2 Environmental Data*

125 Sea surface temperature (SST) and Photosynthetically active radiation (PAR)

126

127 Daily SST values (ºC) were obtained from the National Aeronautics and Space Association and
128 Jet Propulsion Lab (NASA/JPL) Multi-scale Ultra-high Resolution (MUR) SST analysis (Version 4.1,
129 0.01º resolution) which incorporates imaging data from multiple sources as part of the Group for High
130 Resolution Sea Surface Temperature (GHRSST) project. Values from the underlying pixel for each
131 sample site and date were extracted. If no value was present for that pixel, values were averaged from a
132 bounding box containing the nearest tiles surrounding the sample site (0.03 º by 0.03º or 9 pixels total).
133 The MUR SST utilizes multiple instruments to detect SST including microwave sensors which are less
134 prone to interference from clouds which frequently cover Southeast Alaska (Chin et al., 2017).

135 Daily PAR values (μE $m^{-2}s^{-1}$, where 1E = 1 mole photons) were obtained from NASA's Visible
136 and Infrared Imager/Radiometer Suite (VIIRS) aboard the Suomi-NPP satellite (4km resolution). Similar
137 to SST, values were extracted from the underlying pixel unless no value was present, in which case values
138 were averaged over a bounding box of the nearest tiles surrounding the sample site (12km by 12km or 9
139 pixels total). Both SST and PAR datasets were accessed through the NOAA ERDDAP servers (Simons,
140 2019, dataset IDs: *jplMURSST41* and *erdVH2018par1day*).

141 Meteorological Data

142 Air temperature (ºC), precipitation (mm), and wind speed (m s^{-1}) were obtained from the National
143 Centers for Environmental Information using the Global Historical Climatology Network dataset and the
144 R package *rnoaa* (Chamberlain, 2019). Sample sites were associated with the nearest weather station; all
145 sample sites had stations within 100km. Precipitation was measured as 24-hour accumulation total of all
146 precipitation types (snow is included as water equivalent). Wind speed and air temperature were averaged
147 (mean) for the 24-hour period.

148 Freshwater discharge

149 Discharge (m^3 s^{-1}) values were obtained from stream gauge locations from the USGS with the R
150 package *dataRetrieval* (De Cicco et al., 2018). Stream gauges were selected from drainages with > $5km^2$

151 area. Since the scale of gauged streams in Alaska varies by several orders of magnitude, discharge values
152 were scaled (unitless) for each individual gauged stream to enable comparisons between sites.

153 Tidal range

154 Southeast Alaska has mixed tides, with each tidal cycle having a high-high (HH), low-high, high-
155 low, and a low-low tide (LL). Maximum tidal range (HH – LL for a 24-hour period) was calculated using
156 buoy data aggregated by NOAA (Chamberlain, 2019). 24-hour periods that did not contain the maximum
157 tidal range (i.e. did not have a HH and LL) were interpolated from the previous and subsequent day's tidal
158 data.

159 Upwelling

160 Indices of upwelling were accessed from NOAA's Pacific Fisheries Environmental Laboratory
161 (PFEL) which provides daily estimates for 15 locations in the northeastern Pacific. PFEL calculates daily
162 indices based on surface pressure fields (a proxy for wind stress), where positive values indicate
163 upwelling while negative values indicate downwelling. PST sample sites were spatially matched to the
164 nearest PFEL upwelling site.

165 Salinity

166 Salinity (unitless) was measured approximately weekly at each sample site by SEATOR partners
167 using a hand-held refractometer. Due to variations in surface water salinity from precipitation (freshwater
168 lensing) as well as variability due to human and refractometer imprecision (Millero, 2013) salinity values
169 were processed to minimize the effects of outlying values. Salinity values were calculated as a 14-day
170 rolling mean from observed refractometer data for each sample site.

171 In addition to environmental data we included a season variable (winter, spring, summer, fall)
172 based on Julian day which was included to tease apart variation in spring and fall blooms. The following
173 delineations were used for each year: winter: December 21 – March 21, spring: March 22 – June 21,

summer: June 22-September 22, and fall: September 23 – December 20. The locations of environmental sample stations are presented in Figure S1.

2.3 Missing data handling and processing

PST concentrations were linearly interpolated on a daily scale to fill in gaps of less than one month between samples. Gaps of more than 1 month were not filled with any PST values, thus only observed PST values were incorporated into the model from sites with inconsistent sampling. PST concentrations were then classified as either above or below the action threshold of 80µg $100g^{-1}$ tissue set by the US Food and Drug Administration (FDA) (Finnis et al., 2017). For all environmental variables, missing values were filled via interpolation based on the Kalman filter using the R package *imputeTS* (Harvey, 1990; Moritz and Bartz-Beielstein, 2017) after which further missing values were filled using the mean of the three nearest sites.

2.4 Random forest modeling

We constructed random forest models to classify a binary toxicity outcome of PST concentrations above or below 80µg $100g^{-1}$ shellfish tissue using the R package *randomForest* (Liaw and Wiener, 2002). Since the classes were unbalanced (86% of samples were below 80µg $100g^{-1}$) we utilized random over sampling examples using the R package *ROSE* (Menardi and Torelli, 2014) to bootstrap resample until we achieved balanced classes. We randomly subsampled 70% of the balanced data set (~7200 observations including interpolated data) as a training data set and evaluated the performance of the models using the remaining observations. Confusion matrices were utilized to calculate specificity and sensitivity of candidate models, and receiver operating characteristics (ROC) curves were used to evaluate model performance.

The optimal number of randomly sampled features at each node was determined by minimizing the out of bag (OOB) error rate via repeated cross validation (CV). Further hyperparameter tuning is presented in the supplemental materials (Figure S2). The importance of each variable was determined

using the mean decrease in accuracy (MDA) which is determined by randomly permuting each feature individually and assessing the increase in error of the overall model. We also report the mean decrease in Gini (Gini, i.e., decrease in node impurity) which has been shown to be more robust to small perturbations in the data (Calle and Urrea, 2011). Relative MDA and Gini were calculated for ease of graphical interpretation by scaling to the highest value reported for each metric (scale of 0 to 1).

Partial effects plots were examined in the form of accumulated local effect (ALE) plots to visualize model dependencies across the distribution of each variable. ALEs have been shown to be capable of more accurately assessing non-independent variables, which is desirable for this dataset since a number of variables (e.g. PAR and air temperature) are highly correlated (Molnar, 2020).

2.5 Lagging environmental data and OOS hindcasting

To test the accuracy of potential forecast models, we first lagged all environmental variables with respect to PST concentrations. We did not evaluate optimal lags for each variable independently since the efficiency of these variables as predictors is low in single variable models and would be an inappropriate for this system. We then we created random forest models as described above for non-lagged data, holding constant the number of variables tried at each node and number of trees in each forest for comparative purposes. Following model pruning and OOB validation we selected the model with the highest accuracy predicting samples > 80µg $100g^{-1}$ (lowest type II error rate). (false negative) to hindcast OOS PST concentrations from April to October in 2017, 2018, and 2019. We analyzed probabilities (either above or below 80µg $100g^{-1}$) which calculates the proportion of votes from each tree in the forest (Figure 5). By assigning classes to probabilities using a cutoff of 0.5 (i.e., probabilities > 0.5 classified as above 80µg $100g^{-1}$), we calculated accuracies of summer hindcasts compared to actual PST concentrations. These 3 years were selected from the larger dataset because consistent sampling from multiple sample sites was not available until after the spring of 2016 (only 36 samples collected in April-June 2016 as opposed to > 100 in each of 2017, 2018, and 2019).

From the OOS hindcasts we also obtained the earliest date for each year and sample site when the model predicted a greater than 50% likelihood of PST concentrations > 80µg $100g^{-1}$. Using daily interpolated PST concentrations from these sample sites, we then calculated the difference in the predicted date of PST concentrations > 80µg $100g^{-1}$ and the actual date in which PST concentrations crossed that threshold. The difference between the predicted date of the spring/summer bloom and the actual date of the bloom was used as a metric to evaluate the efficacy of the model at predicting the timing of the initial bloom each year at the different sample locations. *Post hoc*, we examined differences in environmental conditions between these three years by averaging environmental variables for April, May and June at each sample site and conducting a principal component analysis (PCA).

General data analysis and plots were done in the R programming environment (R Core Team, 2019) using the *tidyverse* packages (Wickham et al., 2019).

3. Results

Since sampling began in 2016 the STAERL network has tested 986 blue mussel samples from 40 sample sites across Southeast Alaska (Figure 1). PST concentrations from samples ranged from below detection to over 4,400 µg $100g^{-1}$. Monthly mean PST concentrations are presented in Figure 2. Concentrations exceeded 80µg $100g^{-1}$ in 14% of samples, with several sample sites experiencing one or two blooms per year (see Figure 5).

3.1 Variable importance

The full random forest model generated using all features resulted in a balanced accuracy (accuracy averaged across both classes) of 79%, and the ROC for the full model shows good sensitivity (81%, specificity = 73%, Figure 3a). The calculated area under the curve (AUC) of 0.86 for the full model indicates excellent model performance (Hosmer and Lemeshow, 2004). Variable importance measures MDA and Gini showed that SST, air temperature, season, and salinity were among the most important features of the model. By excluding variables with relative MDA and Gini < 0.5 (Figure 3b), we created a

246 partial model which included SST, salinity, air temperature, discharge, season, and PAR which had a
247 balanced accuracy of 74% and an AUC 0.83 (Figure 3a). Sensitivity (true positive rate) for the partial
248 model was 79% and specificity (true negative rate) was 69%.

249 *3.2 Accumulated local effects (ALE)*

250 ALE plots (Figure 4) illustrate the marginal effects of each variable across its distribution.
251 Increasing probabilities of PSTs > 80µg $100g^{-1}$ were associated with warmer air temperatures, high PAR,
252 higher wind speeds, and lower discharge. SST and salinity display maximum marginal effects at
253 approximately 12ºC and 27 respectively.

254 *3.3 Summer hindcasts*

255 By lagging PST concentrations 7 days we created a lagged model with the lowest type II error
256 rate (Figure S3) to hindcast OOS summer concentrations of PSTs. Overall accuracies were 43% (2017),
257 50% (2018), and 37% (2019) (Figure 5).

258 From the hindcasts of daily PST classification, we selected the dates each summer when the
259 model first predicted a >50% probability that the blue mussels would be > 80µg $100g^{-1}$ (Figure 6a). We
260 found that predicted date of the first blue mussel > 80µg $100g^{-1}$ averaged within 10 days of the actual date
261 for most sample sites that experienced blooms in 2017 and 2019 (Figure 6b), while 2018 most predicted
262 dates were far earlier than the first observed bloom.

263 **4. Discussion**

264 Despite the paucity of data compared to regions with state sponsored monitoring programs, we
265 observed blooms of *Alexandrium* (as evidenced by seasonally elevated PST concentrations) in each year
266 of this study, with peaks in concentrations usually occurring during the late spring/early summer (Figure
267 2). However, it should be noted that blue mussel samples that registered > 80µg $100g^{-1}$ were reported in
268 each month from March – November. While concentrations of PSTs in shellfish are difficult to associate

with an acute toxicity reference dose or an LD_{50} due to variation in body weight and portion size, blue mussel concentrations > 1,300 µg $100g^{-1}$ have been associated with PSP cases in Alaska (Alaska Department of Health and Human Services, unpublished data).

4.1 Variable importance

The importance of salinity, SST, air temperature, and discharge in our full model (Figure 3b) is in good agreement with what other modelling efforts in the North Pacific have found (Finnis et al., 2017; Moore et al., 2009; Tobin et al., 2019). While several of these variables are highly correlated on annual timescales (e.g. SST and air temperature), by randomly sampling three or more variables at each split random forests have been shown to be somewhat resistant to multicollinearity (Breiman, 2001; Liaw and Wiener, 2002).

Interestingly, while salinity and discharge were important factors in our model, precipitation was not a strong predictor of PST events, which is similar to models generated for British Columbia (Finnis et al., 2017) and Puget Sound (Moore et al., 2009). Precipitation in Southeast Alaska is frequent and year-round, and many of the sample sites had measurable precipitation on more than 50% of the dates analyzed in this study. This finding is similar to the conclusions of Brandenburg et al. (2017) that while precipitation can indirectly influence the strength or timing of *Alexandrium* blooms, it is likely that the effect of precipitation (and by extension discharge) on sea surface salinity is a more important driver and therefore a better predictor. It is also important to note that in the absence of precipitation during warm, dry summer conditions in Southeast Alaska glacial discharge increases, which can affect the sea surface salinity at some coastal communities with significant glacial input (e.g., Juneau, Haines, Skagway). While salinity was an important variable in our model, Vandersea et al. (2018) found only a weak correlation between salinity and *Alexandrium* cell concentrations. However, we note that the exercise of determining feature importance for random forest classification does not necessitate a strong correlation, thus the endpoint of predicting PST concentrations might foster a different model than one predicting cell concentrations.

294 *4.2 Accumulated local effects*

295 Overall, ALE plots were consistent with the idea that warmer temperatures and minimal FW
296 discharge are associated with high PST concentrations (Figure 4). Similar to other studies, SST was an
297 important factor in determining *Alexandrium* and PST dynamics (Finnis et al., 2017; Moore et al., 2009;
298 Tobin et al., 2019; Valbi et al., 2019; Vandersea et al., 2018). ALE plots (Figure 4) show the highest
299 prediction values for SST between 8°C and 15°C, with a maximum around 12°C. In laboratory
300 conditions, optimal growth rate of *Alexandrium* isolates from the Salish Sea (Puget Sound and Strait of
301 Juan de Fuca) occurred between 10°C and 24°C (Bill et al., 2016), and other studies have found limited
302 growth below 13°C (Nishitani and Chew, 1984). SSTs in Southeast Alaska only rise above 10°C in the
303 summer months, and only rose above 13°C at some sites during the midsummer months (July and
304 August). Recently Tobin et al. (2019) found blooms of *Alexandrium* in Southeast Alaska occurred when
305 SSTs were ≥ 7°C, and Vandersea et al. (Vandersea et al., 2018) found a similar threshold of 7-8°C for
306 *Alexandrium* in Kachemak Bay, Alaska. It is important to note here key differences between our approach
307 and other studies in the region (Tobin et al., 2019; Vandersea et al., 2018), namely the differences in our
308 modeled outcomes (PST concentrations in shellfish vs. *Alexandrium* cell counts) and the source of our
309 SST data (satellite vs. in situ). We therefore interpret these comparisons of SST ranges and thresholds
310 with caution.

311 Several models have predicted the expansion of HAB species globally and a widening of the
312 bloom window into the early spring and late fall (Gobler et al., 2017; Moore et al., 2008; Wells et al.,
313 2019). Indeed, Vandersea et al. (2018) noted an increase in average temperatures from 2012-2017 in
314 Kachemak Bay, Alaska and correlated observed *Alexandrium* blooms with anomalous winter
315 temperatures. Although we do not have a long enough record to examine the expansion of HABs in
316 Southeast Alaska in relation to climate change, it does highlight the necessity for routine monitoring of
317 shellfish and predictive models since elevated SSTs might allow for blooms of *Alexandrium* during
318 traditionally "safe" harvest seasons.

Interestingly, the effect of wind speed in the full model (Figure 4) shows higher ALE with increasing wind speed, a finding which contradicts what has been described by Moore et al. (2011) in which calmer winds are associated with water column stratification. Our model used average wind speed over a 24-hour period, the values on the upper end of the distribution do not represent extremely strong winds (5 m/s is 18km/h), which may help to explain why higher wind speeds were associated with increased probability of PSTs > 80µg $100g^{-1}$. Further refinement of these models could incorporate wind direction as well as speed, although determining the marginal effects of wind direction would likely be highly site-specific.

4.3 Lagged model hindcasting

Due to the nature of bloom dynamics it is expected that the environmental conditions leading to stratification and bloom formation would precede the initiation and peak of the bloom by a number of days (Franks, 2018; Moore et al., 2009; Valbi et al., 2019). Tobin et al. (2019) found evidence that environmental conditions such as decreased wind speed and precipitation 1-10 days prior favored blooms of *Alexandrium* in Juneau, Alaska. Similarly, Moore et al. (2011) examined environmental conditions up to 20 days prior to peak toxicity for *M. edulis* in Puget Sound, Washington and found warm, dry conditions with stable air mass preceded toxic blooms. For forecasting purposes with respect to seafood safety, it is important to resolve the date at which the concentrations of PSTs are likely to cross the threshold for human consumption rather than the peak toxicity. We incorporated a 7-day lag for our environmental variables (Figure S3) to evaluate the efficacy of a forecast model but the overall accuracy of lagged models did not vary by a wide margin (79-81%, Figure S4) which is unsurprising given the autocorrelative nature of many of these variables. So while a 7-day lagged model performed best with respect to predicting PST concentrations > 80µg $100g^{-1}$, other lagged models from 1-10 days would likely perform similarly in terms of overall accuracy of hindcasts.

Although the shape of forecast probabilities generally tracks with PST concentrations over the course of the bloom window (Figure 5), the hindcast models did poorly at classifying OOS shellfish

concentrations from April-October in 2017-2019 (accuracy <= 50%). An initial problem with this type of model is that the model predicts that shellfish PST concentrations will be above the FDA threshold for the entire course of the summer, since environmental conditions are often amenable for bloom formation (Moore et al., 2011). Since PSTs in blue mussel species are eliminated rapidly following the termination of blooms (Bricelj and Shumway, 1998), the continuing misclassification drives the low overall accuracies for these hindcasts. This model examines only environmental variables and does not incorporate toxicokinetics of uptake and elimination of PST in blue mussels. Both uptake and elimination of PSTs could be affected by environmental variables such as temperature and salinity, however not much is known about these interactions (Bricelj and Shumway, 1998). Future models incorporating toxicokinetics of PSTs in mussels should improve the overall accuracy of the model and the incorporation of more years of data might be able to resolve some of these dynamics.

Interestingly, while many communities experienced blooms in May and June in 2017 and 2019, most communities did not experience an increase in PST concentrations in 2018 until August or later. To examine differences across all environmental datasets between each of 2017, 2018, and 2019, we performed a PCA using scaled environmental variables (monthly averages). Significant differences in environmental conditions year-to-year exhibit as unique groupings in the PCA plots (Figure 7). For instance, June of 2017 was significantly different than 2018 and 2019, with precipitation as the driving variable contributing to that principal component. June 2017 had more precipitation and more days with measurable precipitation than June 2018 and June 2019 across all sample sites (data not shown). Visual inspection of the PCA plots for April and May revealed no large-scale identifiable differences among 2017, 2018 and 2019 across samples sites. With a limited longitudinal data set it is difficult to say for certain why 2018 was different than 2017 and 2019 in terms PST dynamics. More data across more summers should be able to resolve whether this year was truly anomalous or if as yet undetermined environmental controls drive later season blooms.

There are several variables which influence PST dynamics which have not been assessed here. The lifecycle of *Alexandrium* involves the production of cysts which can settle out into the sediment following a bloom, and there is an association between cyst deposition and PST events (Cox et al., 2008). Currents can carry vegetative *Alexandrium* suspended in the water column away from cyst beds, thus it would be conceivable for a site with little or no cysts in the sediment to experience high PST concentrations. Nevertheless, cyst mapping would undoubtedly contribute to our understanding of *Alexandrium* dynamics in Southeast Alaska, and efforts to examine cyst beds are currently underway. It is worth noting however that the germination of cysts is triggered by favorable environmental conditions, notably salinity and temperature (Genovesi et al., 2009). Therefore, a model that includes these variables would be somewhat sensitive to cyst dynamics although it would not be able to discern toxic events following cyst resuspension from events involving novel vegetative blooms.

Nutrient concentrations are also known to be important drivers of HABs, and fixed nitrogen often limits phytoplankton production following initial depletion during spring blooms. Influx of nutrients into the photic zone via wind-driven vertical mixing can induce phytoplankton blooms later in the summer and fall (Eslinger et al., 2001). Similarly, there is evidence to suggest that nutrient concentrations strongly impact phytoplankton assemblages, and phosphate limitation is linked to toxin production in *Alexandrium*, likely due to the stimulation of inter- and intraspecific competition under nutrient limiting conditions (Frangópulos et al., 2004). Thus while some of the environmental variables measured here are related to nutrient availability (upwelling, discharge, wind speed), dynamics such as nutrient uptake and absorption have not been modelled here and could potentially explain some variation in toxin production.

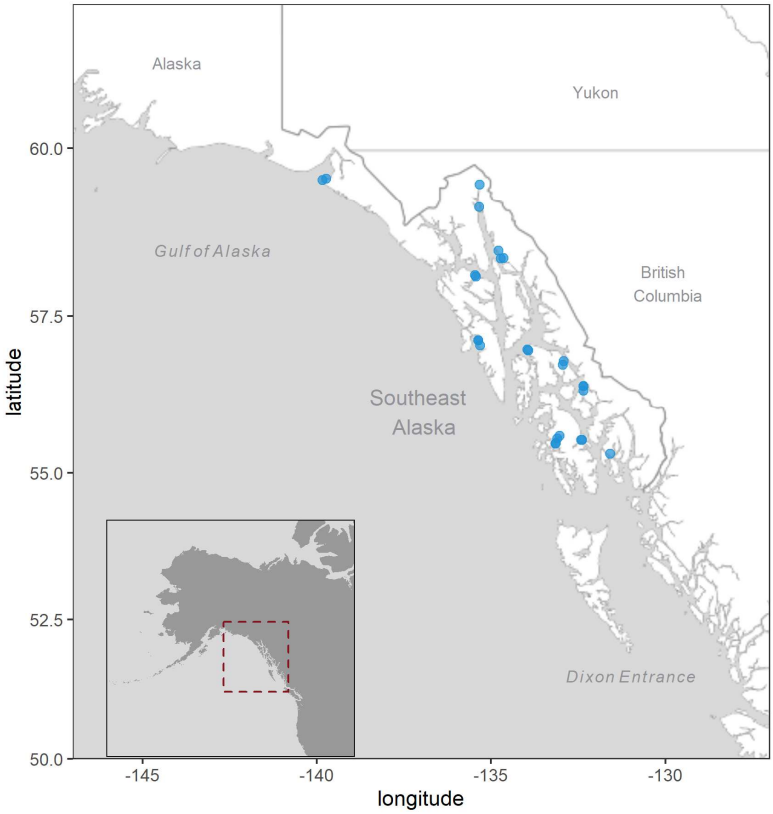
Using random forest models trained on publicly available environmental data we were able to investigate environmental drivers of PST in Southeast Alaska. Our model suggests that, similar to what others have reported in nearby regions, PST dynamics in this region are driven largely by sea surface temperature, air temperature, and salinity. These types of models with high temporal resolution might be able to elucidate more ephemeral aspects of bloom formation which could lead to more accurate forecasts

for harvesters and consumers. Models incorporating additional variables which may drive PST dynamics (e.g., nutrients, toxicokinetics, cyst mapping) will further enhance our ability to predict and forecast PSTs in Southeast Alaska.

5. **Acknowledgements**

Funding for this work was provided by a National Marine Fisheries Service Saltonstall-Kennedy Grant (Award NA17NMF4270238). We would also like to thank the tribal partners of the SEATOR network for their assistance and support including Central Council of Tlingit and Haida Indian Tribes of Alaska, Chilkoot Indian Association, Craig Tribal Association, Hoonah Indian Association, Hydaburg Cooperative Association, Ketchikan Indian Community, Klawock Cooperative Association, Kodiak Area Native Association, Metlakatla Indian Community, Organized Village of Kake, Organized Village of Kasaan, Petersburg Indian Association, Skagway Traditional Council, Sitka Tribe of Alaska, Sun'aq Tribe of Kodiak, Wrangell Cooperative Association, and Yakutat Tlingit Tribe.

405 Figure 1 –



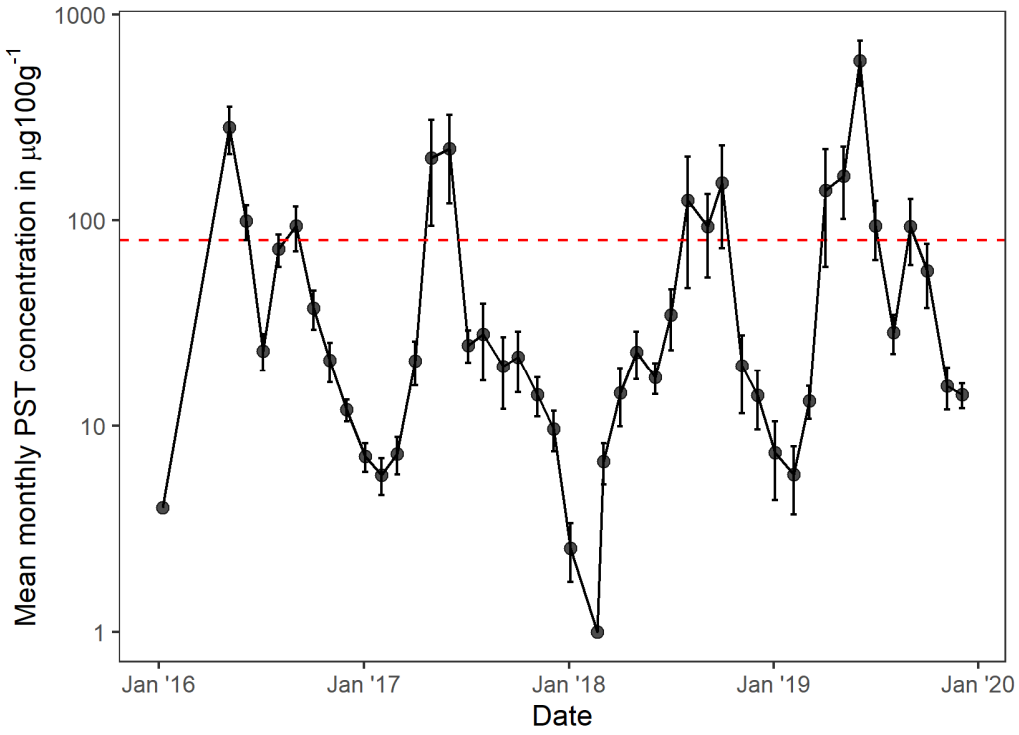
406

407 **Figure 1** Sample sites for shellfish toxins in blue mussels (*M. trossulus*) in Southeast Alaska. Samples
408 were collected by Southeast Alaska Tribal Ocean Research (SEATOR) partners (2016-2019).

409

410

411 Figure 2 –



412

413 **Figure 2 -** Mean monthly PST concentrations (presented on log scale) for samples collected by SEATOR
414 since 2016. The dashed red line indicates the 80 µg $100g^{-1}$ action threshold set by the US FDA. Error bars
415 represent standard error for each month.

Figure 3 –

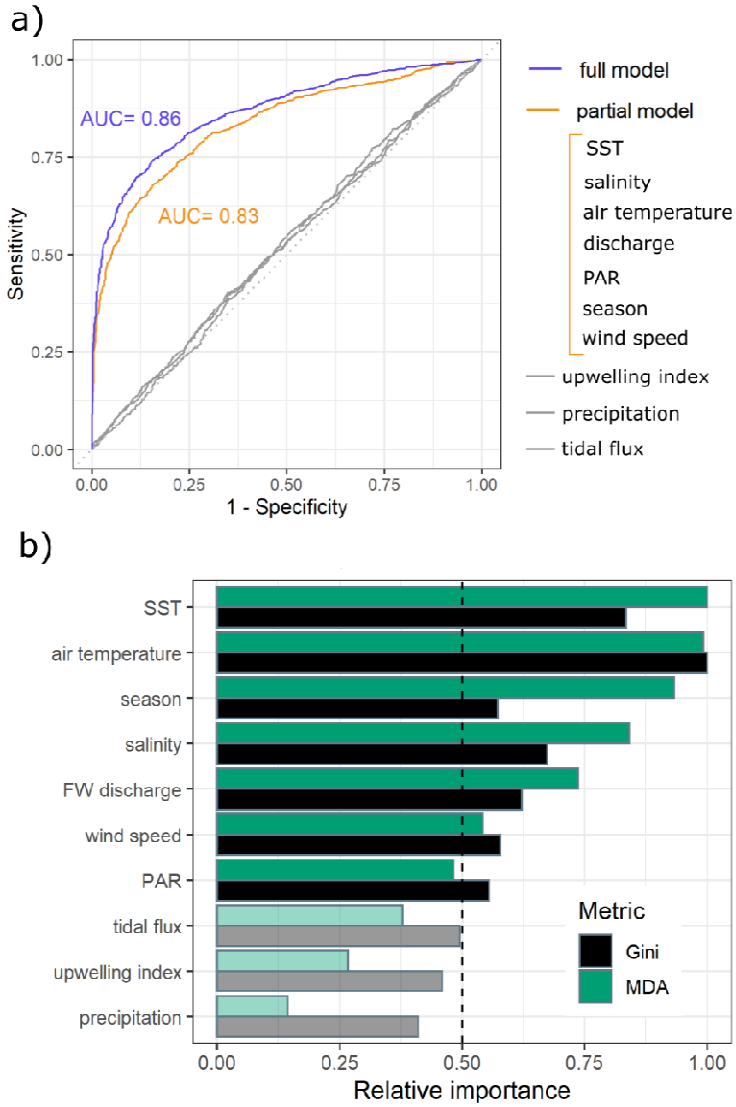
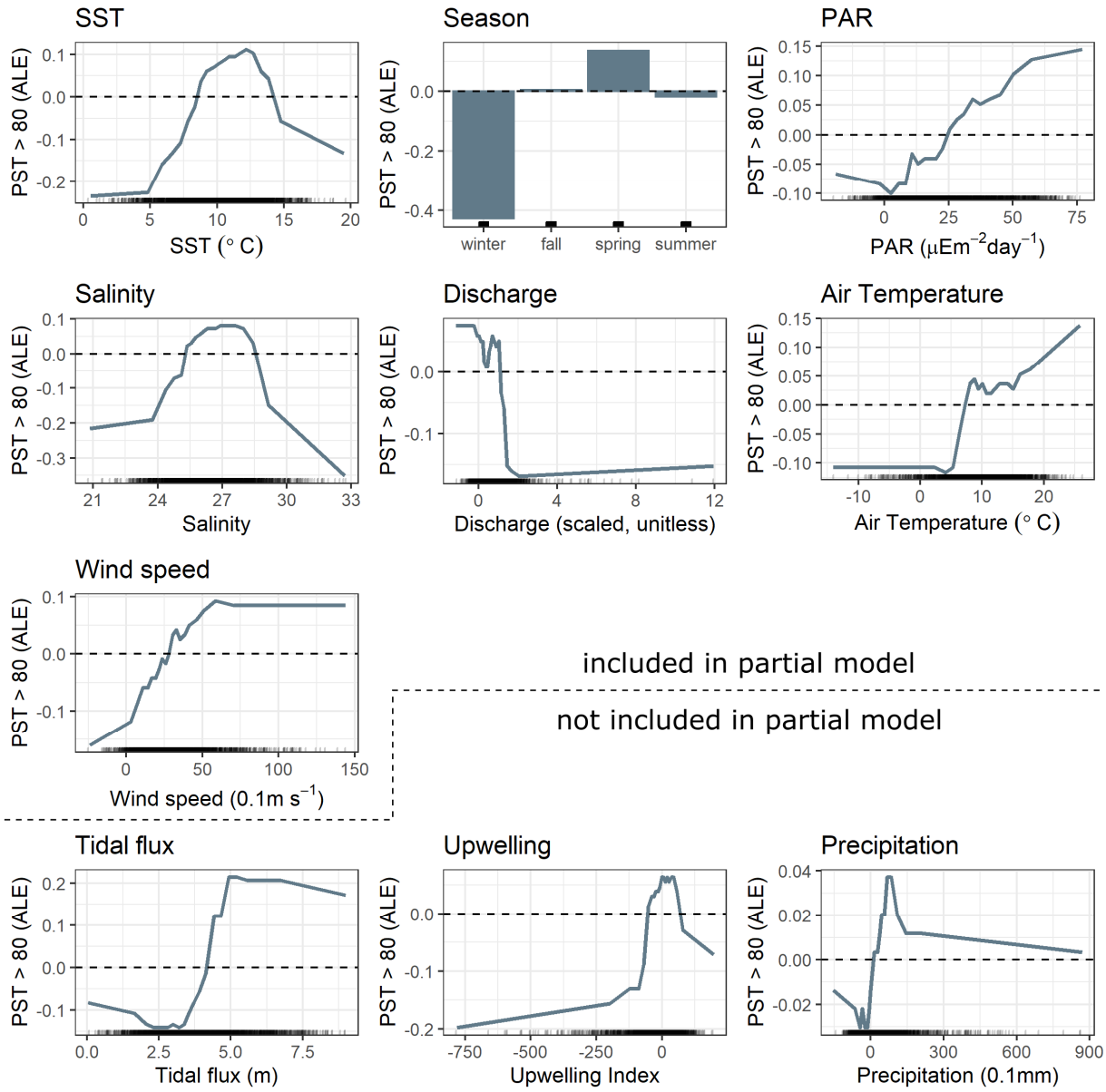


Figure 3 a) Receiver operation characteristic (ROC) curve for the full model (all variables, blue), partial model (select variables, orange), and variables left out of the partial model (grey). Area under the curve (AUC) calculation are presented for both the full. b) Variable importance plots presenting relative importance metrics for mean decrease in Gini impurity (Gini) and mean decrease in accuracy (MDA). For ease of graphical interpretation both metrics were scaled to the highest variable within each metric so that each metric is presented on a scale of 0-1. Variables that fell below the 0.5 relative importance threshold for Gini and MDA were excluded from the partial model presented in panel a). PAR = Photosynthetically available radiation, SST = Sea-surface temperature, FW=freshwater.

427 Figure 4 –

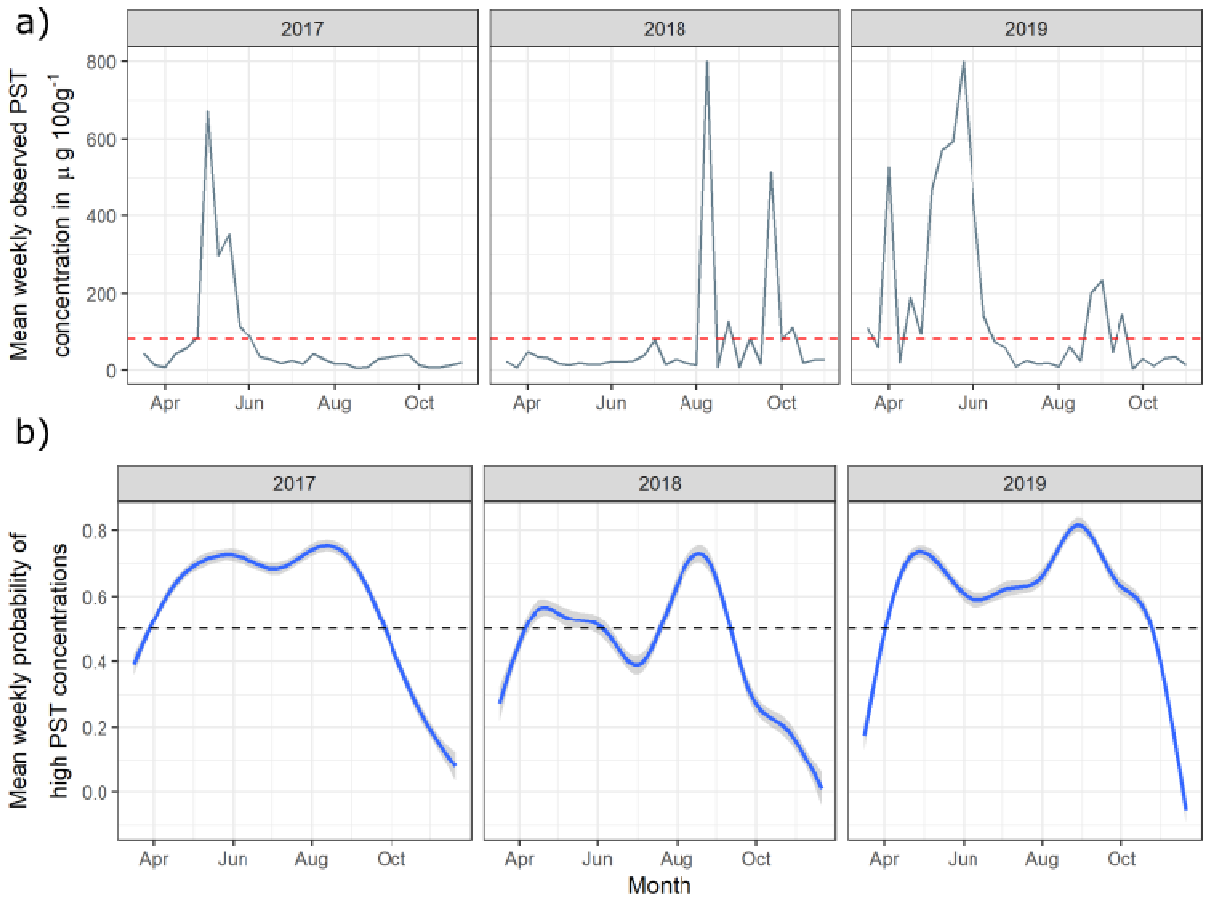


428

429 Figure 4 – Accumulated local effect (ALE) plots for each variable in the full model. The y-axes represent
430 the effect of that variable on predictions of concentrations > 80 μg 100g^{-1} (positive values) and predictions
431 of concentrations < 80 μg 100g^{-1} (negative values). Dashed horizontal lines represent ALE=0 for that
432 variable, indicating values which do not affect predictions significantly. The distribution of each variable
433 is represented visually on the x axis where ticks represent individual data points.

434

435 Figure 5 –

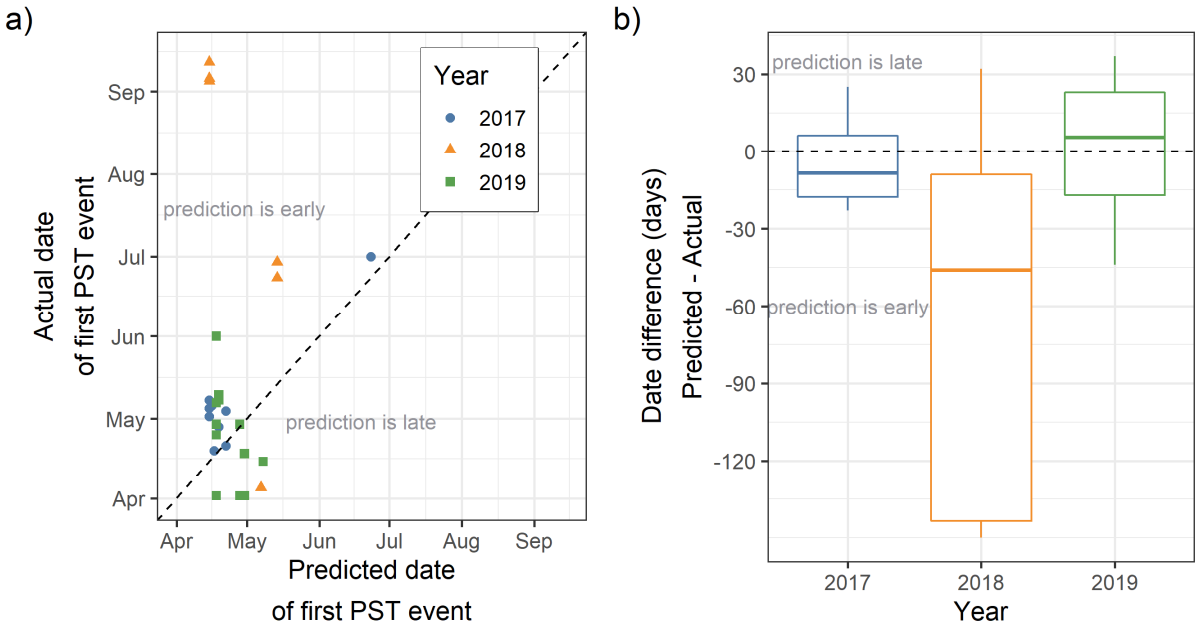


436

437 **Figure 5**a) Weekly average PST concentrations during April-November in 2017, 2018, and 2019. Red
438 dashed line represents the FDA threshold of 80µg $100g^{-1}$. b) Daily average hindcast probabilities for PST
439 concentrations above 80µg $100g^{-1}$. Probabilities approaching 1 indicate high confidence of PST
440 concentrations above 80µg $100g^{-1}$, while probabilities approaching 0 indicate high confidence of PST
441 concentrations less than 80µg $100g^{-1}$.

442

443 Figure 6 –

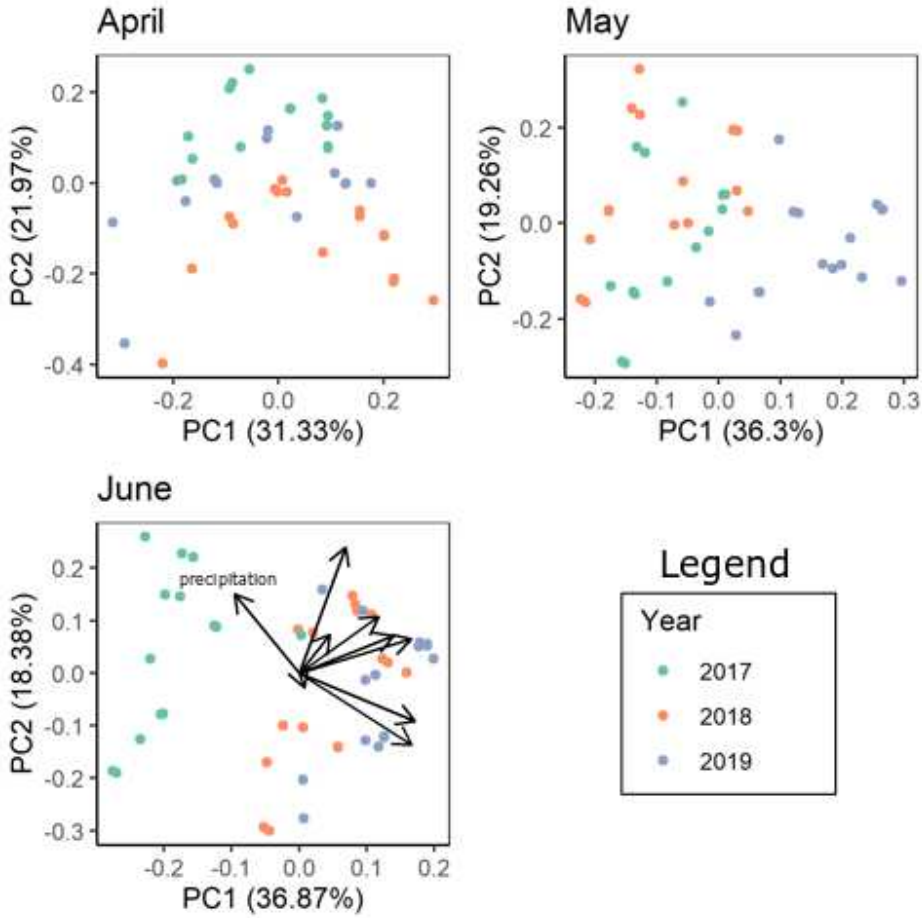


444

445 **Figure 6** – Out-of-sample (OOS) hindcasts for the first date each year when paralytic shellfish toxin
446 (PST) concentrations were given a >50% probability to be > 80µg $100g^{-1}$, an estimate of the timing of the
447 spring bloom. a) The hindcast date is compared to the actual date of first PST > 80µg $100g^{-1}$. Each point
448 represents an individual sample site, and each year is presented as a different color. The dotted line
449 represents the identity line (y=x) between dates on the *x* (predicted) and *y* (actual) axes. b) The difference
450 between the predicted and actual dates (predicted-actual) of the spring bloom.

451

452 Figure 7 –



453

454 Figure 7 - Principal component analysis (PCA) for environmental variables used in this study compared
455 across years for each sample site. Variables were scaled within each month. Eigenvector loadings (using
456 the square of the standard deviation and the rotation is the loading onto the first two principal
457 components) are shown in the June panel and the loading for precipitation is labeled to illustrate
458 differences in precipitation between 2017 and 2018-2019.

459

460

461 **Works Cited**

462 Anderson, D.M., Alpermann, T.J., Cembella, A.D., Collos, Y., Masseret, E., and Montresor, M. (2012).
463 The globally distributed genus Alexandrium: multifaceted roles in marine ecosystems and impacts on
464 human health. Harmful Algae *14*, 10–35.

465 Bean, L.L., McGowan, J.D., and Hurst, J.W. (2005). Annual variations of paralytic shellfish poisoning in
466 Maine, USA 1997–2001. Deep Sea Res. Part II Top. Stud. Oceanogr. *52*, 2834–2842.

467 Bill, B.D., Moore, S.K., Hay, L.R., Anderson, D.M., and Trainer, V.L. (2016). Effects of temperature and
468 salinity on the growth of Alexandrium (Dinophyceae) isolates from the Salish Sea. J. Phycol. *52*, 230–
469 238.

470 Breiman, L. (2001). Random forests. Mach. Learn. *45*, 5–32.

471 Bricelj, V.M., and Shumway, S.E. (1998). Paralytic Shellfish Toxins in Bivalve Molluscs: Occurrence,
472 Transfer Kinetics, and Biotransformation. Rev. Fish. Sci. *6*, 315–383.

473 Calle, M.L., and Urrea, V. (2011). Letter to the Editor: Stability of Random Forest importance measures.
474 Brief. Bioinform. *12*, 86–89.

475 Castrodale, L. (2015). Paralytic Shellfish Poisoning - Alaska, 1993-2014 (Anchorage, AK, USA:
476 Department of Health and Social Services).

477 Chamberlain, S. (2019). Package 'rnoaa.'

478 Chin, T.M., Vazquez-Cuervo, J., and Armstrong, E.M. (2017). A multi-scale high-resolution analysis of
479 global sea surface temperature. Remote Sens. Environ. *200*, 154–169.

480 Cox, A.M., Shull, D.H., and Horner, R.A. (2008). Profiles of Alexandrium catenella cysts in Puget Sound
481 sediments and the relationship to paralytic shellfish poisoning events. Harmful Algae *7*, 379–388.

482 De Cicco, L.A., Lorenz, D., Hirsch, R.M., and Watkins, W. (2018). dataRetrieval: R packages for
483 discovering and retrieving water data available from U.S. federal hydrologic web services (Reston, VA:
484 U.S. Geological Survey).

485 Eslinger, D.L., Cooney, R.T., Mcroy, C.P., Ward, A., Kline, T.C., Simpson, E.P., Wang, J., and Allen,
486 J.R. (2001). Plankton dynamics: observed and modelled responses to physical conditions in Prince
487 William Sound, Alaska. Fish. Oceanogr. *10*, 81–96.

488 Etheridge, S.M. (2010). Paralytic shellfish poisoning: Seafood safety and human health perspectives.
489 Toxicon *56*, 108–122.

490 Finnis, S., Krstic, N., McIntyre, L., Nelson, T.A., and Henderson, S.B. (2017). Spatiotemporal patterns of
491 paralytic shellfish toxins and their relationships with environmental variables in British Columbia,
492 Canada from 2002 to 2012. Environ. Res. *156*, 190–200.

493 Frangópulos, M., Guisande, C., deBlas, E., and Maneiro, I. (2004). Toxin production and competitive
494 abilities under phosphorus limitation of Alexandrium species. Harmful Algae *3*, 131–139.

Franks, P.J.S. (2018). Recent Advances in Modelling of Harmful Algal Blooms. In Global Ecology and Oceanography of Harmful Algal Blooms, P.M. Glibert, E. Berdalet, M.A. Burford, G.C. Pitcher, and M. Zhou, eds. (Cham: Springer International Publishing), pp. 359–377.

Genovesi, B., Laabir, M., Masseret, E., Collos, Y., Vaquer, A., and Grzebyk, D. (2009). Dormancy and germination features in resting cysts of Alexandrium tamarense species complex (Dinophyceae) can facilitate bloom formation in a shallow lagoon (Thau, southern France). J. Plankton Res. *31*, 1209–1224.

Gessner, B.D., and Middaugh, J.P. (1995). Paralytic shellfish poisoning in Alaska: a 20-year retrospective analysis. Am. J. Epidemiol. *141*, 766–770.

Glibert, P.M., Berdalet, E., Burford, M.A., Pitcher, G.C., and Zhou, M. (2018). Global ecology and oceanography of harmful algal blooms (Springer).

Hallegraeff, G.M. (1993). A review of harmful algal blooms and their apparent global increase. Phycologia *32*, 79–99.

Harley, J.R., Lanphier, K., Kennedy, E.G., Leighfield, T.A., Bidlack, A., Gribble, M.O., and Whitehead, C. (2020). The Southeast Alaska Tribal Ocean Research (SEATOR) Partnership: Addressing Data Gaps in Harmful Algal Bloom Monitoring and Shellfish Safety in Southeast Alaska. Toxins *12*, 407.

Harvey, A.C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter by Andrew C. Harvey.

Hosmer, D.W.Jr., and Lemeshow, S. (2004). Applied Logistic Regression (John Wiley & Sons).

Jackley, J., Gardner, L., Djunaedi, A.F., and Salomon, A.K. (2016). Ancient clam gardens, traditional management portfolios, and the resilience of coupled human-ocean systems. Ecol. Soc. *21*.

James, K.J., Carey, B., O'halloran, J., Pelt, F.N.A.M. van, and Škrabáková, Z. (2010). Shellfish toxicity: human health implications of marine algal toxins. Epidemiol. Infect. *138*, 927–940.

Knaack, J.S., Porter, K.A., Jacob, J.T., Sullivan, K., Forester, M., Wang, R.Y., Trainer, V.L., Morton, S., Eckert, G., McGahee, E., et al. (2016). Case diagnosis and characterization of suspected paralytic shellfish poisoning in Alaska. Harmful Algae *57*, 45–50.

Laabir, M., Collos, Y., Masseret, E., Grzebyk, D., Abadie, E., Savart, V., Sibat, M., and Amzil, Z. (2013). Influence of Environmental Factors on the Paralytic Shellfish Toxin Content and Profile of Alexandrium catenella (Dinophyceae) Isolated from the Mediterranean Sea. Mar. Drugs *11*, 1583–1601.

Laanaia, N., Vaquer, A., Fiandrino, A., Genovesi, B., Pastoureaud, A., Cecchi, P., and Collos, Y. (2013). Wind and temperature controls on Alexandrium blooms (2000–2007) in Thau lagoon (Western Mediterranean). Harmful Algae *28*, 31–36.

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

Menardi, G., and Torelli, N. (2014). Training and assessing classification rules with imbalanced data. Data Min. Knowl. Discov. *28*, 92–122.

Millero, F.J. (2013). Chemical oceanography (CRC press).

Molnar, C. (2020). Interpretable Machine Learning A Guide for Making Black Box Models Interpretable (Leanpub).

Moore, S.K., Trainer, V.L., Mantua, N.J., Parker, M.S., Laws, E.A., Backer, L.C., and Fleming, L.E. (2008). Impacts of climate variability and future climate change on harmful algal blooms and human health. Environ. Health *7*, S4.

Moore, S.K., Mantua, N.J., Hickey, B.M., and Trainer, V.L. (2009). Recent trends in paralytic shellfish toxins in Puget Sound, relationships to climate, and capacity for prediction of toxic events. Harmful Algae *8*, 463–477.

Moritz, S., and Bartz-Beielstein, T. (2017). imputeTS: Time Series Missing Value Imputation in R. R J. *9*, 207–218.

Newton, R., and Moss, M. (1984). The subsistence lifeway of the Tlingit people: excerpts of oral interviews (Juneau, Alaska: U.S. Dept. of Agriculture, Forest Service, Alaska Region).

Nishitani, L., and Chew, K.K. (1984). Recent developments in pralytic shellfish poisoning research. Aquaculture *39*, 317–329.

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

Reese, S.L., Estes, J.A., and Jarman, W.M. (2012). Organochlorine contaminants in coastal marine ecosystems of southern Alaska: Inferences from spatial patterns in blue mussels (Mytilus trossulus). Chemosphere *88*, 873–880.

Simons, R.A. (2019). ERDDAP (Monterey, CA: NOAA/NMFS/SWFSC/ERD).

Tobin, E.D., Wallace, C.L., Crumpton, C., Johnson, G., and Eckert, G.L. (2019). Environmental drivers of paralytic shellfish toxin producing Alexandrium catenella blooms in a fjord system of northern Southeast Alaska. Harmful Algae *88*, 101659.

Trainer, V.L., Eberhart, B.-T.L., Wekell, J.C., Adams, N.G., Hanson, L., Cox, F., and Dowell, J. (2003). Paralytic shellfish toxins in Puget Sound, Washington state. J. Shellfish Res. *22*, 213–223.

Valbi, E., Ricci, F., Capellacci, S., Casabianca, S., Scardi, M., and Penna, A. (2019). A model predicting the PSP toxic dinoflagellate Alexandrium minutum occurrence in the coastal waters of the NW Adriatic Sea. Sci. Rep. *9*, 4166.

Van Dolah, F.M., Leighfield, T.A., Doucette, G.J., Bean, L., Niedzwiadek, B., and Rawn, D.F.K. (2009). Single-laboratory validation of the microplate receptor binding assay for paralytic shellfish toxins in shellfish. J. AOAC Int. *92*, 1705–1713.

Van Dolah, F.M., Fire, S.E., Leighfield, T.A., Mikulski, C.M., and Doucette, G.J. (2012). Determination of paralytic shellfish toxins in shellfish by receptor binding assay: collaborative study. J. AOAC Int. *95*, 795–812.

Vandersea, M.W., Kibler, S.R., Tester, P.A., Holderied, K., Hondolero, D.E., Powell, K., Baird, S., Doroff, A., Dugan, D., and Litaker, R.W. (2018). Environmental factors influencing the distribution and abundance of Alexandrium catenella in Kachemak bay and lower cook inlet, Alaska. Harmful Algae *77*, 81–92.

Weingartner, T., Eisner, L., Eckert, G.L., Danielson, S., and Bellwood, D. (2009). Southeast Alaska: Oceanographic Habitats and Linkages. J. Biogeogr. *36*, 387–400.

570 Wells, M.L., Karlson, B., Wulff, A., Kudela, R., Trick, C., Asnaghi, V., Berdalet, E., Cochlan, W.,
571 Davidson, K., De Rijcke, M., et al. (2020). Future HAB science: Directions and challenges in a changing
572 climate. Harmful Algae *91*, 101632.

573 Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A.,
574 Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686.

575