

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

**Performance of Fish-Habitat Classifiers Based On
Derived Predictors from a Coupled Biophysical Model**

Edward D. Weber^{*a}, Yi Chao^b, and Fei Chai^c

^aNOAA Southwest Fisheries Science Center
8901 La Jolla Shores Drive
La Jolla, California 92037-1509, USA

^bRemote Sensing Solutions, Inc.
248 East Foothill Blvd
Suite 200
Monrovia, CA 91016, USA

^cUniversity of Maine
School of Marine Sciences
5706 Aubert Hall
Orono, ME 04469, U.S.A.

*Corresponding author.

E-mail addresses: ed.weber@noaa.gov (E. Weber), ychao@remotesensingsolutions.com (Y. Chao), fchai@maine.edu (F. Chai)

47 **Highlights**

48

- 49 • We compared habitat classifiers based on satellite data or biophysically modeled data.
- 50 • Classifiers from modeled predictors performed similarly to those from satellite data.
- 51 • Modeled zooplankton did not improve accuracy over remotely sensed chlorophyll.
- 52 • Specificity was good but sensitivity was poor for both types of classifiers.
- 53 • Modeled predictors are useful to overcome cloud cover and for forecasting.

54

55 **Abstract**

56

57 Fish habitat models based on remotely-sensed data may be limited by satellite coverage
58 and availability. We compared the fit and predictive power of Random Forest habitat classifiers
59 that were developed using predictors derived from a coupled biophysical model (i.e., modeled
60 predictors) versus similar classifiers that used remotely-sensed satellite data for two data sets
61 (eggs and adults) and four species that occur widely in the California Current system. When
62 tested on independent data, classifiers of spawning habitat that used derived predictors (derived
63 classifiers) had nearly identical accuracies (0-2% difference) to similar classifiers based on
64 satellite data (satellite classifiers). Accuracies of derived classifiers of adult habitat were within -
65 8% to +7% of comparable satellite classifiers. Accuracies of both types of classifiers on test data
66 were much greater for Northern anchovy *Engraulis mordax* and Pacific hake *Merluccius*
67 *productus* (0.75-0.97) than for jack mackerel *Trachurus symmetricus* and Pacific sardine
68 *Sardinops sagax* (0.61-0.72), and generally were greater for classifiers of spawning habitat than
69 for adult habitat. Specificity was very good for both types of classifiers, but sensitivity was poor,
70 because classifiers identified potential habitat which was not fully occupied. Adults of all species
71 used a broader range of habitat conditions during summer than during the spring spawning
72 period. Derived classifiers have some advantages over satellite classifiers; they are not limited by
73 cloud cover and they can make predictions in near real-time or the short-term future. However,
74 there was no consistent improvement in the accuracy of derived predictors that included modeled
75 zooplankton concentrations over comparable satellite classifiers that included
76 reflectance/chlorophyll concentration.

77

78 **Keywords:** Fish, Habitat, Models, ROMS-CoSiNE, Satellite sensing, Random Forest, California
79 Current System

80

81 **1. Introduction**

82

83 Fish habitat models are important for our scientific understanding and management of
84 marine systems because habitat conditions affect fish population dynamics (Hayes et al., 1996;
85 Vasconcelos et al., 2013), community interactions (Lindgren et al., 2016; Pecuchet et al., 2016),
86 and spatial distribution in relation to marine reserves (Thompson et al., 2012) and harvest
87 (Leitão, 2015; Tommasi et al., 2016). Habitat models are also used in the design of surveys to
88 define the sampling frame and allocate sampling effort (Weber and McClatchie, 2010; Zwolinski
89 et al., 2011).

90 Marine fish habitat models have frequently relied on remotely sensed-data from satellites
91 as predictors. Although such models work well for some species, they have limitations. Cloud
92 cover can result in large gaps in geographic coverage where no model prediction is possible for

93 some areas and times of year. Remotely sensed data must be collected and processed before use.
94 Thus, it cannot be used for real-time or near-future prediction. A limited suite of predictor
95 variables is available from satellite data (e.g., sea-surface temperature, chlorophyll, geostrophic
96 current, wind speed, and, recently, salinity) which may not be adequate to characterize the niche
97 of all species with great enough precision and accuracy to be useful. Finally, remotely sensed
98 variables usually are entered directly as “snapshots” into statistical models, which may result in
99 bias due to lack of process. For example, remotely sensed reflectance/chlorophyll predictors in
100 fish habitat models usually are a proxy for secondary production that is available as prey for fish.
101 However, instantaneous chlorophyll measurements may not accurately reflect the recent
102 productivity of an area or the amount of production that has actually been transferred to higher
103 trophic levels.

104
105 An alternative approach to understanding fish habitat associations is to analyze their
106 spatial distributions as part of a much broader mechanistic model of the ecosystem. There has
107 been much recent progress on “end-to-end” models that include hydrodynamics, nutrient flows,
108 primary productivity, and higher trophic levels up to fish or even harvest (Rose et al., 2010;
109 Kishi et al., 2011; Rose et al., 2015). Such models have some advantages over statistical habitat
110 models. They include bottom-up and top-down biological interactions along with environmental
111 conditions. Because they include population dynamics and animal movement, they may be better
112 able to differentiate realized habitat from unoccupied potential habitat (Planque et al., 2007).
113 Finally, such models could be run forward in time to make predictions in the near future.
114 However, end-to-end models currently are most useful for understanding ecosystem interactions
115 via simulation studies rather than for prediction. This is because they require more data than
116 usually are available for adequate calibration and validation, and they are particularly subject to
117 bias due to non-stationarity of the ecosystems they simulate (e.g., environmental regime shifts
118 and changes in genetic structure of component species; Fulton, 2010). For example, a recently
119 published end-to-end model of the California Current (Fiechter et al., 2015; Rose et al., 2015)
120 produced dynamics in the abundance of northern anchovy *Engraulis mordax* and Pacific sardine
121 *Sardinops sagax* that were qualitatively similar to historical data for a 45-year period but could
122 not yet be parameterized to produce annual estimates suitable for “tactical” decision making.

123
124 A hybrid approach to modeling fish habitat that may overcome some of the limitations of
125 both satellite-based statistical models and fully mechanistic models is to use a mechanistic model
126 to provide variables such as environmental conditions and production of lower trophic levels,
127 and then enter modeled conditions as predictors (hereafter derived predictors) into a statistical
128 model or algorithm. Formally, modeled predictors are functions of the parameters used to create
129 the underlying physical model, e.g., wind forcing, boundary conditions, and initial state of the
130 system. They could be described as indices based on the calculations of the mechanistic model.
131 In practice, using modeled variables from mechanistic models as if they were measured can
132 provide a convenient and easily interpretable set of predictors for statistical modeling. There is a
133 trade-off in potential sources of error using this approach. It implicitly incorporates the
134 phenology of oceanic conditions into derived predictors because they are the product of an
135 underlying physical model. However, the approach adds some bias to the statistical model
136 because derived predictors do not perfectly reproduce the conditions they represent. Becker et al.
137 (2016) used this approach to model the distributions of eleven cetacean species in the California
138 Current system. They reported that generalized additive models that used derived predictors of

139 salinity, sea-surface temperature, and log-transformed surface chlorophyll concentration from an
140 ocean circulation model produced nearly identical predictions to those of similar models based
141 on remotely sensed satellite data.

142

143 The objective of this study was to compare the performance of fish habitat models that
144 used derived predictors versus similar models that used remotely sensed satellite predictors,
145 where the type of habitat model was a classification algorithm (i.e., fish presence or absence;
146 hereafter a classifier). Specifically, we wanted to test whether classifiers based on derived
147 predictors from a coupled bio-physical model that provided zooplankton concentrations were an
148 improvement over similar classifiers that used remotely sensed environmental conditions and
149 chlorophyll. That is, the mechanistic model included an ocean circulation model, which provided
150 environmental predictors, coupled with a biological model that provided predictors of secondary
151 production.

152

153 We tested classifier performance using four species that occur widely in the California
154 Current system. They were northern anchovy (hereafter anchovy), Pacific hake *Merluccius*
155 *productus* (hereafter hake), Jack mackerel *Trachurus symmetricus*, and Pacific sardine (hereafter
156 sardine). These species provide good test cases because: 1) they have relatively large geographic
157 ranges in which habitat conditions and species distributions fluctuate greatly; 2) they are
158 important forage fish during all or part of their lifecycle and, thus, provide an important link
159 between secondary production and higher trophic levels; and 3) They have been sampled
160 extensively in scientific surveys as eggs and adults, thereby providing two data sets with which
161 to test classifier performance. The four species occur from the Gulf of Alaska (Jack mackerel and
162 sardine) or northern Vancouver Island (anchovy and hake) to the tip of Baja California and in the
163 Gulf of California (MacCall and Stauffer, 1983; Cohen et al., 1990; Froese and Pauly, 2017).
164 Anchovy usually are found within 20 nm of shore but may extend offshore as far as 250 nm
165 (Froese and Pauly, 2017; U.S. Department of Commerce, 2017a). Hake are primarily restricted
166 to the continental shelf, usually within 200 nm of shore (Cohen et al., 1990). Jack mackerel and
167 sardine are more wide ranging with larger animals commonly moving more than 200 nm
168 offshore (Macewicz and Abramenkoff, 1993; U.S. Department of Commerce, 2017a). All four
169 species spawn during the late winter or spring and release free-floating eggs, which become
170 buoyant shortly after fertilization. Eggs hatch within a few days and typically are captured within
171 short distances of the spawning grounds. Thus, egg surveys of these species have been used with
172 relatively good accuracy to characterize their spawning habitats (e.g., Checkley et al., 2000;
173 Weber and McClatchie, 2010; Zwolinski et al., 2011). We tested classifier performance using
174 relatively data-rich surveys of egg densities during spring and more data-sparse trawl surveys of
175 adults conducted in spring and summer.

176

177 **2. Materials and methods**

178

179 *2.1 Fish Data*

180

181 Eggs were collected using the continuous underway fish-egg sampler (CUFES; Checkley
182 et al., 1997) during spring cruises offshore of the U.S. west coast and Vancouver Island from
183 2001 -2016 as part of the California Cooperative Oceanic Fisheries Investigations program
184 (CalCOFI; cf, McClatchie, 2014) and assessment cruises for coastal pelagic species by the U.S.

185 National Marine Fisheries Service. The CUFES collected eggs through a pump on the ship's hull
186 at a depth of 3 m and rate of 0.63 m³/min, using a sieve of 200-micron mesh. Samples were
187 collected every 30 min as the ship was underway at cruising speeds of 9–12 knots. The area from
188 30–39° N (approximately from San Diego to San Francisco, California) and 117–126° W was
189 sampled consistently, and additional sampling farther north occurred during some years (U.S.
190 Department of Commerce, 2017a). Cruises occurred in March through May but usually were
191 centered in April. The sampling depth was consistent with the near-surface distribution of
192 anchovy, hake, Jack mackerel, and sardine eggs, which typically concentrate in the upper mixed
193 layer (Ahlstrom, 1959). Eggs of these four species were enumerated and identified at sea based
194 on the morphometric characteristics (Moser, 1996).

195 Adults were sampled by trawl as part of assessment cruises for coastal pelagic species by
196 the U.S. National Marine Fisheries Service in spring 2002–2015, as part of the cruises described
197 above. Additional trawl samples were conducted in summer of 2003–2004, 2007–2008, and
198 2012–2015. Summer trawls were conducted primarily off the northwest coast of the U.S. from
199 northern California to Washington state before 2008, and for the entire U.S. west coast from
200 2012–2015. Data were collected using a Nordic 264 surface trawl (NET Systems; Bainbridge
201 Island, WA) with 3.0 m² XL-Lite foam-core alloy midwater doors and a working mouth opening
202 of about 600 m² at towing speed. The trawl had variable-size mesh in the throat to retain a range
203 of animal sizes and an 8-mm mesh liner in the codend. Since 2009, the trawl has been fitted with
204 a marine-mammal excluder device to prevent the capture of dolphins and other large animals
205 (Dotson et al., 2010). Trawls were towed at a target speed of four knots for 30 or 45 min between
206 sunset and sunrise, when coastal pelagic species migrate to the surface to feed. Typically, two or
207 three tows were conducted per night with spacing of at least ten nm. Trawls were targeted to
208 areas where coastal pelagic species were likely to occupy based on acoustic data collected during
209 the day (Zwolinski et al., 2017). That is, trawl samples were not random or necessarily
210 representative of the range of conditions in the overall survey area.

211 *2.2 ROMS-CoSiNE data*

212

213 Ocean simulations were conducted using the Regional Ocean Modeling System (ROMS;
214 Shchepetkin and McWilliams, 2005; Shchepetkin and McWilliams, 2009). The ROMS is a free-
215 surface primitive equation ocean model that is discretized in terrain-following curvilinear
216 coordinates that has frequently been used for modeling transport of larvae. The Pacific basin-
217 wide model used for this study was described by Wang and Chao (2004). It had 30 vertical
218 layers and 1/8 degree horizontal resolution over a domain of approximately 45°S to 65°N latitude
219 and 100°E to 70°W longitude.

220 The Pacific Ocean ROMS was coupled with the Carbon, Silicate, Nitrogen Ecosystem
221 model, known as CoSiNE (Chai et al., 2002; Liu and Chai, 2009). The CoSiNE model consisted
222 of 13 state variables describing plankton (meso-zooplankton, micro-zooplankton, small
223 phytoplankton and diatoms), nutrients (nitrate, silicate, ammonium, detritus nitrogen, detritus
224 silicate, phosphate), and others (dissolved oxygen, total carbon dioxide and total alkalinity). The
225 model was modified to simulate phyto-planktonic photo-acclimation and the dynamic carbon to
226 chlorophyll and carbon to nitrogen ratios with different growth conditions (Xiu and Chai, 2012).

227 The coupled ROMS-CoSiNE model was initiated with climatological temperature,

228 salinity, and nutrients from the World Ocean Atlas 2001 database (Conkright et al., 2002), total
229 carbon dioxide and total alkalinity from the Global Ocean Data Analysis Project database (Key
230 et al., 2004). It was forced with climatological monthly heat and wind from National Centers for
231 Environmental Prediction and National Center for Atmospheric Research (NCEP/NCAR)
232 reanalysis (Kalnay et al., 1996) for 30 years as a spin-up period. Then the coupled ROMS-
233 CoSiNE model was run from 1991-2016 with daily heat flux, evaporation and precipitation from
234 NCEP/NCAR reanalysis and daily sea surface wind (0.25° resolution) from National
235 Oceanographic and Atmospheric Administration multiple-satellite blended sea surface winds
236 (Zhang et al., 2006). No data assimilation was implemented in this ROMS configuration.

237 The performance of the Pacific Ocean ROMS-CoSiNE model has been evaluated and
238 further constrained with different satellite and in-situ data (Bidigare et al., 2009; Chai et al.,
239 2009; Liu and Chai, 2009; Xiu and Chai, 2011; Palacz and Chai, 2012). The model solutions for
240 the central California Current System have been demonstrated to capture spatial variations in the
241 annual means and seasonal cycles of temperature, nutrients, chlorophyll, and primary production
242 based on comparisons with remote sensing (SeaWiFS) and *in situ* observations from CalCOFI
243 (Guo et al., 2014). The mean annual difference between sea-surface temperatures measured at
244 sea in conjunction with CUFES sampling and coincident ROMS predictions was 0.01°C (± 0.01
245 SE). The difference was 0.42 (± 0.01 SE) for salinity (PSS). Santora et al. (2013) also reported
246 that the Pacific ROMS-CoSiNE model captured phytoplankton and zooplankton dynamics
247 reasonably well in the central California Current System.

248 Three-day averaged model output from the ROMS-CoSiNE model in the California
249 Current area was used to provide predictors for the classification algorithm. A subset of discrete
250 solutions from the ROMS-CoSiNE model was extracted as a three-dimensional array at 5 m
251 depth for each year and environmental variable, where environmental variables were small
252 phytoplankton ($< 5\ \mu\text{m}$ diameter), diatoms, microzooplankton, mesozooplankton, salinity,
253 temperature, dissolved oxygen concentration, the zonal component of the current, and the
254 meridional component of the current. The depth component was interpolated from the ROMS
255 vertical layers using the Python module 'roppy' (Ådlandsvik, 2016). The spatial dimensions of
256 the study were 300 latitudes and 144 longitudes that encompassed the area $20\text{-}50^\circ\text{N}$ and 110-
257 130°W . The values of environmental variables to which eggs or adults were exposed were
258 estimated by linear interpolation between the nearest values in space and time.

259 2.3 Satellite Data

260
261 We obtained remotely sensed satellite data for the same geographic area as the ROMS
262 model from the U.S. NOAA-Fisheries Coastwatch ERDDAP data server (U.S. Department of
263 Commerce, 2017b) and European Commission Copernicus Marine Environment Monitoring
264 Service (CMEMS; 2017). Sea-surface temperature (SST) measurements for the period 2001-
265 2012 were obtained from the “Pathfinder Ver 5.2 (L3C), Day and Night, Global, 0.0417° ,
266 Science Quality (Monthly Composite)” data set, Coastwatch ID “erdPH2sstamday”. For the
267 period 2013-2016, we used the “SST, POES AVHRR, GAC, Global, Day and Night (Monthly
268 Composite), Lon \pm 180” data set, Coastwatch ID “erdAGsstamday_LonPM180”. Chlorophyll-
269 estimates based on surface reflectance were obtained from the “Chlorophyll-a, Orbview-2
270 SeaWiFS, R2014.0, 0.1° , Global (Monthly Composite)”, Coastwatch data set, ID
271 “erdSW1chlamday”, for the period 2001-2002, and “Chlorophyll-a, Aqua MODIS, NPP, L3SMI,

272 Global, 4km, Science Quality (Monthly Composite)”, Coastwatch data set ID
273 “erdMH1chlamday” for the period 2003-2016. Bottom depth measurements were obtained from
274 the “SRTM30+ Version 6.0, 30 arc second, Global” Coastwatch data set, ID “usgsCeSrtm30v6”
275 (Becker et al., 2009). Surface currents were obtained from the “Global Observed Ocean Physics
276 Temperature Salinity Heights Geostrophic Currents Sea Surface Salinity and Sea Surface
277 Density Reprocessing” data set, with CMEMS identifier “GLOBAL_REP_PHY_001_021”
278 (Mulet et al., 2012). Sample values were estimated from satellite data by linear interpolation.

279

280 *2.4 Random Forests*

281

282 We used the Random Forest algorithm (Breiman, 2001) to predict two types of habitat
283 responses as a function of two types of predictors for each species (i.e., four classifiers for each
284 of the four species). The random forest algorithm is an ensemble technique that produces
285 predictions based on a collection of classification or regression trees (the “forest”). It is
286 particularly appropriate for fitting a descriptive model of habitat use because the technique can
287 capture simple interactions among predictors without the user explicitly specifying them, is
288 relatively robust to the inclusion of correlated variables, and tends to avoid over fitting. The
289 response variables were the presence of eggs in CUFES samples or adults in trawl samples
290 (hereafter, positive predictions are interpreted to be suitable spawning habitat or adult habitat,
291 respectively). The two predictor types were derived variables from the ROMS-CoSiNE model or
292 remotely sensed satellite data.

293 The derived predictors from the ROMS-CoSiNE model were sea-surface salinity, sea-
294 surface temperature, and concentrations of diatoms, dissolved oxygen, micro-zooplankton, meso-
295 zooplankton, and small phytoplankton, as described above. An additional predictive variable,
296 eddy kinetic energy (EKE), was calculated from the zonal (U) and meridional (V) components of
297 the ROMS model as:

$$298 \quad EKE = \frac{U^2 + V^2}{2} \quad (1)$$

299 The EKE predictor was a measure of turbulent flow used in the sardine habitat model of Nieto et
300 al. (2014), and very similar to the gradient in sea-surface height used by Zwolinski et al. (2011).
301 We also included bottom depth in the ROMS-CoSiNE random forest classifiers. For ROMS-
302 CoSiNE classifiers where trawl captures were the response, month of sample was included as a
303 predictor to fit potential changes in habitat over the relatively long sampling period (spring
304 through summer).

305

306 The predictors for random forests based on satellite data were bottom depth, EKE, the
307 natural log of chlorophyll concentration, and sea-surface temperature. Month of sample was also
308 included in classifiers where trawl captures were the response variable. The zonal and meridional
309 components of the current that were used to calculate EKE in the satellite-based classifiers were
310 from the CMEMS data rather than modeled ROMS data.

311

312 Hereafter, we refer to the random forest classifiers of spawning habitat based on CUFES
313 data that used derived predictors from the ROMS-CoSiNE model as CUFES-ROMS-CoSiNE
314 classifiers and similar random forests that were based on remotely sensed satellite data as

315 CUFES-satellite classifiers. Similarly, we refer to random forest classifiers of adult habitat based
316 on trawls as trawl-ROMS-CoSiNE classifiers or trawl-satellite classifiers.

317 We reserved data from four years, 2003, 2007, 2012, and 2015, as a test data set to
318 evaluate the extent to which the random forest out-of-bag estimates of classifier fit successfully
319 predicted classifier performance on independent data. This left 12 years of CUFES data available
320 for model development, and 11 years of trawl data, including four years of trawl data that were
321 sampled during summer. Available sample sizes were somewhat smaller for satellite-based
322 classifiers than for ROMS-CoSiNE based classifiers due to cloud cover and lack of predictions
323 for nearshore samples. A total of 12,550 CUFES samples were used to develop the CUFES-
324 ROMS-CoSiNE classifiers, and 4,196 to test them. Only 10,401 samples were available to
325 develop the CUFES-satellite classifiers, and 3,735 samples to test them. For trawl-based
326 classifiers, 1,000 development samples and 507 test samples were available for the trawl-ROMS-
327 CoSiNE classifiers. There were 745 development samples, and 407 test samples available for the
328 trawl-satellite classifiers.

329
330 We fitted the random forest classifiers with 5,000 trees using the “ranger” package
331 version 0.9.11 (Wright and Ziegler, 2017) in the R statistical computing environment version
332 3.5.0 (R Core Team, 2018). Each classifier was tuned using several steps before conducting a
333 final fitting. First, we selected the number of predictors sampled for splitting at each node (i.e.,
334 used for a particular tree; ranger parameter “mtry”) by optimizing with respect to out-of-bag
335 error using the “train” function of R package “caret” version 6.0-79 (Kuhn, 2008). Next, we
336 performed predictor selection for each classifier using the Boruta algorithm (Kursa and
337 Rudnicki, 2010), as implemented in in R package “Boruta” version 5.3.0. The Boruta algorithm
338 identifies relevant predictor variables from a candidate set by comparing their predictive abilities
339 to “shadow” variables that are obtained by randomly shuffling the values of original variables
340 across objects. Shadow variables can have non-zero importance only due to random fluctuation
341 (where importance is the normalized mean decrease in accuracy of all random forests without the
342 variable). Variables that have significantly greater importance than the best shadow variable over
343 all permutations are judged to be relevant. We performed the selection using function “Boruta”
344 with 5,000 trees and the “mtry” parameter determined as described above. Variables identified as
345 unimportant were dropped from final classifiers. For all species and gear types, the number of
346 samples where eggs or adults were absent was much greater than the number of samples where
347 they were present. Therefore, we adjusted the “sample.fraction” parameter of the “ranger”
348 function to reduce the number of samples with absences used in each bootstrap resample until
349 the algorithm produced nearly equal numbers of false positives and false negatives. That is, we
350 tuned the model so that the two types of misclassification errors had equal importance (Berk,
351 2008).

352
353 Classifiers were evaluated using three statistics: accuracy, sensitivity, and specificity
354 (Yerushalmy, 1947; Trevethan, 2017). Accuracy was calculated as the proportion of correct
355 classifications as:

$$(True\ positives + True\ negatives) / Number\ samples\ (2)$$

356
357
358
359 Sensitivity, also known as probability of detection or ability to detect a true positive, was
360 calculated as:

361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406

$$\text{True positives} / (\text{True positives} + \text{False negatives}) \text{ (3)}$$

Specificity, also known as the true negative rate or ability to detect a true negative, was calculated as:

$$\text{True negatives} / (\text{True negatives} + \text{False positives}) \text{ (4)}$$

3. Results

All candidate predictor variables were accepted for use in random forest classifiers except one case (Figure 1). The EKE predictor for hake in the satellite-trawl classifier was rejected as having no greater predictive ability than the randomly shuffled shadow variables (as measured by normalized mean decrease in accuracy) and, thus, excluded from the classifier. The EKE predictor was also a relatively weak, but included, predictor for anchovy and sardines in the trawl-ROMS-CoSiNE classifiers. Variable importance estimates for the ROMS-CoSiNE-based classifiers were much more precise than those for satellite-based classifiers due to the larger sample sizes available. Sea-surface temperature was a relatively important predictor for all species. The effects of bottom-up production, as represented by the chlorophyll predictor, were very important in most of the satellite-based classifiers. However, the effects of primary and secondary production were partitioned among the diatom, micro-zooplankton, meso-zooplankton, and small phytoplankton variables in the ROMS-CoSiNE-based models.

Accuracies of ROMS-CoSiNE based classifiers generally were similar to those of their satellite-based counterparts. Accuracies of CUFES-ROMS-CoSiNE classifiers were all within 2% of CUFES-satellite classifiers for the same species when applied to the test data (Figure 2A). The trawl-ROMS-CoSiNE classifiers performed somewhat worse on test data than comparable trawl-satellite classifiers for anchovy and hake (2-8% less accuracy), but somewhat better for Jack mackerel and sardine (4-7%; Figure 2B). Classifiers of adults captured in trawls generally were less accurate than comparable classifiers of spawning habitat from CUFES data, regardless of predictor type. All classifiers had greater out-of-bag accuracies than test accuracies. However, there were relatively large species-specific differences among classifiers. The decline in accuracy between out-of-bag and test estimates was greatest for Jack mackerel and sardine (4-19%) and least for anchovy and hake (1-14%).

The presence of anchovy and hake could be predicted with greater accuracy than the presence of Jack mackerel or sardine for both CUFES and trawl data sets. The CUFES-based classifiers had accuracy greater than 0.90 for anchovy and hake but only 0.66-0.87 for Jack mackerel and sardine. The trawl-based classifiers had accuracies of 0.75-0.95 for anchovy and hake but 0.61-0.75 for Jack mackerel and sardine.

Specificity (i.e., ability to detect a true negative) was very good for most classifiers that used either ROMS-CoSiNE or satellite data (Figure 2) but sensitivity (i.e., probability of detection or ability to detect a true positive) was poor (0.00-0.61 for the test data set). A few classifiers for anchovy and hake had sensitivity values at or near zero with corresponding specificities at or near one. These were the results of very few captures. That is, the classifiers

407 correctly predicted that the species tested would not occur in nearly all samples because the
408 samples occurred in habitat that was poor for these species. However, the classifiers did not
409 correctly predict the species' presence in the small number of samples that occurred within
410 suitable habitat.

411
412 The spatial patterns in habitat predicted by the CUFES-ROMS-CoSiNE and CUFES-
413 satellite classifiers generally were similar for the test data set (Figure 3). Predicted habitats for
414 Jack mackerel and sardine were relatively large and extended far offshore. Lack of satellite
415 coverage due to cloud cover was substantial in 2003 and 2007, and likely precluded the satellite
416 classifiers from identifying some habitat for Jack mackerel and sardine. Predicted habitats for
417 anchovy and hake were constricted to small areas near the coast, with almost no spawning
418 habitat for hake occurring by mid-April. However, the classifiers exhibited some differences for
419 anchovy when extrapolating to the south into unsampled areas along Baja California. The
420 CUFES-satellite classifier for anchovy predicted that suitable habitat extended south along most
421 of Baja California where none was predicted by the CUFES-ROMS-CoSiNE classifier in three
422 of four test years. These differences could not be evaluated with the existing data. No strong
423 trend in the spatial distribution of classifier errors (false positives versus false negatives) was
424 evident for anchovy, Jack mackerel or sardine for any classifier type (Figure 4). Most false
425 positives for hake occurred off the central California coast or in the Southern California Bight.
426 However, there were too few false negatives for hake to evaluate their spatial distribution.

427
428 The temporal patterns in habitat use differed somewhat for adults sampled in trawls.
429 Month of sampling was a moderate to weakly important predictor in the trawl-based classifiers
430 (Figure 1). However, the fact that it was sufficiently important to be retained as a predictor
431 variable indicates that adults occurred in waters with somewhat different habitat conditions
432 during spring versus summer. The distribution of predictors from the ROMS-CoSiNE model
433 where adults were captured (Figure 5) indicated that all species tended to occupy a narrower
434 range of bottom depths in generally shallower water during summer, but a broader range of
435 conditions otherwise. The summer distributions of adults from all species tended to occur in
436 waters with somewhat greater dissolved oxygen, mesozooplankton concentrations, and sea-
437 surface temperatures, but lower sea-surface salinities.

438 439 **4. Discussion**

440
441 The generally similar accuracies between derived classifiers and those of comparable
442 satellite-based classifiers (Figure 2) indicates that derived classifiers may be a useful tool where
443 the limitations of satellite-based classifiers prevent their application. Derived classifiers may be
444 useful where cloud cover severely limits satellite coverage or no lost pixels can be tolerated. The
445 ability to create real-time or near future predictions may be particularly useful for adaptively
446 planning long surveys and dynamic management of fisheries (Becker et al., 2016). These results
447 are consistent with the development process for ocean circulation models. Measured data from
448 satellites and other sensors commonly are used to evaluate the performance of ocean circulation
449 models, as they were for the ROMS-CoSiNE model (*cf.*, section 2.2). Derived predictors were
450 close to their measured counterparts, and produced similar classifiers, because the underlying
451 mechanistic model reproduced measured environmental conditions well. The use of ocean

452 circulation models to predict habitat has now been demonstrated for fish (this study) and
453 cetaceans (Becker et al., 2016).

454
455 The fact that derived classifiers did not have consistently greater accuracies than
456 comparable satellite-based classifiers indicates there was no particular advantage to using
457 zooplankton predictors in the coupled biological portion of the ROMS-CoSiNE model over
458 primary production, as measured by satellite. We hypothesized that the small- and large-
459 zooplankton predictors might provide derived classifiers with greater accuracy than comparable
460 satellite predictors because zooplankton is a more direct measure of fish habitat than chlorophyll.
461 This is because the species tested consume plankton directly, although hake and Jack mackerel
462 switch to piscivory in larger sizes (MacGregor, 1966; Rexstad and Pikitch, 1986). There are
463 several non-exclusive explanations for the lack of model improvement. Inadequate precision of
464 zooplankton predictors may have offset potential gains. Biological interactions may have created
465 differences between the true and predicted concentrations of zooplankton available because
466 modeled zooplankton predictors were calculated based on bottom-up trophic conditions only.
467 Top-down and community interactions also may have been more important in determining
468 habitat conditions than zooplankton. The consistently high importance of variables in trawl
469 models that are not directly consumed by adults (e.g., diatoms) suggests that production variables
470 were, in part, indexing other non-modeled habitat conditions that affected fish distribution rather
471 than prey availability directly. However, we note that importance of predictors related to bottom-
472 up productivity relative to each other (Figure 2) should be interpreted with caution because they
473 are partly correlated and interacting with each other (Kursa and Rudnicki, 2010).

474
475 The degree to which models were overfit, as measured by the decrease in accuracy
476 between out-of-bag- and test-data statistics for comparable classifiers, varied greatly by species,
477 predictor type, and gear (Figure 2). Classifiers for anchovy and hake captured in the CUFES
478 showed little overfitting, as evidenced by declines of 1-4% (e.g., the CUFES-ROMS-CoSiNE
479 classifier for anchovy had out-of-bag accuracy of 0.95 and accuracy on test data of 0.91). To the
480 contrary, Jack mackerel and sardine captured in the CUFES exhibited the greatest degree of
481 overfitting, with declines in accuracy of 4-19%. The pattern for classifiers based on trawl data
482 was opposite that of the CUFES classifiers. Anchovy and hake exhibited the greatest declines in
483 accuracy (7-14%) for trawl-based classifiers, and Jack mackerel and sardine the least (4-7%).
484 These results suggest the degree to which overfitting occurred may have been a function of the
485 number and type of predictors, sample sizes, and species distributions. Even for a relatively long
486 time period with considerable variation in conditions (12,550 observations and 12 years of
487 developmental data over a 16-year period for CUFES data), the algorithm may be fitting false
488 correlations because too few animals have been captured over the entire range of values needed
489 to characterize their tolerance for some predictors. (i.e., the edges of a species' tolerance; Myers,
490 1998). Thus, it may not be obvious when overfitting is likely to be a problem when creating a
491 habitat model or classifier for a novel species. We note that the expected accuracy of a classifier
492 with no predictive ability (i.e., a random guess) is 0.50. A decline in accuracy from 0.87 to 0.68
493 such as that found for sardine in the CUFES-ROMS-CoSiNE classifier is the difference between
494 a very good predictive algorithm and a poor one. These results suggest independent data may be
495 needed to realistically quantify the predictive power of fish habitat models or algorithms if they
496 are to be used for prediction.

497

498 Habitat models that predict animal presence based on environmental conditions have
499 some limitations regardless of whether the predictors are measured or modeled. Even if all
500 important habitat conditions that define a species niche are included in a model, its accuracy will
501 not be perfect because the model predicts potential habitat rather than realized habitat (Planque et
502 al., 2007). The effect of unoccupied potential habitat was evident in this study because classifiers
503 generally had poor sensitivity (i.e., probability of detection) but good specificity (i.e., true
504 negative prediction rate). To achieve greater accuracy, it may be necessary to include animal
505 births, deaths, and movement that determine the subset of appropriate habitat that actually is
506 occupied. The hybrid approach to modeling habitat that uses derived predictors from a
507 mechanistic model could be further refined by including additional population dynamics to the
508 mechanistic model. For example, starting locations of fish could be estimated based on a
509 previous survey, and then movement and survival could be included to improve model accuracy.
510 The limitations of habitat models also have implications for their use in survey design. We
511 assumed that both types of classification errors, false positive or false negative, were equally
512 important in this study and tuned the random forests to balance them (via the sample fractions
513 selected in each class; cf. Methods 2.4). However, if sensitivity is likely to be poor in any case
514 because the model is predicting potential rather than realized habitat, it may be better to train
515 classifiers to maximize specificity at the expense of sensitivity and then stratify by predicted
516 habitat conditions. That is, minimize the area of the stratum in which good habitat exists, and in
517 which the greatest sampling effort is required.

518

519 Classifiers of adults suggested that a broader range of habitat conditions were used during
520 summer than during the spring spawning period for all species (Figures 1 and 5). The effect was
521 most subtle for sardines, whose migration pattern allows them to use similar habitat throughout
522 the year despite seasonal changes (Zwolinski et al., 2011). The other three species tested also
523 stay in appropriate habitat via migration to varying extents. Thus, it has sometimes been assumed
524 that environmental conditions in the spawning habitat represent the same habitat conditions used
525 by the species throughout the year. Our results indicate that habitat models based on spawning
526 habitat (egg presence) should be used with caution to predict the species' presence outside of the
527 spawning season.

528 Our ability to identify adult habitat outside of the spawning season was relatively poor for
529 the species tested because trawl-based classifiers were less accurate than CUFES-based
530 classifiers (Figure 2). This problem could partly be solved with additional trawl sampling. The
531 trawl data set was much sparser than the CUFES data set, and the non-random selection of trawl
532 locations likely reduced the range of habitat conditions sampled (cf. Methods 2.1). However,
533 classifiers based on trawl samples may inherently suffer from worse sensitivity than CUFES-
534 based classifiers because adults actively avoid the trawl to some extent. Classifier accuracy also
535 was relatively poor for the two species with greatest range and mobility, Jack mackerel and
536 sardine. In combination, trawl-based classifiers for Jack mackerel and sardine probably were not
537 accurate enough for most applications. For example, the trawl-satellite classifier for sardine (0.61
538 accuracy) performed only somewhat better than a random choice with expected accuracy of 0.50.
539 It may be most cost effective to survey these species during the spawning season because they
540 are likely to be aggregated in areas that can be better identified *a priori*.

541

542 **5. Conclusions**

543

544 Habitat classifiers developed using derived (modeled) predictors from a coupled
545 biophysical model had similar predictive ability to those of similar satellite-based classifiers for
546 four species of fish. However, there was no improvement in the accuracy of derived predictors
547 that included modeled zooplankton concentrations, which are not available from satellite data,
548 over comparable satellite classifiers that included reflectance/chlorophyll concentration.
549 Derived- and satellite-based classifiers both tended to have good specificity (i.e., ability to
550 identify true negatives) but poor sensitivity (i.e., ability to detect a true positive) because
551 classifiers identified potential habitat which was not fully occupied. Classifiers for two most
552 wide ranging and mobile species tested, Jack mackerel and sardine, were less accurate than those
553 of the species with more restricted ranges of habitat, anchovy and hake. Derived classifiers have
554 advantages over satellite classifiers that make them well suited for dynamic species management
555 and survey planning. They are not limited by cloud cover and they can make predictions in near
556 real-time or the short-term future.

557

558 **Acknowledgements**

559

560 We thank G. DiNardo, T. Garfield, E. Hofmann, S. McClatchie, and one anonymous
561 reviewer for reviewing the manuscript. This research was supported by the U.S. National
562 Oceanic and Atmospheric Administration, Remote Sensing Solutions, Inc., and the University of
563 Maine. Development of the ROMS-CoSiNE has previously been supported by the Jet Propulsion
564 Laboratory of the California Institute of Technology, U.S. National Aeronautics and Space
565 Administration, U.S. National Science Foundation, and others, as described in the cited
566 references.

567

568 **References**

569

570 Ådlandsvik, Bjorn. 2016. Ropyy ROMS post-processing tools in python.

571 <https://github.com/bjornaa/roppy>. Accessed 4/20/2016.

572 Ahlstrom, E. H. 1959. Vertical Distribution of Pelagic Fish Eggs and Larvae, Off California and
573 Baja California. Fishery Bulletin of the Fish and Wildlife Service, 60: 107-143.

574 Becker, E. A., Forney, K. A., Fiedler, P. C., Barlow, J., Chivers, S. J., Edwards, C. A., Moore, A.
575 M., et al. 2016. Moving Towards Dynamic Ocean Management: How Well Do Modeled
576 Ocean Products Predict Species Distributions? Remote Sensing, 8: 26.

577 Becker, J. J., Sandwell, D. T., Smith, W. H. F., Braud, J., Binder, B., Depner, J., Fabre, D., et al.
578 2009. Global Bathymetry and Elevation Data at 30 Arc Seconds Resolution:

579 SRTM30_PLUS. Marine Geodesy, 32: 355-371.

580 Berk, R. A. 2008. Statistical Learning from a Regression Perspective, Springer.

581 Bidigare, R. R., Chai, F., Landry, M. R., Lukas, R., Hannides, C. C. S., Christensen, S. J., Karl,
582 D. M., et al. 2009. Subtropical ocean ecosystem structure changes forced by North
583 Pacific climate variations. Journal of Plankton Research, 31: 1131-1139.

584 Breiman, L. 2001. Random forests. Machine Learning, 45: 5-32.

585 Chai, F., Dugdale, R. C., Peng, T. H., Wilkerson, F. P., and Barber, R. T. 2002. One-dimensional
586 ecosystem model of the equatorial Pacific upwelling system. Part I: model development
587 and silicon and nitrogen cycle. Deep-Sea Research Part II-Topical Studies in
588 Oceanography, 49: 2713-2745.

589 Chai, F., Liu, G. M., Xue, H. J., Shi, L., Chao, Y., Tseng, C. M., Chou, W. C., et al. 2009.
590 Seasonal and Interannual Variability of Carbon Cycle in South China Sea: A Three-
591 Dimensional Physical-Biogeochemical Modeling Study. *Journal of Oceanography*, 65:
592 703-720.

593 Checkley, D. M., Dotson, R. C., and Griffith, D. A. 2000. Continuous, underway sampling of
594 eggs of Pacific sardine (*Sardinops sagax*) and northern anchovy (*Engraulis mordax*) in
595 spring 1996 and 1997 off southern and central California. *Deep-Sea Research Part II-*
596 *Topical Studies in Oceanography*, 47: 1139-1155.

597 Checkley, D. M., Ortner, P. B., Settle, L. R., and Cummings, S. R. 1997. A continuous,
598 underway fish egg sampler. *Fisheries Oceanography*, 6: 58-73.

599 Cohen, D. M., Inada, T., Iwamoto, T., and Scialabba, N. 1990. FAO Species Catalogue. Vol. 10.
600 Gadiform fishes of the world (Order Gadiformes). An annotated and illustrated catalogue
601 of cods, hakes, grenadiers and other gadiform fishes known to date. FAO Fisheries
602 Synopsis No. 125(10). Rome, FAO. 442 pp.

603 Conkright, M. E., Locarnini, R. A., H. E. Garcia, T. D. O. B., Boyer, T. P., and C. Stephens, J. I.
604 A. 2002. World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures, CD-
605 ROM Documentation. p. 17. National Oceanographic Data Center, Silver Spring, MD.

606 Dotson, R. C., Griffith, D. A., King, D. L., and Emmett, R. L. 2010. Evaluation of a Marine
607 Mammal Excluder Device (MMED) for a Nordic 264 Midwater Rope Trawl. ICES
608 Document NOAA-TM-NMFS-SWFSC-455. 14 pp.

609 Fiechter, J., Rose, K. A., Curchitser, E. N., and Hedstrom, K. S. 2015. The role of environmental
610 controls in determining sardine and anchovy population cycles in the California Current:
611 Analysis of an end-to-end model. *Progress in Oceanography*, 138: 381-398.

612 Froese, R. and D. Pauly, editors. 2017. FishBase. www.fishbase.org. Accessed 8/19/2017.

613 Fulton, E. A. 2010. Approaches to end-to-end ecosystem models. *Journal of Marine Systems*, 81:
614 171-183.

615 Guo, L., Chai, F., Xiu, P., Xue, H. J., Rao, S., Liu, Y. G., and Chavez, F. P. 2014. Seasonal
616 dynamics of physical and biological processes in the central California Current System:
617 A modeling study. *Ocean Dynamics*, 64: 1137-1152.

618 Hayes, D. B., Ferreri, C. P., and Taylor, W. W. 1996. Linking fish habitat to their population
619 dynamics. *Canadian Journal of Fisheries and Aquatic Sciences*, 53: 383-390.

620 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., et al.
621 1996. The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American*
622 *Meteorological Society*, 77: 437-471.

623 Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., Feely, R. A., et
624 al. 2004. A global ocean carbon climatology: Results from Global Data Analysis Project
625 (GLODAP). *Global Biogeochemical Cycles*, 18: 23.

626 Kishi, M. J., Ito, S., Megrey, B. A., Rose, K. A., and Werner, F. E. 2011. A review of the
627 NEMURO and NEMURO.FISH models and their application to marine ecosystem
628 investigations. *Journal of Oceanography*, 67: 3-16.

629 Kuhn, M. 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical*
630 *Software*, 28: 1-26.

631 Kursu, M. B., and Rudnicki, W. R. 2010. Feature Selection with the Boruta Package. *Journal of*
632 *Statistical Software*, 36: 1-13.

- 633 Leitão, F. 2015. Time series analyses reveal environmental and fisheries controls on Atlantic
634 horse mackerel (*Trachurus trachurus*) catch rates. *Continental Shelf Research*, 111: 342-
635 352.
- 636 Lindegren, M., Checkley, D. M., Jr., Ohman, M. D., Koslow, J. A., and Goericke, R. 2016.
637 Resilience and stability of a pelagic marine ecosystem. *Proc Biol Sci*, 283.
- 638 Liu, G., and Chai, F. 2009. Seasonal and interannual variability of primary and export production
639 in the South China Sea: a three-dimensional physical-biogeochemical model study. *ICES*
640 *Journal of Marine Science*, 66: 420-431.
- 641 MacCall, A. D., and Stauffer, G. D. 1983. Biology and Fishery Potential of Jack Mackerel
642 (*Trachurus Symmetricus*). *California Cooperative Oceanic Fisheries Investigations*
643 *Reports*, 24: 46-56.
- 644 Macewicz, B. J., and Abramenkoff, D. N. 1993. Collection of Jack Mackerel, *Trachurus*
645 *symmetricus*, off Southern California during 1991 Cooperative U.S.-U.S.S.R. Cruise. p.
646 14. Ed. by N. O. a. A. A. U.S. Department of Commerce. Southwest Fisheries Science
647 Center, La Jolla, California.
- 648 MacGregor, J. S. 1966. Synopsis of the Biology of the Jack Mackerel (*Trachurus*
649 *symmetricus*). p. 16. U.S. Department of the Interior Bureau of Commercial Fisheries,
650 Washington, D.C., USA.
- 651 McClatchie, S. 2014. *Regional Fisheries Oceanography of the California Current System*,
652 Springer, New York. 235 pp.
- 653 Moser, H. G. 1996. *The Early Stage of Fishes in the California Current Region*, Southwest
654 Fisheries Science Center. 1505 pp.
- 655 Mulet, S., Rio, M. H., Mignot, A., Guinehut, S., and Morrow, R. 2012. A new estimate of the
656 global 3D geostrophic ocean circulation based on satellite data and in-situ measurements.
657 *Deep-Sea Research Part II-Topical Studies in Oceanography*, 77-80: 70-81.
- 658 Myers, R. A. 1998. When do environment—recruitment correlations work? *Reviews in Fish*
659 *Biology and Fisheries*, 8: 285-305.
- 660 Nieto, K., McClatchie, S., Weber, E. D., and Lennert-Cody, C. E. 2014. Effect of mesoscale
661 eddies and streamers on sardine spawning habitat and recruitment success off Southern
662 and central California. *Journal of Geophysical Research. C. Oceans*, 119: 6330-6339.
- 663 Palacz, A. P., and Chai, F. 2012. Spatial and temporal variability in nutrients and carbon uptake
664 during 2004 and 2005 in the eastern equatorial Pacific Ocean. *Biogeosciences*, 9: 4369-
665 4383.
- 666 Pecuchet, L., Tornroos, A., and Lindegren, M. 2016. Patterns and drivers of fish community
667 assembly in a large marine ecosystem. *Marine Ecology Progress Series*, 546: 239-248.
- 668 Planque, B., Bellier, E., and Lazure, P. 2007. Modelling potential spawning habitat of sardine
669 (*Sardina pilchardus*) and anchovy (*Engraulis encrasicolus*) in the Bay of Biscay.
670 *Fisheries Oceanography*, 16: 16-30.
- 671 R Core Team 2018. *R: A Language and Environment for Statistical Computing*. Version 3.5.0.
672 Vienna, Austria.
- 673 Rexstad, E. A., and Pikitch, E. K. 1986. Stomach Contents and Food-Consumption Estimates of
674 Pacific Hake, *Merluccius productus*. *Fishery Bulletin*, 84: 947-956.
- 675 Rose, K. A., Allen, J. I., Artioli, Y., Barange, M., Blackford, J., Carlotti, F., Cropp, R., et al.
676 2010. End-To-End Models for the Analysis of Marine Ecosystems: Challenges, Issues,
677 and Next Steps. *Marine and Coastal Fisheries*, 2: 115-130.

678 Rose, K. A., Fiechter, J., Curchitser, E. N., Hedstrom, K., Bernal, M., Creekmore, S., Haynie, A.,
679 et al. 2015. Demonstration of a fully-coupled end-to-end model for small pelagic fish
680 using sardine and anchovy in the California Current. *Progress in Oceanography*, 138:
681 348-380.

682 Santora, J. A., Sydeman, W. J., Messie, M., Chai, F., Chao, Y., Thompson, S. A., Wells, B. K., et
683 al. 2013. Triple check: Observations verify structural realism of an ocean ecosystem
684 model. *Geophysical Research Letters*, 40: 6.

685 Copernicus Marine Environment Monitoring Service. 2017. <http://marine.copernicus.eu>.
686 Accessed 8/11/2017.

687 Shchepetkin, A. F., and McWilliams, J. C. 2005. The regional oceanic modeling system
688 (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model.
689 *Ocean Modelling*, 9: 347-404.

690 Shchepetkin, A. F., and McWilliams, J. C. 2009. Ocean forecasting in terrain-following
691 coordinates: Formulation and skill assessment of the regional ocean modeling system
692 (vol 227, pg 3595, 2008). *Journal of Computational Physics*, 228: 8985-9000.

693 Thompson, A. R., Watson, W., McClatchie, S., and Weber, E. D. 2012. Multi-Scale Sampling to
694 Evaluate Assemblage Dynamics in an Oceanic Marine Reserve. *Plos One*, 7: 12.

695 Tommasi, D., Stock, C. A., Pegion, K., Vecchi, G. A., Methot, R. D., Alexander, M. A., and
696 Checkley, D. M. J. 2016. Improved Management of Small Pelagic Fisheries Through
697 Seasonal Climate Prediction. *Ecological Applications*, 27: 378-388.

698 Trevethan, R. 2017. Sensitivity, Specificity, and Predictive Values: Foundations, Pliabilities, and
699 Pitfalls in Research and Practice. *Front Public Health*, 5: 307.

700 U.S. Department of Commerce, National Oceanic and Atmospheric Administration. 2017a. Egg
701 Distribution Maps for Sardine, Anchovy, and Jack Mackerel.
702 <https://swfsc.noaa.gov/textblock.aspx?Division=FRD&id=1121>. Accessed 8/16/2017.

703 U.S. Department of Commerce, National Oceanic and Atmospheric Administration. 2017b.
704 NOAA Coastwatch West Coast Regional Node. <http://coastwatch.pfeg.noaa.gov/>.
705 Accessed 8/11/2017.

706 Vasconcelos, R. P., Eggleston, D. B., Le Pape, O., and Tulp, I. 2013. Patterns and processes of
707 habitat-specific demographic variability in exploited marine species. *ICES Journal of*
708 *Marine Science*, 71: 638-647.

709 Wang, X. C., and Chao, Y. 2004. Simulated Sea Surface Salinity variability in the tropical
710 Pacific. *Geophysical Research Letters*, 31.

711 Weber, E. D., and McClatchie, S. 2010. Predictive models of northern anchovy *Engraulis*
712 *mordax* and Pacific sardine *Sardinops sagax* spawning habitat in the California Current.
713 *Marine Ecology Progress Series*, 406: 251-263.

714 Wright, M. N., and Ziegler, A. 2017. ranger: A Fast Implementation of Random Forests for High
715 Dimensional Data in C plus plus and R. *Journal of Statistical Software*, 77: 1-17.

716 Xiu, P., and Chai, F. 2011. Modeled biogeochemical responses to mesoscale eddies in the South
717 China Sea. *Journal of Geophysical Research-Oceans*, 116.

718 Xiu, P., and Chai, F. 2012. Spatial and temporal variability in phytoplankton carbon, chlorophyll,
719 and nitrogen in the North Pacific. *Journal of Geophysical Research*, 117.

720 Yerushalmy, J. 1947. Statistical Problems in Assessing Methods of Medical Diagnosis, with
721 Special Reference to X-Ray Techniques. *Public Health Reports*, 62: 1432-1449.

722 Zhang, H.-M., Bates, J. J., and Reynolds, R. W. 2006. Assessment of composite global sampling:
723 Sea surface wind speed. *Geophysical Research Letters*, 33.

724 Zwolinski, J. P., Demer, D. A., Macewicz, B. J., Mau, S., Murfin, D., Palance, D., Renfree, J. S.,
725 et al. 2017. Distribution, Biomass, and Demography of the Central Stock of Northern
726 Anchovy During Summer 2016, Estimated From Acoustic-Trawl Sampling. ICES
727 Document NOAA-TM-NMFS-SWFSC-572. 18 pp.
728 Zwolinski, J. P., Emmett, R. L., and Demer, D. A. 2011. Predicting habitat to optimize sampling
729 of Pacific sardine (*Sardinops sagax*). ICES Journal of Marine Science, 68: 867-879.
730

731 **Figure Captions**

732
733 Figure 1. Box plots of normalized variable importance (decrease in variable importance when the
734 variable is not present divided by its standard deviation) for predictors in 5,000 trees to predict
735 the presence of four species. Black boxplots were included in final classifiers; gray boxplots
736 were not. The area inside each box represents the first and third quartiles. Whiskers extend to the
737 data range. Predictor variables were derived predictors from the ROMS-CoSiNE model (A and
738 B) or remotely sensed satellite data (C and D). Rows A and C indicate Random Forest classifiers
739 to predict the presence of eggs from CUFES samples, and rows (B and D) the presence of adults
740 from trawl samples. The Shadow Min, Mn, And Max columns are the minimum, mean, and
741 maximum distributions of the shadow variables used to perform variable selection. The shadow
742 variable is the best performing randomly shuffled variable whose values were obtained from the
743 original predictors (Kursa and Rudnicki, 2010). That is, they represent the distribution of a
744 variable that has no predictive value. Note the scales of y axes differ.
745

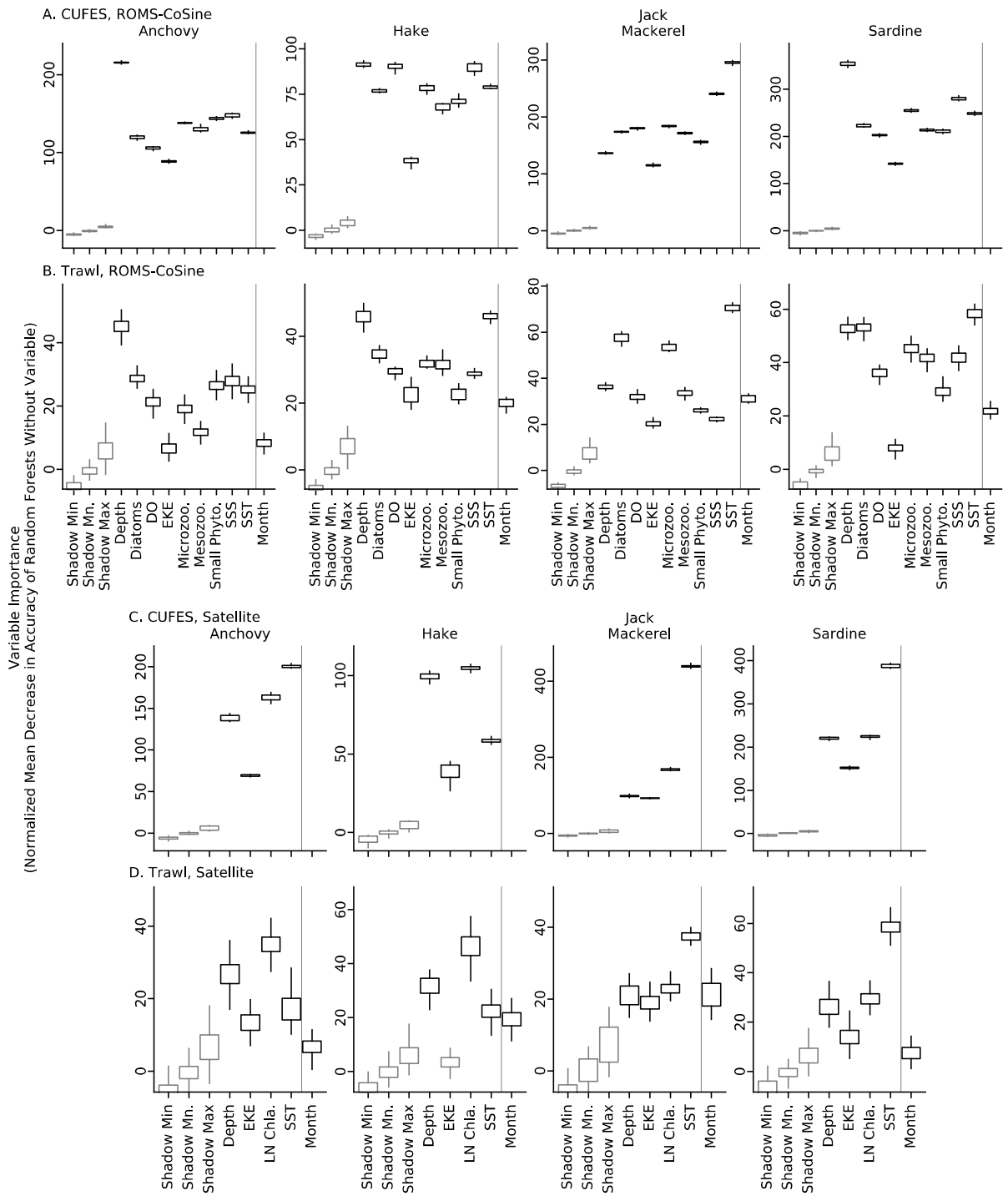
746 Figure 2. Accuracy, sensitivity, and specificity of Random Forests to predict the presence of eggs
747 from CUFES samples (A) and adults in trawls (B) for classifiers of four species (Anch=anchovy,
748 J Mack=Jack mackerel, Sard=sardine) using derived predictors from the ROMS-CoSiNE model
749 or remotely sensed satellite data (Sat). Areas of gray boxes are proportional to the magnitude of
750 the statistic in each cell. The OOB rows indicate statistics for out-of-bag samples in the
751 developmental data sets. The test rows indicate statistics for the independent test data set, years
752 2002, 2006, 2011, and 2015.
753

754 Figure 3. Predicted spawning habitat for four species on April 16 of each year in the test data set
755 (2002, 2006, 2011, and 2015) for Random Forest classifiers using either predictors derived from
756 the ROMS-CoSiNE model or remotely sensed satellite data. The color scale indicates the
757 predicted probability of capturing one or more eggs in a CUFES sample. White areas indicate
758 classifier predictions < 0.5, i.e., no eggs predicted. Blue indicates areas where no prediction
759 could be made for satellite-based classifiers because of cloud coverage or proximity to shore.
760

761 Figure 4. Spatial distribution of classification errors for Random Forests to predict the presence
762 of eggs from CUFES samples applied to the test data set (years 2002, 2006, 2011, and 2015) for
763 four species using derived predictors from the ROMS-CoSiNE model or remotely sensed
764 satellite data. The upper right map in each panel indicates egg presences, and the lower left
765 indicates negative samples. Purple indicates a correct classifier prediction and red indicates an
766 incorrect prediction. Thus, upper purple = true positive, lower purple = true negative, upper red =
767 false negative, and lower red = false positive.
768

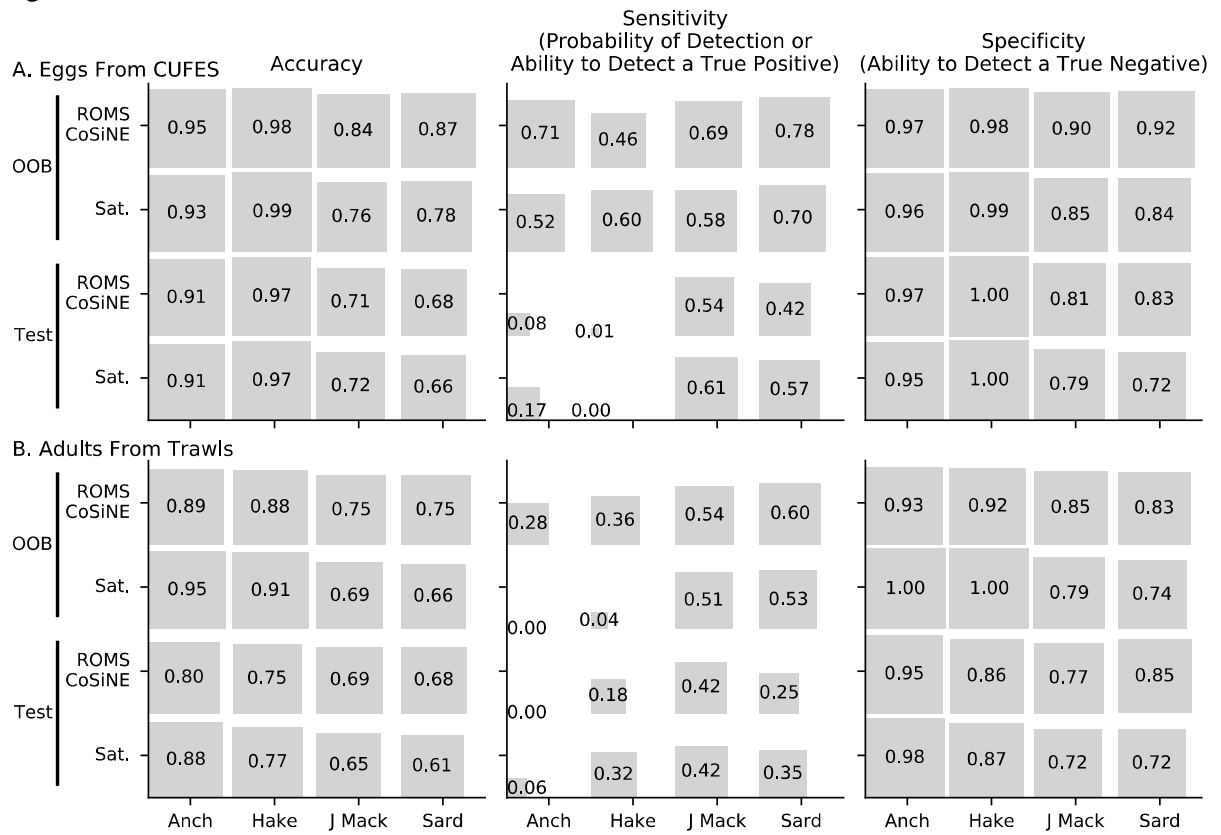
769 Figure 5. Distributions of habitat predictors from the ROMS-CoSiNE model for adults captured
770 in trawls during spring (Spr) and summer (Sum). Half violin plots depict a boxplot in black and
771 kernel density estimate of distribution in gray for each predictor and season.

772 Figure 1.



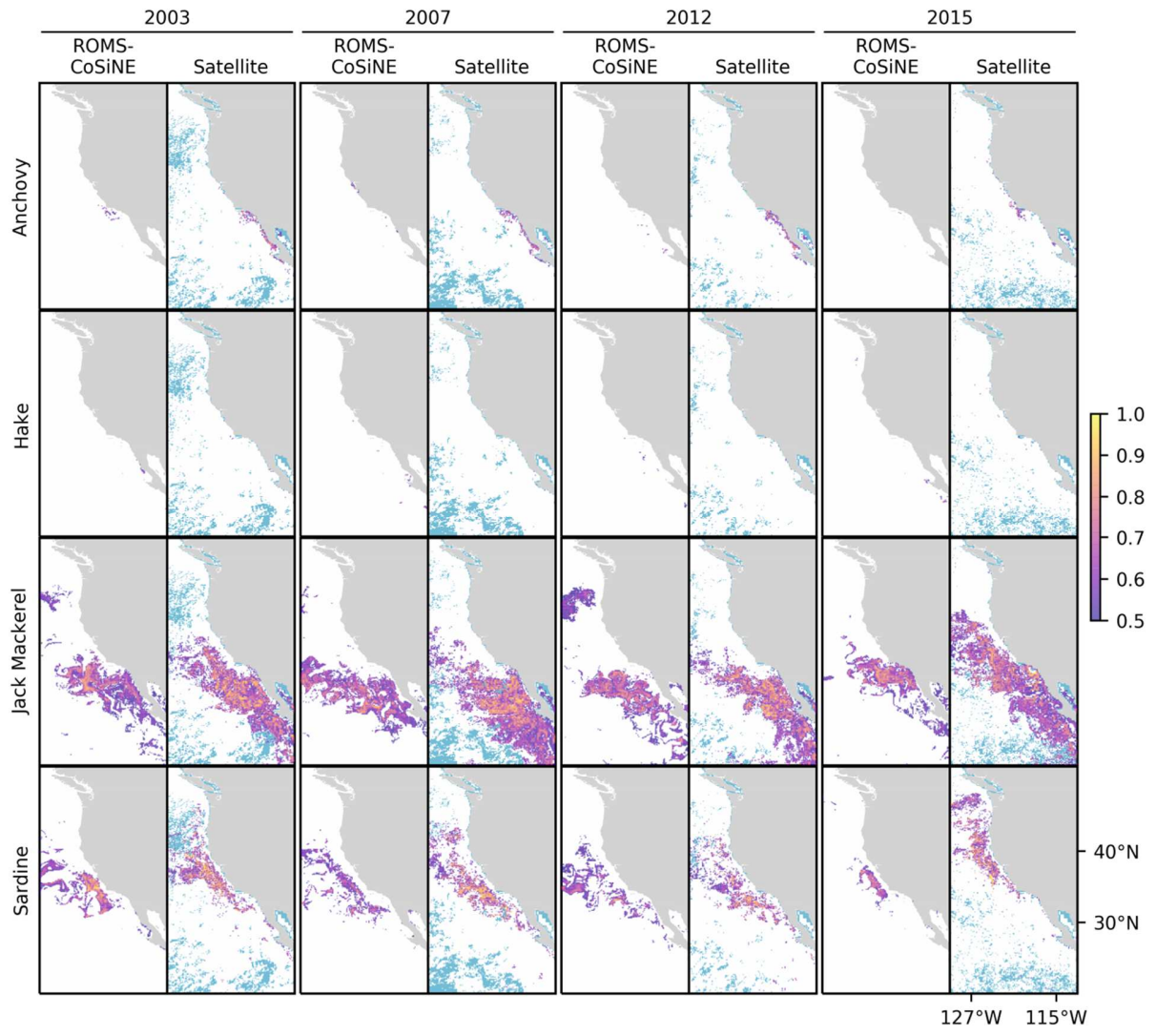
773

774 Figure 2



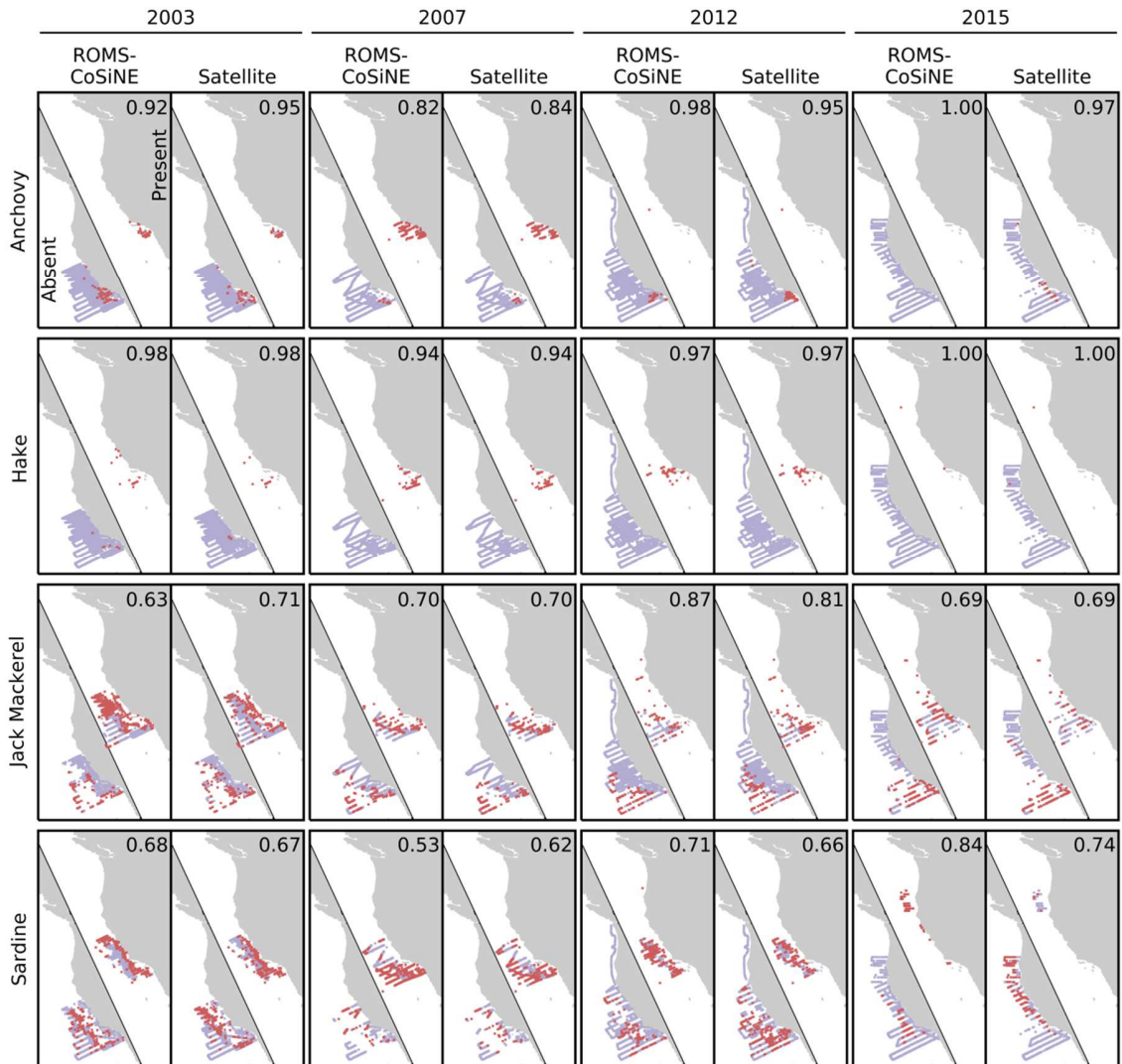
775

776 Figure 3



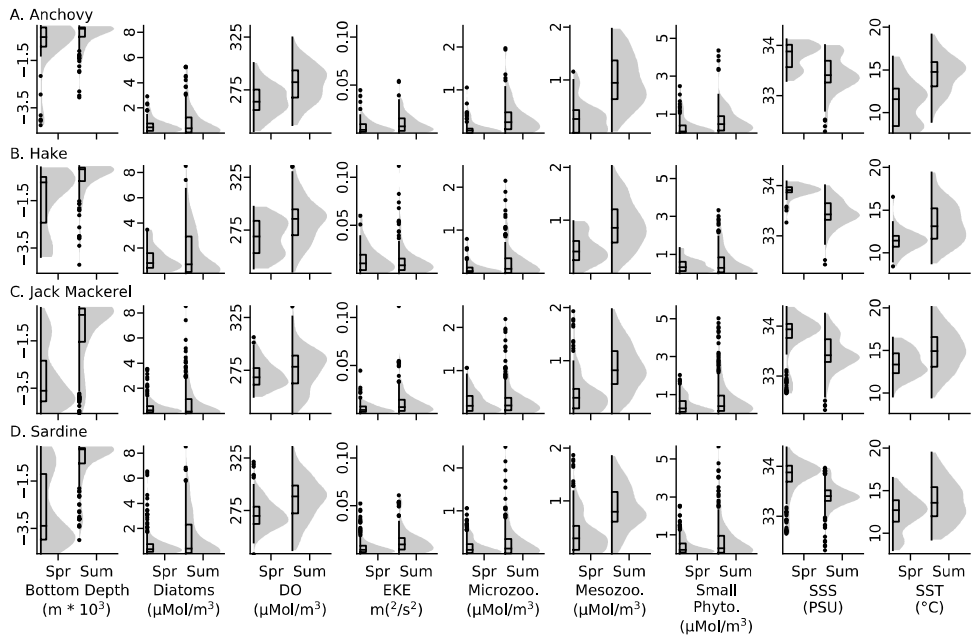
777

778 Figure 4



779

780 Figure 5



781