2 Spatiotemporal modeling of winter sea duck abundance in Nantucket Sound, USA: implica-

3 tions for marine spatial planning

4 ADAM D. SMITH[1,6], BENJAMIN HOFNER[2], JASON E. OSENKOWSKI[3], TABER ALLISON[4],

5 GIANCARLO SADOTI[5], SCOTT R. MCWILLIAMS[1], AND PETER W. C. PATON[1]

6 [1]Department of Natural Resources Science, University of Rhode Island, Kingston, Rhode

7 Island 02881 USA

8 [2]Department of Medical Informatics, Biometry and Epidemiology, Friedrich-Alexander-

9 University Erlangen-Nuremberg, Waldstraße 6, 91054 Erlangen, Germany

10 [3]Rhode Island Department of Environmental Management, 77 Great Neck Road, West

11 Kingston, Rhode Island 02892 USA

12 [4]American Wind Wildlife Institute, 1110 Vermont Avenue, NW, Suite 950, Washington,

13 D.C. 20005 USA

14 [5]Department of Geography, University of Nevada, Reno, Nevada, 89557 USA

[6]Present address: United States Fish and Wildlife Service, National Wildlife Refuge System, Inventory
and Monitoring Branch, 135 Phoenix Rd., Athens, GA 30605 USA; avianmigration@gmail.com

Abstract. Effective marine spatial plans require ecologically meaningful estimates of the distribution and abundance of biological resources that could be affected by anthropogenic change. In North America, there is a pressing need to develop spatially-explicit abundance estimates of marine birds to help assess potential impacts of planned offshore wind energy developments (OWED). Sea ducks are especially important to consider in the context of marine spatial planning because populations of most North American species are below historic levels, and studies conducted in western Europe suggest that OWED could negatively impact sea duck populations. Developers have proposed constructing 130 3.6-MW turbines in Nantucket Sound, Massachusetts where some of the largest concentrations of sea ducks occur during winter in eastern North America. We conducted aerial surveys from 2003 to 2005 to assess spatiotemporal variation in the abundance of sea ducks (Common Eider [*Somateria mollissima*], Black Scoter [*Melanitta americana*], Surf Scoter [*M. perspicillata*], White-winged Scoter [*M. deglandi*], and Long-tailed Duck [*Clangula hyemalis*]) in Nantucket Sound. We related spatiotemporal variation in sea duck occupancy and conditional abundance to potentially relevant biophysical and spatiotemporal covariates. We applied a negative binomial hurdle model that modelled occupancy using a generalized additive model (GAM), and generalized additive models for location, scale, and shape (GAMLSS) to estimate the conditional species abundance. Spatiotemporal effects (i.e., geographic location and time of season) were the dominant explanatory feature in occupancy and conditional abundance estimates. Biophysical covariates also affected occupancy estimates, with occupancy greatest at intermediate monthly sea surface temperatures and in areas with coarser sediments; the effects of biophysical covariates on conditional abundance were less consistent among species. Resulting fine-scale, spatially-explicit models suggested that a planned OWED in Nantucket Sound could displace some sea ducks from potential foraging habitat; however, the largest concentrations of sea ducks were in areas away from the proposed OWED. Our flexible, model-based approach can be used by local and regional planners to prioritize key habitats and areas used by ma-

2

rine birds and other fauna that should be protected from anthropogenic stressors such as

OWED.

Key words. abundance, distribution, Common Eider, generalized additive models, gradient boosting, marine spatial planning, Long-tailed Duck, Nantucket Sound, scoter, sea ducks, wind energy

Introduction

Marine spatial plans (MSP) can provide a rigorous framework to protect marine ecosystems from a variety of anthropogenic stressors including coastal development (Polasky et al. 2008, Foley et al. 2010). A key element of MSPs is accurate spatially-explicit estimates of the abundance and distribution of biological resources, as well as information on the spatial distribution of potential human uses (Douvere 2008, Punt et al. 2009, Bradbury et al. 2014). Along the northwest Atlantic Coast, there are numerous proposals to construct offshore wind energy developments (OWED; Breton and Moe 2009, Musial and Ram 2010), including the first in U.S. waters, the 5-turbine 30 MW facility in Rhode Island Sound to be operational by fall 2016. Therefore there is a pressing need to develop spatially-explicit models of key biological resources that can assist marine spatial planners in siting of OWEDs while minimizing impacts to marine ecosystem integrity.

Understanding the distribution and abundance of wintering marine birds including sea ducks (Tribe: Mergini) represents a particular challenge to marine spatial planners because the environmental factors associated with their distribution and abundance remain notably understudied (Zipkin et al. 2010, Bowman et al. 2015, Flanders et al. 2015). Sea ducks are important to consider in the context of marine spatial planning because populations of most North American species are below historic levels and there is considerable uncertainty in population trajectories (Bowman et al. 2015). Moreover, evidence from Europe's numerous offshore wind facilities suggest that offshore wind energy developments could have negative consequences for sea duck populations (Larsen and Guillemette 2007, Langston 2013, Bradbury et al. 2014).

Model-based approaches are a useful framework to generate spatially-explicit estimates of animal abundance and changes in animal distribution independent of abundance (Borchers et al. 2002, Certain and Bretagnolle 2008, Nur et al. 2011, Kinlan et al. 2012, Winiarski et al. 2014). However, modelling the spatial ecology of marine organisms such as sea ducks

$_{73}$ presents a variety of analytical challenges. First, the spatiotemporal distribution of ma-

$_{74}$ rine organisms can be highly irregular and surveys often produce zero-inflated (Martin

$_{75}$ et al. 2005, Cunningham and Lindenmayer 2005), overdispersed (Richards 2008) count

$_{76}$ data that may vary with biophysical features in complex, non-linear ways (Austin 2007).

$_{77}$ Second, an adequate modeling approach usually must identify a small suite of important

$_{78}$ covariates among many potentially correlated covariates (i.e., high dimensionality; multi-

$_{79}$ collinearity) while avoiding overfitting models, a process that often is complicated by data

$_{80}$ from repeated surveys that regularly exhibit spatial and temporal autocorrelation (Hoeting

$_{81}$ 2009).

$_{82}$ We demonstrate a model-based approach to predict animal distribution and abundance us-

$_{83}$ ing aerial transect survey data of sea ducks in Nantucket Sound, Massachusetts, USA. We

$_{84}$ implemented generalized additive models (GAMs; Hastie and Tibshirani 1990, Wood 2006)

$_{85}$ and their recent extension (GAMLSS - GAMs for location, scale, and shape; Rigby and

$_{86}$ Stasinopoulos 2005) in a gradient descent boosting framework (Friedman 2001). GAMLSS

$_{87}$ methods extend GAMs to allow all parameters of the conditional response distribution

$_{88}$ to be modeled as a function of relevant covariates (Rigby and Stasinopoulos 2005). This

$_{89}$ approach provides spatially-explicit insights into the covariates associated with sea duck

$_{90}$ abundance and its variability. The boosting algorithm accommodates the inclusion of vari-

$_{91}$ ous effect types (e.g., linear, smooth, spatial, spatiotemporal, and random effects; Hofner

$_{92}$ et al. 2014) and correlated covariates (i.e., parameter estimates are regularized), identifies

$_{93}$ the most relevant subset among potentially many covariates (i.e., variable selection), and

$_{94}$ evaluates competing representations of continuous covariates (e.g., linear vs. nonlinear ef-

$_{95}$ fects of covariates; Kneib et al. 2009, Maloney et al. 2012). Our specific objectives were to

$_{96}$ demonstrate a flexible, model-based approach that (1) provides spatially- and temporally-

$_{97}$ explicit estimates of sea duck abundance and distribution while accommodating many of

$_{98}$ the challenges inherent in animal survey data; and (2) describes the associations between

$_{99}$ biophysical features and sea duck distribution and abundance in one of the most impor-

tant migrating and wintering areas for sea ducks in the western Atlantic.

Methods

Study area and species.—We conducted fieldwork throughout Nantucket Sound in Massachusetts, USA (Figure 1). Our study area encompassed ca. 1,500 km^2, was relatively shallow (generally < 20 m deep), and included some of the most important sea duck wintering habitat in the western Atlantic (White et al. 2009, Silverman et al. 2013). The primary species of sea ducks found in Nantucket Sound were Common Eider (*Somateria mollissima*; hereafter eider), Black Scoter (*M. americana*), Surf Scoter (*M. perspicillata*), and White-winged Scoter (*Melanitta deglandi*), and Long-tailed Duck (*Clangula hyemalis*). Additionally, approximately 62 km^2 of of Horseshoe Shoal in northwestern Nantucket Sound is fully permitted for OWED (Figure 1; Santora et al. 2004), which could affect sea duck use of this important wintering area (Drewitt and Langston 2006, Langston 2013, Winiarski et al. 2014).

Aerial strip transect surveys.—During the winters (late October to mid-April) of 2003-2005, we conducted 30 (2003/2004: 13, 2004/2005: 10, 2005/2006: 7) standardized aerial strip-transect surveys (Flanders et al. 2015) throughout an 1,100 km^2 study area in Nantucket Sound (Figure 1). Surveys occurred primarily from November - March (n = 27), with occasional October (n = 1) or April (n = 2) surveys. During each survey, we flew along 15 parallel (ca. 2.5 km apart), roughly north-south transects (Figure 1) using a high-wing, twin-engine aircraft (Cessna Skymaster 337) at an average altitude of 152 m and speed of 167 km/h (90 kts). This altitude allowed us to identify most birds at the sea surface and reduced sea disturbance (i.e., flushing birds to another part of the study area and, thus, potential double counting). The airspeed was the slowest at which the aircraft could safely fly. Surveys occurred only on days with wind speeds ≤ 15 kts and good visibility (>15 km). Surveys occurred between 0900 to 1600 h (average duration of ca. 2.5 h) to

ensure that birds had completed their post-dawn movements (e.g., Davis , W.E. 1997) but had yet to initiate pre-sunset movements from feeding to roosting areas; this time window also reduced glare due to low sun angles.

On each survey flight, two observers used their unaided eyes to continuously detect individuals or flocks, identifying sea ducks to species with the aid of binoculars as needed. Observers monitored the sea surface on their side of the plane in a ca. 91 m wide transect between ca. 56 - 147 m from the plane; observers could not see beneath the plane out to 56 m on each side. The narrow strip width ensured birds were detectable and identifiable with the naked eye and allowed us to avoid situations in which ducks were too abundant or spread over too wide an area to count accurately. This last point was especially relevant because sea ducks occasionally congregated in extremely large aggregations. We thus assumed inconsequential perception bias (i.e., few missed individuals present to be counted; Marsh and Sinclair 1989) given that our flight altitude and narrow transect width (e.g., Certain and Bretagnolle 2008) within the strip transect (Buckland et al. 2012). Our transect dimensions resulted in the sampling of approximately 6% (median; 68.4 km^2) of the study area during a survey.

Observers verbally communicated duck sightings to a recorder, who entered the number of ducks, species, behavior (i.e., on the water or flying), and geographic location into a laptop computer linked to the plane's onboard GPS; we report results only from ducks observed sitting on the water. We subsequently consolidated counts for each species (eider and Long-tailed Duck) or species group (scoter) into 2.25km^2 segments (Figure 1); this resolution (1.5 km x 1.5 km) corresponded approximately to the coarsest level of resolution of biophysical covariates (see below).

Modeling approach.—We related spatiotemporal variation in sea duck occupancy (i.e., probability of presence) and abundance to potentially relevant biophysical and spatiotemporal covariates. We applied a negative binomial hurdle model that first modeled (1) the

7

probability of occurrence of at least one individual (hereafter, the occupancy model) observed on the 180 m wide transect in a given segment using a logistic regression model and then (2) the abundance of sea ducks observed on the transect in that segment conditional on their presence (hereafter, the count model) using a truncated negative binomial model. The prevalence of zero counts (e.g., 75% of eider segment observations) prompted the use of a hurdle model. GAMs flexibly accommodated potential nonlinear effects of covariates on sea duck occupancy; GAMLSS accommodated potential nonlinear effects of covariates on the conditional mean and overdispersion of sea duck abundance in the count models. We generated separate hurdle models for each sea duck group (i.e., eider, scoter, and Long-tailed Duck; Zipkin et al. 2010).

We implemented GAMs and GAMLSS in a gradient descent boosting framework. More specifically, we used component-wise functional gradient descent boosting to fit the models. We first computed the negative gradient of a user-specified loss function (typically the negative log-likelihood of the underlying model); the negative gradient of the loss function can be viewed as working residuals. We used the (negative) binomial log-likelihood as the loss function for occupancy models and the (negative) truncated negative binomial log-likelihood as the loss function for count models.

In the next step, user-specified functional forms of the covariates relative to the response, called base-learners (Hofner et al. 2014), were fitted separately to the negative gradient and a fraction of the single best fitting base-learner was added to the current model fit. The negative gradient was then reevaluated at the current model fit and the procedure iterated for a fixed number of iterations, m_{stop}, specified by the user (see Bühlmann and Hothorn 2007 for additional details). We fit boosted GAMLSS models similarly, although in each iteration the negative gradient was computed separately with respect to each GAMLSS parameter (i.e., mean and overdispersion) while holding the other parameter as a fixed constant. We then fit base-learners to the resulting parameter-specific negative

177 gradient vector and model updates were computed separately for each GAMLSS parame-
178 ter as for boosted GAMs (see Mayr et al. 2012 for additional details).

179 As only the single best-fitting base-learner was selected in each iteration, the algorithm
180 thus integrated intrinsic selection of the most relevant covariates and their functional form
181 (i.e., some base-learners may never be selected). Variable selection was further fostered
182 by stopping the algorithm prior to convergence to maximum likelihood estimates (early
183 stopping; Mayr et al. 2012, Maloney et al. 2012). Furthermore, early stopping aims at
184 maximizing predictive accuracy while avoiding model overfitting. We used 25-fold sub-
185 sampling to determine the optimal stopping iteration for each model. Specifically, we ran-
186 domly drew (without replacement) 25 samples of size n/2 from the original data set. We
187 used the selected sample to estimate the model and the balance of the data in each sample
188 to determine the out-of-bag prediction accuracy (empirical risk) measured by the negative
189 log-likelihood of each model; the optimal stopping iteration ($\hat{m}_{\mathrm{stop}}$) is the iteration with
190 the lowest average empirical risk. In boosted GAMLSS models we used multi-dimensional
191 subsampling to determine the stopping iteration for each of the GAMLSS parameters
192 while allowing for potentially different model complexities in the parameters; a detailed
193 explanation of this cross-validation (subsampling) scheme is given in Hofner et al. (2015b).

194 Despite these agreeable features, boosting methods typically produce "rich" models relying
195 to some extent on many base-learners (Hofner et al. 2015a). Thus, we additionally applied
196 stability selection (Meinshausen and Bühlmann 2010, Shah and Samworth 2013) to iden-
197 tify those base-learners, and thus covariates, most commonly selected in each model while
198 preserving an upper bound of $\alpha \approx 0.06$ for the per-comparison error rate (see Appendix S1
199 and Hofner et al. 2015a for more details).

200 *Covariates.*—We evaluated biophysical covariates expected to influence the distribution,
201 abundance, and movements of sea ducks or, more likely, the distribution and availability
202 of their benthic prey (e.g., mollusks and crustaceans; see Appendix S2); we did not have

9

information related directly to the distribution of preferred prey. Biophysical covariates could be characterized as strictly spatial or temporal effects (i.e., varying only among segments or within or among winters, respectively) or as spatiotemporal effects (i.e., varying among segments and within or among winters; see Appendix S2). Additionally, we included interactions that allowed the effects of two covariates to vary over time within a given winter (see Appendix S2). We standardized (i.e., mean centered and scaled) all continuous covariates. Appendix S3 describes a custom function to visualize the spatial and temporal distribution of covariates in Nantucket Sound.

For each continuous covariate, we specified two base-learners: a linear base-learner and a base-learner for the smooth deviation from the linear effect via penalized splines (i.e., P-splines; Eilers and Marx 1996, Fahrmeir et al. 2004, Schmid and Hothorn 2008). This allows the choice of the modeling alternative for each covariate, i.e., between no effect, linear effect, and smooth effect. We dummy-coded categorical covariates and used a separate linear base-learner for each of the K - 1 linear base-learners for a K-level covariate. To address potential spatial autocorrelation, we included a smooth surface function of the spatial coordinates of segment centers (Kneib et al. 2008); this surface comprised four base-learners — linear base-learners for the easting and northing, their linear interaction, and a penalized nonlinear tensor product P-spline (Kneib et al. 2008, 2009, Maloney et al. 2012). We also allowed this surface to vary over time within a winter (via an interaction). Hofner et al. (2014) describe base-learners in detail.

The decomposition of continuous covariates into centered linear and penalized orthogonal nonlinear base-learners allowed us to weight base-learners equally and thus allowed unbiased model choice (i.e., prevented the preferential selection of smooth base-learners; Kneib et al. 2009, Hofner et al. 2011). Specifically, we restricted each base-learner to a single degree of freedom and omitted the intercept term from each base-learner (Kneib et al. 2009, Hofner et al. 2011). Consequently, we added an explicit linear base-learner for the model

10

intercept. The occupancy and count models had the following structure (see Appendix S2 for descriptions of abbreviated covariates):

$$
\begin{aligned}
g(\cdot) =& int + time + f(time) + SSTw + f(SSTw) + SSTm + f(SSTm) + SSTrel+ \\
& f(SSTrel) + SSTrel \cdot time + f(SSTrel, time) + SBT + f(SBT) + NAOw+ \\
& depth + f(depth) + depth \cdot time + f(depth, time) + d2land + f(d2land)+ \\
& chla + f(chla) + cdom + f(cdom) + f(cdom, chla) + meanphi + f(meanphi)+ \\
& SAR + f(SAR) + tidebmean + f(tidebmean) + tidesd + f(tidesd) + strat+ \\
& f(strat) + ferry + y2004 + y2005 + xkm + ykm + xkm \cdot ykm+ \\
& f(xkm, ykm) + xkm \cdot time + ykm \cdot time + xkm \cdot ykm \cdot time+ \\
& f(xkm, ykm) \cdot time + obs\_window + f(obs\_window).
\end{aligned}
\tag{1}
$$

In our occupancy models, $g(\cdot)$ is $g(\pi_{\text{sea ducks}})$, the occupancy probability of a given duck species in a segment and g is the logit link. In our count models, $g(\cdot)$ took two forms within the GAMLSS framework – the (conditional) mean count of sea ducks, $g(\mu_{\text{sea ducks}})$, and the (conditional) overdispersion in sea duck counts, $g(\sigma_{\text{sea ducks}})$; g is the log link in both cases. Base-learners denoted as $f(\cdot)$ indicate the penalized nonlinear deviations from the corresponding linear base-learner (e.g., $f(time)$). The explicit intercept (int) was a necessary byproduct of our decomposition of base-learners (see above and Kneib et al. 2009, Hofner et al. 2011). Small values of obs_window, our measure of survey effort (see Appendix S2), in some segments impaired the estimability of models including obs_window as an offset; we thus included it as a covariate.

Subsequent to their independent fitting, we consolidated occupancy and conditional count models (see, e.g., Equation 6 in Zeileis et al. 2008) to generate spatially-explicit estimates of unconditional sea duck abundance. These estimates were used to evaluate the approximate explanatory power of our final models using a pseudo R^2 measure of the explained

11

variation (Nagelkerke 1991, Maloney et al. 2012).

All analyses were conducted in R (Version 3.1.3; R Core Team 2014) with the add-on packages gamboostLSS (Hofner et al. 2015b, 2015c), mboost (Hothorn et al. 2010, 2015), and stabs (Hofner and Hothorn 2015). The data and code for reproducing this manuscript and analyses are given as an online electronic supplement at http://github.com/adamdsmith/NanSound_EcolApp.

Results

We fit two models per each species because sea duck occupancy and conditional count were modelled independently. Bootstrapped empirical risk suggested that occupancy models for all species converged to the maximum likelihood estimates (i.e., occupancy models failed to stop early; see Appendix S4). Conversely, bootstrapped empirical risks prescribed early stopping for both the conditional mean and overdispersion parameter in all count models (see Appendix S4). Final occupancy models and models for the conditional mean of the count part included only a subset (12% to 38%) of the 48 base-learners initially specified for selection (see equation 1). Occupancy models generally contained more covariates and their interactions (8-10 of 23) than did count models (3-6 of 23), particularly among stably selected covariates and their interactions (Figure 2, see also Appendices S5 - S7).

Sea duck occupancy.—The covariates associated with occupancy were relatively consistent among sea duck species (Figure 2). The influence of univariate effects on the response is reflected in the range of the effect over the Y-axis and, due to standardization, can be compared among species and covariates within a model. For example, monthly sea surface temperature (*SSTm*) associated more strongly with eider occupancy than did distance to land (*d2land*) because it spanned a larger range of the Y-axis (Figure 2). In contrast, monthly sea surface temperature (*SSTm*) associated much more strongly with occupancy

12

₂₇₀ of Long-tailed Duck than eider and scoter for the same reason (Figure 2). Covariate inter-
₂₇₁ actions, illustrated with bivariate plots, are similarly comparable within a model. Only the
₂₇₂ general association (i.e., positive or negative) with the additive predictor is given for factor
₂₇₃ variables. Comparing univariate, bivariate, and categorical effects is accomplished using
₂₇₄ the detailed covariate plots for eider, scoters, and Long-tailed Duck (Appendices S5 - S7,
₂₇₅ respectively).

₂₇₆ Spatiotemporal effects (i.e., occupancy associated with the xkm-ykm location of segments
₂₇₇ and the change over time within winter [*time*]) were the dominant explanatory feature in
₂₇₈ occupancy models, although these patterns varied considerably among species (Figure 2;
₂₇₉ see Day of season, Northing x Easting). Occupancy increased, but at a decreasing rate,
₂₈₀ with survey effort (*obs_window*) in a given segment (Figure 2). Occupancy estimates in-
₂₈₁ creased at intermediate monthly sea surface temperature (*SSTm*), greater distances from
₂₈₂ land (*d2land*), and in areas with coarser sediments (i.e., smaller meanphi). Eider occu-
₂₈₃ pancy was associated negatively with chromomorphic dissolved organic material (*cdom*)
₂₈₄ and positively with sea floor surface area relative to planimetric area (*SAR*; our measure
₂₈₅ of the topographic variability of the sea floor; Figure 2), whereas scoter occupancy likewise
₂₈₆ related to *SAR* and *cdom*, but in the opposite direction in both cases (Figure 2). Scoter
₂₈₇ occupancy was modestly greater in deeper waters (*depth*), whereas Long-tailed Duck occu-
₂₈₈ pancy was greatest in shallow waters early in the winter but in deeper waters later in the
₂₈₉ winter (Figure 2; *depth* x *time* covariate). Other effects were relatively minor and inconsis-
₂₉₀ tent among species.

₂₉₁ The strong association with occupancy of segment *xkm-ykm* location resulted in distinct
₂₉₂ spatial patterns of occupancy among species (Figure 3, top row) despite the relative sim-
₂₉₃ ilarity of occupancy associations with biophysical covariates (Figure 2). Occupancy was
₂₉₄ typically highest for eider in northwest and southwest Nantucket Sound, in interior Nan-
₂₉₅ tucket Sound for scoter, and in northeast and south Nantucket Sound for Long-tailed Duck

(Figure 3, top row). All species tended to avoid the western edge of the Sound northeast of Martha's Vineyard. Generally, the areas of highest occupancy exhibited the lowest relative variability (Figure 3, bottom row), defined as the median absolute deviation (MAD) of occupancy relative to median occupancy within a segment (a measure analogous to the coefficient of variation).

Sea duck conditional abundance and overdispersion.—Spatial effects (*xkm-ykm*) were the dominant explanatory feature of conditional abundance estimates for scoters and Long-tailed Duck, but they were not selected in the eider model (Figure 2). In contrast with the corresponding occupancy model, scoter conditional abundance decreased with increasing sediment grain size (*meanphi*). Additionally, the relationships between eider conditional abundance and dissolved organic material (*cdom*) and sea floor topography (*SAR*; Figure 2) were more complex than with eider occupancy. The conditional abundance of eider and scoter was also associated with relatively warm or cool sea surface temperatures (SST_{rel}; Figure 2). Biophysical covariates associated with Long-tailed Duck conditional abundance exhibited general agreement with their counterpart in the occupancy models.

Spatially-explicit patterns of median conditional abundance (Figure 4, top row) did not necessarily reflect patterns of occupancy (cf. Figure 3, top row). Some areas of Nantucket Sound exhibited mutually high conditional abundance and occupancy for a given species (e.g., eider in the southwest, scoter in the interior, and Long-tailed Duck in parts of the northeast). However, conditional abundance was low despite relatively high occupancy in some instances (e.g., eider in the northeast and Horseshoe Shoal, scoter in the northeast and southeast, and Long-tailed Duck along the northern margin). Conversely, other areas of Nantucket Sound exhibited lower occupancy but sea ducks, when present, were more abundant (e.g., eider along the eastern margin, and scoter and Long-tailed Duck in the southwest). As in occupancy models, sea ducks were relatively absent from the middle-western margin of Nantucket Sound (i.e., northeast of Martha's Vineyard; see Figure 1).

14

In contrast to sea duck occupancy, however, areas of highest conditional sea duck abundance typically exhibited the highest relative variability over time (Figure 4, bottom row).

Overdispersion in conditional sea duck abundance also varied with biophysical covariates, although there was less consistency in the associated covariates among species (Figure 2; see also Appendices S5 - S7). Variability (i.e., overdispersion) in sea duck counts was heterogeneous in space (Appendix S8; Figure S8.1, top row) and time (Appendix S8; Figure S8.1, bottom row) in Nantucket Sound, particularly for eider and scoter (as indicated by the magnitude of the overdispersion parameter values).

Expected sea duck abundance.—Consolidated occupancy and conditional count models provided estimates of unconditional sea duck abundance in the study area over the survey period (Figure 5, top row). Final models of expected sea duck abundance explained moderate amounts of variation in observed counts of eider, scoter, and Long-tailed Duck (pseudo $R^2 = 0.31$, 0.48, and 0.32, respectively). Conditional abundance (Figure 4) strongly influenced the spatially-explicit patterns of expected abundance. Sea duck species exhibited relatively distinct patterns of abundance in Nantucket Sound. Eider were consistently most abundant in southwestern Nantucket Sound and also relatively abundant in the northeastern part of the study area but less consistently so as evidenced by the relatively high MAD/median abundance over time (Figure 5). Scoter were also most abundant, with occasional extremely large flocks, in southwestern Nantucket Sound, although this was also the area of highest relative variation in scoter abundance; relatively high abundances of scoter also occurred in interior Nantucket Sound (Figure 5). Long-tailed Ducks were consistently most abundant in northeastern Nantucket Sound, as well as along its southern margin (Figure 5). No species' highest abundances occurred in the permitted Nantucket Shoal area, although expected eider and scoter abundances were consistently elevated in some parts of the Shoal (west and southeast, respectively; Figure 5).

Summing the spatially explicit estimates of unconditional sea duck abundance (i.e., Fig-

15

ure 5) provides an estimate of total abundance in a 1.5 km x 180 m transect through all segments in the study area. We compared the total count (summed across all segments) of each sea duck species observed in aerial strip transects with the corresponding estimated total abundance in surveyed segments for each of the 30 aerial surveys (Figure 6). Our models tended to overestimate sea duck abundance when the actual numbers of sea ducks were relatively low, although overestimation was typically less than an order of magnitude. Additionally, scoter abundance was occasionally extreme relative to typical counts and somewhat prone to underestimation during these extreme counts. Nonetheless, the general adherence of observed and predicted abundance to a line of unit slope indicated that it may be reasonable to estimate sea duck abundance for the entire study area based on observed sea duck densities in transects (Figure 6).

Temporal dynamics in wintering sea ducks.—The MAD/median estimates (bottom rows of Figures 3-5, Figure S8.1 in Appendix S8) show that our spatially-explicit estimates of occupancy, abundance, and overdispersion invariably change over time, either explicitly via the selection of a within- or among-winter temporal effect (*time* and *y2004/y2005*, respectively) or implicitly via the selection of biophysical covariates that change within or among winters. The temporal dynamics of the wintering sea duck system in Nantucket Sound was one of its most striking attributes, and we illustrate these dynamics with an animation for scoter occupancy and abundance in Appendix S9.

Discussion

Utility of a boosted GAMLSS modeling framework.—We demonstrated a flexible model-based approach to evaluate the environmental associations of sea duck distribution and abundance based on multiyear replicated surveys. This approach is particularly useful for sea ducks that exhibit considerable variation in their spatial distribution and abundance within and between years (Zipkin et al. 2010, Winiarski et al. 2014, Flanders et

16

al. 2015). The boosted GAMLSS framework offered several useful features including (1) the ability to model all parameters of the conditional distribution (e.g., conditional mean and overdispersion) as a function of covariates, (2) integrated variable reduction and selection among many covariates, and (3) integrated model selection via the simultaneous consideration of competing functional covariate forms (e.g., linear vs. non-linear). Additionally, this framework allowed us to incorporate smooth effects to efficiently account for spatiotemporal trends in the data that were poorly explained by other covariates and to identify those covariates and their functional forms most consistently associated with animal distribution and abundance (via stability selection). The useful features of this modeling framework apply especially to mobile species with non-uniform distributions that vary among and within years, such as the species of sea ducks that we studied. Our estimates of the spatiotemporal abundance of sea ducks in Nantucket Sound were controlled largely by estimates of the conditional abundance and less by spatiotemporal patterns in the occupancy of sea ducks. This suggests that occupancy models alone may be inadequate for assessing risk from anthropogenic disturbances and for describing the fine-scale distribution of marine species (e.g., Winiarski et al. 2014, Flanders et al. 2015).

The importance of spatial scale.—The process whereby migratory animals such as sea ducks select a given area to inhabit during winter involves decisions at multiple spatial scales and the environmental attributes that determine this habitat selection often vary with spatial scale (Johnson 1980, Johnson et al. 2004, 2006). The majority of North American sea ducks migrate from high-latitude arctic and sub-arctic breeding areas to mid-latitude temperate wintering areas where they reside for most of the year (Silverman et al. 2013, Bowman et al. 2015, Flanders et al. 2015). At these large spatial scales, the distribution and abundance of sea ducks during winter may be affected by large-scale ocean characteristics (e.g., Flint 2013), climatic conditions (e.g., Zipkin et al. 2010), and static or persistent habitat features (e.g., bathymetry, substrate, current and frontal systems; Hyrenbach et al. 2000, Nur et al. 2011, Flanders et al. 2015). At regional and local scales,

17

however, most species of sea ducks congregate in large flocks (e.g., up to tens of thousands of birds) at sites where prey are abundant and accessible (Flint 2013, Loring et al. 2013) although the abundance and distribution of these prey, and thus predators, can be extremely ephemeral and dynamic (Hyrenbach et al. 2000, Cisneros et al. 2011).

Given that sea duck distribution and abundance is spatially and temporally dynamic, yet expected to be driven by biophysical covariates (Oppel et al. 2009, Zipkin et al. 2010, Flanders et al. 2015) that may differ in importance depending on spatial scale (Johnson 1980, Johnson et al. 2004, 2006), marine spatial planners must carefully consider the most appropriate information to use when deciding, for example, where to place OWEDs. A larger-scale occupancy model developed by Flanders et al. (2015) suggested that eiders were relatively uniformly distributed across Nantucket Sound, whereas our higher resolution abundance models found that eiders were concentrated in southwestern and central, eastern areas within Nantucket Sound. While large-scale models (Silverman et al. 2013, Flanders et al. 2015) are useful to identify general geographic areas of import to sea ducks, our models provide more detailed estimates of sea duck distribution and abundance within Nantucket Sound. In terms of marine spatial planning, the large-scale models can be used to decide which areas should or should not be considered for OWED and, once an area is zoned for OWED, more localized models can be used to decide where best to place the OWED within the zoned area.

Environmental covariates that best explain sea duck distribution and abundance.—The biophysical associations with sea duck occupancy derived from our models were relatively consistent among species, whereas their associations with sea duck conditional abundance were more species-specific. Distance to land, which often is positively associated with bathymetry, often has a strong influence on sea duck occupancy estimates (Guillemette et al. 1993, Lewis et al. 2008, Winiarski et al. 2014, Flanders et al. 2015). Sediment grain size can also have a strong influence on prey availability for foraging sea ducks (Goudie

18

and Ankney 1988, Lovvorn et al. 2009, Loring et al. 2013) and affected occupancy and conditional abundance in this study. In addition, topographic variability of the sea floor also influenced occupancy and conditional abundance, although the influence of topography on prey availability is less understood. Sea surface temperature and chlorophyll *a* can have a strong influence on occupancy estimates for sea ducks (Zipkin et al. 2010, Flanders et al. 2015), although we found no effect of chlorophyll *a* during this study. Similarly, Zipkin et al. (2010) found the North Atlantic Oscillation (NAO) was important in all these species of sea ducks at a continental scale, whereas we found no support that the NAO affected the distribution of sea ducks at the scale of this study. Certain covariates may associate with marine bird abundance or behavior at specific scales and not at others (Logerwell and Hargreaves 1996) and this may explain the apparent discrepancy between studies in the effect of chlorophyll *a* and NAO.

The unexplained variation in our models and the predominance of marginal spatiotemporal effects suggest that we probably missed important variable(s) relevant to the distribution of sea ducks in Nantucket Sound. In addition, we likely need better biophysical proxies for the distribution of prey eaten by sea ducks or concurrent prey distribution information (e.g., Vaitkus and Bubinas 2001, Kaiser et al. 2006, Žydelis et al. 2009, Cervencl and Fernandez 2012, Cervencl et al. 2014), although this is typically considerably more difficult to characterize at appropriate scales and does not guarantee improved predictive accuracy (Grémillet et al. 2008, Torres et al. 2008, Benoit-Bird et al. 2013).

Marine Spatial Planning: where to place a wind farm in Nantucket Sound given these estimates of the distribution and abundance of sea ducks?—In the past decade, ecosystem-based marine spatial plans have become a reality because comprehensive land use planning is a central component of development plans in North America and Europe (Douvere 2008). One of the biggest challenges facing marine spatial planners is the paucity of relevant information on the spatial distribution and abundance of biological resources includ-

19

ing marine birds (Bradbury et al. 2014, Flanders et al. 2015). We developed fine-scale, spatially- and temporally-explicit maps of the estimated distribution and abundance of sea ducks in Nantucket Sound that could assist marine spatial planners during the zoning process. At a continental scale, Nantucket Sound regularly supports one of the largest concentrations of wintering sea ducks in eastern North America (White et al. 2009, Zipkin et al. 2010, Silverman et al. 2013), therefore if any OWEDs are constructed in Nantucket Sound then steps should be taken to minimize impacts to this key wintering habitat. Bradbury et al. (2014) developed a sensitivity index that suggested that sea ducks were particularly vulnerable to habitat displacement from foraging sites, while they are less vulnerable to collision risk from OWED. Our models suggest that the permitted OWED zone on Horse-shoe Shoal is not located in prime foraging habitat for most species of sea ducks in the Sound, although large numbers of eiders can use this area (see also Flanders et al. 2015). Potential vulnerability to OWED has been incorporated into some recent spatially-explicit planning efforts for seabirds (Garthe and Hüppop 2004, Winiarski et al. 2014, Bradbury et al. 2014) though to our knowledge not yet for sea ducks.

Acknowledgments

Literature cited

Austin, M. 2007. Species distribution models and ecological theory: A critical assessment and some possible new approaches. Ecological Modelling 200:1–19.

Benoit-Bird, K. J., B. C. Battaile, S. A. Heppell, B. Hoover, D. Irons, N. Jones, K. J. Kuletz, C. A. Nordstrom, R. Paredes, R. M. Suryan, C. M. Waluk, and A. W. Trites. 2013. Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies. PLoS ONE 8:e53348.

Borchers, D. L., S. T. Buckland, and W. Zucchini. 2002. Estimating animal abundance: Closed populations. Springer, London.

Bowman, T. D., E. D. Silverman, S. G. Gilliland, and J. B. Leirness. 2015. Status and trends of North American sea ducks: Reinforcing the need for better monitoring. Pages 1–27 *in* J.-P. L. Savard, D. V. Derksen, D. Esler, and J. M. Eadie, editors. Ecology and Conservation of North American Sea Ducks. Studies in Avian Biology. CRC Press, New York, NY, USA.

Bradbury, G., M. Trinder, B. Furness, A. N. Banks, R. W. G. Caldow, and D. Hume. 2014. Mapping seabird sensitivity to offshore wind farms. PLoS ONE 9:e106366.

Breton, S.-P., and G. Moe. 2009. Status, plans and technologies for offshore wind turbines in Europe and North America. Renewable Energy 34:646–654.

Buckland, S. T., M. L. Burt, E. A. Rexstad, M. Mellor, A. E. Williams, and R. Woodward. 2012. Aerial surveys of seabirds: The advent of digital methods. Journal of Applied

Ecology 49:960–967.

Bühlmann, P., and T. Hothorn. 2007. Boosting algorithms: Regularization, prediction and model fitting. Statistical Science 22:477–505.

Certain, G., and V. Bretagnolle. 2008. Monitoring seabirds population in marine ecosystem: The use of strip-transect aerial surveys. Remote Sensing of Environment 112:3314–3322.

Cervencl, A., and S. A. Fernandez. 2012. Winter distribution of Greater Scaup *Aythya marila* in relation to available food resources. Journal of Sea Research 73:41–48.

Cervencl, A., K. Troost, E. Dijkman, M. de Jong, C. J. Smit, M. F. Leopold, and B. J. Ens. 2014. Distribution of wintering Common Eider *Somateria mollissima* in the Dutch Wadden Sea in relation to available food stocks. Marine Biology 162:153–168.

Cisneros, K. O., A. J. Smit, J. Laudien, and D. S. Schoeman. 2011. Complex, dynamic combination of physical, chemical and nutritional variables controls spatio-temporal

510   variation of sandy beach community structure. PLoS ONE 6:e23724.

511   Cunningham, R. B., and D. B. Lindenmayer. 2005. Modeling count data of rare species:

512   Some statistical issues. Ecology 86:1135–1142.

513   Davis , W.E., J. 1997. The Nantucket oldsquaw flight: New England's greatest bird show?

514   Bird Observer 25:16–22.

515   Douvere, F. 2008. The importance of marine spatial planning in advancing ecosystem-

516   based sea use management. Marine Policy 32:762–771.

517   Drewitt, A. L., and R. H. W. Langston. 2006. Assessing the impacts of wind farms on

518   birds. Ibis 148:29–42.

519   Eilers, P. H. C., and B. D. Marx. 1996. Flexible smoothing with B-splines and penalties

520   (with discussion). Statistical Science 11:89–121.

521   Fahrmeir, L., T. Kneib, and S. Lang. 2004. Penalized structured additive regression for

522   space-time data: A Bayesian perspective. Statistica Sinica 14:731–761.

523   Flanders, N. P., B. Gardner, K. J. Winiarski, P. W. C. Paton, T. Allison, and A.

524   O'Connell. 2015. Using a community occupancy model to identify key seabird areas in

525   southern New England. Marine Ecology Progress Series 533:277–290.

526   Flint, P. L. 2013. Changes in size and trends of North American sea duck populations

527   associated with North Pacific oceanic regime shifts. Marine Biology 160:59–65.

528   Foley, M. M., B. S. Halpern, F. Micheli, M. H. Armsby, M. R. Caldwell, C. M. Crain,

529   E. Prahler, N. Rohr, D. Sivas, M. W. Beck, M. H. Carr, L. B. Crowder, J. Emmett

530   Duffy, S. D. Hacker, K. L. McLeod, S. R. Palumbi, C. H. Peterson, H. M. Regan, M. H.

531   Ruckelshaus, P. A. Sandifer, and R. S. Steneck. 2010. Guiding ecological principles for

marine spatial planning. Marine Policy 34:955–966.

Friedman, J. H. 2001. Greedy function approximation: A gradient boosting machine. The
Annals of Statistics 29:1189–1232.

Garthe, S., and O. Hüppop. 2004. Scaling possible adverse effects of marine wind farms
on seabirds: Developing and applying a vulnerability index. Journal of Applied Ecology
41:724–734.

Goudie, R. I., and C. D. Ankney. 1988. Patterns of habitat use by sea ducks wintering in
southeastern Newfoundland. Ornis Scandinavica (Scandinavian Journal of Ornithology)
19:249–256.

Grémillet, D., S. Lewis, L. Drapeau, C. D. Van Der Lingen, J. A. Huggett, J. C. Coetzee,
H. M. Verheye, F. Daunt, S. Wanless, and P. G. Ryan. 2008. Spatial match–mismatch
in the Benguela upwelling zone: Should we expect chlorophyll and sea-surface tempera-
ture to predict marine predator distributions? Journal of Applied Ecology 45:610–621.

Guillemette, M., J. H. Himmelman, C. Barette, and A. Reed. 1993. Habitat selection
by common eiders in winter and its interaction with flock size. Canadian Journal of
Zoology 71:1259–1266.

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. CRC Press, Boca
Raton, FL, USA.

Hoeting, J. A. 2009. The importance of accounting for spatial and temporal correlation in
analyses of ecological data. Ecological Applications 19:574–577.

Hofner, B., and T. Hothorn. 2015. Stabs: Stability selection with error control, R package
version 0.5-1.

Hofner, B., L. Boccuto, and M. Göker. 2015a. Controlling false discoveries in high-
dimensional situations: Boosting with stability selection. BMC Bioinformatics

16:144.

Hofner, B., T. Hothorn, T. Kneib, and M. Schmid. 2011. A framework for unbiased model selection based on boosting. Journal of Computational and Graphical Statistics 20:956–971.

Hofner, B., A. Mayr, and M. Schmid. 2015b. gamboostLSS: An R package for model building and variable selection in the GAMLSS framework. Accepted for publication in Jorunal of Statistical Software.

Hofner, B., A. Mayr, N. Fenske, and M. Schmid. 2015c. gamboostLSS: Boosting methods for GAMLSS models, R package version 1.1-3.

Hofner, B., A. Mayr, N. Robinzonov, and M. Schmid. 2014. Model-based boosting in R: A hands-on tutorial using the R package mboost. Computational Statistics 29:3–35.

Hothorn, T., P. Bühlmann, T. Kneib, M. Schmid, and B. Hofner. 2010. Model-based boosting 2.0. Journal of Machine Learning Research 11:2109–2113.

Hothorn, T., P. Bühlmann, T. Kneib, M. Schmid, and B. Hofner. 2015. Mboost: Model-based boosting, R package version 2.4-2.

Hyrenbach, K. D., K. A. Forney, and P. K. Dayton. 2000. Marine protected areas and ocean basin management. Aquatic Conservation: Marine and Freshwater Ecosystems 10:437–458.

Johnson, C. J., S. E. Nielsen, E. H. Merrill, T. L. Mcdonald, and M. S. Boyce. 2006. Resource selection functions based on use–availability data: Theoretical motivation and evaluation methods. Journal of Wildlife Management 70:347–357.

Johnson, C. J., D. R. Seip, and M. S. Boyce. 2004. A quantitative approach to conservation planning: Using resource selection functions to map the distribution of mountain

⁵⁷⁹ caribou at multiple spatial scales. Journal of Applied Ecology 41:238–251.

⁵⁸⁰ Johnson, D. H. 1980. The comparison of usage and availability measurements for evaluat-

⁵⁸¹ ing resource preference. Ecology 61:65–71.

⁵⁸² Kaiser, M. J., M. Galanidi, D. A. Showler, A. J. Elliott, R. W. G. Caldow, E. I. S. Rees,

⁵⁸³ R. A. Stillman, and W. J. Sutherland. 2006. Distribution and behaviour of Common

⁵⁸⁴ Scoter *Melanitta nigra* relative to prey resources and environmental parameters. Ibis

⁵⁸⁵ 148:110–128.

⁵⁸⁶ Kinlan, B., C. Menza, and F. Huettmann. 2012. Chapter 6: Predictive modeling of

⁵⁸⁷ seabird distribution patterns in the New York Bight. Pages 87–224 *in* C. Menza, B.

⁵⁸⁸ Kinlan, D. Dorfman, M. Poti, and C. Caldow, editors. A biogeographic assessment of

⁵⁸⁹ seabirds, deep sea corals and ocean habitats of the New York Bight: Science to support

⁵⁹⁰ offshore spatial planning. NOAA Technical Memorandum NOS NCCOS 141. NOAA,

⁵⁹¹ Silver Spring, MD.

⁵⁹² Kneib, T., T. Hothorn, and G. Tutz. 2009. Variable selection and model choice in geoaddi-

⁵⁹³ tive regression models. Biometrics 65:626–634.

⁵⁹⁴ Kneib, T., J. Müller, and T. Hothorn. 2008. Spatial smoothing techniques for the assess-

⁵⁹⁵ ment of habitat suitability. Environmental and Ecological Statistics 15:343–364.

⁵⁹⁶ Langston, R. H. W. 2013. Birds and wind projects across the pond: A UK perspective.

⁵⁹⁷ Wildlife Society Bulletin 37:5–18.

⁵⁹⁸ Larsen, J. K., and M. Guillemette. 2007. Effects of wind turbines on flight behaviour of

⁵⁹⁹ wintering common eiders: Implications for habitat use and collision risk. Journal of

⁶⁰⁰ Applied Ecology 44:516–522.

⁶⁰¹ Lewis, T. L., D. Esler, and W. S. Boyd. 2008. Foraging Behavior of Surf Scoters

⁶⁰² (*Melanitta perspicillata*) and White-Winged Scoters (*M. fusca*) in relation to clam

density: Inferring food availability and habitat quality. The Auk 125:149–157.

Logerwell, E. A., and N. B. Hargreaves. 1996. The distribution of sea birds relative to their fish prey off Vancouver Island: Opposing results at large and small spatial scales. Fisheries Oceanography 5:163–175.

Loring, P. H., P. W. C. Paton, S. R. McWilliams, R. A. McKinney, and C. A. Oviatt. 2013. Densities of wintering scoters in relation to benthic prey assemblages in a North Atlantic estuary. Waterbirds 36:144–155.

Lovvorn, J. R., J. M. Grebmeier, L. W. Cooper, J. K. Bump, and S. E. Richman. 2009. Modeling marine protected areas for threatened eiders in a climatically changing Bering Sea. Ecological Applications 19:1596–1613.

Maloney, K. O., M. Schmid, and D. E. Weller. 2012. Applying additive modelling and gradient boosting to assess the effects of watershed and reach characteristics on riverine assemblages. Methods in Ecology and Evolution 3:116–128.

Marsh, H., and D. F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. The Journal of Wildlife Management 53:1017–1024.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: Improving ecological inference by modelling the source of zero observations. Ecology Letters 8:1235–1246.

Mayr, A., N. Fenske, B. Hofner, T. Kneib, and M. Schmid. 2012. Generalized additive models for location, scale and shape for high-dimensional data - a flexible approach based on boosting. Journal of the Royal Statistical Society: Series C (Applied Statis-

27

624   tics) 61:403–427.

625   Meinshausen, N., and P. Bühlmann. 2010. Stability selection (with discussion). Journal of

626   the Royal Statistical Society: Series B (Statistical Methodology) 72:417–473.

627   Musial, W., and B. Ram. 2010. Large-scale offshore wind power in the United States:

628   Assessment of opportunities and barriers. National Renewable Energy Laboratory,

629   Golden, CO, USA.

630   Nagelkerke, N. J. D. 1991. A note on a general definition of the coefficient of determina-

631   tion. Biometrika 78:691–692.

632   Nur, N., J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley,

633   J. A. Wiens, K. Morgan, L. T. Ballance, and D. Stralberg. 2011. Where the wild things

634   are: Predicting hotspots of seabird aggregations in the California Current System. Eco-

635   logical Applications 21:2241–2257.

636   Oppel, S., A. N. Powell, and D. L. Dickson. 2009. Using an algorithmic model to reveal

637   individually variable movement decisions in a wintering sea duck. Journal of Animal

638   Ecology 78:524–531.

639   Polasky, S., E. Nelson, J. Camm, B. Csuti, P. Fackler, E. Lonsdorf, C. Montgomery, D.

640   White, J. Arthur, B. Garber-Yonts, R. Haight, J. Kagan, A. Starfield, and C. Tobalske.

641   2008. Where to put things? Spatial land management to sustain biodiversity and eco-

642   nomic returns. Biological Conservation 141:1505–1524.

643   Punt, M. J., R. A. Groeneveld, E. C. van Ierland, and J. H. Stel. 2009. Spatial planning

644   of offshore wind farms: A windfall to marine environmental protection? Ecological

645   Economics 69:93–103.

646 R Core Team. 2014. R: A language and environment for statistical computing. R Founda-

647   tion for Statistical Computing, Vienna, Austria.

648 Richards, S. A. 2008. Dealing with overdispersed count data in applied ecology. Journal of

649   Applied Ecology 45:218–227.

650 Rigby, R. A., and D. M. Stasinopoulos. 2005. Generalized additive models for location,

651   scale and shape (with discussion). Journal of the Royal Statistical Society: Series C

652   (Applied Statistics) 54:507–554.

653 Santora, C., N. Hade, and J. Odell. 2004. Managing offshore wind developments in the

654   United States: Legal, environmental and social considerations using a case study in

655   Nantucket Sound. Ocean & Coastal Management 47:141–164.

656 Schmid, M., and T. Hothorn. 2008. Boosting additive models using component-wise P-

657   splines. Computational Statistics & Data Analysis 53:298–311.

658 Shah, R. D., and R. J. Samworth. 2013. Variable selection with error control: Another

659   look at stability selection. Journal of the Royal Statistical Society: Series B (Statistical

660   Methodology) 75:55–80.

661 Silverman, E. D., D. T. Saalfeld, J. B. Leirness, and M. D. Koneff. 2013. Wintering sea

662   duck distribution along the Atlantic Coast of the United States. Journal of Fish and

663   Wildlife Management 4:178–198.

664 Torres, L. G., A. J. Read, and P. Halpin. 2008. Fine-scale habitat modeling of a top

665   marine predator: Do prey data improve predictive capacity. Ecological Applications

666   18:1702–1717.

667 Vaitkus, G., and A. Bubinas. 2001. Modelling of sea duck spatial distribution in relation

668   to food resources in Lithuanian offshore waters under the gradient of winter climatic

669 conditions. Acta Zoologica Lituanica 11:288–302.

670 White, T. P., R. R. Veit, and M. C. Perry. 2009. Feeding ecology of Long-Tailed Ducks
671 *Clangula hyemalis* wintering on the Nantucket Shoals. Waterbirds 32:293–299.

672 Winiarski, K. J., D. L. Miller, P. W. Paton, and S. R. McWilliams. 2014. A spatial conser-
673 vation prioritization approach for protecting marine birds given proposed offshore wind
674 energy development. Biological Conservation 169:79–88.

675 Wood, S. N. 2006. Generalized additive models: An introduction with R. Chapman and
676 Hall/CRC, Boca Raton, FL, USA.

677 Zeileis, A., C. Kleiber, and S. Jackman. 2008. Regression models for count data in R.
678 Journal of Statistical Software 27:1–25.

679 Zipkin, E. F., B. Gardner, A. T. Gilbert, A. F. O. Jr, J. A. Royle, and E. D. Silverman.
680 2010. Distribution patterns of wintering sea ducks in relation to the North Atlantic
681 Oscillation and local environmental characteristics. Oecologia 163:893–902.

682 Žydelis, R., D. Esler, M. Kirk, and W. Sean Boyd. 2009. Effects of off-bottom shellfish
683 aquaculture on winter habitat use by molluscivorous sea ducks. Aquatic Conservation:
684 Marine and Freshwater Ecosystems 19:34–42.

Figure Legends

Figure 1. Actual aerial strip transect tracks (gray lines) during winter (October - April, 2003 - 2005) sea duck surveys (n = 30) in Nantucket Sound, Massachusetts, US. The grid indicates the extent of the 1100 km² study area and its division into 504 2.25km² segments. The polygon in northwest Nantucket Sound indicates the 62 km² area of permitted wind energy development on Horseshoe Shoal.

Figure 2. Marginal functional plots for stably selected covariates in the occupancy (probability of presence) and conditional abundance (mean and overdispersion of count models) of Common Eider (COEI), scoter (SCOT), and Long-tailed Duck (LTDU) in Nantucket Sound during three winters, 2003 - 2005. Plots illustrate the partial contribution to the additive predictor (Y-axis) of a covariate holding all other covariates at their mean. Within a model, univariate plots (i.e., lines) share a Y-axis scale, enabling direct comparisons of effect sizes among covariates and species. For bivariate plots, the Y-axis and X-axis reflect the first and second variables listed in the interaction, respectively; colors indicate the direction and magnitude of the partial contribution (blacks = negative, reds = positive; darker colors = larger effect) and are likewise comparable within a model. Northing by easting effects are given only at 31 December. For factor variables, only the general association (i.e., positive or negative) with the additive predictor is given. Covariate abbreviations correspond to Equation 1.

Figure 3. Occupancy probability for Common Eider (COEI), scoter (SCOT), and Long-tailed Duck (LTDU) in Nantucket Sound during three winters, 2003 - 2005. Occupancy probabilities (top row) represent the median expected probability of sea duck presence in a 1.5 km x ca. 180 m transect through a given segment predicted on 10 evenly-spaced dates from 15 November through 1 April in each winter. Spatiotemporal variation in occupancy (%; bottom row) is indicated by the median absolute deviation, MAD, of occupancy probability relative to the median. Predicted values are categorized based on their quartiles;

segments with the highest occupancy or variability (values ≥ 98th percentile) are outlined in black.

Figure 4. Conditional abundance of Common Eider (COEI), scoter (SCOT), and Long-tailed Duck (LTDU) in Nantucket Sound during three winters, 2003 - 2005. Conditional abundances (top row) represent the median expected number of sea ducks, assuming their presence, in a 1.5 km x ca. 180 m transect in each segment predicted on 10 evenly-spaced dates from 15 November through 1 April in each winter. Spatiotemporal variation in conditional abundance (%; bottom row) is indicated by the median absolute deviation, MAD, relative to the median. Predicted values are categorized based on their quartiles; segments with the highest conditional abundance or variability (values ≥ 98th percentile) are outlined in black.

Figure 5. Unconditional abundance of Common Eider (COEI), scoter (SCOT), and Long-tailed Duck (LTDU) in Nantucket Sound during three winters, 2003 - 2005. Median abundances (top row) represent the expected number of sea ducks along a 1.5 km x ca. 180 m transect within each segment predicted on 10 evenly-spaced dates from 15 November through 1 April in each winter. Spatiotemporal variation in abundance (%; bottom row) is estimated from the median absolute deviation, MAD, relative to the median. Predicted values are categorized based on their quartiles; segments with the highest abundance or variability (values ≥ 98th percentile) are outlined in black.

Figure 6. Relationship between observed and predicted total abundance of Common Eider (COEI), scoter (SCOT), and Long-tailed Duck (LTDU) during 30 aerial surveys of Nantucket Sound over three winters, 2003 - 2005. The dashed line indicates a 1:1 relationship between predicted and observed abundances in surveyed segments; points below and above this line indicate underestimates and overestimates of predicted abundances, respectively.
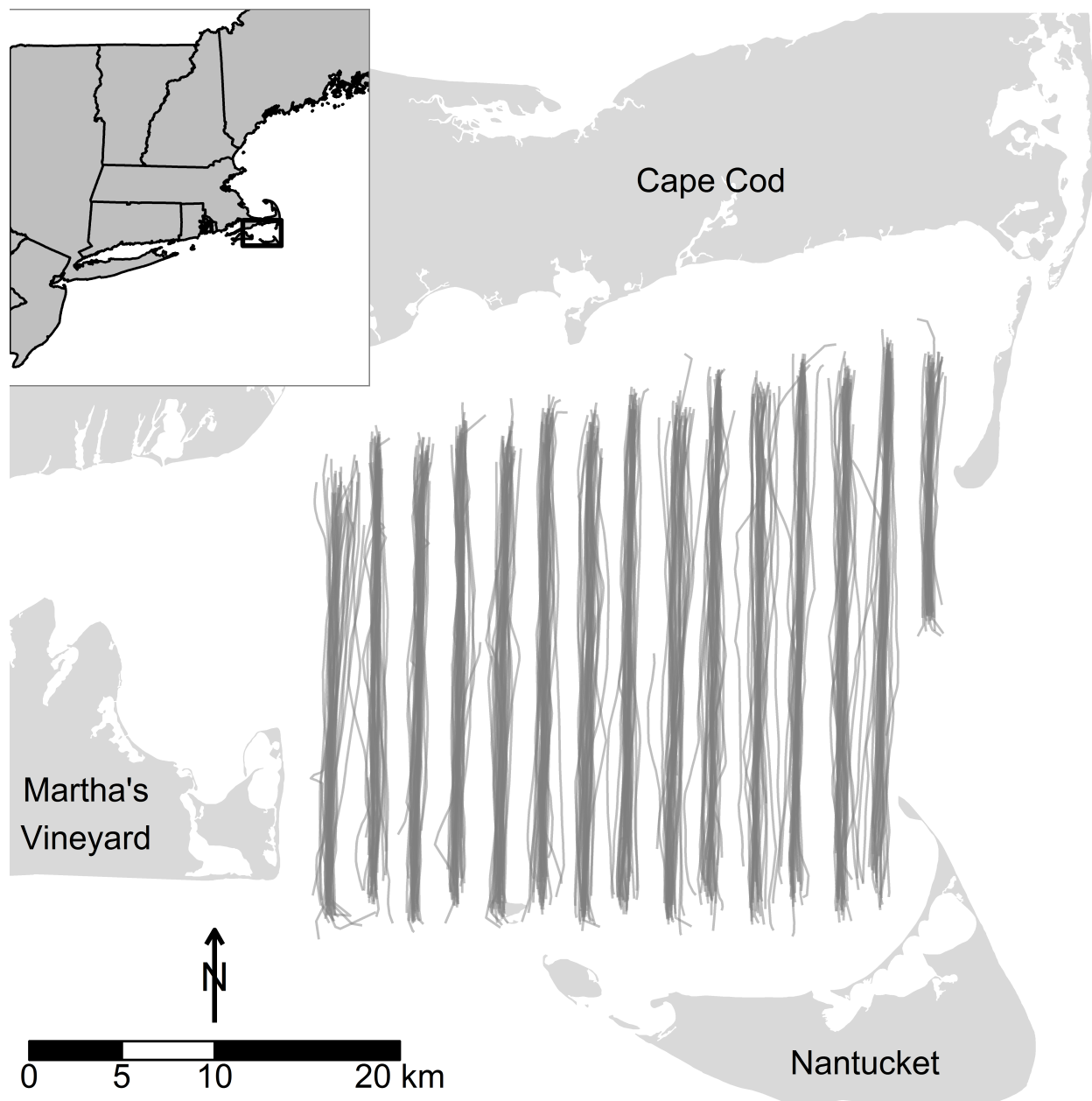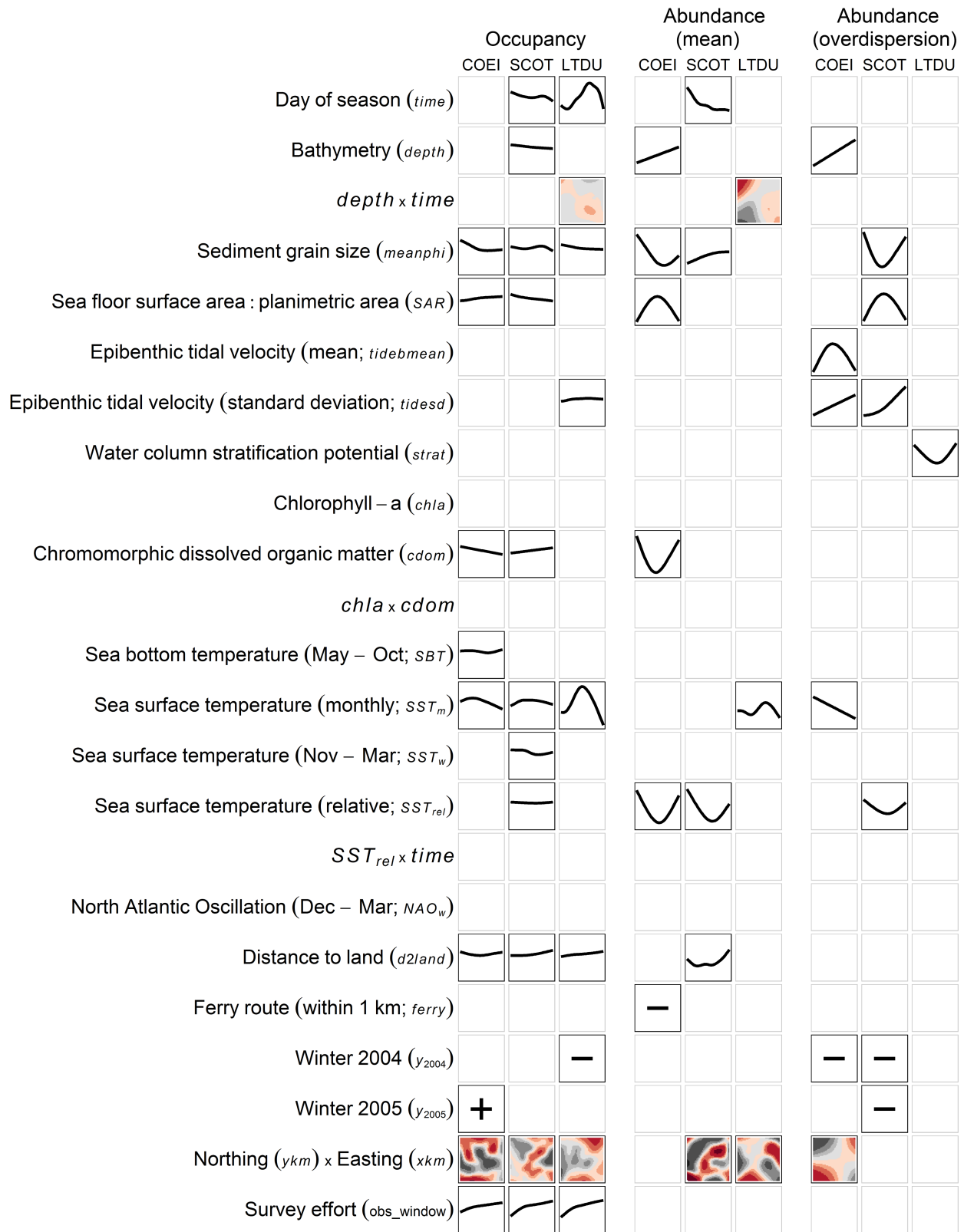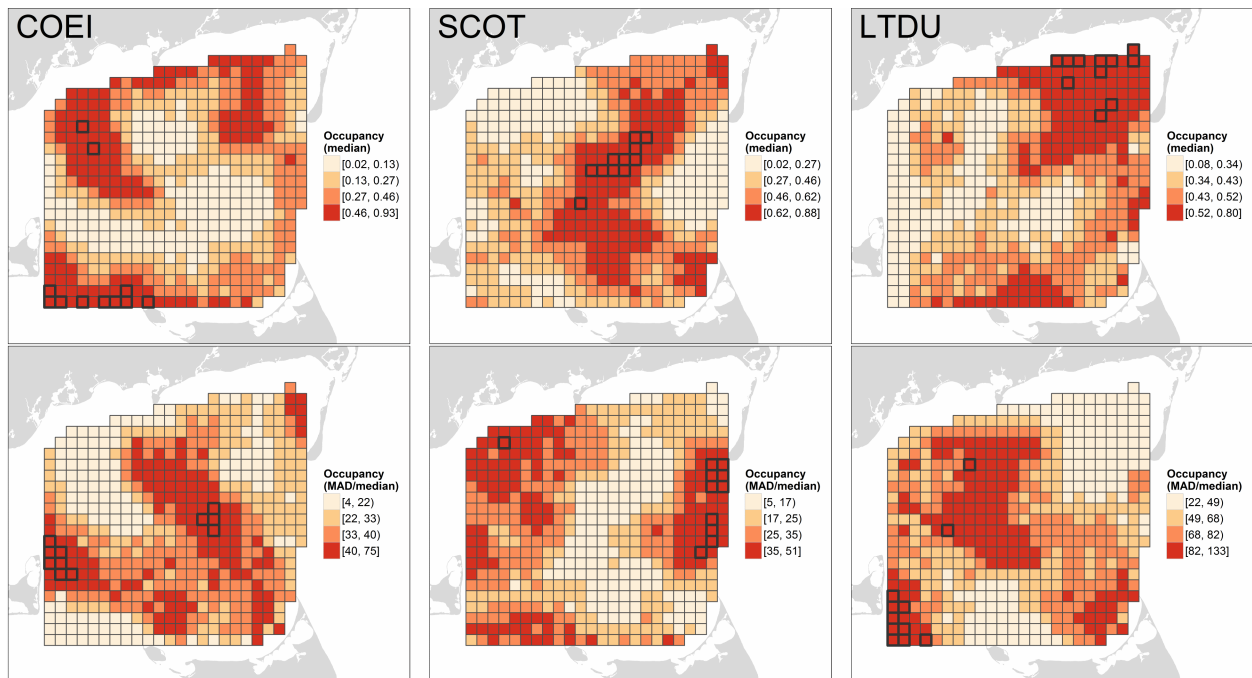
Cape Cod

Martha's
Vineyard

Nantucket

N

| 0 | 5 | 10 | 20 km |

737 Figure 1

738

Figure 2

740

Figure 3

Figure 4

744

745 **Figure 5**

747 **Figure 6**