Ecological Applications, 21(6), 2011, pp. 2241–2257

Where the wild things are: predicting hotspots of seabird aggregations in the California Current System

Nadav Nur,[1] Jaime Jahncke,[1,7] Mark P. Herzog,[1,8] Julie Howar,[1] K. David Hyrenbach,[2] Jeannette E. Zamon,[3] David G. Ainley,[4] John A. Wiens,[1] Ken Morgan,[5] Lisa T. Ballance,[6] and Diana Stralberg[1,9]

[1]*PRBO Conservation Science, 3820 Cypress Drive, #11, Petaluma, California 94954 USA*
[2]*Department of Oceanography, Hawaii Pacific University, 1164 Bishop S., Honolulu, Hawaii 96813 USA*
[3]*NOAA Fisheries–Northwest Fisheries Science Center, Point Adams Research Station, P.O. Box 155, Hammond, Oregon 97121 USA*
[4]*H. T. Harvey and Associates, 983 University Avenue, Building D, Los Gatos, California 95032 USA*
[5]*Canadian Wildlife Service, Environment Canada, P.O. Box 6000, 9860 West Saanich Road, Sidney, British Columbia V8L 4B2 Canada*
[6]*Southwest Fisheries Science Center, NOAA Fisheries, 8604 La Jolla Shores Drive, La Jolla, California 92037 USA*

Abstract. Marine Protected Areas (MPAs) provide an important tool for conservation of marine ecosystems. To be most effective, these areas should be strategically located in a manner that supports ecosystem function. To inform marine spatial planning and support strategic establishment of MPAs within the California Current System, we identified areas predicted to support multispecies aggregations of seabirds ("hotspots"). We developed habitat-association models for 16 species using information from at-sea observations collected over an 11-year period (1997–2008), bathymetric data, and remotely sensed oceanographic data for an area from north of Vancouver Island, Canada, to the USA/Mexico border and seaward 600 km from the coast. This approach enabled us to predict distribution and abundance of seabirds even in areas of few or no surveys. We developed single-species predictive models using a machine-learning algorithm: bagged decision trees. Single-species predictions were then combined to identify potential hotspots of seabird aggregation, using three criteria: (1) overall abundance among species, (2) importance of specific areas ("core areas") to individual species, and (3) predicted persistence of hotspots across years. Model predictions were applied to the entire California Current for four seasons (represented by February, May, July, and October) in each of 11 years. Overall, bathymetric variables were often important predictive variables, whereas oceanographic variables derived from remotely sensed data were generally less important. Predicted hotspots often aligned with currently protected areas (e.g., National Marine Sanctuaries), but we also identified potential hotspots in Northern California/Southern Oregon (from Cape Mendocino to Heceta Bank), Southern California (adjacent to the Channel Islands), and adjacent to Vancouver Island, British Columbia, that are not currently included in protected areas. Prioritization and identification of multispecies hotspots will depend on which group of species is of highest management priority. Modeling hotspots at a broad spatial scale can contribute to MPA site selection, particularly if complemented by fine-scale information for focal areas.

Key words: bagged decision trees; California Current System; habitat-association models; hotspots; marine protected areas; marine reserves; multispecies aggregations; pelagic ecosystems; remote-sensed data; seabirds; spatial predictive models.

Introduction

Across the globe, marine ecosystems have been exploited, disturbed, and altered, putting many species at grave risk (Norse et al. 2005, Polldoro et al. 2008, Halpern et al. 2009, Ojeda-Martínez et al. 2009). In response, marine spatial planning has emerged as a comprehensive approach to improve management and conservation of these ecosystems (Ray 2010). The ultimate objective of marine spatial planning is to analyze and allocate the distribution of human activities, in space and time, in marine areas so as to achieve ecological and economic objectives identified through a political process (Douvere and Ehler 2009). Marine reserves or marine protected areas (MPAs) have been widely viewed as an effective way to advance marine spatial planning and restore ecosystems and populations (Boersma and Parrish 1999, Worm et al. 2003, Norse et al. 2005). To be effective, however, marine spatial

Manuscript received 22 July 2010; accepted 26 October 2010; final version received 21 December 2010. Corresponding Editor: P. K. Dayton.

[7] Corresponding author. E-mail: jjahncke@prbo.org
[8] Present address: U.S. Geological Survey, Western Ecological Research Center, One Shields Avenue, University of California, Davis, California 95616 USA.
[9] Present address: University of Alberta, Department of Biological Sciences, CW 405, Biological Sciences Building, Edmonton, Alberta T6G 2E9 Canada.

planning and the establishment and management of Marine Protected Areas (MPAs) require sound scientific information.

Development of a scientific basis for the establishment of protected areas in marine ecosystems has been lagging compared to that in terrestrial habitats (Hyrenbach et al. 2000, Halpern 2003, Cañadas et al. 2005, Game et al. 2009). The identification and protection of foraging areas for marine predatory species such as seabirds is a high conservation priority (Hooker and Gerber 2004). Because seabirds feed on fish, squid, and plankton species that support other marine top predators (Polldoro et al. 2008), seabirds can also serve as indicators of areas where energy flows rapidly from lower trophic levels to top predators (Furness and Camphuysen 1997, Hooker and Gerber 2004, Durant et al. 2009; but see Grémillet and Charmantier 2010).

Marine birds and mammals aggregate at predictable locations or "hotspots" where food availability is high (Hunt et al. 1999, Cañadas et al. 2005, Piatt et al. 2006, Suryan et al. 2006). Productivity in these areas is often driven by topography and wind patterns, which create oceanographic features that may concentrate prey and are associated with measurements of sea-level height, sea-surface temperature gradients, and chlorophyll concentration (Hyrenbach et al. 2000, 2006, Ballance et al. 2006, O'Hara et al. 2006, Yen et al. 2006, Ainley et al. 2009, Garthe et al. 2009).

Here we ask whether widely available environmental data, including those from remote-sensing satellites, can be used to derive predictive models identifying multi-species aggregations ("hotspots") throughout the highly productive California Current Large Marine Ecosystem (hereafter, CCS; Ekstrom 2008). This large upwelling ecosystem stretches from British Columbia, Canada, to Baja California, Mexico, and seaward for hundreds of kilometers (Fig. 1). The CCS supports many important populations of marine wildlife as well as valuable commercial fisheries (Halpern et al. 2009, Teck et al. 2010). Like most of the world's oceans, the CCS has recently suffered from overexploitation of fish stocks and climatic and oceanographic anomalies (Bakun and Weeks 2004), highlighting the need to protect important wildlife areas. Hotspots for marine organisms have not yet been identified for the CCS on an ecosystem-wide basis. In addition, many areas in the CCS have been inadequately surveyed, so the ability to develop a comprehensive regional perspective on seabird distributions would help to advance marine spatial planning, highlight areas in need of protection, and prioritize future survey efforts.

In this paper we describe the development and application of habitat-association models for a 16-species assemblage of seabirds based on 11 years of survey data spanning the geographic range of the CCS. Model output provides predictive maps covering the full range of the CCS, including areas where little or no survey data are available. We use three criteria to combine predictions across the 16 species to highlight different aspects of potential areas of seabird aggregation, all of which are relevant for marine spatial planning. The three criteria reflect abundance, standardized across species; importance to individual species (as indicated by "core" areas); and persistence of hotspots among years. Because the modeling is conducted over the entire ecosystem, it enables us to consider not only the location and identification of individual MPAs, but also spatial relationships among sites (Lascelles et al. 2009). Model results can therefore inform the designations of MPAs and, more broadly, marine spatial planning (Douvere and Ehler 2009).

There are few examples in which the distribution and abundance of a large number of species have been modeled over a large area and over a considerable time period, especially so with respect to marine ecosystems. Hence, the modeling approach and synthesis of results presented here may provide a valuable approach for application to other marine taxa as well as other regions of the globe.

Methods

Data collection and processing

Bird observational data.—Seabird observation data were obtained from several research and monitoring programs: California Cooperative Oceanic Fisheries Investigation (CalCOFI), National Marine Fisheries Service (NMFS) California Current Ecosystem Study (National Oceanic and Atmospheric Administration [NOAA] and PRBO Conservation Science), CSCAPE and ORCAWALE (NOAA Southwest Fisheries), Line P (Canadian Wildlife Service [CWS] of Environment Canada [EC] and Fisheries and Oceans Canada), and NMFS Rockfish Surveys (NOAA) (Appendix A). Spatial coverage for the analysis lay between 52° N and 30° N latitude and from the shoreline out to 600 km from the mainland (Fig. 2). The northern boundary corresponded to the approximate northern edge of the California Current (Fig. 1); the southern and offshore boundaries were dictated by the extent of survey coverage (Fig. 2). Spatial coverage varied by cruise, with some providing extensive but sparse coverage of the entire west coast from Baja California to Vancouver Island and others providing dense coverage of smaller regions. Temporal coverage included the period from October 1997 to November 2008 (Appendix A). Data from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite were not available prior to October 1997, precluding the use of seabird data before that time for the purposes of modeling.

Surveys used the strip-survey method (Tasker et al. 1984, Buckland et al. 1993) to quantify seabird distribution and abundance (see, for example, detailed descriptions in Clarke et al. 2003 and Yen et al. 2004). The method employed for all surveys recorded birds continuously while the survey vessel was under way: all birds within a survey strip off one forequarter of the

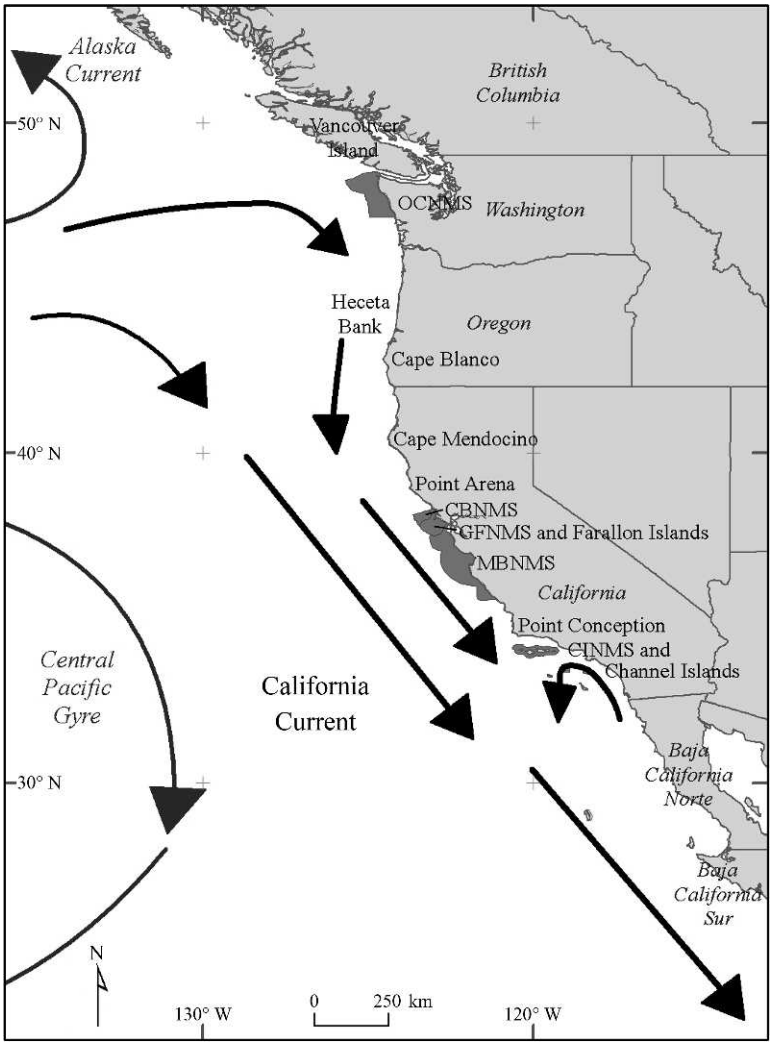


FIG. 1. Portion of Pacific Ocean, depicting the California Current and showing the main direction of flow and important land and ocean features. National Marine Sanctuary (NMS) boundaries are depicted for the Olympic Coast NMS, Cordell Bank NMS, Gulf of the Farallones NMS, Monterey Bay NMS, and Channel Islands NMS.

ship's bow were counted and their behavior was classified. The width of the strip being surveyed depended on the survey vessel and weather conditions, but was between 250 and 500 m, mostly 300 m (84% of survey transects). Although the presumption is that all birds encountered within the strip are detected and recorded, it is likely that detectability varied as a function of species (i.e., size and color) and exogenous conditions (e.g., sea state, cloud cover). Nevertheless, we feel that our approach was warranted because our intention was not to estimate actual absolute density, in which case some correction for bird flight speed and direction would be required (e.g., Spear et al. 1992), but rather to derive an index of relative density that can be combined across species. Furthermore, we standardized abundance across species, thus eliminating species-specific differences in detectability from influencing combined species results.

Each survey transect was divided into 3 km long segments or "bins." We used 3-km bins because previous analysis indicated generally low levels of spatial autocorrelation among adjacent bins at that scale (Yen et al. 2004), although autocorrelation is not a major concern for predictive models such as ours (Diniz-Filho et al. 2003, Ferguson et al. 2006). Seabird observations and survey transect data were represented by points corresponding to the midpoint of a survey bin. Bin center points were used for analysis of distance from land, distance to isobaths, bathymetry, sea-surface temperature (SST), chlorophyll *a* concentration (CHL), and sea-surface height (SSH). For all cruises combined, 58 966 survey bins were used in the statistical analysis. We controlled for area of the survey bin to standardize the number of individuals detected per bin to the number of individuals per square kilometer. In >75% of the surveys the bin area was 0.9 km^2 (i.e., strip width =

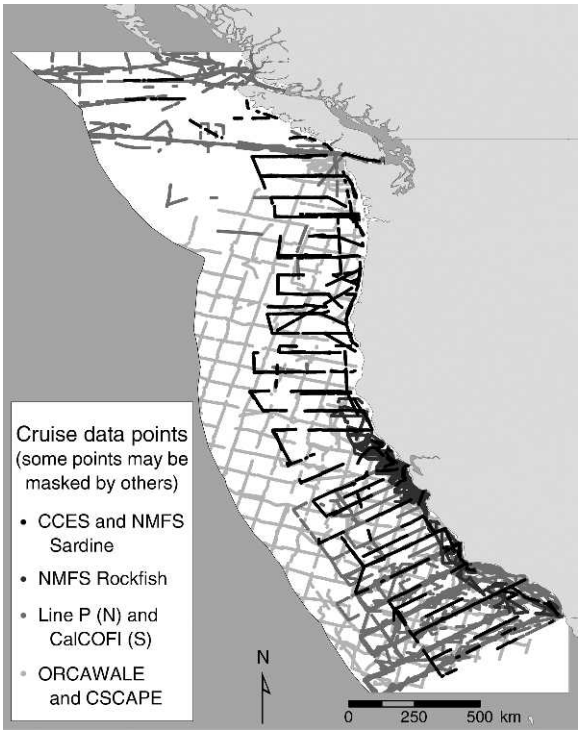


FIG. 2. Map of surveys used in this study of seabird hotspots. Individual data points are shown classified by cruise project or program. See Appendix A for information on cruises.

300 m × 3 km). Each observation included information on GPS position, visibility, and species, number, and behavior of animals observed. Counts of individuals of a single seabird species, by behavior classification, were summed within each bin (Appendix B).

Environmental data.—Bathymetric data were obtained in raster form from the General Bathymetric Chart of the Oceans (GEBCO) with a cell size of 1 × 1 arc-minute (approximately 1.85 × 1.5 km, exact dimensions of the cell varying with latitude). Depth (in meters) for each bin was obtained for each bin center point. We also developed moving-window statistics for depth. For each bathymetric grid cell, we calculated the mean, minimum, maximum, and standard deviation of depth based on the "focal" cell and the eight adjacent cells. Thus, each focal cell was located within a group of nine cells (3 × 3 cells), whose area was ~24 km^2. We also calculated a "contour index" that reflects topographic relief of the sea floor within the same nine cells. Contour index was defined as

$$[(\text{max. depth} - \text{min. depth})/\text{max. depth} \cdot 100].$$

This index varied from 0 to 100, with higher values indicating steeper bathymetry.

Distances from the bird-survey bin center points to the mainland and to islands were calculated in ArcMap 9.2 (ESRI 2006) using shapefiles provided by California Department of Fish and Game and ESRI. Distances from the center points to the 200-m, 1000-m, and 3000-m depth isobaths were calculated from bathymetry data. These three isobaths have been used in similar analyses (e.g., Yen et al. 2005): inside the 200-m isobath is considered the continental shelf, the 1000-m isobath is indicative of the position of the continental shelf slope, and the 3000-m isobath indicates the ocean floor.

Two sources of satellite-derived sea-surface temperature (SST) data were combined to maximize spatial and temporal data coverage: Moderate Resolution Imaging Spectroradiometer (MODIS) (Aqua and Terra) and Advanced Very High Resolution Radiometer (AVHRR) Pathfinder data sets. We used AVHRR SST data when they were available. When AVHRR SST data were not available, we used a predictive equation to estimate AVHRR SST as a function of the available MODIS SST value ($R^2 = 0.96$). The regression equation was based on the set of cells that had both AVHRR and MODIS SST values. The regression-derived estimates of AVHRR SST were used when observed AVHRR SST data were not available. We used daytime composite temperatures over the 8-day period corresponding to the date of the bird survey. Data cell sizes were, on average, 4.6 × 3.7 km.

Chlorophyll (CHL) concentration data (milligrams per cubic meter) were obtained from the SeaWiFS satellite. As with temperature, 8-day composite periods were constructed to minimize missing data, and these were matched with dates and center points of the survey bin. Cell size was, on average, 9.3 × 7.3 km (as with other metrics, exact dimensions varied with latitude). Where SeaWiFS data were unavailable, we used data from MODIS, processed as with SST (i.e., we used a regression equation to predict SeaWiFS values on the basis of MODIS values).

Sea-surface height (SSH) data were obtained from multiple satellites with the same ground-track (including Topex/Poseidon, Jason-1 + ERS, Envisat). We used a composite value corresponding to a 7-day time period. Original spatial coverage was global, with a cell size of 0.25 × 0.25 degrees (~28 km on one side × 22.0 km, depending on latitude). SSH (in centimeters) for each survey bin was obtained by matching the location and date of the survey bin with the 7-day composite SSH raster.

Data for predictive modeling.—Physical and biological data for the modeling process were available from 1997 to 2008; all months were used to develop the predictive models themselves. However, we picked just four focal months for the purposes of model predictions: October, February, May, and July (representing fall, winter, spring, and summer, respectively); predictions were made for each of the four months in each year (1997 to 2008). We chose a "seasonal" approach to prediction rather than producing month-by-month predictions because the survey coverage was adequate at the seasonal scale, but, in general, surveys did not provide adequate month-by-month coverage of the study area to justify 12 monthly predictions per year.

Our choice of months reflected the seasonal cycle of seabirds in the CCS (e.g., February is pre-breeding period for those that breed in the CCS, May is early breeding season, July is late breeding season, and October is the nonbreeding period). With regard to the annual cycle of physical oceanography, October occurs within the "Oceanic" season; February falls within the "Davidson Current" season; May occurs within the middle of the "Upwelling" season; and July falls toward the end of the "Upwelling" season (Ford et al. 2004).

Data used for predictions were similar to those used for analysis except that only four months were used in each year, as just described, and the satellite data were based on monthly average values rather than 8-day values. Overall, 45 months were used for prediction between October 1997 and October 2008. After processing, geodetic datum, spatial extent, and cell size were standardized for all data; the datum was WGS 84; boundaries were limited to −137° W, 30° N, −117° W, 52° N; and the cell size was resampled to the standard cell size of MODIS SST data (~4.6 × 3.7 km).

Single-species statistical analyses

For most of the species initially analyzed, only birds observed foraging or on the water were included in the analyses. For all gulls, terns, kittiwakes, albatrosses, and storm-petrels, however, we included flying birds as well, as these species often forage while flying (following Jahncke et al. 2008) and often are not observed except when flying (Clarke et al. 2003).

We selected species for analysis through a two-step process. First, we selected all seabird species observed in at least 300 bins (i.e., one or more individuals of the appropriate behavior codes were observed in the survey "bin"). This minimum represents ~0.5% of all survey bins analyzed, and was used primarily to filter out rarer species. Twenty-five species met this criterion. We then conducted bagged decision-tree (BDT) analysis on each of the 25 species.

We evaluated predictive adequacy of the resultant BDT model by examining the proportion of deviance explained for each species using test data (data not part of the "training data" used to develop predictions). Species models in which the proportion of deviance explained exceeded 0.40 for test data were considered further. For each of these species, we evaluated goodness of fit by determining whether the drop in predictive ability comparing training data and test data (Hastie et al. 2009) was low to moderate; more specifically, we determined whether the percentage of deviance explained in the test data was between 60% and 100% that of the training data. All species that met the first criterion (i.e., proportion of deviance explained in test data > 0.40), also met the latter criterion. Of the 25 species examined, 16 species produced adequate predictive models and were retained (Table 1; species summary data in Appendix B). All but one of the species retained is considered a species of concern by the Baja to Barrow Initiative of the Audubon Society (Table 1). In addition, two of the 16 species are considered globally threatened, near-threatened, or endangered by the IUCN: Black-footed Albatross and Sooty Shearwater (Table 1).

Bagged decision-tree analysis.—We used bagged decision trees, one type of machine-learning methodology (also referred to as "data mining"; Hochachka et al. 2007). Machine learning algorithms, and decision-tree methodology in particular, have only recently become more widely used in ecological studies (e.g., niche modeling) in marine and terrestrial systems (De'ath 2007, Elith et al. 2008, Leathwick et al. 2008). Decision-tree analysis uses binary "rules" to classify and predict responses based on a set of predictor variables (Breiman et al. 1984, Hastie et al. 2009). However, simple decision trees are highly sensitive to the data at hand and can provide predictions that "overfit" the data. Bootstrap aggregation, or "bagging," enhances the accuracy of the decision tree by using a large number of bootstrapped data sets (an "ensemble"; see Efron and Tibshirani 1993) and performing a separate decision-tree analysis on each data set (Breiman 1996, Hochachka et al. 2007, Hastie et al. 2009). The predictions from each decision tree are then combined and averaged to produce a final prediction. BDT analysis and similar machine-learning methods are discussed in an ecological context by Elith et al. (2008).

There are a number of strengths that make BDT a good choice for this study. (1) The method has been shown to perform well in terms of prediction (Dietterich 2000, Hochachka et al. 2007); (2) the method makes use of observations that have missing values for some of the predictor variables; (3) the method is nonparametric (i.e., no assumption is made about a particular shape of the response function to a predictor variable); (4) a large number of predictor variables can be incorporated into the model (Guyon and Elisseeff 2003); and (5) interactions of predictor variables are incorporated into the final models and these do not have to be specified. This last feature is particularly relevant to this study: because latitude and distance from land were two of the predictor variables, BDT allowed for the effects of other environmental variables (e.g., SST) to vary across the spatial range of the study area. In addition, "day of year" and variables reflecting annual variation were also included in BDT models; hence BDT allowed for the effect of environmental variables to vary within and among years.

We used the recursive partitioning package (rpart) within the R statistical programming language (version 2.10.1, R Development Core Team 2009) for the analysis. Custom code developed by D. Fink was used to implement the bagging, prediction, model averaging, and model fit (Hochachka et al. 2007). We used BDTs to develop predictive models for each of the 16 species. The BDT models were used to make predictions for the entire region of the California Current in each of four months, in each year. As a first step, we partitioned 90%

TABLE 1. Seabird species analyzed and modeled. The table also indicates each species' status as a species of concern in the Baja to Barrow Initiative (B2B) and its IUCN status.

Species code	Common name	Scientific name	B2B	IUCN status
BFAL	Black-footed Albatross	*Phoebastria nigripes*	x	endangered
BOGU	Bonaparte's Gull	*Larus philadelphia*	x	least concern
BRAC	Brandt's Cormorant	*Phalacrocorax penicillatus*	x	least concern
BRPE	Brown Pelican	*Pelecanus occidentalis*	x	least concern
CAAU	Cassin's Auklet	*Ptychoramphus aleuticus*	x	least concern
CAGU	California Gull	*Larus californicus*	x	least concern
COMU	Common Murre	*Uria aalge*	x	least concern
FTSP	Fork-tailed Storm-Petrel	*Oceanodroma furcata*	x	least concern
GWGU	Glaucous-winged Gull	*Larus glaucescens*	x	least concern
HEEG	Heermann's Gull	*Larus heermanni*	x	least concern
HERG	Herring Gull	*Larus argentatus*		least concern
LHSP	Leach's Storm-Petrel	*Oceanodroma homochroa*	x	least concern
RNPH	Red-necked Phalarope	*Phalaropus lobatus*	x	least concern
SAGU	Sabine's Gull	*Xema sabini*	x	least concern
SOSH	Sooty Shearwater	*Puffinus griseus*	x	near-threatened
WEGU	Western Gull	*Larus occidentalis*	x	least concern

Notes: In addition, nine analyzed species did not yield adequate predictive models (see *Methods: Single-species statistical analyses*): Black-legged Kittiwake (*Rissa tridactyla*), Rhinoceros Auklet (*Cerorhinca monocerata*), Black-vented Shearwater (*Puffinus opisthomelas*), Red Phalarope (*Phalaropus fulicarius*), Pacific Loon (*Gavia pacifica*), Pomarine Jaeger (*Stercorarius pomarinus*), Northern Fulmar (*Fulmarus glacialis*), Cook's Petrel (*Pterodroma cookii*), and Pink-footed Shearwater (*Puffinus creatopus*).

of the data for use during model building (i.e., training), leaving 10% available for testing. For each predictive model, we used an ensemble of 500 bootstrapped data sets obtained from the training data (Hochachka et al. 2007). A priori tests with these data showed that there was little improvement in the predictive ability of the bagged ensemble beyond 500 bootstrapped data sets. We assumed a Poisson distribution for the response variable (birds detected per survey bin) in our BDT analysis.

Variable selection.—We identified 20 habitat variables as potentially meaningful for inclusion in the modeling (Table 2). We arrived at this selection of variables by considering previous studies on seabird–habitat relationships, in general and in the CCS in particular (e.g., Yen et al. 2004, Ainley et al. 2009, Tremblay et al. 2009), and the limitations of data availability for the study period and study area.

Date, upwelling transition, and oceanographic indices.—Day of year was included, allowing for seasonal fluctuations in abundance for individual species. We also used "transition date," the date at which the upwelling regime in the CCS changes from a winter pattern to a spring pattern. Transition date demonstrates strong year-to-year variation (Holt and Mantua 2009). Although transition dates in the CCS show regional variation, we were unable to quantify the variations in transition date throughout the system. Instead, we relied on a single annual value based on upwelling and winds (Method 1 of Holt and Mantua 2009), obtained from data near the Farallon Islands in Central California (J. Jahncke and J. Roth, *unpublished manuscript*). BDT allows for an interaction between transition date and latitude, so the effect of transition date in the analysis can vary with latitude. Use of a single transition date in each year for the CCS in the BDT analysis may be an oversimplification. However, it was not our intention to develop a detailed accounting of transition date itself, but rather to characterize how variation in abundance of the target seabird species was related to variation in year, date, and related variables.

Three oceanographic indices were included: Southern Oscillation Index, Pacific Decadal Oscillation, and Northern Pacific Gyre Oscillation (Di Lorenzo et al. 2008). Index values were calculated for two time periods: the three-month period that included the month of survey and the two previous months, as well as the three months prior to that period (i.e., 3–5 months prior to the month of survey).

Multispecies hotspot determination

We used three criteria to identify multispecies hotspots. All three were implemented at the scale of a single "prediction cell," $\sim 4.6 \times 3.7$ km.

Criterion 1: Standardized abundance, summed over all species.—Predicted densities were standardized for each species such that mean density = 0 and SD of density = 1. Standardization was implemented before combining results across species so that each species contributed equal weight to the composite results. Otherwise, species with very high mean abundance would swamp less common species. Species with high variance also could have a large influence on the composite results. We use the term "abundance" henceforth, but recognize that observations and predictions refer to individuals detected during surveys, which only provide an index of abundance. Note that this standardization accounted for possible differences in detectability among species. The standardized abundance values were then summed over all species for the individual predicted cell. We refer to this metric as "summed standardized abundance" to emphasize that it is a multispecies metric and not a

TABLE 2. Variables included in predictive models of seabird hotspots.

Variable	Type of variability	Scale (spatial or temporal)
Latitude	spatial	NA
Depth, minimum	spatial	5.6 km
Depth, mean	spatial	5.6 km
Contour index	spatial	5.6 km
Distance to 200-m isobath	spatial	NA
Distance to 1-km isobath	spatial	NA
Distance to 3-km isobath	spatial	NA
Distance to land	spatial	NA
Chlorophyll	spatial and temporal	9.3 km, 8 days
SST (sea-surface temperature)	spatial and temporal	4.6 km, 8 days
SSH (sea-surface height)	spatial and temporal	28 km, 7 days
Day of year	temporal	within year
Transition date	temporal	annual
Year	temporal	annual
SOI, 0–2 months before	temporal	3 months
PDO, 0–2 months before	temporal	3 months
NPGO, 0–2 months before	temporal	3 months
SOI, 3–5 months before	temporal	3 months
PDO, 3–5 months before	temporal	3 months
NPGO, 3–5 months before	temporal	3 months

Notes: Inside the 200-m isobath is considered the continental shelf; the 1000-m isobath indicates the position of the continental shelf slope; the 3000-m isobath indicates the ocean floor. Transition date is the date at which the upwelling regime changes from a winter pattern to a spring pattern. SOI is the Southern Oscillation Index, PDO is the Pacific Decadal Oscillation, and NPGO is the Northern Pacific Gyre Oscillation. Data sources are: AVHRR Pathfinder v5 SST data from the Jet Propulsion Laboratory's Physical Oceanography Distributed Active Archive Center (http://podaac.jpl.nasa.gov/DATA_CATALOG/avhrr.html); MODIS Aqua and Terra SST and CHL data from OceanColor Web (http://oceancolor.gsfc.nasa.gov); bathymetric data from GEBCO (http://www.gebco.net/data_and_products/gridded_bathymetry_data/); SeaWiFS CHL data from OceanColor Web (http://oceancolor.gsfc.nasa.gov); sea-surface height data from multiple satellites compiled by AVISO (http://www.aviso.oceanobs.com/en/home/index.html). "NA" means not applicable.

single-species metric. These calculations were carried out by month and year, and then results were averaged over years by month as well as averaged over all years and months.

Criterion 2: Important "core" areas for individual species.—For each species, we ranked all prediction cells according to the predicted abundance of that species by month and year. We then identified the smallest set of cells that together constituted 25% of the species' total predicted abundance within the study region. These "highest predicted density" cells were considered to indicate a species' "core area" and were assigned a score of 2. The set of cells that together made up the next quartile of the species' total abundance based on predicted density, were considered important "shoulder" areas and were assigned a score of 1. Thus, the smallest set of cells that constituted 50% of the species' total predicted abundance within the study region received a score of 2 ("core") or 1 ("shoulder). All remaining cells received a score of 0. We chose these two criteria (top 25% and 50% of predicted total abundance), on the basis of previous studies of habitat use (e.g., Hyrenbach et al. 2002) and after preliminary evaluation of several thresholds (e.g., top 10%, top 75%).

Importance was calculated for each species and a weighted average was then calculated over all species. The average was weighted by assigning a weight that was inversely proportional to the number of cells making up the 25% core area for a species. Thus, if species X required 100 cells to reach the 25% criterion whereas species Y required 1000 cells, each "core" cell for species X received 10 times the weight of each "core" cell for species Y. The weighting function was scaled to have mean = 1. We used this weighting function to normalize the contribution of each species to the overall "importance" score. More specifically, the product of the weighting factor and the number of core cells for each species was the same across all species. Species-specific abundances were calculated first (by month, averaged over all years, and overall, averaged over all months and years), and then species-specific importance scores were determined.

Criterion 3: Persistence of hotspots across years.—This index was calculated for individual species for each month separately (February, May, July, and October) and for all months combined. The number of years (out of 11) that a cell was in the top 5% of predicted summed standardized abundance was scored (i.e., in the 95th percentile of all prediction cells for that month or for all months combined). Note that the persistence score itself was only calculated within a specified month or set of months; thus our measure of persistence reflected variation among years but not fluctuations due to seasonal changes in abundance of a species.

Combining all three criteria.—To synthesize information on all three criteria, we first standardized the

TABLE 3. Important and moderately important variables for each species-specific predictive model.

Variable	Proportion of variance explained for each species						
	BFAL	BOGU	BRAC	BRPE	CAAU	CAGU	COMU
Latitude	**0.100**	0.044	0.060	**0.111**	0.039	0.058	0.031
Minimum depth		0.016	0.065	0.022			**0.249**
Mean depth	0.021			0.040	0.034		**0.077**
Contour index		0.040	0.063		**0.156**	0.015	0.020
Distance to 200-m isobath	**0.099**	**0.158**	0.023	0.053	0.044	0.024	0.042
Distance to 1000-m isobath	**0.106**	0.046	0.027	0.033	**0.090**	0.024	0.026
Distance to 3000-m isobath	0.024		0.021	0.014	0.065	0.065	**0.116**
Distance to nearest land	0.021	0.073	**0.275**	**0.248**	0.070	**0.201**	0.044
Chlorophyll		0.051		0.015		0.016	
SST	0.024				0.026	0.015	
SSH	0.026	0.032			0.019	0.042	
Day of year	0.038	**0.106**	0.017	0.015	**0.081**	**0.119**	0.031
Year							
Transition date							
SOI 0–2 months before							0.018
PDO 0–2 months before			0.035	0.025			
NPGO 0–2 months before			0.033				
SOI 3–5 months before							
PDO 3–5 months before							
NPGO 3–5 months before	0.053						
Total proportion deviance explained	0.602	0.653	0.717	0.647	0.716	0.656	0.738

Notes: For each species-specific model, the top 10 variables (out of 20) and the proportion of deviance explained for each variable are shown. Variables with a proportion of deviance explained $\geq$0.075 are in boldface (see *Results: Predictor variables*). Species codes are in Table 1.

hotspot values for each criterion separately. We calculated z scores for each criterion, such that mean z score = 0 and SD = 1. However, because there was strong skew of importance and summed standardized abundance, we first log-transformed these hotspot scores before applying the z-transformation. We added a constant before log-transforming such that minimum log value for summed standardized abundance and for importance was 0. Thus, for these two criteria, we essentially standardized values to a geometric mean of 0 rather than arithmetic mean of 0. Persistence scores were not skewed so were not log-transformed. Finally, we averaged z scores across the three criteria.

RESULTS

Individual species models

Predictor variables.—The proportion of deviance explained in the training data set for the set of predictor variables exceeded 0.520 for all species; for five species it exceeded 0.700 (Table 3). Also shown in Table 3 is the proportion of deviance explained for individual variables, when the other 19 variables were included in the model. For example, the predictive model explained >60% of the deviance in the Black-footed Albatross training data set. The best predictive variable for this species was distance to the 1000-m isobath (10.6% of deviance explained), which reflects proximity to the shelf slope. Proportion of deviance explained for the top 10 most predictive variables for each species are shown in Table 3; results for the other 10 variables, for each species, are not shown. We emphasize that the predictive models themselves include all 20 variables. Table 3 also identifies "important" predictor variables, defined as variables for which the proportion of deviance explained > 0.075, given that the other 19 variables are included in the predictive model. Variables among the top 10 but with proportion of deviance explained <0.075 are considered "moderately important."

Latitude was included in the set of top 10 variables for every species and, for several species, this variable accounted for a large proportion of the total deviance (in particular, for Fork-tailed Storm-Petrel and Glaucous-winged Gull; Table 3). At least one of the three depth variables was an important or moderately important predictor (i.e., among the top 10 predictor variables) for every species except Sooty Shearwater. For Cassin's Auklet, the contour index was the strongest predictor variable among the 20. For Common Murres, minimum depth (within the nine-cell matrix) was the strongest predictor variable.

Distance to nearest land was an important predictor for nearly every species and, for seven species it was the predictor accounting for the highest proportion of deviance among the 20 variables (Table 3). Distances to the 200-m isobath, to the 1000-m isobath, and to the 3000-m isobath were all included among the top 10 variables in terms of proportion of deviance explained, for nearly all species.

Chlorophyll was a moderately important predictor for only eight species, and for SST this was the case for only six species. Among variables derived from remote-sensed satellite data, SSH was the most important predictor. For Sabine's Gull, SSH was the most important predictor of all variables considered, and for Fork-tailed Storm-Petrel, Red-necked Phalarope, and

TABLE 3. Extended.

Proportion of variance explained for each species								
FTSP	GWGU	HEEG	HERG	LHSP	RNPH	SAGU	SOSH	WEGU
0.240	**0.220**	0.065	0.035	0.072	0.045	0.079	0.039	**0.110**
	0.070	0.036	0.032		0.028			0.020
	0.016	0.014	**0.090**	0.024	0.032	0.017		0.027
0.016				0.024		0.016		0.015
	0.023	0.042	**0.083**	0.023	0.059		0.040	0.023
0.065	0.015				0.051	0.041	0.039	0.017
0.014	0.014	0.023	0.039		0.043	0.022	0.038	0.025
	0.040	**0.263**	0.031	**0.079**	**0.124**	0.063	**0.180**	**0.301**
0.061		0.068			0.028	0.017	0.048	
		0.068		0.027		0.017		
0.136	0.051		0.025	0.025	**0.097**	**0.220**	**0.084**	0.017
0.028	0.056	**0.177**	0.059	0.045	0.038	0.036	0.060	0.030
0.042	**0.103**			0.031				
			0.040					
0.019								
0.016							0.030	
		0.047	0.028	0.065			0.028	
0.694	0.654	0.794	0.582	0.527	0.662	0.592	0.710	0.663

Sooty Shearwater, it was the second most important predictor.

Day of year, reflecting seasonal variation in abundance and/or occurrence, was also included in all species' models, although it was never the most important predictor. With the exception of Glaucous-winged Gull, year was generally not an important predictor. Transition date was never included among the top 10 variables for any species. The generally low predictive contribution of these last two variables, year and transition date, which reflect annual variation, may be attributed to the inclusion of other variables in the species models that may have better captured year-to-year variation in relevant ecological influences, specifically CHL, SST, SSH, SOI, PDO, and NPGO.

Among the six oceanographic variables considered (three indices $\times$ two time periods), NPGO 3–5 months previous to the month of the cruise was more often a moderately important predictor variable than the other five variables, but in no case did an oceanographic index demonstrate a large proportion of deviance explained (proportion of deviance < 0.050; Table 3).

To summarize, the variables with highest predictive value were informative on location, especially proximity to land and to the various isobaths.

Single-species predictive maps.—We used the predictive models (summarized in Table 3) to create predictive maps for each of the four months within each year (4×11 years, plus a 12th year for October $= 45$ maps in total) for each species. To illustrate the species-specific predictions, we present results for one species, Cassin's Auklet (*Ptychoramphus aleuticus*). The pattern of predicted abundance of this species differed somewhat between February (averaged over all years; Fig. 3A) and May (averaged over all years; Fig. 3B). In February (immediately before the onset of breeding), only two areas of high predicted abundance for Cassin's Auklet are evident: off the northwest tip of Vancouver Island and surrounding San Nicolas Island, one of the California Channel Islands. By May (middle of the breeding season), areas of high abundance are more widespread, although still close to the coast, with San Nicolas and Santa Catalina Islands in southernmost California being the only hotspots that are some distance from the coast. In May, nearly the entire region of northern California and southern Oregon has high predicted abundance (Fig. 3B); single-species maps for each of the 16 species are *available online*.[10]

We can also use Cassin's Auklet to illustrate the calculation of importance and persistence for a single species (Fig. 4). The pattern of importance was similar for all four months (importance calculated over all seasons shown in Fig. 4A). The "core" areas identified include northern Vancouver Island, northern California (from the Golden Gate north), and southern California (Channel Islands and San Diego region), but excluding the mainland coast of southern California. For this species, areas that, on average, demonstrated high predicted abundance (Fig. 3) also demonstrated high consistency in predicted abundance among years (Fig. 4B), although this was not true for every species.

Multispecies hotspot determination

Predicted hotspots based on standardized abundance, summed over all species, showed strong seasonal variation. For February, the Gulf of the Farallones and nearby Monterey Bay constituted the only clear hotspot region, although the western Channel Islands

[10] ⟨www.prbo.org⟩

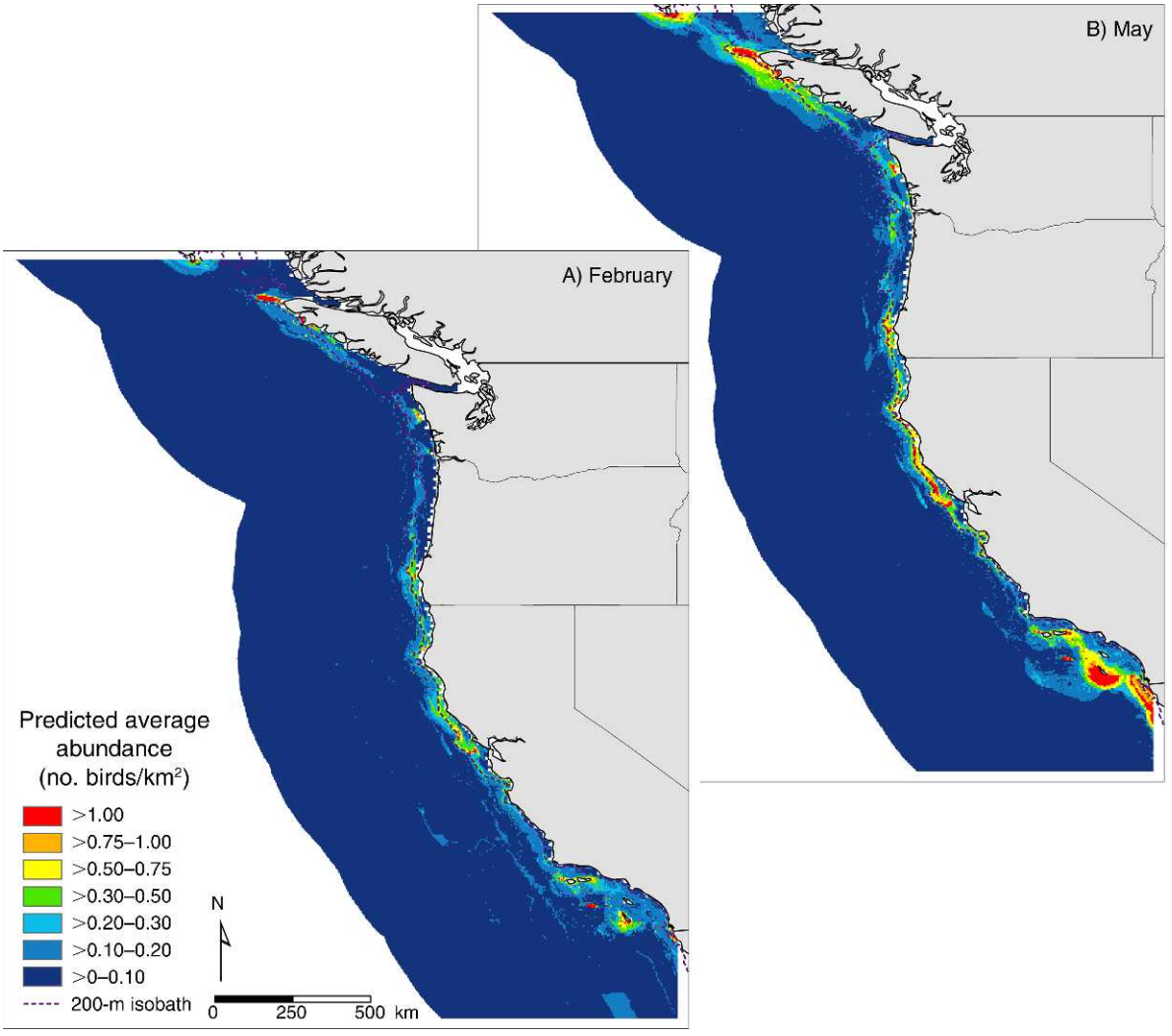


FIG. 3. Cassin's Auklet (*Ptychoramphus aleuticus*) predicted abundance (birds/km^2) by specified month, averaged over all years, for (A) February and (B) May.

appeared as an incipient hotspot (Fig. 5A). The northern California Current demonstrated no hotspots. In contrast, in May the coast of the Olympic Peninsula was a very evident hotspot, as was Heceta Bank, north of Cape Blanco in Oregon (see Fig. 1 for locations). The Gulf of the Farallones and Monterey Bay continued to show high summed standardized abundance in May. Other spatially restricted hotspots were found near Cape Mendocino in northern California and the western Channel Islands (Fig. 5B). In contrast, in October the Gulf of the Farallones/Monterey Bay region demonstrated relatively low summed abundance, with the Olympic Peninsula/southern Vancouver Island being an area of moderately high abundance, together with southern California (Fig. 5D).

Averaging across seasons, areas of high summed, standardized abundance were associated with the Olympic Peninsula coast, Heceta Bank, Cape Mendocino, Gulf of the Farallones, Monterey Bay, the Channel Islands, and the mainland coast of southern California (Fig. 6A). The pattern for hotspots identified by the importance criterion was generally similar to that of summed, standardized abundance (Fig. 6B).

No hotspots, whether for summed abundance or importance, were apparent more than ~90 km offshore (Fig. 6A, B).

There was also substantial overlap between areas identified as highly persistent and those identified as having high summed abundance and/or high species importance (compare Fig. 6C with Fig. 6A, B). However, some areas showing high relative persistence were not as evident using summed abundance or importance criteria, such as the northwest tip of Vancouver Island. Conversely, Heceta Bank demonstrated high summed abundance and high importance, but was not a persistence hotspot (Fig. 6C).

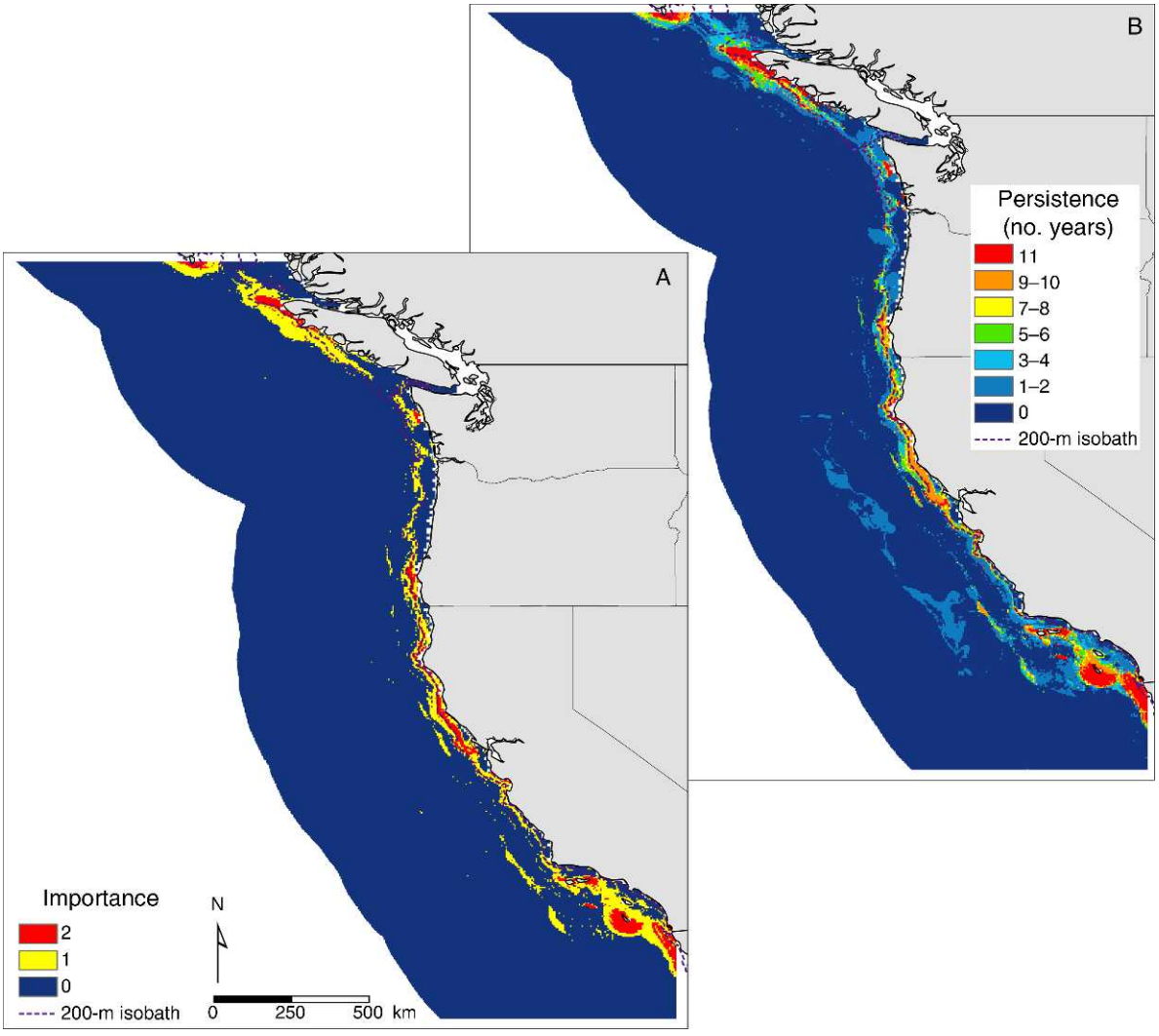
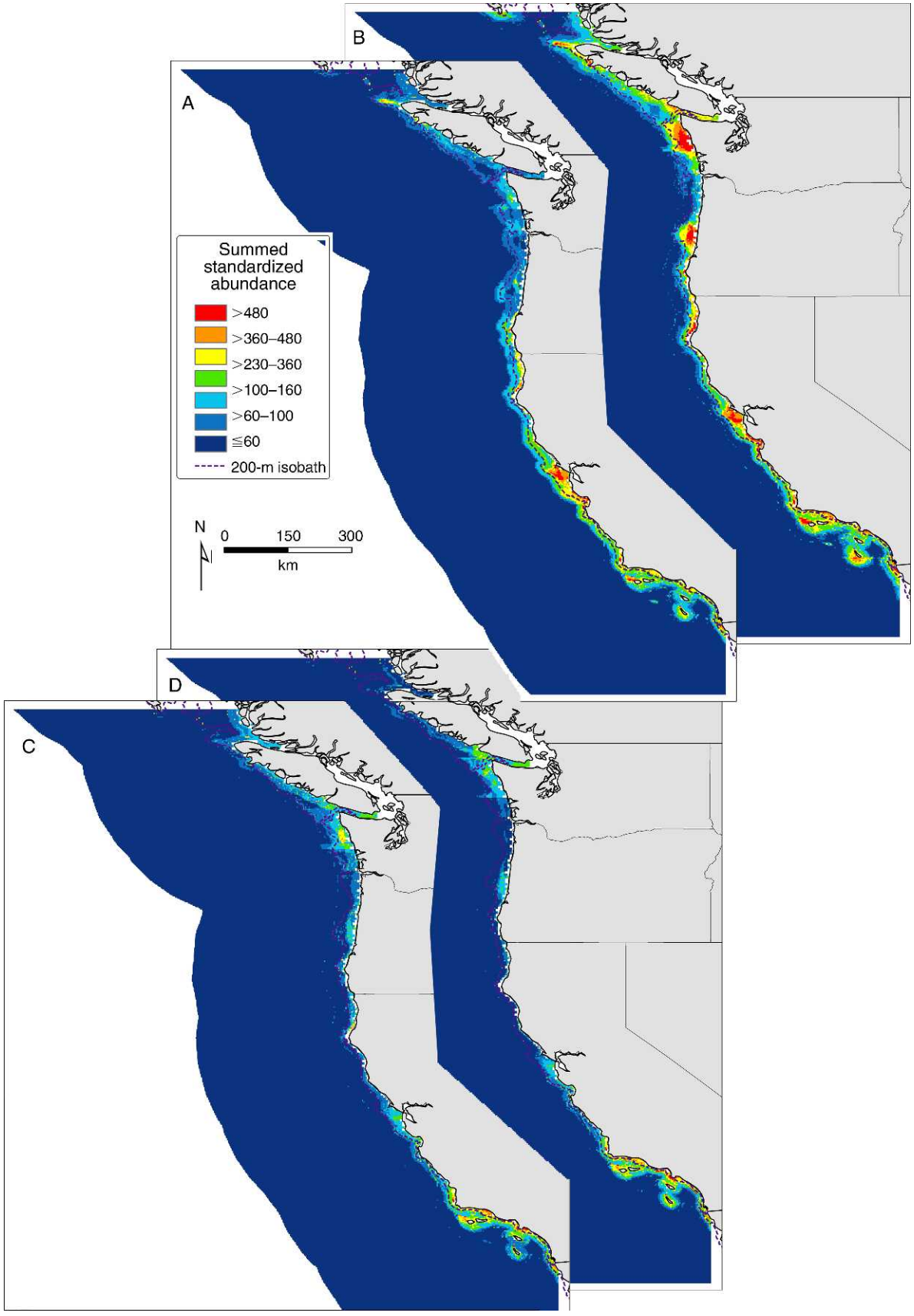


FIG. 4. Predicted (A) importance and (B) persistence for Cassin's Auklet. Importance was calculated over all years and seasons for each species, and a weighted average was then calculated over all species. Cells were sorted according to predicted densities. We identified the smallest set of cells (i.e., cells with the highest predicted density) that together constituted 25% of the species' total predicted abundance; these were scored as "2" ("core" cells). The next-highest set of cells that constituted 50% of the species' total predicted abundance, but were not "core" cells, were scored as "1" ("shoulder" cells). All other cells were scored 0 (see *Criterion 2: Important "core" areas for individual species*). Persistence was calculated over all seasons; shown is the number of years (out of 11) that the prediction cell is in the 95th percentile in terms of predicted abundance. The persistence score for a cell was calculated by individual month, as well as over all seasons. Because the persistence score was only calculated within a specified month, it reflected variation among years, but not seasonal fluctuations in abundance of a species.

Weighting all three criteria equally produced a combined-criterion hotspot map for the California Current (Fig. 7). All hotspots identified using the combined criterion were evident, with one or more of the single criteria. We highlight three areas that appeared as hotspots, but that are not currently included in Marine Protected Areas, or only partially so: Heceta Bank, Cape Mendocino, and the southern Channel Islands (Fig. 7).

DISCUSSION

In this study, we have collected and synthesized at-sea observations of seabird aggregations and combined this information with data on physical and biological features of the marine environment to predict the distributions of multiple species of seabirds over the broad reach of the California Current System. Through the use of data collected in a standardized manner and applying a single set of analytical procedures, we were able to combine information from multiple organizations and investigators to cover a large geographic area and draw on 11 years of seabird monitoring. This has provided a robust basis for identifying aggregation hotspots of seabird species, one that emphasizes patterns


A
B
C
D
Summed standardized abundance
>480
>360–480
>230–360
>100–160
>60–100
≦60
200-m isobath
N
0
150
300
km

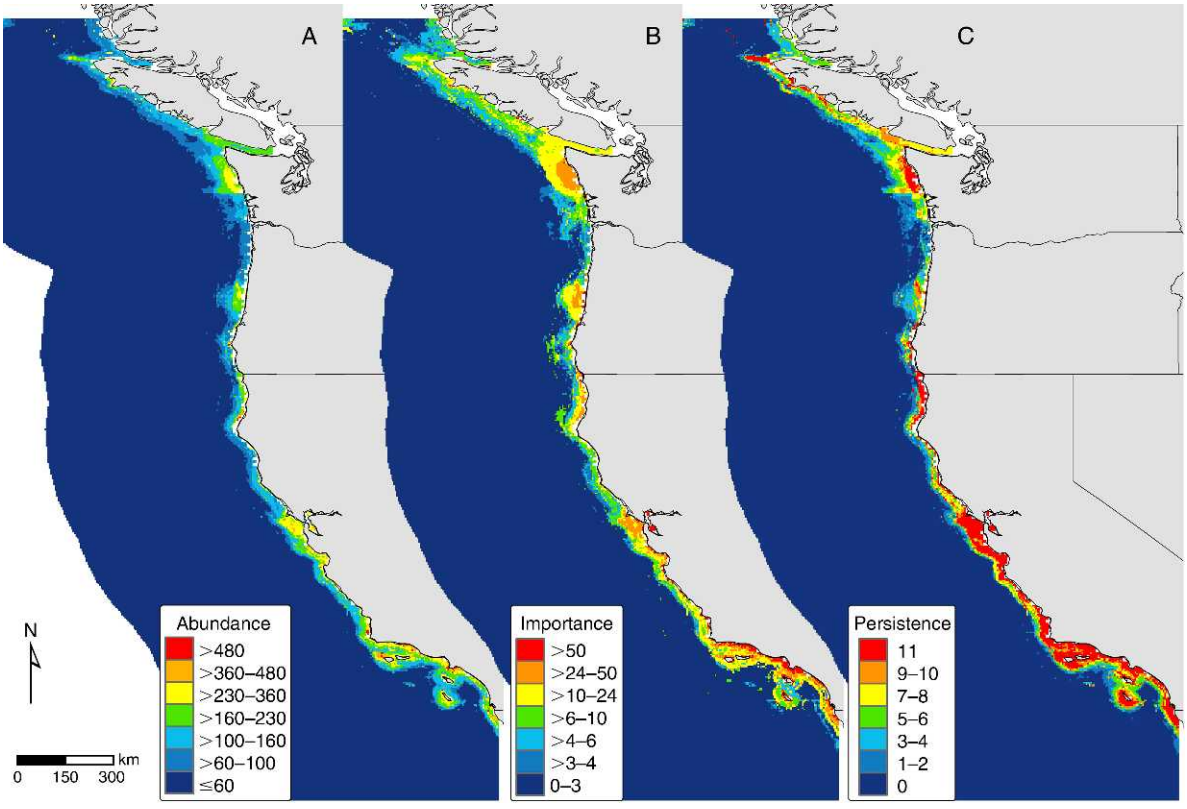


FIG. 6. Hotspot maps for each criterion, calculated over all seasons, for all 16 seabird species. (A) Standardized abundance, averaged over all years, summed over all species. (B) Importance, a weighted average calculated over all years and species (see *Methods: Multispecies hotspot determination: Criterion 2: Important "core" areas for individual species*). (C) Persistence (number of years that the prediction cell is in the 95th percentile in terms of summed standardized abundance).

emerging at a large spatial scale and over a temporal scale that exceeds a decade in duration.

Conducting an analysis at this broad spatial and temporal scale inevitably involves trade-offs. Our information on SSH was obtained at a spatial scale ($28 \times \sim 22$ km cell) that provides only a coarse resolution of oceanic currents and eddies, yet proximity to eddies has been identified as an important predictor in other studies (Ballance et al. 2006, Hyrenbach et al. 2006, Louzao et al. 2006, Ainley et al. 2009). At a fine spatial scale, distributions of seabirds are influenced by a wide array of factors, such as salinity, water-column structure, depth of chlorophyll maximum, fine-scale dynamic height, or prey densities (Ainley et al. 2009, Tremblay et al. 2009), for which data are not available at the scale of the entire CCS.

Our focus in this analysis, however, was not on the fine-scale use of habitats by particular seabird species in localized areas, but on the identification of multispecies foraging aggregations across the entire California Current. This is the scale at which initial decisions are being made regarding priorities for MPA locations and for which information about the likelihood of occurrence of multiple species of interest is critical (Ocean Policy Task Force 2010, Ray 2010).

To identify areas supporting multispecies aggregations of seabirds, we relied on three criteria, reflecting abundance, importance, and persistence of locations. This approach has the advantage of not relying on a single criterion, whose choice may be open to question. Instead, the application of all three criteria provides a robust methodology for combining predictions across species. At the same time, the use of multiple criteria allows the opportunity to identify a hotspot that might be overlooked if one relied upon only a single criterion.

←

FIG. 5. Standardized abundance, by month, summed over all 16 seabird species and averaged over all years for: (A) February, (B) May, (C) July, and (D) October. For each species and cell, mean predicted abundance (estimated number of birds observed per cell, adjusted for survey cell area) was determined (calculated over all years, separately by month). For each species, all mean abundance values were standardized to mean abundance = 0 and SD of abundance = 1. Standardized values were then summed over all species.

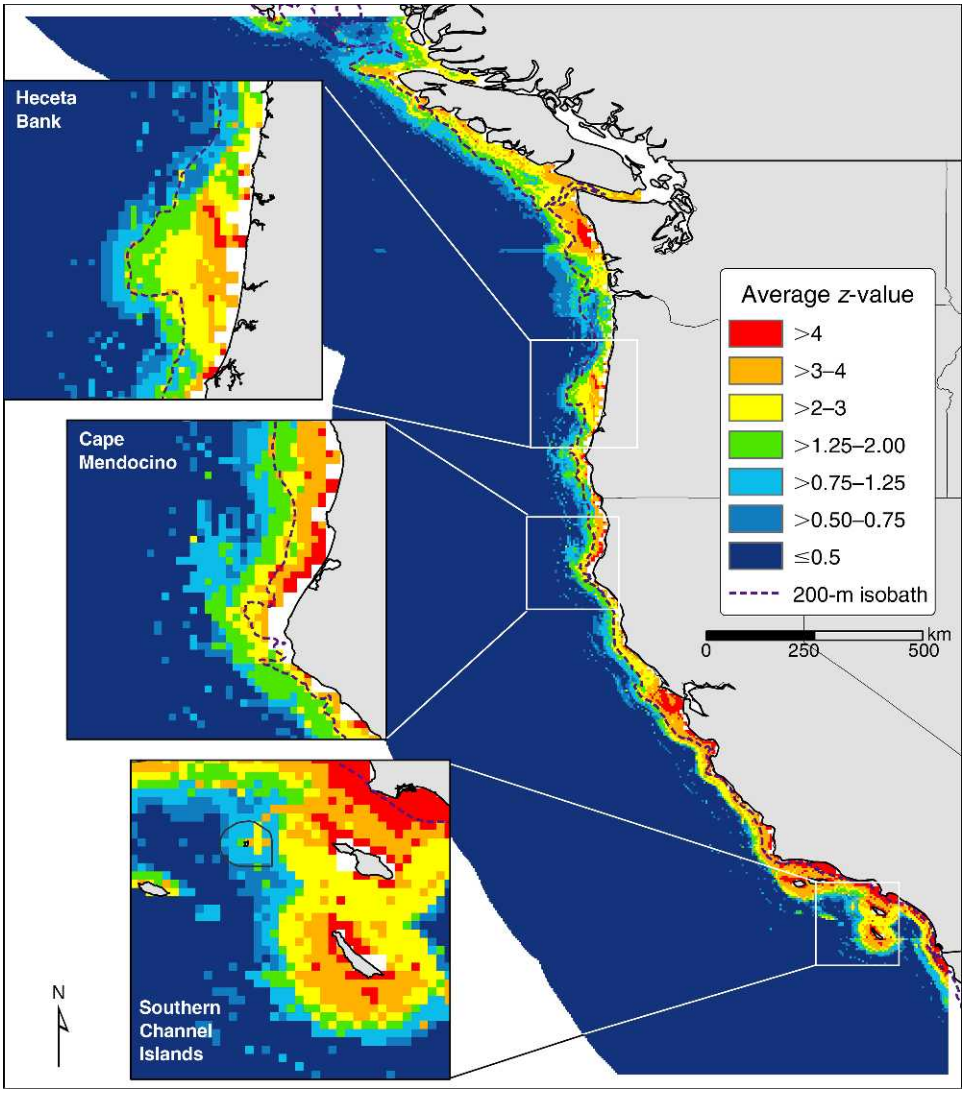


FIG. 7. Hotspot map, combining all criteria. Average *z* scores are shown; the *z* score was calculated for each criterion separately and then averaged (see *Methods: Multispecies hotspot determination; Combining all three criteria*). Insets show hotspot areas that are currently not, or only partially, protected: Heceta Bank, Cape Mendocino region, and southern Channel Islands. The boundary of the Channel Island National Marine Sanctuary is shown in the inset for southern Channel Islands.

Software to identify potential reserve locations (particularly Zonation) has been developed that considers multiple criteria (Moilanen and Wintle 2006, Moilanen 2007). Zonation has been used to model probability of occurrence (e.g., Moilanen 2007), but it can also be implemented using predicted abundances. To implement this program requires specifying importance of connectivity for each species. In contrast, software such as Marxan (Ball and Possingham 2000, Wilson et al. 2005) has an important limitation: it can identify reserve solutions, but it does not provide hierarchic ranking of priority areas, which was one of our goals.

Hyrenbach et al. (2000) identified three types of features to explain aggregations of pelagic predators: static bathymetric features, persistent hydrographic features (e.g., fronts occurring at predictable locations), and ephemeral hydrographic features (cf. Hooker et al. 1999). Our results point to the value of the first class of features in identifying hotspots. At the broad spatial scale of the California Current, we found that seabird foraging hotspots were best predicted by bathymetric features, especially those related to ocean depth and proximity to or distance from land. As a result, the spatial location of predicted hotspots was fairly consistent from season to season and year to year. The consistency of predicted locations supports the value of permanently sited MPAs.

Oceanographic variables, other than SSH, made only modest contributions to the predictive models. However, the rankings of variables in the predictive

models also reflect the spatial scale of the modeling. In explaining variation in abundance of individual species from British Columbia to southern California, it is not surprising that bathymetric or topographic variables predominate over variables such as SST or change in SST, which may contribute to variation in abundance at finer spatial and temporal scales (10 km or less; variation among days or weeks). In fact, SST has often been found to be an important explanatory variable of seabird distribution and abundance in studies conducted at finer spatial scales than that of our analysis (Ainley et al. 2009, Garthe et al. 2009, Tremblay et al. 2009).

Although our focus here is on a multispecies assessment, there was considerable variation among species in the importance of specific predictor variables (Table 3). For example, distance to nearest land was a highly important variable for some species but had low predictive value for other species, such as the Fork-tailed Storm-Petrel. Differences in the importance of predictor variables are not surprising, given the variety of body sizes, food habits, and life-history strategies among the 16 species that we considered. For example, the distribution of some species (e.g., Sooty Shearwaters) may be determined during migration, when pelagic foraging areas are of high value, yet for others (e.g., Common Murres) the critical period may be the breeding season, when foraging areas near breeding colonies are most important (Hooker and Gerber 2004, Lascelles et al. 2009). Information on the location and size of breeding colonies is not yet available for much of the California Current, although other studies (Louzao et al. 2006, Ainley et al. 2009) have incorporated such information into analyses of spatial distribution at a more local scale.

Because species vary in features of their ecology and life history, the hotspots predicted for one species may not be the same as those predicted for other species. The identification of multispecies hotspots as candidate MPAs will therefore depend on which set of species is considered. This, in turn, will reflect management and conservation priorities. If these priorities emphasize a few individual species (e.g., threatened or endangered species), the predicted hotspots will differ from those identified by considering a wide variety of species, as we have done here. MPAs are intended to protect (insofar as possible) entire functioning ecosystems, so a multispecies approach, even one focused on a particular taxon such as seabirds, seems not only appropriate but necessary.

Although current legislation may provide only limited protection for marine wildlife, the National Marine Sanctuary System provides a legal framework that could facilitate development of a network of MPAs. Many of the predicted hotspots that we identified are currently within National Marine Sanctuary boundaries, including the Cordell Bank, Gulf of the Farallones, and Monterey Bay National Marine Sanctuaries (NMS) in central California; Channel Island NMS in southern California; and Olympic Coast NMS in Washington. Other predicted hotspots, however, are not currently protected. In particular, the broader area around the Channel Islands that is not included in the Channel Islands NMS and the coastal area north of Cape Mendocino in northern California to Heceta Bank on the central Oregon coast represent clear gaps in current marine protection. Our analysis also identified two additional gaps farther north. The Swiftsure, Amphetrite, and La Perouse Banks complex immediately north of the Olympic Coast NMS is not protected, nor is the area off the northern end of Vancouver Island. Efforts by CWS/Environment Canada to develop a Marine Wildlife Area (equivalent to an MPA) adjacent to the Scott Islands off the northwest coast of Vancouver Island are currently under way (K. Morgan, *personal communication*).

Human threats to the marine ecosystem may be as great in northern California, Oregon, and Washington as in central and southern California (Halpern et al. 2009). Combined with our results, the analysis of Halpern et al. (2009) indicates that the hotspot region extending from Cape Mendocino to Heceta Bank may represent a high priority for conservation. This is also an area that has not been extensively surveyed for seabirds in comparison to areas farther north in Washington and southern British Columbia or farther south in central and southern California.

The CCS-wide modeling presented here has the advantage of allowing us to compare areas of predicted aggregation in one part of the CCS with areas elsewhere in the CCS, areas that may not have been well studied in the past. For international and national entities, it is useful to compare hotspots in British Columbia with those in California in order to prioritize efforts. Models developed on a finer scale may capture habitat associations that are more relevant to individual species, providing predictions that may be useful in refining the placement of protected areas. Broadscale CCS-wide and fine-scale local modeling efforts can play complementary roles in determining priorities for targeting areas for protection or in identifying areas deserving of systematic surveys to support their designation as MPAs.

Marine ecosystems are dynamic, and are likely to become even more dynamic over the coming decades as climate change and biotic depletion (Hilborn et al. 2003) impact the spatial and temporal distribution of marine hotspots in the CCS. To the extent that hotspots are anchored by fixed bathymetric and geographic features, they may be resilient to such changes. Determining the limits of resiliency may be essential to forward-looking marine spatial planning and making wise investments in MPAs.

Acknowledgments

Financial support was provided by the Resources Legacy Fund Foundation, National Fish and Wildlife Foundation, Moore Family Foundation, Faucett Family Foundation, Hellman Family Foundation, an anonymous PRBO donor,

and Carolyn Johnson and Rick Theis. Numerous agencies have provided long-term support of the at-sea surveys, in particular, NOAA, CWS/EC, and Scripps Institute of Oceanography. Thanks to all those who contributed to data collection and processing for the many at-sea surveys and for keeping the long-term monitoring going, in particular Richard Veit, John McGowan, and Bill Sydeman for their essential roles in data collection. We thank Glenn Ford, Bill Monahan, and Anna Weinstein for discussion. We thank Ben Lascelles and Mark Tasker for incisive review of the manuscript. We thank Daniel Fink (Cornell Lab of Ornithology) for the custom R code for the bagging portion of our analysis. This is PRBO contribution number 1748.

Literature Cited

Ainley, D. G., K. D. Dugger, R. G. Ford, S. D. Pierce, D. C. Reese, R. D. Brodeur, C. D. Tynan, and J. A. Barth. 2009. Association of predators and prey at frontal features in the California Current: competition, facilitation, and co-occurrence. Marine Ecology Progress Series 389:271–294.

Bakun, A., and S. J. Weeks. 2004. Greenhouse gas buildup, sardines, submarine eruptions and the possibility of abrupt degradation of intense marine upwelling ecosystems. Ecology Letters 7:1015–1023.

Ball, I. R., and H. P. Possingham. 2000. Marxan v 1.8.2. Marine reserve design using spatially explicit annealing. Marxan user manual. University of Queensland, Brisbane, Australia. ⟨http://www.uq.edu.au/marxan/docs/marxan_manual_1_8_2.pdf⟩

Ballance, L. T., R. L. Pitman, and P. C. Fiedler. 2006. Oceanographic influences on seabirds and cetaceans of the eastern tropical Pacific: a review. Progress in Oceanography 69:360–390.

Boersma, P. D., and J. K. Parrish. 1999. Limiting abuse: marine protected areas, a limited solution. Ecological Economics 31:287–304.

Breiman, L. 1996. Bagging predictors. Machine Learning 26:123–140.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Wadsworth Publishing, New York, New York, USA.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance sampling. Chapman and Hall, London, UK.

Cañadas, A., R. Sagarminaga, R. De Stephanis, E. Urquiola, and P. S. Hammond. 2005. Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. Aquatic Conservation: Marine and Freshwater Ecosystems 15:495–521.

Clarke, E. D., L. B. Spear, M. L. McCracken, F. F. C. Marques, D. L. Borchers, S. T. Buckland, and D. G. Ainley. 2003. Validating the use of generalized additive models and at-sea surveys to estimate size and temporal trends of seabird populations. Journal of Applied Ecology 40:278–292.

De'ath, G. 2007. Boosted trees for ecological modeling and prediction. Ecology 88:243–251.

Dietterich, T. 2000. An experimental comparison of three methods for constructing ensembles of decision trees: bagging, boosting, and randomization. Machine Learning 40:139–157.

Di Lorenzo, E., N. Schneider, K. M. Cobb, K. Chhak, P. J. S. Franks, A. J. Miller, J. C. McWilliams, S. J. Bograd, H. Arango, E. Curchister, T. M. Powell, and P. Rivere. 2008. North Pacific Gyre Oscillation links ocean climate and ecosystem change. Geophysical Research Letters 35. L08607, [doi:10.1029/2007GL032838]

Diniz-Filho, J. A. F., L. M. Bini, and B. A. Hawkins. 2003. Spatial autocorrelation and red herrings in geographical ecology. Global Ecology and Biogeography 12:53–64.

Douvere, F., and C. Ehler. 2009. New perspectives on sea use management: Initial findings from European experience with marine spatial planning. Journal of Environmental Management 90:77–88.

Durant, J. M., D. O. Hjermann, M. Frederiksen, J. B. Charrassin, Y. Le Maho, P. S. Sabarros, R. J. M. Crawford, and N. C. Stenseth. 2009. Pros and cons of using seabirds as ecological indicators. Climate Research 39:115–129.

Efron, B., and R. Tibshirani. 1993. An introduction to the bootstrap. Chapman and Hall, London, UK.

Ekstrom, J. A. 2008. California Current Large Marine Ecosystem: Publicly available dataset of state and federal laws and regulations. Marine Policy, [doi:10.1016/j.marpol.2008.11.002]

Elith, J., J. R. Leathwick, and T. Hastie. 2008. A working guide to boosted regression trees. Journal of Animal Ecology 77:802–813.

ESRI. 2006. ArcMap 9.2. Environmental Systems Research Institute, Redlands, California, USA.

Ferguson, M. D., J. Barlow, P. Fiedler, S. B. Reilly, and T. Gerrodette. 2006. Spatial models of delphinid (family Delphinidae) encounter rate and group size in the eastern tropical Pacific Ocean. Ecological Modelling 193:645–662.

Ford, R. G., D. G. Ainley, J. L. Casey, C. A. Keiper, L. B. Spear, and L. T. Ballance. 2004. The biogeographic patterns of seabirds in the central portion of the California Current. Marine Ornithology 32:77–96.

Furness, R. W., and C. J. Camphuysen. 1997. Seabirds as monitors of the marine environment. ICES Journal of Marine Science 54:726–737.

Game, E. T., H. S. Grantham, A. J. Hobday, R. L. Pressey, A. T. Lombard, L. E. Beckley, K. Gjerde, R. Bustamante, H. P. Possingham, and A. J. Richardson. 2009. Pelagic protected areas: the missing dimension in ocean conservation. Trends in Ecology and Evolution 24:360–369.

Garthe, S., N. Markones, O. Hüppop, and S. Adler. 2009. Effects of hydrographical and meteorological factors on seasonal seabird abundance in the southern North Sea. Marine Ecology Progress Series 391:243–255.

Grémillet, D., and A. Charmantier. 2010. Shifts in phenotypic plasticity constrain the value of seabirds as ecological indicators of marine ecosystems. Ecological Applications 20:1498–1503.

Guyon, I., and A. Elisseeff. 2003. An introduction to variable and feature selection. Journal of Machine Learning 3:1157–1182.

Halpern, B. S. 2003. The impact of marine reserves: Do reserves work and does reserve size matter? Ecological Applications 13:S117–S137.

Halpern, B. S., C. V. Kapperl, K. A. Selkoe, F. Micheli, C. M. Ebert, C. Kongtis, C. M. Crain, R. G. Martone, C. Shearer, and S. J. Teck. 2009. Mapping cumulative human impacts to California Current marine ecosystems. Conservation Letters 2:138–148.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The elements of statistical learning. Second edition. Springer, New York, New York, USA.

Hilborn, R., T. A. Branch, B. Ernst, A Magnusson, C. V. Minte-Vera, M. D. Scheuerell, and J. L. Valero. 2003. State of the world's fisheries. Annual Review of Environment and Resources 28:359–399.

Hochachka, W. M., R. Caruana, D. Fink, A. Munson, M. Riedewald, D. Sorokina, and S. Kelling. 2007. Data-mining and discovery of pattern and process in ecological systems. Journal of Wildlife Management 71:2427–2437.

Holt, C. A., and N. Mantua. 2009. Defining spring transition: regional indices for the California Current System. Marine Ecology Progress Series 393:285–299.

Hooker, S. K., and L. R. Gerber. 2004. Marine reserves as a tool for ecosystem-based management: the potential importance of megafauna. BioScience 54:27–39.

Hooker, S. K., H. Whitehead, and S. Gowans. 1999. Marine protected area design and the spatial and temporal distribu-

tion of cetaceans in a submarine canyon. Conservation Biology 13:592–602.

Hunt, G. L., F. Mehlum, R. W. Russell, D. Irons, M. B. Decker, and P. H. Becker. 1999. Physical processes, prey abundance, and the foraging ecology of seabirds. Proceedings of the 22nd International Ornithological Congress 22:2040–2056.

Hyrenbach, K. D., P. Fernández, and D. J. Anderson. 2002. Oceanographic habitats of two sympatric North Pacific albatrosses during the breeding season. Marine Ecology Progress Series 233:283–301.

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.

Hyrenbach, K. D., R. R. Veit, H. Weimerskirch, and G. L. Hunt. 2006. Seabird associations with mesoscale eddies: the subtropical Indian Ocean. Marine Ecology Progress Series 324:271–279.

Jahncke, J., L. S. Vlietstra, M. B. Decker, and G. L. Hunt, Jr. 2008. Marine bird abundance around the Pribilof Islands: a multi-year comparison. Deep Sea Research Part II 55:1809–1826.

Lascelles, B., L. Fishpool, J. Croxall, A. Stattersfield, L. Bennun, and M. Aminu-Kano, compilers. 2009. Designing networks of marine protected areas: exploring the linkages between Important Bird Areas and ecologically or biologically significant marine areas. BirdLife International, Cambridge, UK. ⟨http://www.cbd.int/doc/meetings/mar/ewbcsima-01/other/ewbcsima-01-birdlife-02-en.pdf⟩

Leathwick, J., A. Moilanen, M. Francis, J. Elith, P. Taylor, K. Julian, T. Hastie, and C. Duffy. 2008. Design and evaluation of large-scale marine protected areas. Conservation Letters 1:92–101.

Louzao, M., K. D. Hyrenbach, J. M. Arcos, P. Abello, L. G. De Sola, and D. Oro. 2006. Oceanographic habitat of an endangered Mediterranean procellariiform: implications for marine protected areas. Ecological Applications 16:1683–1695.

Moilanen, A. 2007. Landscape zonation, benefit functions and target-based planning: unifying reserve selection strategies. Biological Conservation 134:571–579.

Moilanen, A., and B. A. Wintle. 2006. Uncertainty analysis favours selection of spatially aggregated reserve networks. Biological Conservation 129:427–434.

Norse, E. A., L. B. Crowder, K. Gjerde, K. D. Hyrenbach, C. Roberts, and M. E. Soulé. 2005. The potential for reserves as an ecosystem-based management tool in the open ocean. Pages 302–327 *in* E. A. Norse, and L. B. Crowder, editors. Marine conservation biology. Island Press, Washington, D.C., USA.

Ocean Policy Task Force. 2010. Final recommendations of the Interagency Ocean Policy Task Force, July 19 2010. The White House Council on Environmental Quality, Washington, D.C., USA. ⟨www.whitehouse.gov/files/documents/OPTF_FinalRecs.pdf⟩

O'Hara, P. D., K. H. Morgan, and W. J. Sydeman. 2006. Primary producer and seabird associations with AVHRR-derived sea surface temperatures and gradients in the southeastern Gulf of Alaska. Deep Sea Research II 53:359–369.

Ojeda-Martínez, C., et al. 2009. A conceptual framework for the integral management of marine protected areas. Ocean and Coastal Management 52:89–101.

Piatt, J. F., J. Wetzel, K. Bell, A. R. DeGange, G. R. Balogh, G. S. Drew, T. Geernaert, C. Ladd, and G. V. Byrd. 2006. Predictable hotspots and foraging habitat of the endangered short-tailed albatross (*Phoebastria albatrus*) in the North Pacific: implications for conservation. Deep Sea Research II 53:387–398.

Polldoro, B. A., S. R. Livingstone, K. E. Carpenter, B. Hutchinson, R. B. Mast, N. Pilcher, Y. Sadovy de Micheson, and S. Valenti. 2008. Status of the worlds' marine species. *In* J.-C. Vié, C. Hilton-Taylor, and S. N. Stuart, editors. The 2008 review of the IUCN red list of threatened species. IUCN, Gland, Switzerland. ⟨http://cmsdata.iucn.org/downloads/status_of_the_world_s_marine_species.pdf⟩

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ray, G. C. 2010. Coastal and marine spatial planning: a policy waiting to happen. Aquatic Conservation: Marine and Freshwater Systems 20:363–364.

Spear, L., N. Nur, and D. G. Ainley. 1992. Estimating absolute densities of flying seabirds using analyses of relative movement. Auk 109:385–389.

Suryan, R. M., F. Sato, G. R. Balogh, K. D. Hyrenbach, P. R. Sievert, and K. Ozaki. 2006. Foraging destinations and marine habitat use of short-tailed albatrosses: a multi-scale approach using first-passage time analysis. Deep Sea Research II 53:370–386.

Tasker, M. L., P. Hope Jones, T. Dixon, and B. F. Blake. 1984. Counting seabirds from ships: a review of methods employed and a suggestion for a standardized approach. Auk 101:567–577.

Teck, S. J., et al. 2010. Using expert judgment to estimate marine ecosystem vulnerability in the California Current. Ecological Applications 20:1402–1416.

Tremblay, Y., S. Bertrand, R. W. Henry, M. A. Kappes, D. P. Costa, and S. A. Shaffer. 2009. Analytical approaches to investigating seabird–environment interactions: a review. Marine Ecology Progress Series 391:153–163.

Wilson, K. A., M. I. Westphal, H. P. Possingham, and J. Elith. 2005. Sensitivity of conservation planning to different approaches to using predicted species distribution data. Biological Conservation 122:99–112.

Worm, B., H. K. Lotze, and R. A. Myers. 2003. Predator diversity hotspots in the blue ocean. Proceedings of the National Academy of Sciences USA 100:9884–9888.

Yen, P. P. W., W. J. Sydeman, S. J. Bograd, and K. D. Hyrenbach. 2006. Spring-time distributions of migratory marine birds in the southern California Current: oceanic eddy associations and coastal habitat hotspots over 17 years. Deep Sea Research Part II 53:399–418.

Yen, P. P. W., W. J. Sydeman, and K. D. Hyrenbach. 2004. Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. Journal of Marine Systems 50:79–99.

Yen, P. P. W., W. J. Sydeman, K. H. Morgan, and F. A. Whitney. 2005. Top predator distribution and abundance across the eastern Gulf of Alaska: temporal variability and ocean habitat associations. Deep Sea Research II 52:799–822.

APPENDIX A

Survey coverage (*Ecological Archives* A021-101-A1).

APPENDIX B

Summarization of seabird observations (*Ecological Archives* A021-101-A2).