

## RESEARCH ARTICLE

# An ensemble approach to species distribution modelling reconciles systematic differences in estimates of habitat utilization and range area

J. Harris<sup>1</sup>  | J. L. Pirtle<sup>2</sup>  | E. A. Laman<sup>3</sup>  | M. C. Siple<sup>3</sup>  | James T. Thorson<sup>4</sup> 

<sup>1</sup>Washington Department of Fish and Wildlife, Olympia, Washington, USA

<sup>2</sup>Habitat Conservation Division, Alaska Regional Office, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Juneau, Alaska, USA

<sup>3</sup>Resource Assessment and Conservation Engineering Division, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

<sup>4</sup>Resource Ecology and Fisheries Management, Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, Seattle, Washington, USA

**Correspondence**

J. Harris

Email: [jeremy.harris@dfw.wa.gov](mailto:jeremy.harris@dfw.wa.gov)**Funding information**

NMFS Alaska Region Essential Fish Habitat Research Plan

**Handling Editor:** Verena Trenkel**Abstract**

1. Species distribution models (SDMs) are an important tool for conservation and resource management. However, managers are often interested in derived quantities such as range or area occupied, and how these are calculated can have a large impact.
2. Ecosystem-based management typically requires spatial information about species distributions, which is increasingly generated from SDMs that are then processed to identify occupied habitat. Many types of SDMs exist, but there is little research regarding how this model-choice affects outcomes when defining occupied habitat, in part because these models generate different types of output.
3. We fit a suite of five SDMs to data for 208 species/life stage combinations in three marine ecosystems while ensuring that they all estimate a 'common currency' of numerical abundance. We then calculate out-of-sample predictive performance to weight these constituents in an ensemble SDM.
4. Results show that this approach can reduce bias arising from a priori specification of individual SDMs resulting in a better fit to survey data (constituent SDMs had a median of 7% higher RMSE). The SDMs had a range of responses relative to the ensemble, with MaxEnt typically predicting a median 1.3% higher area occupied, and negative-binomial GAMs predicting 21.4% lower area occupied.
5. Two potential methods of identifying the area of occupied habitat from SDM outputs are compared—probability-based and cumulative density-based methods. We find that cumulative densities result in smaller estimates of area occupied, and we recommend careful consideration of how model-choice affects occupied-habitat estimates in spatial management.
6. *Policy implications:* Finally, we discuss how the patterns identified during the 5-year Review of Essential Fish Habitat for Alaska should be carefully considered by managers using SDMs to identify habitat that may be impacted by anthropogenic activities.

**KEYWORDS**

ecosystem-based management, ensemble models, habitat occupancy, species distribution modelling

## 1 | INTRODUCTION

Species distributions and abundances are shifting worldwide, and quantitative tools like species distribution models (SDMs) can help characterize these shifts so that ecological information can be used in decision-making. Practitioners in several disciplines use SDMs to integrate ecological information within managerial frameworks, including fisheries management, spatial planning, endangered species conservation, invasive species management, climate impact assessment and the estimation of essential biodiversity indicators (Engler et al., 2004; Jetz et al., 2019; Sundblad et al., 2011). Numerous SDMs have been developed to address diverse data types, statistical situations and applications (Guisan & Zimmermann, 2000). The structure of these models ranges widely, from machine learning approaches like boosted regression trees (Elith et al., 2008) to generalized additive models (GAMs; Guisan et al., 2002) and autoregressive state-space models (Anderson et al., 2022; Thorson, 2019). As the options for SDMs expand, there have been calls for additional research examining the consequences of SDM selection for prediction and decision-making (Brodie et al., 2020). A number of papers have compared multiple SDM types based on presence-only data (Valavi et al., 2022) or abundance data (Waldock et al., 2022) and found that ensembles of tuned SDMs or flexible machine learning methods like random forest tend to perform well in terms of predictive power. Practitioners' choices of data type, model structure and outputs impact the advice provided for management, as recognized by frameworks like the Overview, Data, Model, Assessment and Prediction (ODMAP) protocol (Zurell et al., 2020). While the downstream impacts of these modelling decisions have been quantified for some management scenarios using simulations (e.g. Guillera-Aroita et al., 2015), they have not been examined extensively in real-life management applications.

SDMs may be fit to a variety of data such as presence-only, presence-absence and counts. However, a 'common currency' is needed to efficiently compare predictions from models that utilize different underlying data. Cross-validation or other resampling methods can then be used to estimate uncertainty and select a model that fits the data well or provides the greatest predictive ability. Skill testing can also be used to assign weights to each candidate SDM and combine their predictions to produce an ensemble. Model ensembles can help mitigate bias in individual models and incorporate uncertainty about model specification (Dormann et al., 2018). Ensemble methods are widely used and have been shown to produce superior results in a wide variety of SDM applications (Abrahms et al., 2019; Andersen et al., 2019), though well-tuned individual SDMs can sometimes outperform an ensemble (Hao et al., 2020).

Once probability or relative density (hereafter 'density') predictions have been produced by SDMs, they must be post-processed to identify areas where a species is present, often defined as having a probability that the location is suitable habitat, or that

encounter probability is above a given threshold (e.g. Holbrook et al., 2000; Pettorelli et al., 2010). In many cases, the spatial dimension of the predictions may be flattened into a single estimate of total area occupied or total abundance (IUCN, 2012). The method used to determine the threshold for a species presence can have a greater effect on the size of the predicted habitat than the choice of model or data type (De Cubber et al., 2023). Though widely practiced, converting the continuous output of an SDM to binary or categorical output flattens the data and may result in biased inferences in simulations (Guillera-Aroita et al., 2015). The need to post-process model predictions introduces another decision point that may impose further trade-offs and complicate the interpretation of any results.

Applied cases have been recommended as a way to improve decision tools and bridge theory and practice in SDM development (Guisan et al., 2013). In the United States, the Magnuson-Stevens Fishery Conservation and Management Act (1996) requires that management plans account for essential fish habitat (EFH), defined as those waters and substrate necessary to fish for spawning, breeding, feeding or growth to maturity. In most management areas, the EFH process involves generating maps of occupied habitat from one or more SDMs, which are then combined with information about fishing effort and other coastal activities to guide management decisions (e.g. Laman et al., 2018). Thus, EFH serves as a useful case study for how to generate and process SDM output for management applications.

Here, we examine the performance of five different SDMs (maximum entropy (MaxEnt) and four GAM configurations) representing a variety of common SDM types that operate under different data constraints and assumptions, using data from Alaska as a case study. While other approaches such as machine learning methods (e.g. boosted regression trees and random forests) can be used as SDMs, we defined this set of constituent models by augmenting the set of models that have been used individually (and without an ensemble) to designate Essential Fish Habitat by the North Pacific Fisheries Management Council (NPFMC) in 2017 (Laman et al., 2018). Progressively improving this SDM process follows recommendations regarding Translational Ecology (Enquist et al., 2017) wherein management-oriented science (such as our ensemble-SDM application for EFH) must result from co-produced developments between scientists, managers and stakeholders, rather than basing scientific decisions only on technical considerations. We convert the outputs of the different SDMs into a common currency of abundance that allows for models to be compared based on shared measures of predictive performance. We also compare the predictions of the individual SDMs to the performance of a weighted ensemble derived from a combination of all five SDMs. We compare two methods for simplifying SDM predictions into estimates of area occupied. Finally, we share insights from the model development and communication process for practitioners seeking to implement similar ensemble approaches for natural resource management.

## 2 | MATERIALS AND METHODS

### 2.1 | Surveys

We used bottom trawl survey data for groundfish and invertebrate species from three management regions in Alaska (the Gulf of Alaska [GOA]; Aleutian Islands [AI]; and northern and eastern Bering Sea [EBS]; Markowitz et al., 2022; von Szalay & Raring, 2018, 2020; Figure 1). Bottom trawl surveys are conducted in the summer annually in the EBS and biennially or triennially in the GOA and AI by the Groundfish Assessment Program of NOAA's Alaska Fisheries Science Center. These surveys use standardized fishing protocols to document the distribution and abundance of fish and invertebrate species. In all surveys, this study uses the original count of individuals caught, and uses the total area swept by the trawl net to describe effort (Alverson & Pereyra, 1969), which is used as an offset in the models. Fish lengths collected during the survey were used to partition catch into life stages based on literature values, as in Harris et al. (2023). Animal samples used in this study were provided by the National Oceanographic and Atmospheric Administration, Alaska Fisheries Science Center, Groundfish Assessment Program in accordance with the National Marine Fisheries Service Animal Care and Use Policy 04-112 (<https://www.fisheries.noaa.gov/national/laws-and-policies/science-and-technology-policy-directives>).

#### 2.1.1 | Environmental covariates

We included environmental covariates that we expected to influence species and life stage densities (Table S1). These covariates included habitat attributes observed on the bottom trawl survey (e.g. bottom temperature), modelled metrics describing the benthic environment (e.g. bottom depth, slope) and water movement (bottom currents), and structure forming invertebrates (sponge presence). All covariates were assessed for collinearity in each region using variance inflation factors (Zuur et al., 2009), and all covariates presented here scored below five by this metric, indicating low or no collinearity. We developed 1 km<sup>2</sup> scale environmental conditions for early juvenile, subadult and adult life stages. Detailed methods describing the development and processing of the covariate rasters are available in Harris et al. (2022) and the Supplemental Materials.

#### 2.2 | Species distribution models

We created SDMs for 208 species/life stage combinations for groundfish and invertebrates. When published life history studies were available, we apportioned trawl catches into life stages (early juvenile, subadult and adult). This was accomplished by computing the proportional contribution of each stage in a random subsample for fish lengths in that trawl and extending that proportion to each species' total catch. When life history information was not available,

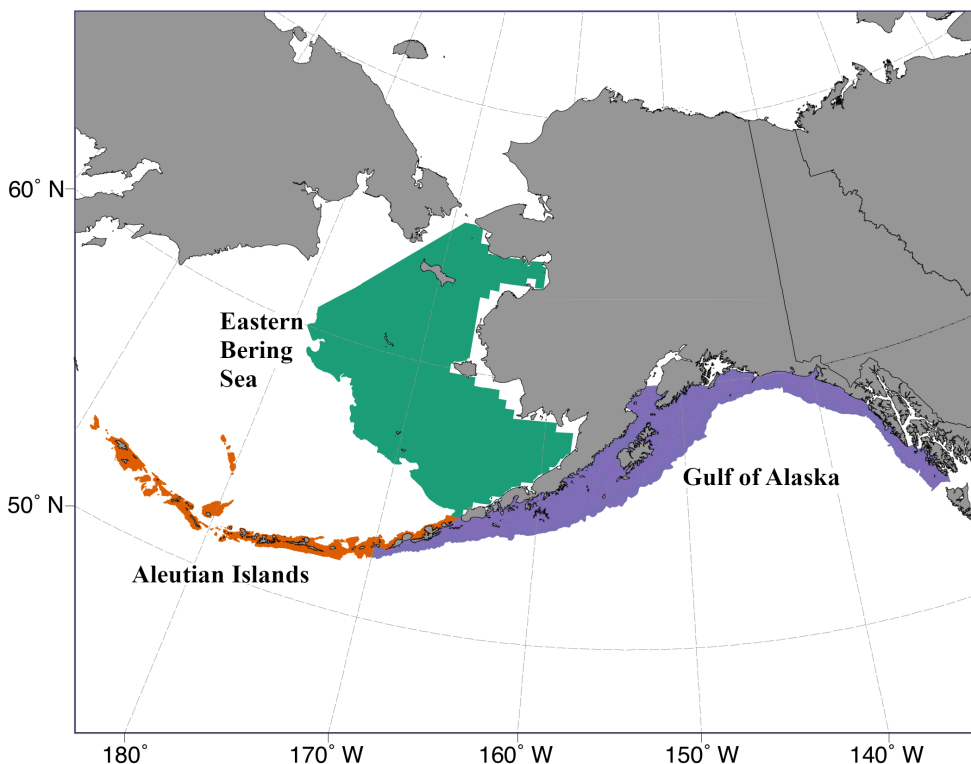


FIGURE 1 Map of survey areas for each of the three regions in this study, showing the Aleutian Islands in orange, Eastern Bering Sea in green and Gulf of Alaska in purple.

SDMs were calculated for all lengths. Constituent SDMs and the SDM ensemble were compared to each other based on model fits and estimates of area occupied, for all the species/lifestage combinations above (Figure 2).

The goal of this paper is to assess the results of many SDMs produced as part of a large regulatory process, so we focus on model performance rather than improving a single species map or SDM. For direct maps of density, encounter probability or prediction uncertainty for individual species, we refer the reader to technical memoranda (Harris et al., 2022; Laman et al., 2022; Pirtle et al., 2023). There has been a significant push to ease the interpretation of SDMs by using a standardized reporting format such as ODMAP (Fitzpatrick et al., 2021; Zurell et al., 2020). A representative example presenting

these maps and other information using the ODMAP reporting format is available in Table S3.

A total of six methods were used to model species distributions: a maximum entropy model (MaxEnt), a binomial presence/absence generalized additive model (paGAM), a hurdle GAM (hGAM), a Poisson GAM ( $GAM_p$ ), a negative-binomial GAM ( $GAM_{nb}$ ) and an ensemble consisting of a weighted average these models (Figure 2). MaxEnt and GAMs are commonly used statistical models for characterizing species distributions in marine environments (Melo-Merino et al., 2020). These models vary in their complexity and in their underlying assumptions about density, and they were chosen for their flexibility with respect to the survey data. They were also selected for their simplicity relative to other approaches, as model structure

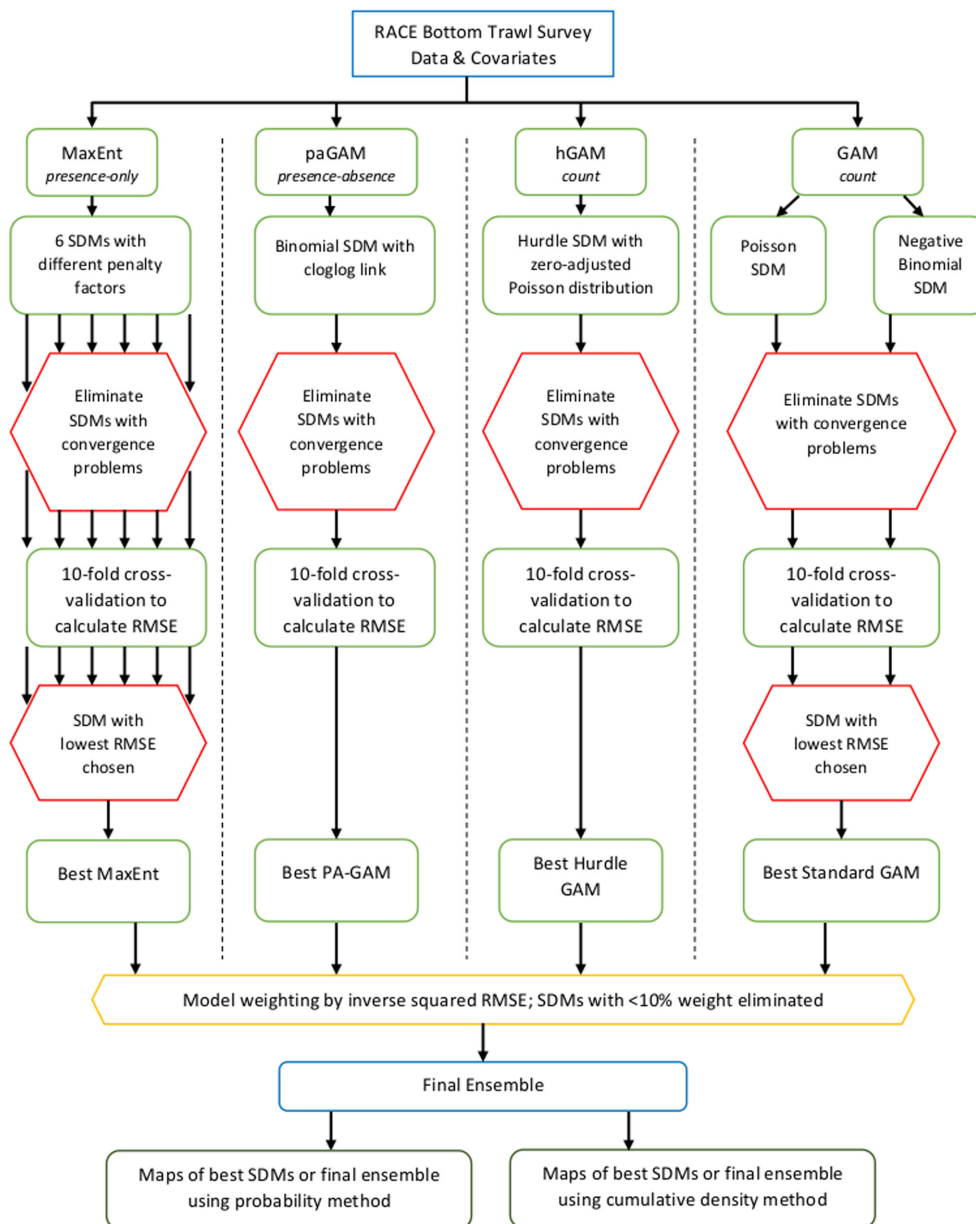


FIGURE 2 Flow chart showing the process for constructing SDM ensembles, with data types in italics. Models that did not converge or were otherwise not selected for the ensemble are eliminated by setting their weight equal to zero.

and results had to be communicated to diverse stakeholders. To facilitate comparisons between models, we use numerical abundance as a 'common currency' for all model predictions. This allowed for direct skill testing of predictions made from different methods and provided a simple avenue for combining multiple predictions into an ensemble.

### 2.2.1 | Maximum entropy models

Maximum entropy (MaxEnt) models use presence data to estimate habitat suitability by finding the probability distribution of maximum entropy, subject to constraints that are based on environmental conditions and information about the target distribution (Phillips et al., 2006). This information consists of a set of 'features', and the constraints are the empirical average of a set of sample points taken from the target distribution. Using the newer implementation in the *maxnet* R package (Phillips et al., 2017), MaxEnt models may be formulated as the result of an inhomogeneous Poisson process using a cloglog link, and so the latent Poisson abundance can be approximated from the model outputs. In practice, this means taking the linear predictor from this model and applying the inverse-log function to predict numerical density. This approximate abundance is then re-scaled so that the mean of the predictions is equal to the mean of observed abundance. While the survey data support more complex presence-absence models, we include MaxEnt because it is a popular tool for identifying presence when data are opportunistic or low quality.

### 2.2.2 | Generalized additive models

We used four types of GAMs in this study: the cloglog-linked presence/absence GAM (paGAM), a cloglog-linked presence/absence GAM combined with a log-linked Poisson-distributed GAM (hGAM; Barry & Welsh, 2002; Potts & Elith, 2006), a log-linked GAM with a Poisson distribution (GAM<sub>p</sub>; Hastie & Tibshirani, 1990); and a log-linked negative-binomial GAM (GAM<sub>nb</sub>; Zuur et al., 2009). The paGAM estimates abundance indirectly, using the cloglog-linked probability to approximate a latent Poisson distribution. The hGAM, GAM<sub>pp</sub> and GAM<sub>nb</sub> each estimate abundance directly, and do not require any

approximation. The abundance predictions from each of the GAMs were scaled so that the mean of the predictions is equal to the mean of observed abundance. Simpler models such as the GAM<sub>p</sub> and paGAM tended to converge more often than complex models like the hGAM. The GAM<sub>p</sub> outperformed the GAM<sub>nb</sub> in 135 out of 208 (65%) cases, though the difference between the two was often small. Additional information on each model is available in the [Supplemental Materials](#).

### 2.2.3 | Model convergence and checks

We employed a consistent framework for selecting and validating models across all species and life stages. Any SDM that failed to converge on stable parameter estimates was discarded (Figure 2). Furthermore, our inference involves extrapolating density to areas beyond the range of fitted covariates, and this extrapolation can be highly inaccurate when covariate responses are nonstationary (Rollinson et al., 2021). We guard against nonstationary responses by eliminating models that provide implausible estimates of density; that is, an SDM was also discarded if it produced any density prediction that was greater than 10 times the highest recorded trawl catch. Other methods can be used to identify the extent of extrapolation in SDMs (Guillaumot et al., 2020), but the 10 times threshold was sufficient to prevent implausible predictions in less sampled areas of the prediction grid from being included in the ensemble. In 189 out of 208 species/life stages all five constituent SDMs converged on stable parameter estimates (Table 1). In 31 out of 208 cases, one or more SDMs failed to converge on a stable estimate, and the hGAM was responsible for 30 out of the 31 failures. The hGAM was also the most prone to predicting implausibly high abundance and was eliminated for a further 36 species/life stages. The MaxEnt was the second most likely to encounter problems with extreme predictions while the paGAM rarely suffered from this particular error.

### 2.2.4 | Model fits

Model fit was assessed by conducting 10-fold cross-validation on all SDMs to estimate RMSE and to assess accuracy and uncertainty

**TABLE 1** Summary of model convergence and other checks. One model was produced for each species/life stage (208 total). SDMs were included in the ensemble only if they passed all checks and had a weight of greater than 0.10 (Figure 2). The final column shows the average weight across all ensemble models, including cases where the weight was set to zero.

Model	Passed convergence check (i.e. algorithm produced stable parameter estimates)	Passed plausibility check (i.e. no prediction greater than 10x max observation)	Included in ensemble (i.e. greater than zero weight)	Average weight in ensemble
MaxEnt	207	161	154	0.20
paGAM	208	204	203	0.31
hGAM	178	142	125	0.17
GAM <sub>p</sub>	208	173	135	0.21
GAM <sub>nb</sub>	208	193	68	0.11

(Figure 2). We used simple-random sampling to partition the data, given that our focus of inference is on predictive performance within the fixed spatial domain (Roberts et al., 2017). When multiple versions of similar models were fit (i.e. MaxEnt with different penalty terms; GAM<sub>p</sub> vs. GAM<sub>nb</sub>), only the version with the best RMSE was retained.

### 2.3 | Ensemble models

An ensemble was constructed as the weighted average of five SDMs (MaxEnt and four GAM-type SDMs), weighted by the inverse squared RMSE (Figure 2), as follows:

$$w_i = \frac{RMSE_i^{-2}}{\sum_{i=1}^m RMSE_i^{-2}}$$

where  $w_i$  is the weight for model  $i$ ,  $RMSE_i$  is the cross-validated RMSE for model  $i$ , and  $m$  is the number of constituent models. The inverse of RMSE-squared is sometimes called 'precision', and precision-weighting is often the optimal weighting method (Gelman, 2014). Additionally, RMSE allows the weights to be based on out-of-sample model performance, which better represents actual prediction uncertainty. While the use of information theoretic weights (i.e. Akaike's information criterion) is widespread, Dormann et al. (2018) caution that the statistical justification for this method is disputed and that cross-validation based weights may be more reliable. The inclusion of poorly performing models may degrade ensemble performance so if any constituent SDM received less than a 10% relative weight, it was eliminated from the ensemble and the weights of the remaining SDMs in the ensemble were recalculated (Figure 2).

### 2.4 | Using SDMs to estimate area occupied

SDM predictions are often used to make inferences about locations of important habitat for a species (e.g. EFH). Total area occupied is commonly used in conservation, and provides a one-dimensional metric of status and/or risk (IUCN, 2012). Because our study compares models with different data types, it is also a good occasion to compare two different potential methods of calculating area occupied, which we refer to as the 'probability' method and the 'cumulative' method.

In the 'probability' method, the area occupied was defined as all locations with an encounter probability greater than 5%. Therefore, the total area is given as:

$$\text{Area} = \sum_{j=1}^n s_j \left\{ \begin{array}{l} s_j = 1 \text{ if } p_j \geq .05 \\ s_j = 0 \text{ if } p_j < .05 \end{array} \right\}$$

where  $s_j$  is the  $j$ th 1km<sup>2</sup> grid location,  $p_j$  is the predicted encounter probability at location  $j$ , and  $n$  is the total number of grid locations.

Because we employed models with different probability distributions and also needed to incorporate the effects of scaling, we approximated the encounter probability for all SDMs and the ensemble as the probability under a Poisson distribution of observing one or more fish given the predicted average trawl abundance. This approximation allows the encounter probability to be calculated in a consistent manner across all the SDMs used in this study and is similar to the methods used to identify EFH in Laman et al. (2018).

In the 'cumulative' method, occupied habitat is defined as all locations with less than 95% cumulative density. All locations are ranked by decreasing estimated density and the cumulative quantile of total density for each location is calculated. The total area is given as:

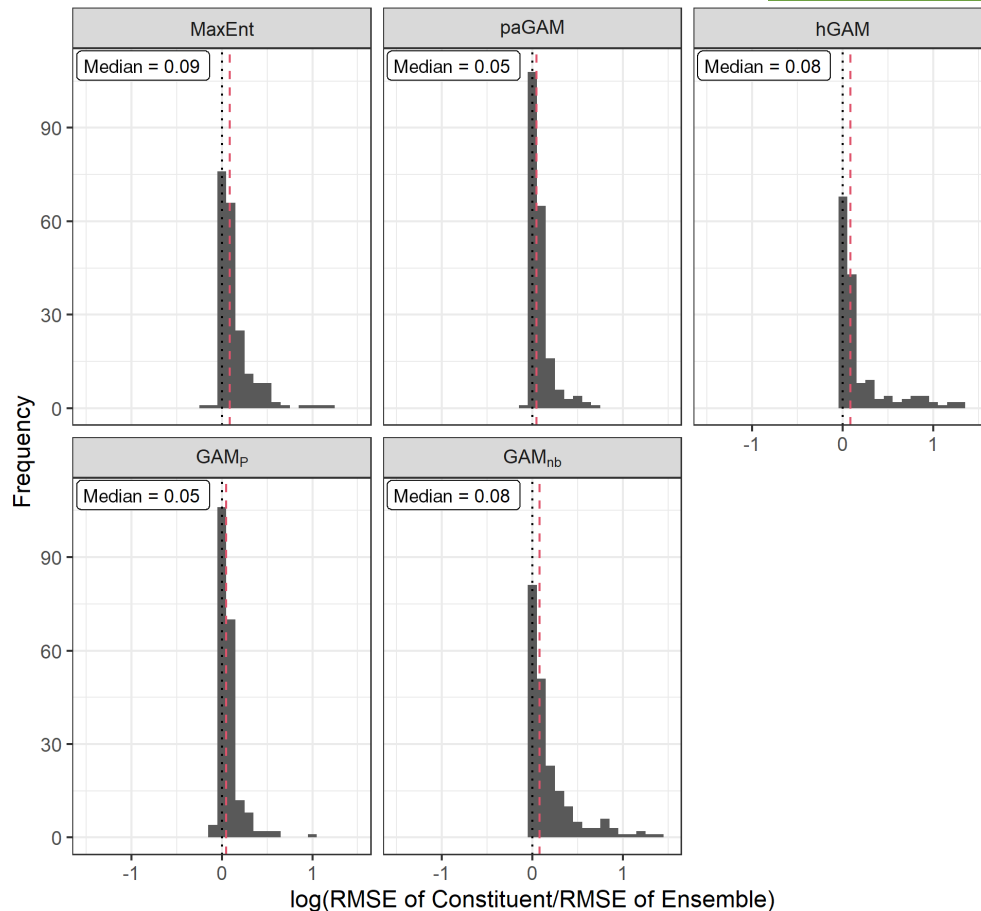
$$\text{Area} = \sum_{j=1}^n s_j \left\{ \begin{array}{l} s_j = 1 \text{ if } \frac{\text{cumulative sum (rank ordered } (d_j))}{\text{sum}(d)} \leq .95 \\ s_j = 0 \text{ if } \frac{\text{cumulative sum (rank ordered } (d_j))}{\text{sum}(d)} > .95 \end{array} \right\}$$

where  $s_j$  is the  $j$ th 1km<sup>2</sup> grid location and  $d_j$  is the predicted density at location  $j$ . The cumulative quantile function is broken into its parts, a cumulative sum of density ordered by descending rank (i.e. highest first), divided by the total sum of all predicted densities. The set of locations ranked below 95% can then be interpreted as the smallest area within which 95% of the population is estimated to occur. This is similar to the methods used by Echave et al. (2012) to describe EFH for Pacific salmon in Alaska. We compare the median area occupied estimates between these two methods based on ensemble predictions.

## 3 | RESULTS

### 3.1 | Model fits

Ensemble models consistently achieved the best fit based on RMSE and were never substantially worse than the best-performing constituent SDM (Figure 3). Overall, the differences between the ensemble and each of the constituent models tended to be small: the GAM<sub>p</sub> had a median of 5% higher RMSE, and MaxEnt had a 9% higher RMSE. MaxEnt, GAM<sub>nb</sub> and hGAM had the largest median difference in RMSE relative to the ensemble, though the models showed largely similar performance overall. The improved fit of the ensemble over the constituent models was consistent across taxonomic groups (Table 2). A full spreadsheet of fit metrics and other information for all 208 species/lifestage combinations is available in the Supplemental Materials. The best constituent model varied by group, though the paGAM and GAM<sub>p</sub> tended to perform well across many situations. In flatfishes, the paGAM performed well, but the GAM<sub>p</sub> demonstrated better model fits for 'roundfishes', which includes many commercially important species. Rockfishes had similar performance using multiple approaches, and the paGAM and GAM<sub>p</sub> performed slightly better than others in chondrichthyan fishes. All the models other than the hGAM



**FIGURE 3** Frequency distribution of log-transformed root-mean-square-error (RMSE) scores across species/life histories from the five constituent models and relative to the RMSE of the ensemble. The ensemble score is indicated by the black dotted line at zero. The red dashed line shows the median score for that model. Values greater than zero mean that the ensemble model had a lower RMSE (better performance) than the constituent.

showed similar performance for the crabs and octopus category. The more complicated hGAM and  $GAM_{nb}$  tended to have higher RMSE than the other models on average. The hGAM was the most likely to fail to converge and hence be excluded from the ensemble, whereas other models appeared to be more stable both in terms of convergence and for producing plausible abundance estimates (Table 1).

### 3.2 | Comparison of SDMs for estimating area occupied

According to the probability method, constituent models differed in their tendency to predict larger or smaller areas occupied in this study. The MaxEnt tended to estimate larger areas of occupied habitat (median 1.2% increase) and the paGAM showed no difference (<1% median difference), whereas the hGAM,  $GAM_{pp}$  and  $GAM_{nb}$  tended towards smaller predicted areas (14%, 15% and 21% median decrease relative to the ensemble, respectively), and hGAMs with poor fit tended to predict very small areas compared to other constituent models. Representative examples of this tendency are shown in the hGAMs for subadult Pacific ocean perch and southern Tanner crab,

which had high RMSEs and were given zero weight in the ensemble (Figure 4). While the ensemble was an average of density models, the area occupied in these examples is calculated from the encounter probability (probability method) and it tended to predict larger occupied habitat areas similar in size to those of the MaxEnt and paGAM. The maps of subadult Pacific ocean perch and tanner crab in the EBS are examples where the ensemble heavily weighted the larger probability predictions from the MaxEnt and paGAM, compared to the more spatially restricted estimates from the other models (Figure 4). The maps of adult rex sole show that this is not always the case; here the ensemble predicts a smaller area occupied than predicted by MaxEnt and similar to the remaining constituent models.

### 3.3 | Comparison of area occupied definitions on predicted habitat

Results also demonstrate that the method ('probability' vs. 'cumulative') used to identify occupied habitat influences the estimate of area occupied. Identifying occupied habitat based on the 'probability' method produced larger estimates of occupied habitat areas

**TABLE 2** Summary of model fit metrics for each species group across regions. The performance of each SDM is summarized as the median RMSE across all species and life stages in its group. Only species/life stages where all five constituent models converged successfully were included (hence the smaller total than in Table 1).

	Model	RMSE		Model	RMSE
Flatfish (Order Pleuronectiformes) N=63	MaxEnt	14.70	Sharks & Skates (Class Chondrichthyes) N=30	MaxEnt	14.70
	paGAM	14.03		paGAM	14.03
	hGAM	21.72		hGAM	21.72
	GAM <sub>p</sub>	14.92		GAM <sub>p</sub>	14.92
	GAM <sub>nb</sub>	16.98		GAM <sub>nb</sub>	16.98
	ensemble	11.30		ensemble	11.30
Roundfish (Families Gadidae, Hexagrammidae, Anaploomatidae) N=28	MaxEnt	58.09	Crabs & Octopus, N=8	MaxEnt	58.09
	paGAM	58.02		paGAM	58.02
	hGAM	88.39		hGAM	88.39
	GAM <sub>p</sub>	57.95		GAM <sub>p</sub>	57.95
	GAM <sub>nb</sub>	58.05		GAM <sub>nb</sub>	58.05
	ensemble	56.57		ensemble	56.57
Rockfish (Family Scorpaenidae) N=48	MaxEnt	11.37	Total, N=177	MaxEnt	11.37
	paGAM	11.25		paGAM	11.25
	hGAM	19.74		hGAM	19.74
	GAM <sub>p</sub>	11.12		GAM <sub>p</sub>	11.12
	GAM <sub>nb</sub>	13.50		GAM <sub>nb</sub>	13.50
	ensemble	9.55		ensemble	9.55

than the 'cumulative' method in most cases (Figure 5). Overall, the probability method resulted in a median of 58% larger estimates of area occupied, with some exceptions (e.g. subadult Dover sole; Figure 5). This pattern varied among species and life stages; 14% of species/life stages had a larger predicted occupied habitat when the cumulative method was used.

The pattern of larger predicted occupied habitat with the probability method was consistent across the three major regions in this study and applied to a wide variety of species and life stages. However, the 14% of cases where the occupied habitat was higher using the cumulative method were predominantly species/life stages with low density in trawl catches (skates, and some less common rockfish and flatfish species). For example, subadult Dover sole are infrequently encountered in the EBS surveys and are rarely found in large numbers. The probability method estimated that the occupied habitat for subadult Dover sole consisted of a narrow band along the continental slope, whereas the cumulative method includes more of the continental slope and a larger area in the south and east of the Bering Sea (Figure 6).

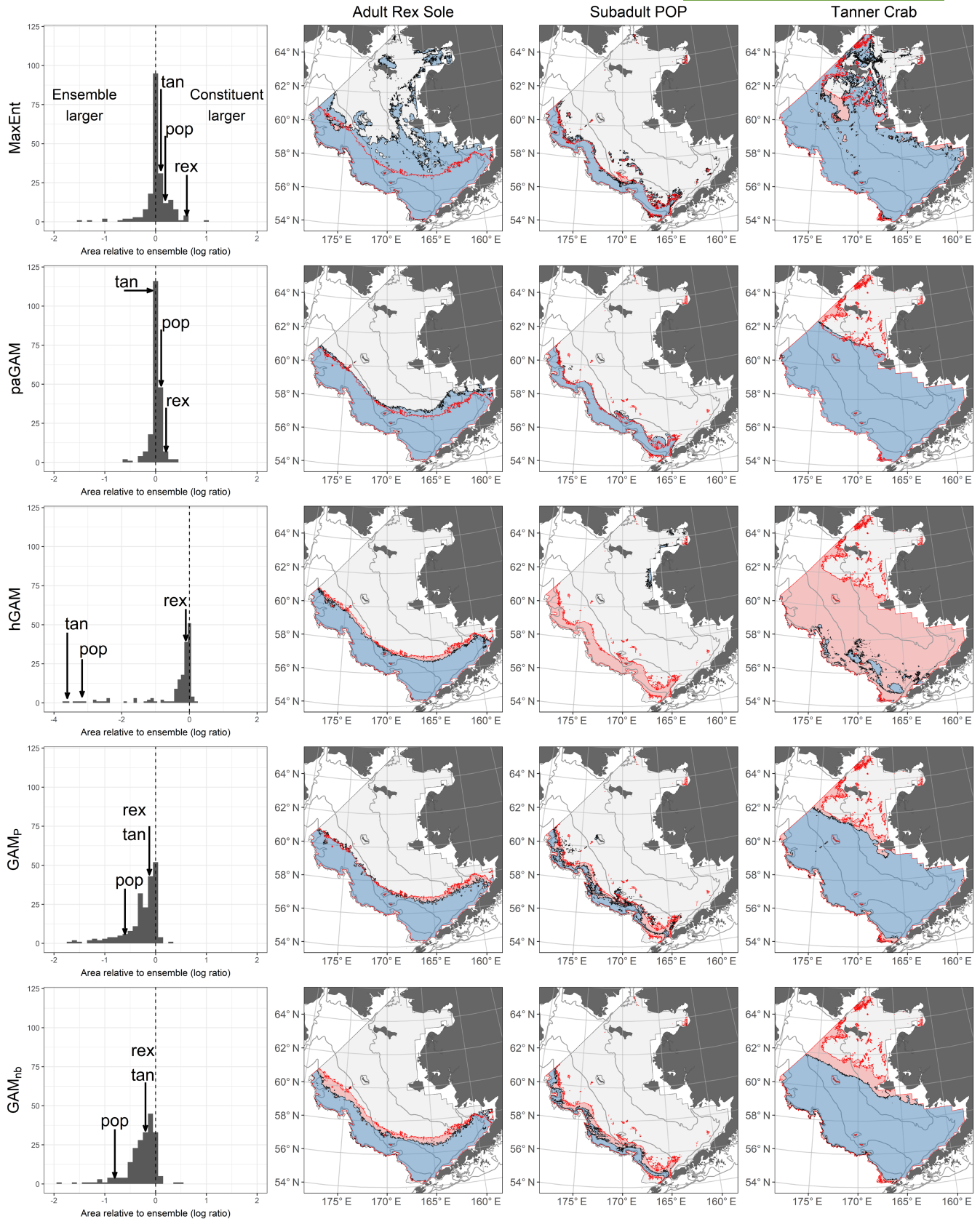
Conversely, the cumulative method tended to produce smaller estimates of area occupied for species/life stages with a patchy distribution (low occurrence in trawls, but high average density). Shortspine thornyhead are found in only 5% of EBS survey stations, but sometimes reach very high density. Using the cumulative method, the adult shortspine thornyhead habitat area was reduced by approximately 75% relative to the probability method and their predicted occupied habitat was restricted to the deeper portions of the Bering slope (Figures 5 and 6). Adult walleye pollock demonstrate the median response, with an occupied habitat area

approximately one-third smaller according to the cumulative method (Figure 6). Walleye pollock are very common in the EBS (present in 86% of tows), and while it is sometimes caught in large numbers, its distribution is less patchy than that of shortspine thornyhead.

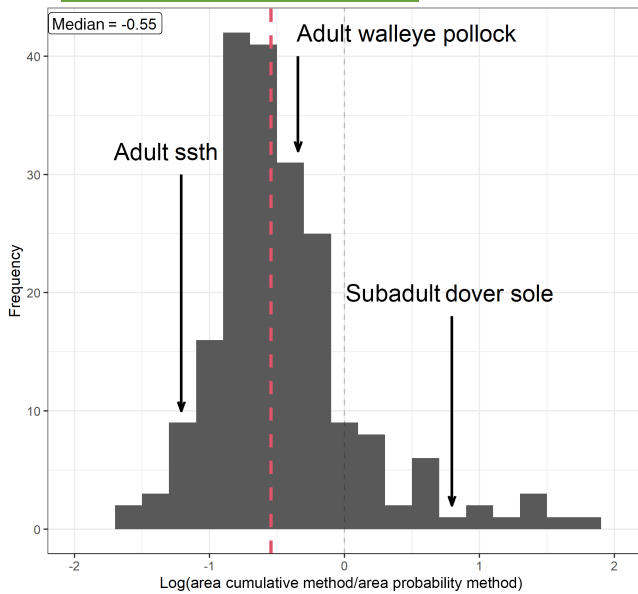
## 4 | DISCUSSION

We set out to evaluate how model structure and choice affect conclusions about species distribution. We compared the performance of five SDMs on 208 species/life stage combinations from three geographic regions. The MaxEnt model was informative for distributions of less common species, like some rockfishes whose patchy distribution and association with rocky areas makes them more difficult to sample with bottom trawl gear, which fits the original purpose of MaxEnt models to identify species distributions from disparate or opportunistic data sources (Elith et al., 2011; Phillips & Dudik, 2008). However, GAMs typically had lower RMSE and were less likely to predict encounters outside of its observed range compared to MaxEnt. The hGAM was the most complex among the GAMs, but it sometimes failed to converge and often had a higher RMSE than the other GAMs. The GAM<sub>p</sub> fit a wide variety of stock/lifestage combinations and typically had a lower RMSE than the more complex GAM<sub>nb</sub>, even though the GAM<sub>nb</sub> is theoretically capable of accounting for the overdispersion. Evaluating model performance with out-of-sample RMSE may disfavour more complex models like the GAM<sub>nb</sub> and hGAM because of their greater potential for overfitting. However, the median RMSE (Table 2) does not capture that these





**FIGURE 4** Comparison of occupied habitat area using the probability method for the five models relative to the ensemble. The left column shows the frequency distribution of the log ratio of occupied habitat with the dotted line at zero representing the ensemble. The maps show adult rex sole (rex), subadult Pacific ocean perch (POP) and Tanner crab (tan) area occupied in the Bering Sea from the distribution, with the blue shaded area with a black outline representing the occupied habitat area for that model and the red shaded area with a bright red outline representing the predicted area occupied in the ensemble.



**FIGURE 5** Histogram showing the difference (log ratio) of the cumulative method (occupied habitat is the smallest area that contains 95% of the total abundance) compared to the probability method (occupied habitat based on areas with >5% encounter probability) for identifying occupied habitat, for all regions combined. A value greater than 0 indicates that the cumulative method produced a larger estimate of the habitat area than the cumulative method for a given species/life stage. The red dashed line shows the median value of 0.55. The abbreviation 'ssth' represents shortspine thornyhead.

models show excellent performance in some particular species and were useful within the ensemble framework. In this study, the ensemble approach mitigated some of the weaknesses of these models, providing better model fits and more stable estimates of area occupied.

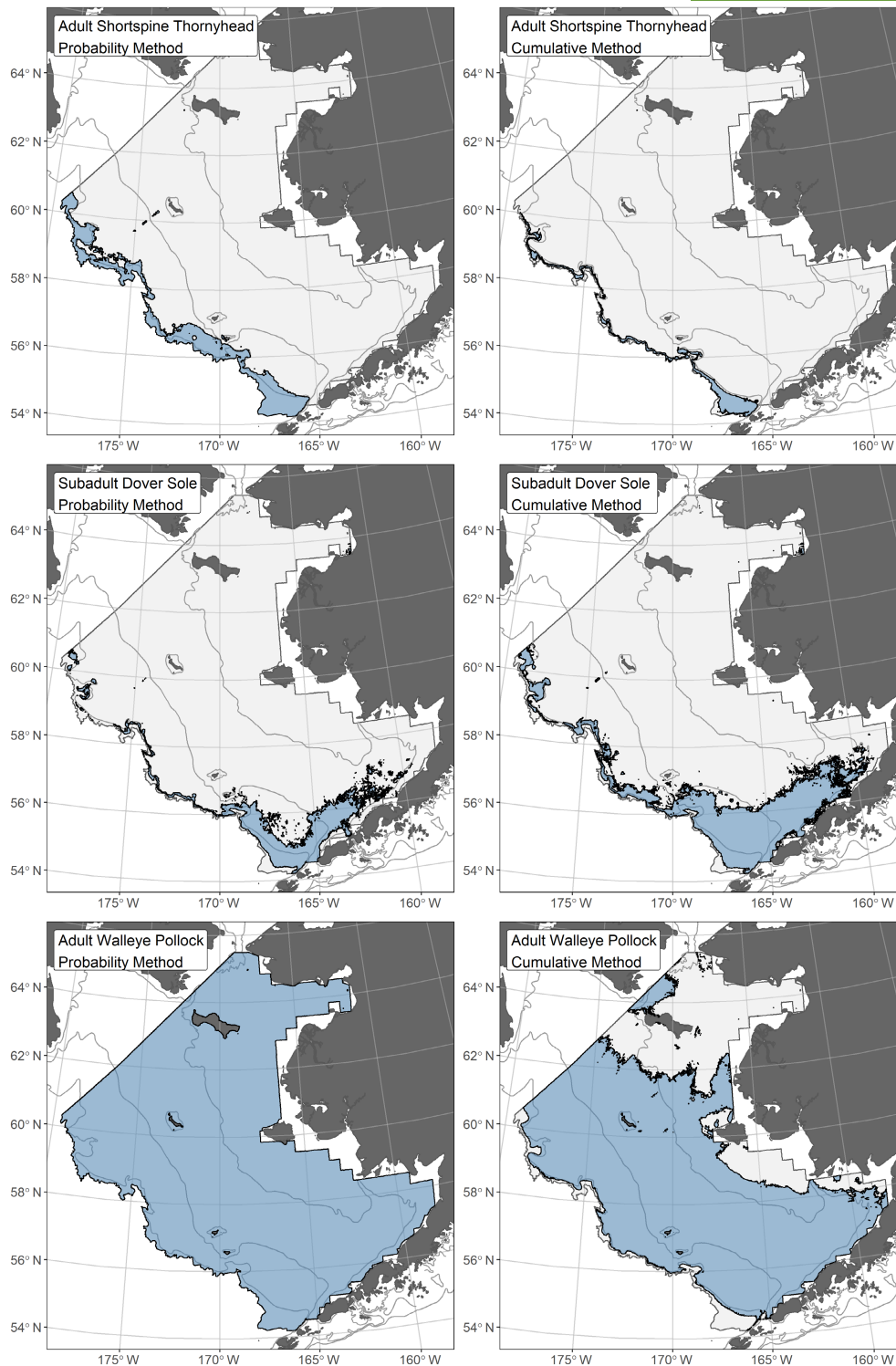
In this study, the two SDMs that are primarily designed around presence-only or presence-absence data (MaxEnt and paGAM) tended to predict larger areas occupied than the others. This is in large part because these models approximate the abundance and must be scaled in order to make accurate predictions of numerical densities. A large scaling factor changes the relationship between predictions of encounter probability and density and seems to inflate probability estimates. While all the models in this study were scaled, the scale factors for the hGAM, GAM<sub>p</sub> and GAM<sub>nb</sub> were typically close to one. In a probability-based approach to defining occupied habitat, this often leads to larger than expected occupied habitat areas. Our results are in agreement with previous studies that found that ensemble models tend to outperform individual models (Seni et al., 2010).

The probability method that we introduce here, in which occupied habitat is defined as areas with a >5% encounter probability, often results in a larger estimate of occupied habitat compared to the cumulative method. This method has a number of potential drawbacks, including that it does not make full use of the density predictions and requires a probability threshold to be specified outside of the

model. By contrast, the cumulative method is more straightforward and can be interpreted as the area that contains 95% percent of the estimated abundance or biomass. While the user must still choose which quantiles of density are relevant for describing key habitat, those quantiles are defined when model predictions are made and will not change afterwards. In abstract terms, the two methods will differ most for species that have a patchy distribution (i.e. species for which most of the total density is found in a small area). Maps produced using the cumulative method will emphasize those areas of high density, whereas maps using the probability method will tend to emphasize the maximum range or geographic extent.

SDM predictions are used in management to describe important habitat for marine species (like EFH), project species habitat-related distribution shifts under climate change scenarios (e.g. Rooper et al., 2021), to obtain indices of abundance and age composition from monitoring data (O'Leary et al., 2020), and to develop stock-specific metrics and status indicators that can be used in stock assessments (Shotwell et al., 2022). The results of this study highlight the importance of keeping management goals in mind when selecting SDM methods and interpreting their results. The 'probability' method is consistent with methods used previously for identifying EFH (Laman et al., 2018). However, this approach generally resulted in a larger estimate of area occupied, which was less precise and harder to communicate to managers and stakeholders in an intuitive way. Density predictions comparable to the 'cumulative' method have been used to quantify whether a range of suitable habitats has shifted, or to characterize certain areas as being important to population productivity (Fredston et al., 2021; Rosenberg et al., 2000). In many of these examples, density quantiles are used to obtain basic information about range shifts or to identify important areas for marine planning (Melo-Merino et al., 2020). The relative utility of these two methods should be dictated by the desired management outcomes and the ecology of the species. A method that estimates a larger area occupied may be less precise but more robust to seasonal shifts in distributions, a relevant concern for survey data that are collected in one season. However, if the chief concern is to delineate a species' core habitat area as opposed to its peripheral habitat, the more precise areas identified using the cumulative method will be more useful.

While the ensemble approach described here is useful for predicting large-scale patterns across several species, we recommend that future work explores SDMs that can handle dynamic environmental covariates (e.g. Barnes et al., 2022) and the mechanistic links between environmental and biological processes (Thorson et al., 2021), that incorporate life history processes including life stage-specific habitat needs and that provide more accurate predictions for species with unique distributions (e.g. a strip of habitat within an ecosystem). Here, we focused on GAMs because they are familiar to stakeholders and are simpler and more easily interpreted than some of the more complex SDMs. MaxEnt models were included here because of their prevalence in habitat analyses for decision-making about rare species, but may soon be replaced by other models that can better account for differences in catchability



**FIGURE 6** Maps showing the difference in predicted area occupied between the probability method (left column; occupied habitat based on areas with >5% encounter probability) and the cumulative method (right column; occupied habitat is the smallest area that contains 95% of the total abundance) for three species/life stages.

(Yackulic et al., 2013). For situations where data are limited in general, and for the Alaska EFH process in particular, machine learning approaches such as random forest or spatio-temporal models may be considered in the future and could readily be included in an ensemble framework like the one described here.

Stakeholder understanding of confidence in scientific information is critical to stakeholder support and the outcomes of natural resource management (e.g. Turner et al., 2016). We have shown that the identification of occupied habitat is affected by a priori choice of correlative models and mapping methods. This ensemble-based

approach to modelling species distributions has several benefits regardless of how the end product is used: ensemble models are more adaptable (allowing the addition of new SDMs as needed) and more stable over time (i.e. because model weights change more slowly with the inclusion of new data than a binary model-selection criterion). When regularly reviewing quantitative maps of occupied habitat, moving from using single SDMs to SDM ensembles in process should reduce the magnitude of the change in area occupied attributable to modelling methods development, so that changes in area occupied due to the environment or other impacts may be more easily detected. The approach described here represents an improvement over previously used methods for describing EFH for groundfish and invertebrates in Alaska (Laman et al., 2018).

## AUTHOR CONTRIBUTIONS

All authors were involved in the design and conception of this study. E. A. Laman and J. L. Pirtle assembled the background data and environmental covariates. J. Harris was responsible for all code produced during this project and led the analysis of the data, with additional advice and assistance from James T. Thorson. J. Harris and M. C. Siple led the writing and editing of the manuscript and figures.

## ACKNOWLEDGEMENTS

The authors thank Lyle Britt and Stan Kotwicki of the NMFS AFSC Groundfish Assessment Program for their support of this work. We thank Gretchen Harrington, Kim Rand and Molly Zaleski for reviews of the initial analysis presented in this report, and stock assessment authors, other species experts, and North Pacific Fishery Management Council EFH 5-year Review reviewers for reviewing the results presented here and providing input on species distributions and modelling approaches. We thank Mark Zimmermann and Chris Rooper for developing bathymetry and invertebrate layers that were used as covariates. This project was funded by the NMFS Alaska Region Essential Fish Habitat Research Plan. Finally, we thank Heather Welch and Sean Rohan at NOAA for reviewing this manuscript before submission.

## CONFLICT OF INTEREST STATEMENT

We declare that we have no conflict of interest.

## DATA AVAILABILITY STATEMENT

All survey data analysed in this manuscript are available publicly through the Fisheries One Stop Shop platform (<https://www.fisheries.noaa.gov/foss/f?p=215:28>). Sources for environmental data used as covariates are described in Table S1. Up to date code used to generate SDMs and create ensemble models is available on GitHub (<https://github.com/alaska-groundfish-efh/EFHSDM>). Archived code and data used for this paper are available from Zenodo: <https://zenodo.org/records/10211730> (Harris et al., 2023).

## ORCID

J. Harris  <https://orcid.org/0000-0001-5465-6409>

J. L. Pirtle  <https://orcid.org/0000-0002-4421-8234>

E. A. Laman  <https://orcid.org/0000-0003-2663-9243>

M. C. Siple  <https://orcid.org/0000-0002-4260-9856>

James T. Thorson  <https://orcid.org/0000-0001-7415-1010>

## REFERENCES

- Abrahms, B., Welch, H., Brodie, S., Jacox, M. G., Becker, E. A., Bograd, S. J., Irvine, L. M., Palacios, D. M., Mate, B. R., & Hazen, E. L. (2019). Dynamic ensemble models to predict distributions and anthropogenic risk exposure for highly mobile species. *Diversity and Distributions*, 25(8), 1182–1193. <https://doi.org/10.1111/ddi.12940>
- Alverson, D. L., & Pereyra, W. T. (1969). Demersal fish explorations in the northeastern Pacific Ocean—An evaluation of exploratory fishing methods and analytical approaches to stock size and yield forecasts. *Journal of the Fisheries Research Board of Canada*, 26, 1985–2001. <https://doi.org/10.1139/f69-188>
- Andersen, P. S., Andersen, E., Graversgaard, M., Christensen, A. A., Vejre, H., & Dalgaard, T. (2019). Using landscape scenarios to improve local nitrogen management and planning. *Journal of Environmental Management*, 232, 523–530. <https://doi.org/10.1016/j.jenvman.2018.11.023>
- Anderson, S. C., Ward, E. J., English, P. A., & Barnett, L. A. K. (2022). sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *bioRxiv*, [Preprint] <https://doi.org/10.1101/2022.03.24.485545>
- Barnes, C. L., Essington, T. E., Pirtle, J. L., Rooper, C. N., Laman, E. A., Holsman, K. K., Aydin, K. Y., & Thorson, J. T. (2022). Climate-informed models benefit hindcasting but present challenges when forecasting species–habitat associations. *Ecography*, 2022(10), e06189.
- Barry, S. C., & Welsh, A. H. (2002). Generalized additive modeling and zero inflated count data. *Ecological Modeling*, 157, 179–188.
- Brodie, S. J., Thorson, J. T., Carroll, G., Hazen, E. L., Bograd, S., Haltuch, M. A., Holsman, K. K., Kotwicki, S., Samhoury, J. F., Willis-Norton, E., & Selden, R. L. (2020). Trade-offs in covariate selection for species distribution models: A methodological comparison. *Ecography*, 43(1), 11–24. <https://doi.org/10.1111/ecog.04707>
- De Cubber, L., Trenkel, V. M., Diez, G., Gil-Herrera, J., Pabon, A. M. N., Eme, D., & Lorange, P. (2023). Robust identification of potential habitats of a rare demersal species (blackspot seabream) in the Northeast Atlantic. *Ecological Modeling*, 477, 110255.
- Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoň, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J. J., Pollock, L. J., Reineking, B., Roberts, D. R., Schröder, B., Thuiller, W., Warton, D. I., ... Hartig, F. (2018). Model averaging in ecology: A review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecological Monographs*, 88(4), 485–504. <https://doi.org/10.1002/ecm.1309>
- Echave, K., Eagleton, M., Farley, E., & Orsi, J. (2012). A refined description of essential fish habitat for Pacific salmon within the U.S. Exclusive Economic Zone in Alaska. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-236, 104 p.
- Elith, J., Leathwick, J. R., & Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4), 802–813.
- Elith, J., Phillips, S. J., Hastie, T., Dudík, M., Chee, Y. E., & Yates, C. J. (2011). A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17, 43–57.
- Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, 41(2), 263–274. <https://doi.org/10.1111/j.0021-8901.2004.00881.x>
- Enquist, C. A., Jackson, S. T., Garfin, G. M., Davis, F. W., Gerber, L. R., Littell, J. A., Tank, J. L., Terando, A. J., Wall, T. U., Halpern, B., Hiers,

- J. K., Morelli, T. L., McNie, E., Stephenson, N. L., Williamson, M. A., Woodhouse, C. A., Yung, L., Brunson, M. W., Hall, K. R., ... Shaw, M. R. (2017). Foundations of translational ecology. *Frontiers in Ecology and the Environment*, 15(10), 541–550.
- Fitzpatrick, M. C., Lachmuth, S., & Haydt, N. T. (2021). The ODMAP protocol: A new tool for standardized reporting that could revolutionize species distribution modeling. *Ecography*, 44(7), 1067–1070.
- Fredston, A., Pinsky, M., Selden, R. L., Szuwalski, C., Thorson, J. T., Gaines, S. D., & Halpern, B. S. (2021). Range edges of North American marine species are tracking temperature over decades. *Global Change Biology*, 27(13), 3145–3156. <https://doi.org/10.1111/gcb.15614>
- Gelman, A. (2014). *Bayesian data analysis* (3rd ed.). CRC Press.
- Guillaumot, C., Moreau, C., Danis, B., & Saucède, T. (2020). Extrapolation in species distribution modelling. Application to Southern Ocean marine species. *Progress in Oceanography*, 188, 102438.
- Guillera-Aroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., ... Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3), 276–292.
- Guisan, A., Edwards, T. C., Jr., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, 157(2–3), 89–100.
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M., ... Buckley, Y. M. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16(12), 1424–1435. <https://doi.org/10.1111/ele.12189>
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2–3), 147–186. [https://doi.org/10.1016/S0304-3800\(00\)00354-9](https://doi.org/10.1016/S0304-3800(00)00354-9)
- Hao, T., Elith, J., Lahoz-Monfort, J. J., & Guillera-Aroita, G. (2020). Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography*, 43, 549–558.
- Harris, J., Laman, E. A., Pirtle, J., Siple, M. C., Rooper, C. N., Hurst, T., & Conrath, C. L. (2022). Advancing model-based essential fish habitat descriptions for North Pacific species in the Aleutian Islands. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-458, 406 p.
- Harris, J., Pirtle, J., Siple, M., Thorson, J., & Laman, E. (2023). Data from: An ensemble approach to species distribution modeling reconciles systematic differences in estimates of habitat utilization and range area. *Zenodo*. <https://doi.org/10.5281/zenodo.10211730>
- Hastie, T. J., & Tibshirani, R. J. (1990). Generalized additive models. *Monographs on Statistics and Applied Probability*, 43, 338.
- Holbrook, S. J., Forrester, G. E., & Schmitt, R. J. (2000). Spatial patterns in abundance of a damselfish reflect availability of suitable habitat. *Oecologia*, 122(1), 109–120. <https://doi.org/10.1007/PL00008826>
- IUCN. (2012). *IUCN red list categories and criteria, version 3.1* (2nd ed.). International Union for the Conservation of Nature. <https://portals.iucn.org/library/node/10315>
- Jetz, W., McGeoch, M. A., Guralnick, R., Ferrier, S., Beck, J., Costello, M. J., Fernandez, M., Geller, G. N., Keil, P., Merow, C., Meyer, C., Muller-Karger, F. E., Pereira, H. M., Regan, E. C., Schmeller, D. S., & Turak, E. (2019). Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology & Evolution*, 3(4), Article 4. <https://doi.org/10.1038/s41559-019-0826-1>
- Laman, E. A., Pirtle, J., Harris, J., Siple, M. C., Rooper, C. N., Hurst, T., & Conrath, C. L. (2022). Advancing model-based essential fish habitat descriptions for North Pacific species in the Bering Sea. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-459, 538 p.
- Laman, E. A., Rooper, C. N., Turner, K., Rooney, S., Cooper, D. W., & Zimmermann, M. (2018). Using species distribution models to describe essential fish habitat in Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 75, 1230–1255. <https://doi.org/10.1139/cjfas-2017-0181>
- Magnuson-Stevens Fishery Conservation and Management Act. (1996). 50 CFR 600.10.
- Markowitz, E. H., Dawson, E. J., Charriere, N. E., Prohaska, B. K., Rohan, S. K., Haehn, R. A., Stevenson, D. E., & Britt, L. L. (2022). Results of the 2018 eastern Bering Sea continental shelf bottom trawl survey of groundfish and invertebrate fauna. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-450, 183 p.
- Melo-Merino, S. M., Reyes-Bonilla, H., & Lira-Noriega, A. (2020). Ecological niche models and species distribution models in marine environments: A literature review and spatial analysis of evidence. *Ecological Modelling*, 415, 108837. <https://doi.org/10.1016/j.ecolmodel.2019.108837>
- O'Leary, C. A., Thorson, J. T., Ianelli, J. N., & Kotwicki, S. (2020). Adapting to climate-driven distribution shifts using model-based indices and age composition from multiple surveys in the walleye pollock (*Gadus chalcogrammus*) stock assessment. *Fisheries Oceanography*, 29(6), 541–557. <https://doi.org/10.1111/fog.12494>
- Pettorelli, N., Lobora, A. L., Msuha, M. J., Foley, C., & Durant, S. M. (2010). Carnivore biodiversity in Tanzania: Revealing the distribution patterns of secretive mammals using camera traps. *Animal Conservation*, 13(2), 131–139. <https://doi.org/10.1111/j.1469-1795.2009.00309.x>
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: An open-source release of Maxent. *Ecography*, 40, 887–893.
- Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190(3–4), 231–259.
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2), 161–175. <https://doi.org/10.1111/j.0906-7590.2008.5203.x>
- Pirtle, J., Laman, E. A., Harris, J., Siple, M. C., Rooper, C. N., Hurst, T., Conrath, C. L., & Gibson, G. (2023). Advancing Model-Based Essential Fish Habitat Descriptions for North Pacific Species in the Gulf of Alaska. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-468, 541 p.
- Potts, J., & Elith, J. (2006). Comparing species abundance models. *Ecological Modelling*, 199, 153–163.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Aroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Rollinson, C. R., Finley, A. O., Alexander, M. R., Banerjee, S., Dixon Hamil, K.-A., Koenig, L. E., Locke, D. H., DeMarche, M. L., Tingley, M. W., Wheeler, K., Youngflesh, C., & Zipkin, E. F. (2021). Working across space and time: Nonstationarity in ecological research and application. *Frontiers in Ecology and the Environment*, 19(1), 66–72. <https://doi.org/10.1002/fee.2298>
- Rooper, C. N., Ortiz, I., Hermann, A. J., Laman, N., Cheng, W., Kearney, K., & Aydin, K. (2021). Predicted shifts of groundfish distribution in the eastern Bering Sea under climate change, with implications for fish populations and fisheries management. *ICES Journal of Marine Science*, 78(1), 220–234. <https://doi.org/10.1093/icesjms/fsaa215>
- Rosenberg, A., Bigford, T. E., Leathery, S., Hill, R. L., & Bickers, K. (2000). Ecosystem approaches to fishery management through essential fish habitat. *Bulletin of Marine Science*, 66(3), 535–542.
- Seni, G., Elder, J., & Grossman, R. (2010). *Ensemble methods in data mining: Improving accuracy through combining predictions*. Morgan and Claypool Publishers.

- Shotwell, S. K., Pirtle, J. L., Watson, J. T., Deary, A. L., Doyle, M. J., Barbeaux, S. J., Dorn, M. W., Gibson, G. A., Goldstein, E. D., Hanselman, D. H., Hermann, A. J., Hulson, P. J. F., Laurel, B. J., Moss, J. H., Ormseth, O. A., Robinson, D., Rogers, L. A., Rooper, C. N., Spies, I., ... Vollenweider, J. J. (2022). Synthesizing integrated ecosystem research to create informed stock-specific indicators for next generation stock assessments. *Deep Sea Research Part II: Topical Studies in Oceanography*, 198, 105070. <https://doi.org/10.1016/j.dsr2.2022.105070>
- Sundblad, G., Bergström, U., & Sandström, A. (2011). Ecological coherence of marine protected area networks: A spatial assessment using species distribution models. *Journal of Applied Ecology*, 48(1), 112–120. <https://doi.org/10.1111/j.1365-2664.2010.01892.x>
- Thorson, J. T. (2019). Guidance for decisions using the Vector Autoregressive Spatio-Temporal (VAST) package in stock, ecosystem, habitat and climate assessments. *Fisheries Research*, 210, 143–161. <https://doi.org/10.1016/j.fishres.2018.10.013>
- Thorson, J. T., Hermann, A. J., Siwicke, K., & Zimmermann, M. (2021). Grand challenge for habitat science: Stage-structured responses, nonlocal drivers, and mechanistic associations among habitat variables affecting fishery productivity. *ICES Journal of Marine Science*, 78(6), 1956–1968. <https://doi.org/10.1093/icesjms/fsaa236>
- Turner, R. A., Addison, J., Arias, A., Bergseth, B. J., Marshall, N. A., Morrison, T. H., & Tobin, R. C. (2016). Trust, confidence, and equity affect the legitimacy of natural resource governance. *Ecology and Society*, 21(3), 18 p. <https://www.jstor.org/stable/26269961>
- Valavi, R., Guillera-Arroita, G., Lahoz-Monfort, J. J., & Elith, J. (2022). Predictive performance of presence-only species distribution models: A benchmark study with reproducible code. *Ecological Monographs*, 92(1), e01486.
- von Szalay, P. G., & Raring, N. W. (2018). Data Report: 2017 Gulf of Alaska bottom trawl survey. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-374, 260 p.
- von Szalay, P. G., & Raring, N. W. (2020). Data Report: 2018 Aleutian Islands bottom trawl survey. U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-AFSC-409, 175 p.
- Waldock, C., Stuart-Smith, R. D., Albouy, C., Cheung, W. W., Edgar, G. J., Mouillot, D., Tjiputra, J., & Pellissier, L. (2022). A quantitative review of abundance-based species distribution models. *Ecography*, 2022(1), 18 p. <https://doi.org/10.1111/ecog.05694>
- Yackulic, C. B., Chandler, R., Zipkin, E. F., Royle, J. A., Nichols, J. D., Grant, E. H. C., & Veran, S. (2013). Presence-only modelling using MAXENT: When can we trust the inferences? *Methods in Ecology and Evolution*, 4(3), 236–243.
- Zurell, D., Franklin, J., König, C., Bouchet, P. J., Dormann, C. F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., & Guisan, A. (2020). A standard protocol for reporting species distribution models. *Ecography*, 43, 1261–1277.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer Science+Business.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**Supporting Information Table S1.** Covariates used for species distribution models. The ‘response function used in GAMs’ column describes the function specified to transform the covariate when estimating the linear predictor, where ‘2D smooth’ refers to a Duchon spline (Duchon, 1977) with a 1st order smoothing penalty, and ‘1D smooth’ may refer to either a univariate thin-plate spline with a 2nd-order smoothing penalty or a cubic regression spline (Wood, 2003). Data source describes the method that was used to obtain each covariate.

**Supporting information S2.** An ODMAP protocol (Fitzpatrick et al., 2021) for adult Rex Sole in the Gulf of Alaska. While this project produced 208 SDM ensembles, this one is provided as an example to show the approach that was used throughout. Detailed information for all 208 models can be obtained from Harris et al. (2022), Laman et al. (2022) and Pirtle et al. (2023).

**Supporting Information Table S3.** Summary statistics from 208 SDM ensembles used in this project. The ‘Region’ column indicates if the model pertained to the Aleutian Islands (AI), Eastern Bering Sea (EBS) or Gulf of Alaska (GOA). The ‘Species’ and ‘Lifestage’ columns give information about the organism being modelled. The ‘Model’ column shows each of the model types explored, and the ‘Converged’ and ‘Plausible’ columns state where the model passed various checks and was included in the ensemble. The ‘Weight’ column provides the fractional contribution of each constituent model to the ensemble. ‘Scale’ shows the scaling factor used for that model. ‘RMSE’ is the root-mean-squared error for each constituent or the ensemble. The columns ‘Probability Area’ and ‘Cumulative Area’ show the predicted area of occupied habitat for that model or ensemble for using the probability method or cumulative method as described in the Materials and Methods section.

**How to cite this article:** Harris, J., Pirtle, J. L., Laman, E. A., Siple, M. C., & Thorson, J. T. (2023). An ensemble approach to species distribution modelling reconciles systematic differences in estimates of habitat utilization and range area. *Journal of Applied Ecology*, 00, 1–14. <https://doi.org/10.1111/1365-2664.14559>