Science Advances

◤AAAS

Supplementary Materials for

The environmental niche of the global high seas pelagic longline fleet

Guillermo Ortuño Crespo*, Daniel C. Dunn, Gabriel Reygondeau, Kristina Boerder, Boris Worm, William Cheung, Derek P. Tittensor, Patrick N. Halpin

*Corresponding author. Email: gortunocrespo@gmail.com

This PDF file includes:

Supplementary Materials and Methods

Fitting and projecting boosted regression trees

The use of boosted regression trees incorporates the strengths of tree-based models (which can utilize multiple types of response variables and missing data) and boosting (an adaptive method of sequentially modeling the residuals of each model iteration) (*39*). Many of the relationships between the distribution of marine species and their environment have proven to be non-linear (*26*). BRTs have the ability to capture non-linear responses in environmental predictors and to identify the variables with the highest explanatory power. The machine learning dimension of BRTs allows the model to progressively learn the response through boosting which, unlike single regression treqes, averages the outputs of a multitude of simple tree models iteratively placing more emphasis on unfitted observations from the first tree by fitting the residuals to a second tree and so on (*39*). Another reason why we chose BRTs as our modeling approach is because they incorporate an important element of stochasticity that improves model performance by fitting each new tree to a random subset of the data that is predetermined by the researcher and known as the 'bag fraction'(*40*); this improves model performance by reducing the variance of the final model. The contribution of each individual tree (hundreds to thousands) is then downscaled by assigning it a 'learning rate' weight. Our final models were the output of the summation of all trees multiplied by the predetermined learning rate (*39*). The complexity or depth of the tree must also be determined, this controls the number of splits or nodes in each tree. At each splitting point, the variable which has the most explanatory power is partitioned in a way that reduces the prediction error, this is done sequentially for however many splits the tree had or until a predetermined minimum number of observations was reached in a branch of a tree. The learning rate and tree complexity parameters are then used in conjunction to determine the total number of trees that should be built.

When fitting BRTs, we paid special attention to finding a balance between model accuracy and model overfitting. We used the R package '*caret*' to optimize the four important parameters affecting BRTs: (1) tree complexity; (2) learning rate (or shrinkage); (3) number of trees; and (4) minimum number of observations in terminal nodes. An exploratory grid was set up where different values for these parameters would be assessed through k-fold cross-validation. These included, four different tree depths, two shrinkage rates, two minimum number of observations at terminal nodes and 10 different number of trees from 500 to 5000 in increments of 500 trees. A 10-fold cross validation of each of the 160 different model combinations was repeated five times to select the parameters that resulted in the highest accuracy rate. There are different opinions about the use and number of background

pseudoabsence points in environmental niche models, we selected a number which is consistent with what is recommended in the literature (*37*) (Table S11 & S12). The optimal combination of parameters was found for each of the 24 individual monthly models and BRTs were fitted to the monthly classified fishing effort data using the generalized boosted model (gbm) function in the R package 'gbm'. The distribution of the models was set to *bernoulli* given binomial nature of the response. After fitting each of the models, 1° by 1° degree raster layers of each of the environmental variables in ABNJ used in the monthly models were superposed and stacked using ArcGIS 10.4.1 and the *'raster'* package on R to project the model results onto two-dimensional geographic space. Through our study, we also address the use of different thresholds by selecting and comparing two cutoff thresholds, which are supported in the environmental niche modeling literature (*38*). The code used to fit the BRT models is available upon request to the authors.

Environmental niche model assessment and interpretation

The binary (suitable vs. non-suitable fishing habitat) monthly habitat maps were then used to calculate the annual persistence of longline fishing habitat in ABNJ; these were created to identify different levels of intra-annual suitability throughout the high seas (Fig. 3 & 4). Persistence is a measure of the number of months that a location was classified as suitable habitat throughout the year. To identify the areas of the high seas with the highest variation in fishing habitat suitability, we calculated the spatial coefficient of variation (Fig. 5) using all of the monthly prediction maps. In addition to monthly persistence, estimates of suitable fishing habitat occupancy by the pelagic longline fleet in ABNJ were calculated my measuring what proportion of the monthly environmental niches (determined by the binary maps) were fished each month (Fig. 3 & 4). This was calculated separately for the maps derived from each of the cutoff thresholds. These outputs were then used to assess how much of the 'fishable' environment in ABNJ is used by the high seas longline fleet each month (Table S6).

For BRTs, the VI scores are calculated by computing the number of times an environmental predictor is selected for splitting at each node of a tree and is then weighted by the squared improvement of the BRT model that results from each split; this is then averaged for each of the trees that was fitted for the BRT(*39*). VI scores for each BRT are scaled to add up to 100, these values were obtained for each of the monthly environmental niche models; these scores allow for the identification of important variables within and between years (Fig. 6; Table S7 & S8). Variables with VI scores >10 were considered meaningful for explaining the distribution of fishing effort.

We generated partial dependence plots for each of the response variables. Partial dependence plots are graphical representations of the relationship between the probability of fishing and an environmental variable. The partial dependence plot of a given variable was calculated by averaging all other predictors in the model except the chosen predictor of interest.

Supplementary Figures
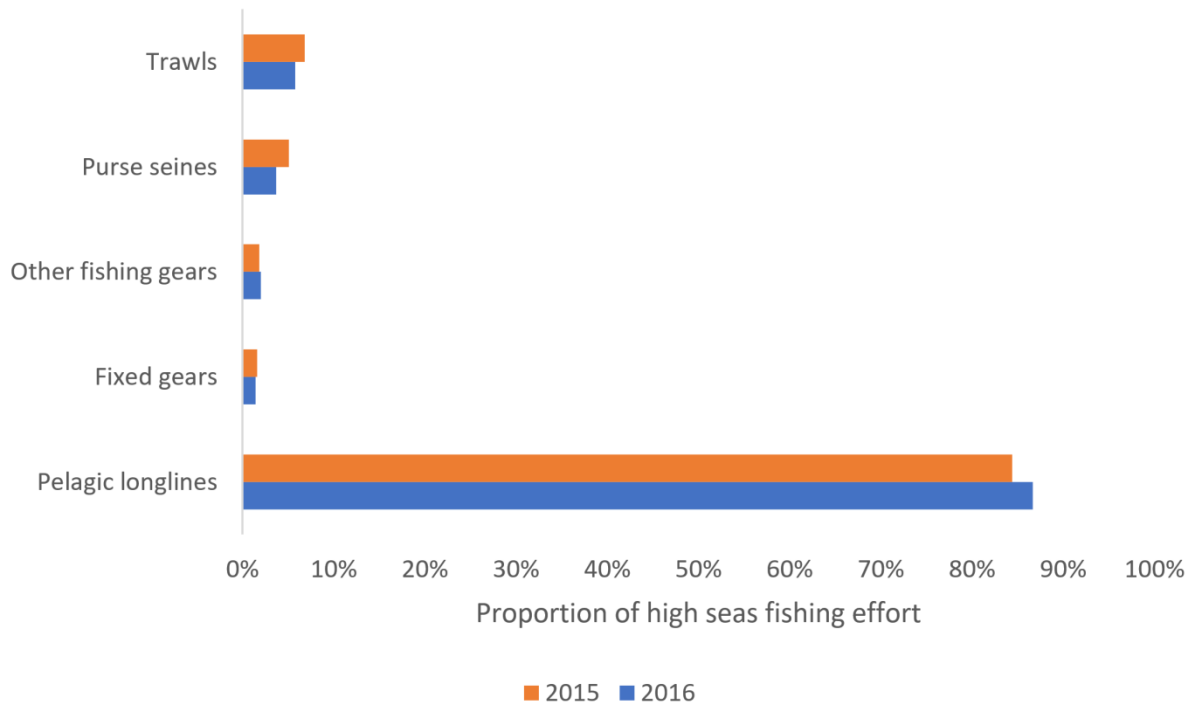
Fig. S1. The proportion of 2015 and 2016 fishing effort (hours) in ABNJ by gear. The fishing effort estimates by gear were calculated by global fishing watch using satellite-based automatic identification system data and neural network classification algorithms. Pelagic longlines are the dominant form of fishing in the high seas as calculated by the number of hours fishing.
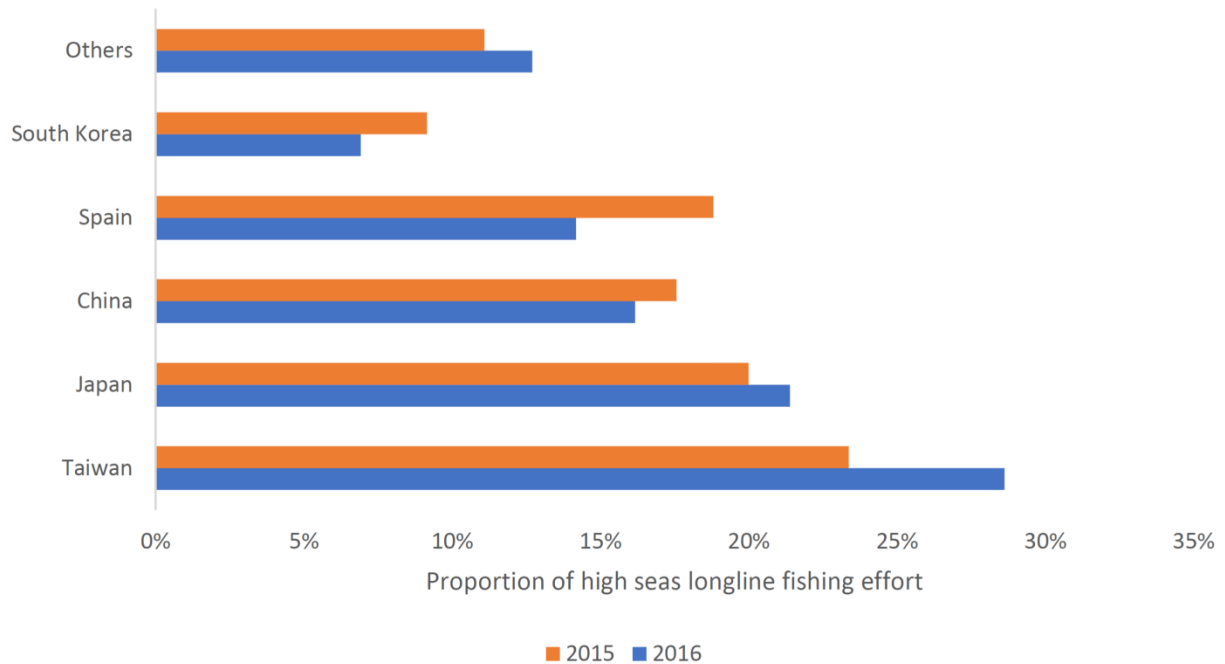
Fig. S2. The proportion of pelagic longline fishing effort attributed to the main fishing States or territories. We break down longline fishing effort by the top-5 longline fishing States or territories in areas beyond national jurisdiction as well as the group comprised by all other longline fishing States in areas beyond national jurisdiction: 'Others'.
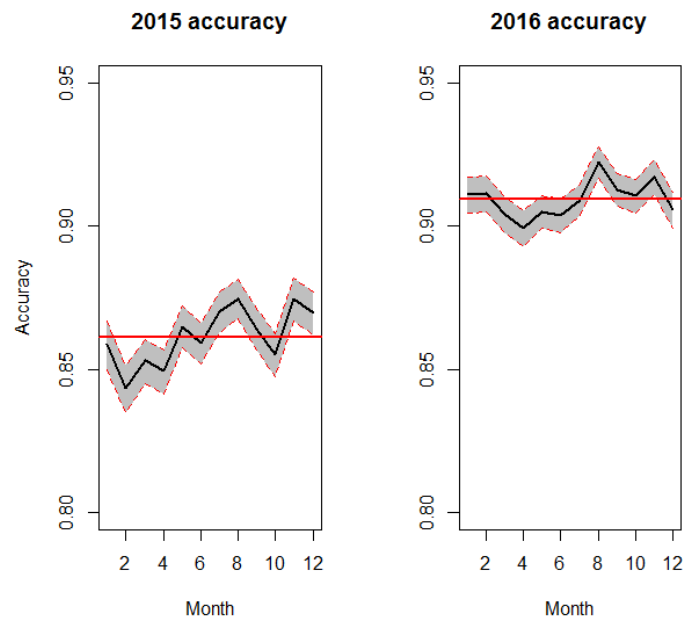
Fig. S3. Accuracy values obtained for the 2015 and 2016 monthly boosted regression tree models after applying an ROC threshold. Shows the accuracy values obtained for the 2015 (left) and 2016 (right) monthly boosted regression tree models (black trend line) after applying an Receiver Operator Characteristic threshold. The monthly accuracy values are bordered by their respective confidence intervals (grey buffer) and the annual average accuracy is overlaid (solid red line).
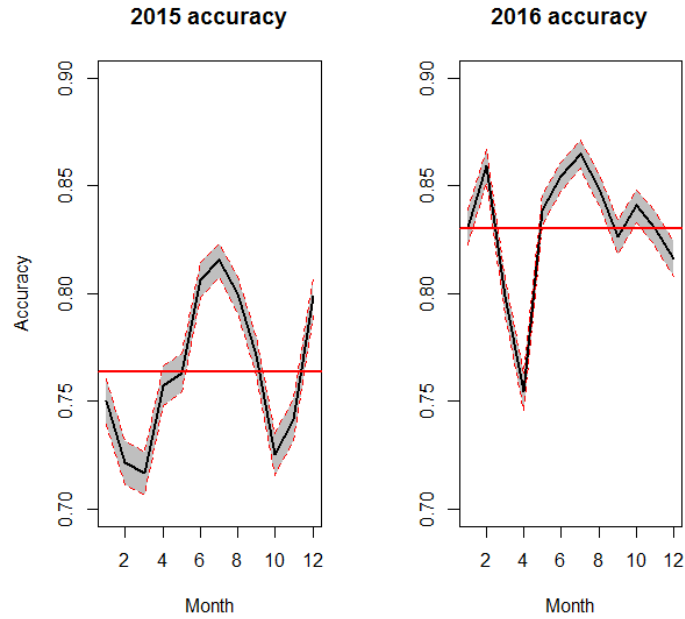
Fig. S4. Accuracy values obtained for the 2015 and 2016 monthly boosted regression tree models after applying an MPD threshold. Shows the accuracy values obtained for the 2015 (left) and 2016 (right) monthly boosted regression tree models (black trend line) after applying a mean probability distribution threshold. The monthly accuracy values are bordered by their respective confidence intervals (grey buffer) and the annual average accuracy is overlaid (solid red line).
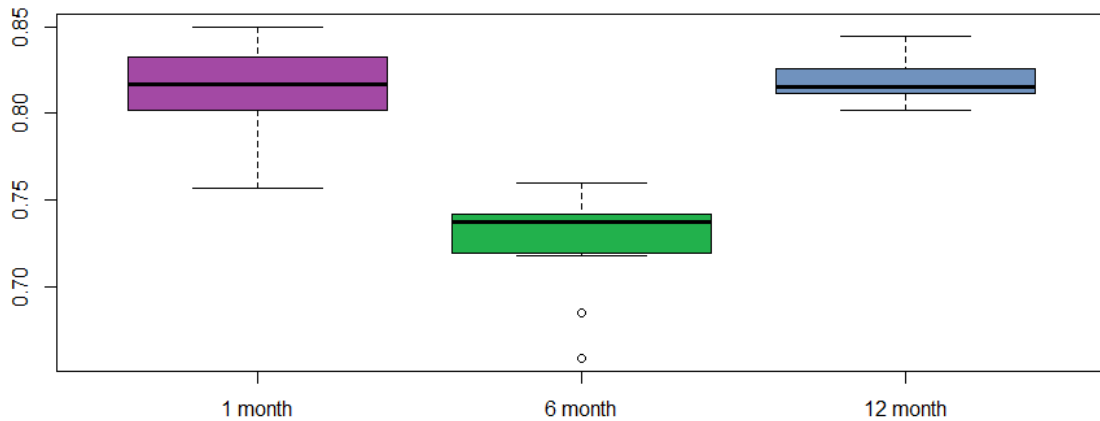
Fig. S5. The predictive accuracy of the monthly BRTs after projecting them onto future environments. The models were projected at three different time-steps: 1 month (n=23), 6 months (n=18) and 12 months (n=12). The accuracy of the models was highest when predicting the distribution of fishing effort for the following month and for the same month of the following year.
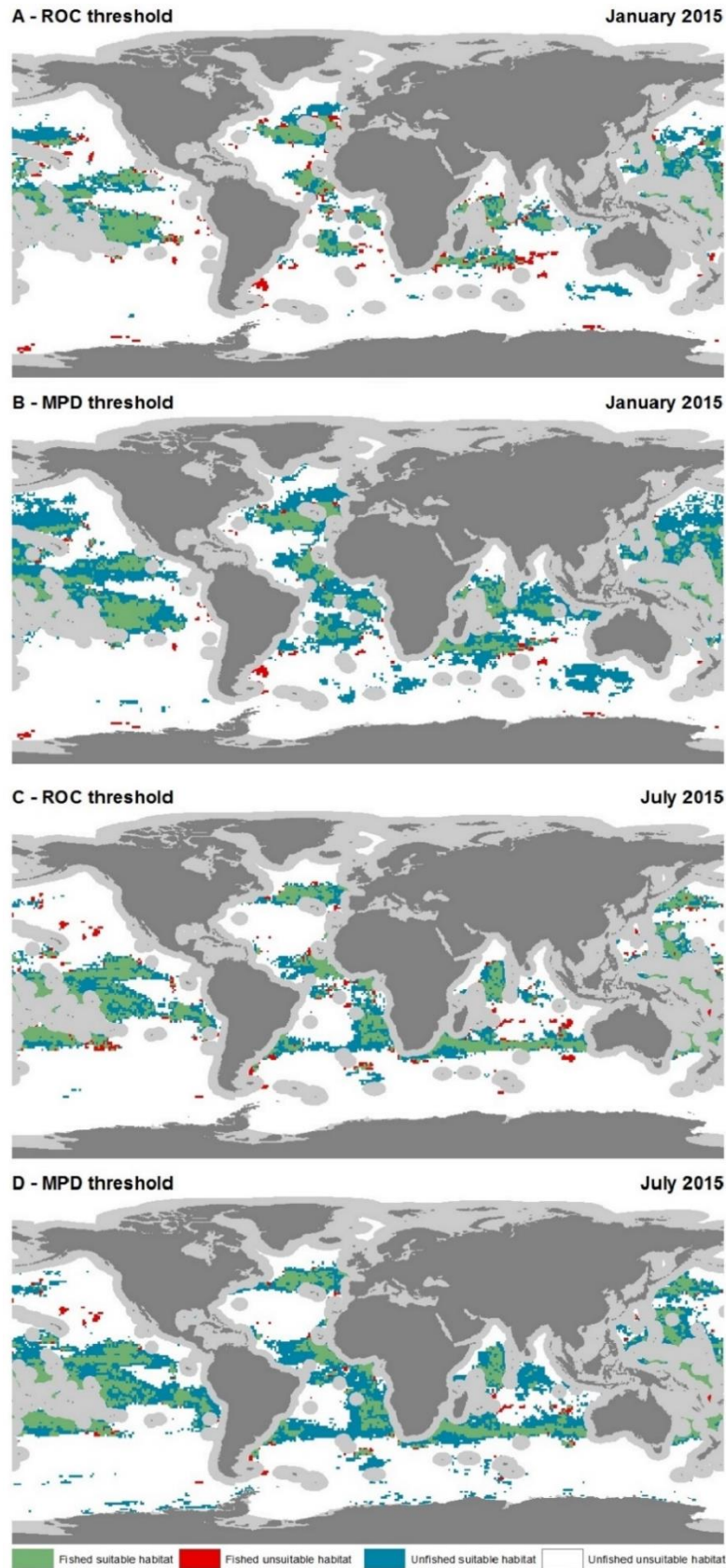
Fig. S6. Distribution of predicted and observed fishing effort in January and July of 2015 using different thresholds: ROC and MPD. (**A** & **C**) ROC. (**B** & **D**) MPD. The true positives (fished suitable habitat) are labeled in *green*; false positives (unfished suitable habitat) are labeled in *blue;* false negatives (fished unsuitable habitat) are labeled in *red;* while the unfished unsuitable habitat is *white*.

2015 Sea Surface Temperature

Fig. S7. The SST partial dependence plots from the monthly 2015 models.
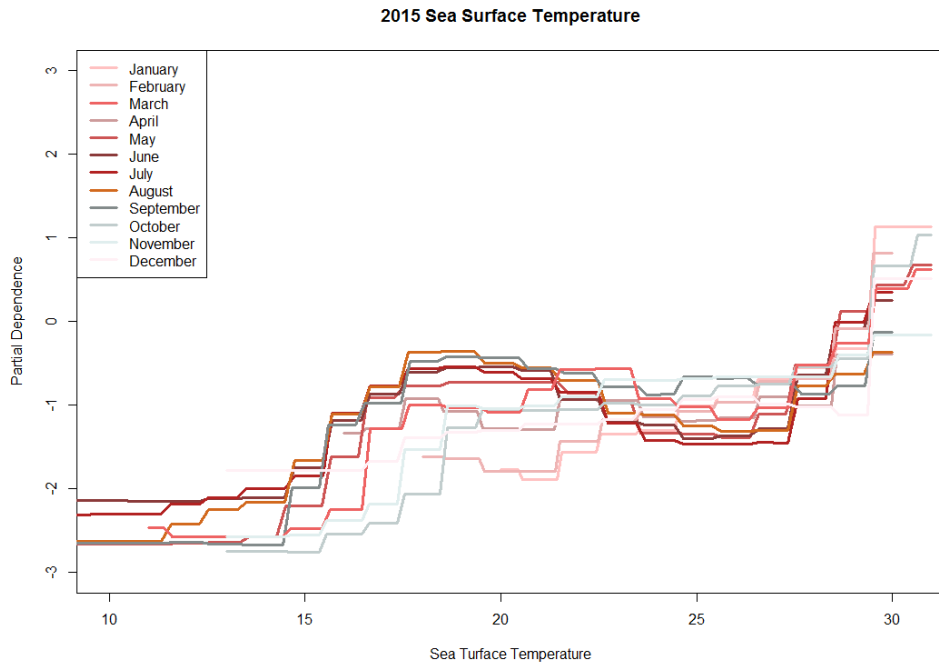
2015 Temperature at 400 meters
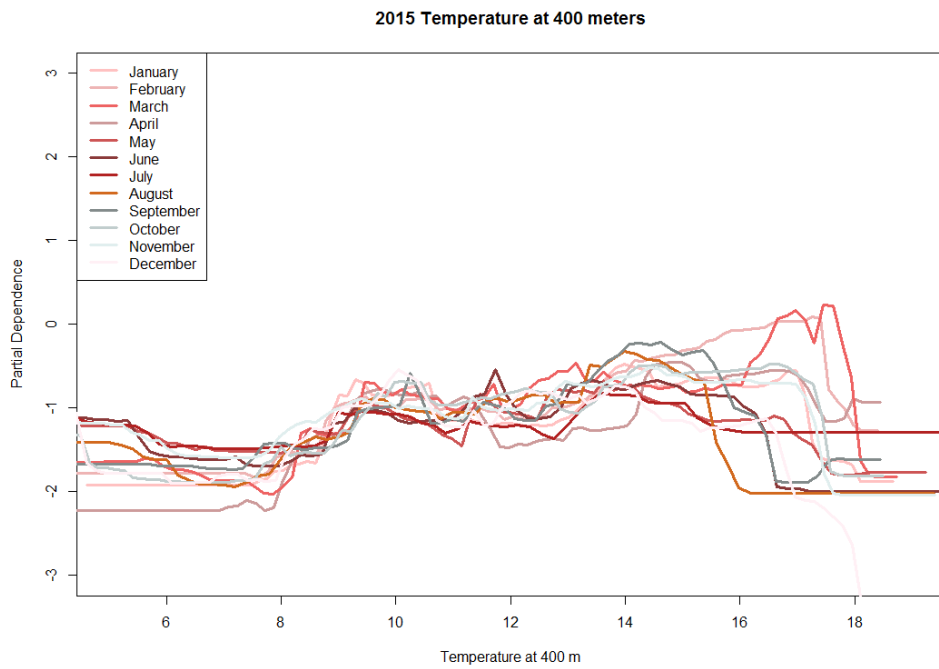
Fig. S8. The temperature at 400-m partial dependence plots from the monthly 2015 models.
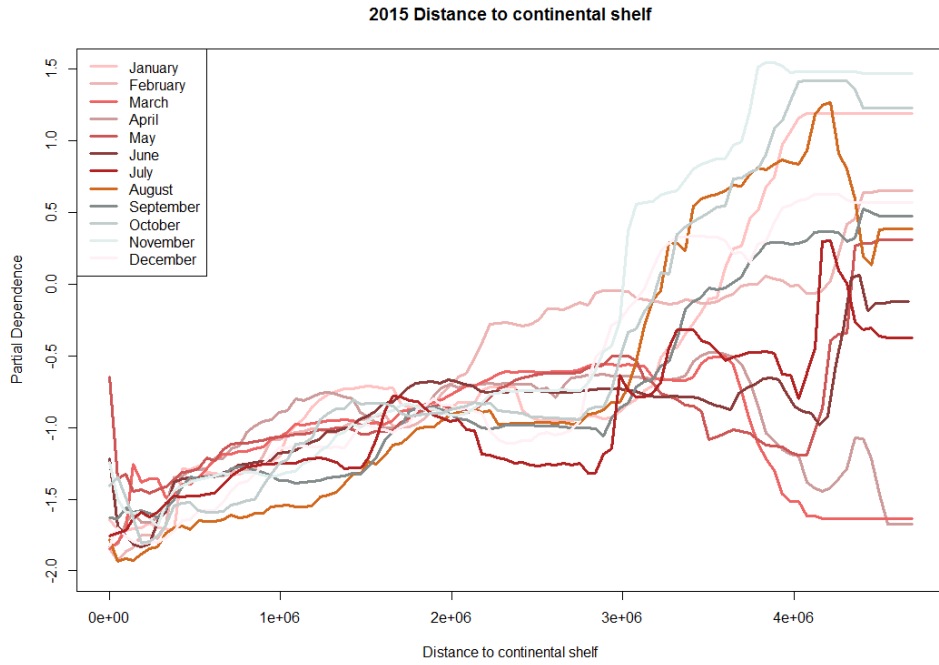
Fig. S9. The DCS partial dependence plots from the monthly 2015 models.
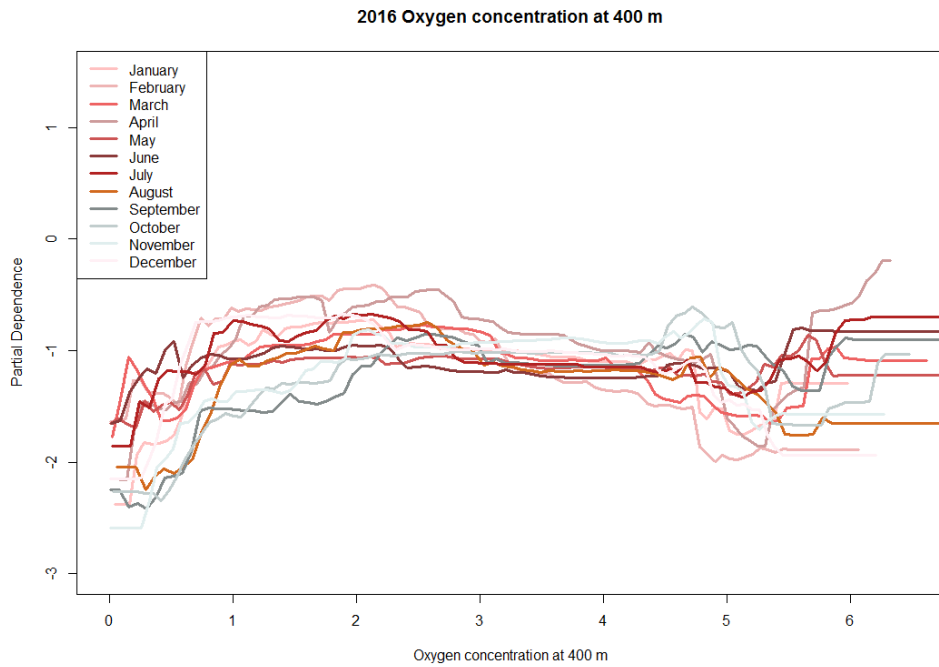
Fig. S10. The oxygen at 400-m partial dependence plots from the monthly 2016 models.
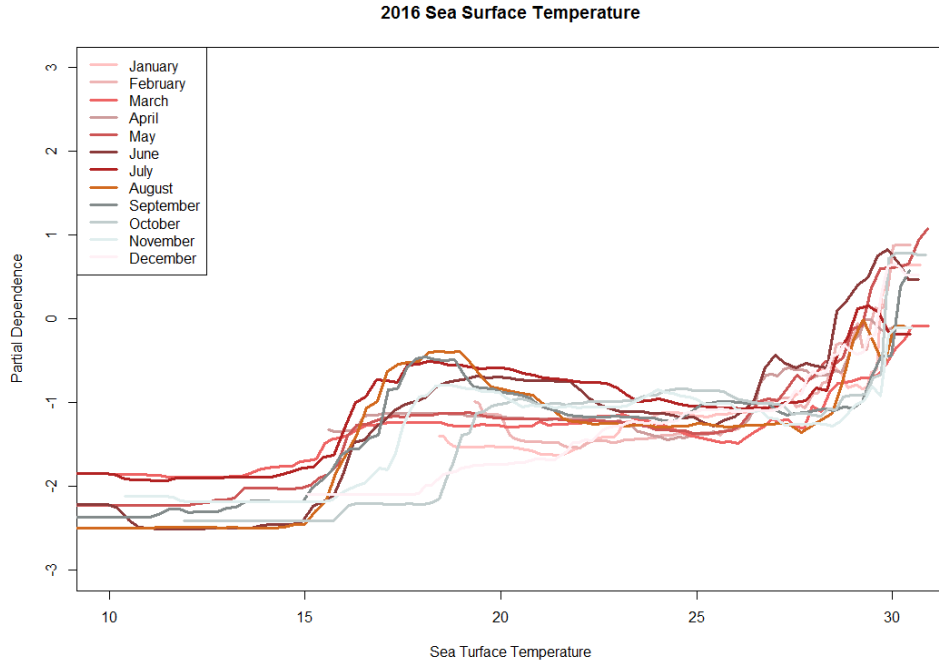
Fig. S11. The SST partial dependence plots from the monthly 2015 models.
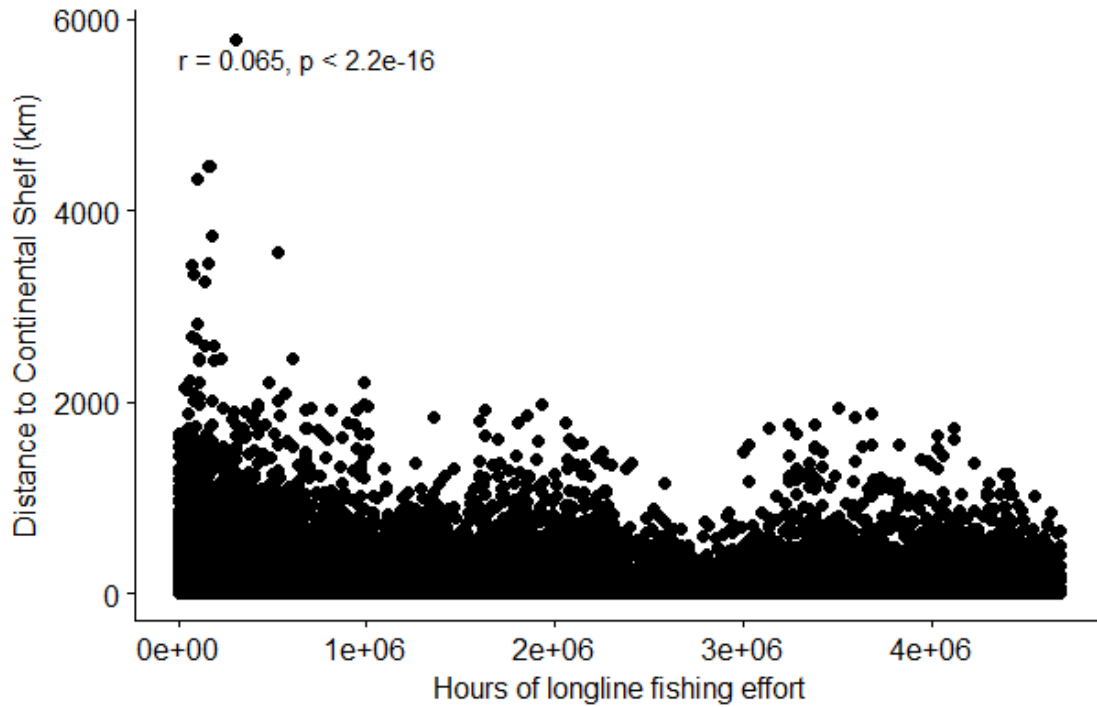
Fig. S12. The distribution of fishing effort intensity as a function of the Euclidean distance (kilometers) to the continental shelf. There is no correlation between distance to continental shelf break and fishing effort intensity in areas beyond national jurisdiction.
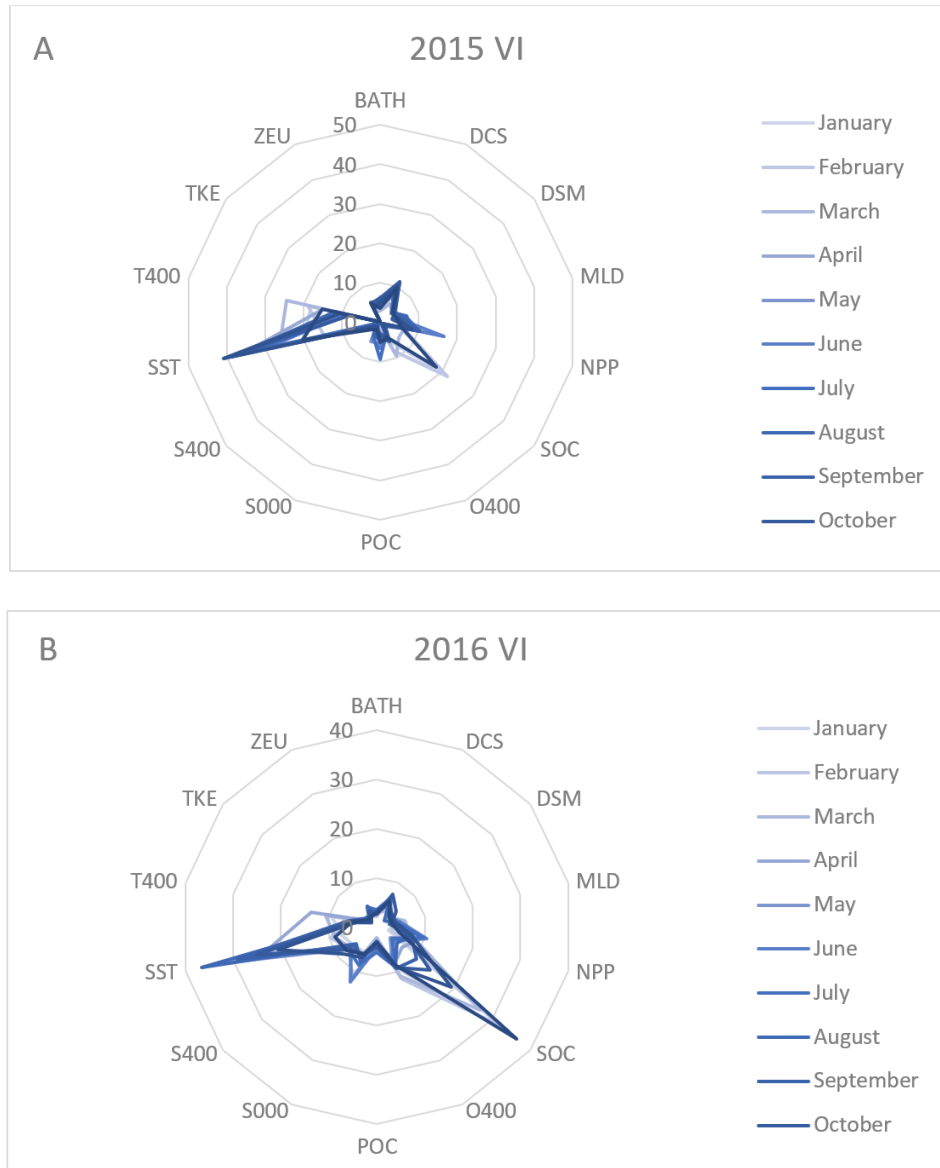
Fig. S13. Monthly variable importance scores for boosted regression trees using background pseudoabsence points from the entire high seas areas for 2015 and 2016. (**A**) 2015 and (**B**) 2016.

Supplementary Tables

Table S1. Various model performance indices of the monthly BRTs for 2015 and 2016. Model performance metrics include the Kappa statistic and the Sensitivity and Specificity values when a Receiver Operator Characteristic cutoff threshold is applied to each of the 24 monthly habitat suitability layers.

	2015				2016			
	Kappa	Sensitivity	Specificity	AUC	Kappa	Sensitivity	Specificity	AUC
January	0.70	0.88	0.85	0.86	0.81	0.93	0.90	0.91
February	0.66	0.86	0.83	0.84	0.81	0.94	0.90	0.91
March	0.68	0.83	0.86	0.84	0.79	0.90	0.90	0.90
April	0.68	0.87	0.84	0.85	0.78	0.91	0.90	0.90
May	0.70	0.81	0.89	0.85	0.79	0.89	0.91	0.90
June	0.70	0.92	0.83	0.87	0.79	0.94	0.88	0.91
July	0.72	0.90	0.86	0.87	0.80	0.93	0.90	0.92
August	0.73	0.91	0.86	0.88	0.83	0.95	0.91	0.91
September	0.70	0.87	0.86	0.86	0.81	0.92	0.91	0.91
October	0.68	0.84	0.86	0.85	0.80	0.93	0.90	0.91
November	0.72	0.85	0.89	0.86	0.82	0.92	0.91	0.91
December	0.72	0.89	0.86	0.87	0.80	0.93	0.89	0.91

Table S2. Various model performance indices of the monthly BRTs for 2015 and 2016. Model performance metrics include the Kappa statistic and the Sensitivity and Specificity values when **a mean probability distribution** cutoff threshold is applied to each of the 24 monthly habitat suitability layers.

	2015				2016			
	Kappa	Sensitivity	Specificity	AUC	Kappa	Sensitivity	Specificity	AUC
January	0.52	0.97	0.64	0.8	0.66	0.99	0.75	0.87
February	0.47	0.97	0.60	0.78	0.71	0.99	0.80	0.89
March	0.47	0.96	0.59	0.77	0.60	0.98	0.71	0.84
April	0.53	0.97	0.65	0.81	0.53	0.99	0.64	0.81
May	0.54	0.97	0.66	0.81	0.67	0.98	0.77	0.87
June	0.61	0.96	0.73	0.84	0.70	0.98	0.79	0.88
July	0.63	0.96	0.74	0.85	0.72	0.98	0.81	0.89
August	0.60	0.97	0.71	0.84	0.69	0.98	0.78	0.88
September	0.56	0.97	0.67	0.82	0.65	0.99	0.74	0.86
October	0.48	0.98	0.60	0.78	0.68	0.98	0.77	0.87
November	0.51	0.98	0.62	0.80	0.66	0.99	0.75	0.87
December	0.60	0.96	0.72	0.84	0.63	0.98	0.73	0.85

Table S3. Various model performance indices of the temporally averaged BRT model. Model performance metrics include the Kappa statistic, the Sensitivity and Specificity values and model accuracy score when a Receiver Operator Characteristic cutoff threshold is applied to each of the 24 monthly habitat suitability layers.

	2015				2016			
	Kappa	Sensitivity	Specificity	Accuracy	Kappa	Sensitivity	Specificity	Accuracy
January	0.38	0.89	0.79	0.80	0.57	0.87	0.88	0.88
February	0.39	0.88	0.79	0.80	0.51	0.93	0.83	0.84
March	0.39	0.86	0.79	0.79	0.42	0.90	0.79	0.80
April	0.35	0.84	0.76	0.77	0.45	0.89	0.80	0.81
May	0.41	0.86	0.79	0.80	0.50	0.91	0.80	0.82
June	0.43	0.87	0.80	0.81	0.54	0.87	0.85	0.86
July	0.43	0.86	0.81	0.81	0.55	0.88	0.85	0.85
August	0.40	0.87	0.78	0.79	0.54	0.90	0.85	0.86
September	0.38	0.86	0.78	0.79	0.45	0.93	0.79	0.81
October	0.37	0.86	0.77	0.78	0.51	0.86	0.85	0.85
November	0.36	0.86	0.78	0.79	0.49	0.88	0.84	0.85
December	0.41	0.85	0.80	0.81	0.54	0.91	0.85	0.86

Table S4. Various model performance indices of the temporally averaged BRT model. Model performance metrics include the Kappa statistic, the Sensitivity and Specificity values and model accuracy score when a mean probability distribution cutoff threshold is applied to each of the 24 monthly habitat suitability layers.

	2015				2016			
	Kappa	Sensitivity	Specificity	Accuracy	Kappa	Sensitivity	Specificity	Accuracy
January	0.29	0.93	0.75	0.72	0.38	0.93	0.75	0.77
February	0.31	0.93	0.75	0.73	0.39	0.93	0.75	0.77
March	0.32	0.88	0.74	0.74	0.35	0.88	0.74	0.76
April	0.28	0.90	0.74	0.70	0.37	0.90	0.74	0.76
May	0.34	0.90	0.76	0.74	0.43	0.90	0.76	0.78
June	0.36	0.91	0.75	0.75	0.40	0.91	0.75	0.77
July	0.33	0.91	0.75	0.74	0.42	0.91	0.75	0.77
August	0.33	0.94	0.75	0.73	0.40	0.94	0.75	0.77
September	0.32	0.92	0.74	0.73	0.37	0.92	0.74	0.76
October	0.32	0.88	0.73	0.73	0.35	0.88	0.73	0.75
November	0.28	0.90	0.74	0.72	0.35	0.90	0.74	0.76
December	0.32	0.93	0.75	0.73	0.39	0.93	0.75	0.77

Table S5. Results from the Wilcoxon signed-rank test comparing the performance of monthly models to the temporally averaged model. Results show that 14 of the 16 model performance metrics were statistically dissimilar between the monthly models and the temporally averaged model using all 24 months of data. The V-statistic is the sum of ranks assigned to the differences with positive sign.

Year	Threshold	Performance metric	V	p-value
2015	MPD	Accuracy	78	4.90E-04
2015	MPD	Kappa	78	4.90E-04
2015	MPD	Sensitivity	0	4.90E-04
2015	MPD	Specificity	78	4.90E-04
2015	ROC	Accuracy	78	4.90E-04
2015	ROC	Kappa	78	4.90E-04
2015	ROC	Sensitivity	55	2.30E-01
2015	ROC	Specificity	78	4.90E-04
2016	MPD	Accuracy	78	4.90E-04
2016	MPD	Kappa	48	5.20E-01
2016	MPD	Sensitivity	78	4.90E-04
2016	MPD	Specificity	78	4.90E-04
2016	ROC	Accuracy	78	4.90E-04
2016	ROC	Kappa	72	6.90E-03
2016	ROC	Sensitivity	78	4.90E-04
2016	ROC	Specificity	78	4.90E-04

Table S6. Amount of fundamental niche occupied by pelagic longliners. This estimate of occupancy is calculated as the percentage of monthly suitable habitat where fishing was detected in 2015 and 2016 using the mean probability distribution and Receiver Operator Characteristic thresholds.

Year	Threshold	January	February	March	April	May	June	July	August	September	October	November	December
2015	ROC	49%	51%	59%	54%	61%	49%	55%	54%	54%	60%	59%	55%
2015	MPD	31%	35%	38%	38%	39%	40%	42%	39%	37%	39%	35%	36%
2016	ROC	62%	57%	68%	72%	66%	61%	67%	63%	66%	54%	66%	68%
2016	MPD	45%	45%	47%	49%	49%	49%	53%	49%	49%	41%	40%	48%

Table S7. The 2015 VI scores. Each of the 14 biophysical covariates assessed in the monthly environmental niche models was assigned a variable importance score, the scores in each month add up to 100.

2015	January	February	March	April	May	June	July	August	September	October	November	December
BATH	6.87	6.66	6.31	7.36	5.83	7.71	6.26	5.75	6.96	8.55	9.16	5.40
DCS	16.73	16.64	11.37	9.92	9.67	9.23	10.46	14.58	10.89	16.01	15.40	14.48
DSM	7.97	6.68	7.53	5.07	6.25	5.72	5.05	4.98	4.95	6.04	4.74	4.16
MLD	8.10	10.20	11.79	9.93	8.37	6.01	7.49	11.45	9.42	8.56	6.06	7.18
NPP	7.69	8.64	8.92	10.65	7.64	11.62	9.28	8.64	10.18	10.06	9.51	8.67
SOC	0.17	0.09	0.47	0.94	0.70	0.50	0.79	0.67	0.84	0.88	0.34	5.87
O400	7.83	11.31	15.20	5.97	5.26	5.35	5.10	4.48	5.67	8.25	6.99	5.72
POC	2.42	2.99	1.95	5.22	6.39	8.18	10.60	6.72	3.77	3.28	6.72	7.28
S000	4.28	2.95	4.44	2.65	5.88	5.73	5.46	3.80	3.29	2.68	2.52	2.67
S400	1.23	1.88	0.83	1.64	1.01	1.68	1.23	0.35	0.84	2.37	1.99	3.00
SST	15.45	14.29	8.93	19.29	22.94	21.92	24.00	22.19	23.76	18.95	19.56	13.72
T400	16.18	12.89	18.03	14.32	12.93	11.93	9.33	12.21	12.70	9.21	11.62	14.03
TKE	0.02	0.03	0.07	0.00	0.00	0.00	0.13	0.02	0.12	0.16	0.02	0.14
ZEU	5.07	4.75	4.17	7.03	7.13	4.44	4.81	4.16	6.60	5.00	5.39	7.66

Table S8. The 2016 VI scores. Each of the 14 biophysical covariates assessed in the monthly environmental niche models was assigned a variable importance score, the scores in each month add up to 100.

2016	January	February	March	April	May	June	July	August	September	October	November	December
BATH	3.74	3.83	3.61	4.37	4.25	3.37	3.38	3.35	3.74	4.09	3.74	3.05
DCS	13.06	12.08	6.56	7.03	6.55	5.41	8.19	8.53	7.36	10.03	9.13	10.77
DSM	3.90	3.97	3.95	3.53	3.88	3.35	3.51	2.60	3.97	5.52	4.65	3.95
MLD	7.15	8.00	8.62	5.32	4.96	6.84	5.23	4.72	4.27	4.01	4.89	6.65
NPP	3.82	4.40	7.80	9.75	10.55	8.54	9.14	8.20	8.09	8.14	8.75	5.81
SOC	13.88	10.85	6.81	7.13	4.19	5.86	4.68	5.42	9.46	10.88	8.20	10.53
O400	13.06	16.14	15.65	12.52	10.12	8.35	10.24	10.85	11.16	12.55	12.76	12.74
POC	4.69	3.49	2.95	4.83	3.31	4.85	4.64	2.87	4.43	4.04	3.00	3.78
S000	6.92	6.22	8.36	8.08	13.03	12.95	10.67	9.35	8.97	6.80	9.57	9.29
S400	7.28	7.94	9.71	9.24	8.71	7.83	7.87	6.53	7.29	9.40	9.16	8.48
SST	8.36	10.81	12.54	11.41	17.55	21.16	19.32	25.07	17.38	11.40	16.03	13.26
T400	8.78	6.49	7.55	9.47	6.50	5.83	5.64	5.20	7.19	6.71	5.55	5.90
TKE	1.80	1.92	2.46	2.45	1.75	2.21	1.91	1.40	2.66	2.39	1.39	2.67
ZEU	3.55	3.87	3.44	4.88	4.64	3.46	5.57	5.90	4.01	4.04	3.19	3.11

Table S9. Average 2016 model performance metrics using different environmental variables. The three models assessed include a model with only static variables (n=3), a model with only variables (n=11) and a final model which included both types (n=14).

	Kappa	Accuracy	Sensitivity	Specificity
Static	0.26	0.61	0.77	0.54
Dynamic	0.53	0.76	0.94	0.67
Static + dynamic	0.66	0.91	0.98	0.75

Table S10. Description of the variable type and source for each of the 14 biophysical and physiographic predictors.

Variable	Code	Variable type	Source
Sea surface temperature	SST	Monthly climatology	Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, C. R. Paver, J. R. Reagan, D. R. Johnson, M. Hamilton, and D. Seidov, 2013. World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp.
Temperature at 400 meters	T400	Monthly climatology	Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, H. E. Garcia, O. K. Baranova, M. M. Zweng, C. R. Paver, J. R. Reagan, D. R. Johnson, M. Hamilton, and D. Seidov, 2013. World Ocean Atlas 2013, Volume 1: Temperature. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 73, 40 pp.
Turbulent kinetic energy	TKE	Monthly average	Roberts, J.J., Best, B.D., Dunn, D.C., Treml, E.A. and Halpin, P.N., 2010. Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environmental Modelling & Software, 25(10), pp.1197-1207.
Particulate organic carbon	POC	Monthly average	www.hermes.acri.fr
Net primary productivity	NPP	Monthly average	www.orca.science.oregonstate.edu
Mixed layer depth	MLD		de Boyer Montégut, C., Madec, G., Fischer, A.S., Lazar, A. and Iudicone, D., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile?based climatology. Journal of Geophysical Research: Oceans, 109(C12).
Surface oxygen concentration	SOC	Monthly climatology	Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, O.K. Baranova, M.M. Zweng, J.R. Reagan, D.R. Johnson, 2014. World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 75, 27 pp.
Oxygen concentration at 400 meters	O400	Monthly climatology	Garcia, H. E., R. A. Locarnini, T. P. Boyer, J. I. Antonov, O.K. Baranova, M.M. Zweng, J.R. Reagan, D.R. Johnson, 2014. World Ocean Atlas 2013, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 75, 27 pp.
Sea surface salinity	SSS	Monthly climatology	Zweng, M.M, J.R. Reagan, J.I. Antonov, R.A. Locarnini, A.V. Mishonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, D.R. Johnson, D.Seidov, M.M. Biddle, 2013. World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.
Salinity at 400 meters	S400	Monthly climatology	Zweng, M.M, J.R. Reagan, J.I. Antonov, R.A. Locarnini, A.V. Mishonov, T.P. Boyer, H.E. Garcia, O.K. Baranova, D.R. Johnson, D.Seidov, M.M. Biddle, 2013. World Ocean Atlas 2013, Volume 2: Salinity. S. Levitus, Ed., A. Mishonov Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.
Euphotic depth	ZEU	Monthly average	www.hermes.acri.fr
Bathymetry	BATH	Static	Arndt, J.E., Schenke, H.W., Jakobsson, M., Nitsche, F.O., Buys, G., Goleby, B., Rebesco, M., Bohoyo, F., Hong, J., Black, J. and Greku, R., 2013. The International Bathymetric Chart of the Southern Ocean (IBCSO) Version 1.0—A new bathymetric compilation covering circum?Antarctic waters. Geophysical Research Letters, 40(12), pp.3111-3117.
Distance to continental shelf	DCS	Static	http://www.continentalshelf.org/
Distance to seamount	DSM	Static	Wessel, P., Sandwell, D.T. and Kim, S.-S., 2010, The global seamount census, Oceanography, 23(1), 24-33.

Table S11. The number of presence and pseudoabsence points in 2015. Each of the 12 monthly environmental niche models in 2015 had double the number of pseudoabsences as well as the percentage of the total number of presence points in each month.

2015	Presence points	Proportion of presence points	Pseudoabsences points
January	2395	7.1%	4790
February	2532	7.5%	5064
March	2820	8.4%	5640
April	2868	8.5%	5736
May	3009	8.9%	6018
June	3054	9.1%	6108
July	3012	8.9%	6024
August	2881	8.5%	5762
September	2890	8.6%	5780
October	2919	8.7%	5838
November	2561	7.6%	5122
December	2772	8.2%	5544

Table S12. The number of presence and pseudoabsence points in 2016. Each of the 12 monthly environmental niche models in 2016 had double the number of pseudoabsences as well as the percentage of the total number of presence points in each month.

2016	Presence points	Proportion of presence points	Pseudoabsences points
January	2956	7.8%	5912
February	3014	8.0%	6028
March	3016	8.0%	6032
April	3216	8.5%	6432
May	3534	9.3%	7068
June	3401	9.0%	6802
July	3563	9.4%	7126
August	3176	8.4%	6352
September	3075	8.1%	6150
October	3083	8.1%	6166
November	2817	7.4%	5634
December	3014	8.0%	6028