Table S1

All the covariates used in the Ensemble Random Forests models, their dynamism (Dyn.) indicating static, temporally dynamic (T. Dyn.), or spatiotemporally dynamic (ST. Dyn.), the data source, the start of the dataset if dynamic as well as their spatial and temporal resolution (Res.) are provided.

Covariate	Dyn.	Source	Data start	Spatial Res.	Temporal Res.
Bathymetry	Static	GEBCO [1]	-	0.0083 °	-
Distance to shore	Static	NASA GSFC [2]	-	0.04 °	-
Distance to seamount	Static	Yesson et al. (2011) [3]	-	0.0083 °	-
Distance to seamount (0-50m)	Static	Derived	-	0.0083 °	-
Distance to seamount (50-100m)	Static	Derived	-	0.0083 °	-
Distance to seamount (100-150m)	Static	Derived	-	0.0083 °	-
Distance to seamount (150-200m)	Static	Derived	-	0.0083 °	-
Lunar phase	T. Dyn.	Lunar package [4]	-	-	1 day
El Niño Southern Oscillation	T. Dyn.	NOAA NWS CPC [5a]	-	Niño 3+4 [5b]	7 day
Sea Level Anomaly	ST. Dyn.	PODAAC-JPL [6]	10/1/1992	0.17 °	5 day
N-S Current Speed (m/s)	ST. Dyn.	PODAAC-JPL [7]	10/12/1992	0.33 °	5 day
E-W Current Speed (m/s)	ST. Dyn.	PODAAC-JPL [7]	10/12/1992	0.33 °	5 day
Current Speed (m/s)	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
Current Direction	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
Distance to current front	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
Distance to log(current) front	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
Current vorticity	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
Current divergence	ST. Dyn.	Derived	10/12/1992	0.33 °	5 day
log(chlorophyll-α)	ST. Dyn.	OceanData - GSFC [8]	7/4/2002	4km	8 day
Strength chl- α front	ST. Dyn.	Derived	7/4/2002	4km	8 day
Distance to chl- α front	ST. Dyn.	Derived	7/4/2002	4km	8 day
Sea Surface Temperature	ST. Dyn.	OceanData - GSFC [8]	7/4/2002	4km	8 day
Strength SST front	ST. Dyn.	Derived	7/4/2002	4km	8 day
Distance to SST front	ST. Dyn.	Derived	7/4/2002	4km	8 day
Depth to minimum DO	ST. Dyn.	Park et al. (2018) [9]	1/1990	1 °	1 month
Mixed Layer Depth	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Isothermal Layer Depth	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Top of Thermocline Layer	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Barrier Thickness	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Temperature Inversion Depth	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Temperature at mixing layer	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
Salinity at mixing layer	ST. Dyn.	NCEP GODAS [10]	1/2005	1 °	1 month
N-S wind speed (m/s)	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
E-W wind speed (m/s)	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Wind speed (m/s)	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Wind direction	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Wind vorticity	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Wind divergence	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Distance to wind front	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day
Distance to log(wind) front	ST. Dyn.	NOAA NCEI [11]	07/1987	0.25 °	1 day

[1] https://www.gebco.net/data_and_products/gridded_bathymetry_data/
[2] https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/
[3] https://doi.pangaea.de/10.1594/PANGAEA.757564
[4] https://cran.r-project.org/web/packages/lunar/index.html
[5a] http://www.cpc.ncep.noaa.gov/data/indices/wksst8110.for
[5b] Used Niño 3+4 zone for SST Anomaly as it is closest to Hawaii
[6]https://podaac.jpl.nasa.gov/dataset/SEA_SURFACE_HEIGHT_ALT_GRIDS_L4_2SATS_5DAY_6THDEG_V_JPL1609
[7] https://podaac.jpl.nasa.gov/dataset/OSCAR_L4_OC_third-deg
[8] https://oceancolor.gsfc.nasa.gov/data/aqua/
[9] https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1002/2017MS001223
[10] http://apdrc.soest.hawaii.edu/projects/argo/
[11] https://www.ncei.noaa.gov/thredds/catalog/uv/daily/2000s/catalog.html

Table S2

Internal model performance metrics, area-under-the-curve (AUC), root mean squared error (RMSE), and the true skill statistic (TSS), for the spatial covariation case study species using Random Forest (RF), RF with downsampling (RF-DS), RF with synthetic minority over-sampling technique (RF-SMOTE), and Ensemble Random Forests (ERF).

Performance Metric	AUC	RMSE	TSS
Giant Manta Ray			
RF	0.61	0.03	0.28
RF-DS	0.91	0.39	0.81
RF-SMOTE	0.95	0.33	0.91
ERF	0.78	0.12	1.00
Scalloped Hammerhead			
RF	0.63	0.03	0.33
RF-DS	0.96	0.38	0.91
RF-SMOTE	0.94	0.31	0.81
ERF	0.75	0.12	1.00
False Killer Whale			
RF	0.55	0.04	0.14
RF-DS	0.61	0.41	0.31
RF-SMOTE	0.96	0.33	0.77
ERF	0.55	0.13	1.00

Fig. S1

(Left) The distribution of probability of presence values for each map, p_n, with the corresponding Moran's I value, a measure of spatial autocorrelation (Moran, 1950), displayed in the legend text. (Right) Ripley's K functions, a measure of spatial clustering (Ripley, 1988), for each probability of presence map, p_n. Observed Ripley's K functions (solid line) that are greater than the theoretical (dashed line) and outside the theoretical envelope (gray polygon) are more clustered than expected.

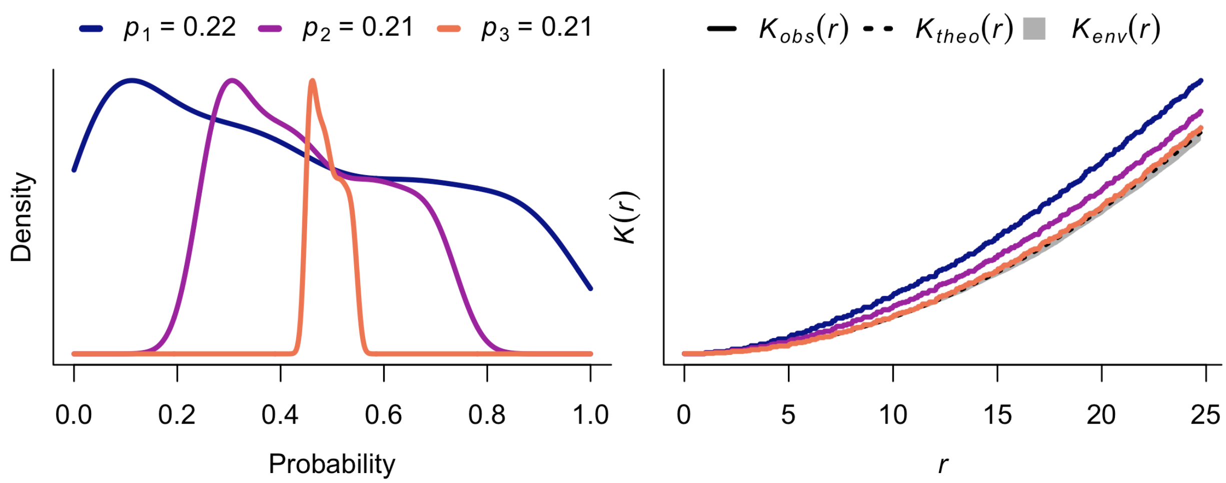
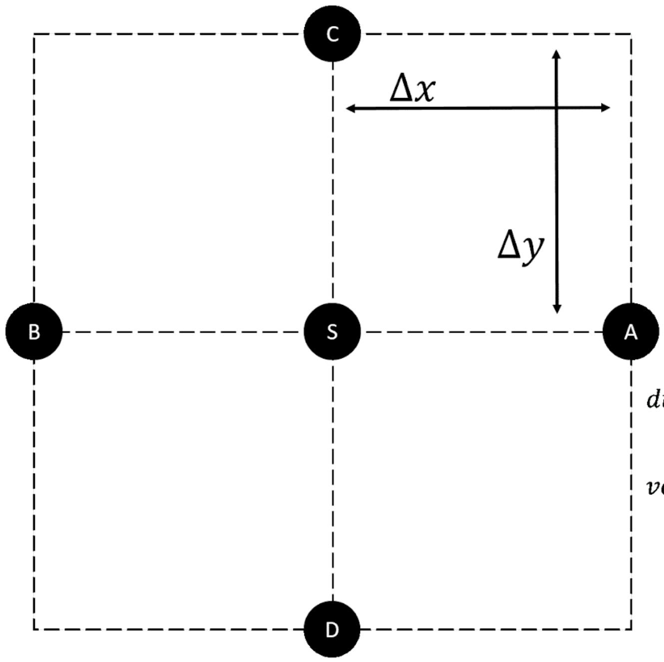


Fig. S2

The derivation of current or wind divergence and vorticity from *u* and *v* direction vectors.



U direction current: E-W
V direction current: N-S

$$divergence(S) = \frac{u(A) - u(B)}{2\Delta x} + \frac{v(C) - v(D)}{2\Delta y}$$

$$vorticity(S) = \frac{v(A) - v(B)}{2\Delta x} - \frac{u(C) - u(D)}{2\Delta y}$$

Fig. S3

Ripley's K functions, a measure of spatial clustering (Ripley, 1988), for each of the point patterns using different detection probabilities (higher detection probabilities are indicated by warmer colors). Observed Ripley's K functions (solid line) that are greater than the theoretical (dashed line) and outside the theoretical envelope (gray polygon) are more clustered than expected.