

Vulnerability of the Critically Endangered leatherback turtle to fisheries bycatch in the eastern Pacific Ocean. I. A machine-learning species distribution model

Table S1. Performance metrics of the full models (all environmental and static variables) after being assessed by additional complementary hold-out cross-validation methods (included, for reference, results of the main K-fold cross-validation method) for the different presence-absence ratios (e.g., 50:50 presence to absence ratio = 50% presence data and 50% absence data in the model). El Nino/La Nina refers to the performance metrics obtained by the model after being trained with data outside El Nino or La Nina periods and tested against them. Similarly, inshore and offshore refer to the performance metrics obtained by the model after being trained with data from inshore or offshore (different limits of 90W and 100W explored) and tested against the other data.

		n.trees	Dev	AUC	TSS
50:50 n = 2176	K-fold	3650	59.37	0.92	0.71
	El Nino	3400	58.04	0.94	0.80
	La Nina	3000	59.36	0.87	0.62
	Inshore (90W)	1500	58.64	0.67	0.29
	Inshore (100W)	1850	58.86	0.67	0.29
	Offshore (90W)	1450	36.78	0.65	0.31
	Offshore (100W)	1150	39.57	0.68	0.36
	33:66 n = 3264	K-fold	2650	61.54	0.94
El Nino		2500	60.69	0.94	0.78
La Nina		2500	63.38	0.87	0.64
Inshore (90W)		1650	63.28	0.65	0.22
Inshore (100W)		1550	61.35	0.67	0.30
Offshore (90W)		2050	39.60	0.70	0.29
Offshore (100W)		1150	35.84	0.70	0.30
25:75 n = 4352		K-fold	3250	60.42	0.96
	El Nino	2700	57.64	0.95	0.79
	La Nina	3050	62.06	0.86	0.59
	Inshore (90W)	1600	58.95	0.66	0.23
	Inshore (100W)	2200	61.79	0.68	0.29
	Offshore (90W)	1750	45.09	0.72	0.33
	Offshore (100W)	1250	33.73	0.66	0.23
	20:80 n = 5440	K-fold	3350	59.14	0.95
El Nino		3150	58.48	0.94	0.76
La Nina		2750	59.37	0.88	0.63
Inshore (90W)		2350	62.27	0.69	0.30
Inshore (100W)		2600	62.64	0.71	0.36
Offshore (90W)		1800	43.62	0.68	0.27
Offshore (100W)		1150	42.28	0.73	0.36

15:85 n = 7250	K-fold	3650	58.96	0.94	0.76
	El Nino	3800	58.87	0.95	0.81
	La Nina	3550	61.69	0.85	0.59
	Inshore (90W)	2800	64.63	0.65	0.27
	Inshore (100W)	2650	62.28	0.64	0.27
	Offshore (90W)	1050	34.64	0.63	0.18
	Offshore (100W)	1300	32.76	0.63	0.20
10:90 n = 10880	K-fold	4800	60.16	0.95	0.78
	El Nino	4150	58.29	0.95	0.81
	La Nina	4550	62.17	0.87	0.65
	Inshore (90W)	3700	65.70	0.65	0.27
	Inshore (100W)	3400	62.28	0.65	0.28
	Offshore (90W)	1350	38.64	0.68	0.26
	Offshore (100W)	1500	34.82	0.62	0.22
5:95 n = 21760	K-fold	5300	56.36	0.95	0.78
	El Nino	4600	54.88	0.94	0.79
	La Nina	5650	59.92	0.86	0.60
	Inshore (90W)	4050	61.18	0.60	0.16
	Inshore (100W)	4450	59.93	0.61	0.21
	Offshore (90W)	750	28.85	0.66	0.28
	Offshore (100W)	1100	38.20	0.58	0.14
2.5:97.5 n = 43520	K-fold	5250	52.08	0.95	0.77
	El Nino	5600	51.88	0.95	0.81
	La Nina	5400	54.28	0.85	0.60
	Inshore (90W)	4850	56.76	0.58	0.17
	Inshore (100W)	5600	56.64	0.57	0.16
	Offshore (90W)	1250	30.86	0.65	0.22
	Offshore (100W)	1200	34.94	0.59	0.19
1:99 n = 108800	K-fold	6650	49.24	0.94	0.78
	El Nino	6250	49.12	0.94	0.79
	La Nina	6950	52.65	0.84	0.58
	Inshore (90W)	5450	55.78	0.60	0.16
	Inshore (100W)	5750	53.47	0.57	0.14
	Offshore (90W)	1050	31.38	0.68	0.30
	Offshore (100W)	1050	36.22	0.60	0.20
0.5:99.5 n = 217600	K-fold	7000	40.57	0.94	0.76
	El Nino	6600	47.13	0.94	0.79
	La Nina	6400	48.37	0.84	0.56
	Inshore (90W)	5850	52.23	0.59	0.16
	Inshore (100W)	6450	50.89	0.55	0.10
	Offshore (90W)	1400	34.81	0.68	0.28
	Offshore (100W)	800	32.02	0.56	0.19

0.19:99.81	K-fold	7000	40.57	0.94	0.76
n = 573889	El Nino	6050	37.64	0.93	0.78
	La Nina	6900	42.72	0.83	0.55
	Inshore (90W)	6250	45.83	0.56	0.13
	Inshore (100W)	7150	45.69	0.53	0.10
	Offshore (90W)	1300	37.22	0.70	0.27
	Offshore (100W)	1600	33.59	0.58	0.20

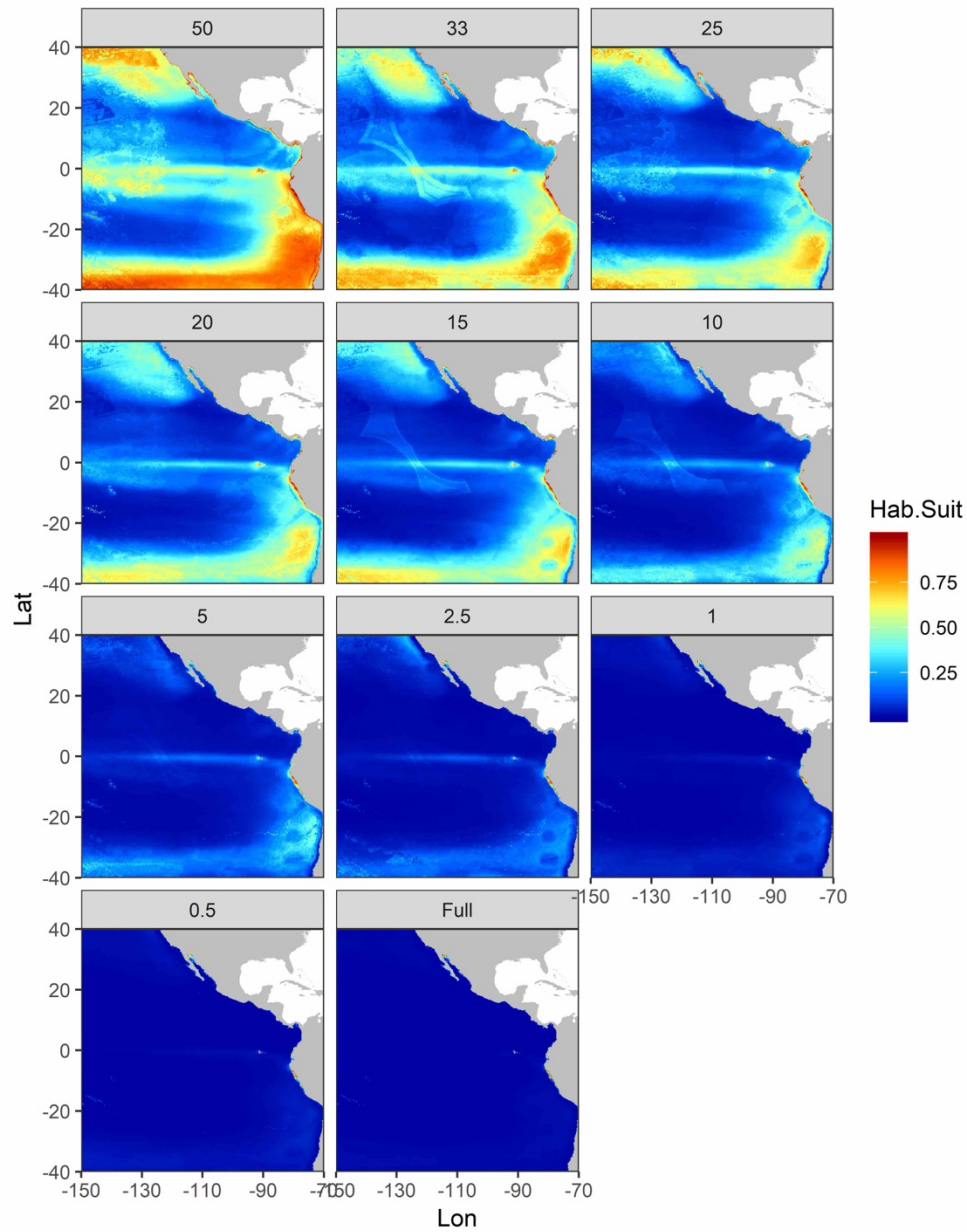


Figure S1. Model-specific predictions (averages for daily predictions for the period 2002–2020). The number on top of each panel denotes the ratio of presence to absences used to build the final model (e.g., 50 refers to a 50:50 presence to absence ratio, 33 to a 33:66 presence to absence ratio, etc.).

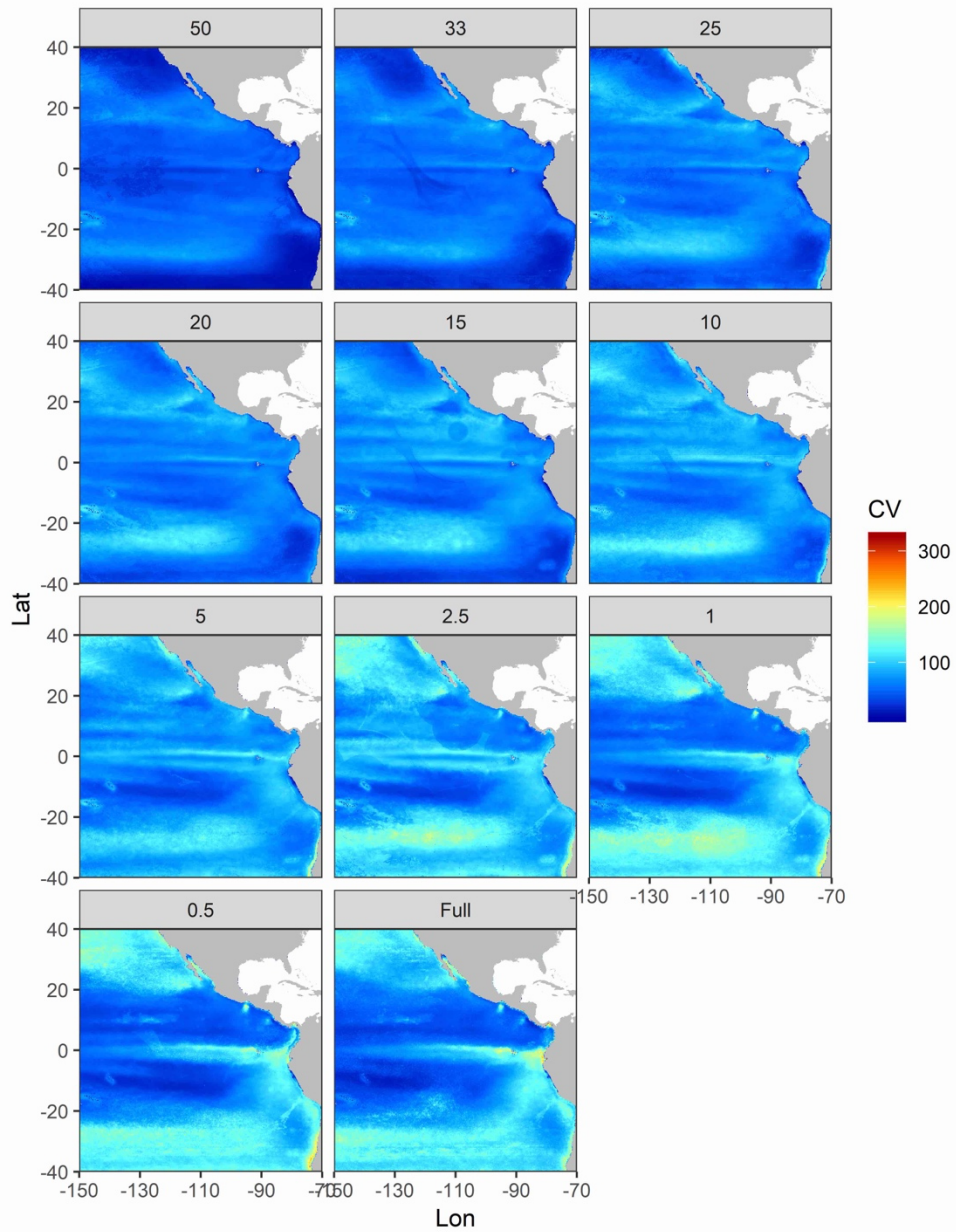


Figure S2. Model-specific coefficient of variation in predictions (averages for daily predictions for the period 2002-2020). The number on top of each panel denotes the ratio of presence to absences used to build the final model (e.g., 50 refers to a 50:50 presence to absence ratio, 33 to a 33:66 presence to absence ratio, etc.).

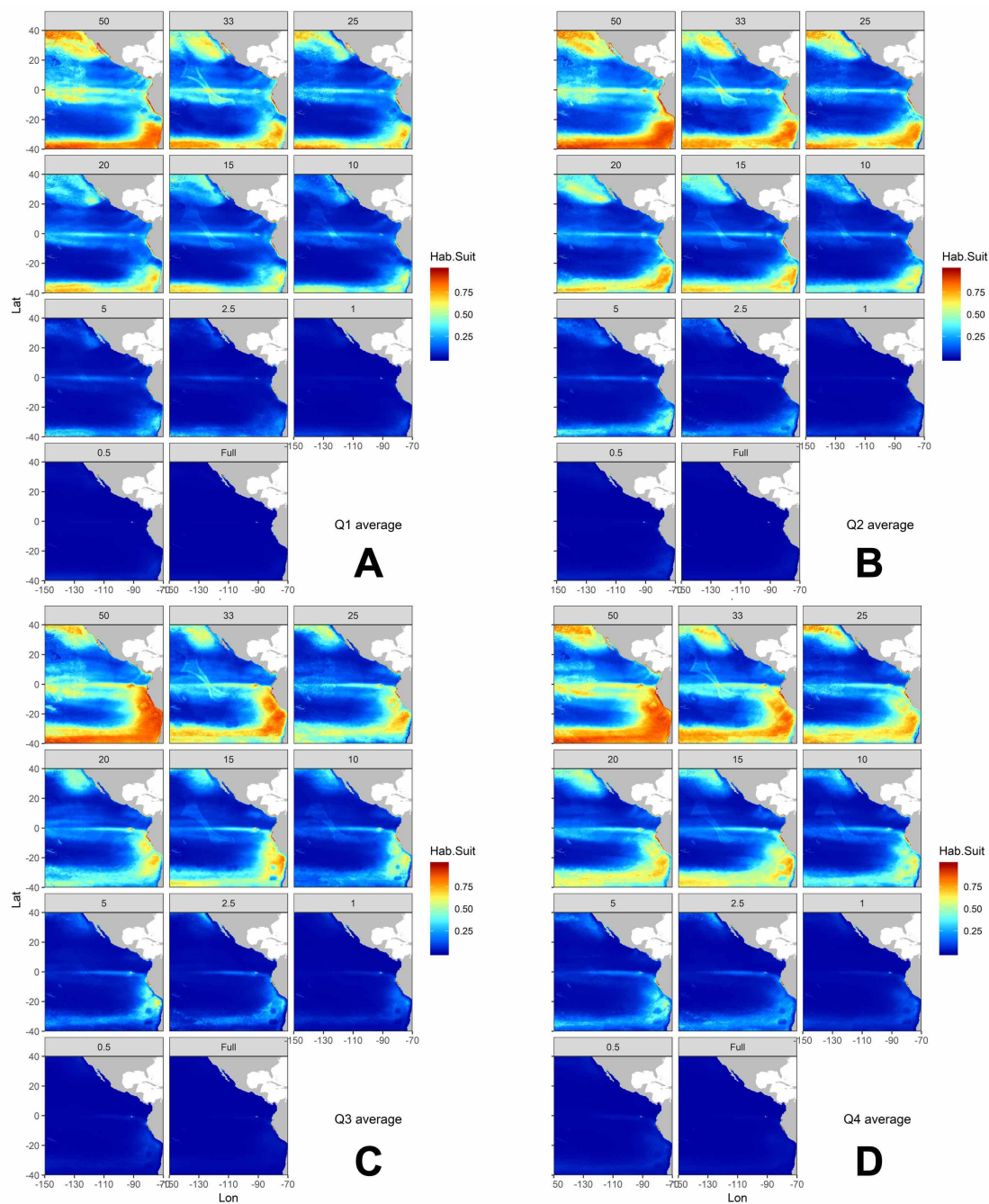


Figure S3. Model-specific predictions (averages for daily predictions for the period 2002-2020) for A) Q1, B) Q2, C) Q3 and D) Q4. The number on top of each panel denotes the ratio of presence to absences used to build the final model (e.g., 50 refers to a 50:50 presence to absence ratio, 33 to a 33:66 presence to absence ratio, etc.).

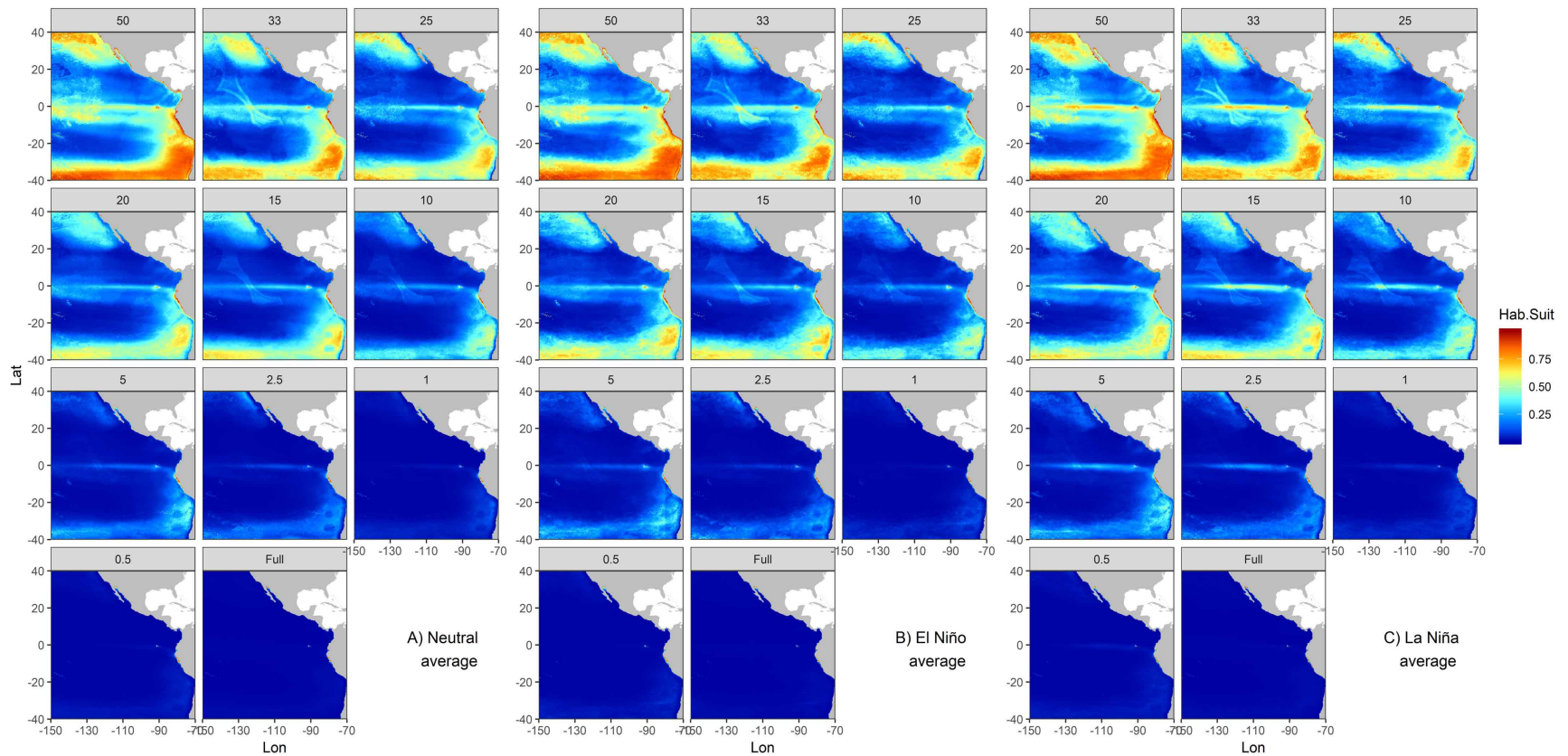


Figure S4. Model-specific predictions (averages for daily predictions for the period 2002-2020) for A) neutral years, B) El Niño years, and C) La Niña years. The number on top of each panel denotes the ratio of presence to absences used to build the final model (e.g., 50 refers to a 50:50 presence to absence ratio, 33 to a 33:66 presence to absence ratio, etc.).

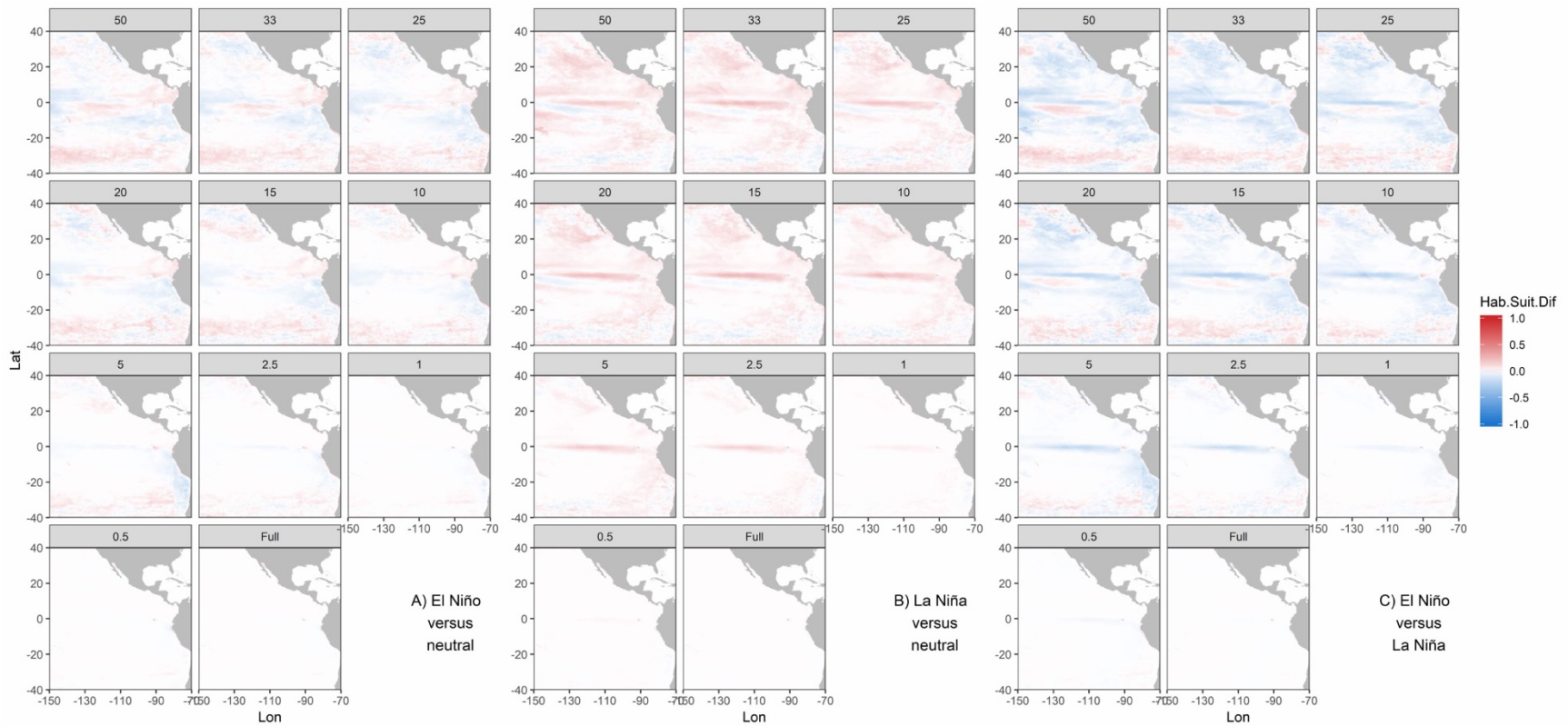


Figure S5. Relative differences in predicted leatherback habitat between A and B) El Niño and La Niña conditions compared to neutral years, and C) El Niño compared to La Niña. Red pixels represent net gain of predicted habitat while blue pixels represent net loss of predicted habitat of one phase relative to the other. The number on top of each panel denotes the ratio of presence to absences used to build the final model (e.g., 50 refers to a 50:50 presence to absence ratio, 33 to a 33:66 presence to absence ratio, etc.).

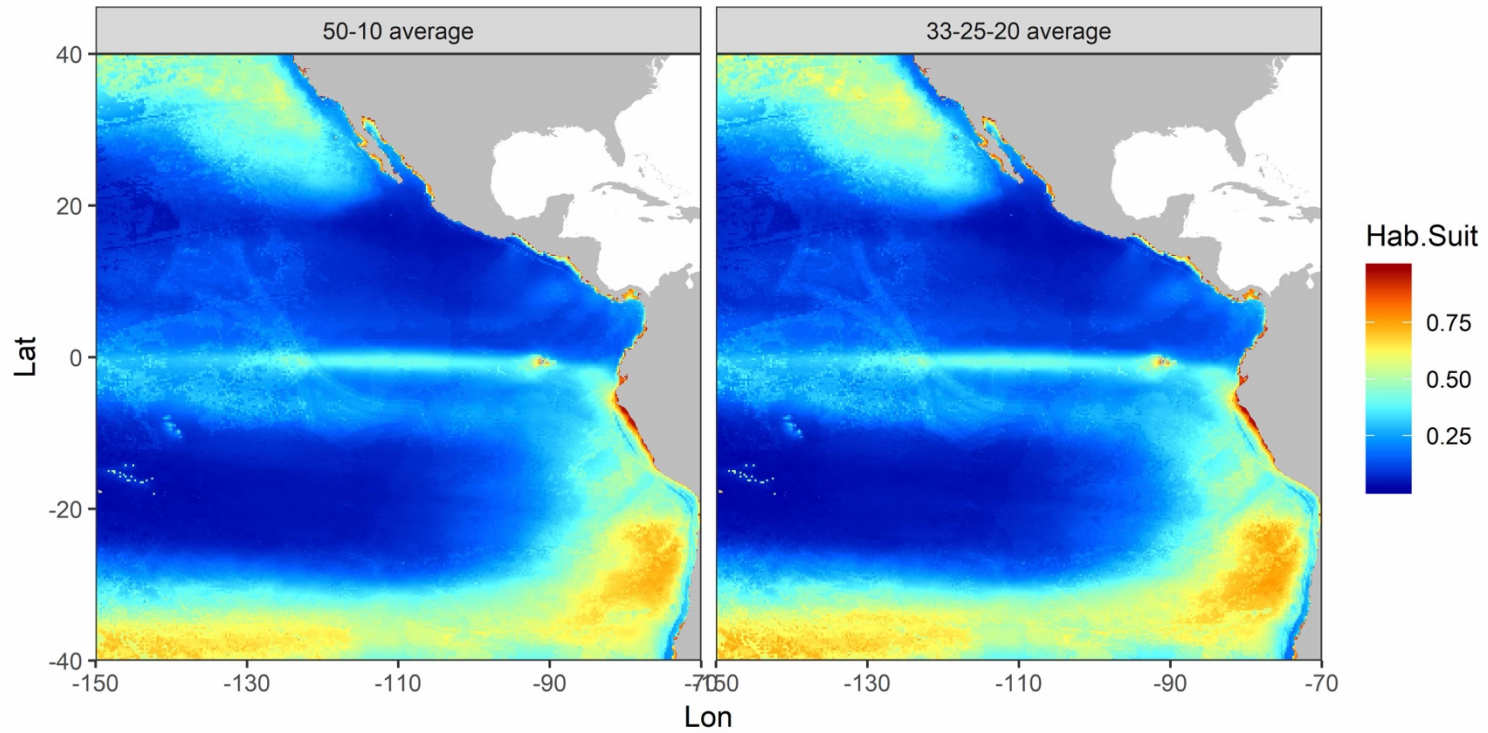


Figure S6. Average predictions from an ensemble of SDMs where the ratio of presence to absences were 50-10 (left panel) and 33-25-20 (right panel)

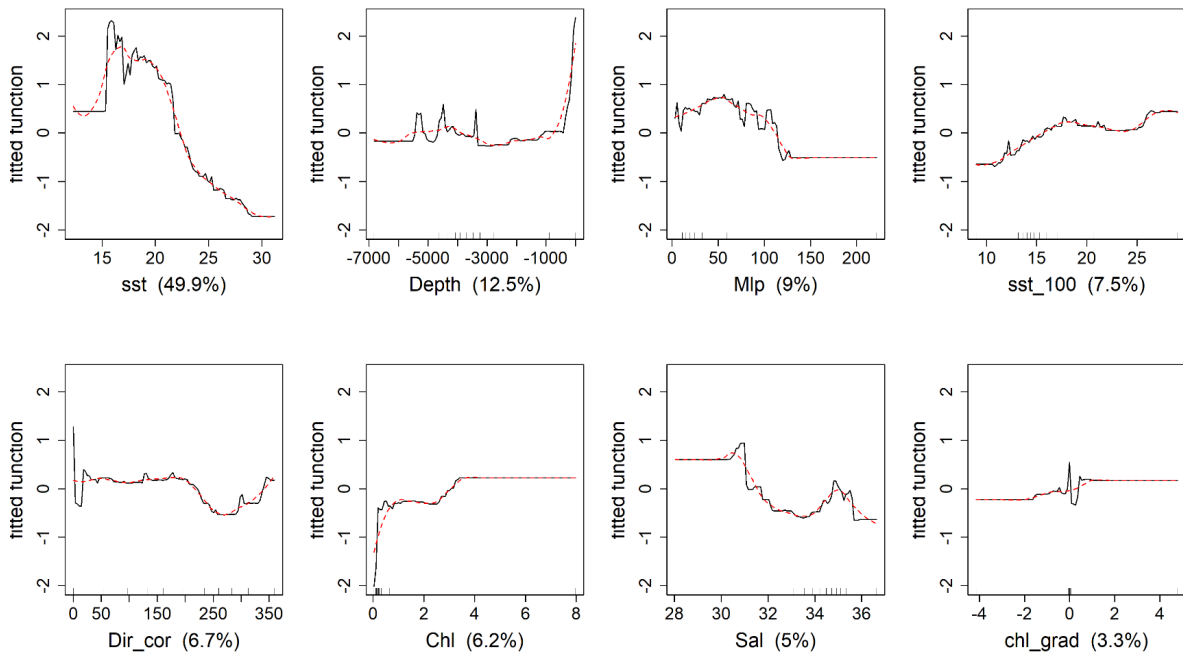


Figure S7. Partial dependence plots for leatherback turtle species distribution models (example of model with 25% of presences vs 75 absences – reference model)