

Appendices: Influence of changing temperature and upwelling intensity on Indian oil sardine (*Sardinella longiceps*) landings

Appendix A: Tests for prior season catch as covariate

Table A1. Model selection tests of time-dependency the log catch during spawning months using F-tests of nested linear models. S_t is the catch during the spawning period (Jul-Sep). N_t is the catch during the non-spawning period (Oct-Jun). S_{t-1} and N_{t-1} are the catch during the prior season during and after the spawning period respectively. S_{t-2} and N_{t-2} are the same for two seasons prior. Test A uses catch during the spawning period as the explanatory variable. Test B uses catch during the non-spawning period as the explanatory variable. The numbers in front of the model equation indicate the level of nestedness. For Test C, there are two nested model sets, each with a different model 3. The Naive model is a model that uses the previous data point in the time series as the prediction; thus the Naive model has no estimated parameters.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Naive Model 1984-2015 data							
$\ln(S_t) = \ln(S_{t-1}) + \epsilon_t$	32	1				122.85	1.599
Time dependency test A 1984-2015 data							
1. $\ln(S_t) = \alpha + \ln(S_{t-1}) + \epsilon_t$	31	0.992	-29			124.83	1.65
2. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	30	0.814	10.3	15.14	0.001	114.14	1.43
3. $\ln(S_t) = \alpha + \beta_1 \ln(S_{t-1}) + \beta_2 \ln(S_{t-2}) + \epsilon_t$	29	0.803	13.6	2.13	0.155	113.88	1.414
Time dependency test B 1984-2015 data							
1. $\ln(S_t) = \alpha + \ln(N_{t-1}) + \epsilon_t$	31	0.856	14.2			111.78	1.346
2. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	30	0.794	22.2	4.06	0.053	109.59	1.308
3. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(N_{t-2}) + \epsilon_t$	29	0.797	19.6	0.01	0.919	111.57	1.346
Time dependency test C 1984-2015 data							
1. $\ln(S_t) = \alpha + \ln(N_{t-1}) + \epsilon_t$	31	0.856	14.2			111.78	1.346
2. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	30	0.794	22.2	4.08	0.053	109.59	1.308
3a. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-1}) + \epsilon_t$	29	0.804	20	0.16	0.688	111.4	1.37
3b. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-2}) + \epsilon_t$	29	0.778	20.8	0.45	0.508	111.09	1.331

Table A2. Model selection tests of time-dependency the catch during spawning months using non-linear or time-varying linear responses instead of time-constant linear responses as in Table A1. See Table A1 for an explanation of the parameters and model set-up.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Time dependency test A 1984-2015 data							
1. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	30	0.814	10.3			114.14	1.43
2. $\ln(S_t) = \alpha + s(\ln(S_{t-1})) + \epsilon_t$	28.2	0.798	19.6	2.74	0.089	111.79	1.371
3. $\ln(S_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	25.5	0.77	20.7	0.97	0.416	113.23	1.382
Time dependency test B 1984-2015 data							
1. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	30	0.794	22.2			109.59	1.308
2. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	28.6	0.761	24.4	1.26	0.287	109.52	1.299
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \epsilon_t$	26.4	0.761	21.2	0.28	0.785	112.42	1.342
Time dependency test C 1984-2015 data							
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	28.6	0.761	24.4			109.52	1.299
2. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \epsilon_t$	26.1	0.698	28.5	1.49	0.242	109.55	1.273
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	25.9	0.724	26.3	1.09	0.367	110.63	1.295
Time varying test D 1984-2015 data							
1. $\ln(S_t) = \alpha_t + \epsilon_t$	29	0.658				114.45	1.373
2. $\ln(S_t) = \alpha_t + \beta_t t + \epsilon_t$	27	0.85				114.24	1.354
3a. $\ln(S_t) = \alpha + \beta_t \ln(S_{t-1}) + \epsilon_t$	28	0.723				115.66	1.49
3b. $\ln(S_t) = \alpha + \beta_t \ln(N_{t-1}) + \epsilon_t$	28	0.794				111.59	1.337

Table A3. Table A2 with 1956-1983 data instead of 1984 to 2015 data. See Table A1 for an explanation of the parameters and model set-up.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Time dependency test A 1956-1983 data							
1. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	24	0.633	-0.7			64.69	0.821
2. $\ln(S_t) = \alpha + s(\ln(S_{t-1})) + \epsilon_t$	22.1	0.614	-0.2	0.78	0.464	65.71	0.844
3. $\ln(S_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	19.9	0.58	3.1	1.19	0.329	66.35	1.053
Time dependency test B 1956-1983 data							
1. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	24	0.634	-3.8			65.48	0.821
2. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	21.6	0.584	8.2	2.24	0.127	63.8	0.783
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \epsilon_t$	18.5	0.495	16.9	1.56	0.231	63.13	0.785
Time dependency test C 1956-1983 data							
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	22.5	0.586	4.3			66.2	0.8
2. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \epsilon_t$	20.7	0.556	4.8	0.91	0.41	67.3	0.829
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	19.5	0.55	12.9	1.42	0.266	63.79	0.967

Table A4. Model selection tests of time-dependency the N_t model using F-tests of nested models fit to 1984 to 2014 log landings data. The years are determined by the covariate data availability and end in 2014 since the landings data go through 2015 and N_{2014} includes quarters in 2014 and 2015. N_t is the catch during the non-spawning period (Qtrs 4 and 1: Oct-Mar) of season t (Jul-Jun). S_{t-1} and N_{t-1} are the catch during the prior sardine season during and after the spawning period respectively. S_{t-2} and N_{t-2} are the same for two seasons prior. Test A uses catch during the spawning period as the explanatory variable. Test B uses catch during the non-spawning period as the explanatory variable. Test C uses both. The numbers next to the model equations indicate the level of nestedness. The Naive model is a model that uses the previous data point in the time series as the prediction; thus the Naive model has no estimated parameters.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Naive Model 1984-2014 data							
$\ln(N_t) = \ln(N_{t-1}) + \epsilon_t$	31	1				90.87	1.015
Time dependency test A 1984-2014 data							
1. $\ln(N_t) = \alpha + \ln(S_{t-1}) + \epsilon$	30	1.363	-20.3			107.36	1.324
2. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	29	1.018	26.2	19.99	0	93.17	1.035
3. $\ln(N_t) = \alpha + \beta_1 \ln(S_{t-1}) + \beta_2 \ln(S_{t-2}) + \epsilon_t$	28	1.009	26.6	1.15	0.292	93.92	1.062
Time dependency test B 1984-2014 data							
1. $\ln(N_t) = \alpha + \ln(N_{t-1}) + \epsilon_t$	30	0.999	24.7			92.87	1.048
2. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	29	0.978	37	6.63	0.016	88.28	1.062
3. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(N_{t-2}) + \epsilon_t$	28	0.97	34.8	0.04	0.843	90.24	1.148
Time dependency test C 1984-2014 data							
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	29	0.978	37			88.28	1.062
2a. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-1}) + \epsilon_t$	28	0.964	35	0.12	0.729	90.15	1.093
2b. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-2}) + \epsilon_t$	28	0.978	34.7	0.01	0.919	90.27	1.208

Table A5. Model selection tests of time-dependency the N_t model using non-linear or time-varying linear responses instead of time-constant linear responses as in Table A4 See Table A4 for an explanation of the parameters and model set-up.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Time dependency test A 1984-2014 data							
1. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	29	1.018	26.2			93.17	1.035
2. $\ln(N_t) = \alpha + s(\ln(S_{t-1})) + \epsilon_t$	27.3	0.992	30.2	1.83	0.185	92.61	1.016
3. $\ln(N_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	24.4	0.94	36.4	1.79	0.177	91.62	1.012
Time dependency test B 1984-2014 data							
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	29	0.978	37			88.28	1.062
2. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	27.6	0.874	45.3	3.88	0.047	84.75	0.966
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \epsilon_t$	25.4	0.805	45.6	0.87	0.441	86.11	1.02
Time dependency test C 1984-2014 data							
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	27.6	0.874	45.3			84.75	0.966
2. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \epsilon_t$	25.1	0.856	43.8	0.53	0.634	87.37	1.081
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	24.8	0.743	56.6	3.39	0.036	79.53	1.062
Time varying test D 1984-2014 data							
1. $\ln(N_t) = \alpha_t + \epsilon_t$	28	0.443				92.98	1.043
2. $\ln(N_t) = \alpha_t + \beta_t t + \epsilon_t$	26	0.455				96.96	1.045
3a. $\ln(N_t) = \alpha + \beta_t \ln(S_{t-1}) + \epsilon_t$	27	0.789				93.96	0.923
3b. $\ln(N_t) = \alpha + \beta_t \ln(N_{t-1}) + \epsilon_t$	27	0.978				90.28	1.031

Table A6. Table A5 with 1956-1983 data instead of 1984 to 2014 data. See Table A4 for an explanation of the parameters and model set-up.

Model	Residual df	MASE	Adj. R2	F	p value	AIC	LOOCV
Time dependency test A 1956-1983 data							
1. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \epsilon_t$	24	0.641	-1.7			44.98	0.574
2. $\ln(N_t) = \alpha + s(\ln(S_{t-1})) + \epsilon_t$	22.1	0.534	16.2	3.53	0.052	41.11	0.542
3. $\ln(N_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	19.9	0.502	18.1	1.09	0.362	42	0.615
Time dependency test B 1956-1983 data							
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \epsilon_t$	24	0.681	-4.2			45.61	0.575
2. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	21.6	0.507	29.1	5.69	0.009	37.12	0.468
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \epsilon_t$	18.5	0.471	32.2	1.14	0.36	37.87	0.506
Time dependency test C 1956-1983 data							
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	21.6	0.507	29.1			37.12	0.468
2a. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \epsilon_t$	19	0.45	34.4	1.49	0.251	36.74	0.498
2b. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \epsilon_t$	19.5	0.465	33.4	1.54	0.24	36.84	0.538

Appendix B: Tests for environmental variables as covariates

Table B1. Model selection tests of GPCP precipitation as an explanatory variable for the catch S_t during monsoon months (Jul-Sep) using 1984 to 2015 data. The data range is determined by the years for which SST was available in order to use a consistent dataset across covariate tests. The base model (M) with prior catch dependency was selected independently (Appendix A). To the base model, covariates are added. V_t is the covariate in same calendar year as the Jul-Sep catch. The specific hypothesis (Table 1) being tested is noted in parentheses. The models are tested as nested sets. Thus 1, 2a, 3a is a set and 1, 2b, 3b is another set. MASE is the mean absolute square error (residuals).

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2015 data						
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	28.6	0.761	24.4			109.52
V_t = Jun-Jul Precipitation (S1)						
2. $\ln(S_t) = M + \beta V_t$	27.6	0.743	23.5	0.67	0.42	110.78
3. $\ln(S_t) = M + s(V_t)$	26	0.734	27	1.51	0.241	110.28
V_t = Apr-May Precipitation (S2)						
2. $\ln(S_t) = M + \beta V_t$	27.6	0.756	23.8	0.72	0.403	110.65
3. $\ln(S_t) = M + s(V_t)$	25.6	0.748	21.1	0.24	0.792	112.98

Table B2. Model selection tests of sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1), and upwelling indices as the explanatory variables (V_t) for the catch during monsoon months (Jul-Sep) using 1984 to 2015 data. The hypothesis tested (Table 1) is noted in parentheses. Three upwelling indices were tested. The nearshore-offshore temperature diferential (UPW), which is the offshore minus nearshore SST, the average nearshore SST along the Kerala coast (boxes 2-5), and the Bakun upwelling index based on wind stress. See Table B1 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2015 data						
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	28.6	0.761	24.4			109.52
V_t = Mar-May r-SST (S5)						
2a. $\ln(S_t) = M + \beta V_t$	27.6	0.772	22.9	0.45	0.509	111.03
3a. $\ln(S_t) = M + s(V_t)$	25.6	0.748	26.5	1.37	0.271	110.68
2b. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.751	23.3	0.57	0.454	110.88
3b. $\ln(S_t) = M + s(V_{t-1})$	25.7	0.722	25.8	1.19	0.32	110.91
V_t = Oct-Dec ns-SST (L1)						
2. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.759	21.8	0	0.952	111.5
3. $\ln(S_t) = M + s(V_{t-1})$	26.1	0.768	21.8	0.66	0.482	112.32
V_t = Jun-Sep UPW (L2)						
\Rightarrow 2a. $\ln(S_t) = M + \beta V_t$	27.6	0.706	33.5	5.01	0.034	106.32
3a. $\ln(S_t) = M + s(V_t)$	25.5	0.682	34.1	0.83	0.455	107.24
2b. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.748	22.6	0.33	0.568	111.15
3b. $\ln(S_t) = M + s(V_{t-1})$	25.1	0.724	26.2	1.28	0.3	111.16
V_t = Jun-Sep ns-SST (L2)						
\Rightarrow 2a. $\ln(S_t) = M + \beta V_t$	27.6	0.745	33.3	5.53	0.027	106.4
\Rightarrow 3a. $\ln(S_t) = M + s(V_t)$	25.9	0.683	41	2.85	0.084	103.43
2b. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.742	23.2	0.54	0.468	110.89
3b. $\ln(S_t) = M + s(V_{t-1})$	25.5	0.715	22.2	0.54	0.599	112.57
V_t = Jun-Sep Bakun-UPW (L2)						
2a. $\ln(S_t) = M + \beta V_t$	27.6	0.776	28.4	3.62	0.069	108.66
\Rightarrow 3a. $\ln(S_t) = M + s(V_t)$	25.5	0.633	47.8	5.66	0.009	99.8
2b. $\ln(S_t) = M + \beta V_{t-1}$	27.7	0.744	26.1	1.67	0.207	109.65
3b. $\ln(S_t) = M + s(V_{t-1})$	25.6	0.728	24.9	0.48	0.628	111.36

Table B3. Model selection tests of the multi-year average nearshore sea surface temperature and ENSO indices as the explanatory variables (V) for the catch during summer months (Jul-Sep) using 1984 to 2015 data. The hypothesis tested (Table 1) is noted in parentheses. The ENSO indices were the ONI index averaged over all months in the calendar year and the DMI index for Sep-Nov. The 2.5-year average nearshore SST is the average for Jan-Jun in the current calendar year and the prior 2 calendar years (30 months total). Thus the average does not include any months during the Jul-Sep catch. See Table B1 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2015 data						
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	28.6	0.761	24.4			109.52
$V_t = 2.5\text{-year ave ns-SST (A1)}$						
$\Rightarrow 2. \ln(S_t) = M + \beta V_t$	27.6	0.723	33.2	5.52	0.027	106.43
$\Rightarrow 3. \ln(S_t) = M + s(V_t)$	26.2	0.653	41	3.22	0.07	103.26
$V_t = \text{ONI (A2)}$						
2. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.758	22	0.08	0.77	111.4
3. $\ln(S_t) = M + s(V_{t-1})$	26.6	0.733	23.6	1.16	0.294	111.28
$V_t = \text{Sep-Nov DMI (A3)}$						
2. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.756	21.9	0.03	0.869	111.42
3. $\ln(S_t) = M + s(V_{t-1})$	24.7	0.761	19	0.41	0.744	114.41

Table B4. Model selection tests of GPCP precipitation as an explanatory variable for the catch (N_t) during post-monsoon months (Oct-May) using 1984 to 2014 data. The data range is determined by the years for which SST was available in order to use a consistent dataset across covariate tests. The base model (M) with prior catch dependency was selected independently (Appendix A). N_{t-1} is the post-monsoon catch in prior season, and S_{t-2} is the catch during Jul-Sep two seasons prior. To the base model, covariates are added. V_t is the covariate in the calendar year, and V_{t-1} is the covariate in the prior calendar year. The specific hypothesis (Table 1) being tested is noted in parentheses. The models are tested as nested sets. Thus 1, 2a, 3a is a set and 1, 2b, 3b is another set.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \epsilon_t$	24.8	0.743	56.6			79.53
V_t = Jun-Jul Precipitation (S1)						
2a. $\ln(N_t) = M + \beta V_t$	23.8	0.755	56.7	1.03	0.318	80.23
3a. $\ln(N_t) = M + s(V_t)$	22.3	0.75	55.3	0.19	0.767	82.02
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.744	54.9	0	1	81.5
3b. $\ln(N_t) = M + s(V_{t-1})$	22.3	0.701	56.4	1.32	0.28	81.18
V_t = Apr-May Precipitation (S2)						
2a. $\ln(N_t) = M + \beta V_t$	23.8	0.742	55.1	0.11	0.735	81.34
3a. $\ln(N_t) = M + s(V_t)$	21.7	0.73	53.7	0.36	0.707	83.39
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.723	56.2	0.74	0.397	80.6
3b. $\ln(N_t) = M + s(V_{t-1})$	22	0.692	55.6	0.5	0.587	81.87

Table B5. Model selection tests of sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1), and upwelling indices as the explanatory variables (V) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data. The hypothesis tested (Table 1) is noted in parentheses. Three upwelling indices were tested. The nearshore-offshore temperature differential (UPW), which is the offshore minus nearshore SST, the average nearshore SST along the Kerala coast (boxes 2-5), and the Bakun upwelling index based on wind stress. See Table B4 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \epsilon_t$	24.8	0.743	56.6			79.53
$V_t = \text{Mar-May r-SST (S5)}$						
2a. $\ln(N_t) = M + \beta V_t$	23.8	0.707	59.1	2.95	0.101	78.46
\Rightarrow 3a. $\ln(N_t) = M + s(V_t)$	22	0.686	63.7	2.39	0.119	75.68
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.765	58.4	2.13	0.159	78.96
3b. $\ln(N_t) = M + s(V_{t-1})$	21.9	0.748	57.6	0.45	0.633	80.52
$V_t = \text{Oct-Dec ns-SST (L1)}$						
2. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.748	54.9	0	1	81.5
3. $\ln(N_t) = M + s(V_{t-1})$	22.5	0.736	56	1.13	0.318	81.37
$V_t = \text{Jun-Sep UPW (L2)}$						
\Rightarrow 2a. $\ln(N_t) = M + \beta V_t$	23.8	0.759	62.2	4.91	0.038	76
3a. $\ln(N_t) = M + s(V_t)$	21.4	0.733	62.3	0.74	0.513	77.2
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.742	54.9	0	0.979	81.49
3b. $\ln(N_t) = M + s(V_{t-1})$	21.4	0.711	56.5	1.12	0.351	81.6
$V_t = \text{Jun-Sep ns-SST (L2)}$						
\Rightarrow 2a. $\ln(N_t) = M + \beta V_t$	23.8	0.717	62.7	5.27	0.033	75.57
3a. $\ln(N_t) = M + s(V_t)$	21.9	0.714	61.8	0.39	0.67	77.33
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.744	55.3	0.23	0.626	81.18
3b. $\ln(N_t) = M + s(V_{t-1})$	21.8	0.76	54.6	0.49	0.616	82.72
$V_t = \text{Jun-Sep Bakun-UPW (L2)}$						
2a. $\ln(N_t) = M + \beta V_t$	23.8	0.758	57.4	1.58	0.221	79.75
3a. $\ln(N_t) = M + s(V_t)$	21.8	0.672	60.3	1.58	0.228	78.55
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.765	56.8	1.17	0.287	80.13
3b. $\ln(N_t) = M + s(V_{t-1})$	22	0.74	57.9	1.08	0.349	80.24

Table B6. Model selection tests of the multi-year average nearshore sea surface temperature and ENSO indices as the explanatory variables (V) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data. The hypothesis tested (Table 1) is noted in parentheses. The ENSO indices were the ONI index averaged over all months in the calendar year and the DMI index for Sep-Nov. The 2.5-year average nearshore SST is the average for Jan-Jun in the current calendar year and the prior 2 calendar years (30 months total). Thus the average does not include any months during the Oct-Mar catch. See Table B4 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \epsilon_t$	24.8	0.743	56.6			79.53
$V_t = 2.5\text{-year ave ns-SST (A1)}$						
$\Rightarrow 2. \ln(N_t) = M + \beta V_t$	23.8	0.667	64.7	7.68	0.012	73.9
$\Rightarrow 3. \ln(N_t) = M + s(V_t)$	22.7	0.594	67.5	2.58	0.12	71.88
$V_t = \text{ONI (A2)}$						
2a. $\ln(N_t) = M + \beta V_t$	23.9	0.794	56.4	0.86	0.351	80.41
3a. $\ln(N_t) = M + s(V_t)$	22.8	0.79	56.9	0.93	0.351	80.5
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.744	54.9	0	1	81.46
3b. $\ln(N_t) = M + s(V_{t-1})$	23	0.748	55.5	0.99	0.313	81.46
$V_t = \text{Sep-Nov DMI (A3)}$						
2a. $\ln(N_t) = M + \beta V_t$	23.9	0.791	55.7	0.42	0.498	80.92
3a. $\ln(N_t) = M + s(V_t)$	21.2	0.77	56.8	1	0.404	81.54
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.746	54.8	0	1	81.52
$\Rightarrow 3b. \ln(N_t) = M + s(V_{t-1})$	21.1	0.678	67.5	4.34	0.018	72.69

Table B7. Model selection tests of GPCP precipitation as an explanatory variable for the catch (N_t) during post-monsoon months (Oct-May) using 1984 to 2014 data and the base model without S_{t-2} . The data range is determined by the years for which SST was available in order to use a consistent dataset across covariate tests. The base model (M) with prior catch dependency was selected independently (Appendix A). N_{t-1} is the post-monsoon catch in prior season, and S_{t-2} is the catch during Jul-Sep two seasons prior. To the base model, covariates are added. V_t is the covariate in the calendar year, and V_{t-1} is the covariate in the prior calendar year. The specific hypothesis (Table 1) being tested is noted in parentheses. The models are tested as nested sets. Thus 1, 2a, 3a is a set and 1, 2b, 3b is another set.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	27.6	0.874	45.3			84.75
V_t = Jun-Jul Precipitation (S1)						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.873	43.3	0	1	86.74
3a. $\ln(N_t) = M + s(V_t)$	24.9	0.849	45.3	1.29	0.288	86.66
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.872	44	0.35	0.556	86.33
3b. $\ln(N_t) = M + s(V_{t-1})$	24.9	0.867	43.4	0.53	0.566	87.69
V_t = Apr-May Precipitation (S2)						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.872	43.5	0.1	0.742	86.59
3a. $\ln(N_t) = M + s(V_t)$	24.5	0.863	41.3	0.24	0.802	89.08
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.86	44	0.35	0.553	86.32
3b. $\ln(N_t) = M + s(V_{t-1})$	24.8	0.833	44.7	0.85	0.431	87.01

Table B8. Model selection tests of sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1), and upwelling indices as the explanatory variables (V) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data and base model without S_{t-2} . The hypothesis tested (Table 1) is noted in parentheses. Three upwelling indices were tested. The nearshore-offshore temperature differential (UPW), which is the offshore minus nearshore SST, the average nearshore SST along the Kerala coast (boxes 2-5), and the Bakun upwelling index based on wind stress. See Table B4 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	27.6	0.874	45.3			84.75
$V_t = \text{Mar-May r-SST (S5)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.854	46.1	1.44	0.242	85.16
3a. $\ln(N_t) = M + s(V_t)$	24.7	0.841	46.6	0.82	0.447	85.97
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.81	49.7	3.6	0.07	82.99
3b. $\ln(N_t) = M + s(V_{t-1})$	24.6	0.777	50.9	1.02	0.379	83.45
$V_t = \text{Oct-Dec ns-SST (L1)}$						
2. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.828	45.6	1.12	0.299	85.47
3. $\ln(N_t) = M + s(V_{t-1})$	25.2	0.839	44.5	0.28	0.684	86.89
$V_t = \text{Jun-Sep UPW (L2)}$						
\Rightarrow 2a. $\ln(N_t) = M + \beta V_t$	26.6	0.863	54	6.63	0.017	80.25
3a. $\ln(N_t) = M + s(V_t)$	24.1	0.834	56	1.22	0.319	80.36
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.876	43.3	0	1	86.74
3b. $\ln(N_t) = M + s(V_{t-1})$	24	0.804	46.9	1.46	0.253	86.25
$V_t = \text{Jun-Sep ns-SST (L2)}$						
\Rightarrow 2a. $\ln(N_t) = M + \beta V_t$	26.6	0.857	51.6	4.71	0.04	81.79
3a. $\ln(N_t) = M + s(V_t)$	24.6	0.86	51.3	0.63	0.545	83.19
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.819	45.2	0.92	0.345	85.67
3b. $\ln(N_t) = M + s(V_{t-1})$	24.5	0.801	43.6	0.36	0.711	87.78
$V_t = \text{Jun-Sep Bakun-UPW (L2)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.825	46.2	1.5	0.231	85.1
3a. $\ln(N_t) = M + s(V_t)$	24.5	0.746	47	0.89	0.428	85.85
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.825	44.2	0.44	0.507	86.22
3b. $\ln(N_t) = M + s(V_{t-1})$	24.7	0.807	44.7	0.83	0.446	87.07

Table B9. Model selection tests of the multi-year average nearshore sea surface temperature and ENSO indices as the explanatory variables (V) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data and base model without S_{t-2} . The hypothesis tested (Table 1) is noted in parentheses. The ENSO indices were the ONI index averaged over all months in the calendar year and the DMI index for Sep-Nov. The 2.5-year average nearshore SST is the average for Jan-Jun in the current calendar year and the prior 2 calendar years (30 months total). Thus the average does not include any months during the Oct-Mar catch. See Table B4 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \epsilon_t$	27.6	0.874	45.3			84.75
$V_t = 2.5\text{-year ave ns-SST (A1)}$						
$\Rightarrow 2. \ln(N_t) = M + \beta V_t$	26.6	0.736	55	8.08	0.009	79.58
$\Rightarrow 3. \ln(N_t) = M + s(V_t)$	25.3	0.664	60.3	3.48	0.064	76.34
$V_t = \text{ONI (A2)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.883	48.4	2.66	0.115	83.81
3a. $\ln(N_t) = M + s(V_t)$	25.5	0.877	47.9	0.37	0.567	84.69
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.875	43.8	0.25	0.614	86.45
3b. $\ln(N_t) = M + s(V_{t-1})$	25.6	0.844	46.3	1.86	0.185	85.53
$V_t = \text{Sep-Nov DMI (A3)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.904	48.1	2.52	0.126	83.99
3a. $\ln(N_t) = M + s(V_t)$	23.6	0.882	48	0.73	0.548	85.86
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.855	43.9	0.26	0.606	86.38
3b. $\ln(N_t) = M + s(V_{t-1})$	23.7	0.853	43.4	0.68	0.571	88.43

Appendix C: Tests for Chlorophyll-a as a covariate

Table C1. Model selection tests of nearshore (0-80km) chlorophyll-a (ns-CHL) as an explanatory variable for the Jul-Sep catch (S_t) using 1998 to 2014 data. The data range is determined by the years for which CHL was available. Only CHL in Oct-Dec and Jan-Mar in the current season (which runs Jul-Jun) is not used since these months are after the Jul-Sep catch being modeled. Non-linearity is modeled as a 2nd-order polynomial due to data constraints and appears as $p()$ in the model equations. The Jul-Sep catch is modeled as a function of Oct-Jun catch in the prior year only, without Jul-Sep catch 2-years prior as in the other covariate analyses (Appendix B). This is done due to data constraints. The models are nested; the Roman numeral indicates the level of nestedness. Models at levels II and higher are shown with the component that is added to the base level model (M) at top.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1998-2014 data						
1. $\ln(S_t) = \alpha + p(\ln(N_{t-1})) + \epsilon_t$	14	0.516	25.3			18.29
V_t = Jul-Sep ns-CHL						
2. $\ln(S_t) = M + \beta V_t$	13	0.515	25.2	0.85	0.377	19.05
3. $\ln(S_t) = M + p(V_t)$	12	0.489	27.8	1.22	0.295	19.1
4. $\ln(S_t) = M + p(V_t) + \beta V_{t-1}$	11	0.484	21.3	0.02	0.898	21.07
5. $\ln(S_t) = M + p(V_t) + p(V_{t-1})$	10	0.479	13.8	0.04	0.851	23.01
V_t = Oct-Dec ns-CHL						
2. $\ln(S_t) = M + \beta V_{t-1}$	13	0.517	20	0.08	0.787	20.2
3. $\ln(S_t) = M + p(V_{t-1})$	12	0.447	27	2.25	0.16	19.29
V_t = Jan-Mar ns-CHL						
2. $\ln(S_t) = M + \beta V_{t-1}$	13	0.515	19.6	0.01	0.934	20.28
3. $\ln(S_t) = M + p(V_{t-1})$	12	0.519	13.9	0.13	0.721	22.09

Table C2. Model selection tests of nearshore (0-80km) chlorophyll-a (ns-CHL) as an explanatory variable for Oct-Jun catch (N_t) using 1998 to 2014 data. The data range is determined by the years for which CHL was available. V_{t-1} is CHL in the prior Jul-Jun season. Non-linearity is modeled as a 2nd-order polynomial due to data constraints and appears as $p()$ in the model equations. The Oct-Jun catch is modeled as a function of Oct-Jun catch in the prior year only, without Jul-Sep catch 2-years prior as in the other covariate analyses (Appendix B). This was done due to data constraints. The models are nested; the numeral indicates the level of nestedness. Models at levels 2 and higher are shown with the component that is added to the base level model (M) at top.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1998-2014 data						
1-M. $\ln(N_t) = \alpha + p(\ln(N_{t-1})) + \epsilon_t$	14	0.875	26.5			18.94
$V_t = \text{Jul-Sep ns-CHL}$						
2. $\ln(N_t) = M + \beta V_t$	13	0.879	24	0.5	0.496	20.26
3. $\ln(N_t) = M + p(V_t)$	12	0.869	18.6	0.13	0.727	22.07
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.832	31.5	2.15	0.17	18.5
3. $\ln(N_t) = M + p(V_{t-1})$	11.4	0.748	35.8	1.29	0.302	18.12
$V_t = \text{Oct-Dec ns-CHL}$						
2. $\ln(N_t) = M + \beta V_t$	13	0.889	24	0.69	0.425	20.26
3. $\ln(N_t) = M + p(V_t)$	12	0.721	35.1	3.57	0.088	18.2
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.718	45.3	5.23	0.043	14.66
3. $\ln(N_t) = M + p(V_{t-1})$	10.8	0.711	39.2	0.11	0.907	17.41
$V_t = \text{Jan-Mar ns-CHL}$						
2. $\ln(N_t) = M + \beta V_t$	13	0.848	32.6	1.86	0.203	18.21
3. $\ln(N_t) = M + p(V_t)$	12	0.846	27	0	0.967	20.21
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.86	21.3	0.07	0.796	20.85
3. $\ln(N_t) = M + p(V_{t-1})$	10.2	0.812	18.2	0.55	0.647	22.7

Appendix D: Covariates along the SE India coast

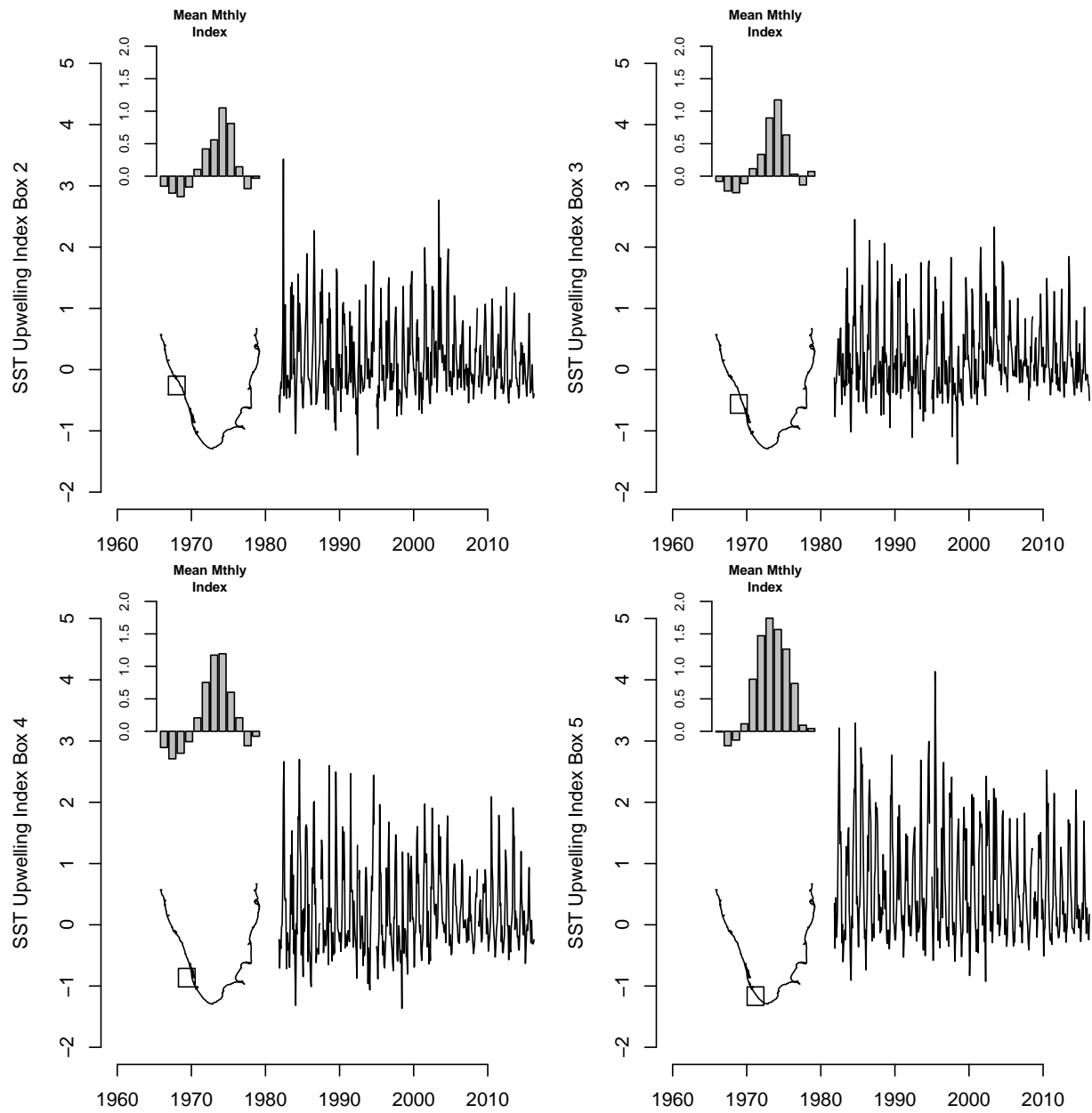


Figure D1. Upwelling index.

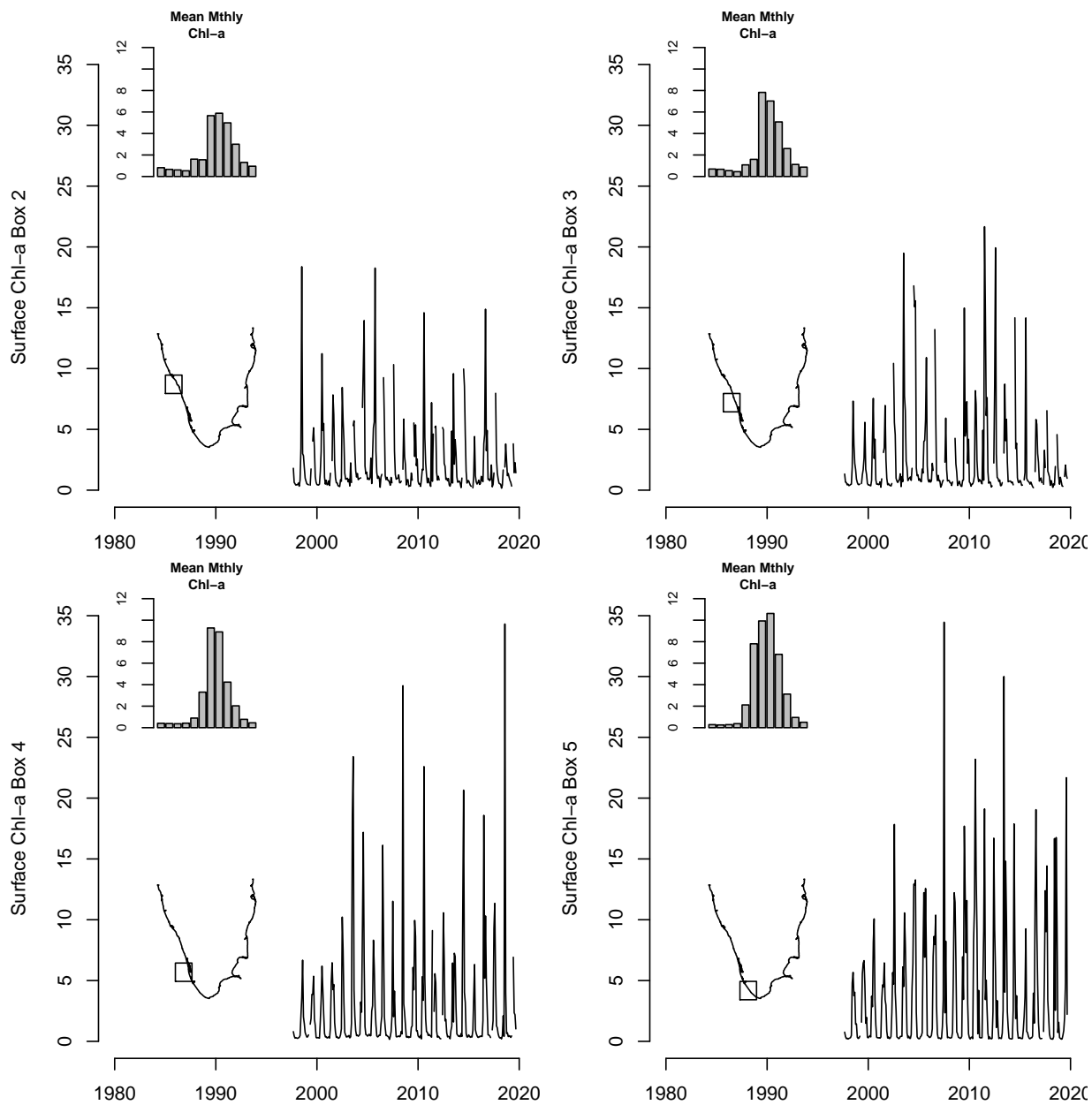


Figure D2. Chlorophyll-a.

Appendix E: Comparison of land and oceanic rainfall measurements

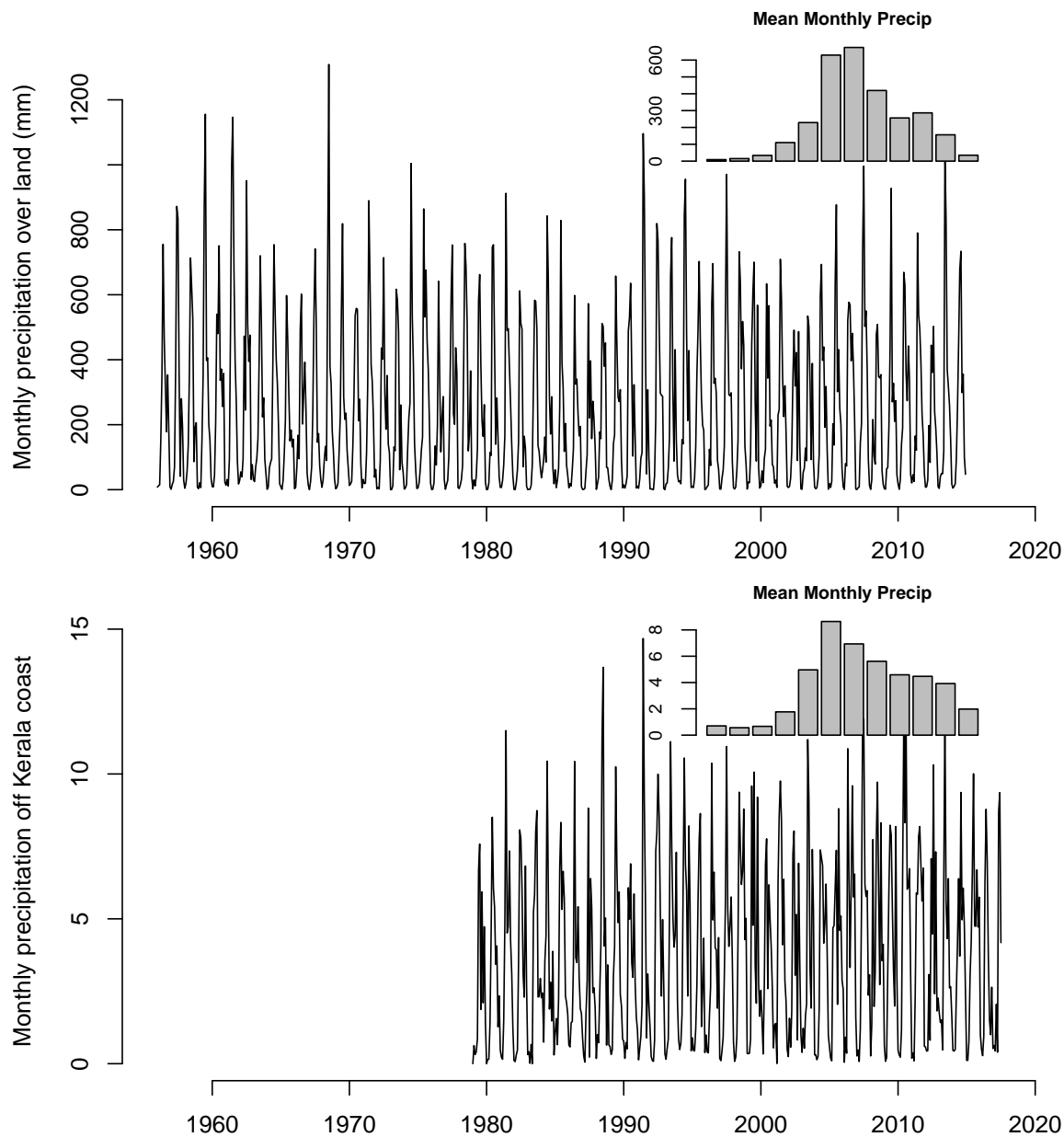


Figure E1

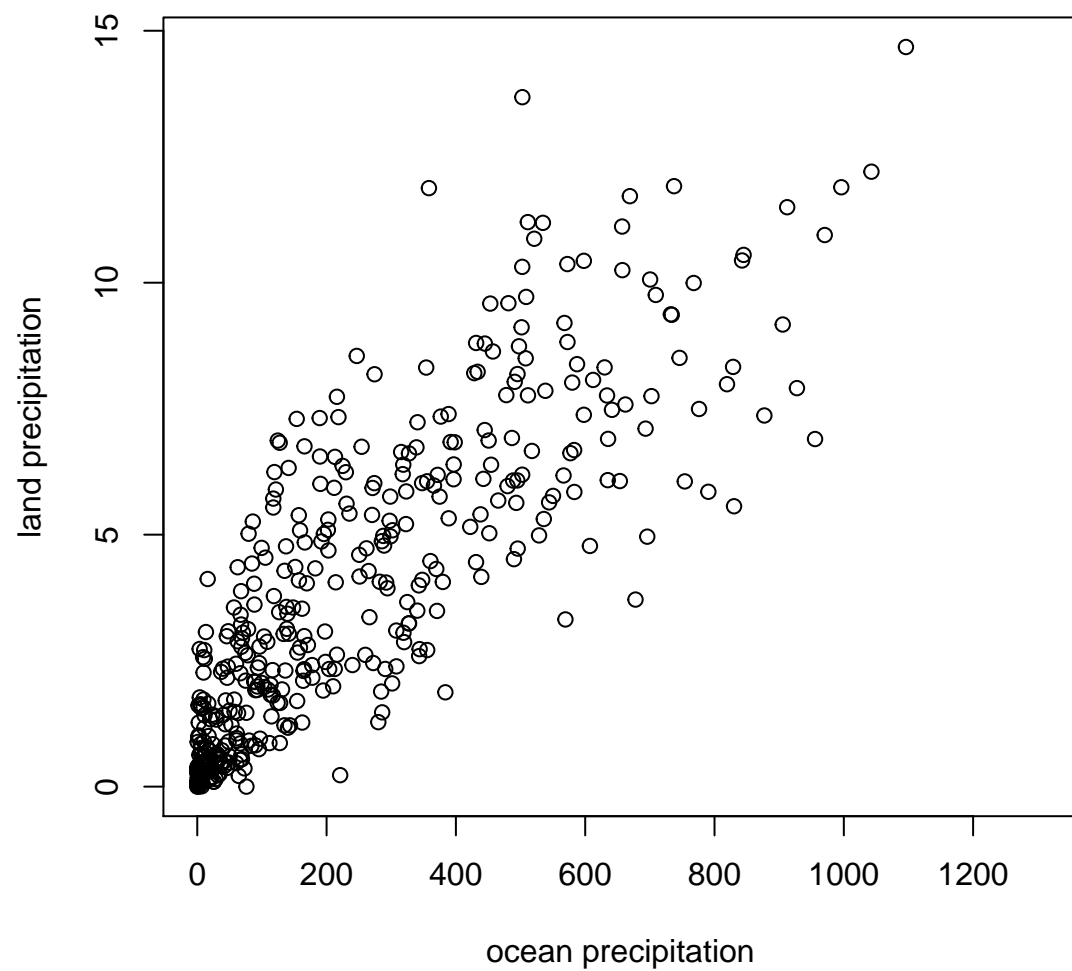


Figure E2. Monthly precipitation measured over land via land gauges versus the precipitation measured via remote sensing over the ocean.

Appendix F: Influential years tests

Validation of the landings base model

This describes a variety of cross-validations used to select the base model for landing. The base model is the model with no environmental covariates only prior landings as covariates.

Three types of base models were fit. The first two were GAM and linear models with Jul-Sep and Oct-Mar in the prior season only or prior season and two seasons prior as covariates. c is the response variable: landings during the two seasons, either Jul-Sep or Oct-Mar.

$$\begin{aligned}\text{GAM t-1 : } X_t &= \alpha + s(c_{t-1}) + e_t \\ \text{Linear t-1 : } X_t &= \alpha + \beta c_{t-1} + e_t \\ \text{GAM t-1, t-2 : } X_t &= \alpha + s(c_{t-1}) + s(d_{t-2}) + e_t \\ \text{Linear t-1, t-2 : } X_t &= \alpha + \beta c_{t-1} + d_{t-2} + e_t\end{aligned}$$

where c_{t-1} was either S_{t-1} (Jul-Sep landings in prior season) or N_{t-1} (Oct-Mar landings in prior season) and d_{t-2} was the same but 2 seasons prior.

These types of models do not allow the model parameters (the intercept α and effect parameter β) to vary in time. The second type of models were dynamic linear models (DLMs). DLMs allow the parameters to evolve in time. Two types of DLMs were used, an intercept only model where the intercept α evolves and a linear model where the effect parameter β is allowed to evolve:

$$\begin{aligned}\text{DLM intercept only : } X_t &= \alpha_t + e_t \\ \text{DLM intercept and slope : } X_t &= \alpha_t + \beta_t t + e_t \\ \text{DLM intercept and effect : } X_t &= \alpha + \beta_t c_{t-1} + e_t\end{aligned}$$

In addition to the GAM, linear and DLM models, three null models were included in the tested model sets:

$$\begin{aligned}\text{intercept only : } X_t &= \alpha + e_t \\ \text{intercept and prior catch : } X_t &= \alpha_t + X_{t-1} + e_t \\ \text{prior catch only : } X_t &= X_{t-1} + e_t\end{aligned}$$

The ‘intercept only’ is a flat level model. The ‘prior catch only’ simply uses the prior value of the time series (in this case landings) as the prediction and is a standard null model for prediction. The ‘intercept and prior catch’ combines these two null models.

The models were fit to the 1956-2015 landings (full data) and 1984-2015 (data that overlap the environmental covariates).

The model performance was measured by AIC, AICc and LOOCV prediction. The LOOCV prediction error is the data point t minus the predicted value for data point t . This is repeated for all data points t . The influence of single data points to on model performance was evaluated by leaving out one data point, fitting to the remaining data and computing the model performance (via AIC, AICc or LOO prediction error).

Results: Jul-Sep landings

The Figure 1D shows the ΔAIC for the models: GAM, linear, and DLM. The figure shows that for the 1984-2015 data with any year left out, the set of models that has the lowest AIC was always the GAM or linear model with Oct-Mar in the prior season. There were cases where deleting a year removed one of these two from the ‘best’ category, but they were still in the ‘competitive’ category with a ΔAIC less than 2.

AIC gives us a measure of how well the models fit the data, with a penalty for the number of estimated parameters. We look at the one-step-ahead predictive performance (Figure 2D), we see that all the GAM, linear and DLM models have a hard time adjusting to shifts in the data (e.g. after 1998). The null models can adjust quickly but has large errors when there are rapid changes. The root mean squared error (which penalizes large predictive errors) is lowest for the models with Oct-Mar in the prior season (Figure 3D).

It should be noted that none of the Jul-Sep models has a particularly high adjusted R^2 . The values are generally less than 0.3. The Jul-Sep landings tend to be highly variable and not related to the catch in prior years. Jul-Sep is during the monsoon during which fishing is not always possible due to sea-state and there is a 6-week fishing ban during this time.

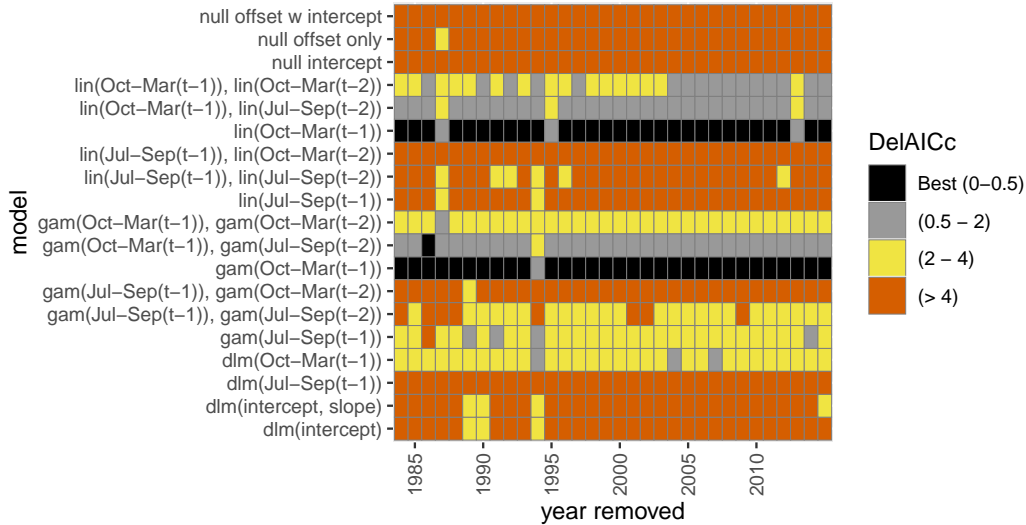


Figure 1D. $\Delta AICc$ for the Jul-Sep landings base models with one year deleted using only the landings data that overlap with the environmental data 1984-2015. See Figure 1D for an explanation of the figure.

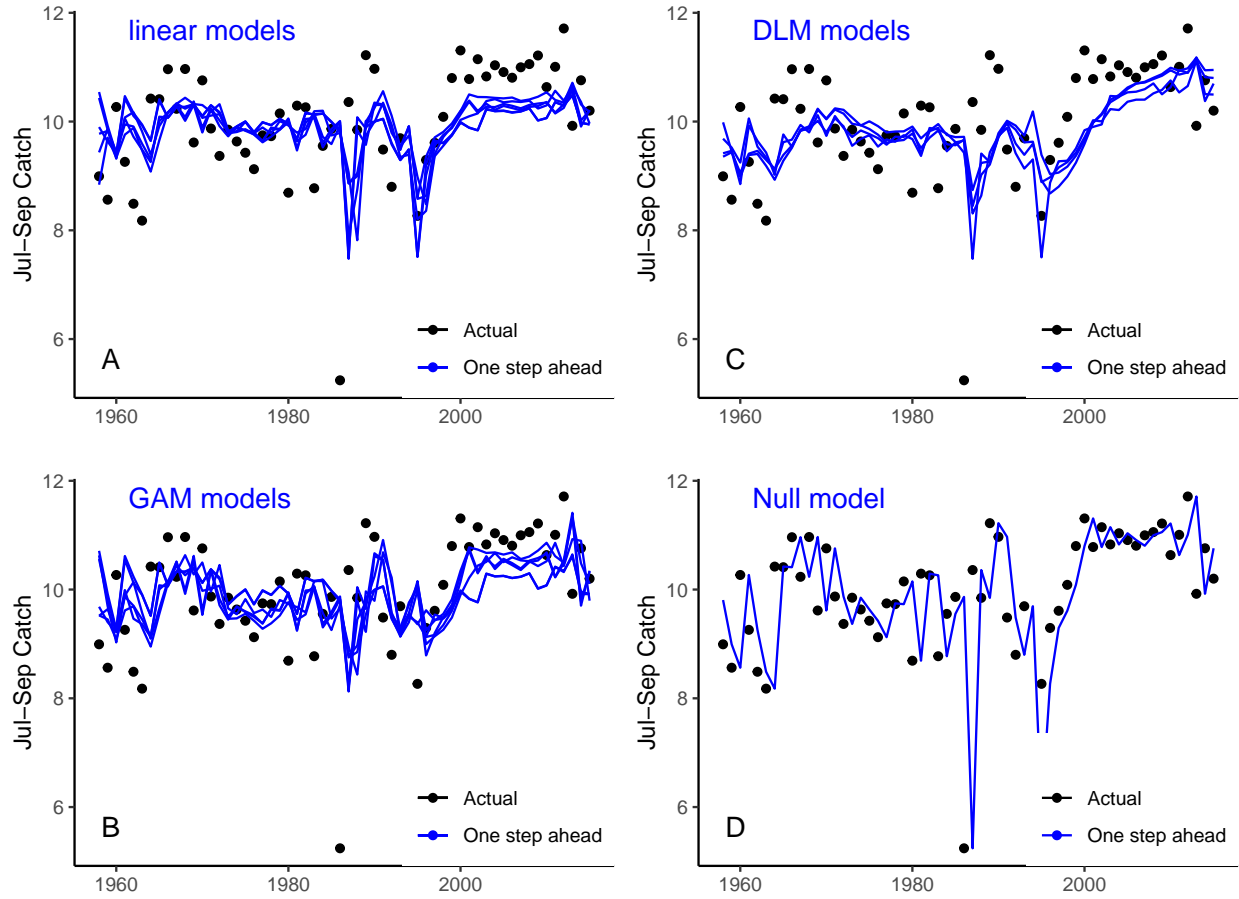


Figure 3D. Leave one out (LOO) one step ahead prediction errors for the linear, GAM, and DLM models of Jul-Sep landings. The data point at year t on the x-axis is predicted from the data up to year $t-1$.

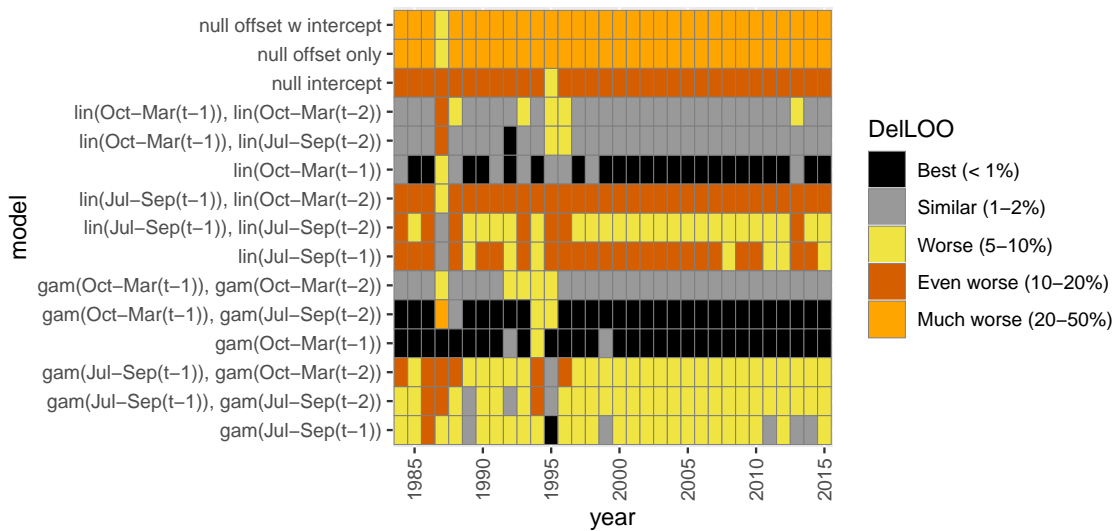


Figure 4D. Leave-one-out predictive performance (leave out a year, fit, predict that year) for the Jul-Sep landings base models. The performance (DelLOO) is the RSME (root mean square error).

Validation of the Oct-Mar landings base models

The Figure 4D shows that for Oct-Mar landings with the 1984 to 2014 data, the best model was always GAM with Oct-Mar in the prior season and Jul-Sep landings two seasons prior. For the one step ahead predictions, simpler models had the lower prediction errors: GAM with Oct-Mar in the prior season for the recent data and linear with Oct-Mar in the prior season for the full data set (Figure 6D).

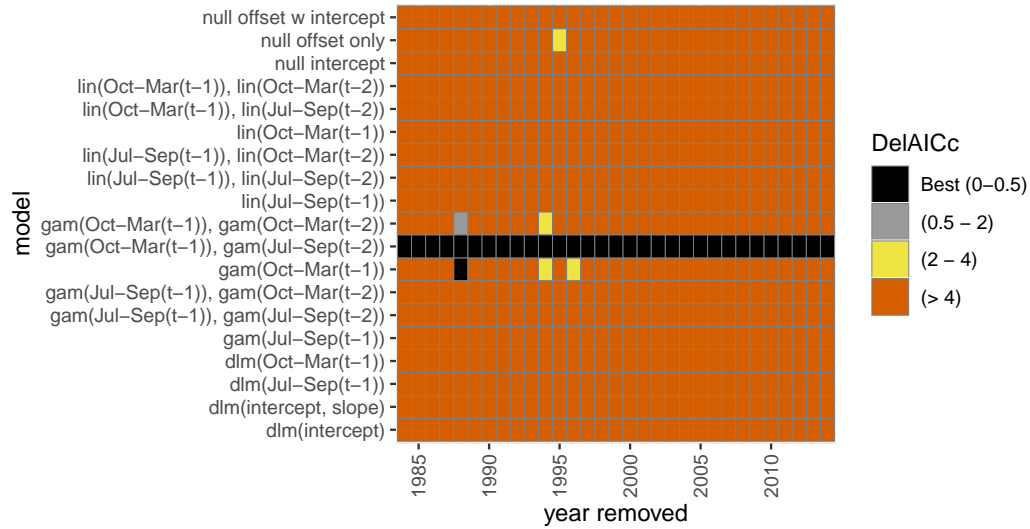


Figure 6D. $\Delta AICc$ for the Oct-Mar landings base models with one year deleted using only the landings data that overlap with the environmental data 1984-2015. See Figure 1D for an explanation of the figure.

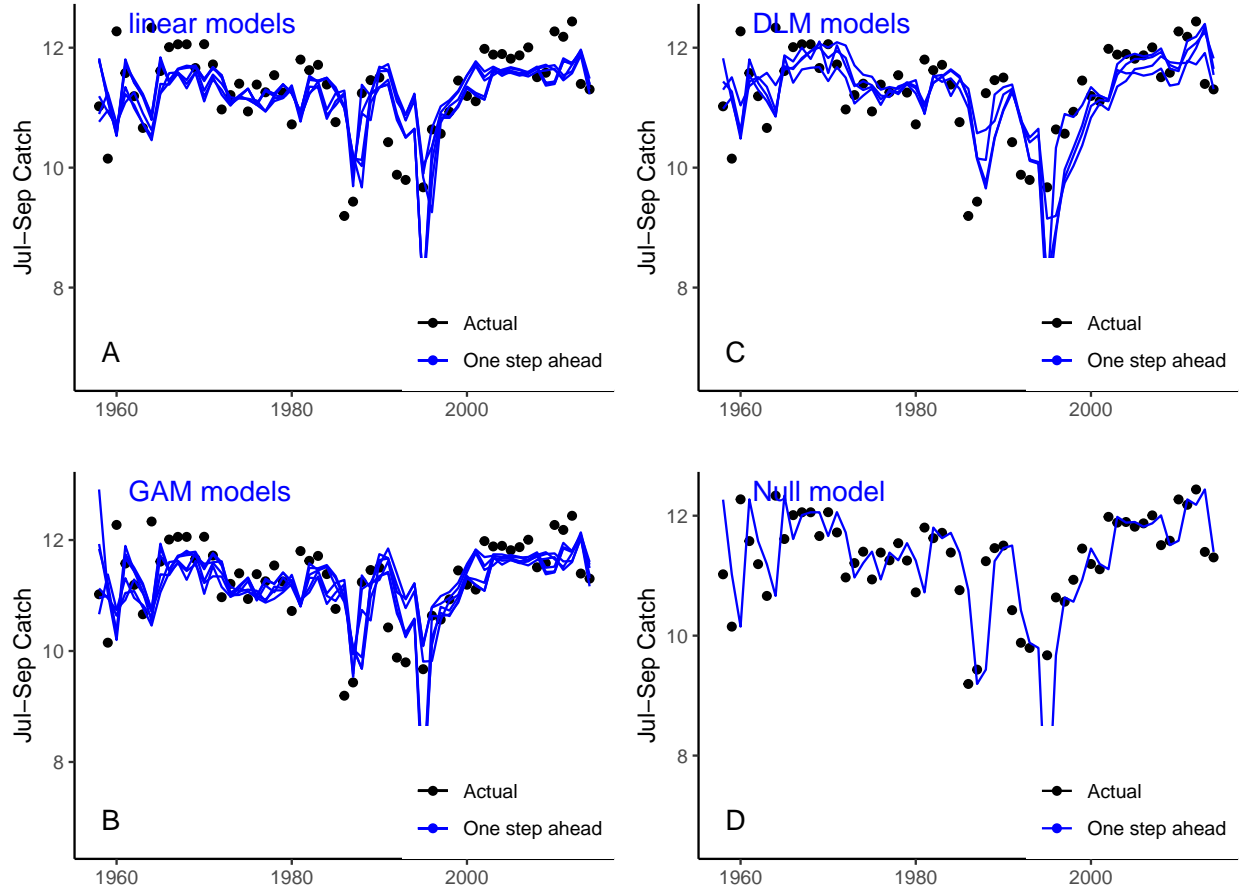


Figure 8D. Leave one out (LOO) one step ahead prediction errors for the linear, GAM, and DLM models of Oct-Mar landings. The data point at year t on the x-axis is predicted from the data up to year $t-1$.

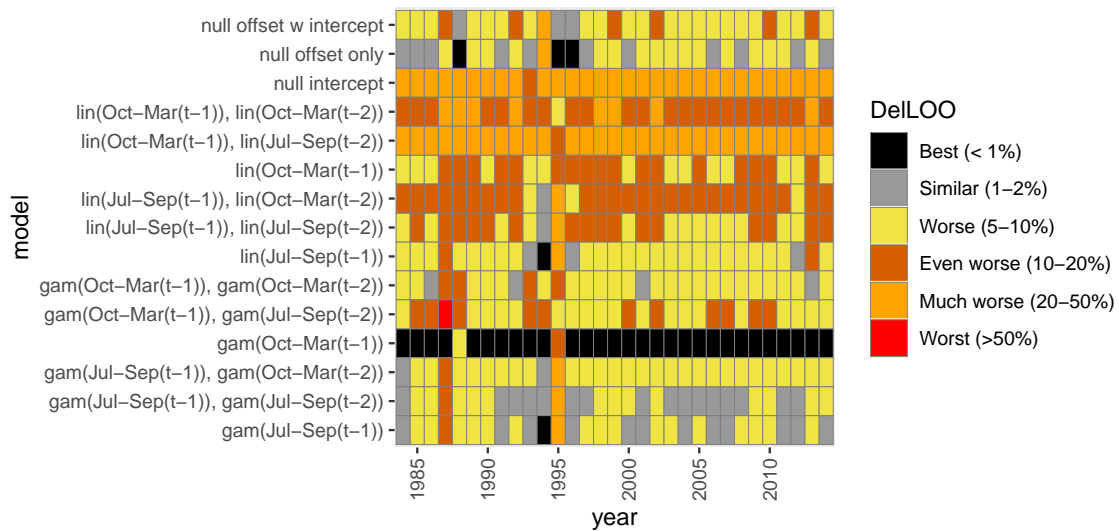


Figure 9D. Leave-one-out predictive performance (leave out a year, fit, predict that year) for the Oct-Mar landings base models. The performance (DelLOO) is the RSME (root mean square error).

Appendix G: Data Sources

Landings Data

Description

A dataset containing the landings (in metric tons) of oil sardines in Kerala state 1956-2016. The data are collected and processed into a total landings estimates based on a stratified sampling of landing sites along the SW coast of India throughout the year. The program is run by the Central Marine Fisheries Research Institute (CMFRI) in Cochin, India. The data were obtained from reports published by CMFRI; see references.

References

CMFRI reports were downloaded from the CMFRI Publication repository <http://www.cmfri.org.in>.

1956-1968 Antony Raja BT (1969). "Indian oil sardine." CMFRI Bulletin, 16, 1-142.

1968-1978 Pillai VN (1982). Physical characteristics of the coastal waters off the south-west coast of India with an attempt to study the possible relationship with sardine, mackerel and anchovy fisheries. Thesis, University of Cochin.

1975-1984 Kerala Jacob T, Rajendran V, Pillai PKM, Andrews J, Satyavan UK (1987). "An appraisal of the marine fisheries in Kerala." Central Marine Fisheries Research Institute.

1975-1984 Karnataka Kurup KN, Nair GKK, Annam VP, Kant A, Beena MR, Kambadkar L (1987). "An appraisal of the marine fisheries of Karnataka and Goa." Central Marine Fisheries Research Institute.

1985-2014 Provided by CMFRI directly via a data request.

SST Data

Description

The SST satellite data were downloaded from the NOAA ERDDAP server using R Mendels `rerddapXtracto` R package which uses the `ropensci rerddap` R package.

Details

For 1981 to 2003, We used the Pathfinder Version 5.2 (L3C) monthly day and night product on a 0.0417 degree grid. These SST data use the Advanced Very-High Resolution Radiometer (AVHRR) instrument on the Pathfinder satellites. These data were provided by GHRSSST and the US National Oceanographic Data Center. This project was supported in part by a grant from the NOAA Climate Data Record (CDR) Program for satellites.

For 2004 to 2016, we used the NOAA CoastWatch sea surface temperature (SST) products derived from NOAA's Polar Operational Environmental Satellites (POES). The SST estimates use the Advanced Very-High Resolution Radiometer (AVHRR) instruments on the POES satellites and are on a 0.1 degree grid.

Both SST data sets were downloaded from the NOAA ERDDAP server:

<https://coastwatch.pfeg.noaa.gov/erddap/info/erdAGsstamday/index.html>

<https://coastwatch.pfeg.noaa.gov/erddap/info/erdPH2sstamday/index.html>.

The SST values were averaged across thirteen 1 degree by 1 degree boxes which roughly parallel the bathymetry (Figure 1 in main text).

References

These data were provided by GHRSSST and the US National Oceanographic Data Center. This project was supported in part by a grant from the NOAA Climate Data Record (CDR) Program for satellites. The data were downloaded from NOAA CoastWatch-West Coast Regional Node and Southwest Fisheries Science Center's Environmental Research Division. To cite these data in a paper, please follow the instructions in the license and at this link: <https://coastwatch.pfeg.noaa.gov/erddap/information.html#citeDataset>

Casey KS, Brandon TB, Cornillon P, Evans R (2010). "The past, present, and future of the AVHRR Pathfinder SST program." In *Oceanography from space*, 273–287. Springer.

Walton C, Pichel W, Sapper J, May D (1998). "The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites." *Journal of Geophysical Research: Oceans*, 103(C12), 27999–28012.

Upwelling Data

Description

Three upwelling indices are in the 'upw' data object: a SST nearshore offshore differential, a wind-based index and the Bakun indices. The upwelling indices and SST data were downloaded from the NOAA ERDDAP server using R Mendels `rerddapXtracto` R package which uses the `ropensci rerddap` R package.

Details

The Wind-based monthly upwelling indices were downloaded from the NOAA ERDDAP server. The first is 1999-2009 on a 0.125 degree grid. The second is 2009 to present on a 0.25 degree grid. and the second is See <https://coastwatch.pfeg.noaa.gov/erddap/info/erdQSstressmday/index.html> and <https://coastwatch.pfeg.noaa.gov/erddap/info/erdQAstressmday/index.html>.

The SST differential upwelling indices were downloaded from the NOAA ERDDAP server. The first is 1981-2012 on a 0.0417 degree grid. The second is 2003-2016 on a 0.1 degree grid. Both are AVHRR so accurate for close to the coast. <https://coastwatch.pfeg.noaa.gov/erddap/info/erdPH2sstamday/index.html> and <https://coastwatch.pfeg.noaa.gov/erddap/info/erdAGsstamday/index.html>. The UPW index is the difference between the coast box (1 to 5) and a box 3 degrees offshore at the same latitude.

The Bakun index (The Bakun 1973) is calculated based upon Ekman's theory of mass transport due to wind stress. The index is computed from the `ektrx` and `ektry`, which are the x- and y- components of Ekman Transport obtained from the ERDDAP link below, and `coast_angle` is 158 degrees for the India west coast near Kochi. <https://coastwatch.pfeg.noaa.gov/erddap/info/erdlasFnWPr/index.html>. The function to compute the Bakun index is at <https://oceanview.pfeg.noaa.gov/products/upwelling/bakun>.

References

SST data: These data were provided by GHRSSST and the US National Oceanographic Data Center. This project was supported in part by a grant from the NOAA Climate Data Record (CDR) Program for satellites. The data were downloaded from NOAA CoastWatch-West Coast Regional Node and Southwest Fisheries Science Center's Environmental Research Division.

Wind-based UPW index: NOAA's CoastWatch Program distributes wind velocity measurements derived from the Seawinds instrument aboard NASA's QuikSCAT satellite. The Seawinds instrument is a dual-beam microwave scatterometer designed to measure wind magnitude and direction over the global oceans. CoastWatch further processes these wind velocity measurements to wind stress and wind stress curl.

Bakun index: The Environmental Research Division (ERD), within NOAA Fisheries, has long been a leader in development and calculation of upwelling and other environmental indices. ERD was originally established as the Pacific Environmental Group at the U.S. Navy Fleet Numerical Meteorology and Oceanography Center (FNMOC) in Monterey, California, to take advantage of the Navy's global oceanographic and meteorological models. FNMOC produces operational forecasts of the state of the atmosphere and the ocean several times daily. Before the advent of satellite oceanography, these forecasts provided global snapshots of ocean

conditions for Navy operations, but were also invaluable for studies of fisheries climatology since they provided long time series of environmental conditions at a much higher resolution than was possible from direct measurement. The FNMOC sea-level pressure became the basis of the Bakun upwelling index calculation, and provides estimates of upwelling for the Northern Hemisphere starting in 1948 and globally since 1981.

Bakun A (1973). "Coastal upwelling indices, west coast of North America." US Department of Commerce, NOAA Technical Report NMFS SSRF-671.

Precipitation Data

Description

Three precipitation datasets off the SW coast of India. Two are satellite derived and one is based on land gauges.

Details

The National Climatic Data Center provides basic information on the GPCP Precipitation dataset. The dataset consists of monthly precipitation estimates (average mm/day) from January 1979 to the present. The precipitation estimates merge several satellite and in situ sources into a final product. Data are provided on a 2.5 degree grid. The GPCP Precipitation data are provided by the NOAA/NCEI Global Precipitation Climatology Project and were downloaded from <https://www.ncei.noaa.gov/data/global-precipitation-climatology-project-gpcp-monthly>. Two boxes were defined, one off the Kerala coast and one off the Karnataka coast, and the average values of all grid points within these boxes were used. The boxes are Kerala Lat(8.75, 11.25), Lon(73.25, 75.75) Karnataka Lat(13.75, 16.25), Lon(71.25, 73.75)

The land gauge data is monthly, seasonal and annual rainfall (in mm) area weighted average for each state in India starting from 1901 onwards. This data set is based on rain gauges. The data are provided by the India Meteorological Department (Ministry of Earth Sciences). The 1901 to 2014 data were downloaded from the Open Government Data Platform India <https://data.gov.in>. The 2015 and 2016 data were extracted from the yearly Rainfall Statistics reports (see references).

NASA's Tropical Rainfall Measuring Mission (TRMM) website provides background on the TRMM data (<https://pmm.nasa.gov/>). 1997 to 2015 monthly precipitation estimates on a 0.25 degree grid were downloaded from the Tropical Rainfall Measuring Mission (TRMM) website. The data are averaged in the 2.5 x 2.5 degree boxes used for the other satellite data.

References

Adler R, Huffman G, Chang A, Ferraro R, Xie P, Janowiak J, Rudolf B, Schneider U, Curtis S, Bolvin D, Gruber A, Susskind J, Arkin P (2003). "The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present)." *Journal of Hydrometeorology*, 4, 1147-1167.

Adler R, Wang J, Sapiiano M, Huffman G, Chiu L, Xie PP, Ferraro R, Schneider U, Becker A, Bolvin D, Nelkin E, Gu G, Program NC (2016). "Global Precipitation Climatology Project (GPCP) Climate Data Record (CDR), Version 2.3 (Monthly)." National Centers for Environmental Information. doi: 10.7289/V56971M6.

Purohit MK, Kaur S (2016). "Rainfall Statistics of India - 2016." India Meteorological Department (Ministry of Earth Sciences). <http://hydro.imd.gov.in/hydrometweb/>.

Kothawale DR, Rajeevan M (2017). Monthly, seasonal and annual rainfall time series for all-India, homogeneous regions and meteorological subdivisions: 1871-2016. Report, Indian Institute of Tropical Meteorology..

NCEI (2017). Global Precipitation Climatology Project Monthly Product Version 2.3. Retrieved from National Centers for Environmental Information website: <https://www.ncei.noaa.gov/data/global-precipitation-climatology-project-gpcp-monthly/access/>.

Chlorophyll Data

Description

The CHL satellite data were downloaded from the NOAA ERDDAP server using R Mendels rerddapXtracto R package which uses the ropensci rerddap R package.

Details

The Chlorophyll-a products are developed by the Ocean Biology Processing Group in the Ocean Ecology Laboratory at the NASA Goddard Space Flight Center.

For 1997 to 2002, we used the Chlorophyll-a 2014.0 Reprocessing (R2014.0) product from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) on the Orbview-2 satellite. These data are on a 0.1 degree grid. See reference below.

For 2003 to 2017, we used the MODIS-Aqua product on a 4km grid. These CHL data are taken from measurements gathered by the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Aqua Spacecraft. See reference below.

Both CHL data sets were downloaded from the NOAA ERDDAP server:

<https://coastwatch.pfeg.noaa.gov/erddap/info/erdSW1chlamday/index.html>

<https://coastwatch.pfeg.noaa.gov/erddap/info/erdMH1chlamday/index.html>.

The CHL values were averaged across thirteen 1 degree by 1 degree boxes which roughly parallel the bathymetry (Figure 1 in main text).

References

NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group; (2014): SeaWiFS Ocean Color Data; NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. https://dx.doi.org/10.5067/ORBVIEW-2/SEAWIFS_OC.2014.0

NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Chlorophyll Data; 2014 Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA. <https://dx.doi.org/10.5067/AQUA/MODIS/L3M/CHL/2014>

Hu C, Lee Z, Franz B (2012). "Chlorophyll algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference." *Journal of Geophysical Research: Oceans*, 117(C1).

ENSO Data

Description

Oceanic Nino Index and Dipole Mode Index.

Details

The ONI index is 3 month running mean of ERSST.v5 SST anomalies in the Niño 3.4 region (5°N-5°S, 120°-170°W)], based on centered 30-year base periods updated every 5 years.

The ONI was downloaded from <http://www.cpc.ncep.noaa.gov/data/indices/oni.ascii.txt>

The DMI is the monthly Dipole Mode Index. The DMI (also IOD index) is defined by the SSTA difference between the western Indian Ocean (10°S-10°N, 50°E-70°E) and the southeastern Indian Ocean (10°S-0°, 90°E-110°E). The data were downloaded from https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/dmi.long.data

References

Saji NH, Yamagata T (2003). "Possible impacts of Indian Ocean Dipole mode events on global climate." *Climate Research*, 25(2), 151-169. doi: 10.3354/cr025151.

Appendix H: Data

Table H1. Landings data and environmental covariates. N=Oct-Mar Kerala landings. S=Jul-Sep Kerala landings. Values have been rounded to 2 digits.

Year	log N	log S	Jun-Jul Precip	Apr-May Precip	Apr-May ns-SST	Oct-Dec ns-SST	Jun-Sep UPW	Jun-Sep ns-SST	Jun-Sep Bakun UPW	2.5 yr ave SST	ONI	Sep-Nov DMI
1983	11.71	8.77	5.53	0.33	29.27	28.76	0.70	27.67	420.36	28.44	0.48	0.11
1984	11.39	9.55	8.25	3.98	29.49	28.39	1.87	25.70	454.20	28.37	-0.51	-0.21
1985	10.76	9.86	6.83	4.50	29.53	28.54	1.44	25.91	542.46	28.24	-0.60	0.12
1986	9.19	5.24	7.05	2.46	29.46	28.23	1.26	25.86	631.25	28.06	0.25	0.19
1987	9.43	10.36	4.53	2.30	30.02	29.54	1.12	26.93	476.04	28.11	1.28	0.53
1988	11.24	9.85	12.44	3.25	29.69	28.60	0.56	27.08	238.77	28.40	-0.81	-0.02
1989	11.46	11.22	8.56	3.53	29.78	28.53	1.24	26.03	316.99	28.58	-0.60	0.02
1990	11.50	10.97	5.94	3.29	29.43	28.92	1.22	26.06	352.55	28.29	0.32	0.13
1991	10.42	9.49	11.92	1.49	29.72	28.13	1.04	26.43	298.60	28.20	0.65	0.36
1992	9.88	8.80	8.99	4.07	29.49	28.47	0.98	26.50	270.65	28.14	0.63	-0.28
1993	9.80	9.69	9.30	2.97	29.47	28.77	0.47	27.04	232.04	28.15	0.31	0.17
1994	6.57	5.91	8.72	4.01	29.36	28.56	1.48	26.11	147.73	28.17	0.48	0.84
1995	9.67	8.27	6.69	1.96	30.12	28.73	0.88	26.79	211.98	28.29	-0.16	0.06
1996	10.64	9.29	7.67	2.16	29.51	28.32	0.78	26.28	174.91	28.33	-0.46	-0.53
1997	10.57	9.61	8.29	1.97	30.19	29.80	0.93	27.64	105.08	28.47	1.17	1.30
1998	10.93	10.09	8.42	2.66	30.60	28.45	-0.10	27.79	-10.37	28.78	-0.06	-0.28
1999	11.45	10.80	7.42	5.22	28.86	28.16	0.74	26.72	69.47	28.83	-1.23	0.22
2000	11.20	11.31	5.18	4.82	29.11	29.13	1.12	26.72	195.69	28.41	-0.84	0.16
2001	11.11	10.78	9.07	7.39	29.38	28.62	1.30	26.19	126.03	28.19	-0.31	0.03
2002	11.98	11.15	5.54	4.52	29.84	28.80	1.12	26.73	134.28	28.26	0.63	0.60
2003	11.88	10.83	10.15	1.87	29.72	28.67	1.15	26.90	94.71	28.31	0.25	0.18
2004	11.90	11.04	6.97	4.65	29.70	29.14	1.07	26.67	387.97	28.42	0.45	0.25
2005	11.82	10.91	6.72	4.73	30.01	28.67	0.78	27.45	429.31	28.53	0.04	-0.07
2006	11.87	10.81	4.97	5.62	29.40	28.96	0.63	27.72	464.25	28.59	0.10	0.83
2007	12.01	11.00	11.90	3.26	29.81	28.60	0.27	27.87	369.20	28.73	-0.58	0.43
2008	11.51	11.06	8.74	2.55	29.18	28.92	0.62	27.49	317.37	28.62	-0.76	0.34
2009	11.58	11.21	8.07	4.61	29.78	29.03	0.66	27.48	405.72	28.63	0.31	0.27
2010	12.27	10.63	10.02	3.86	30.31	28.57	1.01	27.51	99.47	28.73	-0.45	-0.06
2011	12.18	11.01	6.85	4.11	29.57	29.06	1.06	27.45	110.04	28.70	-0.83	0.63
2012	12.44	11.71	5.78	1.94	29.31	29.38	0.69	27.66	131.89	28.68	-0.13	0.49
2013	11.40	9.92	8.88	3.36	29.79	28.64	1.18	26.70	158.15	28.69	-0.29	0.25
2014	11.30	10.76	5.05	3.31	29.88	29.02	0.59	27.78	216.61	28.63	0.14	0.33
2015	NA	10.20	8.32	3.62	30.00	29.55	0.39	28.42	152.07	28.70	1.49	0.70

Table H2. Chlorophyll-a from Box 4.

Year	Jul-Sep log CHL	Oct-Dec log CHL	Jan-Mar log CHL
1983	NA	NA	NA
1984	NA	NA	NA
1985	NA	NA	NA
1986	NA	NA	NA
1987	NA	NA	NA
1988	NA	NA	NA
1989	NA	NA	NA
1990	NA	NA	NA
1991	NA	NA	NA
1992	NA	NA	NA
1993	NA	NA	NA
1994	NA	NA	NA
1995	NA	NA	NA
1996	NA	NA	NA
1997	NA	NA	NA
1998	1.30	-0.20	-0.89
1999	1.46	-0.09	-1.22
2000	1.36	-0.80	-1.15
2001	1.62	-0.33	-1.22
2002	1.84	-0.67	-0.86
2003	2.51	-0.06	-0.76
2004	2.33	-0.35	-0.70
2005	1.67	0.18	-0.84
2006	1.88	-0.45	-1.24
2007	1.52	-0.53	-0.89
2008	1.79	-0.60	-1.03
2009	1.85	0.63	-1.01
2010	1.99	0.03	-0.68
2011	1.31	0.11	-1.12
2012	1.60	-0.26	-1.14
2013	1.46	-0.06	-0.88
2014	1.86	-0.83	-0.87
2015	NA	NA	NA