

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

28 August, 2019

## Introduction

Environmental variability is known to be a key driver of population variability of small forage fish such as sardines, anchovy and herring (Bakun 1996, Alheit and Hagen 1997, Cury et al. 2000, Checkley Jr. et al. 2017). In particular, ocean temperature and upwelling dynamics, along with density-dependent feedback, have been identified as important in affecting recruitment success and biomass of European and Pacific sardines (*Sardina pilchardus* and *Sardinops sagax*) (Jacobson and MacCall 1995, Rykaczewski and Checkley 2008, Alheit et al. 2012, Lindegren and Checkley Jr. 2012, Lindegren et al. 2013). Like other sardines, the Indian oil sardines show strong interannual fluctuations and larger decadal booms and busts. The Indian oil sardine offers an interesting case study to investigate the effects of environmental variability, particularly temperature and upwelling dynamics, as they occupy an ocean system that is warmer than that occupied by other sardines and have a strong seasonal cycle driven by the Indian summer monsoon.

The Indian oil sardine (*Sardinella longiceps* Valenciennes, 1847) is one of the most commercially important fish resources along the southwest coast of India (Figure 1) and historically has comprised approximately 25% of the catch biomass (Vivekanandan et al. 2003). Landings of the Indian oil sardine are highly seasonal and peak during and after the summer monsoon period (June through September), in conjunction with the onset and early relaxation of coastal upwelling. However, the landings of this small pelagic finfish are also highly variable from year to year. Small pelagics are well known to exhibit high variability in biomass due to the effects of environmental conditions on survival and recruitment (Bakun 1996, Alheit and Hagen 1997, Cury et al. 2000, Checkley Jr. et al. 2017). In this fishery, however, environmental conditions also affect exposure of sardines to the fishery. Until recently, the Indian oil sardine

fishery was artisanal and based on small human or low powered boats with no refrigeration. The fishery was confined to nearshore waters, and thus migration of fish schools in and out of the coastal zone greatly affected exposure to the fishery.

Researchers have examined a variety of environmental variables for their correlation with landings of the Indian oil sardine in order to understand the factors that drive landings variability. Precipitation during the southwest monsoon (Antony Raja 1969, 1974, Murty and Edelman 1971, Jayaprakash 2002) and the day of the monsoon arrival (Jayaprakash 2002) is thought to act as either a direct or indirect cue for spawning. Many studies have looked for correlations between precipitation, however the reported effects are positive in some studies and negative in others [Madhupratap et al. (1994);]. Researchers have also looked for and found correlations with various metrics of upwelling intensity, such as sea level at Cochin (Murty and Edelman 1971, Longhurst and Wooster 1990, Madhupratap et al. 1994, Jayaprakash 2002, Thara 2011), salinity and bottom sea temperature [Krishnakumar et al. (2008);], and with direct measures of productivity, such as nearshore zooplankton and phytoplankton abundance (Hornell 1910, Nair 1952, Nair and Subrahmanyam 1955, Madhupratap et al. 1994, George et al. 2012, Piontkovski et al. 2015, Menon et al. 2019). Researchers have also found correlations with near-shore sea surface temperature (SST); SST can affect both somatic growth rates and juvenile survival but also can cause fish to move off-shore and away from the shore-based fishery (Annigeri 1969, Prabhu and Dhulkhed 1970, Pillai 1991).

In this paper, we study the utility of environmental covariates to explain year-to-year variability in landings using the time series of quarterly Indian oil sardine landings from the SW coast of India. This time series is derived from a stratified sampling design that surveys landing sites along the coast and was first implemented in the 1950s (Srinath et al. 2005), however catch-at-length data are not available for the earlier years (prior to 2001) neither are stock size estimates nor fisheries independent data. A number of catch length studies are available but only for a small subset of years at specific landing sites. Thus traditional length- or age-structured models (e.g. Virtual Population Analysis) are not possible. Instead we use time-series models with covariates to model landings. Modeling and forecasting landings data (catch) using time-series models has a long tradition in fisheries and has been applied to many species (Mendelsohn 1981, Cohen and Stone 1987, Nobel and Sathianandan 1991, Stergiou and Christou 1996, Lloret et al. 2000, Georgakarakos et al. 2006, Prista et al. 2011, Lawer 2016), including oil sardines (Srinath 1998, Venugopalan and Srinath 1998). These models can be used to understand the variables associated with catch fluctuations and can be used to provide forecasts that assist fisheries planning. Unlike prior work on landings models with co-

variates, we use non-linear time-series models to allow a flexible effect of covariates and past catch on current landings. We also specifically focus on environmental covariates measured via remote-sensing. Remote-sensing data provide long time series of environmental data over a wide spatial extent at a daily and monthly resolution. In addition, a better understanding of how and whether remote sensing data explains variation in seasonal catch will support future efforts to use remote sensing data to improve catch forecasts.

## **Study Area**

Our analysis focuses on the Kerala coast (Figure 1) region of India, where the majority of the Indian oil sardines are landed and where oil sardines comprise ca. 40% of the marine fish catch (Srinath 1998, Vivekanandan et al. 2003). This area is in the Southeast Arabian Sea (SEAS), one of world's major upwelling zones, with seasonal peaks in primary productivity driven by upwelling caused by winds during the Indian summer monsoon (Madhupratap et al. 2001, Habeebrehman et al. 2008) between June and September. Within the SEAS, the coastal zone off Kerala between 9°N to 13°N has especially intense upwelling due to the combined effects of wind stress and remote forcing (Smitha et al. 2008, Smitha 2010). The result is a strong temperature differential between the near-shore and off-shore and high primary productivity and chlorophyll in this region during summer and early fall (Madhupratap et al. 2001, Habeebrehman et al. 2008, Jayaram et al. 2010, Raghavan et al. 2010, Smitha 2010, Chauhan et al. 2011). The primary productivity peaks subside after September while mesozooplankton abundances increase and remain high in the inter-monsoon period (Madhupratap et al. 2001).

## **Oil sardine life cycle and fishery**

The Indian oil sardine fishery is restricted to the narrow strip of the western India continental shelf, within 20 km from the shore (Figure 1). The yearly cycle (Figure 2) of the fishery begins at the start of spawning during June to July, corresponding with the onset of the southwest monsoon (Chidambaram 1950, Antony Raja 1969) when the mature fish migrate from offshore to coastal spawning areas. The spawning begins during the southwest monsoon period when temperature, salinity and suitable food availability are conducive for larval survival (Chidambaram 1950, Murty and Edelman 1971, Jayaprakash and Pillai 2000, Krishnakumar et al. 2008, Nair et al. 2016). Although peak spawning occurs in June to July, spawning continues into September (Hornell 1910, Hornell and Nayudu 1923, Antony Raja 1969, Prabhu and Dhulkhed 1970) and early- and late-spawning cohorts are evident in the length distributions

of the 0-year fish. Spawning occurs in shallow waters outside of the traditional range of the fishery (Antony Raja 1964), and after spawning the adults migrate closer to the coast and the spent fish become exposed to the fishery.

After eggs are spawned, they develop rapidly into larvae within 24 hrs (Nair 1959). The phytoplankton bloom that provide sardine larvae food is dependent upon nutrient influx from coastal upwelling and runoff from rivers during the summer monsoon and early fall. The blooms start in the south near the southern tip of India in June, increase in intensity and spread northward up the coast (Smitha 2010). Variation in the bloom initiation time and intensity leads to changes in the food supply and to corresponding changes in the growth and survival of larvae and in the later recruitment of 0-year sardines into the fishery (George et al. 2012). Oil sardines grow rapidly during their first few months, and 0-year fish (40mm to 100mm) from early spawning appear in the catch in August and September in most years (Antony Raja 1970, Nair et al. 2016). As the phytoplankton bloom spreads northward, the oil sardines follow, and the oil sardine fishery builds from south to north during the post-monsoon period. Oil sardines remain inshore feeding throughout the winter months, until March to May when the inshore waters warm considerably and sardines move off-shore to deeper waters (Chidambaram 1950). Catches of sardines are correspondingly low during this time for all size classes. The age at first maturity occurs at approximately 150 mm size (Nair et al. 2016), which is reached within one year. When the summer monsoon returns, the oil sardine cycle begins anew.

Overall, catches along the Kerala coast are fairly high throughout the year except during quarter 2 March-May (Figure 3). The age-distribution caught by the fishery varies through the year. The fishery is closed during June to mid-July during the monsoon and peak spawning, and when it resumes in mid-July, it is first dominated by 1-2.5 year old mature fish (Bensam 1964, Antony Raja 1969, Nair et al. 2016). In August or September a spike of 0-year (40mm) juveniles from the June spawning typically appears in the catch (Antony Raja 1969, Nair et al. 2016) and another spike of 0-year fish is sometimes seen in February from the late fall spawning (Prabhu and Dhulkhed 1967, 1970). From October through July, the catch is dominated by fish from 120mm-180mm (Antony Raja 1970, Prabhu and Dhulkhed 1970, Nair et al. 2016) which is a mix of 0-year, 1-year and 2-year fish (Nair et al. 2016, Rohit et al. 2018).

# Materials and Methods

## Sardine landing data

Quarterly fish landing data have been collected by the Central Marine Fisheries Research Institute (CMFRI) in Kochi, India, since the early 1950s using a stratified multi-stage sample design that takes into account landing centers, number of fishing days, and boat net combinations in fishing operations (Srinath et al. 2005). The quarterly landings for oil sardine landed from all gears in Kerala were obtained from CMFRI reports (1956-1984) and online databases (1985-2015) (CMFRI 1969, 1995, 2016, Pillai 1982, Jacob et al. 1987). The quarterly landing data were log-transformed to stabilize the seasonal variance.

## Remote-sensing data

We analysed monthly composites of the following environmental data derived from satellites: sea surface temperature (SST), chlorophyll-a (CHL), upwelling (UPW) and precipitation. The monthly means of the covariate time series are shown in Figure 4.

For sea surface temperature, we used Advanced Very-High Resolution Radiometer (AVHRR) data, which provides accurate nearshore SST values. Although the ICOADS product provides SST values for earlier years, ICOADS does not provide accurate nearshore temperatures. For 1981 to 2003, we used the Pathfinder Version 5.2 product on a 0.0417 degree grid. These data were developed by the Group for High Resolution Sea Surface Temperature (GHRSSST) and served by the US National Oceanographic Data Center. For 2004 to 2016, we used the NOAA CoastWatch SST products derived from NOAA's Polar Operational Environmental Satellites (POES). For chlorophyll-a, we used the chlorophyll-a products 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. For 2003 to 2017, we used the MODIS-Aqua product on a 0.05 degree grid. These CHL data are taken from measurements by the Moderate Resolution Imaging Spectroradiometer (MODIS) on NASA's Aqua Spacecraft. The SST and CHL data were averaged across thirteen 1 degree by 1 degree boxes which roughly parallel the bathymetry (Figure 1). The SST and CHL satellite data were retrieved from the NOAA ERDDAP server (Simons 2017).

For an index of coastal upwelling, we used the sea-surface temperature differential between near shore and 3 degrees offshore as described by Naidu et al. (1999) and Smitha et al. (2008). The index was computed as the average SST in box 4 off Kochi (Figure 1) minus the average SST in box 13. For SST, we used the remote-sensing sea-surface temperature data sets described above. This SST-based upwelling index has been validated as a more reliable metric of upwelling off the coast of Kerala compared to wind-based upwelling indices (Smitha et al. 2008). The SST-based upwelling index and chlorophyll-a blooms are strongly correlated (Figure 5).

Precipitation data were obtained from two different sources. The first was an estimate of the monthly precipitation (in mm) over Kerala from land-based rain gauges (Kothawale and Rajeevan 2017); these data are available from the Indian Institute of Tropical Meteorology and the data are available from the start of our landing data (1956). The second was a remote-sensing precipitation product from the NOAA Global Precipitation Climatology Project (Adler et al. 2016). This provides estimate of precipitation over the ocean using a global 2.5 degree grid. We used the 2.5 by 2.5 degree box defined by latitude 8.75 to 11.25 and longitude 73.25 to 75.75 for the precipitation off the coast of Kerala. These data are available from 1979 forward (NCEI 2017). The land and nearshore ocean precipitation data are highly correlated (Appendix), supporting the use of the land time series as a proxy for the precipitation over the ocean off the Kerala coast.

## Hypotheses

Our statistical analyses were structured around specific hypotheses (Table 1) concerning which remote-sensing covariates in which months should correlate with landings in specific quarters. These hypotheses were based on biological information concerning how environmental conditions affect sardine survival and recruitment and affect exposure of Indian oil sardines to the coastal fishery. The quarter 3 (Jul-Sep) catch overlaps the summer monsoon and the main spawning months. This is also the quarter where small 0-year fish from early spawning (June) often appear in the catch, sometimes in large numbers. Variables that affect or are correlated with movement of sardines inshore should be correlated with quarter 3 landings. In addition, pre-spawning (January-May) environmental conditions should be correlated with the spawning strength as adult oil sardines experience an acceleration of growth during this period along with egg development. The post-monsoon catch (Oct-May) is a mix of 0-year fish (less than 12 months old) and mature fish (greater than 12 months old). Variables that are correlated with

spawning strength and larval and juvenile survival should correlate with the post-monsoon catch both in the current year and in future years, one to two years after.

Our hypotheses (Table 1) focus mainly on two drivers: upwelling and ocean temperature. We also test hypotheses concerning precipitation as this has historically been an environmental covariate considered to influence the timing of oil sardine landings. Salinity has been proposed as responsible for spawning initiation; however we do not have long-term salinity data. In the Appendix, we show tests of chlorophyll-a but as with salinity the CHL time series is short (1997-2015) and the tests have low power.

## Explanatory models

We modeled the catches during the late-spawning season (quarter 3, July-September) separately from the post-monsoon season (October-March). Thus there is no seasonality in our catch time series, as we analyzed a yearly time series of quarter 3 catches separately from a yearly time series of post-monsoon catches. We divided the catch in this way for biological and statistical reasons. Catch in quarter 3 (July-September) captures a mix of spawning age fish as it overlaps with the tail end of the spawning season, is affected by a fishery closure from July to mid-August during the summer monsoon, and is periodically inflated by the appearance of small 0-year fish from early summer spawning. In addition, the covariates that affect the timing of spawning, movement of post-spawning mature fish inshore, and early egg and larval survival may be different than those that affect later growth, survival and shoaling that exposes fish to the inshore fishery. Analyzing catch and covariate time series without seasonality also had an important statistical benefit—we removed the problem of seasonality in the catch and all the covariates. The oil sardine life-cycle is seasonal and driven by the strong seasonality in this monsoon influenced system. A simple statistical model with quarters will explain much of the quarterly catch data since most of the yearly variability is due to seasonality and any environmental covariate with a similar seasonality will also show high correlation with the landings. Our goal was to explain year-to-year variability thus eliminating the confounding effect of seasonality in the data was important.

We tested ARIMA models on both quarter 3 and post-monsoon catch time series and found little support for auto-regressive errors (ARIMA models with a MA component) based on diagnostic tests of the residuals and model selection. The best supported ARIMA models were simple AR models ( $x_t = bx_{t-1} + \eta_t$ ). This lack of strong autocorrelation in residuals has been found in other studies that tested ARIMA models for forecasting small pelagic catch

(Stergiou and Christou 1996). We thus used AR-only models, however we tested both linear and non-linear models using a generalized additive model (GAM) of the form  $x_t = s(x_{t-1}) + \eta_t$ . The landings models were fit using conditional sum of squares (conditioning on the first 1 or 2 landings values in the time series). We investigated correlations between environmental variables and sardine catch using generalized additive models (GAMs, Wood 2017) to allow one to model the effect of a covariate as a flexible non-linear function. It was known that the effects of the environmental covariates were likely to be non-linear, albeit in an unknown way. Our approach is analogous to that taken by Jacobson and MacCall (1995) in a study of the effects of SST on Pacific sardine recruitment.

The first step in our analysis was to determine the catch model: the model for current catch as a function of the past catch. One feature of GAMs is that they allow the smoothing parameter of the response curve to be estimated. However we fixed the smoothing parameter at an intermediate value so that reasonably smooth responses were achieved and to limit the flexibility of the models being fit. Multi-modal or overly flexible response curves would not be realistic for our application. We used GAMs with smooth terms represented by penalized regression splines (Wood 2011, using the mgcv package in R) and fixed the smoothing term at an intermediate value (sp=0.6).

Our catch models took the form

- random walk:  $\ln(C_{i,t}) = \alpha + \ln(C_{j,t-1}) + \varepsilon_t$
- AR-1:  $\ln(C_{i,t}) = \alpha + \phi_1 \ln(C_{j,t-1}) + \varepsilon_t$
- AR-2:  $\ln(C_{i,t}) = \alpha + \phi_1 \ln(C_{j,t-1}) + \phi_2 \ln(C_{k,t-2}) + \varepsilon_t$
- non-linear:  $\ln(C_{i,t}) = \alpha + s_1(\ln(C_{j,t-1})) + s_2(\ln(C_{k,t-2})) + \varepsilon_t$

where  $\ln(C_{i,t})$  is the log catch in the current year  $t$  in season  $i$ . We modeled two different seasons: 3rd quarter catch  $C_{m,t}$  (July-September) which is during the late part of the summer monsoon and post-summer monsoon catch  $C_{p,t}$  (October-June). The catches were logged to stabilize and normalize the variance.  $s()$  is a non-linear function estimated by the GAM algorithm. The model is primarily statistical, meaning it should not be thought of as being a population growth model. We tested models with prior year post-summer monsoon catch ( $C_{p,t-1}$ ) and 3rd quarter catch ( $C_{m,t-1}$ ) as the explanatory catch variable. The catch models were fit to 1982 to 2015 catch data, corresponding to the years where the SST, upwelling and precipitation data were available. F-tests and AIC on nested sets of models (Wood et al. 2016) were used to evaluate the support for the catch models and later for the covariate models.

After selection of the best model with the 1982-2015 data, the fitting was repeated with the 1956-1981 and 1956-2015 catch data to confirm the form of the catch models.

Once the catch models were determined, the covariates were studied individually and then jointly. As with the catch models, F-tests and AIC on nested sets of GAM models were used to evaluate the support for models with covariates. The smoothing term was fixed at an intermediate value ( $sp=0.6$ ) instead of treated as an estimated variable. Our models for catch with covariates took the form

$$\ln(C_{i,t}) = M + s_1(V_{1,t}) + s_2(V_{2,t}) + \varepsilon_t$$

where  $M$  was the best catch model from step 1. The covariate model seeks to explain unexplained variance in the the catch model via a linear or non-linear function of the covariates,  $V_1$  and  $V_2$ . The covariates tested are those discussed in the section on covariates that have been hypothesized to drive the size of the sardine biomass exposed to the fishery.

## Results

### Catches in prior seasons as explanatory variables

Models with  $\ln(N_{t-1})$  or  $\ln(S_{t-1})$  as a covariate were strongly supported over the “naive” model: the “last year’s catch” model (Table 2, time-dependency test). The “naive” model has no estimated parameters and is a standard null model for time series modeling. There was support for including the post-summer monsoon catch in the previous year as an explanatory variable for the quarter 3 catch. Post-monsoon catch in the prior year was more supported as an expanatory covariate than quarter 3 catch in the prior year. Addition of the catch two years prior,  $\ln(N_{t-2})$  or  $\ln(S_{t-2})$ , did not reduce AIC and for  $\ln(N_{t-2})$  led to either no decrease in the residual error (MASE) or an increased the residual error for the model with linearity (Table 2, Linearity test). Addition of  $\ln(S_{t-2})$  did decrease the residual errors, but the was not warranted given the increased number of estimated parameters based on AIC. We also tested the support for non-linearity in the effect of the prior year catch on the quarter 3 catch. This was done by comparing models with  $\ln(N_{t-1})$  or  $\ln(S_{t-1})$  included as a linear term or as a non-linear function  $s()$  (Table 2, Linearity test). The residual error decreased using a non-linear response at the cost of 1.4 degrees of freedom. The result was only weak (non-significant) support for allowing a non-linear response based on AIC and the F-test. The full set of models tested,

including tests using catch during the spawning months in previous seasons as a covariate are shown in Tables A1 and A2. The results were the same if we used the full landings data set from 1956 to 2015 (Table A3). Overall, the landings in prior seasons was only weakly explanatory for the quarter 3 catch, and the maximum adjusted  $R^2$  for these models was less than 30% (Table 2).

The results on model structure were similar for models of the landings ( $N_t$ ) during the post-summer monsoon months (Table 3), but the models explained more of the variance. The most supported model for  $N_t$  (Table 3) used a non-linear response to landings during the post-monsoon months of the previous season  $\ln(N_{t-1})$  with a non-linear response to quarter 3 landings two years prior  $\ln(S_{t-2})$ . There was low support for including landings earlier than two seasons prior or for using the quarter 3 landings during in the immediately prior season (Tables A4, A5, and A6). We did not test models for the October-June catch using the quarter 3 (July-September) catch in the current fishing season, so immediately prior. These data would not be available in a forecasting setting as the data require 12 months to process.

## Environmental covariates as explanatory variables

There was no support for using precipitation during the summer monsoon (June-July) or pre-monsoon period (April-May) as an explanatory variable for the catch during quarter 3 (Table B1) nor the non-spawning months (Table B2). This was the case whether precipitation in the current or previous season was used, if precipitation was included as non-linear or non-linear effect, or if the smoothing term (degree of non-linearity allowed) was estimated and thus not constrained, and if either precipitation during monsoon (June-July) or pre-monsoon (April-May) were used as the covariate. Quarter 3 overlaps with the spawning period and precipitation is often thought to trigger spawning, however we were unable to find any consistent association of catch during these spawning and early-post spawning months with precipitation.

However, we found significant correlation between average SST during the early post-spawning period (July-September) and quarter 3 catch during the spawning season (Table 3, Table B3). Sea surface temperature has been found to be correlated with sardine recruitment in a number of other studies [Jacobson and MacCall (1995); add the others]. The residual error was lowest in a model with SST in both the prior year and two years prior included and with a non-linear response for both (Appendix). This a similar result to Jacobson and MacCall (1995) who found that SST in multiple prior years was supported as explanatory variables for sardine recruitment and productivity. However the reduction in degrees of freedom was high for this

model and it was not supported based on AIC. The model with the lowest AIC was a model with only a non-linear response of current year July-September SST. The response shows a step-response with a negative effect at low temperatures and then an increased effect at higher temperatures (Figure 6). This type of step-response has been found in studies of the effect of SST on recruitment in Pacific sardines (Jacobson and MacCall 1995). In our case, the SST effect is for the same months as the catch thus the effect is associated either with spawning or movement of mature fish into the nearshore area (where they are exposed to the fishery). The  $R^2$  for this model was 0.37 (Table 4).

*eeh: this is not right. Need to fix.* The strongest predictor of the catch during quarter 3 however was the upwelling strength during January-March (4-6 months prior) in the previous season (Tables 4 and B4). January-March is the period when adults are preparing for spawning, however January-March is a period when the young of the year and age-1 fish would be found feeding in large shoals in the coastal region and this time of year has the second highest catches and catches that are dominated by small-sized fish (*citation*). The best supported model included only the upwelling strength in the prior year, without SST included as a covariate (Table 4). The  $R^2$  for this model was 0.60 and the fitted versus observed catches for this model are shown in Figure 7.

*eeh: this is not right. Need to fix.* For catch in the post-summer monsoon months (Oct-Jun), sea surface temperature during July to September in the current season was a significant predictor similar to what was found for catches in the spawning months (quarter 3). Upwelling in the prior season was also a significant predictor, but the important months were Oct-Dec. This is period is important for larval and early juvenile survival and growth, and other studies have also found this to be a critical period for future stock size in sardines (*citation*). The model with the lowest residual variance and lowest AIC was the model which included both the SST and upwelling covariates (Table 4). The  $R^2$  for this model was 0.73 and the fitted versus observed catches for this model are shown in Figure 7.

Chlorophyll-a density is speculated to be an important predictor of larval sardine survival and growth. In addition, sardines shoal in response to coastal chlorophyll blooms, which brings them in contact with the coastal fisheries. Thus chlorophyll-a density is assumed to be an important driver of future or current sardine catches. We had chlorophyll-a remote-sensing data only from 1998 onward. Our simplest covariate model required 5 degrees of freedom, thus we were limited in the analyses we could conduct. In addition, the years, 1998-2014, have relatively low variability in catch sizes; the logged catch sizes during this period range from 10-11 during quarter 3 and 11-12 during the other three quarters. Second degree polynomial

models were fit (Appendix C) to the average log chlorophyll-a density in the current and prior season from quarter 3 (Jul-Sep), 4 (Oct-Dec), and 1 (Jan-Mar). Chlorophyll-a density was not a significant predictor for the spawning catch for any of the tested combinations of current or prior season and quarter. The only significant effect was seen for post-summer monsoon catches using chlorophyll-a density in Oct-Dec of the current and prior season (Table C1). This matches results which found that the upwelling index in October-December of the prior season was a predictor for the post-summer monsoon catch. The SST-based upwelling index and chlorophyll-a density are both indices of low-trophic level productivity.

## Discussion

*I am still trying to decide how to write the discussion. How does this relate to other work on sardines and esp the correlation between SST and landings?*

Sardines in all the world's ecosystems exhibit large fluctuations in abundance (Baumgartner et al. 1992, Schwartzlose et al. 2010). These small forage fish are strongly influenced by natural variability in upwelling driven by both large-scale forces, i.e. El Nino patterns, and by changes in winds and currents and in addition local conditions of temperature, salinity, and oxygen levels have both direct and indirect on sardine recruitment and survival.

Many studies on Pacific sardines have looked at the correlation between sea surface temperature (SST) and recruitment. Temperature can have direct effect, an indirect effect on food availability or affect survival (Houde 1987). Studies in the California Current System, have found that SST explains year-to-year variability in Pacific sardine recruitment (Jacobson and MacCall 1995, Checkley Jr. et al. 2009, 2017, Lindegren and Checkley Jr. 2012). Consistent with these studies, we also found that SST was the covariate that explained variability in catch anomalies (difference between the landing prediction from prior years' catches).

McClatchie et al. (2010) found no SST relationship with SST and Pacific sardine recruitment, however their analysis used a linear relationship while both other studies that found a relationship (Jacobson and MacCall 1995, Checkley Jr. et al. 2017) allowed a non-linear relationship. Both Jacobson and MacCall (1995) and Checkley et al (2017) found a step-like response function for temperature, where lower temperatures were poor and had negative effects and then at a threshold value the effect became positive. Our analysis found a similar step-like effect function.

There were three outlier years when catch were much lower than expected based on prior

catches. For these years, sea surface temperature improved the model fit greatly (Figure 8).

Remote sensing satellites can be used to detect changes in ocean physical, biological and chemical properties, such as surface temperature, winds, surface height, surface waves, rainfall and surface salinity, as well as the ecosystem and water quality changes. Using these covariates to improve landing forecasts has the potential to improve fisheries management for small pelagics (Tommasi et al. 2016). However, improving forecasts using these covariates is not obvious. We found no improvement in forecasts over a simple model using a non-linear effect of past catch using many of the covariates known to be associated with spawning, growth and survival: precipitation, upwelling indices, and chlorophyll-a.

Nonetheless, remote-sensing derived chlorophyll, SST and SSH estimates have been used to estimate sardine habitat of the west coast of the USA (Zwolinski et al. 2011). In India, satellite measurements of chlorophyll have been used to predict high densities of oil sardines. This is used a functional tool for producing forecasts for fisherman and has lead to increases in their efficiency (catch per unit effort) (*citation*). The weak performance of our upwelling covariates may be due to the coarse correspondence between the temperature differential and other aspects of upwelling intensity that are critical to oil-sardines—the spatial extent both along the coast and off the coast and the timing of the start of upwelling.

## References

- Adler, R., J.-J. Wang, M. Sapiano, G. Huffman, L. Chiu, P. P. Xie, R. Ferraro, U. Schneider, A. Becker, D. Bolvin, E. Nelkin, G. Gu, and N. C. Program. 2016. Global Precipitation Climatology Project (GPCP) Climate Data Record (CDR), version 2.3 (monthly). Report, National Centers for Environmental Information.
- Alheit, J., and E. Hagen. 1997. Long-term climate forcing of European herring and sardine populations. *Fisheries Oceanography* 6:130–139.
- Alheit, J., T. Pohlmann, M. Casini, W. Greve, R. Hinrichs, M. Mathis, K. O'Driscoll, R. Vorberg, and C. Wagner. 2012. Climate variability drives anchovies and sardines into the North and Baltic Seas. *Progress in Oceanography* 96:128–139.
- Annigeri, G. G. 1969. Fishery and biology of the oil sardine at Karwar. *Indian Journal of Fisheries* 16:35–50.
- Antony Raja, B. T. 1964. Some aspects of spawning biology of Indian oil sardine Sar-

- Sardinella longiceps* Valenciennes. Indian Journal of Fisheries 11:45–120.
- Antony Raja, B. T. 1969. Indian oil sardine. CMFRI Bulletin 16:1–142.
- Antony Raja, B. T. 1970. Estimation of age and growth of the Indian oil sardine, *Sardinella longiceps* Val. Indian Journal of Fisheries 17:26–42.
- Antony Raja, B. T. 1974. Possible explanation for the fluctuation in abundance of the Indian oil sardine, *Sardinella longiceps* Valenciennes. Proceedings of the Indo-Pacific Fisheries Council 15:241–252.
- Bakun, A. 1996. Patterns in the ocean: Ocean processes and marine population dynamics. Book, California Sea Grant, in cooperation with Centro de Investigaciones Biologicas del Noroeste, La Paz, Mexico.
- Baumgartner, T. R., A. Soutar, and V. Ferreira-Bartrina. 1992. Reconstruction of the history of the Pacific sardine and northern anchovy populations over the past two millennia from sediments of the Santa Barbara basin, California. CalCOFI Report 33:24–40.
- Bensam, P. 1964. Growth variations in the Indian oil sardine, *Sardinella longiceps* Valenciennes. Indian Journal of Fisheries 11 A:699–708.
- Chauhan, O. S., B. R. Raghavan, K. Singh, A. S. Rajawat, U. Kader, and S. Nayak. 2011. Influence of orographically enhanced SW monsoon flux on coastal processes along the SE Arabian Sea. Journal of Geophysical Research. Oceans 116:C12037.
- Checkley Jr., D. M., J. Alheit, Y. Oozeki, and C. Roy. 2009. Climate change and small pelagic fish. Edited Book, Cambridge University Press, Cambridge.
- Checkley Jr., D. M., R. G. Asch, and R. R. Rykaczewski. 2017. Climate, anchovy, and sardine. Annual Review of Marine Science 9:469–493.
- Chidambaram, K. 1950. Studies on the length frequency of oil sardine, *Sardinella longiceps* Cuv. & Val. and on certain factors influencing their appearance on the Calicut coast of the Madras Presidency. Proceedings of Indian Academy of Sciences 31:352–286.
- CMFRI. 1969. Marine fish production in India 1950-1968. Bulletin of the Central Marine Fisheries Research Institute 13:1–171.
- CMFRI. 1995. Marine fish landings in India during 1985-93. Marine Fisheries Information Service, Technical and Extension Series 136:1–33.
- CMFRI. 2016. Marine fishery landings 1985-2015. Report, Central Marine Fisheries

Research Institute.

Cohen, Y., and N. Stone. 1987. Multivariate time series analysis of the Canadian fisheries system in Lake Superior. *Canadian Journal of Fisheries and Aquatic Sciences* 44:171–181.

Cury, P., A. Bakun, R. J. M. Crawford, A. Jarre, R. A. Quinones, L. J. Shannon, and H. M. Verheyen. 2000. Small pelagics in upwelling systems: Patterns of interaction and structural changes in “wasp-waist” ecosystems. *ICES Journal of Marine Science* 57:603–618.

Georgakarakos, S., D. Doutsouvas, and V. Valavanis. 2006. Time series analysis and forecasting techniques applied on loliginid and ommastrephid landings in Greek waters. *Fisheries Research* 78:55–71.

George, G., B. Meenakumari, M. Raman, S. Kumar, P. Vethamony, M. T. Babu, and X. Verlecar. 2012. Remotely sensed chlorophyll: A putative trophic link for explaining variability in Indian oil sardine stocks. *Journal of Coastal Research* 28:105–113.

Habeebrehman, H., M. P. Prabhakaran, J. Jacob, P. Sabu, K. J. Jayalakshmi, C. T. Achuthankutty, and C. Revichandran. 2008. Variability in biological responses influenced by upwelling events in the eastern Arabian Sea. *Journal of Marine Systems* 74:545–560.

Hornell, J. 1910. Report on the results of a fishery cruise along the Malabar coast and to the Laccadive Islands in 1908. *Madras Fishery Bulletin* 4:76–126.

Hornell, J., and M. R. Nayudu. 1923. A contribution to the life history of the Indian sardine with note, on the plankton of the Malabar coast. *Madras Fishery Bulletin* 17:129.

Houde, E. D. 1987. Fish early life dynamics and recruitment variability. *American Fisheries Society Symposium* 2:17–29.

Jacob, T., V. Rajendran, P. K. M. Pillai, J. Andrews, and U. K. Satyavan. 1987. An appraisal of the marine fisheries in Kerala. Report, Central Marine Fisheries Research Institute.

Jacobson, L. D., and A. D. MacCall. 1995. Stock-recruitment models for Pacific sardine (*Sardinops sagax*). *Canadian Journal of Fisheries and Aquatic Sciences* 52:566–577.

Jayaprakash, A. A. 2002. Long term trends in rainfall, sea level and solar periodicity: A case study for forecast of Malabar sole and oil sardine fishery. *Journal of the Marine Biological Association of India* 44:163–175.

Jayaprakash, A. A., and N. G. K. Pillai. 2000. The Indian oil sardine. Pages 259–281 in V. N. Pillai and N. G. Menon, editors. *Marine fisheries research and management*. Book

Section, Central Marine Fisheries Research Institute, Kerala, India.

Jayaram, C., N. Chacko, K. A. Joseph, and A. N. Balchand. 2010. Interannual variability of upwelling indices in the southeastern Arabian Sea: A satellite based study. *Ocean Science Journal* 45:27–40.

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

Krishnakumar, P. K., K. S. Mohamed, P. K. Asokan, T. V. Sathianandan, P. U. Zacharia, K. P. Abdurahiman, V. Shettigar, and N. R. Durgekar. 2008. How environmental parameters influenced fluctuations in oil sardine and mackerel fishery during 1926-2005 along the southwest coast of India? *Marine Fisheries Information Service, Technical and Extension Series* 198:1–5.

Lawer, E. A. 2016. Empirical modeling of annual fishery landings. *Natural Resources* 7:193–204.

Lindgren, M., and D. M. Checkley Jr. 2012. Temperature dependence of Pacific sardine (*Sardinops sagax*) recruitment in the California Current Ecosystem revisited and revised. *Canadian Journal of Fisheries and Aquatic Sciences* 70:245–252.

Lindgren, M., D. M. Checkley, T. Rouyer, A. D. MacCall, and N. C. Stenseth. 2013. Climate, fishing, and fluctuations of sardine and anchovy in the California current. *Proceedings of the National Academy of Sciences* 110:13672–13677.

Lloret, J., J. Leonart, and I. Sole. 2000. Time series modelling of landings in Northwest Mediterranean Sea. *ICES Journal of Marine Science* 57:171–184.

Longhurst, A. R., and W. S. Wooster. 1990. Abundance of oil sardine (*Sardinella longiceps*) and upwelling on the southwest coast of India. *Canadian Journal of Fisheries and Aquatic Sciences* 47:2407–2419.

Madhupratap, M., T. C. Gopalakrishnan, P. Haridas, and K. K. C. Nair. 2001. Mesozooplankton biomass, composition and distribution in the Arabian Sea during the fall intermonsoon: Implications of oxygen gradients. *Deep Sea Research Part II: Topical Studies in Oceanography* 48:1345–1368.

Madhupratap, M., S. R. Shetye, K. N. V. Nair, and S. R. S. Nair. 1994. Oil sardine and Indian mackerel: Their fishery, problems and coastal oceanography. *Current Science* 66:340–

McClatchie, S., R. Goericke, G. Auad, and K. Hill. 2010. Re-assessment of the stock-recruit and temperature-recruit relationships for Pacific sardine (*Sardinops sagax*). Canadian Journal of Fisheries and Aquatic Sciences 67:1782–1790.

Mendelssohn, R. 1981. Using Box-Jenkins models to forecast fishery dynamics: Identification, estimation and checking. Fishery Bulletin 78:887–896.

Menon, N. N., S. Sankar, A. Smitha, G. George, S. Shalin, S. Sathyendranath, and T. Platt. 2019. Satellite chlorophyll concentration as an aid to understanding the dynamics of Indian oil sardine in the southeastern Arabian Sea. Marine Ecology Progress Series 617-618:137–147.

Murty, A. V. S., and M. S. Edelman. 1971. On the relation between the intensity of the south-west monsoon and the oil-sardine fishery of India. Indian Journal of Fisheries 13:142–149.

Naidu, P. D., M. R. R. Kumar, and V. R. Babu. 1999. Time and space variations of monsoonal upwelling along the west and east coasts of India. Continental Shelf Research 19:559–572.

Nair, P. G., S. Joseph, V. Kripa, R. Remya, and V. N. Pillai. 2016. Growth and maturity of Indian oil sardine *Sardinella longiceps* (Valenciennes, 1847) along southwest coast of India. Journal of Marine Biological Association of India 58:64–68.

Nair, R. V. 1952. Studies on the revival of the Indian oil sardine fishery. Proceedings of Indo-Pacific Fisheries Council 2:1–15.

Nair, R. V. 1959. Notes on the spawning habits and early life-history of the oil sardine, *Sardinella longiceps* Cuv. & Val. Indian Journal of Fisheries 6:342–359.

Nair, R. V., and R. Subrahmanyam. 1955. The diatom, *Fragilaria oceanica* Cleve, an indicator of abundance of the Indian oil sardine, *Sardinella longiceps* Cuv. and Val. Current Science 24:41–42.

NCEI. 2017. Global Precipitation Climatology Project Monthly Product Version 2.3. Report, National Centers for Environmental Information.

Nobel, A., and T. V. Sathianandan. 1991. Trend analysis in all-India mackerel catches using ARIMA models. Indian Journal of Fisheries 38:119–122.

Pillai, V. N. 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.

Pillai, V. N. 1991. Salinity and thermal characteristics of the coastal waters off southwest coast of India and their relation to major pelagic fisheries of the region. *Journal of the Marine Biological Association of India* 33:115–133.

Piontkovski, S., H. Al Oufi, and S. Al Jufaily. 2015. Seasonal and interannual changes of Indian oil sardine, *Sardinella longiceps*, landings in the governorate of Muscat (the Sea of Oman). *Marine Fisheries Review* 76:48–58.

Prabhu, M. S., and M. H. Dhulkhed. 1967. On the occurrence of small-sized oil sardine *Sardinella longiceps* Val. *Current Science* 35:410–411.

Prabhu, M. S., and M. H. Dhulkhed. 1970. The oil sardine fishery in the Mangalore zone during the seasons 1963-64 and 1967-68. *Indian Journal of Fisheries* 17:57–75.

Prista, N., N. Diawara, M. J. Costa, and C. Jones. 2011. Use of SARIMA models to assess data-poor fisheries: A case study with a sciaenid fishery off Portugal. *Fisheries Bulletin* 109:170–185.

Raghavan, B. R., T. Deepthi, S. Ashwini, S. K. Shylini, M. Kumaraswami, S. Kumar, and A. A. Lotlike. 2010. Spring inter monsoon algal blooms in the eastern Arabian Sea: Shallow marine encounter off Karwar and Kumbla coast using a hyperspectral radiometer. *International Journal of Earth Sciences and Engineering* 3:827–832.

Rohit, P., M. Sivadas, E. M. Abdussamad, A. Margaret Muthu Rathinam, K. P. S. Koya, U. Ganga, S. Ghosh, K. M. Rajesh, K. Mohammed Koya, A. Chellappan, K. G. Mini, G. George, S. K. Roul, S. Surya, S. Sukumaran, E. Vivekanandan, T. B. Retheesh, D. Prakasan, M. Satish Kumar, S. Mohan, R. Vasu, and V. Supraba. 2018. Enigmatic Indian oil sardine: An insight. Report, ICAR-Central Marine Fisheries Research Institute.

Rykaczewski, R. R., and D. M. Checkley. 2008. Influence of ocean winds of the pelagic ecosystem in upwelling regions. *Proceedings of the National Academy of Science* 105:1965–1970.

Schwartzlose, R. A., J. Alheit, A. Bakun, T. R. Baumgartner, R. Cloete, R. J. M. Crawford, W. J. Fletcher, Y. Green-Ruiz, E. Hagen, T. Kawasaki, D. Lluch-Belda, S. E. Lluch-Cota, A. D. MacCall, Y. Matsuura, M. O. Nevárez-Martínez, R. H. Parrish, C. Roy, R. Serra, K. V. Shust, M. N. Ward, and J. Z. Zuzunaga. 2010. Worldwide large-scale fluctuations of sardine and

anchovy populations. *South African Journal of Marine Science* 21:289–347.

Simons, R. A. 2017. ERDDAP. <https://coastwatch.pfeg.noaa.gov/erddap>. Report, Monterey, CA: NOAA/NMFS/SWFSC/ERD.

Smitha, B. R. 2010. Coastal upwelling of the south eastern Arabian Sea — an integrated approach. Thesis.

Smitha, B. R., V. N. Sanjeevan, K. G. Vimalkumar, and C. Revichandran. 2008. On the upwelling of the southern tip and along the west coast of India. *Journal of Coastal Research* 24:95–102.

Srinath, M. 1998. Exploratory analysis on the predictability of oil sardine landings in Kerala. *Indian Journal of Fisheries* 45:363–374.

Srinath, M., S. Kuriakose, and K. G. Mini. 2005. Methodology for estimation of marine fish landings in India. Page 57 in CMFRI special publications. Book Section, Central Marine Fisheries Research Institute.

Stergiou, K. I., and E. D. Christou. 1996. Modeling and forecasting annual fisheries catches: Comparison of regression, univariate and multivariate time series methods. *Fisheries Research* 25:105–138.

Thara, K. J. 2011. Response of eastern Arabian Sea to extreme climatic events with special reference to selected pelagic fishes. Thesis.

Tommasi, D., C. A. Stock, K. Pégion, G. A. Vecchi, R. D. Methot, M. A. Alexander, and D. M. C. Jr. 2016. Improved management of small pelagic fisheries through seasonal climate prediction. *Ecological Applications* 27:378–388.

Venugopalan, R., and M. Srinath. 1998. Modelling and forecasting fish catches: Comparison of regression, univariate and multivariate time series methods. *Indian Journal of Fisheries* 45:227–237.

Vivekanandan, E., M. Srinath, V. N. Pillai, S. Immanuel, and K. N. Kurup. 2003. Marine fisheries along the southwest coast of India. Pages 759–792 in G. Silvestre, L. Garces, I. Stobutzki, C. Luna, M. Ahmad, R. A. Valmonte-Santos, L. Lachia-Alino, P. Munro, V. Christensen, and D. Pauly, editors. Assessment, management, and future directions for coastal fisheries in Asian countries. Book Section, WorldFish Center Conference Proceedings 67, WorldFish Center, Penang.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood es-

timation of semiparametric generalized linear models. *Journal of the Royal Statistical Society B* 73:3–36.

Wood, S. N. 2017. Generalized additive models: An introduction with R. 2nd editions. Book, Chapman; Hall/CRC.

Wood, S. N., N. Pya, and B. Safken. 2016. Smoothing parameter and model selection for general smooth models (with discussion). *Journal of the American Statistical Association* 111:1548–1575.

Zwolinski, J. P., R. L. Emmett, and D. A. Demer. 2011. Predicting habitat to optimize sampling of Pacific sardine (*Sardinops sagax*). *ICES Journal of Marine Science* 68:867–879.

## Figure Legends

Figure 1. Close up of Kerala State with the latitude/longitude boxes used for the satellite data. Kerala State is marked in grey and the oil sardine catch from this region is being modeled.

Figure 3. Quarterly catch data 1956-2014 from Kerala. The catches have a strong seasonal pattern with the highest catches in quarter 4 Note that quarter 3 is July-Sept and that the fishery is closed July 1 to Aug 15, thus the fishery is only open 1.5 months in quarter 3. The mean catch (metric tonnes) in quarters 1 to 4 are 38, 19.2, 30.9, and 59.9 metric tonnes respectively.

Figure 2. The sardine life-cycle in the SE Indian Ocean and how it interacts with the fishery.

Figure 4. Remote sensing covariates used in the analysis. All data are monthly averages over Box 4 in Figure 1 on the Kerala coast off of Kochi. Panel A) Upwelling Index. The upwelling index is the difference between the near-shore sea surface temperature (SST) and the off-shore SST defined as 3 degrees longitude offshore. Panel B) Surface chlorophyll-a (CHL). The CHL data are only available from 1997 onward. Panel C) Sea surface temperature constructed from Advanced Very High Resolution Radiometer (AVHRR). Panel D) Average daily rainfall (mm/day) off the Kerala coast.

Figure 5. Key oil sardine life-history events overlaid on the monthly SST in the near-shore and off-shore and the near-shore CHL.

Figure 6. Effects of covariates estimated from the GAM models. Panel A) Effect of SST during the spawning months (Jul-Sep) on catch during the spawning months. Low SST is associated with lower than expected catch during the spawning months. Panel B) Effect of upwelling (inshore/off-shore SST differential) during Jan-Mar of the prior season on catch during the next spawning months (Jul-Sep). The index is the difference between offshore and inshore SST, thus a negative value indicates warmer coastal surface water than off-shore. Warm coastal water during Jan-Mar when sardines are foraging along the coast, is associated with lower catch during the next spawning season. Panel C) Effect of SST during the spawning months (Jul-Sep) on catch during the subsequent non-spawning months (Oct-Jun). Low SST is associated with lower than expected catch during the following non-spawning months. Panel D) Effect of upwelling (inshore/off-shore SST differential) during Oct-Dec of the prior season on catch during the non-spawning months (Oct-Jun) the next season. Strong upwelling (positive upwelling index) in the early larval and juvenile period (Oct-Dec) is associated with higher than expected catch in the next season.

Figure 8. Fitted versus observed catch with models with and without environmental covariates. Panel A) Fitted versus observed log catch in the spawning months with only non-spawning catch in the previous season as the covariate:  $S_t = s(N_{t-1}) + \varepsilon$ . Panel B) Fitted versus observed log catch in the spawning months with Jan-Mar upwelling in the prior season added as a covariate to the model in panel A. This model was:  $S_t = s(N_{t-1}) + s(V_{t-1}) + \varepsilon$ . Panel C) Fitted versus observed log catch in the non-spawning months with only non-spawning catch in the previous season and spawning catch two season prior as the covariates:  $N_t = s(N_{t-1}) + s(S_{t-2}) + \varepsilon$ . Panel D) Fitted versus observed log catch in the non-spawning months with current season SST in Jul-Sep ( $V$ ) and Oct-Dec upwelling in the prior season ( $W$ ) added as covariates. This model was  $N_t = s(N_{t-1}) + s(S_{t-2}) + s(V_t) + \beta W_{t-1} + \varepsilon$ .  $W$  was added linearly since the data were insufficient to estimate four non-linear effects.

\begin{document}

Table 1. Hypotheses for covariates affecting landings.  $S_t$  is quarter 3 (July-September) catch in the current season,  $S_{t-1}$  is quarter 3 catch in the previous season.  $N_t$  is the post-monsoon October-March catch in the current season and  $N_{t-1}$  is the October-March catch in the prior season. Because the fishing season is July-June,  $N_t$  spans two calendar years. DD = hypotheses related to effects of past abundance (landings) on current abundance. S = hypotheses related to spawning. L = hypotheses related to larval and juvenile growth and survival. A = hypotheses affecting all ages.

Hypothesis	Resp.	Covariates
DD1. $S_t$ is dominated by mature age 2+ fish, thus abundance of the 1-yr and 2-yr ages in the prior season (Oct-Mar catch) should be correlated with the abundance of mature fish this year.	$S_t$	$N_{t-1}$
DD2. Abundance of 1-yr and 2-yr fish should be correlated with strength of the cohorts from the previous two seasons. The quarter 3 catch, dominated by mature fish, in the prior two years is expected to be correlated with post-monsoon catch.	$N_t$	$S_{t-1}$ and $S_{t-2}$
DD3. Because age 2 fish appear in the post-monsoon catch, we also expect the post-monsoon catch (dominated by age 1 and 2) in the previous season to be correlated with the post-monsoon catch in the current season. Post-monsoon catch two seasons prior should be minimally correlated with current post-monsoon catch.	$N_t$	$N_{t-1}$
S1. The onset of monsoon precipitation triggers movement of adults from offshore to spawning areas due to changes in salinity, turbulence or noise. Spent adults migrate inshore and are exposed to the fishery. Strong spawning affects post-monsoon catch in current and future seasons.	$S_t$	Jun-Jul precipitation in year $t$
S2. The level of precipitation in pre-monsoon months predicts spawning strength.	$S_t$	Apr-May precipitation in year $t$
S3. Precipitation initiates and supports spawning. Spawning affects post-monsoon catch in current and future seasons.	$N_t$	Apr-May and Jun-Jul precipitation in year $t$ and $t - 1$
S4. Extremely high upwelling brings poorly oxygenated water and very low temperatures to the surface causing mature fish to avoid nearshore areas. Avoidance of nearshore leads to lower exposure to the fishery and lower catch.	$S_t$	Jun-Sep upwelling index in year $t$
S5. Extreme heat events in the pre-spawning months cause mature fish to move offshore away from productive feeding areas leading to poor spawning condition. Poor recruitment leads to few 0-age in current season catch and 1-age fish in next season's catch.	$S_t$ and $N_t$	Nearshore Mar-May SST in year $t$ and $t - 1$



Table 1. Continued.

Hypothesis	Resp.	Covariates
L1. The prior year post-monsoon larval survival and growth is associated with higher future biomass. Larval growth and survival is highest in an intermediate temperature window. Low SST at this time is also indicative of strong upwelling which advects larvae into offshore waters where productivity is lower.	$N_t$ and $S_t$	Nearshore SST during Oct-Dec in year t-1
L2. Upwelling is associated with higher productivity and higher density of zooplankton, which leads to better larval and juvenile growth and survival. The strength of summer upwelling should be associated with higher biomass in future years and the appearance of 0-age fish in post-monsoon catch. However, extremely strong upwelling brings poorly oxygenated water to the surface causing larval mortality and offshore advection and causing mature fish to move offshore.	$N_t$ and $S_t$	Jun-Sep upwelling index in year $t - 1$ and $t$
L3. Chlorophyll blooms are signatures of high productivity from nutrient influx either due to upwelling or coastal inputs. The monsoon bloom intensity should be associated with 0-year fish abundance in year $t$ and future sardine biomass.	$N_t$ and $S_t$	Chl-a density Jun-Sep in year $t - 1$ and $t$ (for $N_t$ )
A1. The multi-year average sea surface temperature (SST) has been found to correlate with recruitment strength in Pacific sardine. Presumably the long-term average SST is associated with a variety of factors which affect spawning and early survival (Checkley et al. 2017).	$N_t$ and $S_t$	3-year average SST
A2. The changes brought about by the El Niño Southern Oscillation (ENSO) cycle have a variety of effects on environmental parameters (precipitation, SST, thermal fronts, Wind) which impacts spawning and early survival. This in turn impacts the overall abundance (Rohit et al. 2018).	$N_t$ and $S_t$	ONI in year t-1

Table 2. Model selection tests of time-dependency and linearity for the monsoon (Jul-Sep) catch model using F-tests and AIC of nested models fit to log of landings data.  $S_t$  is the catch during the monsoon (Jul-Sep) of season  $t$ .  $N_{t-1}$  is the post-monsoon (Oct-Mar) catch in the prior sardine season.  $N_{t-2}$  is the same for two seasons prior.  $s()$  is a non-linear function of the response variable. MASE is mean absolute squared error of the residuals. The tests are nested and the numbers before the models indicates the nest level. In the list with numbers 1, 2, 3a, 3b, two nested tests were done, one on 1, 2, 3a and one on 1, 2, 3b. See tables in Appendix A for the full set of time-dependency and linearity tests. This table only shows the nested model sets that were most supported.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
Naive Model 1984-2015 data						
$\ln(S_t) = \ln(S_{t-1}) + \varepsilon_t$	32		1			122.85
AR-1 Model						
$\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	30	0.814				114.14
Time dependency test						
1. $\ln(S_t) = \alpha + \ln(N_{t-1}) + \varepsilon_t$	31	0.856	14.2			111.78
2. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	30	0.794	22.2	4.06	0.053	109.59
3a. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(N_{t-2}) + \varepsilon_t$	29	0.797	19.6	0.01	0.919	111.57
3b. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-2}) + \varepsilon_t$	29	0.778	20.8	0.45	0.508	111.09
Linearity and time-dependency test						
1. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	30	0.794	22.2			109.59
2. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	28.6	0.761	24.4	1.26	0.287	109.52
3a. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(N_{t-2})) + \varepsilon_t$	26.4	0.761	21.2	0.28	0.785	112.42
3b. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \varepsilon_t$	25.9	0.724	26.3	1.09	0.367	110.63

Table 3. Model selection tests for the post-monsoon (Oct-Mar) catch model ( $N_t$ ) using F-tests and AIC.  $S_t$  is the catch during the monsoon (Jul-Sep).  $N_t$  is the catch during the post-monsoon period (Oct-Mar) of season  $t$ ; note the fishing season is defined as Jul-Jun not calendar year.  $S_{t-1}$  and  $N_{t-1}$  are the catch during the prior sardine season during and after the monsoon respectively.  $S_{t-2}$  and  $N_{t-2}$  are the same for two seasons prior.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
Naive Model 1984-2014 data						
$\ln(N_t) = \ln(N_{t-1}) + \varepsilon_t$	31	0.606				90.87
Linearity and time-dependency test						
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	29	0.592	37			88.28
2. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	27.6	0.529	45	4.87	0.026	84.75
3a. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \varepsilon_t$	25.4	0.488	46	1.09	0.357	86.11
3b. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	24.8	0.45	57	11.05	0.007	79.53

Table 4. Top covariates for the monsoon (Jul-Sep) and post-monsoon (Oct-Mar) catch ( $S_t$  and  $N_t$ ) models. The models are nested; the number indicates the level of nestedness. Models at levels 2 and higher are shown with the component that is added to the base level model (M0 or M1) at top. The full set of covariate models tested are given in Appendix B. The fitted versus observed catches from the covariate models are shown in Figure 8.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
Jul-Sep catch models with covariates						
$V_t$ = Jun-Sep SST current season						
$W_t$ = Jun-Sep UPW current season						
1-M0. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon$	28.6	0.761	24			109.52
2a. $\ln(S_t) = M0 + s(V_t)$	25.9	0.683	41	3.84	0.025	103.43
2b. $\ln(S_t) = M0 + \beta W_t$	27.6	0.706	33	4.96	0.034	106.32
Oct-Mar catch models with covariates						
$V_t$ = Mar-May SST current season						
$W_t$ = Jun-Sep upwelling current season						
1-M1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \varepsilon$	24.8	0.45	57			79.53
2a. $\ln(N_t) = M1 + s(V_t)$	22	0.413	63	2.53	0.089	76.01
3a. $\ln(N_t) = M1 + s(V_t) + \beta W_t$	21.1	0.434	64	1.45	0.239	76.08
2b. $\ln(N_t) = M1 + \beta W_t$	23.8	0.46	62	4.91	0.037	76

## **Appendices**

## 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
Naive Model 1984-2015 data						
$\ln(S_t) = \ln(S_{t-1}) + \varepsilon$	32	0.986				122.85
Time dependency test A 1984-2015 data						
1. $\ln(S_t) = \alpha + \ln(S_{t-1}) + \varepsilon_t$	31	0.978	-29			124.83
2. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	30	0.803	10.3	15.14	0.001	114.14
3. $\ln(S_t) = \alpha + \beta_1 \ln(S_{t-1}) + \beta_2 \ln(S_{t-2}) + \varepsilon_t$	29	0.791	13.6	2.13	0.155	113.88
Time dependency test B 1984-2015 data						
1. $\ln(S_t) = \alpha + \ln(N_{t-1}) + \varepsilon_t$	31	0.844	14.2			111.78
2. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	30	0.783	22.2	4.06	0.053	109.59
3. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(N_{t-2}) + \varepsilon_t$	29	0.786	19.6	0.01	0.919	111.57
Time dependency test C 1984-2015 data						
1. $\ln(S_t) = \alpha + \ln(N_{t-1}) + \varepsilon_t$	31	0.844	14.2			111.78
2. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	30	0.783	22.2	4.08	0.053	109.59
3a. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-1}) + \varepsilon_t$	29	0.792	20	0.16	0.688	111.4
3b. $\ln(S_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-2}) + \varepsilon_t$	29	0.767	20.8	0.45	0.508	111.09

Table A2. Model selection tests of time-dependency the catch during spawning months using non-linear responses instead of 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
Time dependency test A 1984-2015 data						
1. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	30	0.803	10.3			114.14
2. $\ln(S_t) = \alpha + s(\ln(S_{t-1})) + \varepsilon_t$	28.2	0.787	19.6	2.74	0.089	111.79
3. $\ln(S_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	25.5	0.759	20.7	0.97	0.416	113.23
Time dependency test B 1984-2015 data						
1. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	30	0.783	22.2			109.59
2. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	28.6	0.75	24.4	1.26	0.287	109.52
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \varepsilon_t$	26.4	0.751	21.2	0.28	0.785	112.42
Time dependency test C 1984-2015 data						
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	28.6	0.75	24.4			109.52
2. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \varepsilon_t$	26.1	0.688	28.5	1.49	0.242	109.55
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	25.9	0.714	26.3	1.09	0.367	110.63

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
Time dependency test A 1956-1983 data						
1. $\ln(S_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	24	0.624	-0.7			64.69
2. $\ln(S_t) = \alpha + s(\ln(S_{t-1})) + \varepsilon_t$	22.1	0.605	-0.2	0.78	0.464	65.71
3. $\ln(S_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	19.9	0.572	3.1	1.19	0.329	66.35
Time dependency test B 1956-1983 data						
1. $\ln(S_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	24	0.625	-3.8			65.48
2. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	21.6	0.576	8.2	2.24	0.127	63.8
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \varepsilon_t$	18.5	0.488	16.9	1.56	0.231	63.13
Time dependency test C 1956-1983 data						
1. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	22.5	0.577	4.3			66.2
2. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \varepsilon_t$	20.7	0.548	4.8	0.91	0.41	67.3
3. $\ln(S_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	19.5	0.542	12.9	1.42	0.266	63.79

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
Naive Model 1984-2014 data						
$\ln(N_t) = \ln(N_{t-1}) + \varepsilon$	31	0.606				90.87
Time dependency test A 1984-2014 data						
1. $\ln(N_t) = \alpha + \ln(S_{t-1}) + \varepsilon$	30	0.826	-20.3			107.36
2. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	29	0.617	26.2	19.99	0	93.17
3. $\ln(N_t) = \alpha + \beta_1 \ln(S_{t-1}) + \beta_2 \ln(S_{t-2}) + \varepsilon_t$	28	0.611	26.6	1.15	0.292	93.92
Time dependency test B 1984-2014 data						
1. $\ln(N_t) = \alpha + \ln(N_{t-1}) + \varepsilon_t$	30	0.605	24.7			92.87
2. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	29	0.592	37	6.63	0.016	88.28
3. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(N_{t-2}) + \varepsilon_t$	28	0.588	34.8	0.04	0.843	90.24
Time dependency test C 1984-2014 data						
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	29	0.592	37			88.28
2a. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-1}) + \varepsilon_t$	28	0.584	35	0.12	0.729	90.15
2b. $\ln(N_t) = \alpha + \beta_1 \ln(N_{t-1}) + \beta_2 \ln(S_{t-2}) + \varepsilon_t$	28	0.592	34.7	0.01	0.919	90.27

Table A5. Model selection tests of time-dependency the  $N_t$  model using non-linear responses instead of 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
Time dependency test A 1984-2014 data						
1. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	29	0.617	26.2			93.17
2. $\ln(N_t) = \alpha + s(\ln(S_{t-1})) + \varepsilon_t$	27.3	0.601	30.2	1.83	0.185	92.61
3. $\ln(N_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	24.4	0.57	36.4	1.79	0.177	91.62
Time dependency test B 1984-2014 data						
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	29	0.592	37			88.28
2. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	27.6	0.529	45.3	3.88	0.047	84.75
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \varepsilon_t$	25.4	0.488	45.6	0.87	0.441	86.11
Time dependency test C 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	27.6	0.529	45.3			84.75
2. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \varepsilon_t$	25.1	0.519	43.8	0.53	0.634	87.37
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	24.8	0.45	56.6	3.39	0.036	79.53

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
Time dependency test A 1956-1983 data						
1. $\ln(N_t) = \alpha + \beta \ln(S_{t-1}) + \varepsilon_t$	24	0.388	-1.7			44.98
2. $\ln(N_t) = \alpha + s(\ln(S_{t-1})) + \varepsilon_t$	22.1	0.323	16.2	3.53	0.052	41.11
3. $\ln(N_t) = \alpha + s_1(\ln(S_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	19.9	0.304	18.1	1.09	0.362	42
Time dependency test B 1956-1983 data						
1. $\ln(N_t) = \alpha + \beta \ln(N_{t-1}) + \varepsilon_t$	24	0.412	-4.2			45.61
2. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	21.6	0.307	29.1	5.69	0.009	37.12
3. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(N_{t-2})) + \varepsilon_t$	18.5	0.285	32.2	1.14	0.36	37.87
Time dependency test C 1956-1983 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	21.6	0.307	29.1			37.12
2a. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-1})) + \varepsilon_t$	19	0.273	34.4	1.49	0.251	36.74
2b. $\ln(N_t) = \alpha + s_1(\ln(N_{t-1})) + s_2(\ln(S_{t-2})) + \varepsilon_t$	19.5	0.282	33.4	1.54	0.24	36.84



## 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 spawning 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 B4. Model selection tests of sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1), upwelling and ONI 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. Two upwelling indices were tested. The nearshore-offshore temperature differential (UPW), which is the offshore (box 13) minus nearshore (box 4) SST, and the average nearshore SST along the Kerala coast (boxes 2-5). These are highly correlated but not identical. The ONI index is the average over all months in the calendar year. The 2.5-year average SST is the average for Jan-Jun in the current calendar year and the prior 2 calendar years. Thus the average does not include any months during the Jul-Sep catch (response variable). 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. $\ln(S_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	28.6	0.761	24.4			109.52
$V_t = \text{Ave Mar-May SST (S4)}$						
2a. $\ln(S_t) = M + \beta V_t$	27.6	0.772	23.3	0.59	0.449	110.87
3a. $\ln(S_t) = M + s(V_t)$	25.5	0.753	26.4	1.26	0.303	110.8
2b. $\ln(S_t) = M + \beta V_{t-1}$	27.6	0.754	22.8	0.39	0.533	111.07
3b. $\ln(S_t) = M + s(V_{t-1})$	25.7	0.723	26.1	1.35	0.275	110.74
$V_t = \text{Ave Oct-Dec SST (L1)}$						
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
$V_t = \text{Ave. Jun-Sep UPW (S4 and L2)}$						
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
$V_t = \text{Ave. Jun-Sep SST (S4 and L2)}$						
2a. $\ln(S_t) = M + \beta V_t$	27.6	0.745	33.3	5.53	0.027	106.4
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

Model	Residual df	MASE	Adj. R2	F	p value	AIC
$V_t = 2.5\text{-year average SST (A1)}$						
2. $\ln(S_t) = M + \beta V_t$	27.6	0.723	33.2	5.52	0.027	106.43
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

Table B5. 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. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \varepsilon_t$		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	NA	NA	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 B6. Model selection tests of sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1), upwelling and ONI as the explanatory variables ( $V_t$ ) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data. The hypothesis tested (Table 1) is noted in parentheses. Two upwelling indices were tested. The nearshore-offshore temperature differential (UPW), which is the offshore (box 13) minus nearshore (box 4) SST, and the average nearshore SST along the Kerala coast (boxes 2-5). These are highly correlated but not identical. The ONI index is the average over all months in the calendar year. The 2.5-year average SST is the average for Jan-Jun in the current calendar year and the prior 2 calendar years. Thus the average does not include any months during the Oct-Mar catch. See Table B5 for an explanation of the models.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + s(\ln(S_{t-2})) + \varepsilon_t$	24.8	0.743	56.6			79.53
$V_t = \text{Ave Mar-May SST (S5)}$						
2a. $\ln(N_t) = M + \beta V_t$	23.8	0.701	59	2.84	0.107	78.53
3a. $\ln(N_t) = M + s(V_t)$	22	0.682	63.2	2.29	0.13	76.01
2b. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.762	57.1	1.33	0.26	79.93
3b. $\ln(N_t) = M + s(V_{t-1})$	22	0.747	57.4	0.79	0.455	80.61
$V_t = \text{Ave Oct-Dec SST (L1)}$						
2. $\ln(N_t) = M + \beta V_{t-1}$	23.8	0.748	54.9	NA	NA	81.5
3. $\ln(N_t) = M + s(V_{t-1})$	22.5	0.736	56	1.13	0.318	81.37
$V_t = \text{Ave. Jun-Sep UPW (L2)}$						
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
$V_t = \text{Ave. Jun-Sep SST (L2)}$						
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

Model	Residual df	MASE	Adj. R2	F	p value	AIC
$V_t = 2.5\text{-year average SST (A1)}$						
2. $\ln(N_t) = \mathbf{M} + \beta V_t$	23.8	0.667	64.7	7.68	0.012	73.9
3. $\ln(N_t) = \mathbf{M} + s(V_t)$	22.7	0.594	67.5	2.58	0.12	71.88
$V_t = \text{ONI (A2)}$						
2. $\ln(N_t) = \mathbf{M} + \beta V_{t-1}$	23.8	0.744	54.9	NA	NA	81.46
3. $\ln(N_t) = \mathbf{M} + s(V_{t-1})$	23	0.748	55.5	0.99	0.313	81.46



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. 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). In the Appendix C tables, the 2nd best model is used. This model is simpler than the best model (based on AIC) and uses only  $N_{t-1}$ . 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. MASE is the mean absolute square error (residuals).

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	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	NA	NA	86.74
3a. $\ln(N_t) = M + s(V_t)$	24.9	0.849	45.3	1.24	0.303	86.66
4a. $\ln(N_t) = M + s(V_t) + \beta V_{t-1}$	23.9	0.844	43.1	NA	NA	88.59
5a. $\ln(N_t) = M + s(V_t) + s(V_{t-1})$	22.4	0.844	42.9	0.61	0.507	89.49
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	0.744	86.59
3a. $\ln(N_t) = M + s(V_t)$	24.5	0.863	41.3	0.24	0.805	89.08
4a. $\ln(N_t) = M + s(V_t) + \beta V_{t-1}$	23.5	0.856	39.5	0.21	0.643	90.78
5a. $\ln(N_t) = M + s(V_t) + s(V_{t-1})$	21.7	0.831	40.3	0.89	0.415	91.18
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 average sea surface temperature off the Kerala coast (up to 80km offshore in boxes 2-5 in Figure 1) as the explanatory variable ( $V_t$ ) for the catch during post-monsoon months (Oct-May) using 1984 to 2014 data. The hypothesis tested (Table 1) is noted in parentheses. See Table B7 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
$V_t = \text{Ave Mar-May SST (S5)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.854	46.2	1.73	0.202	85.12
3a. $\ln(N_t) = M + s(V_t)$	24.8	0.839	46.7	0.96	0.393	85.9
4a. $\ln(N_t) = M + s(V_t) + \beta V_{t-1}$	23.8	0.835	48.7	2.41	0.136	85.41
5a. $\ln(N_t) = M + s(V_t) + s(V_{t-1})$	22.3	0.781	54.6	2.77	0.095	82.39
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.813	48	2.54	0.124	84.04
3b. $\ln(N_t) = M + s(V_{t-1})$	24.7	0.778	49	0.96	0.395	84.54
$V_t = \text{Ave Oct-Dec SST (L1)}$						
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

Table B9. Model selection tests of upwelling intensity off Cochi as the explanatory variable. See Table B7 for an explanation of the models. Two upwelling indices were tested. The nearshore-offshore temperature differential (UPW), which is the offshore (box 13) minus nearshore (box 4) SST, and the average nearshore SST along the Kerala coast (boxes 2-5). These are highly correlated but not identical.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1984-2014 data						
1. $\ln(N_t) = \alpha + s(\ln(N_{t-1})) + \varepsilon_t$	27.6	0.874	45.3			84.75
$V_t = \text{Ave. Jun-Sep UPW (L2)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.863	54	6.42	0.02	80.25
3a. $\ln(N_t) = M + s(V_t)$	24.1	0.834	56	1.18	0.335	80.36
4a. $\ln(N_t) = M + s(V_t) + \beta V_{t-1}$	23.1	0.829	55.7	0.76	0.389	81.35
5a. $\ln(N_t) = M + s(V_t) + s(V_{t-1})$	20.8	0.797	54.6	0.5	0.64	83.29
2b. $\ln(N_t) = M + \beta V_{t-1}$	26.6	0.876	43.3	NA	NA	86.74
3b. $\ln(N_t) = M + s(V_{t-1})$	24	0.804	46.9	1.46	0.253	86.25
$V_t = \text{Ave. Jun-Sep SST (L2)}$						
2a. $\ln(N_t) = M + \beta V_t$	26.6	0.857	51.6	4.33	0.05	81.79
3a. $\ln(N_t) = M + s(V_t)$	24.6	0.86	51.3	0.58	0.573	83.19
4a. $\ln(N_t) = M + s(V_t) + \beta V_{t-1}$	23.6	0.839	49.6	0.11	0.73	85
5a. $\ln(N_t) = M + s(V_t) + s(V_{t-1})$	21.7	0.838	47	0.1	0.904	87.55
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
#						

## Appendix C: Tests for Chlorophyll-a as a covariate

Table C2. Model selection tests of Chlorophyll-a (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.  $V_t$  is CHL in the current season which spans two calendar years from July to June in the next year.  $V_{t-1}$  is CHL in the prior Jul-Jun season. Only CHL in Oct-Dec and Jan-Mar in the prior season is used since for the current season, 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. $\ln(S_t) = \alpha + p(\ln(N_{t-1})) + \varepsilon$	14	0.516	25.3			18.29
$V_t$ = Jul-Sep Chlorophyll						
2. $\ln(S_t) = M + \beta V_t$	13	0.503	24.6	0.69	0.427	19.2
3. $\ln(S_t) = M + p(V_t)$	12	0.48	19.5	0.16	0.699	20.94
4. $\ln(S_t) = M + p(V_t) + \beta V_{t-1}$	11	0.5	13.7	0.17	0.688	22.65
5. $\ln(S_t) = M + p(V_t) + p(V_{t-1})$	10	0.497	5.1	0.01	0.935	24.64
$V_t$ = Oct-Dec Chlorophyll						
2. $\ln(S_t) = M + \beta V_{t-1}$	13	0.516	19.6	0	0.99	20.29
3. $\ln(S_t) = M + p(V_{t-1})$	12	0.456	21.5	1.33	0.272	20.51
$V_t$ = Jan-Mar Chlorophyll						
2. $\ln(S_t) = M + \beta V_{t-1}$	13	0.522	20.6	0.16	0.697	20.08
3. $\ln(S_t) = M + p(V_{t-1})$	12	0.526	16.7	0.4	0.541	21.52

Table C1. Model selection tests of Chlorophyll-a (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.  $V_t$  is CHL in the current season which spans two calendar years from July to June in the next year.  $V_{t-1}$  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. $\ln(N_t) = \alpha + p(\ln(N_{t-1})) + \epsilon$	14	0.875	26.5			18.94
$V_t$ = Jul-Sep Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.893	23.1	0.32	0.587	20.45
3. $\ln(N_t) = M + p(V_t)$	12	0.874	17.9	0.15	0.709	22.21
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.86	25	0.69	0.422	20.03
3. $\ln(N_t) = M + p(V_{t-1})$	11.7	0.839	21.7	0.27	0.677	21.36
$V_t$ = Oct-Dec Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.883	23.9	0.59	0.458	20.29
3. $\ln(N_t) = M + p(V_t)$	12	0.744	29.5	2.22	0.167	19.62
4. $\ln(N_t) = M + p(V_t) + \beta V_{t-1}$	11	0.679	40.8	2.99	0.114	17.16
5. $\ln(N_t) = M + p(V_t) + p(V_{t-1})$	10	0.68	34.9	0	0.976	19.16
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.764	39.4	3.87	0.074	16.41
3. $\ln(N_t) = M + p(V_{t-1})$	11.3	0.728	37.7	0.49	0.595	17.62
$V_t$ = Jan-Mar Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.901	23.6	0.4	0.541	20.34
3. $\ln(N_t) = M + p(V_t)$	12	0.829	23.9	0.89	0.367	20.92
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.866	21.2	0.05	0.829	20.88
3. $\ln(N_t) = M + p(V_{t-1})$	11.1	0.873	15.2	0.23	0.791	22.97

Table C3. Model selection tests of Chlorophyll-a as an explanatory variable for the catch during the non-spawning months (Oct-Jun) using box 5.

Model	Residual df	MASE	Adj. R2	F	p value	AIC
base model (M) 1998-2014 data						
1. $\ln(N_t) = \alpha + p(\ln(N_{t-1})) + \epsilon$	14	0.875	26.5			18.94
$V_t$ = Jul-Sep Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.865	22.5	0.24	0.635	20.6
3. $\ln(N_t) = M + p(V_t)$	12	0.904	20.4	0.61	0.451	21.69
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.839	28.5	1.33	0.271	19.22
3. $\ln(N_t) = M + p(V_{t-1})$	12	0.837	25.2	0.07	0.789	20.42
$V_t$ = Oct-Dec Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.864	28.4	1.4	0.265	19.25
3. $\ln(N_t) = M + p(V_t)$	12	0.844	24	0.26	0.62	20.91
4. $\ln(N_t) = M + p(V_t) + \beta V_{t-1}$	11	0.666	35.6	2.9	0.119	18.62
5. $\ln(N_t) = M + p(V_t) + p(V_{t-1})$	10	0.649	29.9	0.11	0.743	20.42
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.739	35.5	2.88	0.116	17.48
3. $\ln(N_t) = M + p(V_{t-1})$	11.7	0.732	34.2	0.52	0.534	18.39
$V_t$ = Jan-Mar Chlorophyll						
2. $\ln(N_t) = M + \beta V_t$	13	0.847	29.5	1.56	0.24	18.98
3. $\ln(N_t) = M + p(V_t)$	12	0.804	31.6	1.33	0.276	19.11
2. $\ln(N_t) = M + \beta V_{t-1}$	13	0.89	21.4	0.09	0.769	20.84
3. $\ln(N_t) = M + p(V_{t-1})$	8.9	0.682	27.9	1.07	0.427	20.97

## **Appendix D: Covariates along the SE India coast**

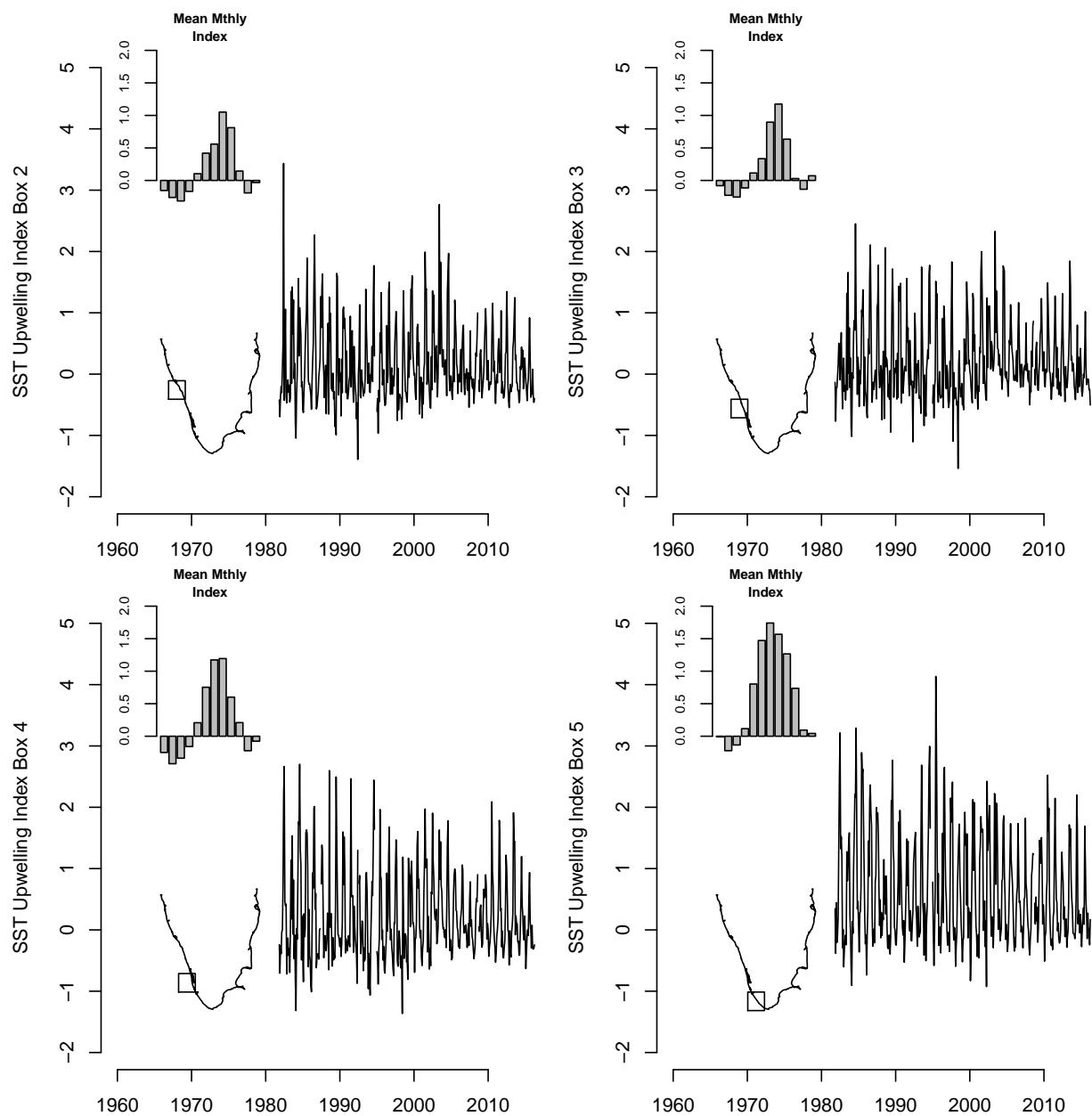


Figure D1. Upwelling index.



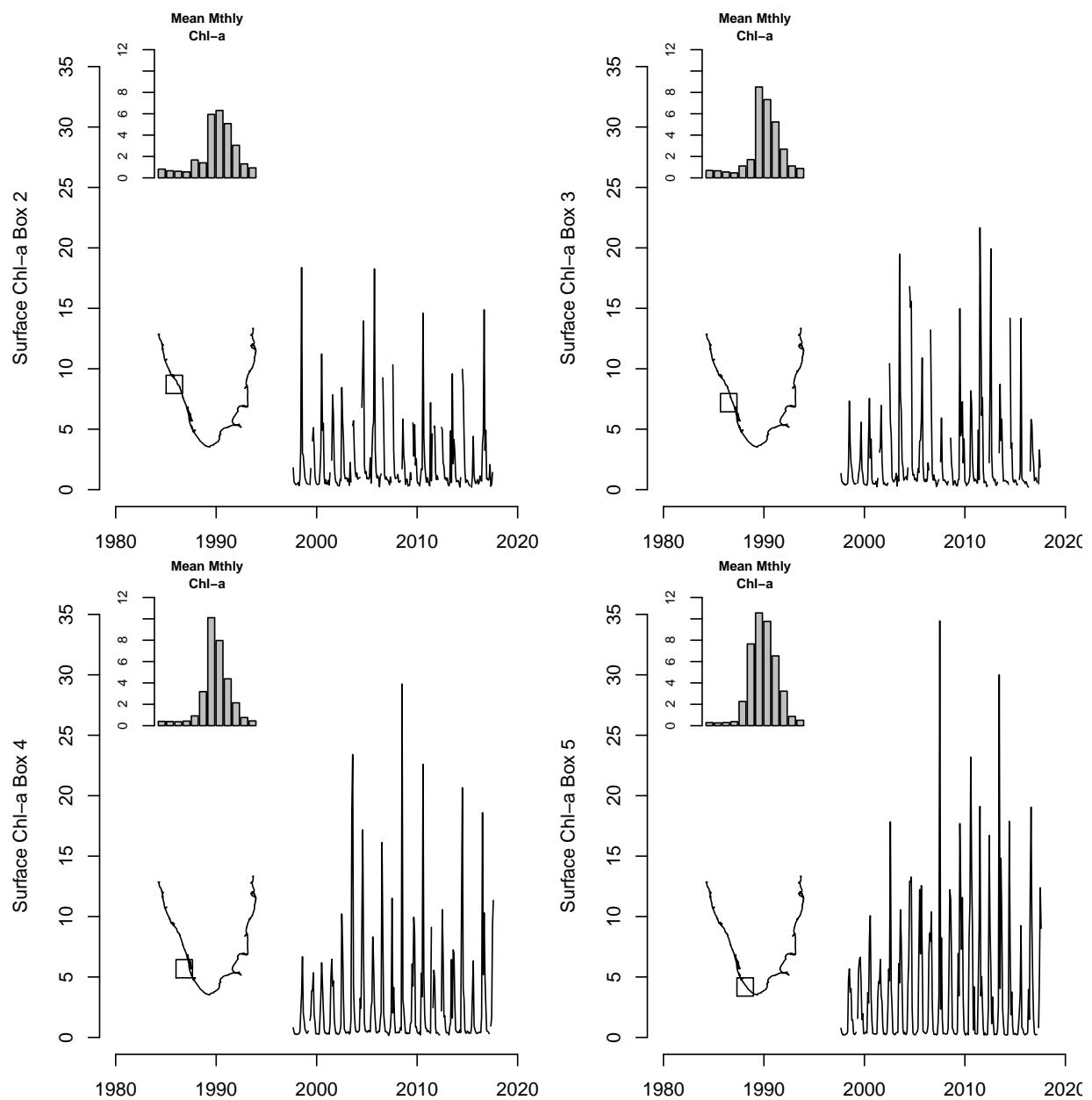


Figure D2. Chlorophyll-a.

## **Appendix E: Comparison of land and oceanic rainfall measurements**

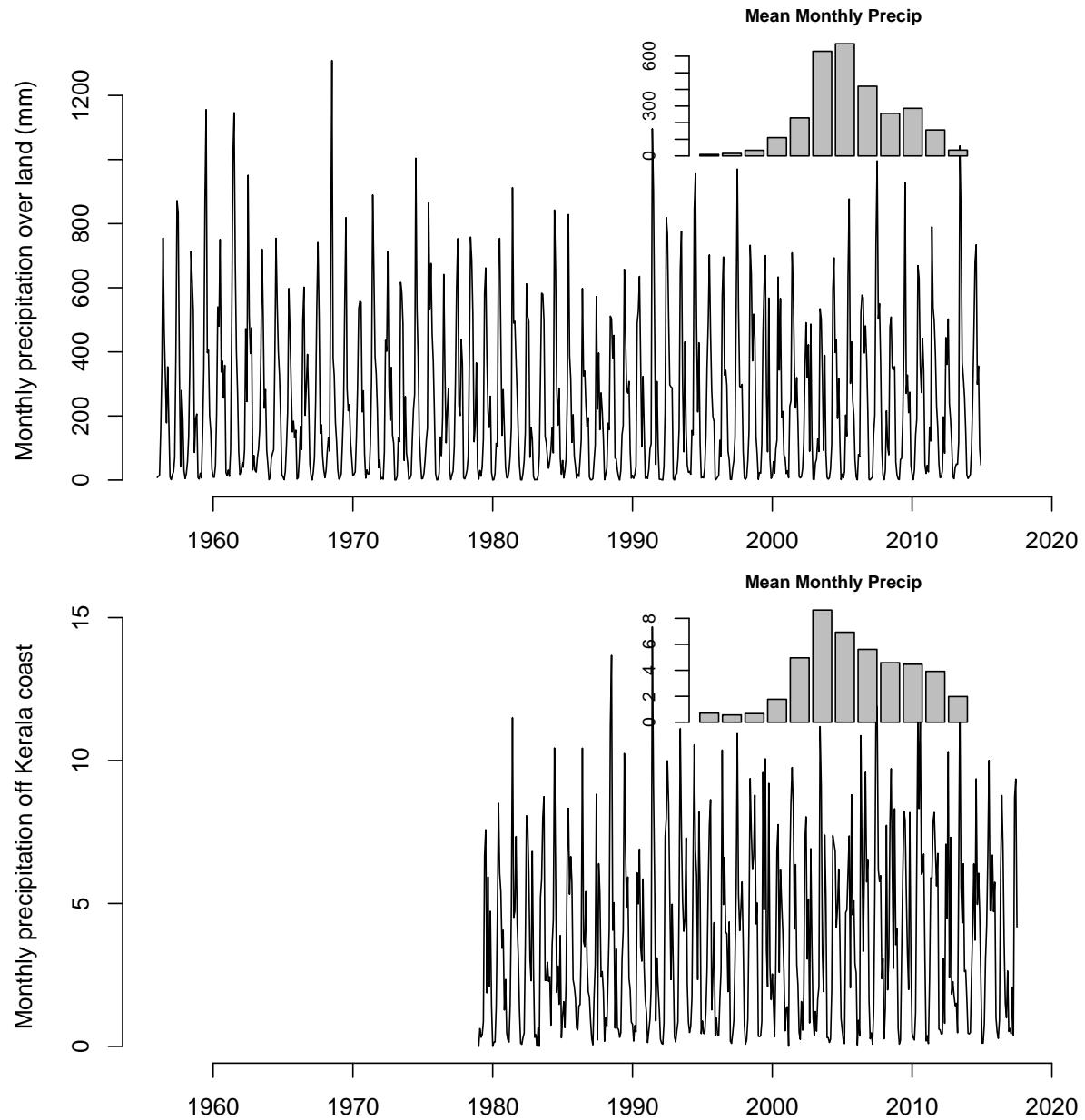


Figure E1



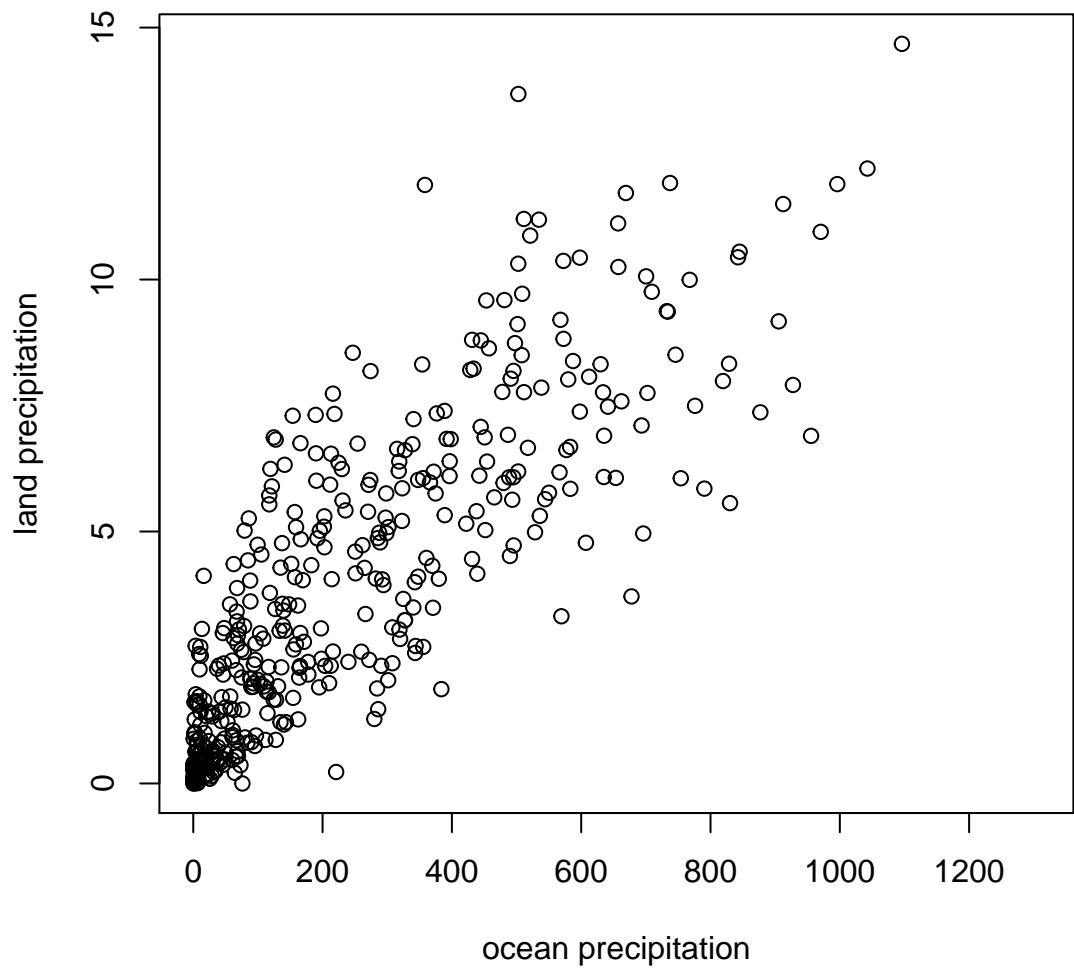


Figure E2. Monthly precipitation measured over land via land gauges versus the precipitation measured via remote sensing over the ocean.

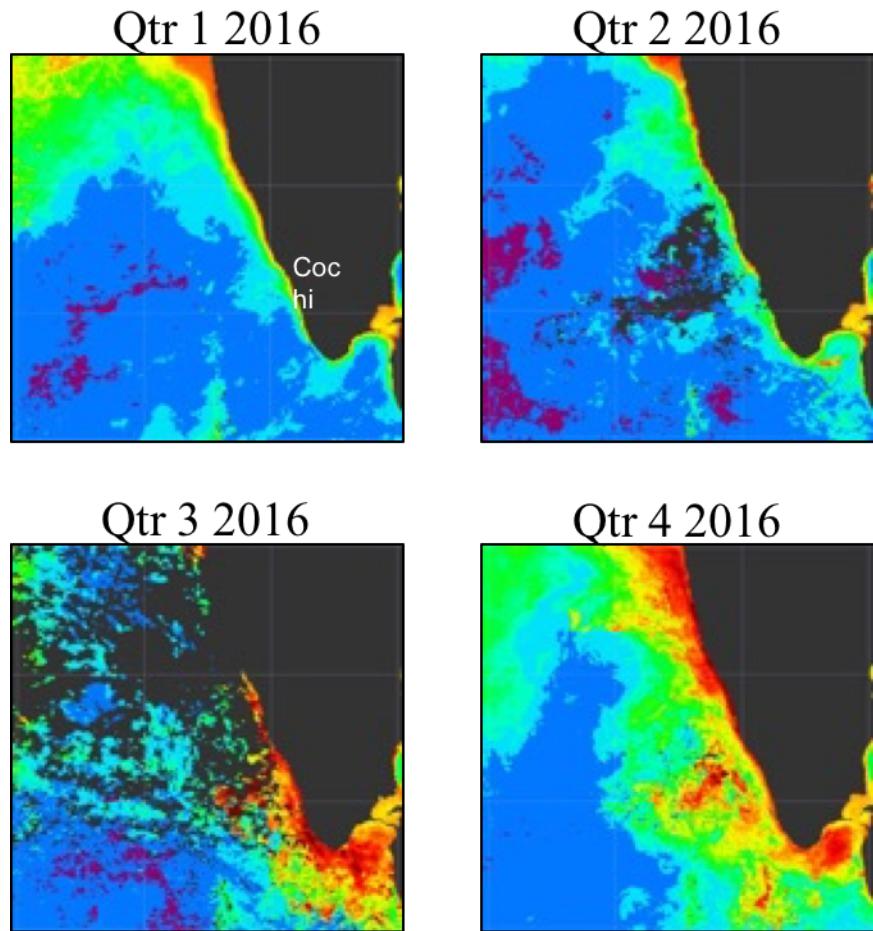


Figure 9

## Appendix F: Chlorophyll-a images in 2016

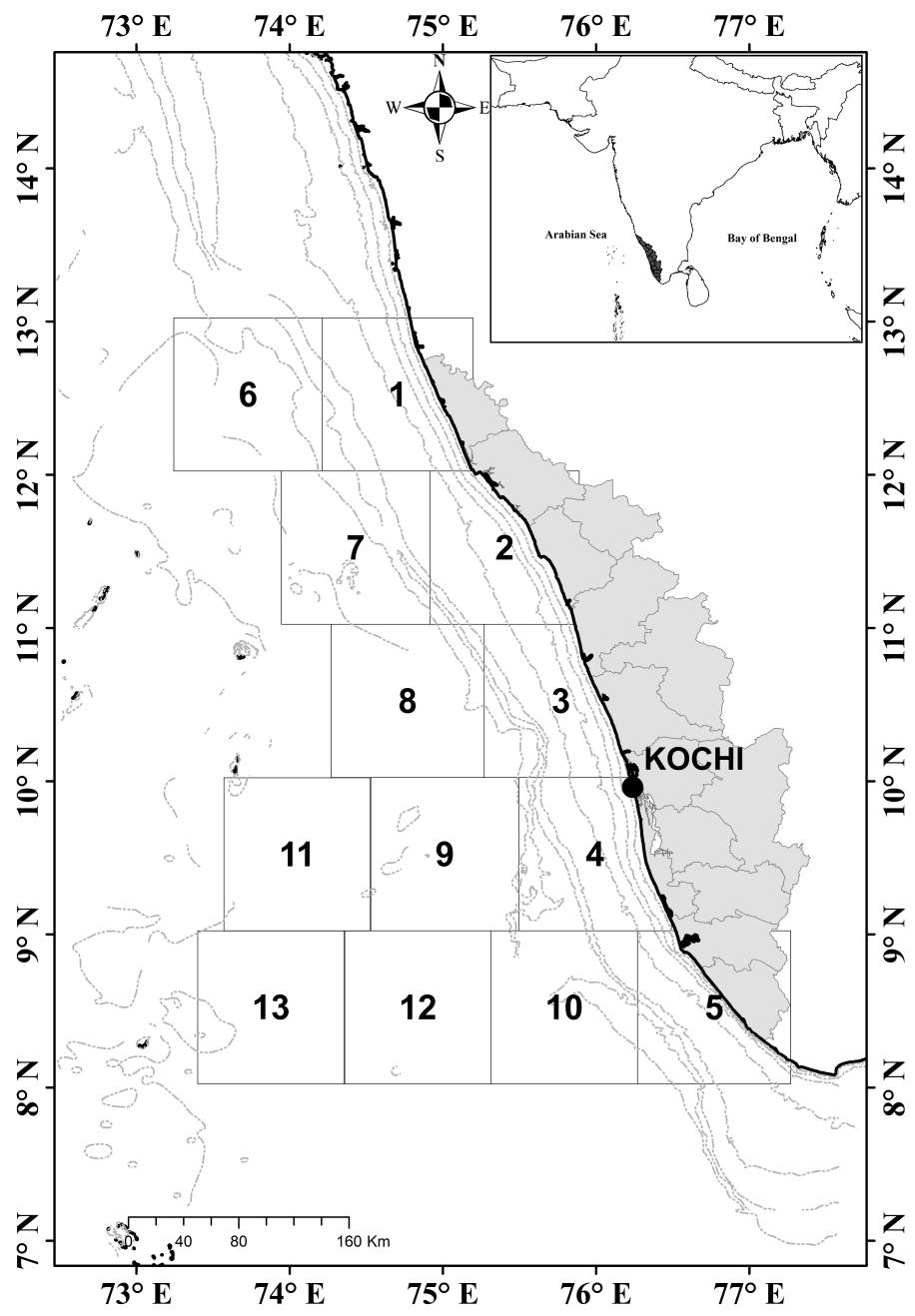


Figure 1

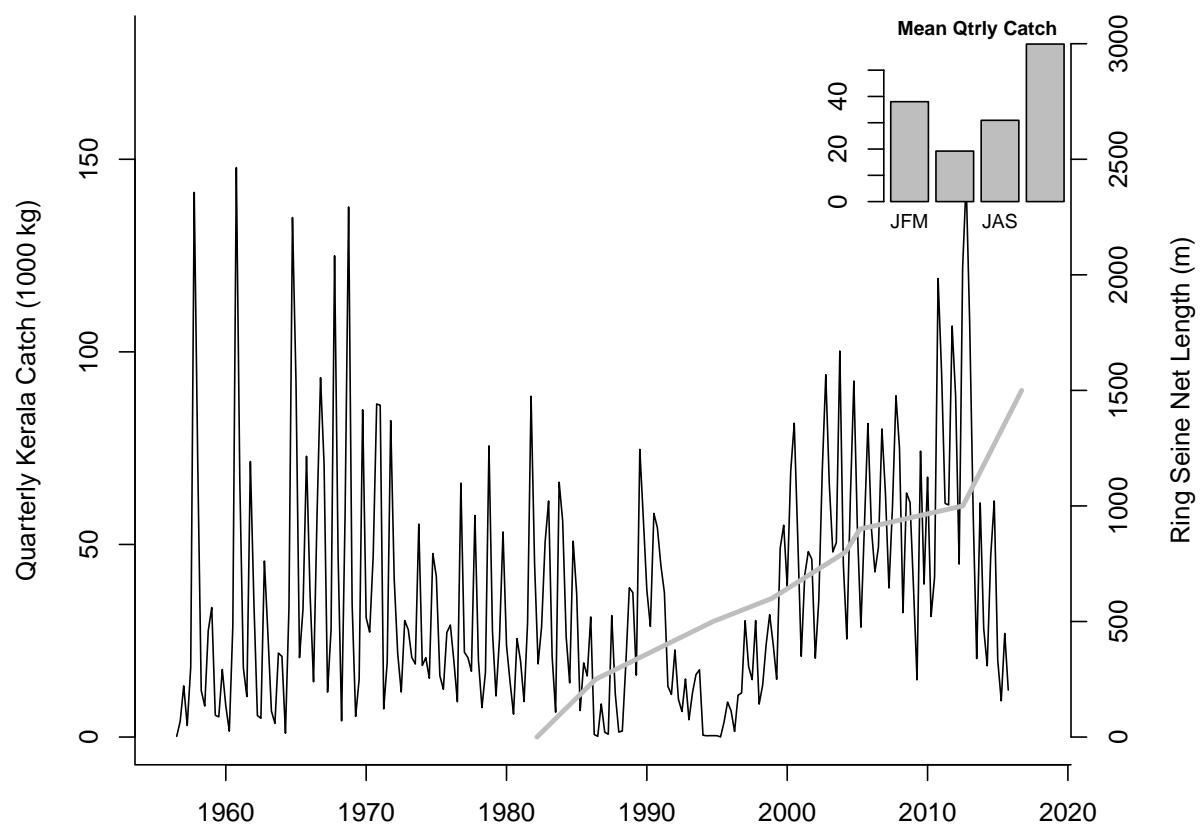


Figure 3

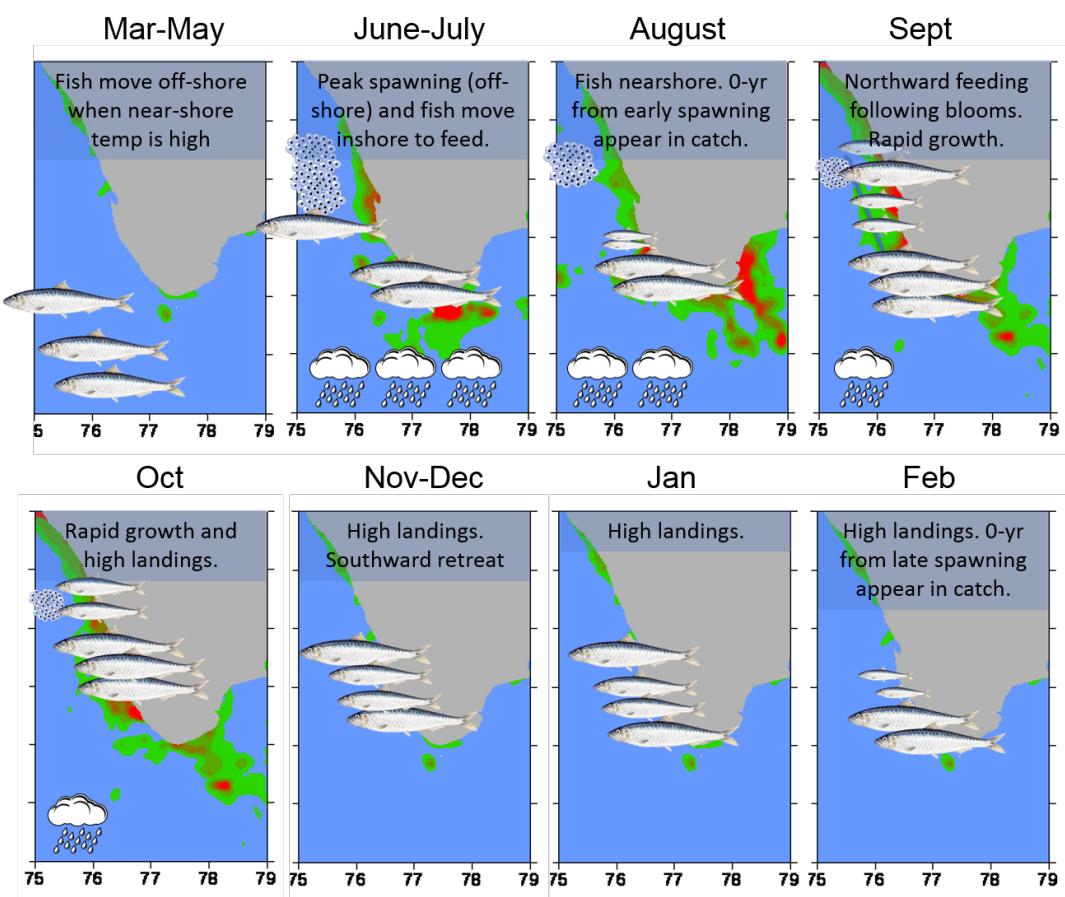


Figure 2

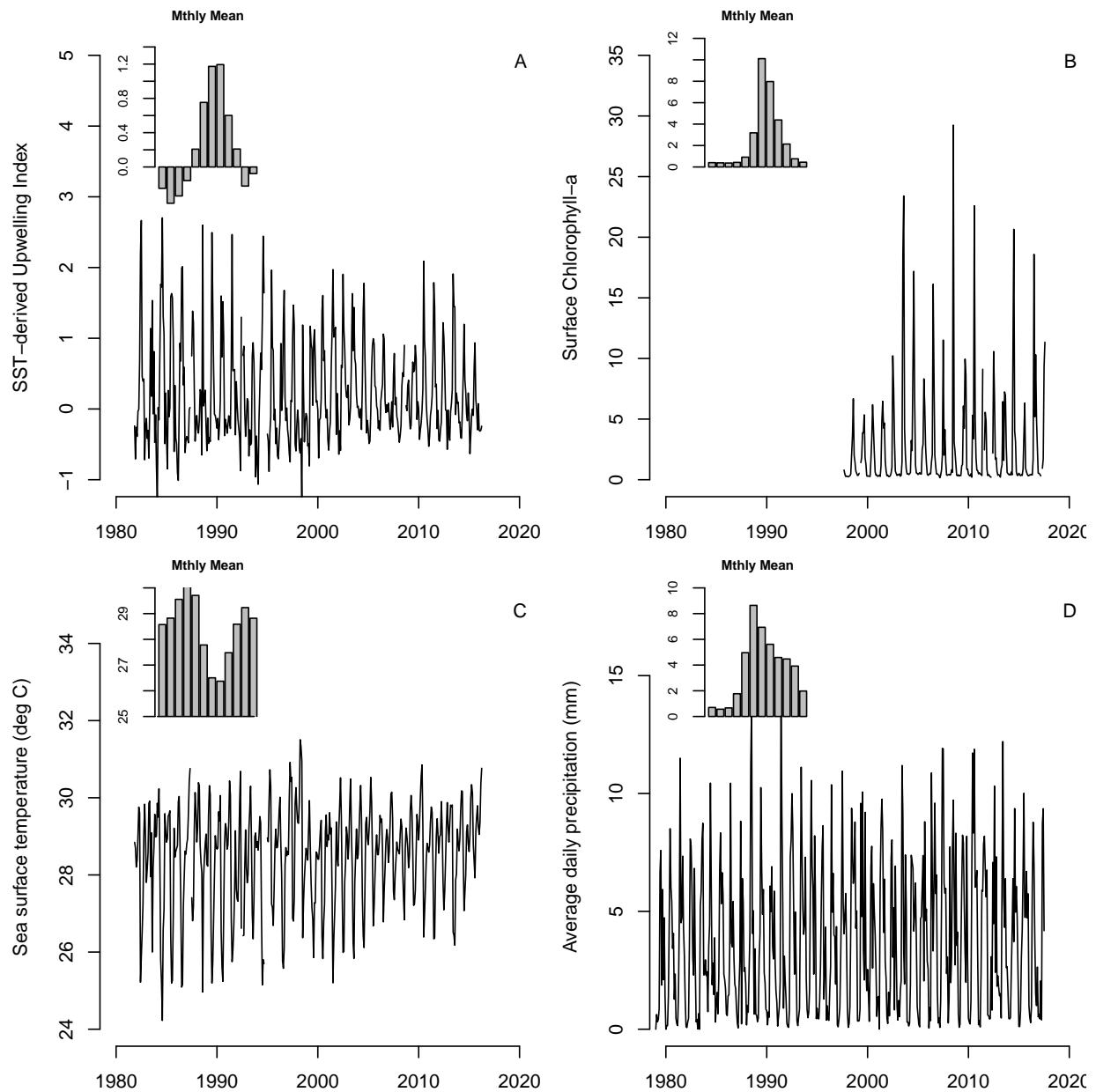


Figure 4

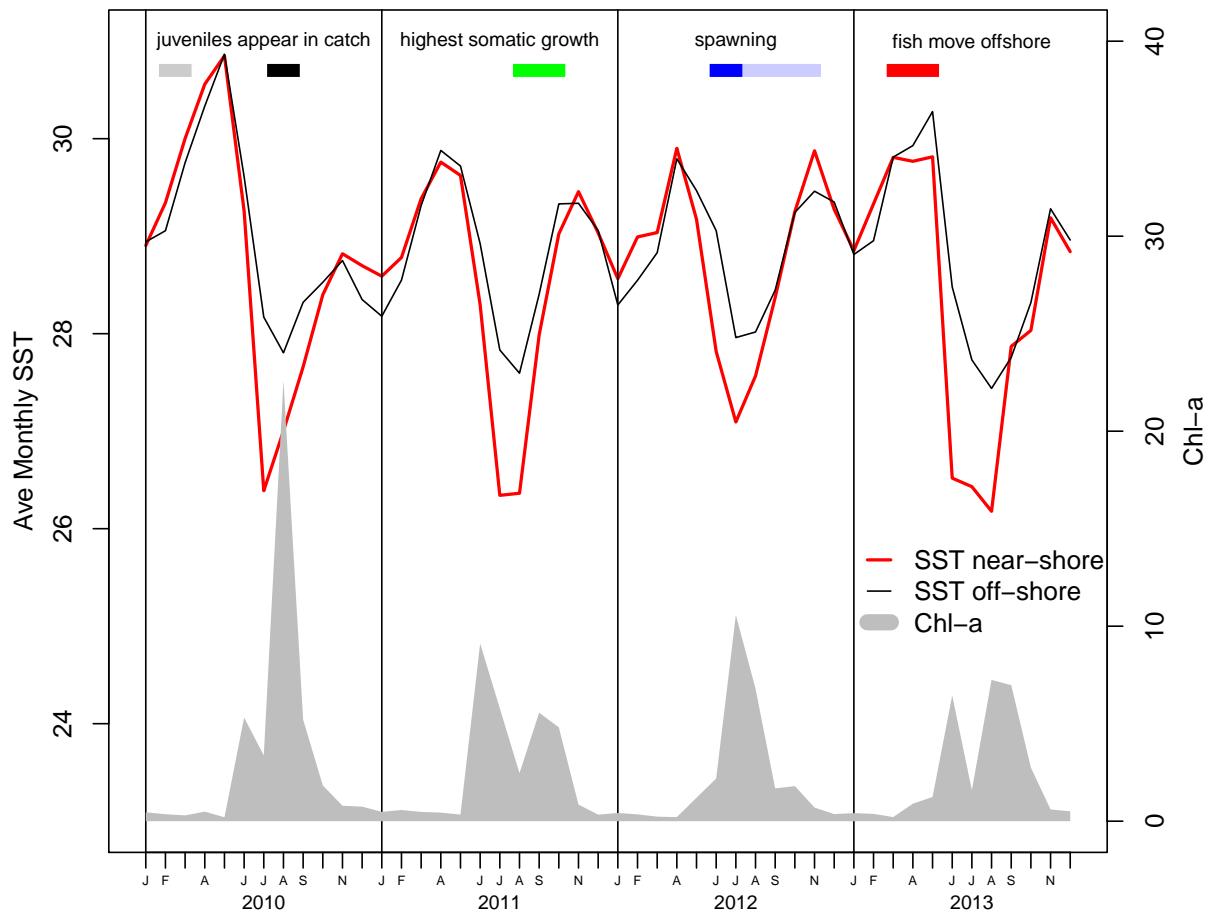


Figure 5

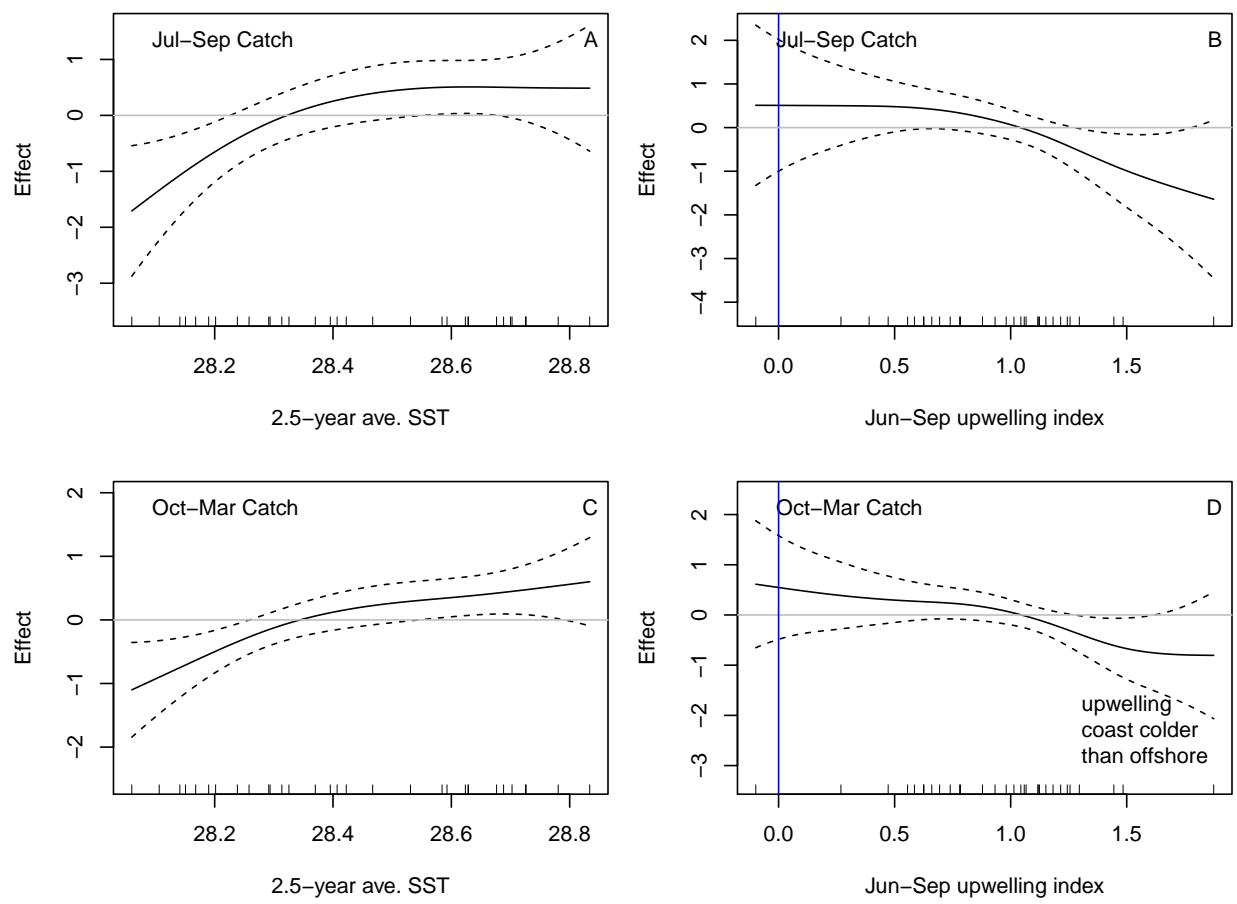


Figure 6

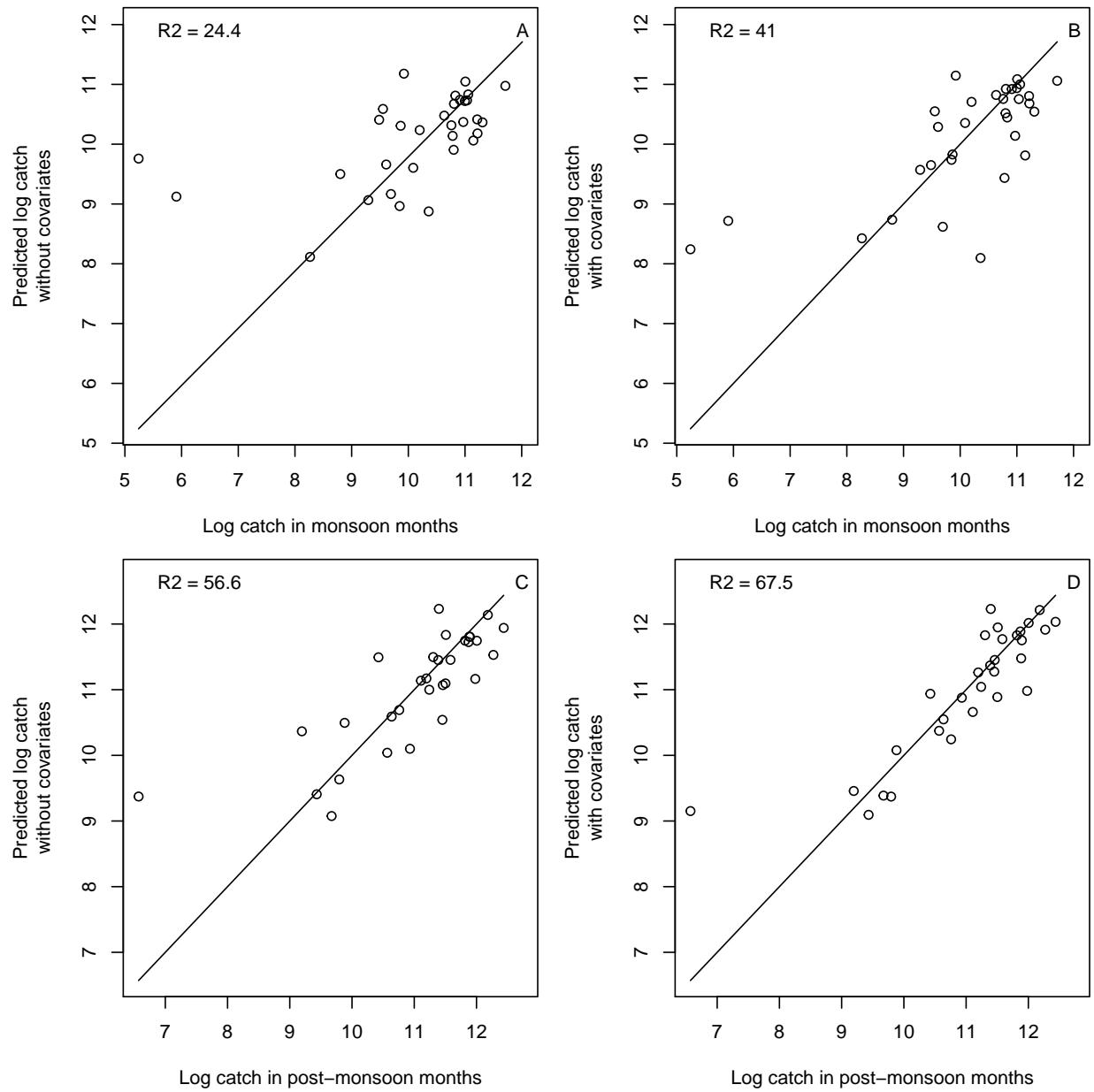


Figure 8