Use of satellite data for understanding and predicting oil sardine (*Sardinella longiceps*) catch variability along the southwest coast of India

14 November, 2017

Introduction

The Indian oil sardine (*Sardinella longiceps* Valenciennes, 1847) is one of the most commercially important fish resources along the southwest (Malabar) coast of India and contributes to about 40% of the marine fish catch of the state of Kerala (Srinath 1998). Landings of the oil sardine peak during, (June-Sept) and after (Oct-Dec), the summer monsoon period, in conjunction with the onset and early relaxation of coastal upwelling. However, the landings of this pelagic finfish are highly variable for reasons not well understood. Earlier researchers have indicated that the variable yields may be due to changing food sources and variable offshore habitats (needs reference and better explanation). However, Longhurst and Wooster (1990) described a decadal variability in sardine abundance on the Malabar coast and established a significant positive correlation with sea level, upwelling intensity, sea surface temperature, and rainfall.

The life cycle (include graphic of the life-cycle) of the oil sardine fishery begins with the entry of adult fish into the inshore areas during June – July, corresponding with the onset of the southwest monsoon (Chidambaram 1950, Antony Raja 1969) and restricted along the narrow strip of the western India continental shelf, within 20 km from the shore. The fish migrate from offshore to coastal waters and vice versa coinciding with the prevailing wind conditions (Hornell 1910). A gradual increase in temperature ranging from 26 to 28°C is favourable for their inshore migration (Chidambaram 1950). According to Nair et al. (2016), the exact location of oil sardine spawning grounds along the Indian coastline is still unclear; however, it is generally believed that the spawning of oil sardine occurs during the southwest monsoon period (Jun-Jul) when temperature, salinity and suitable food availability are conducive for larval survival (Murty and Edelman 1966, Jayaprakash and Pillai 2000, Krishnakumar et al. 2008), although some studies suggest that spawning can occur as late as September (Hornell 1910, Hornell and Nayudu 1923, Antony Raja 1969, Prabhu and Dhulkhed 1970).

After eggs are spawned, they develop rapidly into larvae within 24 hrs (Nair 1959). Variation in the bloom initiation time and intensity lead to changes in the food supply to sardine larvae and corresponding changes in their growth rate and survival and the later recruitment of sardines into the fishery (George et al. 2012). The phytoplankton bloom that provides for sardine larvae food is dependent upon coastal upwelling. Oil sardines grow rapidly during the first few months and mature to full size during the first 2.5 years.

The age at first maturity occurs at less than one year, at about 150 mm size, and the range of sardines that are fished typically range from 150-220 mm in size and 1-2.5 years old. On the southwest coast, the oil sardine fishery commences soon after the outbreak of monsoon in June and continues until March-April. Along the Kerala coast catches are fairly high throughout the year except during March-May (Figure 1). Along the west coast, the length composition of catches ranges between 50- 220mm. Virgin spawners (140-160 mm) enter the fishery on the Kerala coast during June/July and as the season progresses to September-December, adults occur in reduced quantities and catches are dominated by age-1 juveniles (120-140 mm) that were spawned in the prior season (citation). Seasons of rainfall extremes (weak or intense) cause recruitment failure, while a daily rainfall of 20-30 mm during June-August along the southwest coast may indicate good recruitment to the fishery (citation).

As the success of commercial fishery for each season is determined by the number of juveniles recruited at the beginning of the same season, rainfall, which affects spawning success, has been used to forecast the strength of juvenile brood entering the fishery. Though earlier studies have shown that most of spawning activity occurs in the near-shore area or in the mouths of estuaries, there are observations of eggs and larvae in the shelf

region even at 200m depth (citation). Also, studies from other upwelling ecosystems (citation), have shown that the planktonic stage larvae accumulate in frontal zones as well as in river plumes. Studies from Bakun (citation), Bakun and English (citation), Cury and Roy (citation) etc. explain the linkage to the atmospheric and oceanic processes based on number of hypothesis like Optimal environment window, Match-mismatch hypothesis etc. Also, the changes in the fishery in either short term or long term is explained for other similar theories such as population expansion (Mac Call's hypothesis) and the Wasp waist hypothesis.

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. These tools are relevant for fisheries management (IOCCG 2009; citation?) as well as development of predictive capabilities. Add more discussion regarding prior work/references on using satellite data to predict oil sardine catch. Satellite chlorophyll, SST and SSH have been used to estimate sardine habitat of the west coast of the USA (Zwolinski et al. 2011). In the last few years, 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)

In this paper, we revisit the analysis of environmental covariates that correlate with oil sardine catch using remote-sensing data and generalized additive modelling (Wood 2017). Based on biological information concerning how environmental conditions in the coastal waters affect spawning density, exposure of sardines to the coastal fishery, and larval and juvenile sardine survival, we develop a set of hypotheses about which satellites covariates should correlate with current or future landings. We use generalized additive modelling of the long-term landing data with long-term remote-sensing data to test for these correlations.

Study Area

Being land-locked to the north, the northern Indian Ocean is the only part of the world oceans that is not connected to the polar region and lies entirely in the tropics. It also experiences exclusively the seasonally reversing monsoon that is coupled with processes in the Pacific Ocean (Webster et al. 1999). The western Arabian Sea that experiences high atmospheric pressure during most of the year, lies adjacent to the Horn of Africa as well as the Arabian peninsula. This results in the Aeolian transport (referred to as Shamal activity) to the basin as well as persistent upwelling off-Somalia and along the coast of Yemen (Preusser et al. 2002, Rao et al. 2003, Singh et al. 2008). High plankton biomass, resulting in elevated productivity in this region is advected towards central Arabian Sea through Ekman pumping. This process is intensified during ISM and during the early spring-bloom period in February through March. A persistent Oxygen Minimum Zone (OMZ) in the north and central Arabian Sea basin results in large net-export of CO2 from waters to the atmosphere (Sarma et al. 1998, Naqvi et al. 2000, Kumar et al. 2001). Often this coastal productivity affects to the OMZ and vice-a-versa (Devol et al. 2006). Monsoonal reversal of winds and, in turn, coastal currents, along with high amount of river discharge, generates high degrees of variability in the productivity induced from either mixing or upwelling in the eastern Arabian Sea (Sarupria and Bhargava 1993, Srinivas and Dineshkumar 2006, Habeebrehman et al. 2008, Jayaram et al. 2010, Chauhan et al. 2011, Joshi and Rao 2012, Gupta et al. 2016).

In this work, we focus on the the southeastern Arabian Sea (SEAS) (Figure 2). The SEAS has relatively narrow continental shelf on the western shoreline of the Indian peninsula and is influenced by ISM as well as IWM and tidal interactions as well as discharge from multiple rivers (Shetye et al. 1995, Oasim 2003, Menon et al. 2005). This region also recurrently witnesses intense spring blooms of Trichodesmium spp. This, along with coastal current dynamics, upwelling and high rainfall, results in high coastal productivity with surface chlorophyll values up to 32-39 mg m-3 off the Kerala coast (Madhupratap et al. 2001, Jayaram et al. 2010, Raghavan et al. 2010, Chauhan et al. 2011). The SEAS also experience high chlorophyll concentration s(~40 mg m-2) and net primary productivity (~1300 mgC m-2 d-1) off Kochi during summer upwelling (Gupta et al. 2016). In the southern SEAS coastal waters off Kochi, extreme variability in productivity during the ISM (40-1225 mgCm-2d-1), post-ISM (21-263 mgCm-2d-1) and during intermonsoon period (12-648 mgCm-2d-1)

have been observed (Sarupria and Bhargava 1993, Srinivas and Dineshkumar 2006, Habeebrehman et al. 2008, Jayaram et al. 2010, Joshi and Rao 2012, Gupta et al. 2016).

Hypotheses concerning correlation between covariate and catches

Spawning occurs during the summer monsoon, primarily during quarter 3. During this time, sardines move inshore from offshore and become exposed to the primarily coastal fishery. Factors that affect inshore migration or spawning should affect the catch in quarter 3, which is primarily of spawning fish. The following are the major factors that are hypothesized to affect the catch in quarter 3 due to either a stronger spawning class or higher (or lower) exposure of the spawning age fish to the fishery due to higher (or lower) inshore migration.

Quarter 3 (spawning) catch correlates

- Precipitation low or high during either the summer monsoon (June-July) or pre-monsoon (Apr-May) is associated with low spawning.
- Low sea surface temperature is associated with delayed and limited spawning in Pacific sardines (Jacobson and MacCall 1995). The reasons for this may be behavioral as low temperature as associated with poor survival of larvae. Larval mortality rates may be higher in colder water due to increased zooplankton predation and upwelling (associated with cold surface temperatures) advects larvae into offshore waters.
- In addition to advection of larvae offshore, upwelling brings low oxygen water to the surface which discourages inshore migration to coastal spawning areas. Thus due to both lower SST and lower dissolved oxygen, strong upwelling during the spawning season is expected to be negatively associated with the spawning biomass in the nearshore region.
- Salinity is also associated with spawning initiation; however we do not have long-term salinity data.
- Spawners are age 2+ fish, thus the biomass in the previous two seasons should be correlated with the biomass of spawners this year. This should especially be the case for the biomass in quarters 4, 1 and 2 of the previous season as these were age 1-2 fish. A stronger cohort of age 1-2 fish in the previous year should be associated with a stronger spawner class in the current year. Note that it could also be the case that higher catch of age 1-2 fish in the previous season would be associated with lower spawners in the current year due to depletion. However during the majority of our study period, the fishery was largely artisanal and is not thought to be sufficient to deplete the population.

After spawning, during the 4th quarter of the current calendar year and the 1st and 2nd quarters of the next calendar year, the length distribution of the catch shifts to smaller fish as juveniles begin shoaling inshore (citation). The catch during the non-spawning part of the season should be correlated with factors associated with increased juvenile shoaling and survival of eggs and larvae from previous seasons' spawning. The former is associated with higher exposure of the non-spawning biomass to the coastal fishery and the latter is associated with a stronger (larger) non-spawning biomass.

Post-spawning catch correlates

The post-spawning period corresponds to catch in quarter 4 of the current calendar year and quarters 1 and 2 of the next calendar year. The main environmental covariates that should be important are those associated with egg and larval survival as these have been found to be the most important drivers of recruitment (citation) and covariates associated with shoaling and inshore migration of juvenile fish.

- Sea surface temperature is associated with egg and larval survival (Jacobson and MacCall 1995).

- Upwelling is associated with higher productivity and higher density of zooplankton, which leads to better larval and juvenile growth and survival. Thus the strength of upwelling during the post-spawning quarters should be associated with higher non-spawning biomass in subsequent years.
- Chlorophyll blooms are signatures of high productivity from nutrient influx either due to upwelling or coastal inputs. Thus the bloom intensity in prior years should be associated with future non-spawning biomass. In addition, juveniles shoal in response to chlorophyll blooms (citation) and thus the near shore chlorophyll bloom intensity in the current season should be associated with higher post-spawning catches due to higher exposure of the non-spawning biomass to the coastal fishery.
- Fish age 1-2 should be correlated with spawners in the previous two seasons, since a larger spawner class should be associated with higher recruitment. Thus catch in quarter 3 the previous two calendar years is expected to be correlated with the catch after the summer monsoon (quarter 4 of the current calendar year with quarter 1 and 2 of the next calendar year).
- Because age 2 fish also appear in the catch after quarter 3, we also expect the post-spawning catch in the previous season to be correlated with the post-spawning catch in the current season.

Data

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. These data are used to produce the quarterly total landings estimates with a precision of ±5% (*CMFRI, 1969-1994; Bal and Rao, 1990, Srinath, 2005*). The quarterly landing data were log-transformed to stabilize the seasonal variance.

Add discussion of the composition of the fishery since the early 1950s. Fishery largely artisanal until recently. Add discussion of why researchers consider the catch to be an index of total biomass and that fishery is not having a large impact on the standing biomass (because fishery has been traditionally coastal and non-motorized).

Covariate data

We analysed monthly composites of the following satellite data: Sea Surface Temperature (SST), chlorophyll (CHL), Sea Surface Height (SSH), upwelling (UPW) and precipitation. Specifically the following datasets were used. For sea surface temperature, Advanced Very-High Resolution Radiometer (AVHRR) SST data. For 1981 to 2002, we used Pathfinder Version 5.2 product on a 4km grid. For 2003 to 2016, we used the AVHRR product from the POES. These data are at a 0.1° spatial scale. For he chlorophyll-a, we used SeaWiFS data on a 0.1° spatial scale.for 1981 to 2002, and MODIS data from 2003 to 2017 at 0.05° spatial scale. The upwelling index was based on the sea-surface temperature differential between near shore and 3 degrees offshore (citation) using our sea-surface data from 1981 to 2016. The index has been validated as more reliable metric of upwelling off the coast of Kerala compared to wind-based upwelling indices (Smitha citation). All these satellite data were retrieved from the NOAA ERDDAP server (See Appendix for full details). The satellite data were averaged across thirteen 2.5 degree by 2.5 degree boxes which roughly parallel the bathymetry (Figure 2). The precipitation data were downloaded from two different sources. The first was an estimate of the monthly precipitation over Kerala from land-based gauges. This time-series is available from the start of our landing data (1956). The second was a remote-sensing precipitation product, the NOAA Global Precipitation Climatology Project. This provides a precipitation estimate using a global 2.5 degree grid. We used a 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. Full details on the remote-sensing data sets are available in the appendices.

Models

We investigated correlations between environmental variables and sardine catch during spawning and post-spawning periods using generalized additive models (GAMs, Wood 2017). GAMs allow one to model the effect of a covariate as a flexible non-linear function and it was known that the effects of the environmental covariates were likely to be non-linear, albeit in an unknown way. Our approach is similar to that taken by Jacobson and MacCall (1995) in a study of the effects of SST on Pacific sardine recruitment.
The model for catch in quarter 3 was of the form

$$ln(S_t) = \alpha + f_1(c_1) + f_2(c_2) + g_1(ln(C_{t-1})) + g_2(ln(C_{t-2})) + \epsilon$$

where $ln(S_t)$ is the log catch in the 3rd quarter of year t. Catch in the 3rd quarter (July-Aug) captures mainly spawning age fish as it overlaps with the tail end of the spawning season—the fishery is closed July to mid-August during the height of the summer monsoon and spawning season. The model seeks to explain the variability in catch in the 3rd quarter via a non-linear function of the covariates, c_1 and c_2, plus a non-linear function of total catch in prior years. The catches were logged to stabilize and normalize the variance. The model is primarily statistical, meaning it should not be thought of as being a population growth model. It is a form of model that is often used for modelling the effects of covariates on the number of fish that recruit into the fishery (e.g. Jacobson and MacCall 1995, add others or a textbook). The covariates tested are those discussed in the section on covariates that have been hypothesized to drive the size of the spawning biomass exposed to the fishery. The relevant covariate may be in the concurrent with the catch or in a prior year.

The model for catch in the non-spawning quarters (October to June) took a similar form

$$ln(N_t) = \alpha + f_1(c_1) + f_2(c_2) + g_1(ln(S_{t-1})) + g_2(ln(N_{t-1})) + \epsilon$$

where $ln(N_t)$ is the log of the total catch in the post-spawning season October to June (a nine-month period that spans two calendar years). We included the catches during the spawning season (S_t) separately from the non-spawning season (N_t). The covariates that affect the egg and larval survival may be quite different from those affecting the age-1 to age-2 or age-2 to age-3 survival. We wanted to be able to model these separately. The covariates tested are those discussed in the section on covariates that have been hypothesized to drive both survival and growth of juvenile fish and the factors that lead migration of fish inshore and thus exposed to the coastal fishery.

Model selection was conducted in a step-wise fashion. The length of the covariate data was 33-35 years (1982 to 2015 or 2017) for most covariates and this was not sufficient for fitting all covariates simultaneously. The model for the effect of the past of biomass on current biomass (density-dependence) was first determined. Once a sufficient model for the density-dependence was determined, the covariates were studied individually and then jointly. F-tests and AIC on nested sets of GAM models (Wood et al. 2016) were used to evaluate the support for models. One feature of GAMs is that they allow the 'flexibility' or smoothing parameter of the response curve to be estimated. However we fixed the degree of flexibility so that reasonably smooth responses were achieved. Multi-modal or overly wiggly response curves would not make sense for our covariates. We used GAMs with smooth terms represented by penalized regression splines (Wood 2011, using mgcv package in R) and fixed the smoothing term at an intermediate value.

Results

There was support for including the post-spawning catch in the previous year as an explanatory variable for the catch during the spawning season (3rd quarter). Models with $ln(N_{t-1})$ were strongly supported over an intercept only model (Table 1, time-dependency test). However the addition of the catch two years prior, $ln(N_{t-2})$, lead to either no decrease in the residual deviance (i.e. increase in the explained variance) and in fact, increased the residual deviance for the model with non-linearity (Table 1, Linearity test). We also tested the support for non-linearity in the effect of the prior year catch on the catch in the spawning season. This was done by comparing models with $ln(N_{t-1})$ included as a linear term or as a non-linear

function $s(ln(N_{t-1}))$ (Table 1, Linearity test). The residual deviance decreased using a non-linear response however the cost was 1.4 degrees of freedom. The rest was only weak (non-significant) support for allowing a non-linear response. The results were the same if we used the full landings dataset (1956-2015) or only the data that used for the covariate analysis (1982-2015) (Tables A1 and A2).

There was no support for using precipitation during the summer monsoon (Jun-Jul) or pre-monsoon period (Apr-May) as an explanatory variable for the 3rd quarter catch (Table B1). 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 the 1982-2015 remote-sensing precipitation was used or the 1956-2015 land-based precipitation was used.

However, we found significant correlation between average sea surface temperature during the early post-spawning period (Jul-Dec) and catch during the spawning season (Table 2). Sea surface temperature has been found to be correlated with sardine biomass in a number of other studies [Jacobson and MacCall (1995); add the others]. The residual deviance 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. 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, despite having the lowest residual degrees of freedom, given the cost (loss of degrees of freedom). The model with the lowest AIC and with a p-value of 0.108 for the F-test was a model with only the past year and non-linear response. The response shows a step-response with a negative effect at low temperatures and then an increased effect a higher temperatures (Figure 3). This type of step-response has been found in studies of the effect of SST on recruitment in Pacific sardines (*citations from Jacobson and MacCall*). The R2 for this model was 0.50 (Table 2).

The strongest predictor of the catch during the spawning season however was the upwelling strength during Jan-Mar (Table 3), during a period when the young of the year and age-1 fish would be found feeding in large shoals in the coastal region. The is the time of year with the second highest catches that are dominated by small-sized fish (*citation*). The model with lowest residual deviance included a non-linear effect of upwelling in the prior year and two years prior. As for SST, the cost in terms of low of degrees of freedom led to this model not being the best supported. The best supported model included only the upwelling strength in the prior year, not two years prior, and included a non-linear response (Table 3). The R^2 for this model was 0.63 (Figure 4).

Still to do. Test interactions between the variables that turned out to be important.

Still to do. Same analyses for the catch in the post-spawning months.

Discussion

Needs to be written. The results make sense biologically and are supported by other studies.

Discuss why precipitation did not come up. Implications. Supports the use of SST and upwelling indices for forecasting.

Do a pilot analysis of chlorophyll however the data available are short and will be inconclusive.

References

Antony Raja, B. T. 1969. Indian oil sardine. CMFRI Bulletin 16:1–142.

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.

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.

Devol, A. H., A. G. Uhlenhopp, S. W. A. Naqvi, J. A. Brandes, D. A. Jayakumar, H. Naik, G. S., C. L. A., and T. Yoshinari. 2006. Denitrification rates and excess nitrogen gas concentrations in the arabian sea oxygen deficient zone. Deep Sea Research Part I: Oceanographic Research Papers 53:1533–1547.

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.

Gupta, G. V. M., V. Sudheesh, K. V. Sudharma, N. Saravanane, V. Dhanya, K. R. Dhanya, G. Lakshmi, M. Sudhakar, and S. W. A. Naqvi. 2016. Evolution to decay of upwelling and associated biogeochemistry over the southeastern arabian sea shelf. Journal of Geophysical Research: Biogeosciences 121:159–175.

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.

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., 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.

Joshi, M., and A. D. Rao. 2012. Response of southwest monsoon winds on shelf circulation off kerala coast, india. Continental Shelf Research 32:62–70.

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 south-west coast of india? Marine Fisheries Information Service T&E Series 198:1–5.

Kumar, S. P., N. Ramaiah, M. Gauns, V. V. S. S. Sarma, P. M. Muralidharan, S. Raghukumar, M. D. Kumar, and M. Madhupratap. 2001. Physical forcing of biological productivity in the northern arabian sea during the northeast monsoon. Deep Sea Research Part II 48:1115–1126.

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.

Menon, H. B., A. Lotliker, and S. R. Nayak. 2005. Pre-monsoon bio-optical properties in estuarine, coastal and lakshadweep waters. Deep Sea Research Part II: Topical Studies in Oceanography 48:1345–1368.

Murty, A. V. S., and M. S. Edelman. 1966. 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.

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. 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.

Naqvi, S. W. A., D. A. Jayakumar, P. V. Narvekar, H. Naik, V. V. S. S. Sarma, W. D'souza, S. Joseph, and M. D. George. 2000. Increased marine production of n2o due to intensifying anoxia on the indian continental shelf. Nature 408:346–349.

Oasim, S. Z. 2003. Indian estuaries. Book, Allied, Mumbai, India.

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.

Preusser, F., D. Radies, and A. Matter. 2002. A 160,000-year record of dune development and atmospheric circulation in southern arabia. Science 296:2018–2020.

Raghavan, B. R., T. Deepthi, S. Ashwini, S. K. Shylini, M. Kumarswami, S. Kumar, and A. A. Lotliker. 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.

Rao, P. G., H. R. Hatwar, M. H. Al-Sulaiti, and A. H. Al-Mulla. 2003. Summer shamals over the arabian gulf. Weather 58:471–478.

Sarma, V. V. S. S., M. D. Kumar, and M. D. George. 1998. The central and eastern arabian sea as a perennial source of atmospheric carbon dioxide. Tellus B 50:179–184.

Sarupria, J. S., and R. M. S. Bhargava. 1993. Seasonal primary production in different sectors of the eez of india. Mahasagar 26:139–147.

Shetye, S. R., A. D. Gouveia, S. Y. Singbal, C. G. Naik, D. Sundar, G. S. Michael, and G. Nampoothiri. 1995. Propagation of tides in the mandovi-zuari estuarine network. Proceedings of the Indian Academy of Sciences-Earth and Planetary Sciences 104:66–682.

Singh, R. P., A. K. Prasad, V. K. Kayetha, and M. Kafatos. 2008. Enhancement of oceanic parameters associated with dust storms using satellite data. Journal of Geophysical Research: Oceans 113:C11008.

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

Srinivas, K., and P. K. Dineshkumar. 2006. Atmospheric forcing on the seasonal variability of sea level at cochin, southwest coast of india. Continental Shelf Research 26:1113–1133.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben. 1999. Coupled ocean–atmosphere dynamics in the indian ocean during 1997–98. Nature 401:356–360.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation 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. 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 per month catch rate is second highest in quarter 3 although total catch is lower in quarter 3.

Figure 2. 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 modelled.

Figure 5. Key currents and circulation patterns in the Indian Ocean that affect winter and summer monsoon associated upwelling.

Figure 6. Cartoon of the sardine life-cycle in the SE Indian Ocean and how it interacts with the fishery. [*missing*]

Figure 7. Average Chlorophyll-a bloom intensity in 2016 by quarter. Red is high density chlorophyll-a. Black is missing data due to cloud cover. Data from NOAA.

Figure 8. Monthly rainfall 1956-2014 from Kerala.

Figure 9. Monthly Upwelling Index 1981-2012 off the Kerala coast. The upwelling index is the difference between the near-shore and offshore sea surface temperature.

Figure 10. Monthly Chl-a 1997-2017 off the Kerala coast.

Figure 11. Monthly SST 1981-2012 off the Kerala coast. [*missing*]

Figure 3. Non-linear effect of temperature on catch during the spawning season (3rd quarter). [*missing*]

Figure 4. Fitted versus observed catch in the spawning period from the best covariate model. [*missing*]

Table 1. Model selection tests of time-dependency and linearity for the S_t model using F-tests of nested models fit to log landings data. S_t is the catch during the spawning period (Qtr 3 = July-Sep) of season t (Jul-Jun). N_{t-1} is the catch during the prior sardine season after the spawning period (the 9-months following Qtr 3, Oct-Jun, of the previous sardine season). N_{t-2} is the same for two seasons prior. $s()$ is a non-linear function of the response variable.

Model	Residual df	Residual deviance	F	p value	AIC
Time dependency test 1982-2015 data					
$ln(S_t) = \alpha + \epsilon$	33	65.13			122.59
($\text{Var}(\epsilon) = 1.97$)					
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	32	49.32	9.94	0.004	115.14
($R^2adj = 22\%$, $\text{Var}(\epsilon) = 1.59$)					
$ln(S_t) = \alpha + \beta_1 ln(N_{t-1}) + \beta_2 ln(N_{t-2}) + \epsilon$	31	49.28	0.02	0.878	117.11
($R^2adj = 19\%$, $\text{Var}(\epsilon) = 1.59$)					
Linearity test 1982-2015 data					
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	32	49.32			115.14
($R^2adj = 22\%$, $\text{Var}(\epsilon) = 1.49$)					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + \epsilon$	30.6	44.36	2.3	0.131	113.51
($R^2adj = 28\%$, $\text{Var}(\epsilon) = 1.49$)					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + s(ln(N_{t-2})) + \epsilon$	28.3	43.39	0.29	0.777	116.41
($R^2adj = 25\%$, $\text{Var}(\epsilon) = 1.49$)					

Appendix A: Tests for the structure of the Qtr 3 and non-Qtr 3 catch models

Table A1. F-tests of nested models of the time-dependency of the catch in Qtr 3 using the full catch data (1956-2015) or only the data that overlap with the satellite covariate data (1982-2015). See Table 1 for definitions.

Model	Residual df	Residual deviance	F	p value	AIC
Time dependency test 1956-2015 data					
$ln(S_t) = \alpha + \epsilon$	57	80.84			187.85
$(\text{Var}(\epsilon) = 1.42)$					
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	56	67.28	11.16	0.002	179.21
$(R^2adj = 15\%, \text{Var}(\epsilon) = 1.22)$					
$ln(S_t) = \alpha + \beta_1 ln(N_{t-1}) + \beta_2 ln(N_{t-2}) + \epsilon$	55	66.83	0.37	0.545	180.82
$(R^2adj = 14\%, \text{Var}(\epsilon) = 1.22)$					
Time dependency test 1982-2015 data					180.82
$ln(S_t) = \alpha + \epsilon$	33	65.13			122.59
$(\text{Var}(\epsilon) = 1.97)$					
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	32	49.32	9.94	0.004	115.14
$(R^2adj = 22\%, \text{Var}(\epsilon) = 1.59)$					
$ln(S_t) = \alpha + \beta_1 ln(N_{t-1}) + \beta_2 ln(N_{t-2}) + \epsilon$	31	49.28	0.02	0.878	117.11
$(R^2adj = 19\%, \text{Var}(\epsilon) = 1.59)$					

Table A2. F-tests of nested models of the linearity of the catch in Qtr 3 using the full catch data (1956-2015) or only the data that overlap with the satellite covariate data (1982-2015). See Table 1 for definitions.

Model	Residual df	Residual deviance	F	p value	AIC
Linearity test 1956-2015 data					
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	56	67.28			179.21
$(R^2adj = 15\%, \text{Var}(\epsilon) = 1.17)$					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + \epsilon$	54.4	63.46	2.07	0.145	177.98
$(R^2adj = 19\%, \text{Var}(\epsilon) = 1.17)$					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + s(ln(N_{t-2})) + \epsilon$	51.9	62.04	0.48	0.66	180.74
$(R^2adj = 17\%, \text{Var}(\epsilon) = 1.17)$					
Linearity test 1982-2015 data					180.74
$ln(S_t) = \alpha + \beta ln(N_{t-1}) + \epsilon$	32	49.32			115.14
$(R^2adj = 22\%, \text{Var}(\epsilon) = 1.49)$					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + \epsilon$	30.6	44.36	2.3	0.131	113.51
$(R^2adj = 28\%, \text{Var}(\epsilon) = 1.49)$					
$ln(S_t) = \alpha + s(ln(N_{t-1})) + s(ln(N_{t-2})) + \epsilon$	28.3	43.39	0.29	0.777	116.41
$(R^2adj = 25\%, \text{Var}(\epsilon) = 1.49)$					

Table A3. Model selection tests of time-dependency the NS_t model using F-tests of nested models fit to log landings data. NS_t is the catch during the non-spawning period (Qtrs 3, 1 and 2: Oct-Jun) of season t (Jul-Jun). S_{t-1} and NS_{t-1} are the catch during the prior sardine season during and after the spawning period respectively. S_{t-2} and NS_{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. For Test C, the nestedness is lines 1-3 and lines 1-2 and 4.

Model	Residual df	Residual deviance	F	p value	AIC
Time dependency test A 1982-2015 data					
$ln(NS_t) = \alpha + \epsilon$	32	46.47			108.95
$(\text{Var}(\epsilon) = 1.45)$					
$ln(NS_t) = \alpha + \beta ln(S_{t-1}) + \epsilon$	31	32.57	13.16	0.001	99.22
$(R^2adj = 28\%, \text{Var}(\epsilon) = 1.06)$					
$ln(NS_t) = \alpha + \beta_1 ln(S_{t-1}) + \beta_2 ln(S_{t-2}) + \epsilon$	30	31.68	0.84	0.366	100.3
$(R^2adj = 27\%, \text{Var}(\epsilon) = 1.06)$					
Time dependency test B 1982-2015 data					
$ln(NS_t) = \alpha + \epsilon$	32	46.47			108.95
$(\text{Var}(\epsilon) = 1.45)$					
$ln(NS_t) = \alpha + \beta ln(NS_{t-1}) + \epsilon$	31	27.42	21.17	0	93.54
$(R^2adj = 39\%, \text{Var}(\epsilon) = 0.9)$					
$ln(NS_t) = \alpha + \beta_1 ln(NS_{t-1}) + \beta_2 ln(NS_{t-2}) + \epsilon$	30	27	0.46	0.503	95.03
$(R^2adj = 38\%, \text{Var}(\epsilon) = 0.9)$					
Time dependency test C 1982-2015 data					
$ln(NS_t) = \alpha + \beta ln(NS_{t-1}) + \epsilon$	31	27.42			93.54
$(\text{Var}(\epsilon) = 0.88)$					
$ln(NS_t) = \alpha + \beta_1 ln(NS_{t-1}) + \beta_2 ln(S_{t-1}) + \epsilon$	30	27.33	0.09	0.763	95.43
$(R^2adj = 37\%, \text{Var}(\epsilon) = 0.91)$					
$ln(NS_t) = \alpha + \beta_1 ln(NS_{t-1}) + \beta_2 ln(S_{t-2}) + \epsilon$	30	27.41	0.01	0.907	95.52

Model	Residual df	Residual deviance	F	p value	AIC
($R^2adj = 37\%$, $\mathrm{Var}(\epsilon) = 0.91$)					

Table A4. Model selection tests of time-dependency the NS_t model using non-linear responses instead of linear responses as in Table A3 See Table A3 for an explanation of the parameters and model set-up.

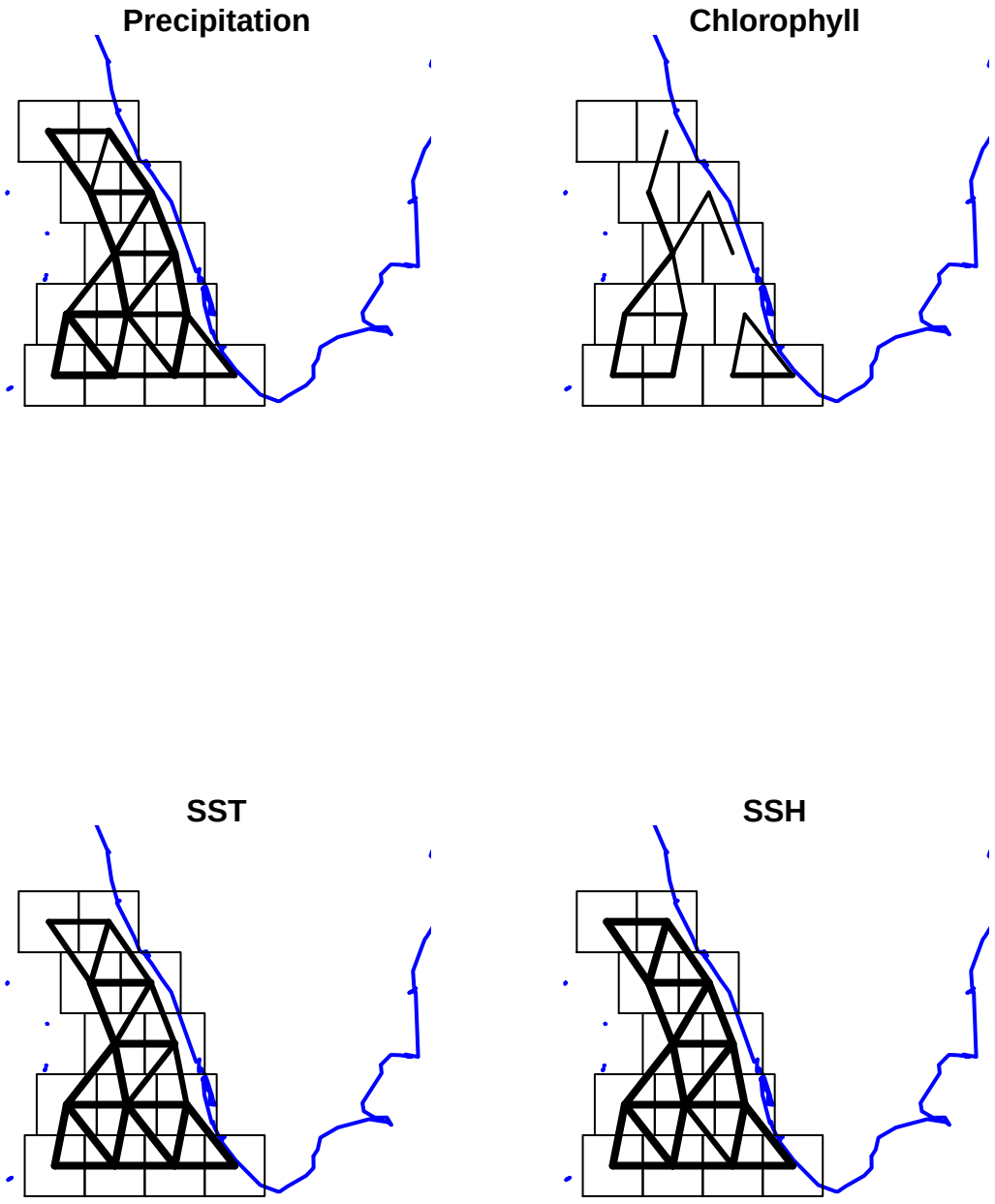
Model	Residual df	Residual deviance	F	p value	AIC
Time dependency test A 1982-2015 data					
$ln(NS_t) = \alpha + \epsilon$	32	46.47			108.95
($\text{Var}(\epsilon) = 1.45$)					
$ln(NS_t) = \alpha + s(ln(S_{t-1})) + \epsilon$	29	29.54	5.83	0.003	98.92
($R^2adj = 31\%$, $\text{Var}(\epsilon) = 0.96$)					
$ln(NS_t) = \alpha + s_1(ln(S_{t-1})) + s_2(ln(S_{t-2})) + \epsilon$	26	26.12	1.19	0.334	99.77
($R^2adj = 34\%$, $\text{Var}(\epsilon) = 0.96$)					
Time dependency test B 1982-2015 data					
$ln(NS_t) = \alpha + \epsilon$	32	46.47			108.95
($\text{Var}(\epsilon) = 1.45$)					
$ln(NS_t) = \alpha + s(ln(NS_{t-1})) + \epsilon$	29.5	23.04	13.46	0	89.79
($R^2adj = 47\%$, $\text{Var}(\epsilon) = 0.71$)					
$ln(NS_t) = \alpha + s_1(ln(NS_{t-1})) + s_2(ln(NS_{t-2})) + \epsilon$	27.2	19.95	1.87	0.169	88.81
($R^2adj = 51\%$, $\text{Var}(\epsilon) = 0.71$)					
Time dependency test C 1982-2015 data					
$ln(NS_t) = \alpha + s(ln(NS_{t-1})) + \epsilon$	29.5	23.04			89.79
($\text{Var}(\epsilon) = 0.77$)					
$ln(NS_t) = \alpha + s_1(ln(NS_{t-1})) + s_2(ln(S_{t-1})) + \epsilon$	26.8	21.78	0.59	0.607	92.42
($R^2adj = 46\%$, $\text{Var}(\epsilon) = 0.78$)					
$ln(NS_t) = \alpha + s_1(ln(NS_{t-1})) + s_2(ln(S_{t-2})) + \epsilon$	26.5	15.97	4.04	0.017	82.62
($R^2adj = 60\%$, $\text{Var}(\epsilon) = 0.58$)					

Table A5. Model selection tests of time-dependency for the NS_t model using 1956-2015 data. See Table A3 for definitions.

Model	Residual df	Residual deviance	F	p value	AIC
Time dependency test B linear 1956-2015 data					
$ln(NS_t) = \alpha + \epsilon$	56	54.2			162.89
($\mathrm{Var}(\epsilon) = 0.97$)					
$ln(NS_t) = \alpha + \beta ln(NS_{t-1}) + \epsilon$	55	36.81	25.5	0	142.84
($R^2adj = 31\%$, $\mathrm{Var}(\epsilon) = 0.68$)					
$ln(NS_t) = \alpha + \beta_1 ln(NS_{t-1}) + \beta_2 ln(NS_{t-2}) + \epsilon$	54	36.81	0	0.976	144.84
($R^2adj = 30\%$, $\mathrm{Var}(\epsilon) = 0.68$)					
Linearity test 1956-2015 data					
$ln(NS_t) = \alpha + \beta ln(NS_{t-1}) + \epsilon$	55	36.81			142.84
($\mathrm{Var}(\epsilon) = 0.67$)					
$ln(NS_t) = \alpha + s(ln(NS_{t-1})) + \epsilon$	53.6	35.24	1.75	0.191	142.2
($R^2adj = 33\%$, $\mathrm{Var}(\epsilon) = 0.65$)					
$ln(NS_t) = \alpha + s_1(ln(NS_{t-1})) + s_2(ln(NS_{t-2})) + \epsilon$	51.3	33.96	0.87	0.439	143.8
($R^2adj = 33\%$, $\mathrm{Var}(\epsilon) = 0.65$)					

Appendix B: Model selection for the covariates

Appendix C: Correlation of covariates across the boxes



Correlation of the covariates across boxes. Correlation is shown by the width of the lines between neighboring boxes.