Received: 2 October 2017
IOTC-2017-WPTT19-INFO1

Bayesian CPUE Standardisation Model for Maldives Pole and Line Skipjack Tuna 1970-2015

Paul A. H. Medley[1], Mohamed Ahusan[2] & M. Shiham Adam[2]

Abstract

We outline the problems and progress in finding solutions for extending an abundance index for skipjack tuna back from 2004 to 1970 when catch and effort data collection began. We report the potential solutions for two problems, being the reconstruction of vessel length information using a vessel survival regression (described in Medley et al. 2017) and a random effects component used to account for missing mechanisation information on the fleet 1974-1979. Preliminary results indicate that remaining problems are still significant and will need to be resolved before an extended index can be proposed.

Introduction

The first Indian Ocean skipjack model-based stock assessment conducted in 2011 used the Maldives pole and line standardised CPUE as an abundance index. These data were only used from 2004 when monthly information was recorded for each vessel. Earlier data exist from 1970, but these are compiled monthly records by atoll and do not record the activities of individual vessels. In addition, significant corollary information about the fleet operations of these earlier data is missing, making it difficult to use all data in a single consistent index (Kolody et al. 2010). Here we report progress in dealing with various problems in combining the different data sources into a single coherent index. A number of problems need to be resolved, and it is not yet clear than acceptable solutions can be found for all of these.

We give here two solutions for two separate problems; accounting for missing covariates on vessel size and whether vessels are motorized or not. For the first, we used a reconstruction of fleet size composition described in Medley et al. 2017. This gives a probabilistic account of fleet length composition which can be used to account for important changes in fishing efficiency. Secondly, we use a random effects model to estimate the probability for vessels being mechanised for the initial period after mechanisation occurred. Both these techniques look promising and it is hoped similar approaches might be found for other problems that have been identified.

[1] Alne, United Kingdom, paulahmedley@gmail.com
[2] Marine Research Centre, Malé, Maldives; mahusan@mrc.gov.mv, msadam@mrc.gov.mv

Methods

Data

The Maldives catch and effort data collection has undergone extensive revisions since it began in 1970 (Table 1). Fishery data collection began in 1959 when landings were reported in numbers of tuna reported to the island offices or collected by the island clerks at the time and site of landing. These data were compiled and monthly reports were sent to the Ministry of Fisheries. Initially, the data collection system did not distinguish between gears. This was because traditionally, the Maldivian vessels would be gear specific to the type of fishing vessel:

- *Bokkuraa* (small wooden boats 3-5 m. originally powered with oars now mostly by outboard engines) used for trolling and handling within atolls and on coral reefs.
- *Vadhu dhoni* (5–8m originally sail now motorised) used for troll fishing near the islands and within atoll lagoons as well as general coral reef fishing.
- *Masdhoni* (10–12m standard pole-and-line vessels) which use livebait to catch predominantly skipjack and yellowfin tuna. The *masdhoni* underwent transformation following mechanization and with the introduction of the handline fishery for large yellowfin tuna, these vessels are also used to conduct both fisheries.

From 1959 data were only recorded from *masdhoni* vessels. In 1966 the system was expanded to include the *vadhu dhoni* fleet. At this time, numbers of tuna were only recorded in three categories: large skipjack; small skipjack and yellowfin; little (kawakawa) and frigate tuna. The system was expanded again in 1970 to record five categories of tuna separately in addition to catches of reef fish. From this point, with landings recorded by species, it is possible to estimate a standardised CPUE index for each species. The primary problem with the recorded data 1970-2004 is individual trips were not recorded, but landings and effort are reported combined. So much of the information is missing. In some cases, additional information was reported tracking fleet changes. Notably, the Ministry required island chiefs to report catches of sailing and mechanized *masdhonis* separately from 1979 after much of the fleet had already transitioned. Other changes to the fleet which may well have increased efficiency but have not been recorded include changes in fleet size composition, improved design and engine power.

Table 1 A summary of the history of data collection and issues for interpreting

Year	Notes
1970	Reported catches may have been inflated 1970-71 because a number of fishermen may have inflated catches in the hope of qualifying for a government prize. Although this incentive existed from mid 1950s to 1981, the problem was most apparent in 1970-1971 when cash prizes were given directly to top crews (Anderson, 1986).
1974	Vessel mechanisation starts, but is not recorded consistently
1979	Mechanized vessels begin to be recorded separate from sailing vessels.
1981	FAD installation begins. Prize money for high catches ceases.
1989	Vessel type and number of *dhoni* begin to be recorded, but mixed gear trips are not identified in data.
2004	Trip landing data begins to be recorded.
2010	Log-book data begins, but does not cover the entire fleet. Landings begin to be reported as numbers and/or weight rather than numbers.
2014	Detailed log book data on trip begins to be recorded, including bait, set type, fishing times by gear and location. Weight rather numbers becomes commonest data to report landings.

In addition, during the latter years to the early 1980s, fishing vessels which completed a certain number of fishing days were exempt from annual fishing vessel registration fee. This may have prompted the over-reporting of effort to avoid the fee. Details of the fee system are being sought to make a possible correction for this.

The largest potential source of errors for the catch weight data may be the conversion factors used to estimates the weight from recorded fish numbers (Cook 1995, Anderson, 1986). Several factors have been derived over the years. For the standard data, mean weights have been estimated as 2.1kg for small skipjack and 5.7kg for large skipjack. There appears to be little supporting evidence for these values and they are fixed over all years 1970-2016. This problem will eventually have to be addressed as all recent catch data are collected as weights.

Data were combined from three sources. The "per trip"" report data and logbook data are already used in the current 2004-2015 index. These data are ostensibly compatible as they have the same covariates recorded and have been previously combined into a single data set. The IPTP/MOFA Merged data 1970-2007 were drawn from previous work (Adam, 1999) and represent the monthly catch and effort by vessel type. The structure of these data is different and the data set was organised separately.

Fitting Method

To deal with the various issues arising for the different data sources, it was decided to use a Bayesian approach as the only way to deal with the problems in a consistent and transparent manner. To achieve this, the model is being developed in Stan (Stan 2016), which provides a good, robust platform for fitting Bayesian models using MCMC. Stan is designed to improve MCMC performance by using Hamiltonian Monte Carlo (HMC) sampling.

The following fits the weight model because these data have already been produced for the 1970-2004 data set.

The missing data (proportion vessels motorised 1975-1978) is estimated as a latent variable ("random effect") within the model. The probability for the proportion is based on a binomial for the approximate number of vessels in the fishery. Although this is calculated from the effort, effort is not used to calculate the variance because clearly effort days are not independent. The number of vessels contributing to the observed effort is not known, so this is estimated as effort / 20 (assuming each vessel fishes 20 days in a month). Lower values for the number of trials are preferred to ensure a reasonable variance around the expected proportion.

Standard Log-linear models are used to provide good first guesses for parameter estimates.

MCMC Model

This preliminary Stan model deals with two problems:

1. Unknown proportion of motorised vessels 1974-1979. The model describes the proportion motorised vessels where data is missing using the beta distribution with the same mean and variance as the binomial for the proportion motorised, so the Beta distribution parameters are calculated as: $\alpha = p * (n-1)$ and $\beta = (alpha/p) - alpha$, where p is the expected proportion motorised and n = number of vessels operational in that quarter.

2. Unknown fleet length composition 1970-2003.

The model structure is still under investigation. For this trial fit, the model had the following components for the 1970-2007 data: Atolls are treated as a categorical variable. Atoll classification had changed between the old and new data sets, so separate atoll factors were fitted to each data set. It is not clear how the atoll should affect the skipjack catch rate, so the current formulation will need to be revisited. Vessel length is fitted as a covariate in the same way as for the 2004-2015 index. Vessel power: sail vs mechanised. No other terms were fitted. Only main effects and no interaction effects were fitted.

For the random effects model, the proportion motorised is fitted as a binomial to the observed motorised / non-motorised trips for those quarters where the data exist with the expected proportion as a logistic function of time. The random effect variable is then estimated using the beta probability function consistent with a binomial having mean and variance taken from this logistic function and overall number of trips in that quarter. For the vessel length model, the proportion of trips undertaken by vessels at each length is assumed to be proportional to the vessel fleet size composition estimated separately (Medley et al. 2017). This proportion is multiplied by a vessel length effect to generate the expected overall effect for each quarter.

The model was fitted with a Gamma likelihood, where separate scale (beta) parameters were fitted for each data source.

Results and Discussion

All parameters showed MCMC convergence based on 4 chains, each with 6500 iterations and warmup=1500, thinning=4.

Overall upward trend unlikely, suggesting that vessel fishing efficiency has not been adequately accounted for. The danger is any trend will be induced from the standardisation and the trend will likely be removed or reversed by the standardisation. Such corrections unless carefully justified will undermine confidence in the use of the data series as an abundance index.

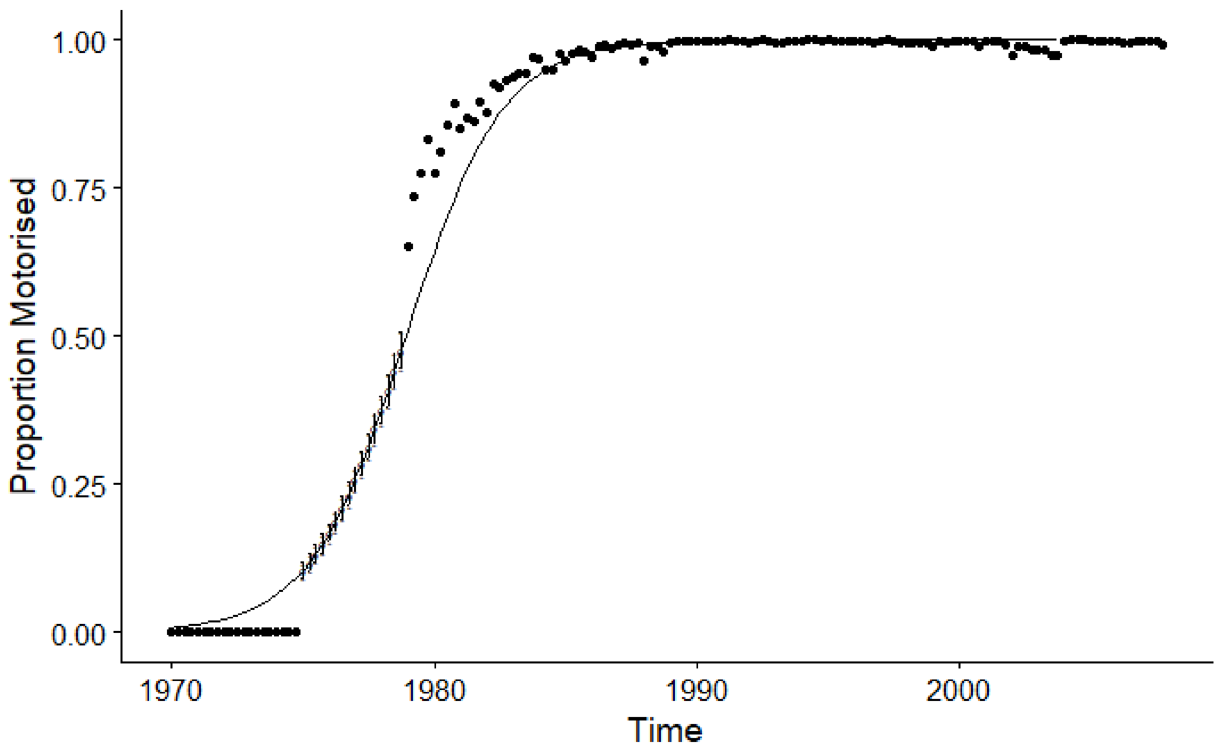

Figure 1 Proportion mechanised with observations, logistic curve defining expected proportion and "random effect" estimates with 90% confidence interval.

The two main parameters that adjust fishing power are the terms for vessel length and the sail vessel power category. The vessel length parameter (0.064 s.e. 0.001) was primarily estimated from the increase in catch rate present in the recent trip data 2004-2015. This adjustment was effectively applied through to 1970 based on the reconstructed vessel sizes.

The vessel power effect for sail was negative (-0.004 s.e. 0.004), but the estimate was very close to zero. This implied mechanisation had no significant effect on catch rates, which seems unlikely and points to a problem in the data.

It has been shown that the proportion of vessels being mechanised through the period 1974-1979 can be estimated, and the random effects approach can be used to account for errors in these estimates (Figure 1). However, a better model than the logistic is required to fit the observations better and, with the power effect being so close to zero, this part of the model had no impact on the final index estimates.

Any adjustments from the model on the nominal indices is therefore due to changes in vessel length only (Figure 2). At this stage, the model looks promising but significant problems remain which need to be addressed before the index could be put forward for consideration.

- Realistic estimates of the impact of mechanisation on fishing power.
- Other changes to vessel fishing efficiency over time which might be inferred from their registration date and other information.
- Mean fish weight will need to be estimated within the model working from recorded numbers of fish to develop a biomass index consistent with the most recent data.
- Gear used and mixed gear trips (indicated by other catches) which will lead overestimated effort directed at tuna.
- Finally, government initiatives to encourage fish production may have affected data records.

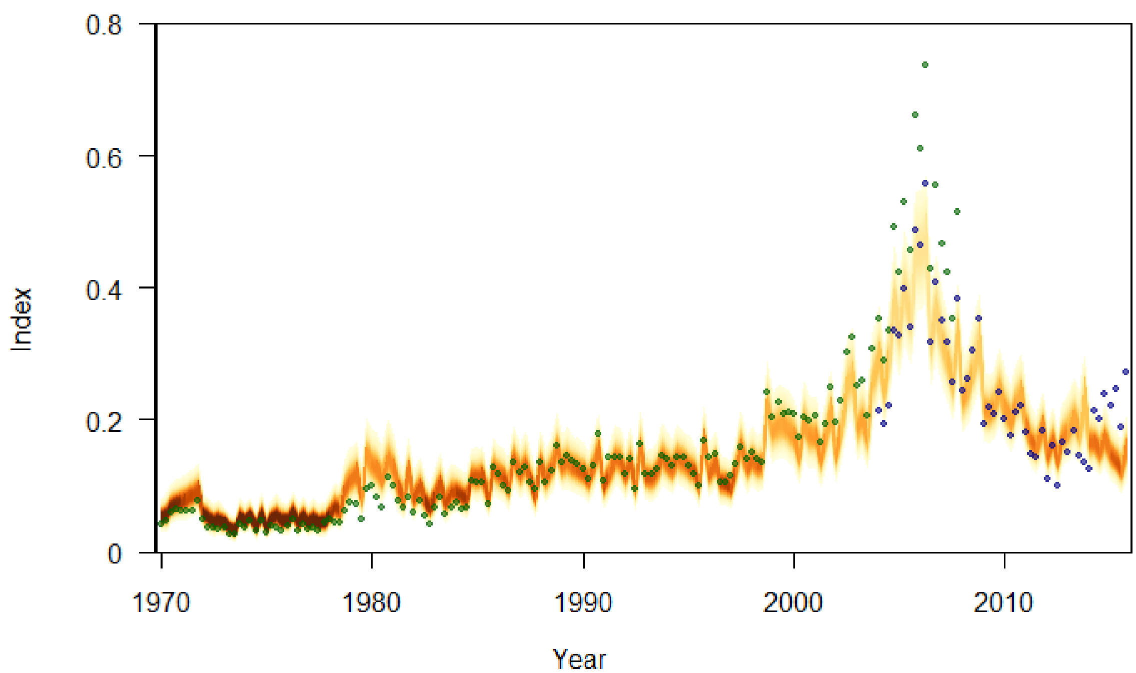

Figure 2 Watercolour plot of the index estimate for generalised linear model MCMC fit and scaled nominal indices for the monthly summary data (green) and new trip based data (blue).

References

Adam, M. S. (1999). Population dynamics and assessment of Skipjack tuna (Katsuwonus pelamis) in the Maldives. Unpublished Doctoral Thesis submitted to T. H. Huxley School of Environment, Earth Sciences and Engineering. London, Imperial College: 303 pages.

Anderson, R. C. (1986). Republic of Maldives tuna catch and effort data 1970-1983. Colombo, IPTP.

Cook, J. (1995). CPUE and conversion factors. Economic Planning and Coordination Section, Ministry of Fisheries and Agriculture, Malé. Maldives, unpublished report: 5 pages

Kolody, D., M. S. Adam, C. Anderson. 2010. Catch rate standardization for the Maldivian skipjack pole and line fishery 1970-2007. IOTC-2010-WPTT-05.

Medley, P., Ahusan, M., Adam M.S. 2017. Reconstruction of Maldives Historic Fleet Size Composition from Partial Register Data 1970-2004. This meeting.

Stan 2016 Stan 2016. Stan Development Team. Stan Modeling Language: User’s Guide and Reference Manual. Version 2.11 mc-stan.org