Received: 2 October
IOTC-2017-WPTT19-44

Maldives Pole and Line Skipjack Tuna CPUE standardization 2004-2015

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

Abstract

Abundance indices are an important requirement for reliable stock assessment. Maldives pole and line fishery provides a CPUE-based abundance index for the Indian Ocean skipjack stock assessment. The CPUE index needs standardization because the fishery has improved its efficiency over time. Generalized linear models are fitted consistent with previous standardization models using new data which has been carefully reviewed and corrected. The model was restructured based on a review of the available categorical and covariate variables. The final recommended index was produced from the new linear model fitted using a Markov chain Monte Carlo to the available data, with numbers of skipjack caught as the response variable.

Introduction

Indian Ocean skipjack catches have increased substantially since the late 1980s with the arrival of industrial purse seiners in the mid-1980s, and development of the fishery in association with drifting floating objects or Fish Aggregating Devices (FADs) since the 1990s (IOTC 2016: Skipjack Tuna Supporting Information). The WPTT has previously concluded that the stock was probably not overexploited, but analyses of indicators of stock status suggested close monitoring of the stock would be required in the following years (IOTC-2009-WPTT-R[E]).

Maldives pole and line (PL) CPUE series have been standardized and updated on several occasions in the past. Kolody, Adam and Anderson (2010) attempted the first standardization on atoll aggregated catch and effort data using three sets of data; i) operational catch and effort data from the pole-and-line fleet 2004-2007, ii) aggregated catch and effort data 1970-2007, and iii) the registry of new vessels 1958-2010. It was concluded that a meaningful standardization can only done on the vessel specific catch and effort dataset available from 2004 onwards.

Kolody and Adam (2011) conducted the first proper standardization of the vessel specific catch and effort data 2004-2010 using vessel registry information and a subset of daily observations from logbooks introduced in 2010. The authors attempted to study the effect

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

the large number of records with zero skipjack, prevalent in the dataset. The index of abundance was subsequently used for the first model-based assessment of the Indian Ocean skipjack tuna stock (WPTT 2011 Report). Before 2011, a formal model-based stock assessment was not conducted due to the lack of a reliable index of abundance.

Sharma, Geehan and Adam (2013), standardized the vessel specific dataset (2004-2011), updated in 2014, and incorporating Maldives' anchored FAD (aFADs) data into the analysis using the number of active aFADs associated with the nearest atoll that the landing data was collected from (Sharma *et al.*, 2013). The authors also attempted to estimate the standardized CPUE back to 1985 by using vessel lengths available from the vessel registry dataset, but the results were not considered sufficient reliable to be used in the stock assessment.

Despite the extensive fishing by industrial fleets in the western India Ocean, the assessment of the Indian Ocean skipjack stocks has so far depended upon the index of abundance from the Maldives pole-and-line CPUE series. It has been stated that purse seine effort data is difficult to standardize due to rapid changes in use of electronics in searching and the difficulty of obtaining standardized operation data (IOTC-WPTT 2012). This paper presents the efforts to update the 2014 Maldives PL CPUE standardization for use in stock assessment of 2017.

Methods

Data

The following datasets are used in the analysis:

- Vessel specific monthly catch and effort data 2004-2009; cleaned dataset (Ahusan, Sharma and Adam 2015) addressing issues reported in previous standardizations (e.g. zero skipjack catch records, duplicate records, records with sum of effort exceeding days of month).
- Vessel specific monthly Catch and Effort data for 2010-2015 filtered for pole and line (PL) and handline (HL) gear types and mechanized *masdhoni* vessel types. These data were originally recorded in numbers (thanks to the enumeration system of data collection established in the Maldives), but also given as weight (using conversion factors (Cook 1995).
- Logbook data from 2014 and 2015 aggregated over vessel, month and gear strata.
- Vessel registry datasets with date of registration, vessel descriptors such as vessel length, gross tonnage, engine horsepower etc.
- Anchored FAD dataset (1981-2017) with 473 records of location, deployment and lost date.

The Maldives Ministry of Fisheries introduced fishery logbooks to replace the island office reported catch and effort data collection system. Reliance on the island offices to report fishery data has been gradually replaced as the new system became established and reporting improved. The key difference between the two is the greater level of information reported in the new system (e.g. effort reporting, baiting, bycatch, ETP interaction, etc.). Despite its introduction in 2010 (and revision in 2012), the data has been scarce and unreliable for the initial years. As a result, only logbook data from 2014 and 2015 are used. Furthermore, it is observed that more of the catch data are being reported in weights unlike for the previous system, and only landings weights were being reported in 2016. It should be noted that 2016 data have been not used in this analysis. The data was made available very early September and needed cleaning and cross checking with logbooks.

Data Processing and Statistical Analysis

Data tables are loaded from the text files. These consist of the FAD database, with number of FADs by region and month, the associated atoll including their regional allocation and the monthly catch effort data. These are read into R dataframe objects and joined/mutated to create a single data frame with the relevant fields for analysis. Analyses were documented using RMarkdown scripts (http://rmarkdown.rstudio.com/) so that they are fully reproducible from the raw data.

FAD data consists of the deployment date, the location, and date of loss if the FAD has been lost. The date of loss was not always available for the earlier part of the time series (1981-1999 inclusive; one record with no loss date exists in 2005). To enable the FAD information to still be used, some loss date had to be assumed. For this purpose, the mean survival time for all FADs for which the loss date exists over the period 1981-1999 (mean longevity 630 days) was used to provide a loss date for those FADs missing this date. The survival time appears to have changed over the years, so the mean of all FAD survival times 1981-2015 (683 days) was not thought to be a suitable estimate. One FAD in 1986 had no deployment date. The date was estimated based on the FAD number as the mean deployment date for the FADs before and after it in the FAD register.

All additions and corrections were small changes and would not, it was believed, have a significant impact on the final results. FADs have increased over time, but from around 2000, the Maldives has maintained around 50 active FADs (Figure 1).

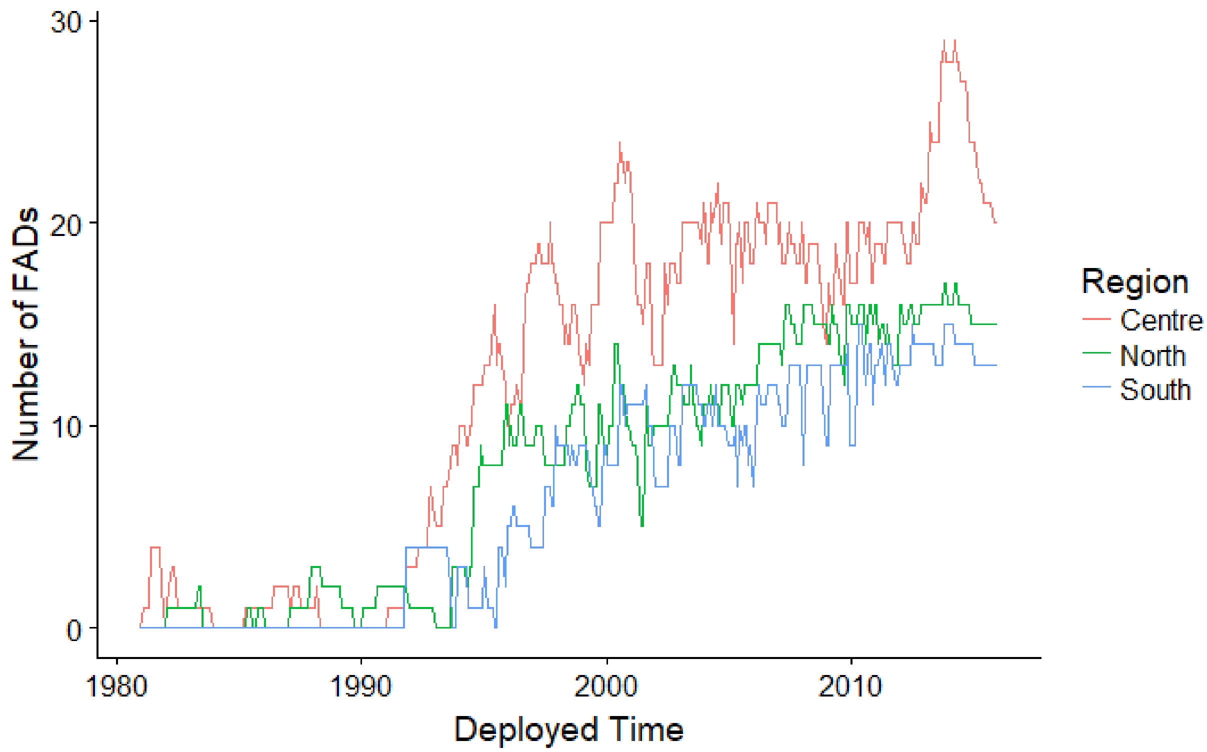

Figure 1: Number of active FADs over time and region.

Fitting Methods

Two methods were used to fit the models. The model structure, including the error model, was explored using maximum likelihood generalized linear models (glm package in R). This provides a robust method for model fitting with good diagnostics. However, the model structure is constrained, which will make it difficult to continue to use this approach when dealing with difficulties arising with the 1970-2004 data set, and it has limited options for accounting for uncertainty. The final model was therefore fitted using Markov chain Monte Carlo (MCMC) in the modelling software Stan (mc-stan.org; rstan package in R), which provides more flexible modelling and explicitly deals with uncertainty in a consistent way.

Stan software (Stan 2016) provides a robust approach to MCMC which allows it to fit models that defeat most other methods. The root cause for problems in MCMC is the slow convergence of conditional sampling when parameters are highly correlated in the posterior, a common problem for time-series models and hierarchical models with interacted predictors. Stan is designed to improve MCMC performance by using Hamiltonian Monte Carlo (HMC) sampling (as opposed to Gibbs sampling in BUGS, for example), which requires more complex calculations. HMC uses the gradient of the log posterior as well as the function value. Stan incorporates reverse-mode algorithmic differentiation, in much the same way ADMB (Methot, 2015, Fournier *et al.*, 2012) does, as well as various techniques to speed up MCMC simulations. Similarly to BUGS (Lunn et al., 2009) and ADMB, Stan uses a modelling language to define models in a familiar notation that is transformed to C++ code before being compiled into an efficient executable program. Importantly, Stan provides good diagnostic tools that all but guarantee a sample of draws from the posterior is fully representative, and applies more rigorous MCMC procedures than other MCMC software may do.

Results and Discussion

Response Variable

Two response variables were available: the catch in numbers and the catch in weight. Catch in numbers was used in the legacy index. It was apparent that the majority of catch weights are not observed in the monthly data, but have been calculated assuming 2.1kg per fish for the small skipjack and 5.7kg per fish for the large skipjack. The resulting catches in weight are therefore not independent observations and less suitable for modelling. Allocation between "small" and "large" skipjack may not be reliable, so combining their numbers should result in a more consistent index.

Records were filtered to remove records with zero skipjack catch or zero effort. The majority of observations were primarily for vessels 12m-27m in length and took place in the central region.

Legacy Index

The “legacy” generalized linear model (GLM) was fitted with the same structure as the standardization model fitted to the log-CPUE used in the 2014 stock assessment (Sharma, Geehan and Adam, 2014), but using data to the end of 2015.

This new fit showed similar residual patterns and issues as the previous fit. Only one record relates to the largest vessel class in the North region, so estimates of interaction terms was unreliable. Compared to the nominal index (unstandardized CPUE) the standardized legacy index shows significant differences which may be the result of overfitting. Residuals are clearly not normally distributed, with lower catch rates being over-represented. There is also some evidence of heteroscedasticity with perhaps the log transform leading to an underestimated variance at higher CPUE.

Inspection of the parameter correlation matrix (not presented here as it is large) does not suggest time parameters are unacceptably correlated with others in the model.

The predicted values for the index are calculated based on each year (2004-2015), the month at the start of each quarter, the "Centre" region and a fixed number of FADs based on the monthly average across regions for this time period (9.1 FADs).

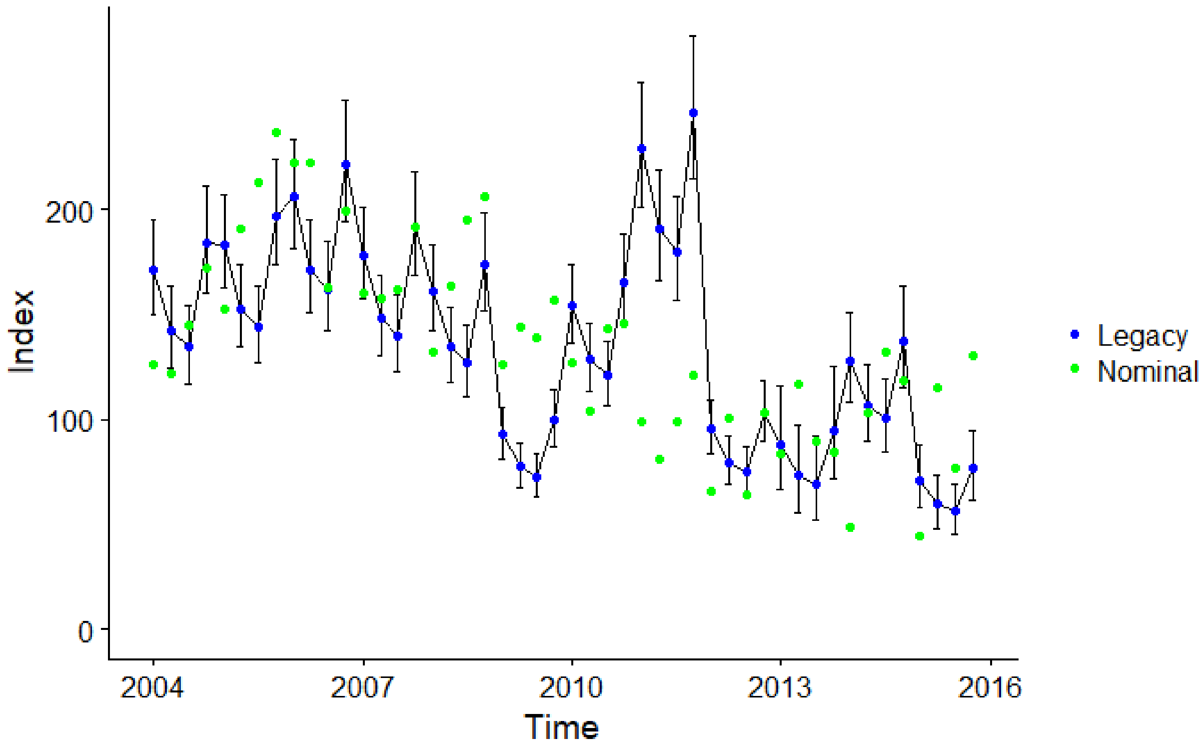

Figure 4 Maximum likelihood index with 95% confidence intervals based on the parameter covariance matrix (inverted Hessian) for the legacy model from Sharma, Geehan and Adam (2014) with more recent corrected data.

The updated index has similarities to the original legacy index. Both are primarily dependent on the main year effect and these have a similar scale to each other, although there are detailed differences in estimates. Changes in the fit are due to changes in the underlying data. Many problems have been removed (incorrect effort recording, duplicate records etc.).

Although the abundance index should have improved, issues remained with both estimates:

1. Models show a significant negative relationship with the quantity of FADs in the North and South regions. This is questionable. FADs show some correlation with the year effect parameters which is undesirable.

2. The intercept value for the previous index appears very low (average <1 fish per day for the base category), whereas the updated model has reasonable intercept value (>130 fish a day).

3. Both models suggest a much higher catch rate for the smallest sized vessels (<7m) than the next higher category, which otherwise show an increasing trend with vessel size. It is possible this is to do with mis-classification of vessels or other problems with data recording.

New Data

The next task was to add the new logbook data before further exploring alternative models. The new logbook system started in 2010, but the current and revised version was rolled out in 2014. It required that fishers complete a more detailed trip form based on individual days fishing. Note that requiring numbers of fish be reported resulted in most of the new logbook data being removed from the analysis. In future, catch data may be recorded as weight resulting in a significant change in the CPUE time series. This would require further review of the existing data to provide a smooth transition between numbers and weight.

The data table was manipulated so that it can be added to the monthly data with the same format. In addition, new explanatory variables were added. Because the response variable is on a log scale, some potential covariates were also converted to the same scale. Covariates on log and linear scales consisted of catch in numbers and weight of other species, vessel length and number of FADs. Other categories were the data source (old monthly reports and the new logbook data) and the quarter.

Model Structure

The model structure justification here is mostly provided for the catch numbers response variable only. The catch weight response variable model is only provided in those cases where a significant difference is possible. In general, the numbers and weight variables produce very similar models because, for the majority of the data, the catch weight is calculated from the catch numbers.

As part of the exploration of the model, alternative likelihood functions to the legacy model's log-normal were tried, consisting of the Gaussian, Poisson and Gamma. The log was the only reasonable GLM link function for these data.

Using the log-normal likelihood resulted in many observations with strong negative residuals, indicating that the records have catch rates much lower than expected. This can result from an inappropriate choice of likelihood which allows for greater dispersion than the data indicate. To explore this, a Box-Cox transform was applied to the response variable to see whether this suggested an alternative likelihood for a minimum model with only time series effects of year and quarter.

The optimum transform for both numbers and weight response variables (numbers: 0.35, weight: 0.31) were found greater than zero (zero being the equivalent to a log-transform used), but less than 0.5 (approximate Poisson equivalent). This suggested the Gamma as the best alternative likelihood. This was confirmed with other exploratory fits using the Gaussian and Poisson, which were unable to explain the over-dispersion and indicated worse diagnostics.

The Gamma likelihood and Box-Cox transforms produced similar results. Inspection of the Gamma model indicated that lower residuals were much better behaved, but the positive residuals were over-represented compared to what might be expected, a reversal to the log-normal. For both models, the diagnostics demonstrated poorer performance in different areas, so there was no clear preference between the Gamma and log-normal. There are theoretical reasons to prefer the Gamma however (McCullagh and Nelder 1989), so this was further explored with other model structures.

Using a Gamma model prevents calculating an *F*-statistic to measure the statistical significance of different model terms. This results in some subjectivity in selecting terms to retain in the standardization model. Use of statistics, such as *F,* to decide whether to include terms in GLMs with large data sets will tend to result in overfitting and including parameters with no clear justification. In general, the approach used here was to retain only terms which resulted in clear, high changes in deviance rather than rely on a measure of statistical significance. Overfitting could result in a poorer abundance index than the nominal CPUE.

The minimum model for the skipjack catch response variable consists of a log effort offset and the time series terms. Quarters (qu) are preferred to months (mn) because the stock assessment time step is quarters and this still accounts for any monsoon effects. Month does explain more of the variation, but this factor may also be confounded with other explanatory variables and was much less parsimonious (Table 1). It was likely that month terms in the legacy model caused some problems for the index due to aliasing with other parameters. As a result, the base model is the year (yr) and quarter (qu) main effects with interaction terms (yr+qu+yr:qr = yr*qu).

Table 1 Analysis of deviance for the numbers Gamma model with time categorical variables year, month and quarter (yr, mn, qu). ":" indicates an interaction term.

Terms	Residual		Terms	
	Degrees of Freedom	Deviance	Degrees of Freedom	Deviance
Yr	40998	72650		
yr + qu	40995	72321	3	328.82
yr + qu + yr:qu	40962	71652	33	668.92
yr + qu + yr:qu + mn	40954	71528	8	124.20
yr + mn + yr:mn	40866	70616	88	912.01

There is a small difference between the logbook and trip landing data, but it is probably not significant at least for the model of catch numbers (Table 2). The catch weight data are predominantly in the new log books and for this reason there is a larger difference. There is some overlap in the data collection scheme, so it is possible to apply a correction if required. However, overall it appears that the data sources are broadly compatible and an adjustment may not be required.

Table 2 Analysis of deviance for the Gamma model data source categorical variable (so) "trip landing record" or "logbook".

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
Numbers				
yr*qu	40962	71652		
yr*qu + so	40961	71650	1	2.46
Weight				
yr*qu	43064	80493		
yr*qu + so	43063	80354	1	138.43

Various other covariates are available: small yellowfin catch, large yellowfin catch, number of FADs in the region, other tuna catch, other non-tuna catch. It was checked whether they were better added on the linear or log scales. In some cases the Gamma model likelihood failed to converge, which is an indication of a poor fit. So for consistency in these exploratory fits, the log-normal likelihood was used.

Table 3 Analysis of deviance for the numbers lognormal model for various covariates on the log and linear scales.

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
yr*qu	40962	111680		
Small yellowfin (linear: yfs_n; log: lyfs)				
yr*qu + lyfs	40961	111607	1	73.00
yr*qu + yfs_n	40961	107622	0	3984.30
Large yellowfin (linear: yfl_n; log: lyfl)				
yr*qu + lyfl	40961	110019	1	1661.30
yr*qu + yfl_n	40961	111668	0	-1649.40
Number of FADs (linear: nFAD; log: lFAD)				
yr*qu + lFAD	40961	110446	1	1233.29
yr*qu + nFAD	40961	110032	0	414.23
Other tuna catch numbers (linear: otu_n; log: lotu)				
yr*qu + lotu	40961	109897	1	1782.50
yr*qu + otu_n	40961	111265	0	-1368.10
Other non-tuna catch numbers (linear: oth_n; log: loth)				
yr*qu + loth	40961	104269	1	7410.90
yr*qu + oth_n	40961	111301	0	-7031.80

These results suggested that "other tuna", "other fish" and "large yellowfin" are better covariates on a log-scale, whereas small yellowfin and number of FADs are better explanatory variables on a linear scale within the linear predictor (Table 3). Given the log link function for the model, log covariates would seem to be justified, but there is no strong theoretical reason to choose one over the other.

Table 4 Analysis of deviance for the lognormal model with step-wise addition of different covariates as main effects (see Table 3 for terms).

Terms	Residual		Terms	
	Degrees of Freedom	Deviance	Degrees of Freedom	Deviance
Numbers				
yr*qu	40962	111680		
yr*qu +loth	40961	104269	1	7410.90
yr*qu +loth +yfs_n	40960	98092	1	6177.10
yr*qu +loth +yfs_n +lyfl	40959	96611	1	1480.80
yr*qu +loth +yfs_n +lyfl +nFAD	40958	95608	1	1003.10
yr*qu +loth +yfs_n +lyfl +nFAD +lotu	40957	94757	1	851.30
Weight				
yr*qu	43064	122252		
yr*qu +loth	43063	113800	1	8452.2
yr*qu +loth +yfs_n	43062	107778	1	6021.8
yr*qu +loth +yfs_n +lyfl	43061	106818	1	960.5
yr*qu +loth +yfs_n +lyfl +lotu	43060	105873	1	944.8
yr*qu +loth +yfs_n +lyfl +lotu +nFAD	43059	105114	1	758.6

A stepwise introduction of main effects (Table 4) suggests that all covariates have a significant effect and are all candidates for inclusion. However, while some covariates might help indicate how much effort was directed at skipjack, it is not clear that the functional relation between the response and explanatory variables can be justified in this form. In addition, these variables may well be confounded with others (e.g. vessel length) and it may be inappropriate to use variables which themselves may be related to fishing effort and fluctuate in response to past fishing activity and other effects. This is mainly an issue for the skipjack catch weight model where small yellowfin catch is a significant effect.

The most important term is the "log other fish catch" (loth) which probably helps the CPUE of skipjack adjust for mixed gear use and targeting of other fish within trips. Numbers of FADs does explain some variation, but surprisingly it is relatively small. FADs may serve a greater purpose in reducing costs, fuel use for example, rather than raising catch rates, at least for skipjack, but this needs to be investigated further. Including yellowfin and other tuna catch as a covariate is unsafe.

The spatial effect is defined by the atoll and sets of atolls defined by region. Atolls explain a great deal more variation than region (Table 5). It is likely that regional groups for the atolls can be defined better than the three regions defined here and this should be examined in future.

Table 5 Analysis of deviance for the Gamma numbers model spatial effect of atoll or region (at, reg).

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
yr*qu	40962	71652		
yr*qu + reg	40960	70636	2	1015.70
yr*qu + at	40942	54886	18	15751.00

The vessel size is important as has been already determined. The vessel size class was defined as a category variable in Sharma *et al*. 2014. Alternatively its use was explored here as a covariate on the log and linear scale. The linear-scaled variable explained a large amount of the data variance for both the numbers and weight models (Table 6). The category variable did a little better, but at the expense of an additional 5 parameters, making it less parsimonious. The categorical variable did suggest that there were departures from a simple linear relationship, but there was insufficient information to provide a better functional relationship.

Table 6 Analysis of deviance of the lognormal model for vessel length as linear, log and by size class category (vl, lv, vc). Size classes are: <7m, 7-12m, 12-17m, 17-22m, 22-27m, 27-32, >32m.

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
Numbers				
yr*qu	40962	111680		
yr*qu + vl	40961	86847	1	24832.60
yr*qu + lv	40961	86827	0	20.30
yr*qu + vc	40956	85144	5	1683.20
Weight				
yr*qu	43064	122252		
yr*qu + vl	43063	94077	1	28175.60
yr*qu + lv	43063	94301	0	-224.50
yr*qu + vc	43058	92342	5	1959.30

Combining the main effect models suggested that the number of FADs is the weakest term (Table 7). Given FADs have increased over time and could be correlated with abundance, it was considered wise to remove this from the final model. This decision was supported, particularly for the catch numbers model where number of FADs was a very poor explanatory variable.

Table 7 Analysis of deviance for Gamma model with all main effects.

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
Numbers				
yr*qu	40962	71652		
yr*qu + vl	40961	62873	1	8779.10
yr*qu + vl + at	40941	50330	20	12542.80
yr*qu + vl + at + loth	40940	50184	1	146.10
yr*qu + vl + at + loth + nFAD	40939	50183	1	0.90
Weight				
yr*qu	43064	80493		
yr*qu + vl	43063	69179	1	11313.60
yr*qu + vl + at	43043	55915	20	13264.00
yr*qu + vl + at + loth	43042	55795	1	120.70
yr*qu + vl + at + loth + nFAD	43041	55759	1	35.90

The first level interaction term between vessel length and other species catch may be worth including, although as both values are covariates it is difficult to interpret (Table 8). Other interaction terms only explain a small amount of the data variation considering the number of parameters fitted. These are relatively unimportant parameters. The conclusion is that any joint effect of location and vessel size and other catch being landed is relatively small.

Table 8 Analysis of deviance for the Gamma numbers model with first level interaction effects.

Terms	Residual Degrees of Freedom	Residual Deviance	Terms Degrees of Freedom	Terms Deviance
yr*qu +vl +at +loth	40940	50184		
yr*qu +at +vl +loth +loth:vl	40939	49812	1	372.46
yr*qu +at +vl +loth +loth:vl +at:vl	40919	49103	20	709.23
yr*qu +at +vl +loth +loth:vl +at:vl +at:loth	40899	47577	20	1525.21

An alternative response variable would be the total catch of small yellowfin and skipjack combined instead of including yellowfin as an explanatory variable. It was thought combining these catches for the catch weight model might reduce the lowest residual outliers, but there is little apparent improvement over the skipjack model (Table 9).

Table 9 Analysis of deviance for lognormal numbers model with the response variable consisting of the total catch of skipjack and small yellowfin combined.

Terms	**Residual**		**Terms**	
	Degrees of Freedom	**Deviance**	**Degrees of Freedom**	**Deviance**
yr*qu	43064	102546		
yr*qu + at	43044	72560	20	29985.80
yr*qu + at + lv	43043	64687	1	7873.10
yr*qu + at + lv + loth	43042	64268	1	419.00
yr*qu + at + lv + loth + nFAD	43041	64266	1	1.60

Checking for an effort parameter (e.g. effort is sometimes thought to be non-linearly related to catch) suggests little is to be gained by allowing for a non-linear adjustment of effort (Table 10). The effort parameter (0.834) indicates there is some diminishing of returns on catch rates for longer trips. This may be an artefact since longer trips with more fishing days may be the direct result of lower catch rates.

Table 10 Analysis of deviance for the Gamma numbers model with and without (default) a non-linear parameter for fishing effort.

Terms	**Residual**		**Terms**	
	Degrees of Freedom	**Deviance**	**Degrees of Freedom**	**Deviance**
offset(lef) +yr*qu +vl +at +loth	40940	50184		
lef +yr*qu +at +lv +loth	40939	50520	1	-335.71

With no strong justification for more terms in the model, the final model for both catch numbers and catch weight has main effect terms for the quarter time series with atoll, vessel length and log "other catch".

Comparing the new proposed maximum likelihood index to the legacy index, the main difference is the removal of the increase in 2011 to create a smoother index. Other changes are relatively minor and both indices show a similar downward trend.

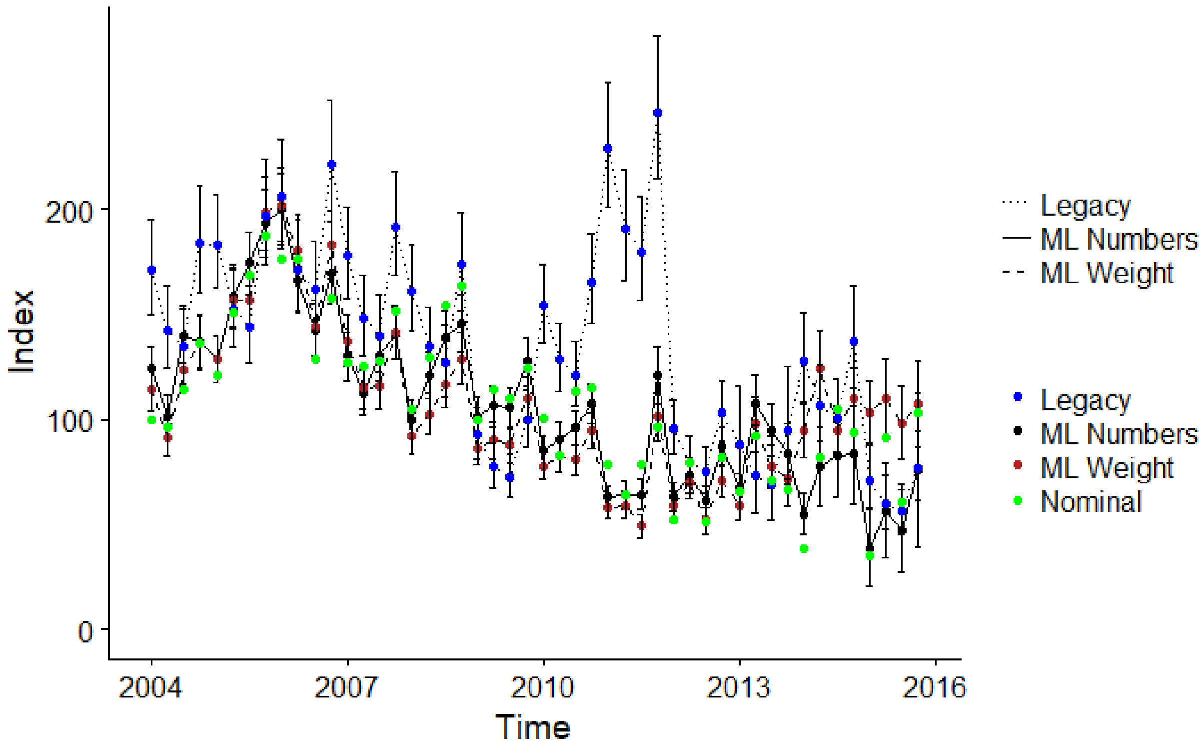

Figure 7 Maximum likelihood indices based on skipjack catch numbers and weight, and the legacy index estimated for the updated data set.

MCMC Fit

Stan software was used to fit the same model using a Bayesian MCMC approach. The fitted models used the same data and model structure as the maximum likelihood model (Gamma likelihood and same main effects). Priors were used, but were uninformative (uniform on the log scale) for all parameters, except the Gamma beta (scale) parameter which was given a weakly informative half Cauchy prior favouring lower variance for the observation error. Priors constrained parameters to a reasonable range and avoided unwarranted prior weight on nonsensical values, which also aids convergence. The original maximum likelihood linear models were used provide initial estimates for the MCMC simulations. However, the Stan Hamiltonian MCMC method should be robust to the simulations start point.

Four Markov chain simulations were run to obtain 5000 draws from the posterior. Non-informative priors were used. All priors were uniform densities except the Gamma scale parameter, which was weakly informative (half Cauchy density: see Stan 2016). Convergence was tested by comparing the chains. Stan provides extensive diagnostics to test convergence, the simplest being the $\hat{R}$ statistic, which indicates convergence for each parameter with values close to 1.0. The model estimates a standard index value for an 18m vessel operating from Malé atoll for each year and quarter (Table 11, 12). Results indicated that the MCMC had converged for these and the other parameters.

Table 11 Standardised index estimates from MCMC runs for the numbers model. Indices are on a log-scale. The MCMC simulation was run with 4 chains with 3500 iterations each, of which 1000 were used as "warm-up" and thinning of 2, resulting in 1250 draws per chain.

Year	Quarter	Mean	SD	Confidence Interval 2.5 – 97.5%		Effective Sample Size	$\hat{R}$
2004	1	4.958	0.028	4.904	5.012	3672	1.001
2004	2	4.717	0.031	4.659	4.779	3380	1.001
2004	3	5.017	0.028	4.963	5.072	3686	1.000
2004	4	5.105	0.029	5.049	5.161	3574	1.000
2005	1	4.979	0.027	4.927	5.032	3566	1.000
2005	2	5.023	0.028	4.968	5.078	3641	1.001
2005	3	5.181	0.025	5.132	5.230	3789	1.000
2005	4	5.213	0.026	5.162	5.264	3781	1.000
2006	1	5.243	0.026	5.191	5.295	3637	1.000
2006	2	5.136	0.027	5.083	5.188	3760	1.000
2006	3	4.945	0.028	4.890	4.999	3736	1.001
2006	4	5.022	0.027	4.968	5.076	3722	1.000
2007	1	4.858	0.028	4.804	4.914	3601	1.000
2007	2	4.787	0.030	4.729	4.844	3575	1.000
2007	3	4.819	0.029	4.763	4.875	3427	1.000
2007	4	4.956	0.027	4.902	5.009	3579	1.000
2008	1	4.715	0.030	4.654	4.774	2979	1.000
2008	2	4.769	0.029	4.712	4.825	3629	1.000
2008	3	4.927	0.027	4.875	4.980	3625	1.000
2008	4	5.001	0.029	4.944	5.057	3583	1.000
2009	1	4.732	0.029	4.673	4.789	3615	1.000
2009	2	4.697	0.030	4.637	4.754	3450	1.000
2009	3	4.765	0.030	4.708	4.826	3084	1.001
2009	4	4.872	0.029	4.814	4.928	3625	1.001
2010	1	5.017	0.025	4.967	5.065	3903	1.000
2010	2	4.720	0.030	4.662	4.778	3416	1.001
2010	3	4.893	0.029	4.837	4.948	3751	1.000
2010	4	4.915	0.029	4.855	4.970	3632	1.001
2011	1	4.753	0.031	4.693	4.814	3621	1.000
2011	2	4.526	0.036	4.455	4.597	3609	1.000
2011	3	4.517	0.038	4.440	4.590	3760	1.000
2011	4	4.651	0.037	4.579	4.722	3279	1.000
2012	1	4.585	0.036	4.514	4.656	3742	1.001
2012	2	4.592	0.038	4.517	4.666	4018	1.000
2012	3	4.393	0.042	4.311	4.474	3873	1.000
2012	4	4.710	0.036	4.639	4.779	3521	1.001
2013	1	4.800	0.037	4.728	4.871	3798	1.000

Year	Quarter	Mean	SD	Confidence Interval 2.5 – 97.5%		Effective Sample Size	$\hat{R}$
2013	2	4.847	0.044	4.760	4.930	4477	1.000
2013	3	4.757	0.044	4.668	4.839	4196	1.000
2013	4	4.731	0.061	4.609	4.846	4438	1.000
2014	1	4.535	0.065	4.406	4.662	4799	1.000
2014	2	4.685	0.083	4.522	4.840	4968	1.000
2014	3	4.775	0.083	4.611	4.932	4837	1.000
2014	4	4.628	0.090	4.444	4.799	4887	1.000
2015	1	4.262	0.169	3.900	4.577	4640	1.000
2015	2	4.343	0.127	4.074	4.575	4549	1.000
2015	3	4.235	0.143	3.936	4.501	5000	1.000
2015	4	4.486	0.146	4.192	4.756	4851	1.000

Table 12 Standardized index estimates from MCMC runs for the weight model. See Table 11 for explanation.

Year	Quarter	Mean	SD	Confidence Interval 2.5 – 97.5%		Effective Sample Size	$\hat{R}$
2004	1	-0.957	0.029	-1.013	-0.902	3699	1.000
2004	2	-1.209	0.032	-1.273	-1.148	3796	1.000
2004	3	-0.933	0.030	-0.991	-0.874	3715	1.000
2004	4	-0.788	0.029	-0.845	-0.732	3588	1.000
2005	1	-0.890	0.027	-0.944	-0.834	3426	1.001
2005	2	-0.855	0.028	-0.911	-0.800	3540	1.000
2005	3	-0.796	0.027	-0.848	-0.743	3694	1.000
2005	4	-0.670	0.026	-0.721	-0.620	3862	1.000
2006	1	-0.606	0.026	-0.658	-0.555	3793	1.000
2006	2	-0.672	0.027	-0.724	-0.620	3476	1.001
2006	3	-0.944	0.029	-0.999	-0.888	3777	1.001
2006	4	-0.844	0.027	-0.898	-0.793	3911	1.000
2007	1	-0.984	0.029	-1.042	-0.928	3612	1.000
2007	2	-1.111	0.030	-1.171	-1.054	3825	1.001
2007	3	-1.145	0.029	-1.200	-1.086	3295	1.000
2007	4	-0.930	0.028	-0.986	-0.876	3781	1.000
2008	1	-1.227	0.030	-1.287	-1.169	3728	0.999
2008	2	-1.206	0.030	-1.264	-1.151	3854	1.001
2008	3	-1.084	0.029	-1.141	-1.029	3695	1.000
2008	4	-0.972	0.029	-1.029	-0.917	3457	1.000
2009	1	-1.246	0.030	-1.304	-1.187	3536	1.000
2009	2	-1.287	0.031	-1.350	-1.227	3658	1.000

Year	Quarter	Mean	SD	Confidence Interval 2.5 – 97.5%		Effective Sample Size	$\hat{R}$
2009	3	-1.244	0.030	-1.302	-1.185	3645	1.000
2009	4	-1.112	0.029	-1.168	-1.053	3697	1.000
2010	1	-0.940	0.026	-0.993	-0.890	3991	1.000
2010	2	-1.228	0.030	-1.288	-1.170	3624	1.001
2010	3	-1.103	0.029	-1.161	-1.045	3569	1.001
2010	4	-1.054	0.030	-1.111	-0.996	3749	1.000
2011	1	-1.169	0.031	-1.230	-1.108	3662	1.001
2011	2	-1.411	0.037	-1.484	-1.338	3985	1.000
2011	3	-1.484	0.039	-1.560	-1.409	4028	1.000
2011	4	-1.338	0.038	-1.413	-1.263	3971	1.000
2012	1	-1.364	0.037	-1.437	-1.291	3828	1.000
2012	2	-1.383	0.039	-1.458	-1.306	3740	1.001
2012	3	-1.572	0.040	-1.651	-1.495	3954	1.000
2012	4	-1.263	0.035	-1.332	-1.193	3903	1.000
2013	1	-1.141	0.037	-1.216	-1.069	3737	1.000
2013	2	-1.140	0.044	-1.229	-1.056	4244	1.001
2013	3	-1.236	0.045	-1.324	-1.149	4436	1.000
2013	4	-1.216	0.061	-1.334	-1.095	4918	1.000
2014	1	-1.022	0.048	-1.118	-0.928	4313	1.000
2014	2	-0.840	0.049	-0.937	-0.744	4126	1.000
2014	3	-0.744	0.044	-0.831	-0.658	3963	1.000
2014	4	-0.715	0.048	-0.809	-0.622	4363	1.000
2015	1	-0.702	0.051	-0.803	-0.603	4250	1.000
2015	2	-1.012	0.057	-1.125	-0.901	4289	1.000
2015	3	-0.986	0.061	-1.107	-0.872	4497	0.999
2015	4	-0.801	0.064	-0.930	-0.679	4332	1.000

Index Comparison

The MCMC and maximum likelihood fit produces similar results, although the MCMC is flatter for both the weight and numbers indices (Figures 8 & 9). Standardised indices show a significant reduction of the noise in the nominal indices, as might be expected, but follow the same trends.

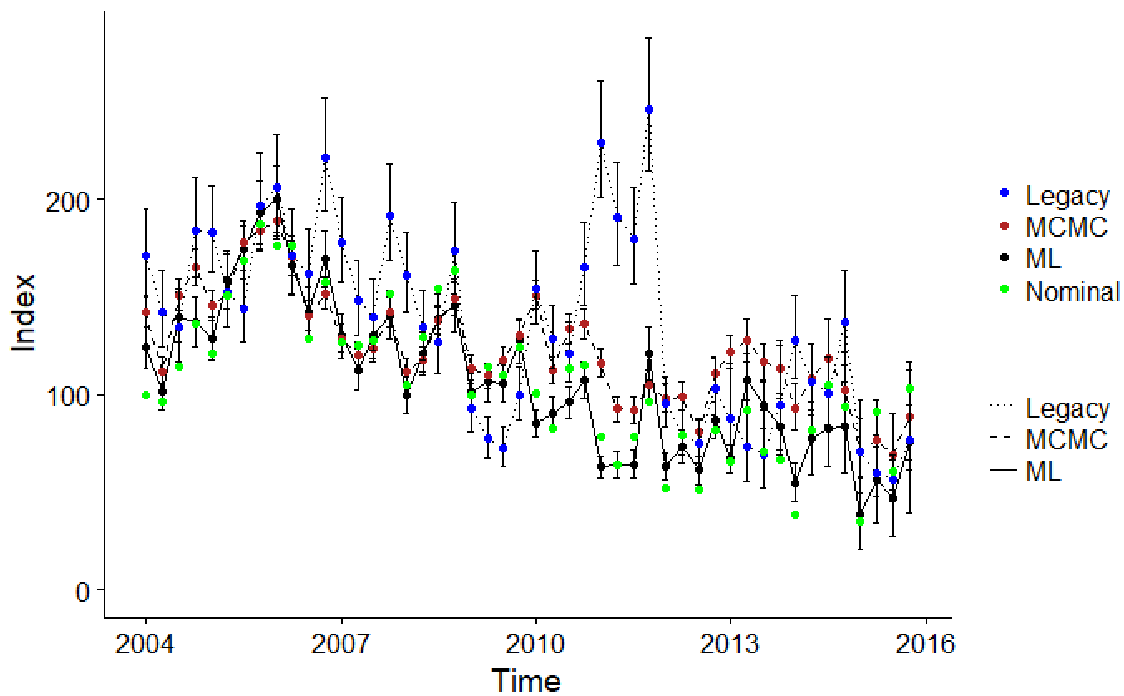

Figure 8 Alternative abundance indices based on numbers for the legacy model, and the new model fitted using maximum likelihood (ML) and Bayesian Markov Chain Monte Carlo (MCMC).

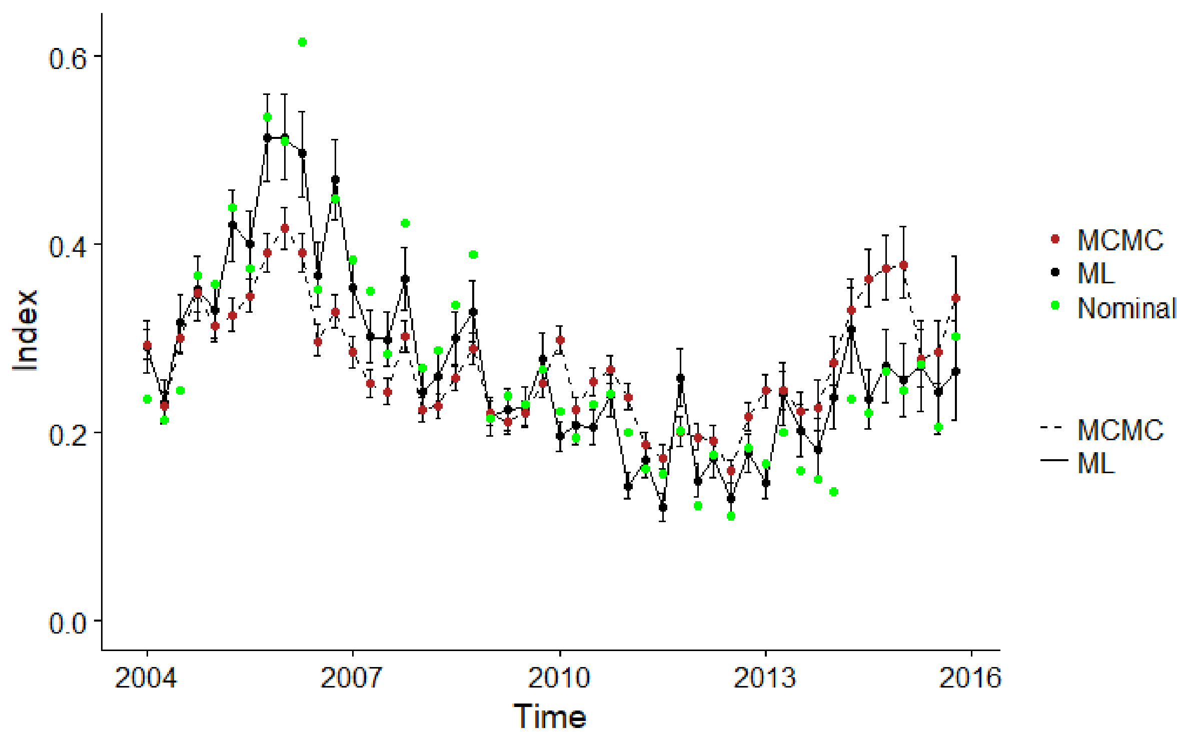

Figure 9 Alternative abundance indices with 95% confidence interval based on weight for the new model fitted using maximum likelihood (ML) and Bayesian Markov Chain Monte Carlo (MCMC).

Concluding Remarks

The philosophy behind the approach to this standardization has been to be as parsimonious as possible. Parameters were only included in the model where it appeared justified. The primary objective of standardization is to remove effects on the trend of the index that are not related to abundance. A secondary aim is to reduce the variance of the indices so that they measure the mean index more precisely. The models here attempt to do this without risking degradation of the indices which can result from overfitting.

Stan provides the marginal estimates for indices (mean catch rates), integrating over other parameters. The mean of the MCMC indices therefore accounts for the spread in the posterior probability density, unlike the maximum likelihood estimates which are located at the likelihood mode. Given the posterior uses non-informative priors, the posterior and likelihood should be very similar. Difference between the MCMC and ML estimates can be attributed to the consideration of the entire probability density rather than its mode. This should result in a smoother more reliable index.

The model based on weight data departs most from the numbers model in the most recent years. The Maldives is in the process of changing the reported data from numbers to weight, so it might be expected errors during this period to increase. Although the numbers data apply to most of the time series, removal of records that did not report numbers left the recent years less well estimated. As more data become available, transfer from numbers to weight should become more assured, although the standardisation may need to account for this change.

While the model estimates the errors for the indices, this only represents the estimated observation error for the mean catch rate. The error associated with the difference between the relative stock size and the index is not estimated and would need to be included in the stock assessment model.

Further analysis of more detailed information available from log-books may indicate how the standardisation model might be improved. One way in which the model could improve is to make greater use of random effects modelling, which can help solve a range of issues. Random effects could be useful for partition CPUE variance among various sources. For the current CPUE this could serve two purposes.

1. Catchability for individual vessels could be estimated as a random effect based on a mean fishing power determined by vessel characteristics (length primarily). Given the number of vessels and various data problems, this may be difficult to achieve using maximum likelihood.

2. Interaction terms between the time series variables and other factors, notably factors based on area. In this latter case, it might be hypothesized that catch rates vary dependent on fish location rather than overall abundance, so high catch rates in any area may imply lower catch rates in other areas. Random effects can be used to model this sort of factor. This might be most useful when joining indices that have some spatial overlap. The lack of spatial information make this approach probably unnecessary at this stage.

References

Ahusan, M., Sharma, R., Adam, M.S. 2015. Cleaning the Maldives Catch and Effort Dataset (2004-2009). IOTC-2015-WPNT-INF03.

Cook, J. (1995). CPUE and conversion factors. Economic Planning and Coordination Section, Ministry of Fisheries and Agriculture, Malé. Maldives, unpublished report: 6 pages

Fournier, D.A., Skaug, H.J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M.N., Nielsen, A., and Sibert, J. 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. 27:233-249.

IOTC 2016: Skipjack Tuna Supporting Information. Status summary for species of tuna and tuna-like species under the IOTC mandate, as well as other species impacted by IOTC fisheries. http://www.iotc.org/science/status-summary-species-tuna-and-tuna-species-under-iotc-mandate-well-other-species-impacted-iotc. Retrieved on: 24 September 2017.

IOTC-WPTT11 2009[E]. Report of the Eleventh Session of the Working Party on Tropical Tunas. Mombasa, Kenya.

Kolody, D., Adam, M.S., Anderson, C. 2010. Catch Rate Standardization for the Maldivian Skipjack Pole and Line Fishery 1970-2007. IOTC-2010-WPTT-05.

Kolody, D., Adam., M.S. 2011. Maldives Skipjack Pole and Line Fishery Catch Rate Standardization 2004-2010. IOTC-2011-WPTT12-29.

Lunn, D., Spiegelhalter, D., Thomas, A., Best, N. 2009. The BUGS project: Evolution, critique and future directions. Statistics in Medicine. 28 (25): 3049–3067.

McCullagh, P. and Nelder, J.A. 1989. Generalized linear models, Second Edition. Chapman and Hall, New York.

Methot, R.D. 2015. User manual for Stock Synthesis, model version 3.24s. Updated February 11, 2015.

Sharma, R., Geehan, J., Adam, M.S. 2013. Maldives Skipjack Pole and Line Fishery Catch Rate Standardization 2004-2011: Reconstructing Historic CPUE till 1985. IOTC-2013-WPTT15-32.

Sharma, R., Geehan, J., Adam, M.S. 2014. Maldives Skipjack Pole and Line Fishery Catch Rate Standardization 2004-2012: Reconstructing Historic CPUE till 1985. IOTC-2014-WPTT16-42.

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