Scientific Committee
Fifteenth Regular Session

Pohnpei, Federated States of Micronesia
12-20 August 2019

Standardized CPUE for skipjack tuna *Katsuwonus pelamis* from the Papua New Guinea archipelagic purse seine fishery

---

WCPFC-SC15-2019/SA-IP-05

Tiffany Vidal [1], Graham Pilling [1], Laura Tremblay-Boyer [2], and Thomas Usu [3]

[1]Oceanic Fisheries Programme (OFP), The Pacific Community, Nouméa, New Caledonia
[2]Dragonfly Data Science, Wellington, New Zealand
[3]National Fisheries Authority, Port Moresby, Papua New Guinea

Executive Summary

Following the decline in the pole and line fishery, the purse seine fishery operating within Papua New Guinea (PNG) archipelagic waters has been used to provide an index of abundance for the Papua New Guinea/Solomon Islands region of the skipjack assessment model since 2014. The purse seine fishery within PNG archipelagic waters is well established, but the fishing strategies have changed dramatically through time. During the 1990s and early 2000s, the fishery was largely focused around fish aggregation devices (FADs). Gradually, fishing effort has shifted away from FADs and more towards unassociated sets, and over the past five years the fleet has predominantly targeted free-schooling aggregations. An important challenge when standardizing catch per unit effort (CPUE) is accounting for differences in fishing strategies, and for that reason we chose to use CPUE data from associated sets only, as there has been a consistent time series of fishing activity focused on floating objects since 1997. This document describes the generalized linear modeling approach used to standardize CPUE from the PNG archipelagic purse seine fleet for the period 1997-2018. The standardized index suggests the fishery has been relatively stable for the past decade with increasing variability in the most recent time periods. We have highlighted uncertainties associated with the index and have provided recommendations for future improvements.

Introduction

In the western and central Pacific Ocean (WCPO) the skipjack tuna *Katsuwonus pelamis* fishery is largely a purse seine fishery, with about 79% of the skipjack harvest in 2017 landed by the purse seine fleet (Williams and Reid, 2018). Purse seiners concentrate fishing effort on aggregations of fish, and for that reason standardized purse seine catch per unit effort (CPUE) metrics may be hyperstable and relatively insensitive to changes in stock abundance over time (Hoyle et al., 2014). In addition, the continual incorporation of new technologies and strategies by the fishing fleet in an effort to increase efficiency and maximize profit margins suggests that the traditional unit of effort, the set, has not remained constant through time (i.e. effort creep; Pilling et al., 2016). Despite these challenges, purse seine fishing is a major component of the skipjack fishery in the WCPO, and the catch and effort data associated with this fishery likely contain valuable information regarding changes in abundance through time. The standardization approach described below has been developed to account for factors that may have influenced CPUE, for the Papua New Guinea/Solomon Islands region of the skipjack assessment model, but which are unrelated to changes in abundance.

In 2014, a standardized index of abundance from the Papua New Guinea (PNG) archipelagic purse seine fishery was developed (Pilling et al., 2014), and deemed appropriate for use in the assessment given the relative stability of the fleet and anchored-FAD (aFAD) fishing strategy through time (Sokimi, 2009). By 2015, the fishery had shifted from fishing almost exclusively associated sets to about 40% of the overall purse seine effort being performed on floating objects.

In this document, associated sets are defined as those that target schooling aggregations of fish associated with floating objects, whereas unassociated sets target free-schooling fish aggregations. Floating objects, in this context, include man-made anchored and drifting FADs (e.g., buoys or rafts), as well as natural floating objects, such as logs, whales, and whale sharks, around which fish may aggregate. The trend towards unassociated sets has continued into 2018, with approximately 15-25% of the sets made between 2016 and 2018 in PNG archipelagic waters performed on floating objects, with the remaining sets targeting free-schooling aggregations. The shift in fishing strategies employed through time, combined with continual improvements in vessel and fishing technologies, are expected to impact abundance trends derived from CPUE. To address these changes, we modeled catch rates from associated sets made by a core fleet of vessels with a consistent history of fishing activity in the region.

The objective of this analysis is to provide a standardized index of skipjack abundance from purse seine CPUE for the Papua New Guinea/Solomon Islands region of the 2019 skipjack stock assessment model. We used a Tweedie generalized linear modeling approach to model skipjack catch rates through time from purse seine sets associated with floating objects. The associated fishery has persisted since 1997 and represents a consistent time series of relatively stable fishing practices, and was therefore deemed appropriate for the development of a standardized index of skipjack abundance.

Methods

Preparation of the dataset

The time series of operational (logsheet) data available from purse seine vessels fishing within PNG archipelagic waters (i.e. the Bismarck Sea) between 1997 and 2018, were examined (Figure 1). Papua New Guinea (PNG) and Philippines (PH) flagged vessels had the most consistent time series, and therefore were used in the CPUE analysis. In previous analyses, PNG flagged vessels were used exclusively, but there was evidence of similar fishing patterns between PNG and PH vessels as well as flag switching between these nations, justifying retention of the PH vessels for this analysis. Data, including the set type and catch (in mt), were available from 132 vessels of varying time series lengths. Some vessels had an intermittent time series, or one of relatively short duration, due to fleet evolution (e.g., decommissioning of older vessels and addition of newer vessels). Logsheet records with an associated set type of anchored FAD (aFAD), drifting FAD (dFAD), or other associated (e.g., drifting logs, whales, whale sharks; ASSOT) were retained.

During the initial standardization of purse seine CPUE for this region (Pilling et al., 2014), only associated sets (aFAD, dFAD, or ASSOT) were used because the predominant fishing strategy of the fleet was associated with floating objects. However, through time, the fishing strategy in the region has largely shifted towards unassociated sets, and in 2016, the standardization analysis incorporated unassociated sets as well (Tremblay-Boyer et al., 2016). This issue was again revisited in 2019, and we have elected to include only associated sets in this analysis, as changes in fishing strategy may confound changes in abundance, even when accounting for the set type in the standardization model.

We further subset the data to remove extreme catch observations, determined as being outside the 99.9th quantile of observed catch by set (> 217 mt). Skunk sets, defined as sets with less than one mt of total tuna catch (approximately 3.5% of the logsheet sets), were omitted from the analysis as they were assumed to represent a failed set and were not expected to be representative of abundance. We did however, retain sets with total tuna catch greater than one mt even if the total amount of skipjack was relatively as low (e.g., primarily a yellowfin set), as these sets were considered informative for the estimation of trends in abundance. We identified a 'core fleet' of vessels from the filtered data set, deemed representative of the fishery over time, for the analysis. The selection criteria for the core fleet is described below.

Core fleet determination

A 'core' group of fishing vessels was selected, based on a combination of fishing history and catch rates, and was assumed to be representative of the purse seine fleet, fishing associated sets, through time. The core fleet consisted of 33 vessels, selected from the initial 132 vessels in the dataset, which provided a reasonable time series of information for modeling catch rates. Using criteria similar to that implemented in 2016 (Tremblay-Boyer et al., 2016), vessels were retained only if they had been active for at least 18 quarters ($\sim$ 20% of the time series; with at least one set in each quarter), and had at least one catch in the top 95% of catches for a given year. Data for a given quarter were retained for each vessel only if they performed at least five sets during that three month period. In the previous analysis, a criterion of six active quarters was used; we increased the selection criterion to 18 quarters in an attempt to retain a fleet of vessels that had a reasonably consistent history of fishing activity in the region, throughout the time period of interest. The resulting dataset represented 44% of the effort data and of the 46% catch data, for the time series.

Clustering

A k-means clustering algorithm was used to create an additional explanatory variable to describe species composition, at the set level. Clustering allows for a reduction in the variability among individual sets by classifying sets into a pre-defined number of groups, based on the species composition of the catch. Because the purse seine fleet in PNG waters may switch between targeting and/or catching schools that are predominantly of one species or another (i.e., skipjack or yellowfin), we wanted to account for the variability in catch rates associated with the dominant species harvested in a given set. The dominant species in the catch composition from a given set, based on the proportion of each tuna species (in mt), was used to determine the cluster association. Cluster groupings of two, three, and four were evaluated, with two groups being preferred — a skipjack cluster and a yellowfin tuna cluster.

Tweedie GLM

We used a Tweedie generalized linear modeling (GLM) approach to model skipjack CPUE (mt/set) as a function of covariates, due to potential pitfalls associated with delta-lognormal (a.k.a hurdle) models, described in detail by Thorson (2017). Delta-lognormal models are commonly used for fisheries CPUE standardization (Maunder and Punt, 2004; Lynch et al., 2012) by modeling the probability of non-zero catch, assuming a binomial error structure, and the positive CPUE data with a log-normal distribution, assuming statistical independence between the two model components. This assumption of independence may be unrealistic for ecological systems because regions of high density are likely to have high encounter probabilities (Royle and Nichols, 2003). Similarly, abundant species may be distributed more widely, increasing encounter probability throughout their range while less abundant species may exhibit patchier distributions and overall lower catch rates.

The Tweedie GLM is suitable for non-negative continuous data with a high density of observations at zero, and addresses the concerns described above in a unified framework by modeling skipjack CPUE from set i (C_i) as,

$$C_i \sim Tw_p(\mu_i, \phi)$$

where

$$\mu_i = exp(\eta_i)$$

$$\eta_i = \beta_0 + \beta_x X$$

and η_i is the linear predictor for the observed skipjack CPUE, at the set level, using the log-link function. Here, β_x is a vector of regression coefficients and X is a matrix of predictor variables. The variance of C_i is assumed to be a function of the estimated dispersion ϕ and power p parameters (see Jorgensen, 1997 and Bonat and Kokonendji, 2017 for details).

$$V(C_i) = \phi\mu_i^p$$

We fit the GLMs using the R package *glmmTMB* (Magnusson et al., 2016).

A suite of candidate models was developed based on expert knowledge of fishery dynamics through time. We evaluated four predictor variables: year-quarter, vessel identifier, set type (aFAD, dFAD, ASSOT), and cluster grouping; all predictor variables were treated as factors. We included one interaction term in the suite of candidate models, the interaction between set type and cluster. The preferred model was selected through an evaluation process that considered the Akaike Information Criterion (Akaike, 1973), residual patterns over space and time, and contribution of covariates to final indices, based on step and influence plots. These criteria were applied in light of retaining as simple a model structure as possible to facilitate the interpretation of year-quarter effects.

Results

The nominal CPUE trends for skipjack and yellowfin tuna were generally positive throughout the most recent years in the time series (Figure 2). There has been a marked shift in fishing strategy from associated sets to unassociated sets and a transition in the vessels that comprise the PNG purse seine fleet. The core fleet used in this analysis included 33 vessels, a decrease from the 37 vessels used in 2016 (Tremblay-Boyer et al., 2016) but an increase from the 13 used in 2014 (Pilling et al., 2014; Figure 3). In recent years, the core vessels have fished primarily on ASSOT sets (Figure 4), although the species composition by set type has remained relatively stable through time (Figure 5). The core fleet was absent from the most recent three quarters of the time series, a period during which the third largest median nominal CPUE value was observed (preceded by two observations from 1998).

We utilized two clusters for the model cluster variable associated with set-level species composition. The two clusters were defined as sets predominantly harvesting skipjack or yellowfin (Figure 6), and explained 79% of the variation in species composition. The proportion of sets harvesting primarily yellowfin has generally increased through time.

The final GLM included year-quarter, cluster, and vessel identifier as factors; no interaction terms were included. The resulting standardized CPUE skipjack index has been reasonably stable through time, with increasing variability in recent years (Figure 7). The step plots illustrate the relative change in the standardized index with the addition of covariates, while the influence plots depict the influence a covariate has on the index, given the change in its distribution through time (Bentley et al., 2012). For example, if a covariate in a given time step has a large positive influence value, it suggests that the standardized index would be higher if that covariate were omitted. Cluster has been an influential variable through time, with a negative trend, such that in recent years the CPUE would be lower without the species composition variable included in the model (Figure 7). There is a positive trend in the influence of the vessel effect on the model through time, a potential indication of effort creep, with newer vessels demonstrating higher rates of skipjack catch in the more recent years as compared to vessel sets from the earlier part of the time series.

The dispersion and power parameters were estimated to be 2.94 and 1.54, respectively. Model diagnostics for the preferred model indicated approximately normally distributed quantile residuals, but with some consistently large, positive residuals throughout the time series (Figure 8). The spatial residuals do not display any concerning clustering in space.

Discussion

The purse seine fishery in PNG archipelagic waters has demonstrated dynamic and adaptive fishing behaviors throughout this time series. The standardized index of abundance for skipjack tuna from the purse seine fleet fishing in PNG archipelagic waters (Figure 7; bottom-left panel) has been relatively stable through time, although with greater fluctuations in recent years. The PNG flagged fleet was initially selected for standardization due to a relatively consistent time series of purse seine effort focused on aFADs (Pilling et al., 2014); however, the fishery has been gradually moving away from aFAD focused effort and increasingly targeting free schooling skipjack unassociated with floating structures. This pattern continued through 2018 with the proportion of effort and harvest coming from unassociated sets increasing over time (Figures 3 and 4). The vessels participating in the fishery have also evolved throughout the time series as the more historical vessels are departing the fishery and newer vessels are entering. We have used a conservative approach by modeling only associated sets and by extending the time period requirement of active fishing effort, for inclusion in the core fleet, to 18 quarters. These decisions were intended to minimize the impact newer vessels might have on the standardized index through time, but this approach is not sustainable in the long-term as vessel replacement and upgrades are expected to continue as long as the fishery is profitable, and fewer vessels are fishing schools associated with floating objects. In the interest of maintaining a stable fleet for the analysis, we have been unable to provide a standardized CPUE index for the three most recent time periods, due to inactivity by the core fleet. In the future, standardizing purse seine CPUE in the Papua New Guinea/Solomon Islands region of the skipjack assessment model will likely require an analysis of unassociated catch rates. Further research into the appropriate effort metrics for free-school sets and standardization approaches is recommended to produce an index capable of reliably estimating changes in abundance over time.

Changes in species composition (as represented by the cluster variable) have been influential through time, as the number of sets harvesting primarily yellowfin as has been slowly increasing. The cluster variable was included in the model to control for variability associated with targeted/harvested species, as catch characteristics of skipjack are expected to change as a result. Set type (aFAD, dFAD, ASSOT) had very little influence on the overall index, and was therefore dropped from the model. Here, we have modeled associated set catch rates from logsheet data only; however, there remain important questions about the changes in fishing behaviors over time. For example, has the shift towards unassociated sets been related to market drivers, management scheme, or abundance of skipjack versus yellowfin? Although the dominant species in a set has been used to explain some of the variation in catch rates, it should be noted that most sets are not exclusively of one species or another and most often the harvests are a mixture of skipjack and yellowfin. These changes in fishing strategies over a relatively short time period is notable, and worthy of further investigation. Additional information regarding FAD densities and tuna behavior around FADs is an important research area related to purse seine CPUE analyses.

The influence of the vessel effect shows a positive trend through time, and may be suggestive of changes in efficiency, an aspect of the analysis that requires additional attention. A reliable database documenting vessel and gear characteristics (e.g., vessel length, engine power, net length, number of support vessels) as well as technologies employed by the skipper (e.g., echo-sounder equipped FADs, remote sensing software) is of great importance for purse seine fisheries more generally to control for changes in effective effort through time (e.g., Torres-Irineo et al., 2014). In addition, in the wider fishery it may prove valuable to explore observer and vessel monitoring system (VMS) data to investigate alternative approaches for quantifying effort beyond the set. These data elements may help to mitigate the impact of effort creep (Pilling et al., 2016) on standardized CPUE indices. Effort creep has the potential to mask true trends in abundance causing hyperstability of the index by assuming effective effort has remained constant when in fact, less effort is required to find, catch, and process the same amount of fish through time due to technological innovation. In addition, a vessel attributes database may enable expansion of or elimination of the core fleet. It may be possible to include a mixture of newer and older vessels, assuming we can statistically control for the differences in efficiency amongst them. The proposed database should help to quantify these important changes through time.

We acknowledge the uncertainties regarding the use of this purse seine CPUE index of abundance for Pacific skipjack tuna, and recommend continued research into standardization approaches (see recommendations in Hoyle et al., 2014). As pole and line indices are no longer reliable for the fishery in PNG archipelagic waters, there is a salient need to improve upon the datasets describing the characteristics of this purse seine fishery and technologies fishers are using to prosecute it.

Acknowledgments

We would like to thank the Pacific-European Union Marine Partnership (PEUMP) programme for funding TV's position, which facilitated the analysis.

References

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, 60(2):255–265.

Bentley, N., Kendrick, T. H., Starr, P. J., and Breen, P. A. (2012). Influence plots and metrics: tools for better understanding fisheries catch-per-unit-effort standardizations. *ICES Journal of Marine Science*, 69(1):84–88.

Bonat, W. H. and Kokonendji, C. C. (2017). Flexible Tweedie regression models for continuous data. *Journal of Statistical Computation and Simulation*, 87(11):2138–2152.

Hoyle, S. D., Langley, A. D., and Campbell, R. A. (2014). Recommended approaches for standardizing CPUE data from pelagic fisheries. *Western and Central Pacific Fisheries Commission 10th Regular Session*, WCPFC-SC10-2014/SA-IP-10.

Jorgensen, B. (1997). *The theory of dispersion models*. CRC Press.

Lynch, P. D., Shertzer, K. W., and Latour, R. J. (2012). An evaluation of methods for standardizing catch rates of highly migratory by-catch species. *Collect. Vol. Sci. Pap. ICCAT*, 68(4):1498–1509.

Magnusson, A., Skaug, H., Nielsen, A., Berg, C., Kristensen, K., Maechler, M., van Bentham, K., Bolker, B., and Brooks, M. (2016). glmmTMB: Generalized Linear Mixed Models using Template Model Builder. *R package version 0.0*, 2.

Maunder, M. N. and Punt, A. E. (2004). Standardizing catch and effort data: a review of recent approaches. *Fisheries Research*, 70(2):141–159.

Pilling, G., Tidd, A., the PNA Office, Norris, W., and Hampton, J. (2016). Examining indicators of effort creep in the WCPO purse seine fishery. *Western and Central Pacific Fisheries Commission 12th Regular Session*, WCPFC-SC12-2016/MI-WP-08.

Pilling, G., Usu, T., Kumasi, B., Harley, S., and Hampton, J. (2014). Purse seine CPUE for skipjack and yellowfin in the PNG purse seine fishery WCPFC SC10-SA-WP-03. *Majuro, Republic of the Marshall Islands*, pages 6–14.

Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence–absence data or point counts. *Ecology*, 84(3):777–790.

Sokimi, W. (2009). Purse seine fishing around moored FADs in Papua New Guinea. *SPC Fisheries newsletter*, 129.

Thorson, J. T. (2017). Three problems with the conventional delta-model for biomass sampling data, and a computationally efficient alternative. *Canadian Journal of Fisheries and Aquatic Sciences*, 75(9):1369–1382.

Torres-Irineo, E., Gaertner, D., Chassot, E., and Dreyfus-León, M. (2014). Changes in fishing power and fishing strategies driven by new technologies: The case of tropical tuna purse seiners in the eastern Atlantic Ocean. *Fisheries Research*, 155:10–19.

Tremblay-Boyer, L., Pilling, G., Kumasi, B., and Usu, T. (2016). Standardized CPUE for skipjack tuna (*Katsuwonus pelamis*) from the Papua New Guinea archipelagic purse seine fishery. In *Western and Central Pacific Fisheries Commission 12th Regular Session.*

Williams, P. and Reid, C. (2018). Overview of Tuna Fisheries in the Western and Central Pacific Ocean, including Economic Conditions – 2017. *Western and Central Pacific Fisheries Commission 14th Regular Session.*

Figures

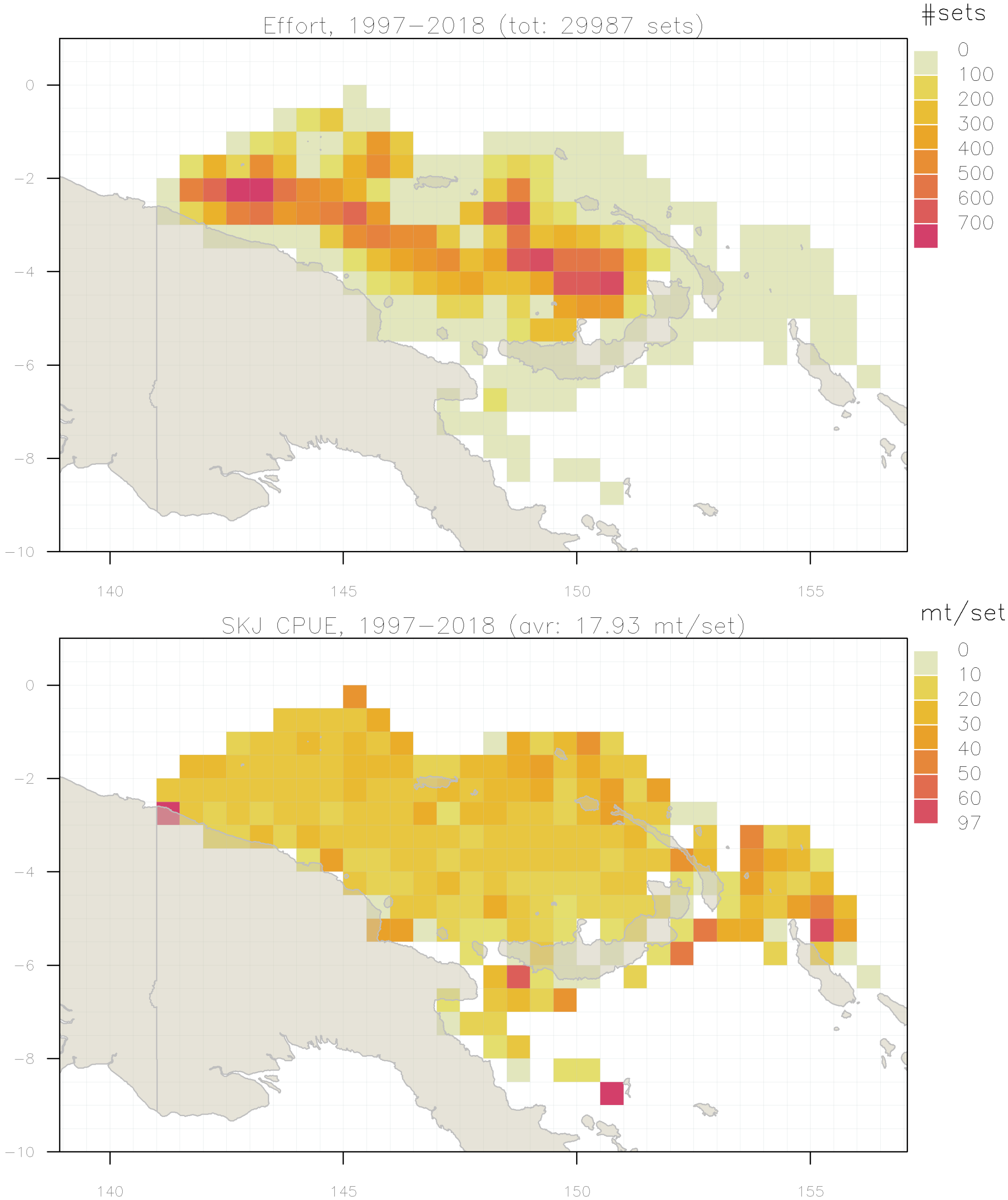

Figure 1: Aggregated effort (number of sets) and skipjack CPUE (mean CPUE, mt/set) by 0.5 degree cell, in PNG archipelagic waters, from 1997 to 2018.

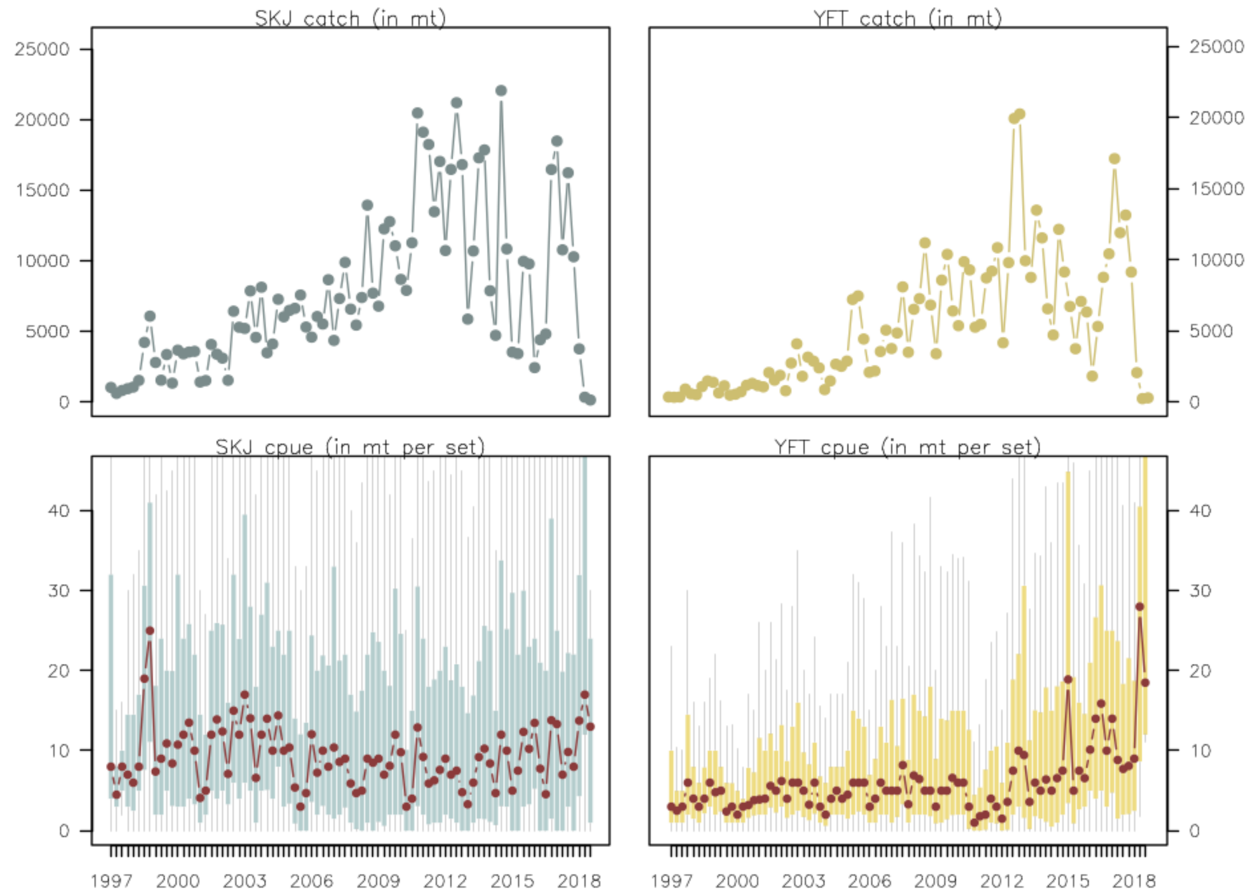

Figure 2: Summary of quarterly catch (top) and CPUE (bottom; nominal, mt/set) for skipjack and yellowfin tuna from 1997 to 2018. The boxplots highlight the median in red with the colored boxes covering the 25th ot 75th quartiles of CPUE observed for that quarter.

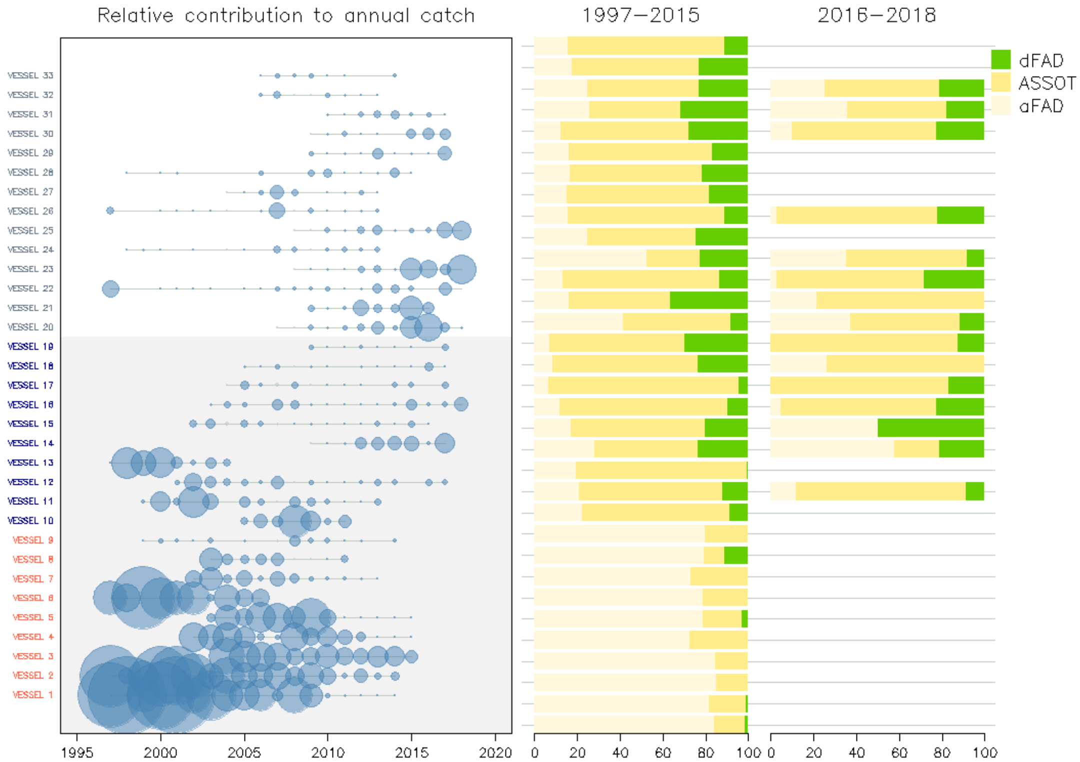

Figure 3: Key vessel statistics for the core fleet: relative contribution to the annual catch over time (left) and distribution of set type over the 1997-2015 and 2016-2018 periods. The period split was chosen to highlight the new data incorporated since the last CPUE standardization in 2016. The vessels are split between those that belonged to the core fleet for previous standardizations (gray background; red text for 2013 core vessels, blue text for additional 2016 core vessels, and gray text for vessels new to the analysis in 2019). Within each category the vessels are ranked by total catch 1997-2018, with greatest catch at the bottom; bubble size is scaled by the vessel contribution to annual catch. The figure on the right shows all sets made by the vessels selected to be in the core fleet. The core fleet contributes much less to the overall proportion of total skipjack in recent years, than it did in the original analysis.

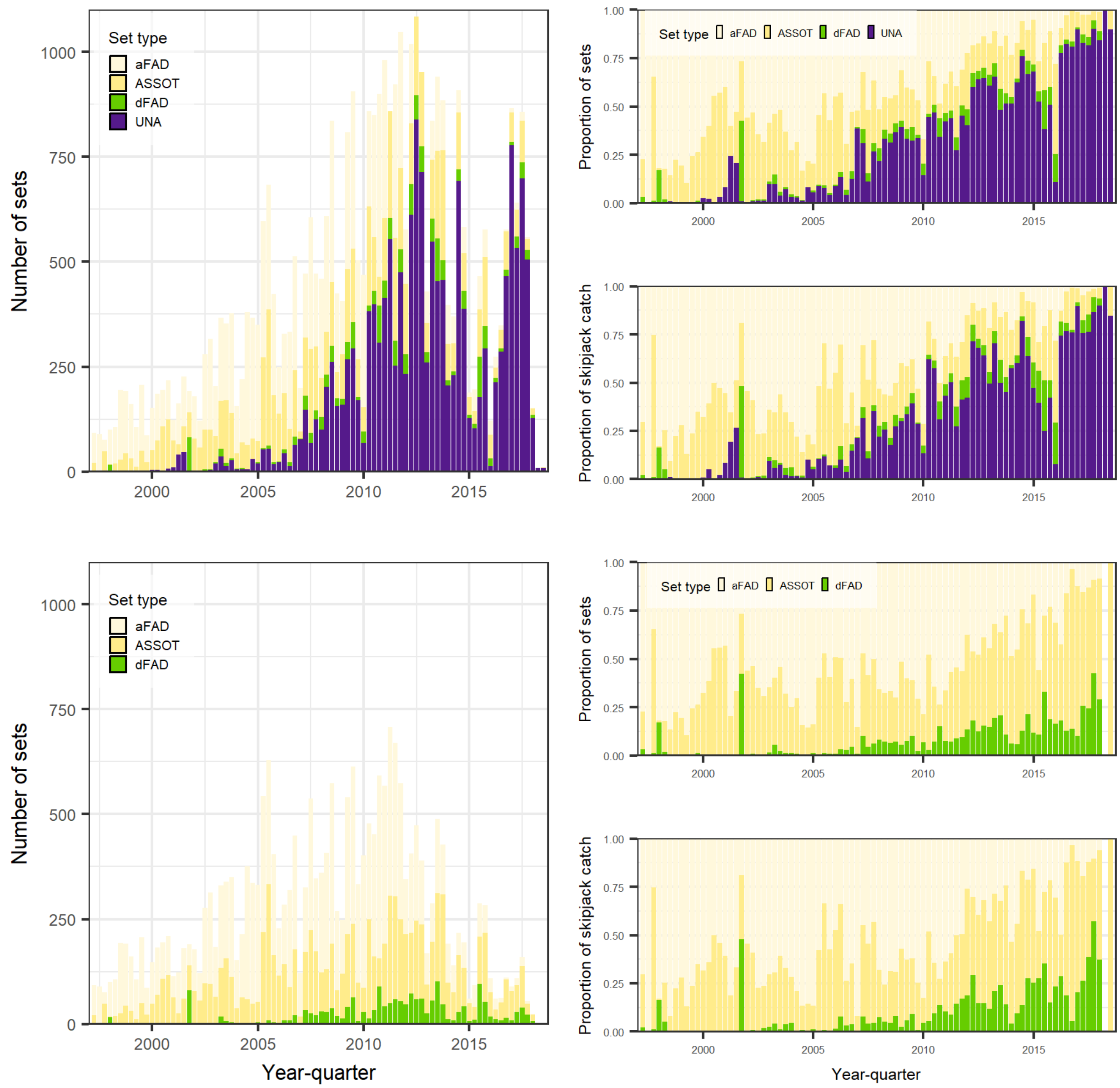

Figure 4: Number of quarterly purse seine sets alongside the proportion of sets and skipjack catch, by set type, over the time series. The top figures illustrate the number and proportion of sets and proportion of skipjack catch over the time series from the unfiltered data set. The bottom figures show the number and proportion of sets and proportion of skipjack catch, by set type, for the filtered data used in this analysis (i.e., sets associated with floating objects).

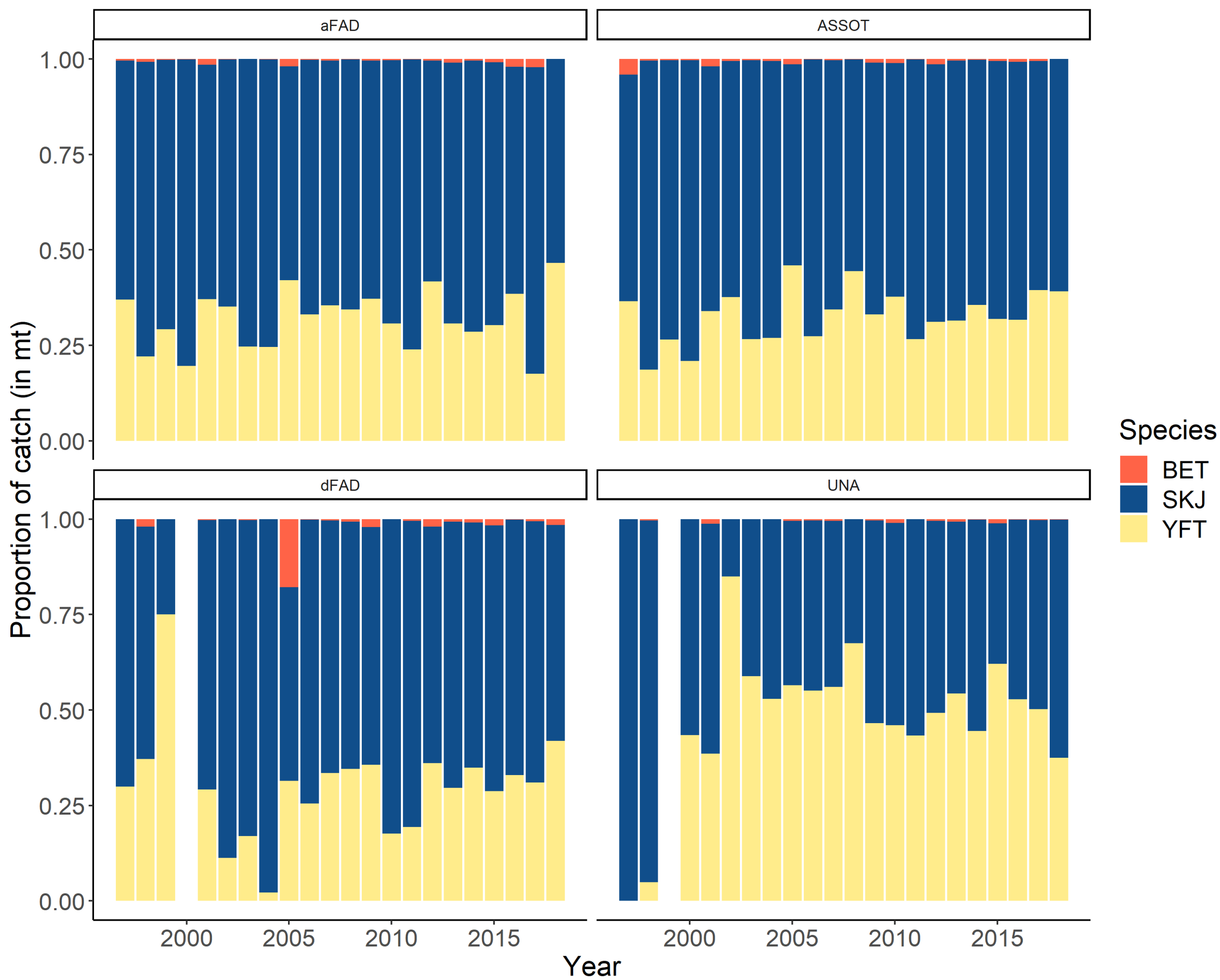

Figure 5: Species composition of purse seine harvest, represented as a proportion, by set type and year from 1997-2018.

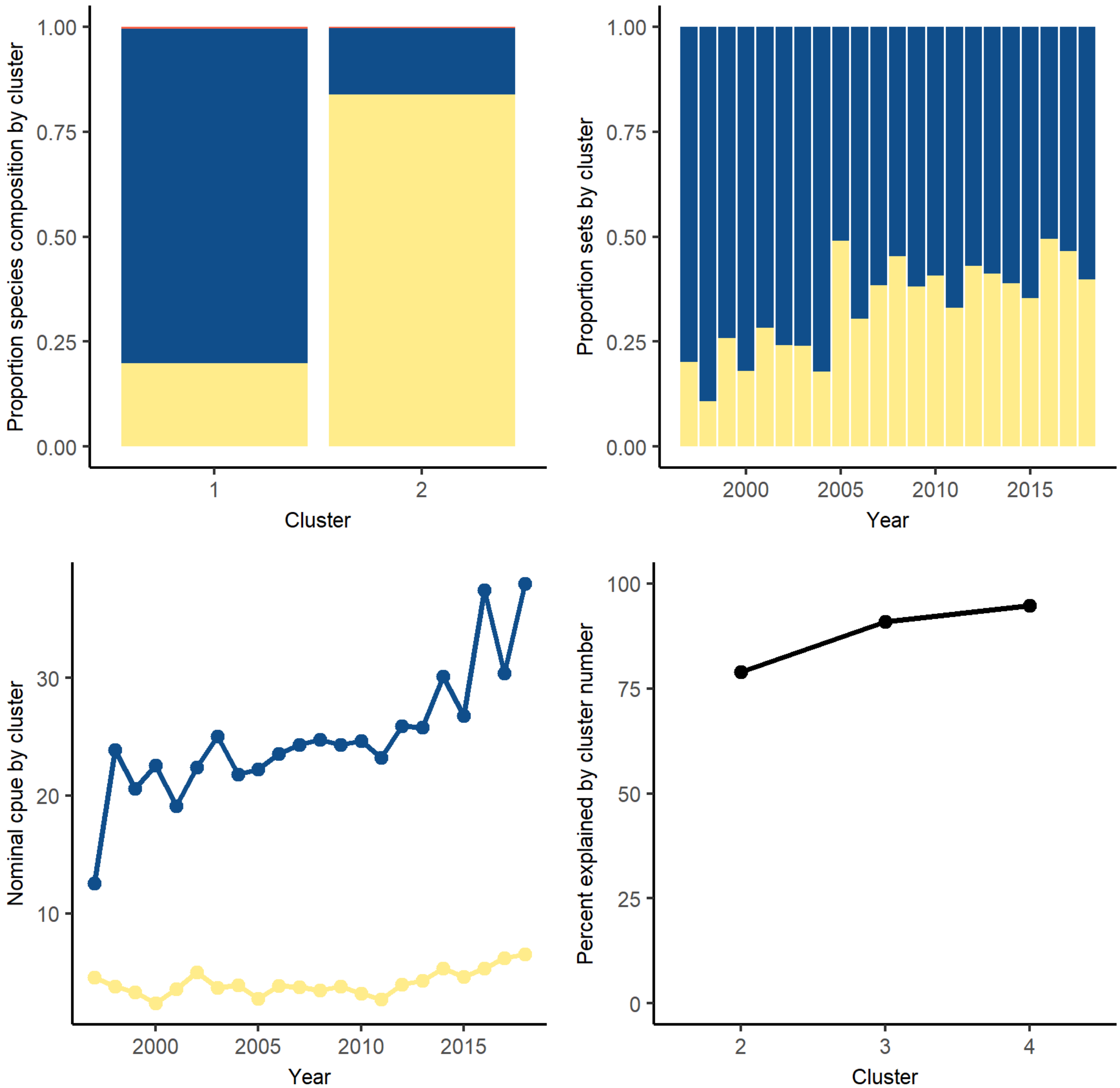

Figure 6: Cluster characteristics and catch summaries for the final species composition cluster variable. In each plot blue represents skipjack, yellow represents yellowfin tuna, and red indicates bigeye tuna. The two cluster groupings are representative of sets that harvested primarily skipjack (blue) or yellowfin (yellow).

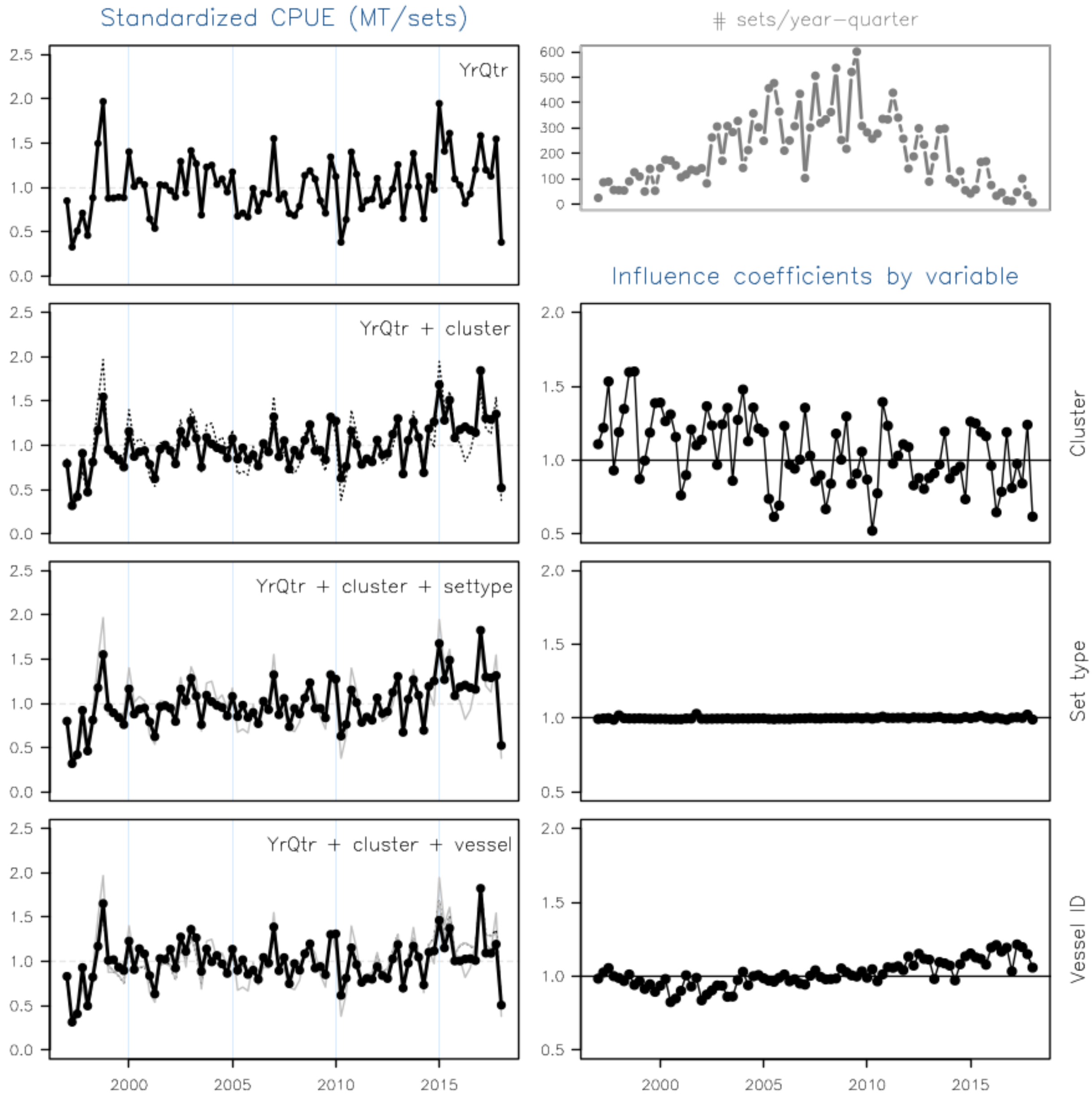

Figure 7: Step (left) and influence (right) plots for each explanatory variable, building up to the final model. In the step plot, the current standardized index is shown in bold, the index from the previous step is drawn with a dashed line and earlier indices are in gray. The right-hand panel depicts the influence of each covariate on the index through time.

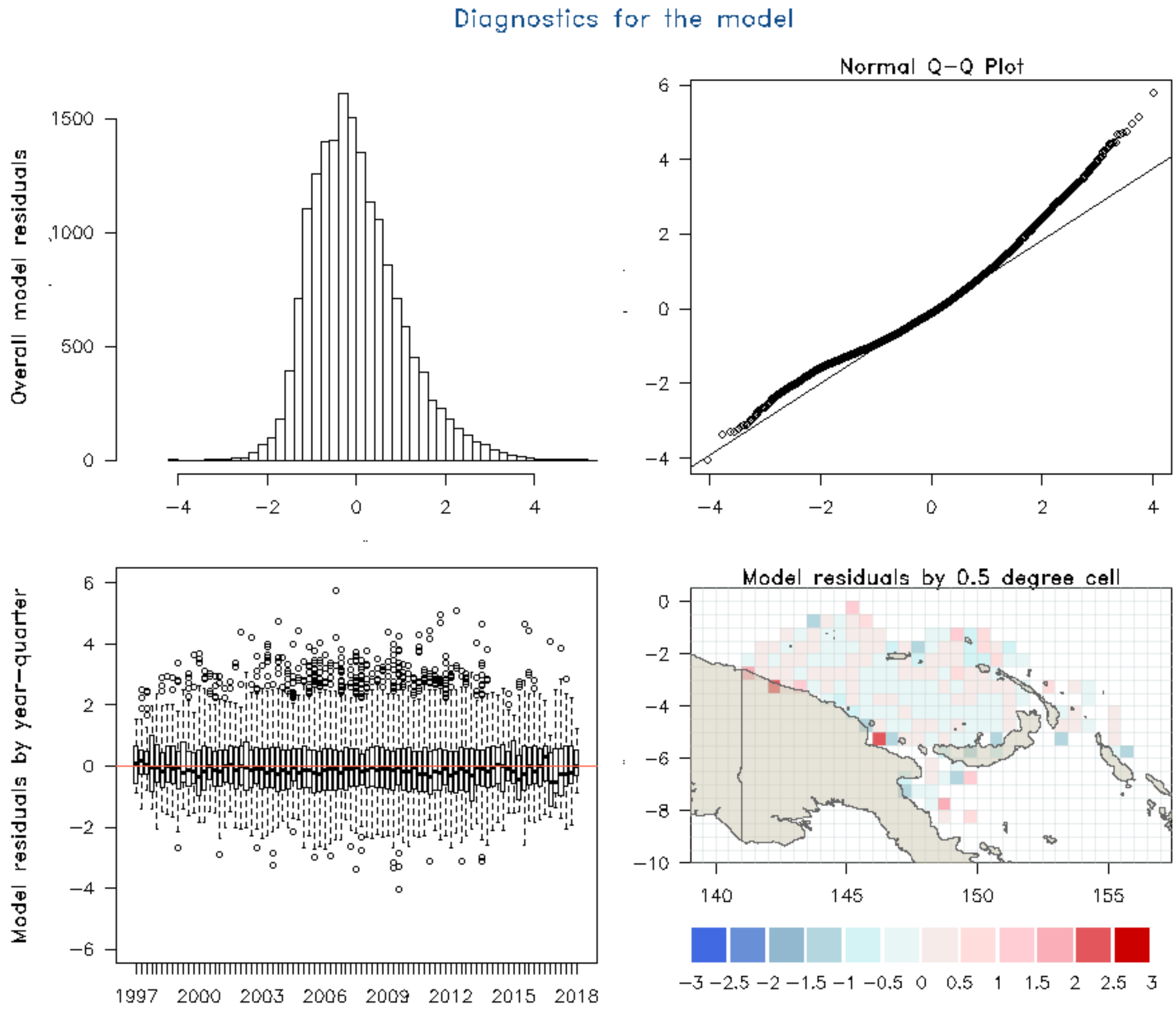

Figure 8: Key diagnostics for the final Tweedie GLM.

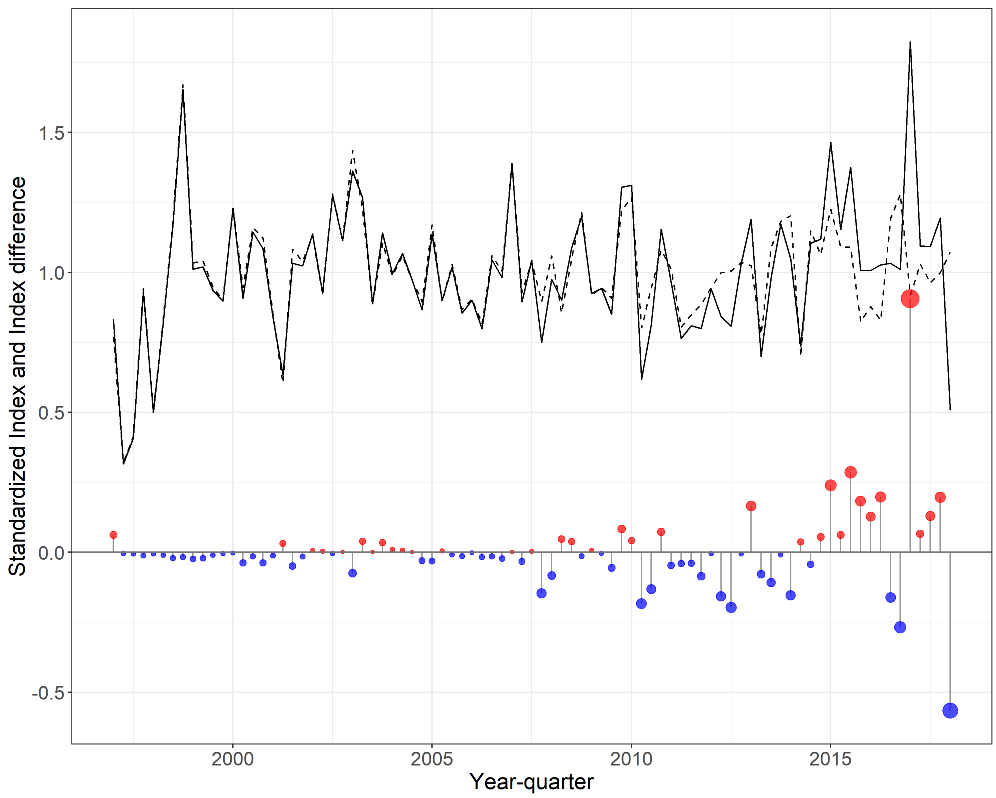

Figure 9: Standardized CPUE indices from the filtered data set with associated sets (solid line) and all set types (dashed line). The difference between the two indices (all set types index subtracted from the associated set only index) is shown with segments branching off the horizontal line at an index value of 0. When the standardized CPUE index estimated from associated sets is higher than the standardized CPUE index estimated from all set types (including unassociated sets), the points are red (point sizes are scaled according to the absolute difference between the indices), and when the associated index is lower, the points are blue.