SCIENTIFIC COMMITTEE
TWELFTH REGULAR SESSION
Bali, Indonesia
3–11 August 2016

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

WCPFC-SC12/SA-IP-04

Laura Tremblay-Boyer[1], Graham Pilling[1], Brian Kumasi[2], Thomas Usu[2]

[1]Oceanic Fisheries Programme, Secretariat of the Pacific Community
[2]National Fisheries Authority, Port Moresby, Papua New Guinea

1 Introduction

Indices of abundance are notoriously challenging to derive from purse seine catch and effort due to the efficiency of the traditional unit of effort, the set, being influenced by a suite of vessel attributes that evolve over time ('effort creep', see also Pilling et al. 2016). To circumvent this, stock assessments for the skipjack fishery of the Western and Central Pacific Ocean (WCPO), for which purse seiners account for about 90% of the catch (Williams and Terawasi, 2016), have relied on indices from pole-and-line fisheries. The effort and spatial extent of these fisheries has reduced over time, however, compromising the validity of the indices in recent years.

In 2014, a new abundance series was presented to the Scientific Committee based on the PNG archipelagic purse-seine fleet (Pilling et al., 2014) for the WCPO skipjack stock assessment (Rice et al., 2014). This purse seine fishery was deemed appropriate for standardization due to the nature of the fleet and stability of the anchored FAD-based fishing strategy over time (see Sokimi, 2009, for more details).

At the time of the previous standardization (Pilling et al., 2014), 55.9% of the sets and 56.9% of the skipjack catch (in mt) were on anchored FADs over the standardization period of 1997-2012. A core fleet of 13 representative vessels had been identified, which fished mostly on anchored FADs and accounted for almost 60% of the skipjack catch. The situation has evolved over the three years of data accumulated since (2013-2015), with the vessels designated under the 2012 core fleet now taking 26.2% of the SKJ catch and 21.7% of overall sets being performed on anchored FADs. The annual catch for both skipjack and yellowfin tuna have declined, but the yellowfin decline is less severe, such that the overall proportion of yellowfin in the catch has increased. In parallel, unassociated sets have become the prevalent fishing method (Figure 3). The transition from anchored FADs to unassociated sets started to occur in 2007 but became especially pronounced over the last 5 years. Since unassociated sets tend to include less skipjack than sets on anchored FADs (–18.1%), this transition is likely to impact abundance trends inferred from CPUE. Lastly, the fishing grounds for this fleet have remained mostly stable over time, with the bulk of the effort occurring within the Bismarck sea (Figure 2).

The objective of this analysis is to generate standardized indices of abundance for skipjack to be used in the 2016 stock assessment. We used a similar approach to that used in the 2014 analysis but we both expanded the definition of the core fleet and modified the model structure to account for the changes in the fishery.

2 Methods

2.1 Preparation of the dataset

The dataset was assembled from SPC-available logsheets of sets performed within PNG archipelagic waters between 1997 and 2015 (see map in Figure 2). Only records with school type anchored FADs, drifting FADs, other associated (drifting logs, whales, whale sharks) or unassociated (free school) were retained (99.9% of entries). The dataset was further processed for the standardization by removing sets with skipjack catch greater than 175 mt (corresponding to the 99.9th quantile of observed catch by set). Skunk sets, where there was no tuna catch (¡1mt), were also removed. This was done as a zero skipjack set resulting from a skunk (failed) set should reflect a manipulation issue, while a zero skipjack set where other species of tuna were caught should reflect a targeting and/or abundance trend (but see discussion). Finally, we only kept records belonging to vessels within a 'core fleet', as described below.

2.2 Core fleet definition

The core fleet was defined in order to retain vessels that were representative of the fleet and had enough records through time to justify the addition of a vessel effect that could be distinguished from a year-quarter effect. From the 75 vessels that occurred in the initial dataset, we kept only vessels that had been active for 6 quarters or more (at least one set) and had been at least once in the top 95% of skipjack catch for the year, leaving 37 vessels in total. For these, vessel records were retained for a particular year-quarter only if at least 5 sets had been performed by this vessel in the 3 months period.

2.3 Clustering

We used clustering to create an additional 'species targeting' variable to use in the CPUE standardization. Clustering then allows classification of sets into a set number of groups ('clusters') where the distribution of caught species by set are similar. The most-caught species in the cluster is used to identify the cluster's main target species, based on the assumption that sets that target a specific species should have a higher proportion of that species in their catch. Cluster analysis has been recommended by several authors to formally generate this set-wise classification (He et al., 1997; Hoyle et al., 2014), and has been used in recent standardizations of bigeye and albacore CPUE in the WCPO (Bigelow and Hoyle, 2012; Tremblay-Boyer et al., 2015b,a).

We used k-means clustering with the Hartigan-Wong algorithm from the R package 'stats' (R Core Team, 2013). The k-means algorithm assigns observations to a user-defined number of clusters to maximize the Euclidean distance between the group means ('cluster centre', the mean proportion of each species in the catch of sets belonging to the cluster) based on the proportion of the three tropical tuna species (BET, SKJ and YFT) caught by a vessel in a set. We ran the algorithm specifying 2, 3 and 4 clusters, and with 20 different initial values each time to ensure that the final cluster configuration was stable.

There is no formal criterion to select the 'best' number of individual clusters for a dataset. Instead, we based our selection of cluster numbers on the number of species most frequently caught, differentiation of CPUE trends between clusters (as otherwise this variable would not yield additional insight) and visual inspection of the relationship between the number of clusters and the proportion of variance explained by including an additional cluster.

2.4 Delta-lognormal model

We used a delta-lognormal approach to model CPUE as a function of covariates, where the probability of having a set with non-zero catch and the CPUE of the catch when positive are modelled separately using a binomial and a log-normal distribution, respectively. The binomial GLM uses a binary response variable (y_i ; $1 =\geq 1$ fish caught, or a $0 =$ zero fish caught in set i)

$$y_i \sim \text{Bernoulli}(p_i) \tag{1}$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_{yq[i]} + ... \tag{2}$$

where p_i is probability of at least one skipjack individual being caught in set i, the logit link function is used to express this probability in terms of the linear predictor, β_{yq} are the year-quarter coefficients, and '...' are coefficients for the levels of additional factor variables included for set i.

For the lognormal response variable,

$$\log c_i \sim \text{Normal}(\log \mu_i, \sigma^2) \tag{3}$$

$$\log \mu_i = \beta_0 + \beta_{yq[i]} + ... \tag{4}$$

where $\log \mu_i$ is the log-CPUE of the number of skipjack caught in a set, and the parameters in the linear predictor are interpreted as above. The same covariate structure was used for both the lognormal and the binomial component.

We explored the following covariates for the standardization: year-quarter, set type, vessel name, vessel attributes, cluster type, cell (lon-lat) and other species catch. The final model structure was based on traditional model diagnostics like the improvement in AIC, residual patterns over space and time, 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.

3 Results

The expanded core fleet (37 vessels *vs.* the initial 13 of Pilling et al. 2014) delineates two main types of fishing strategy, anchored FADs and unassociated sets, with the prevalence of the latter increasing over time (Figure 3). Skipjack predominates in set composition by set type, except for unassociated sets where yellowfin is dominant (Figure 4). For all set types there are no clear trends over time in species composition (Figure 4). Core fleet vessels appear to mostly use the anchored FAD or unassociated sets approaches, with vessels active earlier in the time-series predominantly using anchored FADs and vessels active later on predominantly focusing on unassociated sets (Figure 5).

We retained two clusters for the final clustering variable, as clustering identified groups that clearly target skipjack or yellowfin but not bigeye tuna. The two clusters explained 80% of the variation in species composition between clusters. When more than two clusters were specified, clusters with intermediate mixes of the two main species were identified but did not otherwise yield additional information. The main trends for the two clusters configuration are summarized in Figure 6. The yellowfin cluster increased in prevalence over time and the nominal skipjack CPUE was different between clusters, which underscored the importance of this variable for standardization.

The final CPUE model included year-quarter, set type, cluster and vessel ID, with no interactions. Step and influence plots (Figure 7) show the relative change in standardized year-quarter effects by the sequential addition of new explanatory factors to the GLM model. Influence plots assess the influence of a covariate on the standardized index given the change in its relative distribution over time. A positive influence means that the unstandardized indices are higher when the covariate is not accounted for (see Bentley et al., 2012, for more details). In the current standardization, the variable that had the most influence was the cluster, which increased the decline of the nominal CPUE over time. Vessel ID was also important in more recent years as some vessels achieving higher SKJ tonnage per set joined the fleet.

Key residual diagnostics are included in Figures 8 and 9 for the lognormal and binomial components respectively. Residuals for the lognomal model show some deviation from normality due to the different shape of the distribution for the response variable (e.g. tail length) between some levels of the covariates, most especially the set type. We explored alternative probability distributions but could not achieve significantly better diagnostics, and year-effects were not affected by the use of alternative distributions.

4 Discussion

This paper presents standardized indices of abundance for skipjack tuna for the archipelagic purse seine fishery of Papua New Guinea. The final series (Figure 7, bottom panel) is less variable than the nominal, mostly due to the inclusion of the targeting/cluster variable, but otherwise shows a similar trend in time. The effect of set type was not as pronounced as expected despite the strong shift over time from anchored FADs to unassociated sets, but that shift is also captured by the cluster variable (since there is a higher prevalence of unassociated sets in the yellowfin targeting clusters).

There appears to be a trend for larger catches per set in recent years, but these larger sets are also associated with vessels newly-arrived to the fleet. This underscores the need for a reliable database of vessel attributes as a prelimanary dataset of vessel GRT explained some of the variation in set size. The accuracy of the assembled vessel attribute variables was not considered reliable enough to formally include in the analysis as it was missing for many vessels and the existing values often differed between databases.

The relative increase of yellowfin tuna in the catch and the shift away from anchored FADs requires further investigation, notably to identify whether this increase is driven by yellowfin abundance or is a function of lower skipjack abundance. There is also some effort occurring east of the Bismark sea in recent years, mostly with newer vessels using unassociated sets, and this is the only area that showed somewhat higher skipjack catch rates compared to the main Bismarck sea area.

Skunk sets were removed from the analysis (since these should typically not reflect abundance), but we note that there has been a slight increase in the frequency of these sets over time. In parallel, there was a strong ENSO event in 2015 which could have impacted catch rates for skipjack tuna due, potentially, to an eastern displacement of the stock (e.g. Lehodey et al., 1997). However, while such oceanographical effects can be used to informally interpret trends in local abundance, they should not be standardized against since the abundance index reflects one model region only within the stock assessment.

We recommend careful consideration of alternative sources of abundance metrics for Pacific skipjack tuna. If we cannot rely on pole-and-line indices, and the current PNG archipelagic fleet continues to change as rapidly as it has in the last years, that fleet will be impacted by the same issues that prevent us from using other purse seine fleets for indices of abundance. A reliable dataset of vessel attributes (e.g. tonnage, power, ...) could inform our interpretation of the high-catch sets observed for some vessels in recent years, and allow us to isolate the vessel effect from the temporal trend in abundance. This would be especially important if a different class of vessel is entering the fleet, as appears to be the case from the examination of the data presented here.

References

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.

Bigelow, K. A. and Hoyle, S. D. (2012). Standardized CPUE for South Pacific albacore. WCPFC-SC8-2012/SA-IP-14, Busan, Republic of Korea, 7–15 August 2012.

He, X., Bigelow, K. A., and Boggs, C. H. (1997). Cluster analysis of longline sets and fishing strategies within the hawaii-based fishery. *Fisheries Research*, 31:144–158.

Hoyle, S. D., Langley, A. D., and Campbell, R. A. (2014). Recommended approaches for standardizing CPUE data from pelagic fisheries. WCPFC-SC10-2014/SA-IP-10, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Lehodey, P., Bertignac, M., Hampton, J., Lewis, A., and Picaut, J. (1997). El Nino Southern Oscillation and tuna in the western Pacific. *Nature*, 389:715–718.

Pilling, G., PNA Office, Norris, W., and Hamp (2016). Examining indicators of effort creep in the WCPO purse seine fishery. WCPFC-SC12-2016/MI-WP-08, Bali, Indonesia, 3–11 August 2016.

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

R Core Team (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Rice, J., Harley, S., Davies, N., and Hampton, J. (2014). Stock assessment of skipjack tuna in the Western and Central Pacific Ocean. WCPFC-SC10-2014/SA-WP-05, Majuro, Republic of the Marshall Islands, 6–14 August 2014.

Sokimi, W. (2009). Purse seine fishing around moored fads in papua new guinea. Technical report, SPC Fisheries newsletter #129.

Tremblay-Boyer, L., McKechnie, S., and Harley, S. J. (2015a). Spatial and fisheries structure and regional weights for the 2015 south Pacific albacore tuna (*Thunnus alalunga*) assessment. WCPFC-SC11-2015/SA-IP-07, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Tremblay-Boyer, L., McKechnie, S., and Harley, S. J. (2015b). Standardized CPUE for south Pacific albacore tuna (*Thunnus alalunga*) from operational longline data. WCPFC-SC11-2015/SA-IP-03, Pohnpei, Federated States of Micronesia, 5–13 August 2015.

Williams, P. and Terawasi, P. (2016). Overview of tuna fisheries in the Western and Central Pacific Ocean, including economic conditions–2015. WCPFC-SC12-2016/GN-WP-01, Bali, Indonesia, 3–11 August 2016.

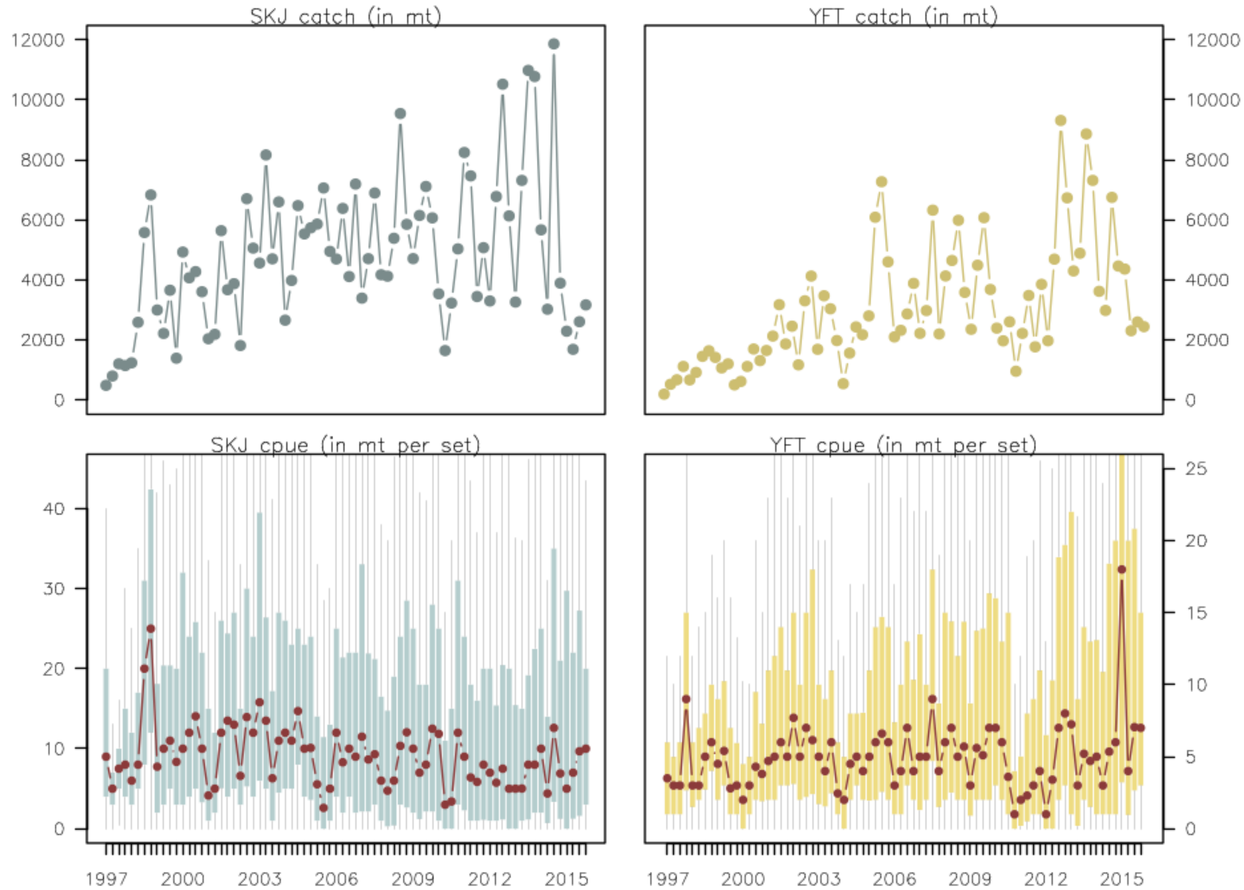

Figure 1: Summary of quarterly catch (top) and CPUE (bottom) (nominal, mt/set) for skipjack (left) and yellowfin (right) tuna from 1997 to 2015. The boxplot highlight the median in red and the box covers the 25th to 75th quartiles of CPUE observed for that quarter approximately.

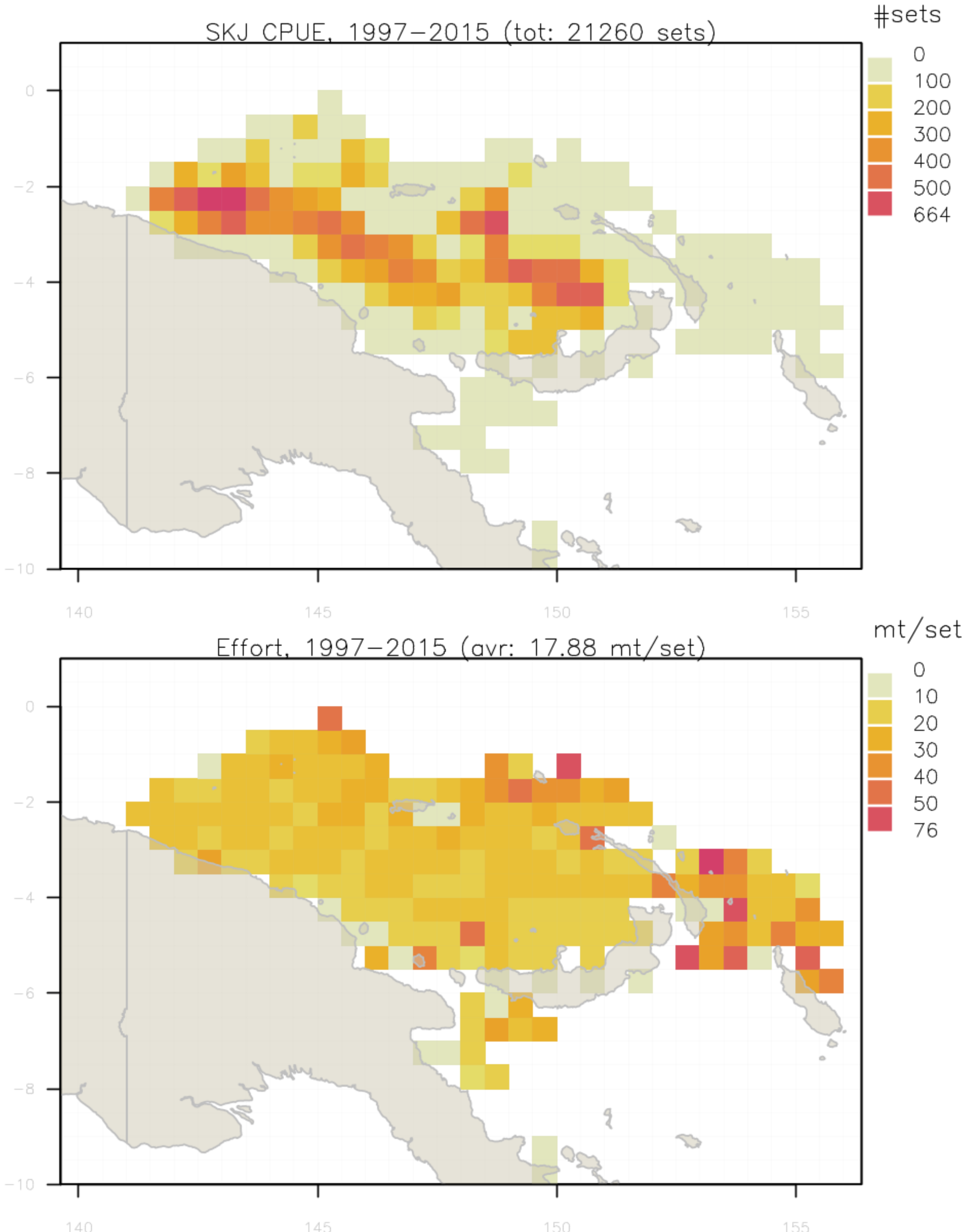

Figure 2: Aggregated effort by 0.5 degree cell in number of sets (top) and skipjack CPUE (nominal, mt/set) from 1997 to 2015.

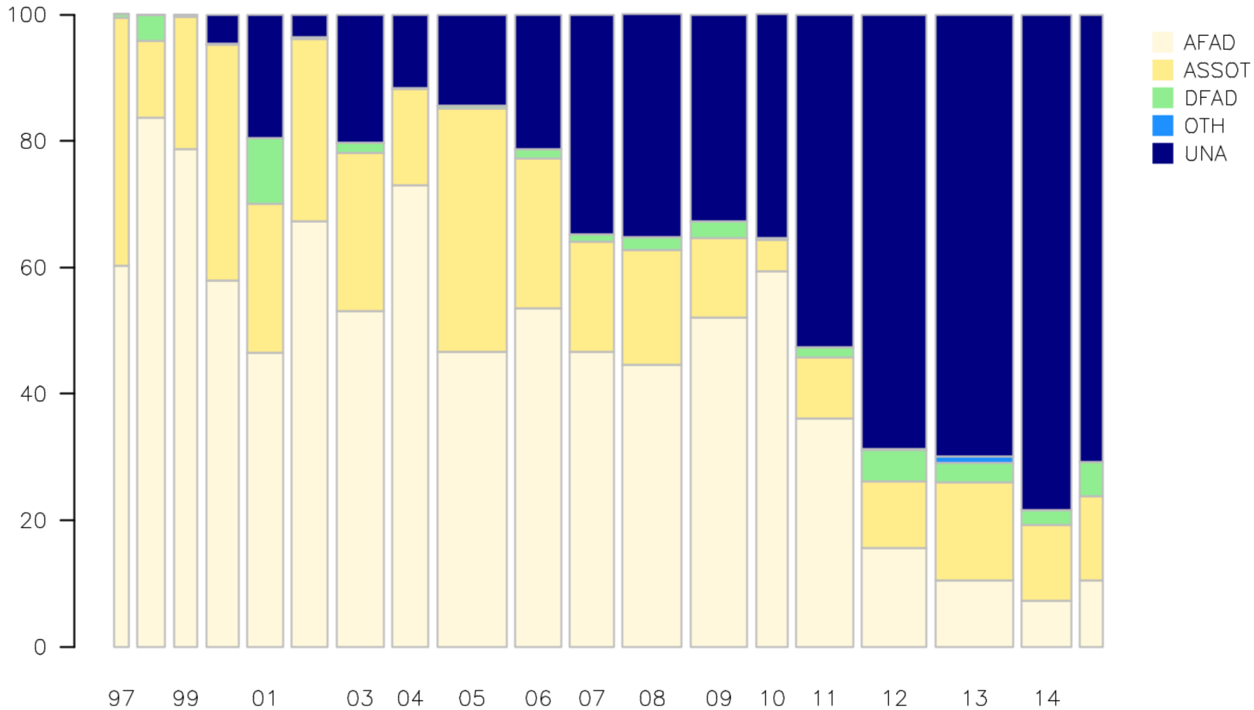

Figure 3: Relative distribution of effort by set type from 1997 to 2015. The width of the bar is proportional to the number of sets recorded in that year.

Figure 4: Relative distribution of key tuna species caught by set type from 1997 to 2015 for anchored FADs sets (top left), associated sets (top right), drifting FADs sets (bottom left) and unassociated sets (bottom right).

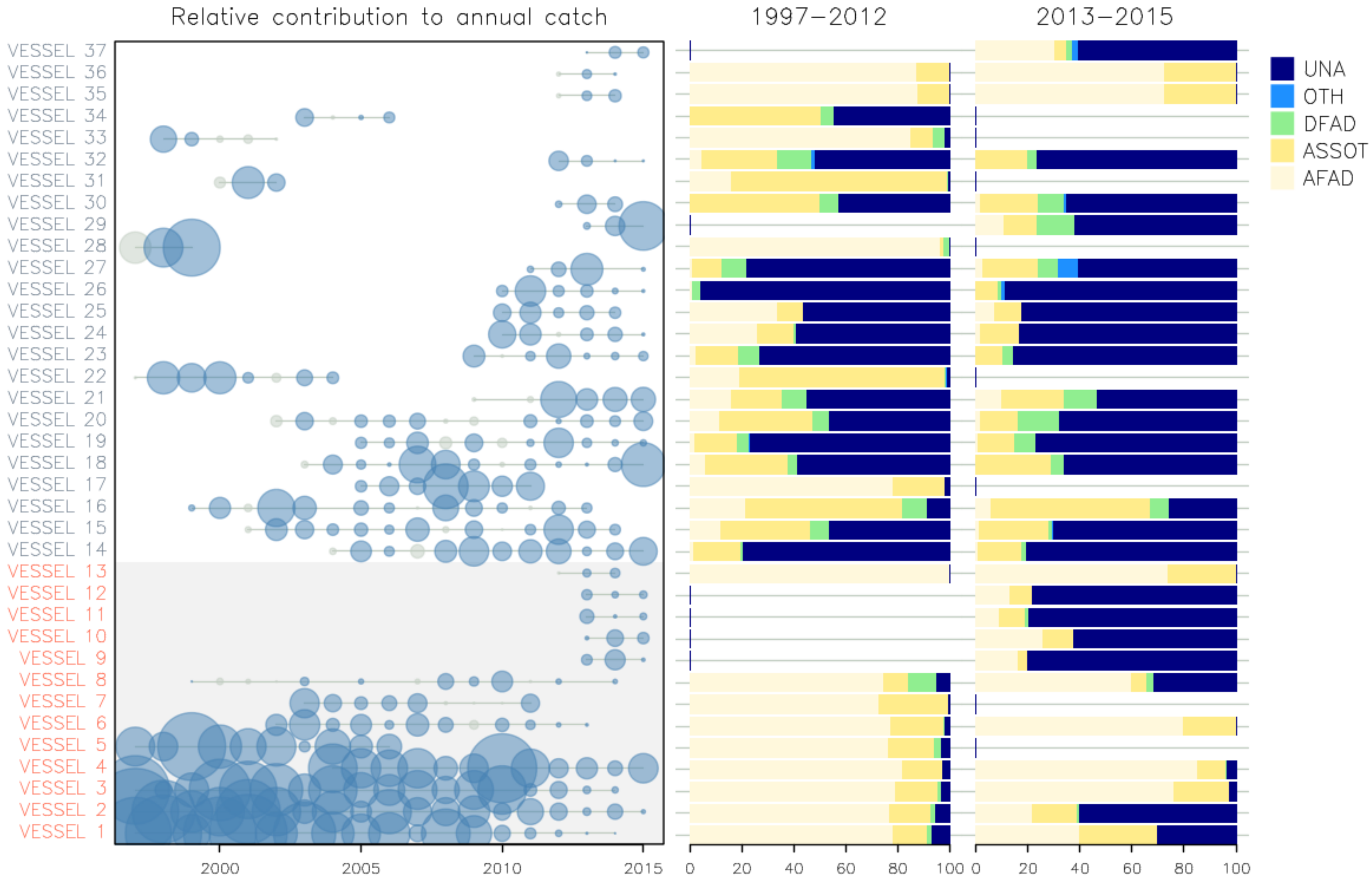

Figure 5: Key vessel statistics for the core fleet: relative contribution to the annual catch over time (left) and distribution of set type over the 1997-2012 and 2013-2015 period. The period split was chosen to highlight the new data coming since the last CPUE standardization for these fleets. The vessels are split between those that belonged to the core fleet for the previous standardization (bottom, vessel ID in red) and those which are new to the core fleet in this year's analysis (top, vessel ID in grey). Within each category the vessels are ranked by total catch 1997-2015, with greatest catch at the bottom.

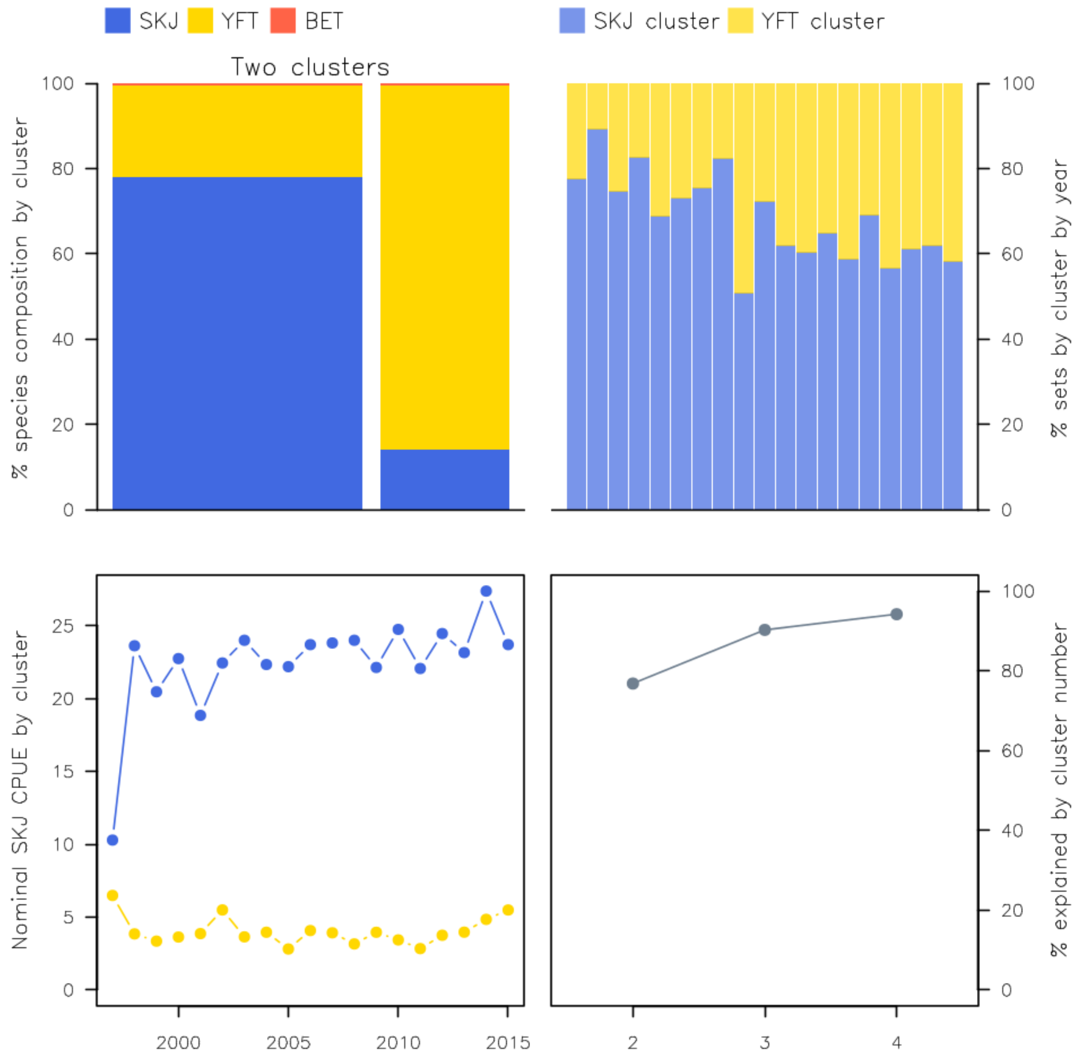

Figure 6: Key summaries for the two clusters configuration chosen for the final clustering variable: top left, average proportion of catch by species by cluster, with width of bar proportional to cluster size (the skipjack cluster is on the left); top right, temporal trend in the distribution of sets between the skipjack and yellowfin clusters; bottom left, annual nominal skipjack CPUE by cluster (average mt/set); bottom right, proportion of variation explained between clusters under the two, three and four clusters configuration.

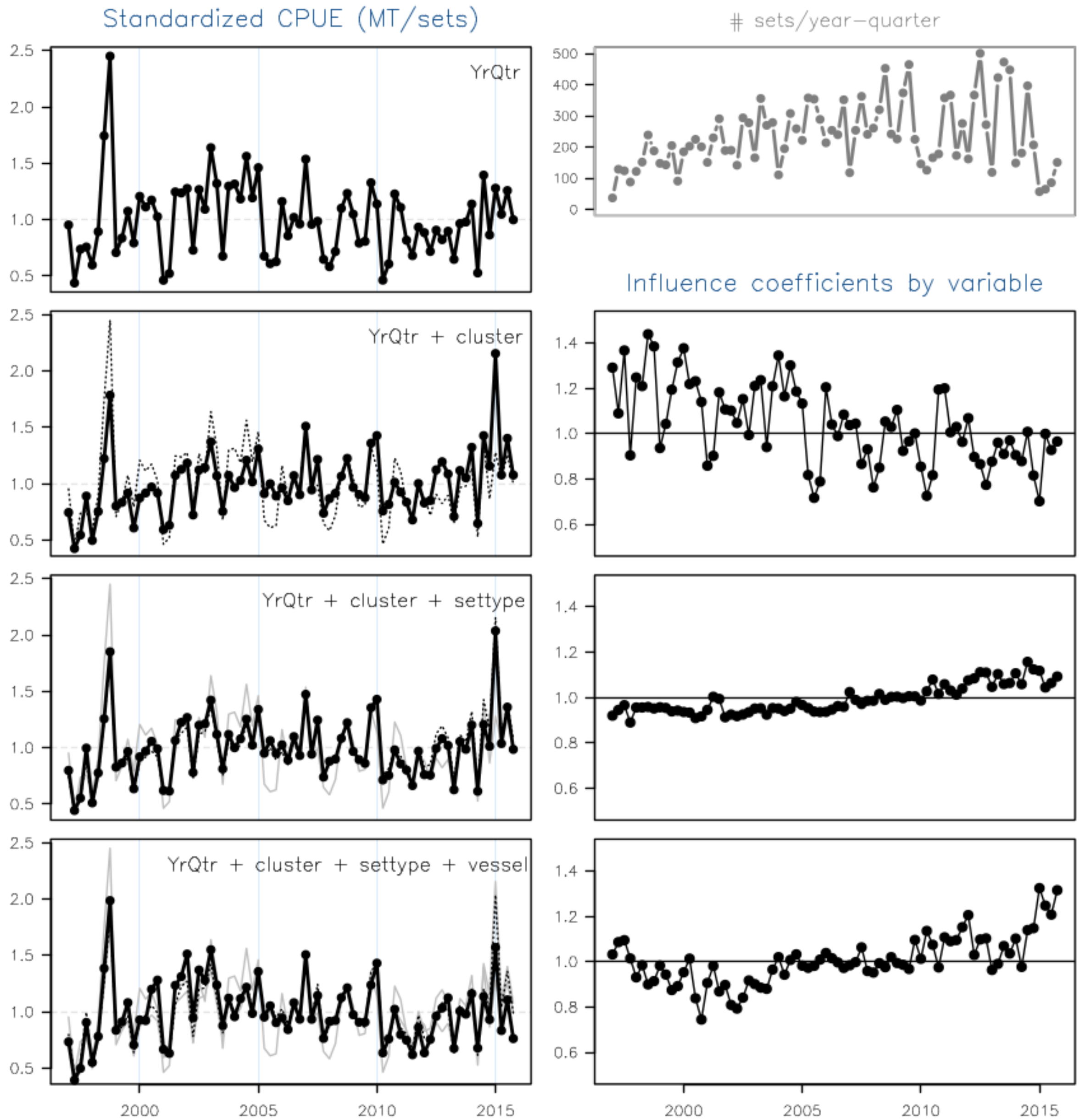

Figure 7: Step (left) and influence (right) plots for each explanatory variable, building up to the final model. In the step plot, the current standardized index is shown in bold, the index from the previous step is drawn with a dashed line and earlier indices are in grey. The influence plots are shown for the lognormal component of the model.

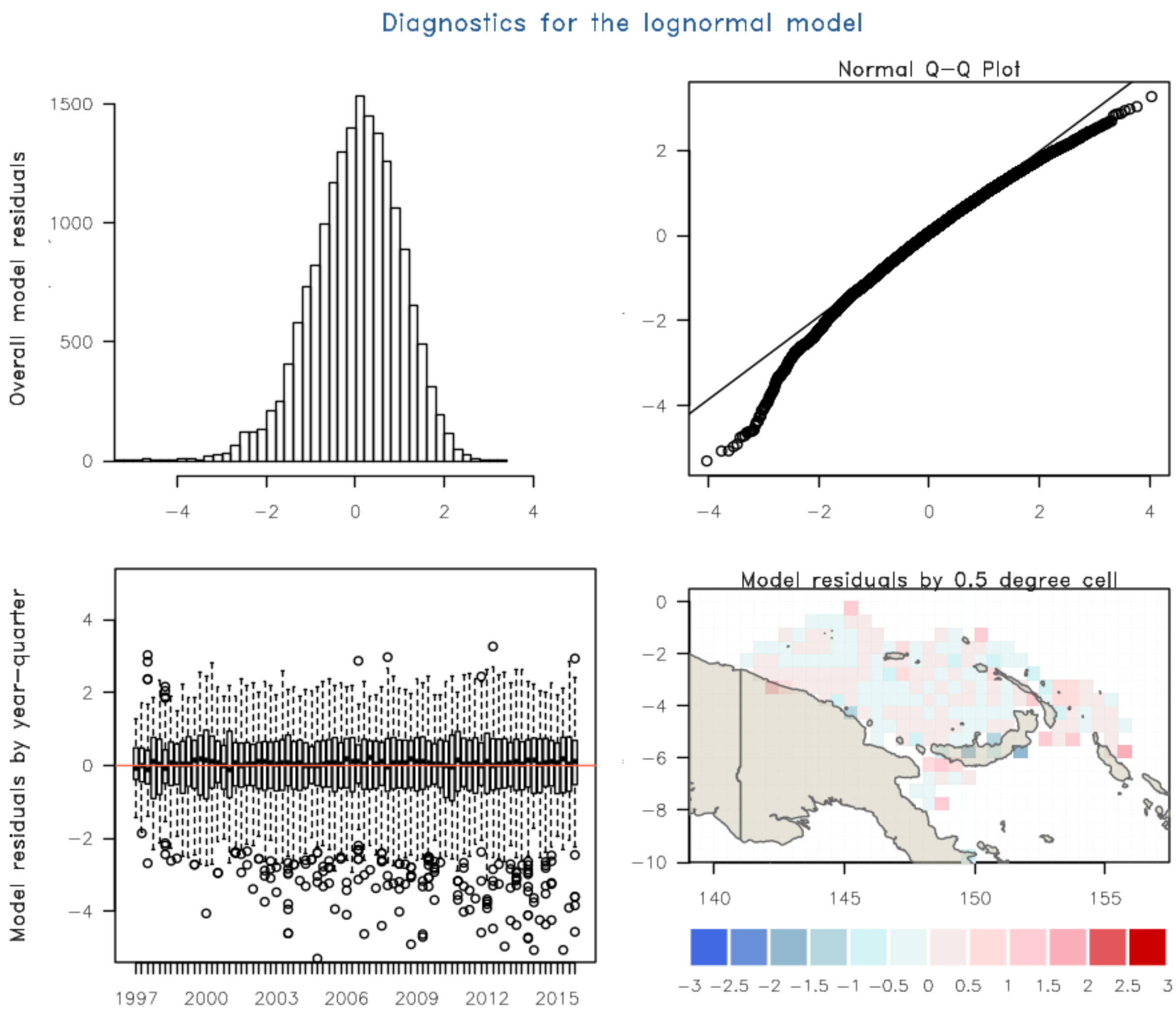

Figure 8: Key diagnostics for the lognormal component of the final model.

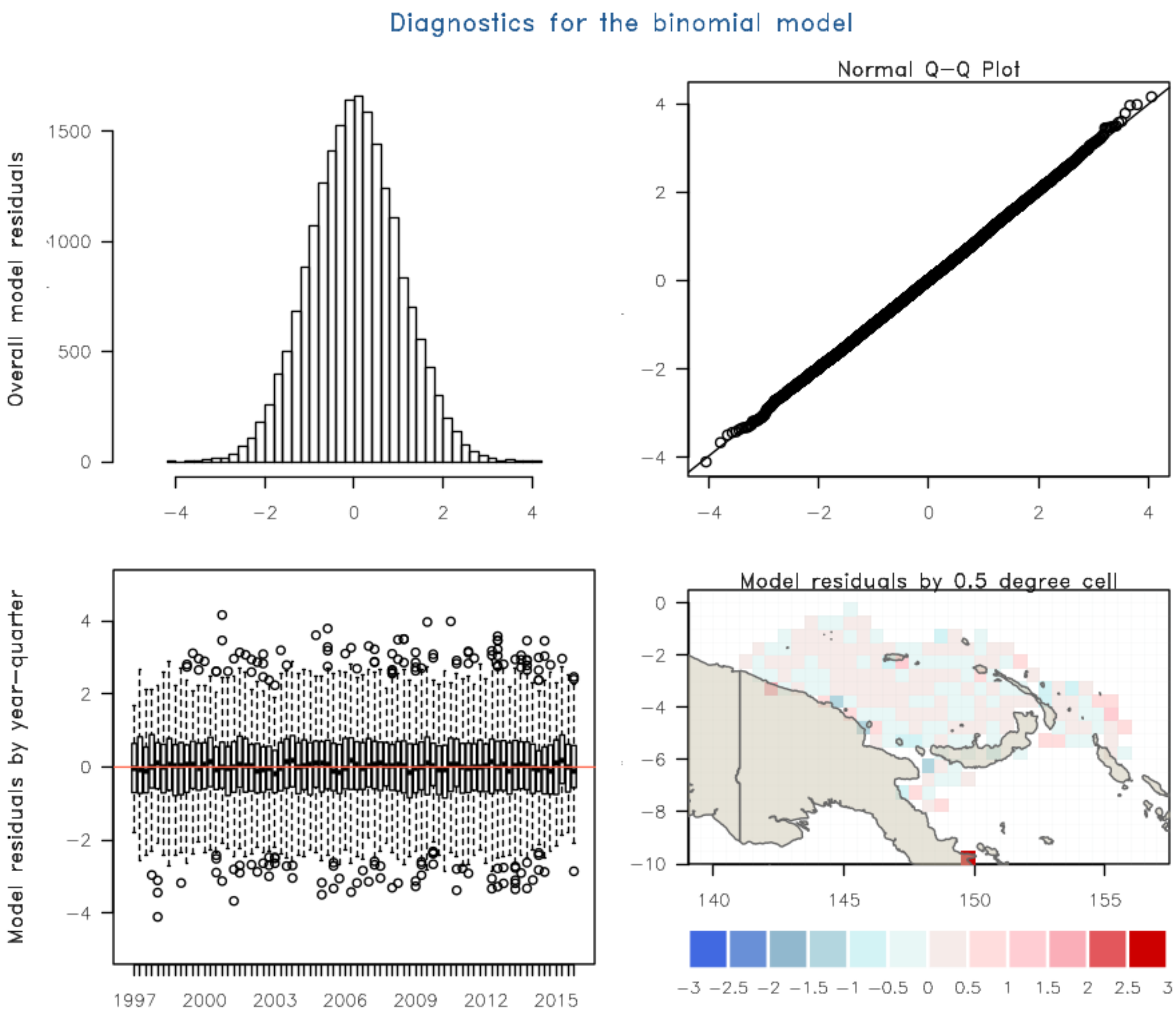

Figure 9: Key diagnostics for the binomial component of the final model.