SCIENTIFIC COMMITTEE
THIRTEENTH REGULAR SESSION

9-17 August 2017
Rarotonga, Cook Islands

Report of the workshop on analysis of CPUE for stock assessments, Noumea, April 2017

Graham Pilling and Steve Brouwer

OFP

Pacific Community (SPC), Noumea, New Caledonia

Report of the workshop analysis of CPUE for stock assessments, Noumea, April 2017

Oceanic Fisheries Programme, Pacific Community

Introduction

A workshop on operational longline data was held at SPC HQ, Noumea, on the 20th and 21st April 2017. The objectives were to collate the operational longline catch and effort data in preparation for the 2017 stock assessments of WCPO bigeye and yellowfin tuna, and southwest Pacific swordfish. The workshop was a result of the agreements reached in 2016.

Fourteen scientists from eleven organizations participated in the workshop, along with SPC staff. A list of participants is provided in Appendix 1.

John Hampton (OFP, SPC) opened the meeting and introduced the workshop. Steve Brouwer (OFP, SPC), who chaired the meeting, described general arrangements. The agenda is presented in Appendix 2.

This report briefly describes the various presentations made and focuses on important issues discussed by participants, and any specific suggestions made. The report does not attribute comments to countries except where the comment related to discussions of particular data issues or implied the undertaking of particular analyses.

The outcomes of this meeting will be reflected in the papers submitted to WCPFC-SC. Copies of most of the PowerPoint presentations prepared by SPC can be provided on request.

Operational data holdings and data rescue discussions

John Hampton provided an overview of the history of the current process, which was viewed as a collaborative project between countries and SPC. He noted that the agreement was outside the WCPFC, being between SPC and the individual countries. He highlighted the current status of operational data, noting that the 2015 workshop identified a work plan for country consideration to continue to improve the data sets.

Laura Tremblay-Boyer (OFP, SPC) provided a brief review of the data issues and recommendations identified at the 2015 workshop, and the current status of the Pacific-wide data set to be used within 2017 analyses. The presentation covered in particular patterns of hooks between floats (HBF), issues with consistent vessel identifications (vessel IDs), and set start time.

Workshop discussions clarified that the high frequency of 5 hooks between floats recorded early in the data set was reasonable for the 1960s, before deep longlining became more common in the mid-1970s. Participants were comfortable with the HBF patterns seen.

The vessel ID information provided in the data set suggested that those IDs frequently implied a number of very short periods of fishing (often 1 quarter of the year) by some vessels. In general, countries felt that the patterns for their fleet were reasonable. However, concerns were raised for specific fleets, in particular Japan and Australia. In general, the assigned vessel IDs within the data set were unique to an individual vessel and maintained through time (e.g. US, JP which were based on licence # or call sign), while in other fleets the vessel ID might change with a change of ownership (e.g. KR), or result from data recovery efforts that lead to changes in vessel ID early in the data set (e.g. TW).

The progress against each suggested action point from the 2015 workshop was reviewed.

On data rescue, Japan noted that no substantial progress had been achieved, but given the recent arrival (April 2017) of additional staff, future progress on for example vessel IDs prior to 1979 would be possible in the future.

Proxies could be used to fill particular gaps in information. It was suggested that hooks between floats information could be used to correlate with targeting (e.g. shallow fishing likely to occur at night). The influence of changing pattern of fishing activities to avoid bird bycatch was another key issue that would need to be considered in the targeting and standardisation processes.

Set start time information is now a mandatory data field (although the unit, e.g. UTC, home port or local, would be important additional information to collect and specify). It represents a potentially important variable when examining targeting by vessels. This field was not historically present within the Japanese logbook, but was added in 2016, although the unit needs to be confirmed. However, information on night-time and day-time setting for the Japanese fleet should be available historically, and the inclusion of the gear type in CPUE analysis could be investigated with SPC in future. In turn, for Japanese data after 1994, the gear type may include information on the intended target species (e.g. swordfish, shark or tuna) and time of set (e.g. SWO at night, shark/tuna during the day).

Korea noted that set time information could be available historically, but in UTC rather than local time. A conversion to local time from UTC could be undertaken given that the lat/long position of set was known. The workshop observed that the current data agreement indicated local time should be provided, so this might need to be reviewed in future (e.g. provide set time and the units). For DWFN fleets, the sub-set of information already available to SPC through the activities of these fleets within SPC member EEZs might provide further information. Collaborative work between the countries and SPC would be required.

SPC and China had worked on data reconciliation, and the backlog of data had largely been resolved.

Hooks between float information in some of the Pacific island fleets had been examined by SPC. Previously recorded at the trip level, some data actually represented the number of baskets rather than HBF, which led to data issues. Data had been adjusted to correct this, and information was improved. There was the potential to undertake a data rescue exercise for historical data when capacity was available.

The workshop suggested:

- **Given that some DWFN operational data are available from fishing activities within EEZs, SPC and JP/KR collaborate to analyse both EEZ operational data and that provided under the scientific data agreement, to supplement available information. This would require that vessel IDs in the two data sets be matched in some way.**
- **Given uncertainties in the consistency of vessel IDs for some fleets, SPC and those countries should work together to improve the information.**
- **That Korea investigate the provision of a time series of set time information and the units e.g. UTC or local time.**
- **That SPC and Japan collaborate on the potential use of classification methods for estimating 'set start time' information prior to 1979, including through the data on night- and day-time setting, and to investigate provision of set target species information after 1994.**

Laura Tremblay-Boyer presented an overview of spatial and temporal trends in the available Pacific-wide data by fleet and species. The workshop noted that the coverage of logsheet information had improved over time, which would not necessarily affect the standardisation process but did influence some of the plots presented. There was the potential to disaggregate domestic and 'charter' elements for some fleets, but the use of vessel ID within analyses would avoid this issue. The workshop also noted an increase in the nominal CPUE trend for both bigeye and yellowfin in recent years.

The workshop suggested:

- **That SPC confirm that the 1960s Pago Pago data were present in the data set, and to confirm the overall time series of data for American Samoa as that data should in theory be available throughout the time series.**

- **The data for Australia be checked as the time series should peak in 2003 (around 13 million hooks).**
- **That to ensure consistent data holdings, countries produce diagnostics of the distribution of quarters actively fishing to compare against SPC held data, given the latter shows a very high proportion of vessels that are active for only one quarter.**

Multi-species CPUE standardisation approaches

Rob Campbell (CSIRO) presented recent research results on standardisation approaches for multispecies CPUE, as used within Australian harvest strategies for key fisheries. Simulation results indicated that delta-lognormal approaches and (dependent upon the catch data generation approach) clustering at the trip or set level worked best, while inclusion of gear information notably improved model fits. However, no particular model performed best across all species included in the simulation study. The discussion suggested that the consistent application of the two-stage modelling approach across all collaborators would have been useful to allow a more direct comparison to be made. Given the challenges in developing 'realistic' data for simulations such as this, the distributions of simulated catches were shown to compare well with those of observed catches. It was also noted that the potential to include the catch level of other non-target species within the GLM would be interesting, although trends in the abundance of those covariate species might bias the result.

Overview of proposed approach for 2017 CPUE standardisation

Laura Tremblay-Boyer and Sam McKechnie (OFP, SPC) provided a review of the CPUE standardisation approaches used in previous stock assessments, and a summary of the proposed approaches for the 2017 assessments. These included an approach comparable to that used within the previous assessments, that is, clustering of a core-fleet-filtered data set for each assessment model region to identify species targeted by fishing trips, and the application of a range of GLMs to develop standardised indices. The challenge of potentially combining time series where incorporation of vessel effects is possible through vessel ID (e.g. later in the time series), and where it is not (e.g. earlier in the time series) was highlighted. The workshop clarified that vessel ID (which is used as a proxy for the fishing power effect associated with individuals vessels) does not capture all vessel effects – e.g. those due to technical change, skipper, etc.

The workshop discussed how the 'optimum' number of clusters for inclusion within the analysis was identified. It was noted that while this was region specific, generally 2 or 3 clusters were used. While more clusters were 'better' than fewer, there was the potential to have unbalanced clusters through time. Alternative approaches to identifying the 'best' number of clusters exist (e.g. scree plots, deviance versus the number of clusters), but there appeared to be no clear 'preferred' statistical approach to the selection of cluster number.

Bigeye and yellowfin targeting

Laura Tremblay Boyer presented the data filtering approach to identify a core fleet of vessels for bigeye and yellowfin tuna. The criteria to define the 'core fleet', based on year/quarter of data and number of sets made during the time series, were described. The data filtering approach also took into account the potential challenges of being able to run statistical analyses on very large data sets.

The workshop noted that changes in both regulation and data provision in the Japanese fleet could influence identified patterns of targeting over time, notably, smaller Japanese vessels were allowed to operate further offshore later in the time series in Region 1, while the provision of coastal vessels data only started after 1994 and these vessels may have quite different gears and targeting. There would be a mix of targeting by Japanese vessels between tropical species and north Pacific albacore between 20°N and 15°N, in particular between Regions 1 and 3 in the 2000s. It was felt that this issue would be captured within the identification of targeting through cluster analysis.

For Region 6, a lower threshold for core fleet inclusion (20 yr/qtr, compared to 30yr/qtr) will be attempted. The potential to define a 'core' area of the fishery based upon e.g. representation in the catch could be examined. The cluster analysis to define targeting could also address this.

To improve the filtering process in the future, the potential to 'track' vessels, particularly those that re-flag during the time series, was queried. This is now possible through the TUFMAN 2 'vessel instances' information, but will not be possible historically without national assistance.

The workshop suggested for the core fleet analysis:

- **Exploring either the removal of spatial cells that have limited effort, or aggregate those cells into larger areas, as this could potentially ease computational burden and thus allow improvements to models fitted e.g. allowing the inclusion of interactions between time and space. To help this process, examination of CPUE patterns (for core fisheries) over space, as well as the catch and effort separately, may indicate 'higher abundance' areas.**
- **For Region 1 specifically, the pattern by vessel size could also be considered, given the smaller coastal fleet will be smaller day trip vessels. SPC and Japan to collaborate on the potential use of corresponding vessel size for the list of Japanese vessel IDs.**
- **For Region 3, investigate whether the proposed cut off of 30 yr/qtr corresponds to some key emerging fleets, which may affect the time series in particular in the later years, and consider whether a different cut off (e.g. 20 yr/qtr) could be used to ensure that the spatial/temporal pattern is appropriate.**
- **Consider sampling by the effort distribution of each fleet, rather than the cut off approach, to preserve spatial coverage of data.**

Laura Tremblay-Boyer presented the proposed clusters to be included within the standardisation approaches for each model region. It was suggested that sub-setting the data by specific cluster could be a 'scenario' within the CPUE analysis approach. However, the workshop favoured an alternative approach where all cluster information was retained, and the cluster group was used in the analysis as a covariate. This would retain more information, and ensure that sufficient data for each vessel was maintained.

Temporal clustering, where data sets are separated by e.g. decade, was suggested by the workshop. This had been examined previously, and it was found that as long as sufficient cluster numbers were used within the overall clustering analysis, temporal patterns were captured. However, the potential for the

long-time series approach to wrongly capture patterns that relate to changes in relative abundance between species was noted. This was a reason behind retaining both yellowfin and bigeye clusters for the analysis in particular regions, since the pattern could be confounded between those two species, Similarly, albacore and swordfish clusters might be more closely related to targeting of those species.

For the clusters in Region 5, there was no specific 'targeting' of albacore on the east coast of Australia. The albacore 'cluster' likely emerges due to the high abundance of the species in the catch in early years, despite those vessels targeting e.g. yellowfin. The consideration of two clusters was suggested.

For Region 9, the nominal CPUE is likely to be used within the assessment.

The workshop suggested for the clustering analysis:

- **That an alternative scenario be used that retains all cluster information, but uses the cluster group as a covariate in the analysis.**
- **That a table be developed for each region to indicate the number of vessels excluded at each stage of the process, where table columns were: total number of vessels; number of vessels after yr/qtr filter; number of vessels after exclusion of clusters; and the rows presented those numbers over decades starting in 1960. This table should also incorporate the corresponding catch and effort information.**
- **Incorporating information on the gear deployed, where available, would provide better direct information on targeting than inferring targeting through the clustering analysis.**

Swordfish targeting

Yukio Takeuchi summarised the CPUE time series used within the 2013 swordfish stock assessment, and presented the proposed approach for the 2017 assessment and preliminary results for identifying targeting of swordfish within the operational data set.

The workshop noted that EU, NZ, AU fleets are likely targeting SWO but have short time series. For the assessment, those fleets could provide abundance indices for the recent period, but the model will require a longer time series to cover the temporal span of the assessment, likely from one or more of the distant water fleets.

The workshop suggested:

- **Given the EU fleet in Region 2C are known to target swordfish, they should be removed from the clustering analysis and that analysis re-run on the remaining data.**
- **Following an additional 5 years of data, there is the potential to develop a standardised time series for the NZ CPUE given that the period of change in targeting of that fleet will have passed.**

Bigeye and yellowfin CPUE standardisation

Laura Tremblay-Boyer introduced the geo-statistical approach to CPUE standardisation proposed for the 2017 stock assessments of bigeye and yellowfin. The workshop noted that following the choice of the number of knots, there were alternative approaches to placing those knots on the spatial surface and that the number of knots on the spatial surface was the biggest driver of computing time for this modelling approach. The uneven concentration of sets for particular species over space was observed, and potential alternative subsampling schemes were discussed including by flag and by knot. The potential to arrange knots differently in areas considered to be of low biomass/uniform CPUE compared to 'core' regions was also discussed, but geostatistical models make the assumption that sampling intensity (e.g. number of sets) is independent from the modelled variable (e.g. CPUE). The workshop was supportive of the geostatistical approach and suggested that efforts in this area be continued.

The workshop suggested:

- **Examine the potential to aggregate the data at an appropriate stratum to reduce the computational constraints (e.g. aggregate 1°x1°, hooks per basket).**
- **Examine whether knots could be concentrated in e.g. the tropical area where potential abundance information is high for these species, and fewer knots placed in temperate regions where the information content within the data is less.**

Francisco Abascal Crespo (EU) presented findings of the potential influence of oceanography on bigeye behaviour and catch rates in the WCPO, based upon archival tag data. He noted that over the longer time series (back to 1960s), the correlation with oceanographic variables was very low, compared to that where the more recent time series was examined. The potential to also use these results as a source of regional weighting within the assessments, where required, was also suggested. The workshop observed that, given the likely influences of oceanography on catch rates, it was important to continue to investigate this area.

Laura Tremblay-Boyer presented the proposed approach to incorporate oceanographic covariates within the 2017 geostatistical CPUE standardisation, focussing on the use of a time series of Pacific-wide thermocline depth. It was highlighted that the only source of oceanographic variables at the appropriate resolution going back to the 1960s was the ORAS4 model developed by the European Centre for Medium-Range Weather Forecasts.

The workshop suggested:

- **noting that the reliability of oceanography reanalysis models before the 1980s was likely to be lower than the more recent period, it may be appropriate to examine a model fit with a shorter, more recent, time series of data.**
- **Considering the use of an isotherm depth that indicates the 'preferred' temperature conditions for each species: e.g. using a 15° isotherm for bigeye, and a higher temperature**

value for yellowfin. It was noted that inclusion of thermocline value AND the difference in depth between two thermocline levels would also potentially capture this.

- **Examining the use of SST, which is 'measured' directly from satellites, rather than temperature depth that is a model-derived value.**
- **That there was the potential to include the outputs of the Peatman/Abascal Crespo model on BET depth distributions within the model, although again, this meant using a modelled value in the GLM.**

Laura Tremblay-Boyer presented the preliminary indices developed both through GLM where data were filtered through clustering, and the geostatistical approach. These were compared to the standardised indices (annual mean) used within the 2014 assessment. Current time series were shortened for display due to the inclusion of vessel ID within the models. The workshop noted that the trends were influenced by the clusters selected for exclusion from the data set for both species, dependent upon the model region. Retention of 'large data clusters' and use of a cluster effect may be preferable. For the geostatistical analysis, the results from alternative data sub-sampling approaches (by knot, by flag) were compared. Vessel or cluster effects were not included. A strong thermocline effect was indicated. The potential to use model selection (e.g. AIC) was suggested, but the use of model diagnostics was preferred, given that although the inclusion of non-significant covariates might affect the CVs, CV values are not currently used within the assessment. In turn, the impact of parameters is examined through influence plots.

The need for a long standardised CPUE time-series was highlighted. Given the issues with consistent vessel IDs within the data set, alternative approaches to developing those series were raised for discussion at the PAW.

<u>The workshop suggested:</u>

- **That the 'standard' GLM standardisation approach be the 'primary' approach within the 2017 assessments, and the geostatistical approach be used as an alternative test for bigeye in particular, and yellowfin if time allows.**
- **All CPUE standardisation approaches will use the Pacific wide operational data sets, including the Pago Pago data.**
- **For the GLM approach, retention of all data with inclusion of a cluster effect was recommended as the primary approach. As a potential scenario where clusters are excluded, the removal of only clusters that are small and clearly indicate targeting of another species should be considered.**
- **For the geostatistical approach, evaluating the impact of sub-sampling on the resulting CPUE index was recommended by performing the same analysis with repeated sub-samples. With sufficient time and computing power, standard errors could be estimated using this approach.**
- **The potential to use the geostatistical 'areas' identified through the geostatistical knots as a covariate within the 'standard' CPUE approach was suggested.**

- **To compare the influence of approaches and allow more direct comparison of results, the same general model used for the geo-statistical model should be used within the 'standard' GLM approach.**
- **Results from the geostatistical approach should be compared with and without the thermocline variable included.**
- **Examining whether vessel effects (as a random effect) can be included within the geostatistical approach (with a modification of the sub-sampling approach).**

Eric Chang (Taiwan) presented the preliminary results of standardisation analyses of Chinese Taipei longline catch rate data for bigeye and yellowfin. The data reconciliation approach between Chinese Taipei and SPC was described. As a practical approach, that analysis suggested that for small longline vessels, SPC data should be used, and for larger longline vessels, TW data should be used. Given the issues with the vessel ID within SPC holdings, the need for data to be recompiled to re-code the vessel ID was noted. The use of hooks per basket (HBP) as a targeting proxy is affected by gear design (weight of line material), current strength, etc. In turn, 'American-style' longline gear, with greater HPB, is now used to target albacore, in contrast to previous operational approaches. Hence a clustering approach may be more appropriate. He also highlighted that the annual coverage of logsheets varied over time. Comparisons of observer data-defined set targeting and species-based clustering results generated comparable standardised CPUE time series.

 The workshop suggested:

- **That as the GLMs had a number of factors, it would be interesting to re-run the GLM without either the observer or cluster effects included within the model, to identify which covariates led to the changes from the nominal CPUE series. The use of influence plots would help in this regard.**
- **That for the Taiwanese fleet, SPC data be used for small vessels and Chinese Taipei data be used for large vessels.**

SW Pacific swordfish

Rob Campbell (CSIRO) presented the results of CPUE standardisations from the Australian longline fishery for swordfish. He highlighted that the fishery has changed considerably over time, in particular in the species targeted. The workshop noted the potential for spatial serial depletion within the data, which had been shown by the author at an earlier SC meeting, however, the aim of the approach was to get a homogeneous area for the GLM effects. The potential for co-varying parameters within the GLM was highlighted, but the fishery has a diverse setting approach and the covariates allowed that to be captured. A sharp early decline within the CPUE time series was seen, which may strongly influence the assessment results. The potential for serial depletion of seamounts within the fishery to affect the time series (e.g. leading to hyperstability), and the relationship between the index and overall population biomass was discussed. The seamount influence would imply a much more complex spatial structure than is included within the assessment, and previous analyses that included a 'seamount' effect within the GLM showed little effect.

The workshop suggested:

- **that it would be useful to plot the standardised CPUE vs the effort.**

Other issues

Under other issues, John Hampton highlighted the potential to develop a collaborative peer-reviewed publication this year on the approaches used on this important and unique data set. Noting that data confidentiality would be maintained, and that outputs would be no more disaggregated than those presented to the Workshop, he highlighted that an article in a high impact peer reviewed journal would be a useful and tangible output from the scientific collaboration under the MOUs.

The workshop participants:

- **were in agreement with the approach, but noted that participants would need to discuss this and any paper drafts at the national level prior to submission.**

Final remarks

John Hampton and Steve Brouwer thanked participants for a fruitful workshop. A draft workshop report was circulated for comment among meeting participants prior to finalization and submission to SC13.

APPENDIX 1: List of Participants

Name	Affiliation
Rob Campbell	CSIRO, AU
Sangaa Clark	PNAO
Keith Bigelow	NOAA, US
Eric Chang	National Sun Yat-sen University, TW
Hung-I Liu	Overseas Fisheries Development Council, TW
Keisuke Satoh	National Research Institute of Far Seas Fisheries, JP
Takayuki Matsumoto	National Research Institute of Far Seas Fisheries, JP
Francisco Abascal Crespo	IEO, Spain
Sung Il Lee	National Institute of Fisheries Science, KR
Doo Nam Kim	National Institute of Fisheries Science, KR
Elaine Garvilles	Bureau of Fisheries and Aquatic Resources - National Fisheries Research and Development Institute, PH
SungKwon Soh	WCPFC Secretariat
Lilis Sadiyah	Center for Fisheries Research and Development, ID
Fayakun Satria	Center for Fisheries Research and Development, ID
John Hampton	SPC
Laura Tremblay-Boyer	SPC
Sam McKechnie	SPC
Yukio Takeuchi	SPC
Rob Scott	SPC
Graham Pilling	SPC
Steve Brouwer	SPC
Lauriane Escalle	SPC
Peter Williams	SPC
Tom Peatman	SPC

APPENDIX 2: Agenda

2017 Workshop on analysis of CPUE for Stock Assessments

Agenda

SPC, Noumea, 20th-21st April 2017

Thursday 20th		
09:00 – 09:15	Welcome and general meeting arrangements	JH
09:00 – 10:00	Operational data holdings • Summary of current status • Review of issues and progress compared to previous workshop	JH LTB
10:30 – 12:00	Operational data holdings • Round table update on data and data rescue activities • Overview of spatial and temporal trends by fleet and by species	JH LTB
Lunch		
13:30 – 14:30	Multi-species CPUE standardisation approaches (AU) Overview of proposed approach for CPUE	RC LTB/SM
15:00 – 16:30	Bigeye and yellowfin targeting • Summary of preliminary results Swordfish targeting • Summary of preliminary results	LTB/SM YT
Friday 21st		

09:00 – 10:00	Bigeye and yellowfin CPUE standardization
	• Introduction and rationale for new approach: Geostatistical GLM
	   - Importance of oceanography
	   - Inclusion of oceanography covariates
		LTB
		FAC
		LTB

10:30 – 12:00	Bigeye and yellowfin CPUE standardization • Summary of preliminary results • Chinese Taipei CPUE standardisation	LTB/SM EC
Lunch		
13:30 – 14:30	Swordfish CPUE standardization • Australian CPUE analysis	RC
14:30 – 15:00	Other Issues • Potential publications	JH
15:30 – 16:30	TBC	