CCSBT-ESC/0409/27

Data post-processing for input to the 2004 stock assessments and comparisons of 2001 and 2004 assessment datasets

Ann Preece
Scott Cooper
Jason Hartog

Prepared for the CCSBT 5th Meeting of the Stock Assessment Group (SAG5) and the 9th Meeting of the Scientific Committee (SC9)
6-11 and 13-16 September 2004, Seogwipo KAL Hotel in Jeju, Korea

Table of Contents

List of Figures

Abstract

This paper describes the data processing by Australia to provide inputs to the 2004 stock assessment models (including the management procedures operating model which has been used as a stock assessment model). For the first time we have predominantly used the data from the CCSBT database. Processing methods used by the members to prepare their data for the CCSBT database are described elsewhere (e.g. CCSBT-ESC/0409/26).

Some of the data provided to the CCSBT database are different to what has been distributed to the members in the past. There are differences in the actual "raw" data, and differences in how much pre-processing has been done. Some new methods have been used for processing of the CCSBT data to create the inputs to the stock assessments because of these differences, and these methods are described here. In addition to the CCSBT held data, Japan has provided data called 'L5M' for the years 1965-1994, and these data have been used instead of the CCSBT data.

Introduction

This paper describes the data processing by Australia to provide inputs to the 2004 stock assessment models. The stock assessment models anticipated at the 2004 Scientific Committee meeting include age based methods (eg VPAs), length based assessments (e.g. Kurota et al 2001) and stock assessments that use the management procedures operating model and data. All these assessments require data to be processed in slightly different ways.

For the first time we have predominantly used the data from the CCSBT database. The methods used by the members to prepare their data for the CCSBT database are described elsewhere (e.g. CCSBT-ESC/0409/26). Some of the data provided to the CCSBT database are different to what has been distributed to the members in the past. There are differences in the actual "raw" data, and differences in how much pre-processing has been done. Some new methods have been used for processing of the CCSBT data because of these differences, and these methods are described here. In addition to the CCSBT held data, Japan has provided data called 'L5M' for the years 1965-1994, and these data have been used instead of the CCSBT data.

In 2002 data were prepared for the initial management procedures runs and these data have not been updated until this year. The general description of the methods to be used were agreed by the members. Where these descriptions are not detailed enough to replicate or where there has been a change in procedure, the methods are described here. For stock assessment inputs in 2001 changes to methods were highlighted in Preece et al 2001 (CCSBT-SC/0108/21), and changes to these methods are described here.

The notes here are provided to assist in the gathering together of all data processing methods used and agreed in the past between members. However, this paper is not seen as a replacement for the full technical documentation of methods that have been presented or evolved at CCSBT SC, SAG and various workshops and tri-lateral meetings in the past. Some issues need further discussion or analysis to be resolved and these are highlighted. The details are provided to enhance transparency in the data processing and exchange arrangements and to enable others to replicate the methods.

Data inputs to stock assessment models provided by Australia

CSIRO processed the CCSBT and L5M data to produce the following inputs for the various stock assessments models. The inputs were provided to the CCSBT secretariat and to the members in electronic tables, and are not reproduced here.

Catch at age (CAA) summaries:

1) CAA for the Australian Purse Seine (PS) and Longline (LL) fisheries,
2) CAA for the Australian Surface fishery by season (July-June),
3) Indonesian CAA from direct age data (by season),
4) Indonesian CAA from length data by calendar year,
5) The Australian surface fishery CAA by season and cohort using the new growth curve (VB log *k*) and 2 alternative growth curves for hypotheses for pre-1960s growth.

Length Frequency (LF) summaries:

6) LF for the New Zealand (NZ) domestic fishery
7) LF for the NZ Joint Venture fishery (annual summary)
8) LF for the Indonesian fishery by season

Catch, Effort and Age Frequency data for the input file for the CPUE analyses:

9) Australian Joint Venture Catch, effort and age-frequency data by 5x5xmonth
10) Japanese LL catch, effort and age-frequency data by 5x5xmonth
11) NZ Joint Venture catch, effort and age-frequency data by 5x5xmonth

CPUE series:

12) Nominal CPUE
13) Laslett Core Area CPUE

Other inputs:

14) Mean length at Age data (by season) for the new growth curve (VB log *k*) and the two alternative growth curve for the pre-1960s growth.
15) Tag data.

Data Preparation – details of methods used this year

Australian data

The underlying data for the Australian fisheries were updated for only the most recent years (see CCSBT-ESC/0409/26). Bureau of Rural Sciences (BRS) provided:

1) total catch by gear, in numbers and weight, by calendar year (2002-2003) and quota years (2001/2002 and 2002/2003).
2) Catch in weight for the Aus surface fishery by season (2002 and 2003).
3) Length frequency data by 5x5xm by gear for Jan 2002 to Dec 2003.
4) Length frequency by half month was also provided for the purse-seine fishery.

CSIRO processed these data to provide:

1) CAA for the Surface Fishery by season and cohort for 2000-2003 seasons

2) Longline LF summary for 2000-2003 (calendar years)
3) CAA for the Australian longline fishery 2002-2003 5x5xm (calendar years)
4) CAA for the Surface Fishery (using half month data) for 2002-2003 calendar years
5) CAA domestic (Surface and longline data combined) for 2000-2001 calendar years
6) CAA for the surface fishery using the VB log *k* growth curve and 2 alternative growth curves

The Australian total catch and size data provided to the CCSBT for 1951-2001 were the same data used in the 2001 assessment and in earlier years, and in the MP data exchange in 2002. Therefore the stock assessment inputs using these data were not provided again this year, because they were unchanged (this includes the Australian Joint venture LF summary for the years 1989-1995).

CAA for the Australian surface fishery by season and cohort

The CAA for the Australian surface fishery by season and cohort for 2000-2003 were calculated from the half-month LF data, using gear type "PS" for 2002 and 2003 and state "SA" for 2000 and 2001. Proportional ageing was used (see Appendix 1 for details of this method). The age of a fish caught in Jul-Dec is incremented to be the age on the 1st Jan of the next year – this gives the cohort-age for the season (July-June). The data provided for the management procedures operating model for previous years were not updated because the data provided to the CCSBT were the same data used in 2002.

Separating the Australian domestic length frequency data by gear

The method for splitting the Australian-surface fishery from the LL fishery LF data for years 1951-2001 where gear type is not specified has not been documented elsewhere so is included here for completeness. The predominant gear type used in each state in each half year was used to determine the split into the LL or surface fishery for that state and half-year, as follows:

- All catches up to June 1989 are included in the surface fishery.
- All catches from SA are included in the surface fishery.
- All catches from NSW from 1989 onwards are included in the longline fishery.
- Tasmanian catches in July 1989 to June 1990 are included in the surface fishery, all other Tasmanian catches (after June1990) are included in the longline fishery.
- WA, Esperance and Albany catches in from Jan 1999 onwards are included in the longline fishery, all other catches are included in the surface fishery.

Calendar year CAA

The Australian CAA by gear was calculated for 2002 and 2003 because information on gear type has been provided by BRS for these years only. The CAA for the Australian domestic fishery (longline and surface combined) was calculated for the years 2000 and 2001. The LF and CAA summaries for earlier years are unchanged. The LF data are processed to create CAA using the proportional ageing method (Appendix 1).

Australian Joint Venture data for the CPUE input file

The input file for the Japanese LL CPUE analysis contains an Australian Joint Venture data component. These data (catch, effort, and age frequency) are unchanged from previous years. The Australian Joint Venture operated for the years 1989-1995. The underlying length frequency data are from Australian observer records and the catch and effort data from

Australian logbooks. These data were extracted from the data held at CSIRO. In the future, we should check to see whether these data can be extracted from the data provided to the CCSBT database.

Ageing using the new growth curves

The Australian surface fishery data was aged using the new growth curve (which we have dubbed "VB log *k*" (Polacheck et al, 2003)). The mean length at age, from this growth curve, for the 2 fishery seasons in the operating model were also exchanged.

Two alternative hypotheses for the pre-1960s growth and corresponding growth curves were also provided (the VB log *k* hypothesis assumes the same growth in the 1950s as the 1960s). These alternative hypothesis growth curves were dubbed "VBLK_SLSK" for the small *L_inf* and small *k* hypothesis, and the "VBLK_SL" for the small *L_inf* and same *k* hypothesis. The Australian surface fishery was aged using these growth curves and the mean length at age for the two fishery seasons were provided. For these two alternative hypotheses the boundary between the 1950s and 1960s growth curves needed to be smoothed for the ageing functions to work. There was no smoothing between the other decades.

NZ data

The NZ data provided to the CCSBT are different to the data used in stock assessments in the past. The CCSBT data were used. In addition, size data for the early years of the NZ domestic fishery (NZ_Dansol data) have been used (as in the past) but these data are not currently in the CCSBT database.

NZ domestic length frequency summaries

NIWA provided the NZ domestic fishery length frequency summary data for use in the operating model for the years 2000-2003. As was past practice, these data were not stratified at the 5x5xmonth level before being raised to equal the total catch in numbers in a year. Instead, the LF from across all areas were added together and then raised. The NZ domestic data have not been stratified at 5x5xm level prior to raising in the past (i.e. there was no stratification for production of the NZ domestic CAA for stock assessment in 2001 or LF for the MP operating model in 2002). The adequacy of the sample sizes at the 5x5xmonth level of aggregation should be analysed and proposals for substitution and raising methods developed.

CSIRO provided the NZ domestic fishery length frequency for the years prior to 2000, following the methods used in the past for the NZ domestic CAA. Two data sources for the raw length frequency data were used: NZ_Dansol data held at CSIRO for the years 1980-1993, and the CCSBT held size frequency data for 1994 onwards. The length frequency for all areas and months combined for each year was raised to equal the total annual catch in numbers in the RaisedCatch table in the CCSBT database (as advised by T.Murray).

The NZ_Dansol data includes weight and length measurements for individual fish, and sampling is greater than 50% of the catch for the years 1982-1990 (see Figure 1). For the LF in 1980-1981 the 1982 LF is used. The 1990 LF is used for the LF in 1990-1993. Length or weight data are used (length is used if both length and weight have been measured). These are the same practices used in the past to the created the NZ domestic CAA and LF summaries.

The CCSBT size frequency tables contain both length and weight data for the NZ domestic fishery with significant sampling from 1994 onwards (Figure 1). There are more weight measurements than length measurements each year, so weight data only was used, to avoid double counting of the same fish.

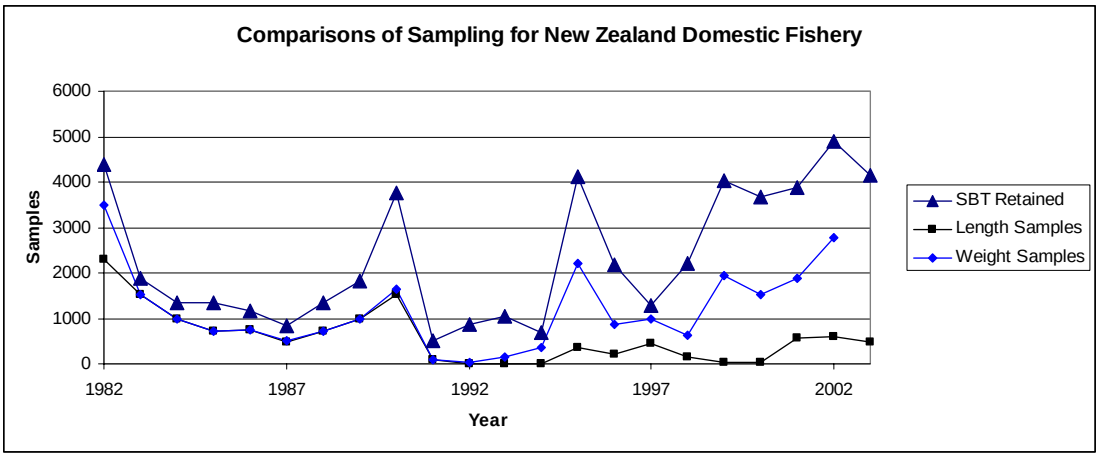

Figure 1: The number of length and weight measurements and numbers of SBT caught by the New Zealand domestic fishery.

NZ joint venture (charter) fishery length frequency summary

The NZ Joint Venture (JV) length frequency summary was provided by NIWA for the years 2000-2003 and by CSIRO for the years 1989-1999 (there was no charter fishery in 1996). The NZ JV LF data provided for the operating model were also not stratified at the 5x5xm level. Weight data were used for 1989-1997 and translated to lengths, length data were used for 1998-2003, because of higher sampling using these methods in these years (see Figure 2). Both length and weight data were from the CCSBT size frequency tables.

The length frequency from across all areas was added together in each year before it was raised (or lowered) to equal the total annual catch in numbers in the RaisedCatch table in the CCBST database. Numbers of samples are greater than the catch count in the early 1990s, which may indicate problems in the method used to estmate the total catches in the CCSBT database RaisedCatch table for this fishery.

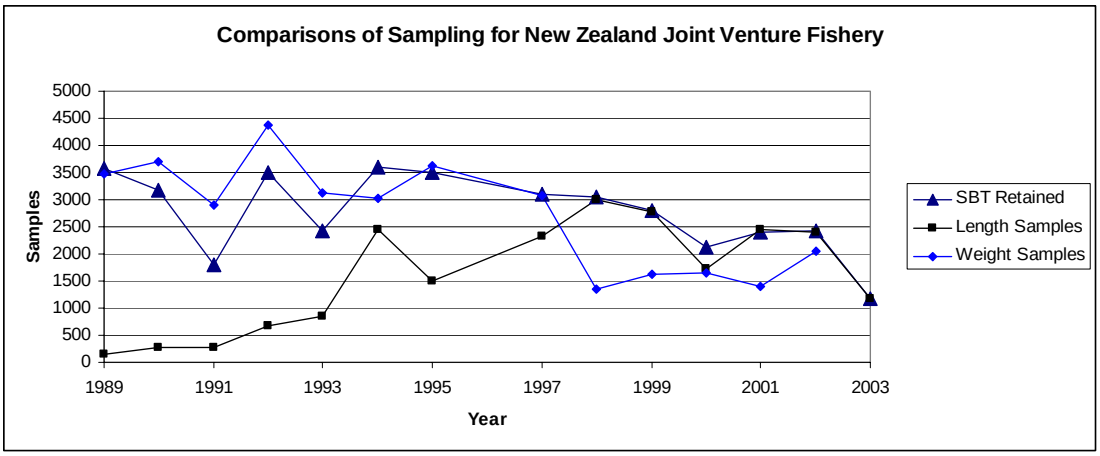

Figure 2: Numbers of length and weight samples and the total number of SBT caught each year by the New Zealand Joint Venture Fishery.

NZ joint venture data for use in the CPUE analysis

The NZ joint venture data is a component of the Japanese LL CPUE input data, used to create the CPUE series provided for use in the stock assessments. Only NZ joint venture Catch and Effort (CE) data where the target species was "SBF" were included. Only the LL data from the NZ_TLC logbook are included. The NZ joint venture operations with the Philippines are excluded.

These data (catch, effort and age frequency) are extracted at the 5x5xmonth stratification, and the underlying size frequency data are raised (or lowered) to equal the catch in the RaisedCatch table for each 5x 5xmonth strata, the practice in the past has been to use the higher of the 2 numbers. Weight data were used for 1989-1997 and length data were used for 1998-2003, because there were more samples of this type in these years. For strata where there is catch but no LF data we have used the average LF from all years in which the strata was fished.

The length frequencies are translated to age using cohort-age method (see Appendix 1). These data were not proportionally aged because this has not been the practice in the past. This was because there was no information in the past or current datasets about the width (or precision) of the length class.

Indonesian data

Size sampling and catch monitoring of the Indonesian catches commenced in 1993 in a collaborative program between CSIRO and Indonesia (Andamari et al, 2004). Catch totals (weight) were provided by the CCSBT in the global catch table. The size data from years 1994 and 1995 were used to provide the length or age frequency distribution for years 1976-1993.

Indonesian direct ageing data and length frequency data

The full time series for the Indonesian direct ageing data (by season and cohort 1995-2003) and length frequency (by season 1994-2003) was provided this year. The raw size data underlying both of these items is from the CSIRO/Indonesia collaborative research program,

and was changed to only include measurements from one qualified measurer (see CCSBT/SC/0318).

Indonesian CAA from length data

The calendar year CAA from size data (1976-2003) were requested by Japan for their stock assessment. The average length frequency for the years 1994-1995 were used as the length frequency for the years prior to 1994. The average of the mean weight of a fish in 1994 and 1995 was used to translate catch in tonnes to numbers for the years (1976-1993) where there was no size sampling.

Indonesian total catch in weight and average weight of a fish by season

The Indonesian catches before 1993 are not available by month, only calendar year. Examination of the post 1993 data indicated that about half the catch was taken in each half of the calendar year, and therefore a 50:50 split of the calendar year catch was used to create the total catch in weight by season.

The average weight of a fish in a season was provided to Japan for their length based assessment model (e.g. Kurota 2001), so that they could calculate the numbers of fish (from total catch in weight in a season) and create the LF by season for the early years using the average of the 1994 and 1995 LF.

Tag data

The Recruitment Monitoring Program (RMP) tag data (1991-1997) were re-extracted to allow for any updating of the data in these years. The method used was documented in CCSBT/SC/0108/21. The database extract query was provided to the CCSBT for future reference.

Japanese data

Japanese LL data for the CPUE input file

The Japanese component of the CPUE input file was extracted separately by CSIRO and NRIFSF. A combined file which included all three components (Aus JV and NZ JV extracted by CSIRO, and the Japan LL component) was provided by both groups to the members. These combined files were similar but not exactly the same. The causes of the differences appear to be in the ageing method used on the Japanese component of the file. We were not able to resolve the differences prior to the meeting, but concluded that the effects on the 4+ aggregated CPUE were very small. The Australian extraction of the Japanese LL data for the CPUE input file used the L5M data for the years 1969-1994 and the CCSBT data for the years 1995-2003. Proportional ageing was used for translating length frequency to age frequency. The raw data were already raised and substituted (by Japan).

Non-retained catches in 1995 and 1996

In 1995 and 1996 significant numbers of small fish caught during Japanese commercial operations were released. An estimate of the percentage mortality was agreed between the members (see CCSBT/SC/0108/21). These data are not yet in the CCSBT database and therefore have not been included in inputs used in this year's assessments.

CPUE series

The Laslett Core CPUE series and Nominal CPUE series were recalculated with the L5M and CCSBT data.

Comparison of 2001 and 2004 assessment datasets

In July, 2004, Japan provided their L5M data for use in this year's assessment. The L5M data have been used for the years 1969-1994. The Japanese data provided to the CCSBT were used for the years 1995-2003. The CCSBT data were used for the extractions for the other countries. Both of these data sets (L5M and CCSBT) have differences from the data used in the stock assessments in previous years. The data used in the 2001 assessment were provided by the members and updated each year. The data held at CSIRO, referred to here as "CSIRO copy of data", are a copy of the assessment data provided by the members. A brief analysis of the differences are provided here to assist stock assessment and management procedure data users in their understanding of where data may have changed and altered model fits, and to identify changes in the data that need to be clarified by the members.

For the Japanese data, at the gross scale (annual and all areas), the L5M and 2001 assessment data (CSIRO copy of data) are similar for total hooks and catch (in numbers) per year apart from hooks and catch counts in the 1990s (see Figures 3,4,5 and 6). In 1995 where we have used the CCSBT held data, there appears to be a major discrepancy (50%) in total hooks, and a smaller difference in the total catch in number in this year. There are other small differences for the late 1990s.

A finer scale look at the 1995 data shows only a 3% difference in hook count between the datasets and an 8% difference in numbers of SBT caught when the spatial and temporal areas are restricted to those used in the CPUE analysis: statistical areas 4-9 and quarters 2 and 3. Further finer scale (e.g. 5x5x month) comparisons of differences for the Japanese data in hooks, total catch and the length frequencies, could not be completed in time for this paper.

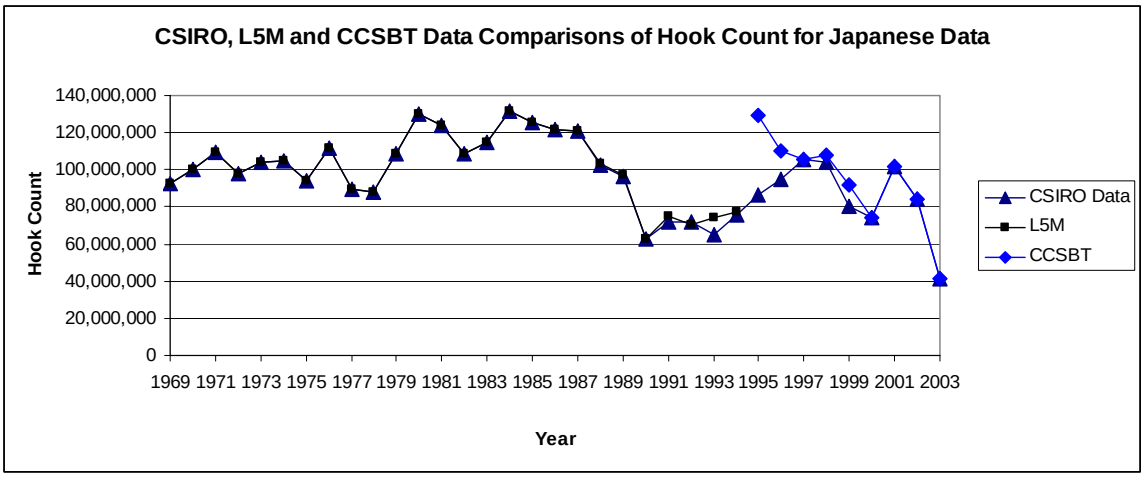

Figure 3. Comparison of the number of hooks per year for the 2001 (CSIRO copy of data) and 2004 (L5M and CCSBT) datasets.

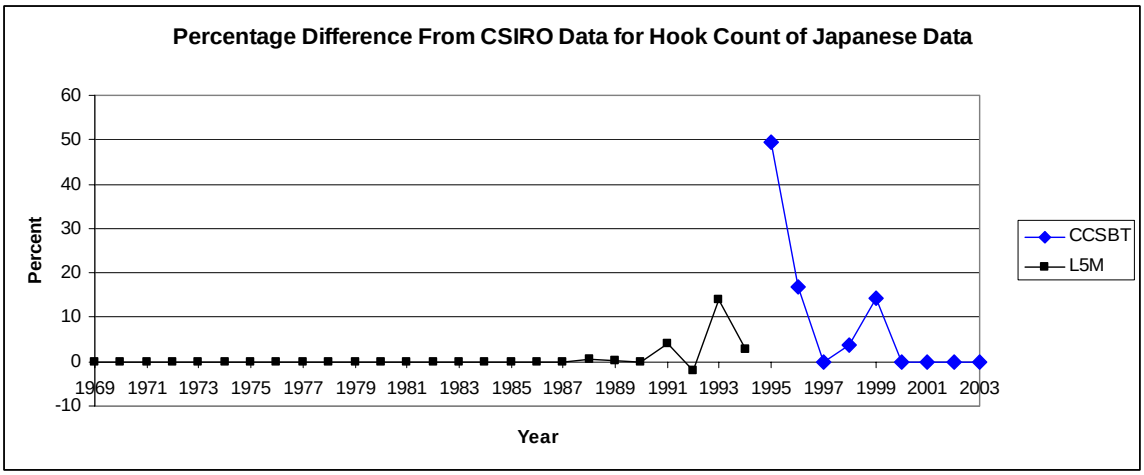

Figure 4: Percentage differences between the 2001 and 2004 datasets for the Japanese hook counts per year.

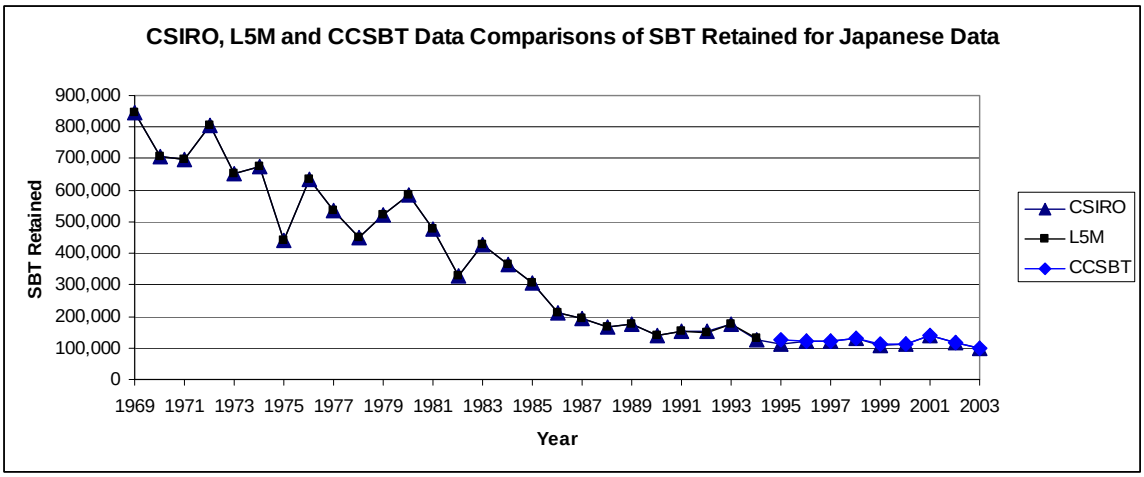

Figure 5: Comparison of the total catch in numbers per year for the 2001 (CSIRO copy) and 2004 (L5M and CCSBT) datasets

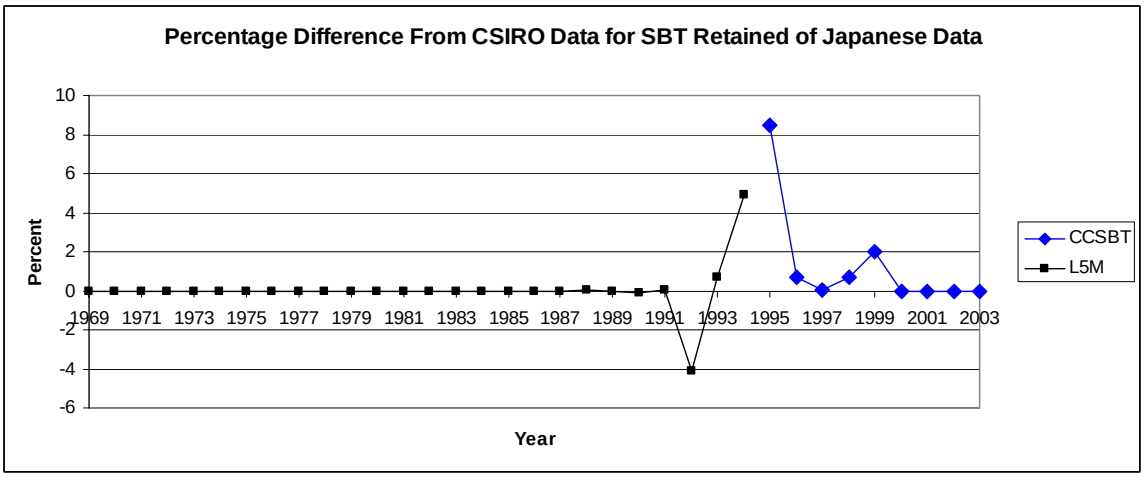

Figure 6: Percentage differences between the 2001 and 2004 datasets for the Japanese catch in numbers per year.

The New Zealand data used in the stock assessments have also changed. The CCSBT data were used this year. At the gross scale (annual and all areas), there are some differences in hooks and catch in numbers in some years for the Joint Venture Fishery (see Figures 7 and 8). The catch in number were taken from the CCSBT RaisedCatch table. In 1996 there was no Joint Venture. In 1991 there is a large difference in the number of hooks in the 2 datasets. Further analyses have not been undertaken because of time constraints.

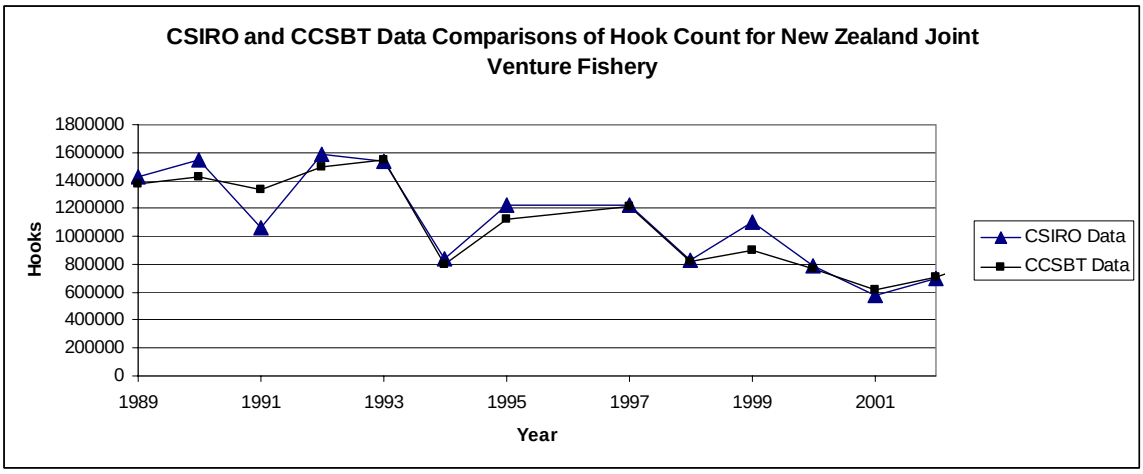

Figure 7: Comparison between the 2001 (CSIRO copy) and 2004 (CCSBT) assessment data sets for the New Zealand Joint Venture count of hooks per year.

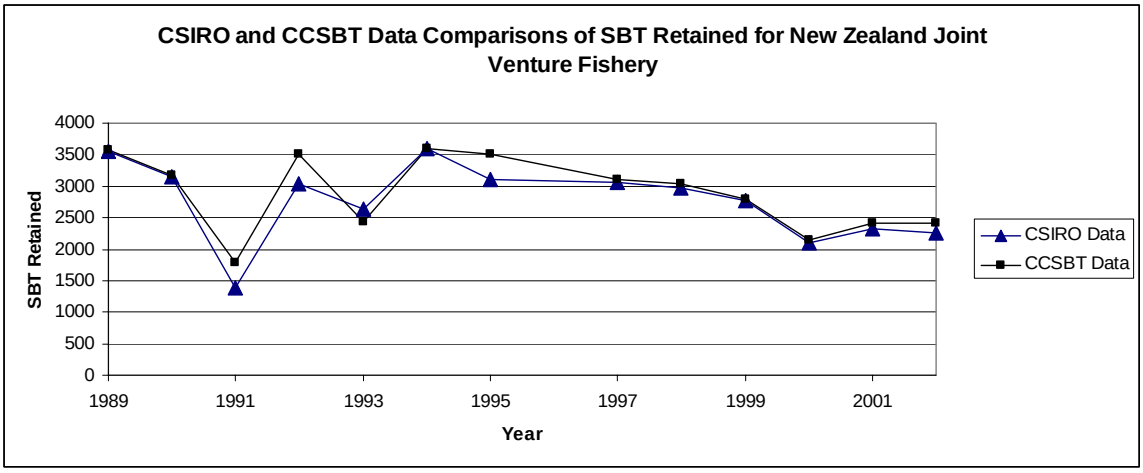

Figure 8: Comparison between the 2001 (CSIRO copy) and 2004 (CCSBT) datasets of the number of fish caught by the New Zealand Joint Venture operations.

Comparisons of the 2 datasets for the New Zealand domestic fishery show differences in the hook counts up to 1997 and then only very small differences in the most recent years (Figure 9). For the NZ domestic catches there have been some significant changes, possibly due to the method for estimating total number of fish from catch in weight and average weight of a fish (Figure 10).

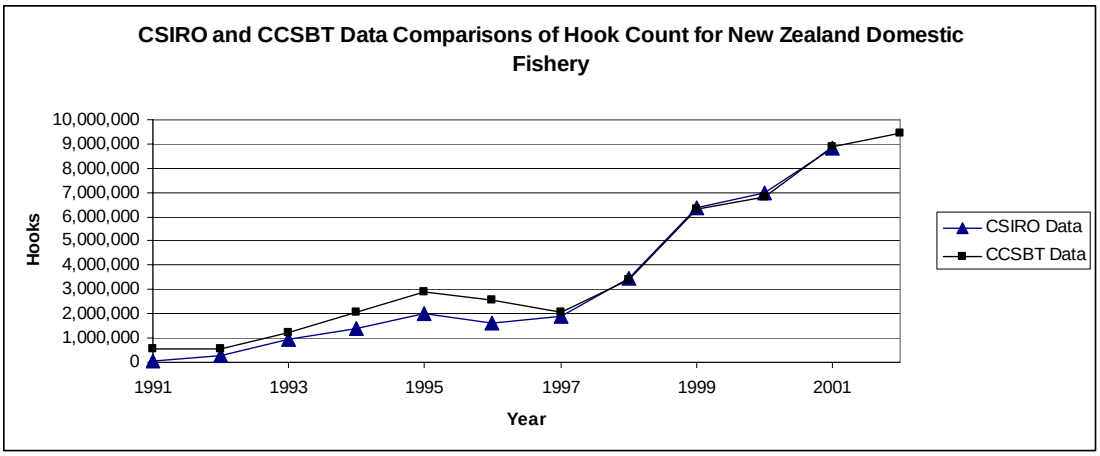

Figure 9: Comparison of the 2001 (CSIRO copy) and 2004 (CCSBT) assessment datasets for the New Zealand domestic fishery count of number of hooks set.

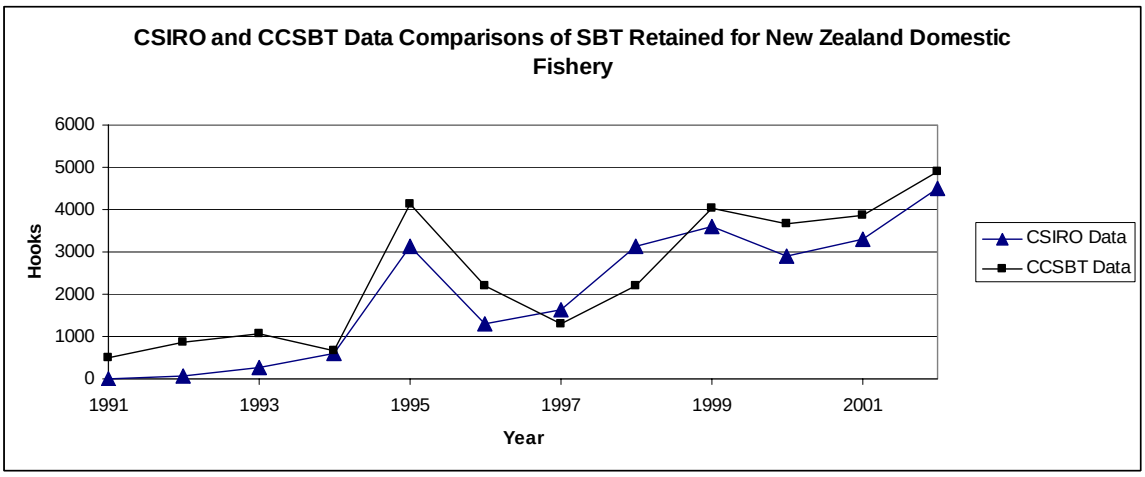

Figure 10: Comparison of the 2001 (CSIRO copy) and 2004 (CCSBT) assessment datasets for the New Zealand domestic fishery count of number of SBT caught.

There were also changes to the Taiwanese and Korean data used in the assessments this year, but these have not been examined.

Summary

Details of the methods used and issues to be discussed further have been provided here as an adjunct to the existing CCSBT literature on data processing methods that has been provided in the past. The CCSBT database which has been developed should become increasingly useful as the raw data provided to it is verified and methods used for post-processing the data are increasingly mechanised.

References

Andamari, R., Retnowati, D., Davis, T.L.O., Herrera, M., Poisson, F. and Proctor, C.H. (2004). The catch of SBT by the Indonesian Longline Fishery Operating Out of Benoa, Bali in 2003. CCBST-ESC/0409/11.

Farley, J.H. and T.L.O. Davis. 2003. Length and age distribution of SBT in the Indonesian longline catch on the spawning ground. CCSBT-SC/0309/18.

Hobsbawn, P.I., P.C. Sahlqvist and K.J. McLoughlin. (2004). Data Preparation for Australia's submission to the CCSBT Stock Assessment Group and Management Procedure Workshop 2004. CCSBT-ESC/0409/26

Kurota, H., S. Tsuji, N. Takahashi, K. Hiramatsu and T. Itoh. (2001). Exploration of cohort analysis based on catch at length data for southern bluefin tuna. CCSBT-SC/0108/32.

Polacheck, T., Laslett, G.M., and Eveson, J.P. 2003. An integrated analysis of the growth rates of southern bluefin tuna for use in estimating the catch at age matrix in the stock assessment. Final Report. FRDC Project 1999/104. ISBN 1 876996 38 2

Preece, A., T. Polacheck, D. Kolody, P. Eveson, D. Ricard, P. Jumppanen, J. Farley and T. Davis. Summary of the primary data inputs to CSIRO's 2001 stock assessment models. CCSBT-SC/0108/21

Appendix 1

Proportional Ageing:

The proportional ageing method takes the length of a fish and width of the length class (i.e. 1cm, 2cm, etc length classes) and where the length range (minimum to maximum lengths) cover more than 1 age class, partitions a fish into proportions in each age class.

Ages are calculated to be the age of the fish at the 1st January of the year. Cutpoints are linearly interpolated to adjust them for the time of year when the fish was caught. Middle of the time period is used (ie. for monthly data the middle of the month is used).
Minimum length and maximum length of a fish is calculated from the length and length class width. Weight measurements are first translated to lengths.

For example:

If the length of a fish is 100cm and the length class width (precision) is 1cm, then the length range is from 99.5 to 100.5 cm. If a cutpoint sits between within this range, then the catch (or single fish) is proportionally allocated to the two cohorts on either side of the cutpoint. If the cutpoint is at 100.2 cm, then 70% of the catch is allocated to the lower cohort and 30% to the upper cohort. In the case that more than 1 cutpoint is within the range of lengths then the catch is proportionally allocated to all of the cohorts.

Cohort Ageing:

Our method for cohort ageing is the same as the method described above, for adjustment of cutpoints for time of the year of the catch. Actual date of the catch or midpoint of the period is used. For the Cohort Age method, the width of the length class is not used. The catch is allocated to a single age class for the length given. There is no ageing across several age classes. This method is used where length class width is not known.