SC21 Length data explore for climate indicators

## Exploration of WCPFC longline length frequency data

This document extracts length frequency data from SPC databases to explore it in preparation for further analysis and potential use in SPC’s upcoming climate indicators report. Here, we focus on the longline data given the complexity and issues associated with the use and handling of purse seine length frequency data.

## Database extract

Below, is the database extract from the updated fishmaster SPC database.

# db1 <- "driver=SQL Server;server=NOUFAMESQL04;database=FISH\_MASTER"  
# channel <- odbcDriverConnect(db1)  
#   
# LF\_raw <-sqlQuery(channel,"  
# SELECT lf.ORIGIN\_ID,  
# ORIGINDESC,  
# YR,  
# QTR,  
# MON,  
# GR,   
# FLAG\_ID,  
# FLEET\_ID,  
# LAT\_SHORT,  
# CAST(LON\_SHORT as CHAR(3)) as lon\_c,  
# RIGHT(LON\_SHORT,1) as Lon\_h,  
# AREA\_ID,  
# TSTRAT,  
# ASTRAT,  
# SCHTYPE\_ID,  
# SP\_ID,  
# LEN,  
# LSTRAT,  
# FREQ  
# FROM lf.LF\_MASTER lf inner join ref.origin o on lf.origin\_id = o.origin\_id  
# WHERE sp\_id in ('BET', 'YFT', 'SKJ')  
# order by 1,2,3,4,5,6,7,8,9,10,14,16", max=0, stringsAsFactors=FALSE)   
# odbcCloseAll()  
  
#save(LF\_raw, file = paste0(data\_wd, "biology/bio\_len\_freqs\_raw\_07\_07\_25.Rdata"))

This data is then cleaned by the following filtering steps:

* WCPFC convention area
* Longline gear only
* BET and YFT only

## Exploratory plots

Note that length data is collected in 2cm size bins.

|  |
| --- |
| Raw length frequency data by species. |

### Year

We can see in @raw-len-by-yr, that the distribution and number of samples collected for both YFT and BET improves over time.

@raw-len-by-area-yr shows the spatial distribution of length samples across the Pacific Ocean where an upper limit of 10000 samples per grid cell was applied for improved plotting. We can see that sampling is not evenly distributed throughout the Pacific Ocean and especially in the early years is at a low resolution leading to the gridding of the data. This reflects different resolutions with which the data was collected. We can also see that in most years there is a high number of samples that come from very few grid cells. We can also see an improvement in the distribution of samples throughout the Pacific Ocean over time, showing improvements in observer and port sampling programs.

|  |
| --- |
| Raw length frequency data by year. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data by year. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data map by year. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data map by year. A) BET, B) YFT. |

### Sampling program

@len-origin-tbl shows the different programs and their sample sizes that have collected length sames for YFT and BET. In @raw-len-by-orig, we can see that some of these programs have more data than others, with some being very sparse. The distributions and temporal coverage of this data would likely need to be considered when modelling this data. This is especially apparent when we look at this data in space, @raw-len-by-area-origin shows strong temporal variability and resolution across the various sampling programs. The main source of length data include the SPC port sampling program (SPLL), Japan (JPLL), Korea (KRLL), Taiwain (TWLL) and Hawaii (HWOB).

| **Origin ID** | **Sampling program** | **Sample n** |
| --- | --- | --- |
| AUOB | AFMA (AFZ) Observer Data | 103,215 |
| CKOB | Cook Islands Observer Programme | 40,638 |
| CS | Coral Sea Tagging | 974 |
| FJCT | Fiji In-country Tagging | 4 |
| FJOB | Fiji Observer Programme | 257,440 |
| FMOB | FSM Observer Data | 52,098 |
| HWOB | Hawaii observer programme | 380,412 |
| HWTP | Hawaiian Tagging Project (PFRP) | 35 |
| JPLL | Japanese longline size data | 1,843,536 |
| KIOB | Kiribati Observer Programme | 55,421 |
| KRLL | Korean Longline Length data | 277,688 |
| MHOB | Marshall Islands Observer Data | 47,072 |
| NCOB | New Caledonia Observer Programme | 27,407 |
| NZOB | New Zealand Observer programme | 8,294 |
| PFOB | French Polynesia Observer Programme | 70,065 |
| PGOB | PNG Observer Data | 87,771 |
| PKSI | Indonesia - P4KSI samlpling under WPEA-OFM | 20,563 |
| PWOB | Palau Observer Programme | 3,828 |
| RTTP | Regional Tuna Tagging Project | 5 |
| SBOB | Solomon Islands Observer Data | 107,467 |
| SPLL | Regional Port sampling data - Longline | 3,743,767 |
| SPOB | SPC Observer Data | 16,245 |
| TOOB | Tonga Observer Programme | 112,038 |
| TWLL | Taiwanese Longline length data | 1,520,089 |
| VNWP | Vietnam -- WPEA Port sampling project -- LENGTHS | 100,463 |

|  |
| --- |
| Raw length frequency data by sampling program. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data by sampling program. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data mapped by sampling program. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data mapped by sampling program. A) BET, B) YFT. |

### Flag

@raw-len-by-flag shows the distribution of length frequencies by flag. Similar to above, length frequency distributions vary heavily across flag with some bimodal distributions present potentially suggesting changes in targeting.

|  |
| --- |
| Raw length frequency data by flag. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data by flag. A) BET, B) YFT. |

|  |
| --- |
| Violin plots of length frequncies by year and flag. A) BET, B) YFT. |

|  |
| --- |
| Violin plots of length frequncies by year and flag. A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data mapped by flag A) BET, B) YFT. |

|  |
| --- |
| Raw length frequency data mapped by flag A) BET, B) YFT. |

ORIGIN\_ID ORIGINDESC YR QTR MON GR FLAG\_ID  
1 ARTP Albacore Research Tagging Project size data 1990 4 12 P FJ  
2 ARTP Albacore Research Tagging Project size data 1990 4 12 P FJ  
3 ARTP Albacore Research Tagging Project size data 1990 4 12 P FJ  
4 ARTP Albacore Research Tagging Project size data 1990 4 12 P FJ  
5 ARTP Albacore Research Tagging Project size data 1990 4 12 P FJ  
6 ARTP Albacore Research Tagging Project size data 1990 4 12 T FJ  
 FLEET\_ID LAT\_SHORT lon\_c Lon\_h AREA\_ID TSTRAT ASTRAT SCHTYPE\_ID SP\_ID LEN  
1 <NA> 35S 170 E NZ M 5 SKJ 44  
2 <NA> 35S 170 E NZ M 5 SKJ 45  
3 <NA> 35S 170 E NZ M 5 SKJ 46  
4 <NA> 35S 170 E NZ M 5 SKJ 48  
5 <NA> 35S 170 E NZ M 5 SKJ 50  
6 <NA> 35S 170 E NZ M 5 SKJ 45  
 LSTRAT FREQ  
1 1 1  
2 1 1  
3 1 1  
4 1 1  
5 1 1  
6 1 2