Two-compartment models of Main Hawaiian Islands Yellowfin Tuna Population

Interim Report

John Sibert*
Joint Institute of Marine and Atmospheric Research
University of Hawai'i at Manoa
Honolulu, HI 96822 U.S.A.

March 25, 2015

Introduction

The Yellowfin Tuna (YFT) population in main Hawaiian Islands (MHI) is embedded in a larger pan-Pacific stock. Nevertheless, local fishermen believe that the MHI supports a "resident" yellowfin population. Some scientific observations are consistent with this belief. Recent tagging, tracking and studies show that the rate of exchange between the MHI population and the larger stock is low (Itano and Holland 2000). Analysis of YFT otoliths sampled from throughout the Pacific conclude that 90% or more of the MHI

*sibert@hawaii.edu

population was reared in the MHI (Wells et al 2012). The Hawaii based longline along with various small-scale fisheries land a combined catch approximately *???? mt* of YFT annually. Management of these fisheries is an important an important local issue deserving of scientific support.

This paper explores some potential models that might be used to inform development of options for the management of fisheries for YFT in the MHI.

The principle assumptions for modeling the MHI YFT population are:

1. The Pacific Ocean near Hawaii is divided into two regions: MHI (region 1) and elsewhere (region 2).
2. Fish emigrate from region 1 to region 2, but emigrant fish have no effect on region 2 population dynamics. Region 2 is an "infinite sink".
3. Fish immigrate from region 2 to region 1, and mix completely.
4. Immigrant fish are indistinguishable from "resident" fish (i.e. both groups of fish have the same population dynamics) and interact with the fishing gear identically.
5. Immigration into the MHI is dependent on the biomass of the yellowfin population outside of the MHI as estimated by some other model, e.g., MULTIFAN-CL or SEAPODYM.
6. The fishery comprises several gear types, each with characteristic fishing mortality.

7. The evolution of fishing mortality over time is a random walk where the current fishing mortality in the current time step is equal to the fishing mortality in the previous time step plus a random deviation.

Data

Several sources of data used are used in this analysis:

1. Yellowfin catch weights reported to the Hawaii Department of Aquatic Resources (HDAR) from 1949 to 2014.

2. Longline yellowfin catch weights reported to the National Oceanic and Atmospheric Administration, National Marine Fisheries Service (NOAA) under the federally mandated log book program from 1995 through 2013.

3. Measurements of average yellowfin fish weights sampled at the Honolulu auction by NMFS staff from 2000 through 2013.

4. Model estimates of yellowfin biomass by MULTIFAN-CL for regions 2 and 4 from the most recent WCPFC stock assessment (Davies et al 2014).

All data are reported by quarter of the year.

Catch Time Series

The HDAR data comprise catch reports for the following gear categories: “Aku boat”, “Bottom/inshore HL”, “Longline”, “Troll”, “Tuna HL”, “Casting", "Hybrid", "Shortline", "Other", and "Vertical line". For this analysis catches by "Casting", "Hybrid", "Shortline", "Other", and "Vertical line" are combined into a new category, "Misc.". The "Misc" catches are highest after year 2000, but comprise a very small proportion of the catch. The catch time series for the HDAR data are shown in Figure 1. These time series exhibit marked quarterly cycles suggesting a strong seasonal signal in the catches by all fleets. Some series also contain sustained periods of zero catch which document the development and subsequent shift away from a specific gear type. It is assumed that these declines in catches represent a "collapse" of a fishery due more to social and economic factors than to a decline of stock. Some time series are punctuated by brief, one or two quarters in length, of zero catches. Again it is assumed that these zero catches are not caused by low stock levels. The "Aku Boat" fishery was a small pole and line fishery targeting skipjack tuna in Hawaii, for which yellowfin was an incidental catch.

The longline fishery has changed drastically over time. In the late 1940s, it was a relatively small fishery using traditional Okinwawan-style gear *(reference)* , usually labeled "flag line". Participation in this fishery generally declined from 1950 to late 1980s, as can be seen in the time series plots. In the 1990s, the longline fishery expanded rapidly with the introduction of US fishing boats from the Atlantic ocean using modern monofilament longline gear and deeper sets.

Figure 1 also displays the first order differences between successive quarters. This statistic emphasizes the seasonal periodicity of all fishing fleets and helps to identify potential anomalies in the data such as changes in reporting protocol. Potential anomalies can be seen in the late 1980s in the Tuna HL fleet, the early 1990s in the longline fleet, and the mid 1950s in the Aku Boat fleet.

The partial autocorrelations within each time series are shown in Figure 2. These correlations confirm the quarterly periodicity of the catch time series. The catch time series for Inshore HL,Longline, Troll, and Tuna HL show similar patters with significant positive autocorrelations at lags of 1,3 and 4 quarters. The autocorrelation pattern appears some what different for the Aku boat catch time series.

NOAA began to collect data from the longline fleet under a federally mandated logbook program in 1990, and in 1995 the deep and shallow setting are distinguished in the data. The HDAR data does not distinguish between deep and shallow sets. Figure 3 shows the correspondence between the HDAR and NOAA time series. The combined deep plus shallow catches from NOAA line up fairly well with the overlaping HDAR data. The simple average of the HDAR data with the combined NOAA deep plus shallow data appears to have iroughtly the same autocorrelation structure as the constituent time series, Figure 4. *Data are also available from NOAA for the period 1990-1995, but have not yet been included in this analysis.*

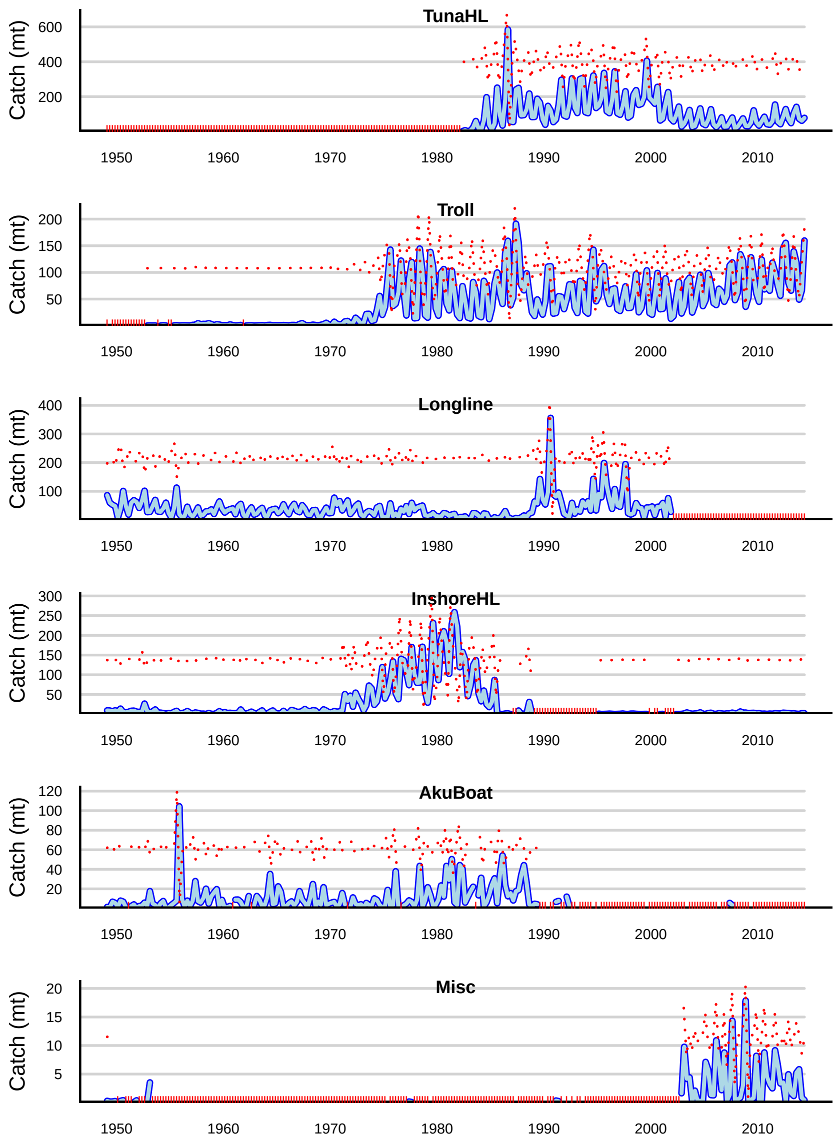

Figure 1: Yellowfin catch in metric tonnes by principle fisheries operating in the Main Hawaiian Islands from the HDAR data. The dotted red line superimposed on each time series is the difference in catch between successive quarters. The red tick marks on the abscissa indicate quarters where catches were zero.

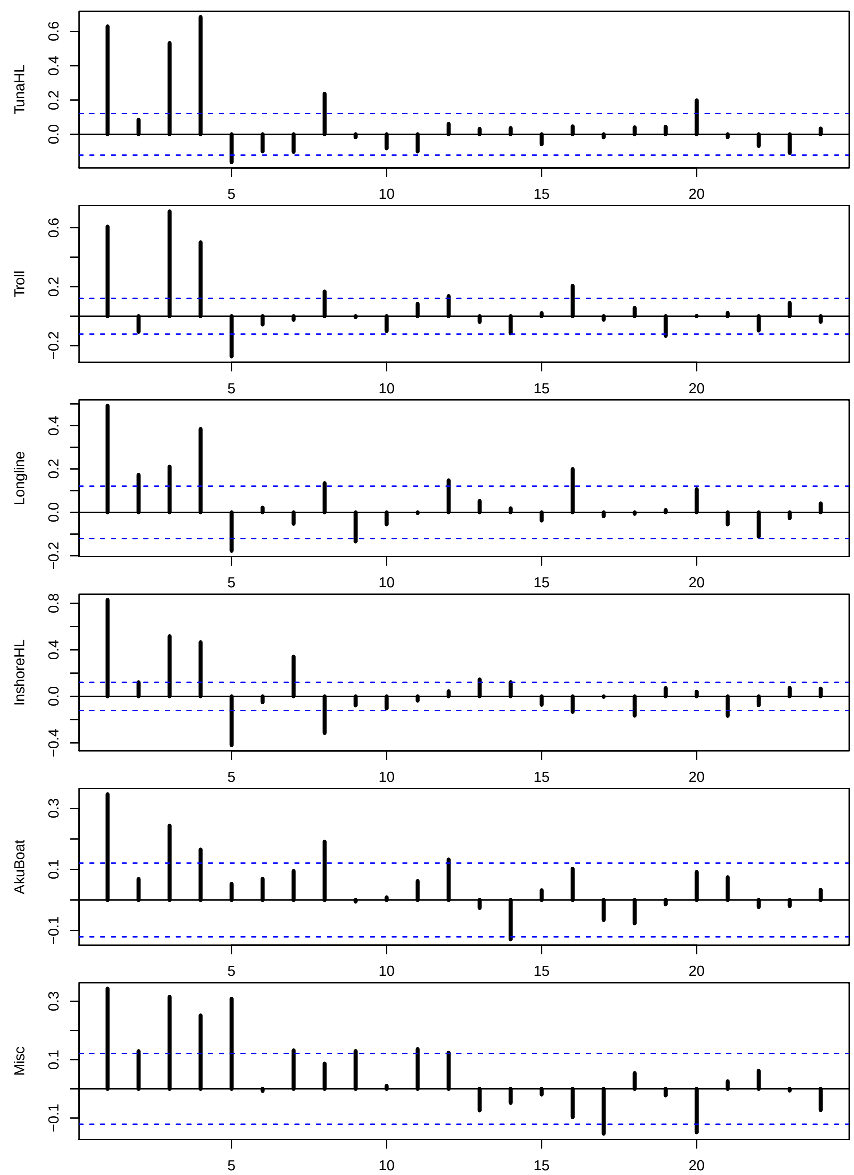

Figure 2: Partial autocorrelation coefficients of the catch time series. The dashed blue lines indicate approximate 95% confidence limits of the correlations.

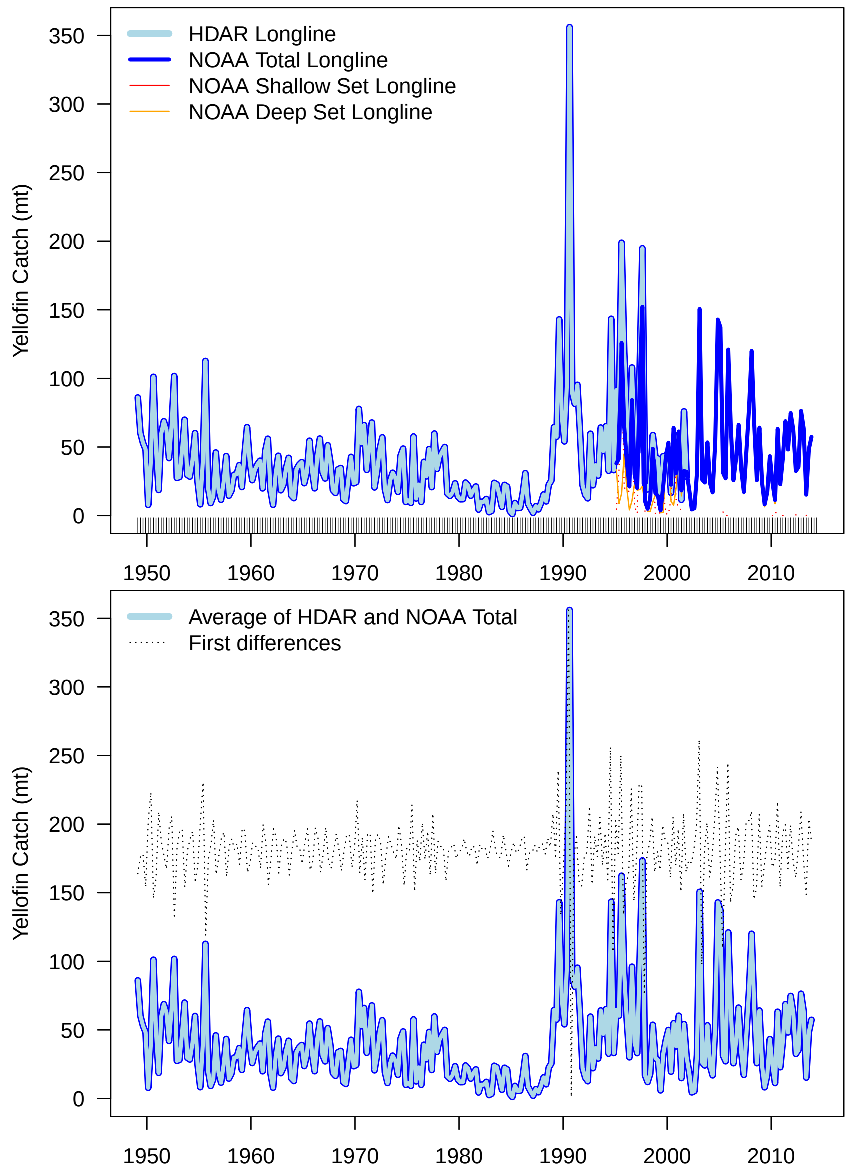

Figure 3: Comparison between HDAR and NOAA longline time series. The upper panel shows the NOAA deep and shallow set data superimposed on the HDAR data. The lower panel shows the time series produced by a simple average of the HDAR data and the sum of the NOAA deep and shallow catches.

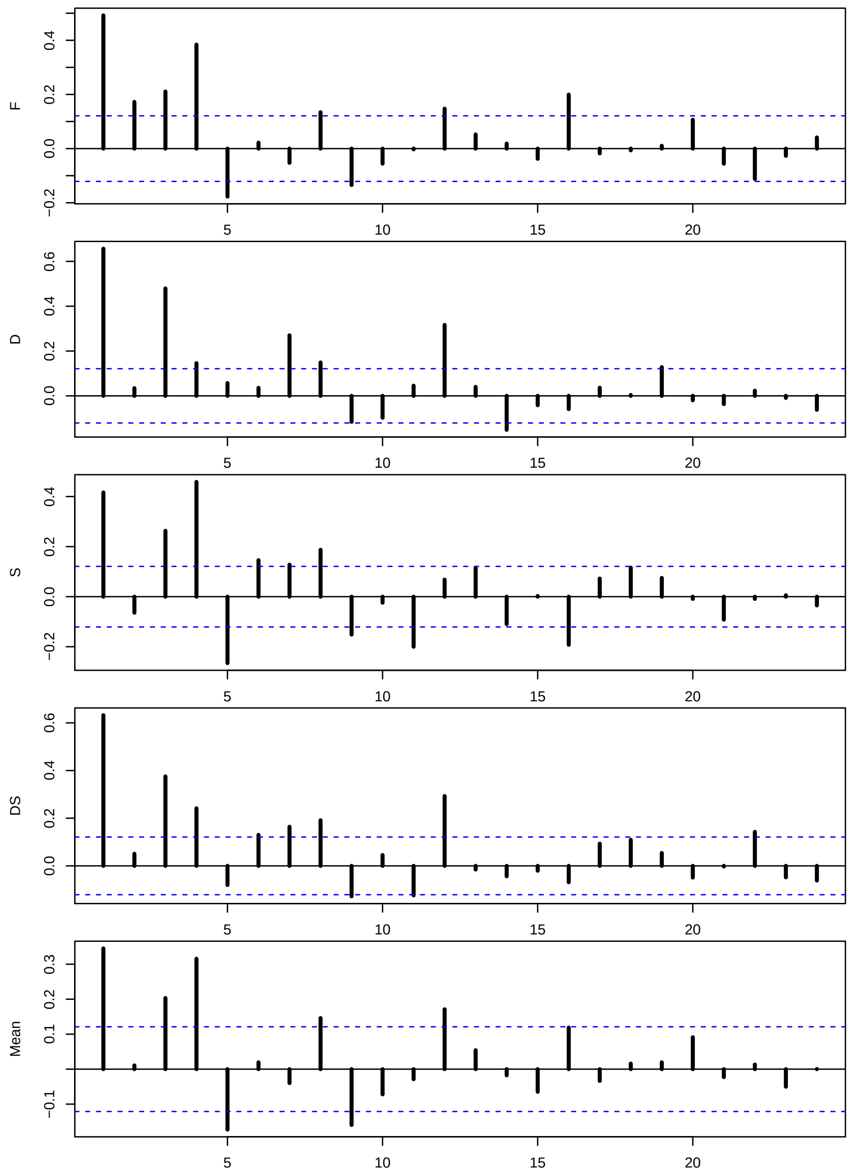

Figure 4: Partial autocorrelations within the NOAA and combined time series. F refers to the HDAR longline data, D refers to the NOAA deep set data, S to the shallow set data, DS to the combined deep and shallow set data, and Mean to the average of the HDAR and NOAA deep plus shallow.

Fish Weight Time Series

Models

The general modeling approach is to develop a state-space model similar to the that utilized by Nielsen and Berg (2014). State-space models separate variability in the biological processes in the system (transition model) from errors in observing features of interest in the system (observation model). The general form of the transition equation is

$$\alpha_t = T(\alpha_{t-1}) + \eta_t \tag{1}$$

where the function T embodies the stock dynamics mediating the development of the state at time t from the state at the previous time with random process error, η. The general form of the observation equation is

$$x_t = O(\alpha_t) + \varepsilon_t \tag{2}$$

where the function O describes the measurement process and with error ε

Age-aggregated model: logistic population dynamics

Let $N_{1,1}$ equal the biomass of fish originating in region 1 and residing in region 1 and $N_{2,1}$ equal the biomass of fish originating in region 2 but residing in region 1. The total biomass of fish residing in region 1 is $N_{1,1} + N_{2,1}$, and the dynamics of the population in region 1 is represented as a modified Schaefer Model:

$$\frac{d}{dt}\big(N_{1,1} + N_{2,1}\big) = \big(N_{1,1} + N_{2,1}\big)\Big[r\Big(1 - \frac{N_{1,1} + N_{2,1}}{K}\Big) - F - T_{12}\Big] + T_{21} \tag{3}$$

where r is the per capita logistic growth rate per year, K is the asymptotic biomass, F is the total fishing mortality in region 1, T_{12} is the emigration rate from region 1 to region 2, and T_{21} is the annual rate of immigration of biomass from region 2 to region 1.

An equivalent differential equation could be devised for the dynamics of fish residing in region 2 (i.e., $\frac{d}{dt}\big(N_{2,2} + N_{1,2}\big)$), but the dynamics of the fish population in region 2 is external to this model. T_{21} can be considered to be form of population "forcing" by the larger stock in which the MHI population is embedded. T_{21} could a prediction from other models, such as MULTIFAN-CL or SEAPODYM, in a sort of "off-line" coupling, or perhaps represented by an autocorrelated stochastic process.

The appearance of $N_{1,1} + N_{2,1}$ in the numerator of the logistic term reflects the assumption that the population dynamics immigrant fish depends on the population dynamics in region 1. This assumption leads to an important non-linearity in the model that predicts the possibility of overwhelming of the local stock by a more numerous immigrant stock as is often characteristic of mixed-stock fisheries.

The proportion of "local" fish $\frac{N_{1,1}}{N_{1,1}+N_{2,1}}$ is of potential interest. Equation 3 can be expanded and rearranged to become

$$\begin{aligned} \frac{d}{dt}\big(N_{1,1} + N_{2,1}\big) &= N_{1,1}\Big[r\Big(1 - \frac{N_{1,1}}{K}\Big) - F - T_{12}\Big] \\ &+ N_{2,1}\Big[r\Big(1 - \frac{N_{2,1}}{K}\Big) - F - T_{12}\Big] + T_{21} \qquad (4) \\ &- 2r\frac{N_{1,1}N_{2,1}}{K} \end{aligned}$$

The non-linear term, $2r\frac{N_{1,1}N_{2,1}}{K}$, in equation 4 represents the reduction in biomass of one population by the presence of the other population. In order to represent the two populations by separate equations, this nonlinear term must be appropriately apportioned. A new parameter, q, is introduced to accomplish the apportionment, leading to a set of simultaneous non-linear differential equation for the components of the population inhabiting region 1.

$$\begin{aligned}
\frac{dN_{1,1}}{dt} &= N_{1,1}\Big[r\Big(1-\frac{N_{1,1}}{K}\Big)-F-T_{12}\Big]-(1-q)2r\frac{N_{1,1}N_{2,1}}{K} \\
\frac{dN_{2,1}}{dt} &= N_{2,1}\Big[r\Big(1-\frac{N_{2,1}}{K}\Big)-F-T_{12}\Big]-q2r\frac{N_{1,1}N_{2,1}}{K}+T_{21}
\end{aligned} \tag{5}$$

Does this system have any meaningful equilibrium?

The log transformed equivalent of equation 5 is

$$\begin{aligned}
\frac{d\log(N_{1,1})}{dt} &= r\Big(1-\frac{N_{1,1}}{K}\Big)-F-T_{12}-(1-q)2r\frac{N_{2,1}}{K} \\
\frac{d\log(N_{2,1})}{dt} &= r\Big(1-\frac{N_{2,1}}{K}\Big)-F-T_{12}-q2r\frac{N_{1,1}}{K}+\frac{T_{21}}{N_{2,1}}
\end{aligned} \tag{6}$$

Transition Equation $T(\alpha_{t-1})$. The state space transition equation for the two components of the MHI population is developed by solving equations 6 by finite differences approximation with explicit time stepping and adding process error terms.

$$
\begin{aligned}
\log N_{1,1_t} &= \log N_{1,1_{t-\Delta t}} \\
&+ \Delta t\left(r\left(1 - \frac{N_{1,1_{t-\Delta t}}}{K}\right) - F_{t-\Delta t} - T_{12} - (1-q)2r\frac{N_{2,1_{t-\Delta t}}}{K}\right) + \eta_{1,t}
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
\log N_{2,1_t} &= \log N_{2,1_{t-\Delta t}} \\
&+ \Delta t\left(r\left(1 - \frac{N_{2,1_{t-\Delta t}}}{K}\right) - F_{t-\Delta t} - T_{12} - q2r\frac{N_{1,1_{t-\Delta t}}}{K} + \frac{T_{21}}{N_{2,1_{t-\Delta t}}}\right) + \eta_{2,t}
\end{aligned}
$$

where $\eta_{k,t} \sim N(0, \Sigma_\eta), k = 1, 2$. The covariance matrix, $\Sigma_\eta = \left[\begin{array}{cc} \sigma^2_{\eta,1} & \rho_\eta \\ \rho_\eta & \sigma^2_{\eta,2} \end{array}\right]$, expresses individual process errors for both population as well as a correlation between the errors.

Fishing Mortality. Fishing mortality is modeled explicitly as a random walk without recourse to effort standardization and estimation of catchability coefficients. The logarithm of fishing mortality is assumed to follow a random walk with normal increments, i.e.

$$
\log F_{g,t} = \log F_{g,t-1} + \xi_{g,t}; \qquad \xi_{t,g} \sim N(0, \sigma^2_{\xi,g}) \tag{8}
$$

where $\sigma^2_{\xi,g}$ is the variance of the fishing effort random walk for fleet g. *Distribution of increments?*

Population Forcing T_{21}. Immigration of stock from region 2 to region 1 is a form of forcing whereby events outside of the model influence the model dynamics. The term T_{21_t} is the biomass immigrating into the MHI stock at each time step. For the model under development I assume that source of immigrants is MFCL region 2, so that

$$
T_{21_t} = T^*_{21} \cdot B_{2,t} \tag{9}
$$

where $B_{2,t}$ is biomass in region 2 at time t taken from the MFCL output file `plot-12.par.rep` from the 2014 WCPFC stock assessment (Davies et al 2014). T^*_{21} is the proportion on $B_{2,t}$ which migrates at each time step. Alternatively $B_{2,t}$ may be assumed to be constant at the average of the estimate biomass in region 2, $\overline{B_2}$.

Parameterization of K. K is the asymptotic biomass in a population growing according to logistic dynamics. It is the population size to which the population tends at it equilibrium approaches and is often dubbed carrying capacity. Its main role in a logistic population is to scale the population size. It is not clear that a coupled logistic model such as 3 has a non-trivial equilibrium. Furthermore the $B_{2,t}$ is time dependent, Figure ??. Several alternative parameterizations of K can be envisaged, for example, constant tat the maximum of $B_{2,t}$, an unknown parameter, or even a random walk. For testing purposes I have assumed the maximum of $B_{2,t}$.

Given that the model forcing $B_{r,t}$ is not constant and that the fishery is developing in a period of profound change in the oceans, it is difficult to justify the assumption of constant K. An alternative parameterization for K is to assume a random walk, for example,

$$\log K_t = \log K_{t-1} + \omega_t; \qquad \omega_t \sim N(0, \sigma^2_\omega) \tag{10}$$

where σ^2_ω is the characteristic variance of the assymptoc population size random walk.

Observation Equation, $O(\alpha)$. Predicted catch in region 1 is the product of fishing mortality and the total population in region, $N_{1,1} + N_{2,1}$. Thus the state-space observation model predicting catch in the region 1 under this model is

$$\log C_{g,t} = \log\Bigg(F_{g,t} \cdot \Big(\frac{N_{1,1_{t-\Delta t}} + N_{1,1_t}}{2} + \frac{N_{2,1_{t-\Delta t}} + N_{2,1_t}}{2}\Big)\Bigg) + \varepsilon_{t,g} \qquad (11)$$

where the total population in region 1 is the sum of the average population over the time step (Quinn and Deriso, 1999), and $\varepsilon_{t,g} \sim N(0, \sigma^2_{\varepsilon,g})$.

Constraint on "Proportion Local". Proportion local is defined as $p = \frac{N_{1,1}}{N_{1,1}+N_{2,1}}$. The logit transformed proportion local, $L(p)$, is assumed to be normally distributed around a mean value, $\bar{p}$.

$$L(p) \sim N(L(\bar{p}), \sigma^2_{L(p)}). \qquad (12)$$

p is generally assumed to be approximately 0.9. By varying the value of the variance, $\sigma^2_{L(p)}$, p, can be made as close to 0.9 as is prudent. In principle, it may be possible to estimate $\bar{p}$ and $\sigma^2_{L(p)}$, but interaction with the parameter q in equation 5 needs to be explored.

Estimation. The model states, α_t, are assumed to be random effects (Skaug and Fournier, 2006). Model parameters are estimated by maximum likelihood by defining the joint likelihood of the random effects and the observations.

$$L(\theta, \alpha, x) = \prod_{t=2}^{m} \left[\phi\big(\alpha_t - T(\alpha_{t-1}), \Sigma_\eta\big)\right] \prod_{t=1}^{m} \left[\phi\big(x_t - O(\alpha_t), \Sigma_\varepsilon\big)\right] \qquad (13)$$

Here, m is the number of time steps in the catch time series and ϕ is a vector of model parameters. A complete list of parameters is found int Table **??**. The actual number of parameters to be estimated depends on the model

Table 1: Model variables.

Variable	Definition
m	Number of quarterly time steps
n	Number of fishing fleets
r	Instantaneous growth rate (q^{-1})
K	Asymptotic biomass (mt)
T_{12}	Emigration rate (q^{-1})
T^*_{21}	Immigration rate (q^{-1})
$\sigma_{1,\eta}, \sigma_{2,\eta}$	Population growth SD
ρ_η	Correlation between population growth precess errors
$\sigma_{\xi,g}\ g = 1 \ldots n$	Fishing mortality random walk SD
$c_g\ g = 1 \ldots n$	Proportion of F random walk contaminated by fat-tailed distribution
$\sigma_{\varepsilon,g}\ g = 1 \ldots n$	Observation error SD
$\bar{p}$	Mean proportion local
$\sigma_{L(p)}$	SD logit transformed p
q	Nonlinear apportionment proportion

configuration. The model is implemented in AD Model Builder (Fournier et al 2012).

Age-structured models

Results & Discussion

Next steps

Acknowledgements. This work was funded by the Western Pacific Regional Fisheries Managment Council. I thank the Council for its generous support and Council Staff Paul Dalzell and Eric Kingma for encouraging me

Table 2: Model "estimates". Initial values of model parameters and final values just prior to program exit with "Matrix not positive definite in Ln_det_choleski" message.

Vatiable	Initial Value	Final Value
m	244	
n	5	
r	0.100	0.292
K	144,800	—
T_{12}	0.018	0.0116
T^*_{21}	0.018	0.00171
$\sigma_{1,\eta}$	0.378	0.120
$\sigma_{2,\eta}$	0.378	0.216
ρ_η	0.0	—
$\sigma_{\xi,1}$	0.368	0.368
$\sigma_{\xi,2}$	0.368	0.368
$\sigma_{\xi,3}$	0.368	0.368
$\sigma_{\xi,4}$	0.368	0.368
$\sigma_{\xi,5}$	0.368	0.368
$\sigma_{\varepsilon,1}$	1.051	1.042
$\sigma_{\varepsilon,2}$	1.051	1.045
$\sigma_{\varepsilon,3}$	1.051	1.045
$\sigma_{\varepsilon,4}$	1.051	1.045
$\sigma_{\varepsilon,5}$	1.051	1.050
$\bar{p}$	0.9	—
$\sigma_{L(p)}$	2.72	—
q	0.54	—

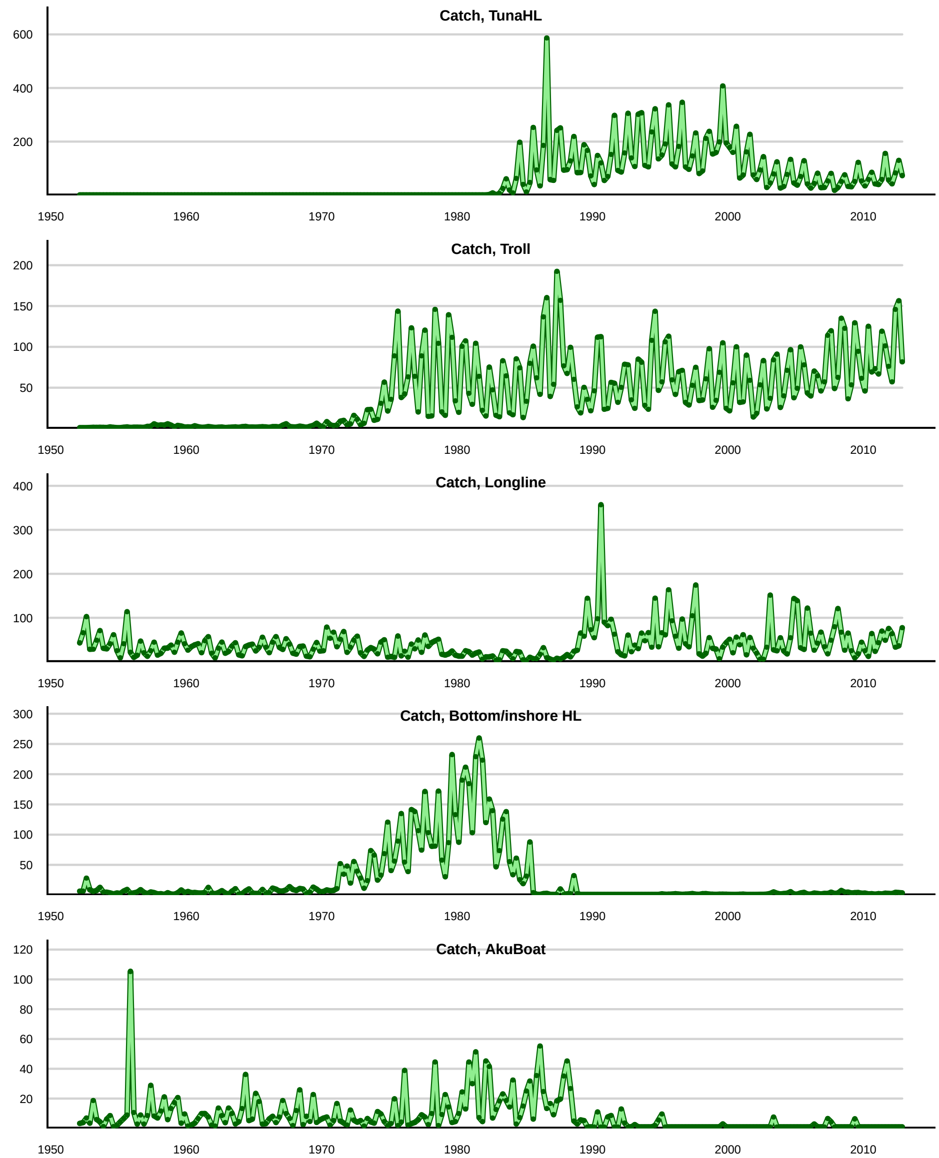

Figure 5: Estimated catch by fleet. The dark green dots indicate observed catch and light green lines indicate predicted catch,

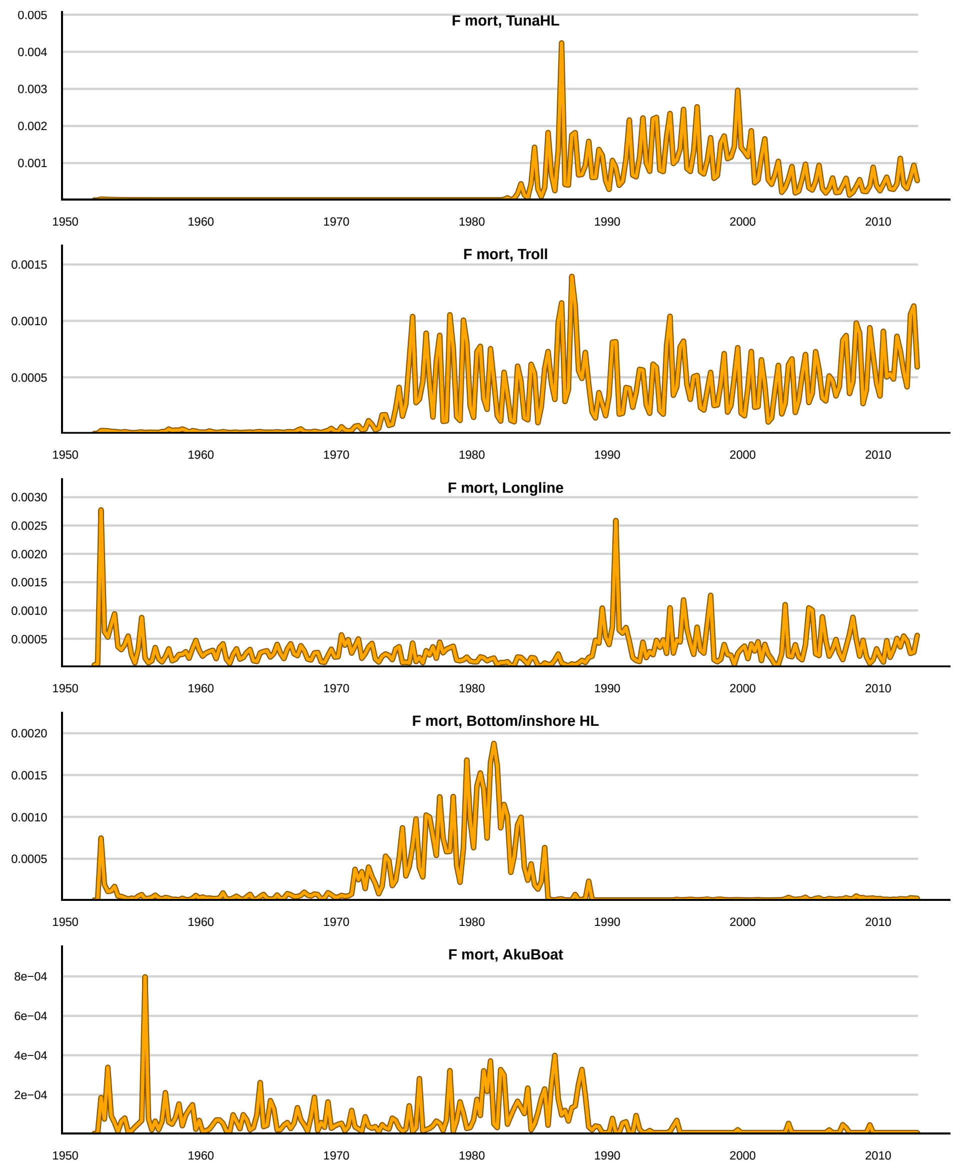

Figure 6: Estimated fishing mortality by fleet.

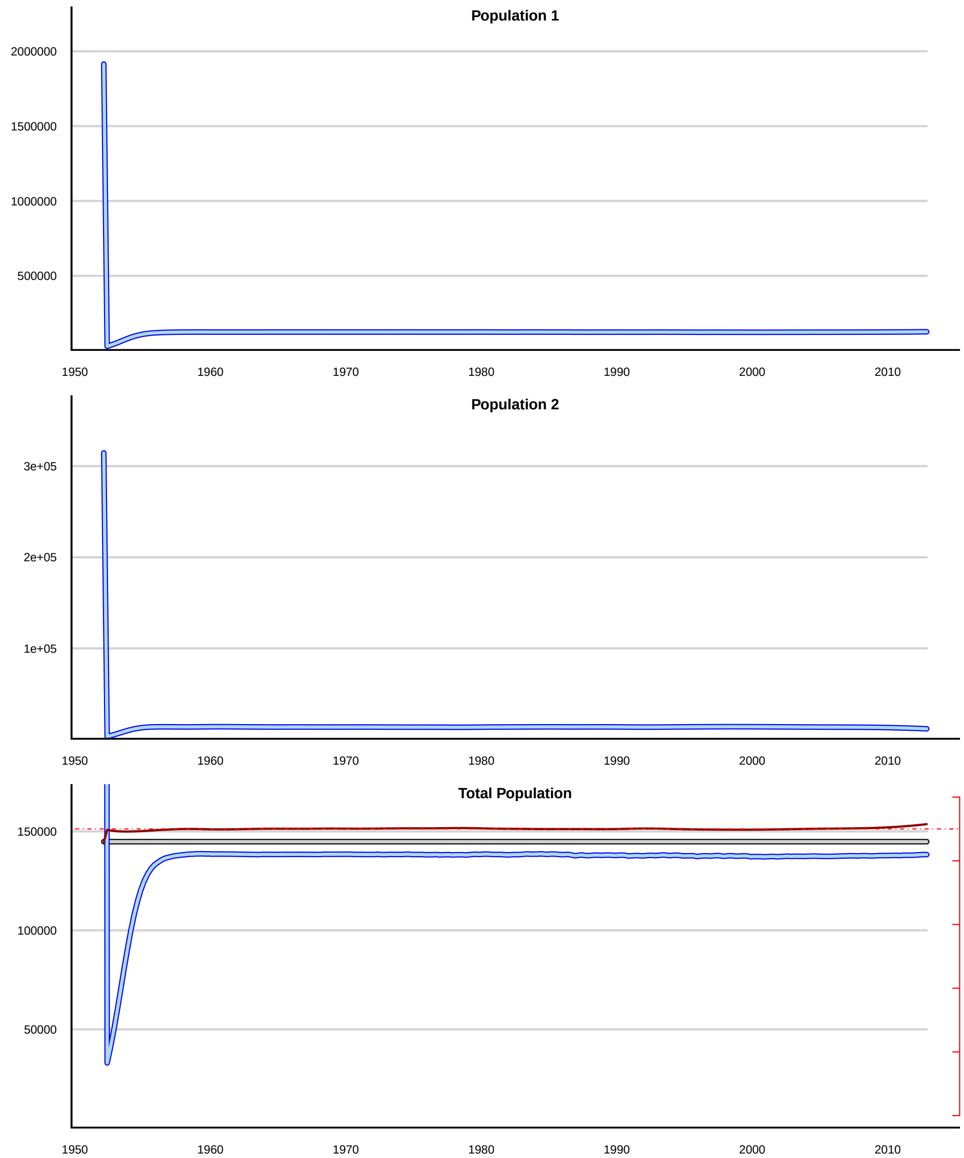

Figure 7: Estimated biomass. Light blue lines indicate predicted $N_{1,1}$ (upper panel), $N_{2,1}$ (middle panel) and $Nsum$ bottom panel. Red line in bottom panel indicates proportion local; dashed red line indicates $p = 0.9$. Grey line in bottom panel is K.

to actually take on this challenging project and for their on-going collaboration. Martha Maciasz, student intern at the Council, assisted with processing weight frequency data. Thanks to Mr. Reginald Kokubun of the Hawaii Division of Aquatic Resources for supplying catch report data from the HDAR commercial fisheries data base. Thanks to Mr. Keith Bigelow and Ms. Karen Sender of NOAA Pacific Island Fisheries Science Center for supplying logbook reporting data and weight-frequency data from the PIFSC data base. Thanks also to Dr. John Hampton of the Secretariat of the Pacific Community, Oceanic Fisheries Programme, for making available MULTICAN-CL output files from the latest Western and Central Pacific Fisheries Commission yellowfin tuna stock assessment, and to Mr. Nick Davies for sharing R scripts and advice to decode the MFCL output files.

References

Davies, N., S. Harley, J. Hampton, S. McKechnie. 2014. Stock assessment of yellowfin tuna in the western and central pacific ocean. WCPFC-SC10-2014/SA-WP-04.

Fournier, D.A., Skaug, H.J., Ancheta, J., Sibert, J., Ianelli, J., Magnusson, A., Maunder,M.N., Nielsen, A., 2012. AD Model Builder: using automatic differentiation for forstatistical inference of highly parameterized complex nonlinear models. Opti-mization Methods and Software 27, 233249.

Itano, D., K. Holland. 2000. Movement and vulnerability of bigeye (Thunnus obesus) and yellowfin tuna (Thunnus albacares) in relation to FADs and natural aggregation points.Aquat. Living Resour. 13: 213-223.

Kleiber, P., J. Hampton, N. Davies, S. Hoyle, D. Fournier. 2014. MULTIFAN-CL Users Guide

Nielsen, A. and C. Berg. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158:96-101.

Quinn, T and R. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Skaug, H., Fournier, D., 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. Computational Statistics & Data Analysis 51, 699709.

Wells, D., J. Rooker, D. Itano. 2012. Nursery origin of yellowfin tuna in the Hawaiian Islands. Mar. Ecol. Prog. Ser. 461:187-196.