Stock Assessment Models for the Main Hawaiian Islands Yellowfin Tuna Fishery

John Sibert[*]

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

October 9, 2015

Abstract

The yellowfin tuna fishery in the Main Hawaiian Islands (MHI) is an important commercial and recreational fishery and has a documented 65 year history. Local management of this fishery is hampered by lack of stock assessment results specific to the MHI. The feasibility of developing state-space Schafer model with ties to the larger Pacific yellowfin population, external forcing, Bayesian constraints, and robust estimation specifically for the MHI was investigated. The models are feasible, but parameter estimation is hindered by possible lack of an exploitation signal in the data.

1 Introduction

The Main Hawaiian Islands (MHI) yellowfin tuna, *Thunnus albacares*, (YFT) fishery is a small, but very important important fishery within the State of Hawaii, and sustainable management of this fishery is a local issue that deserves scientific support. The documented history of the YFT fishery begins

---

[*]sibert@hawaii.edu

in 1949, predating statehood by 10 years, and is possibly a longer catch history than that of the Japanese distant water longline fishery. The MHI YFT population is, of course, embedded in the much larger pan-Pacific stock, which also has a long exploitation history and is the subject of on-going stock assessment efforts by the Western and Central Pacific Fisheries Commission (WCPFC). Many local fishermen in Hawaii believe that the MHI supports a "resident" yellowfin population which spawns and resides in the MHI. Some scientific observations are consistent with this belief. Recent tagging and tracking 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 population was spawned and reared in the MHI (Wells et al 2012).

Contemporary fisheries management often depends on results of stock assessment models to establish quantitative fishery reference points such as "acceptable biological catch". Such models do not exist for the MHI YFT population, and it is not clear that creating stock assessment model for a single component of a shared stock is feasible. The purpose of the work described in this report is to asses the feasibility of developing stock assessment models for the MHI YFT that link the MHI population to the larger Pacific population.

2 Data

Two basic sources of catch data were used in this analysis: quarterly time series of YFT catch weights reported to (1) Hawaii Department of Aquatic Resources (HDAR) for several fishing gears from 1949 to 2014 and (2) to the National Oceanic and Atmospheric Administration (NOAA), National Marine Fisheries Service for the Hawaii-based longline fleet from 1995 through 2013 under the federally mandated log book program. The quarterly data from HDAR show clear seasonal variation and include occasional zero observations between quarters with substantial reported catch (Figure A.1). Whether these seasonal fluctuations and zero catch observations represent changes in abundance, changes in catchability, or changes in fishery participation is unknown. In preliminary tests, the model appeared to interpret catch variability as variability in stock abundance predicting large fluctuations in biomass around the isolated zero observations. Therefore, the quarterly data were aggregated by year and used for all further model runs; see
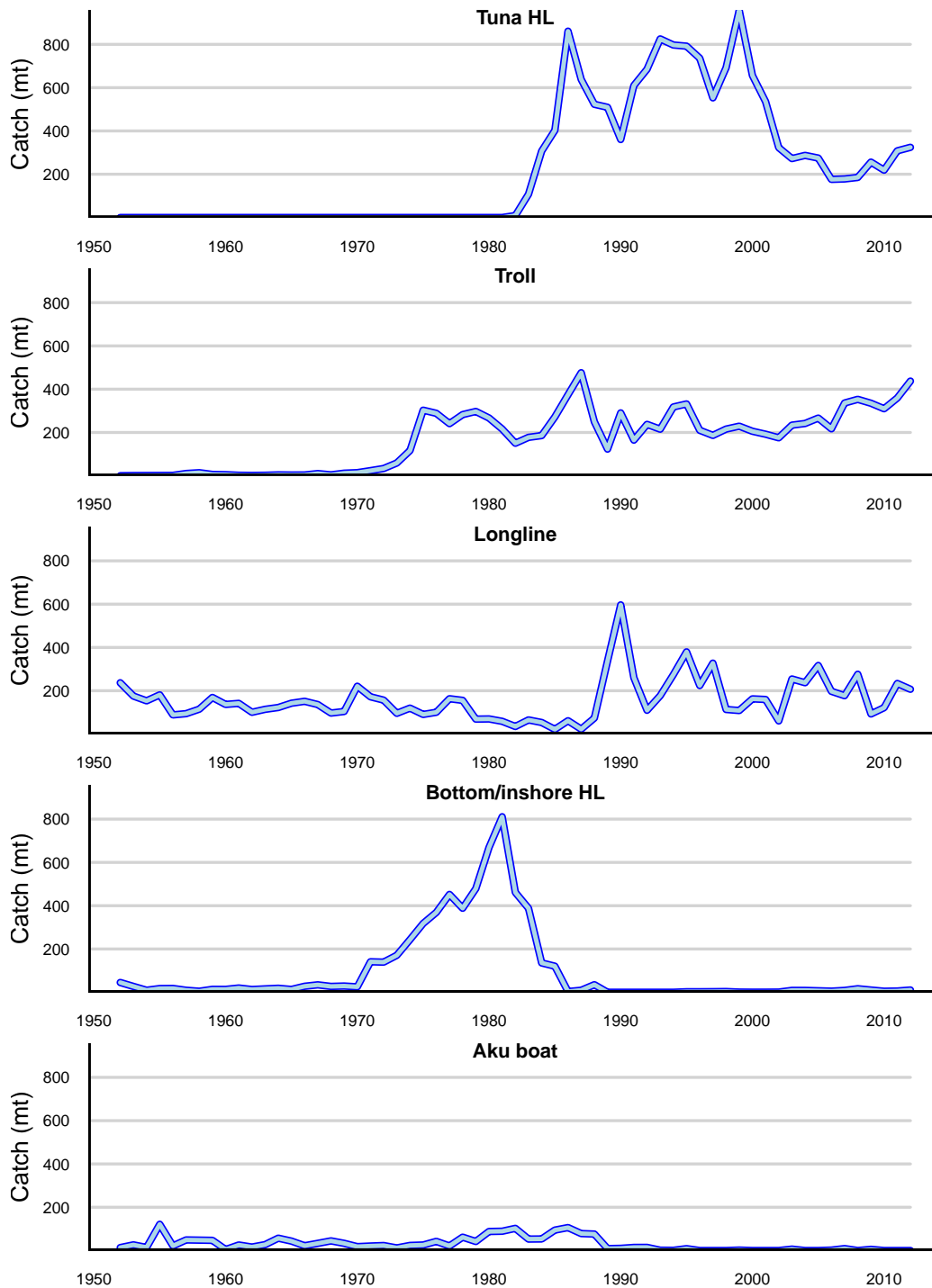
Figure 2.1: Reported annual yellowfin catch in metric tonnes by fisheries operating in the Main Hawaiian Islands from combined HDAR and NOAA data.
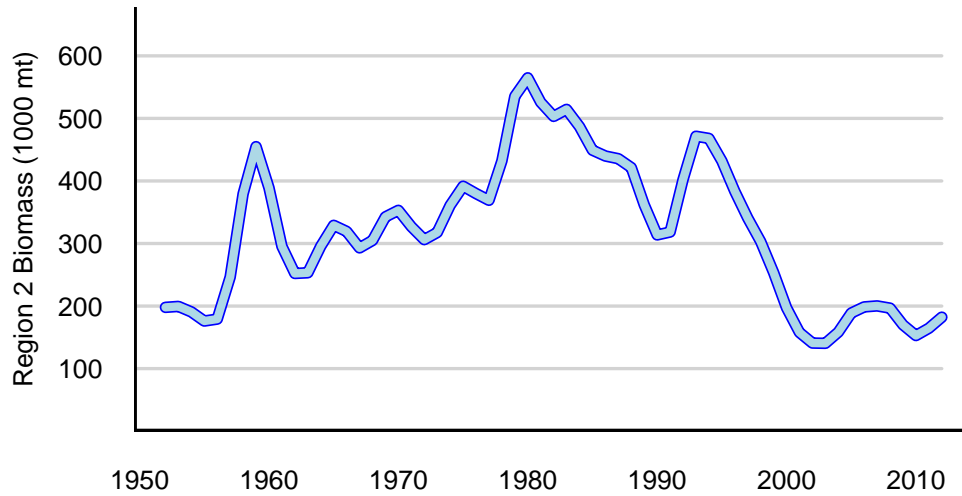
Figure 2.2:   Annual estimated biomass trends in MFCL Region 2 from the 2014 WCPFC YFT stock assessment, Davies et al. 2014.

Figure 2.1. Statistical properties of the catch data and the method of joining the HDAR and NOAA longline time series are elaborated in Appendix A.

More background and discussion of data deficiencies is needed here, e. g. the "ahi" problem, HDAR reporting requirements, changes in the database, and "recreational" data.

A biomass exchange model presupposes a source source of biomass. Model estimates of yellowfin biomass by MULTIFAN-CL (MFCL) from the most recent WCPFC stock assessment (Davies et al 2014) provide convenient estimates of immigrant biomass. The MHI straddle the boundary between MFCL regions 2 and 4. Region 2 is arguably more similar ecologically to the MHI than the more equatorial region 4, and the population in region 4 is much larger and exhibits a greater level of depletion than the population in region 2, Figure A.6. The quarterly estimates for region 2 were aggregated by year and used as the immigrant source for model testing, Figure 2.2. Since

the MFCL assessment uses data for the period 1952 through 2012, catch data outside of this period were excluded from this analysis.

Further properties of the data are presented in Appendix A.

3 Models

Two similar state-space models, both derived from a basic Schaefer biomass dynamics model (Schaefer, 1954), were tested. 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

$$\alpha_t = T(\alpha_{t-1}) + \eta_t \tag{3.1}$$

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

$$x_t = O(\alpha_t) + \varepsilon_t \tag{3.2}$$

where the function O describes the measurement process with error ε in observing the population. Age-aggregated models were developed in preference to a more complex age-structured model for examining the feasibility of a MHI YFT assessment model.

Fishing mortality in both models is treated explicitly as a random walk without recourse to effort standardization, estimation of catchability coefficients, or creation of an index of abundance based on catch per unit of effort (CPUE). The logarithm of fishing mortality is assumed to follow a random walk with normal increments, i.e.,

$$\log F_{g,t} = \log F_{g,t-1} + \xi_t; \quad \xi_t \sim N(0, \sigma_F^2) \tag{3.3}$$

where σ_F^2 is the variance of the fishing mortality random walk.

3.1 Two populations with explicit exchange

The principle model assumptions are:

1. The Pacific Ocean is divided into two regions: MHI (region 1) and elsewhere (region 2).

2. Fish immigrate from region 2 to region 1, and mix completely. Immigrant fish are indistinguishable from "local" fish, and the two components interact identically with the fishing gear in region.

3. Fish emigrate from region 1 to region 2, but emigrant fish have no effect on region 2 population dynamics, i.e., region 2 is an infinite sink. Emigration is considered to be a source of mortality analogous to "natural" mortality.

4. 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 (MFCL) or SEAPODYM. In the present model, MFCL region 2 is assumed to be the source of immigrants (see previous paragraph).

5. The fishery comprises several gear types or fleets, each characterized by a distinct fishing mortality time series. Fishing practices have evolved continuously over the 66 year history of the fishery making estimation of satisfactory measures of fishing effort difficult and contentious.. The evolution of fishing mortality for each fleet over time is represented by a random walk, i.e., the fishing mortality in the current time step is equal to the fishing mortality in the previous time step plus a random deviation.

6. Ninety percent of the fish resident in region 1 are assumed to be locally spawned and reared (Wells et al 2012). This assumptions requires that the model to be able to track the size of both the local and the immigrant subpopulations separately so that the proportion local can be computed. It is further assumed that this ratio has been nearly invariant since 1952.

Full derivations of and all model equations arising from these assumptions are presented in Appendix B. The resulting model is something like a state-space Schafer model with emigration, external forcing, Bayesian constraints, and robust estimation. 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 following pair of simultaneous non-

linear differential equations describe the dynamics of the two subpopulations.

$$
\begin{aligned}
\frac{dN_{1,1}}{dt} &= N_{1,1}\left[r\left(1 - \frac{N_{1,1}}{K}\right) - F - T_{12}\right] - (1-q)2r\frac{N_{1,1}N_{2,1}}{K} \\
\frac{dN_{2,1}}{dt} &= N_{2,1}\left[r\left(1 - \frac{N_{2,1}}{K}\right) - F - T_{12}\right] - q2r\frac{N_{1,1}N_{2,1}}{K} + T_{21}
\end{aligned}
\tag{3.4}
$$

where r is the 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 rate of immigration of biomass from region 2 to region 1. The nonlinear mortality term $2r\frac{N_{1,1}N_{2,1}}{K}$ is a consequence of assuming two subpopulations constrained by the same logistic process. The parameter, q, apportions the effects of the nonlinear mortality term between the two subpopulation.

3.2   Single population with indexed abundance

The principle model assumptions are:

1. The dynamics of the population of YFT in the MHI follows a simple Schaefer model is MHI-specific growth parameters.

2. The local dynamics are "forced" by assuming that the local abundance is approximately proportional to the abundance of the larger Pacific population.

3. Fishing mortality is represented by a random walk (as in the two-population model).

Thus the local dynamics follow the classic Schaefer form:

$$
\frac{dN}{dt} = rN(1 - \frac{N}{K}) - FN
\tag{3.5}
$$

where N is the biomass of YFT in the MHI, r is the logistic growth rate per year, K is the asymptotic biomass, and F is the total fishing mortality per year in the MHI.
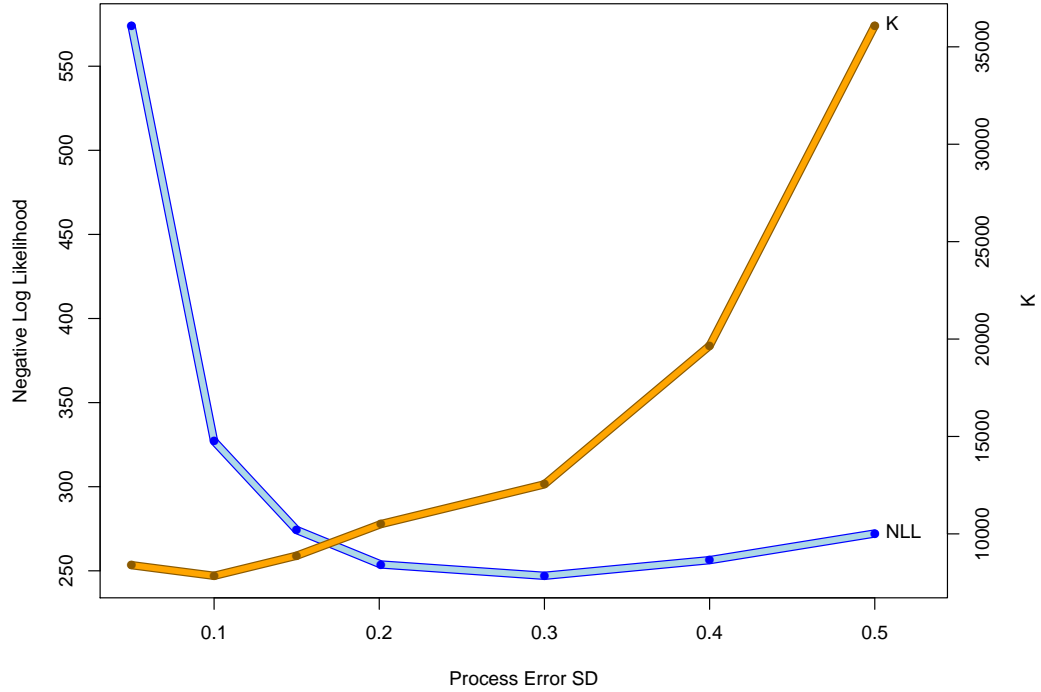
Figure 4.1: catches.

4   Results

4.1   Two populations with explicit exchange

Preliminary results of the two population model were presented in a previous report. More work has been done on model interpretation and these previous results need revision.

4.2   Single population with indexed abundance

The complete list of paramters to be estimated in this model are presented in Table B.2. Thre of these parameters are standard deviations of the state space transition equation process errors, σ_N, σ_F, and σ_Q. It proved impossible to estimate all three of these standard deviations simultaneously. Therefor estimations were carried out for a range of fixed values of these parameters. The results are shown in Figure

Table 4.1: Parameter estimates from the single population model.

Variable	Definition	
r	Instantaneous growth rate	$0.4y^{-1}$ (fixed)
K	Asymptotic biomass	12500 mt
Q	Abundance index proportionality constant	0.00329
p_0	Proportion of zero catch observations	$p_0 = 0.15738$ (fixed)
σ_N	Population growth SD	0.3 (fixed)
σ_F	Fishing mortality random walk SD	0.3 (fixed)
σ_Q	Abundance index proportionality SD	0.3 (fixed)
σ_Y	Observation error SD	0.0452

5 Discussion

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., H.J. Skaug, J. Ancheta, J.Sibert, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen. 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., C. Berg. 2014. Estimation of time-varying selectivity in stock assessments using state-space models. Fisheries Research 158:96-101.

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

Schaefer, M. B. 1954. Some aspects of the dynamics of populations important to the management of the commercial marine fisheries. IATTC Bull. 1(2):27-56.

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.

Appendix

A   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 A.1. These time series exhibit marked annual cycles suggesting a strong seasonal signal in the catches by all gears. 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 episodes (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 longline fishery has changed drastically over time. In the late 1940s, it was a relatively small fishery using traditional tarred rope gear making relatively shallow sets. Participation in this fishery generally declined from 1950 to late 1980s. In the 1990s, the longline fishery expanded rapidly with the introduction of US fishing boats from the Atlantic ocean using modern monofilament longline gear capable of making both deep and shallow sets.

Figure A.1 also displays the first order differences between successive quarters. This statistic emphasizes the seasonal periodicity of all fishing gear and helps to identify potential anomalies in the data such as changes in reporting protocols.

The partial autocorrelations within each time series are shown in Figure A.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 to be somewhat 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, began to distinguish deep

and shallow sets in the data. The HDAR data do not distinguish between deep and shallow sets. Figure A.3 shows the correspondence between the HDAR and NOAA time series. The combined deep plus shallow catches from NOAA align fairly well with the overlapping HDAR data. The simple average of the HDAR data with the combined NOAA deep plus shallow data appears to have roughly the same autocorrelation structure as the constituent time series, Figure A.4. Data are also available from NOAA for the period 1990-1995, but have not yet been included in this analysis.

Histograms of first differences of the logarithm of catch time series are shown in Figure A.5. These differences are clearly not normally distributed, indicating that a robust catch likelihood might be useful.
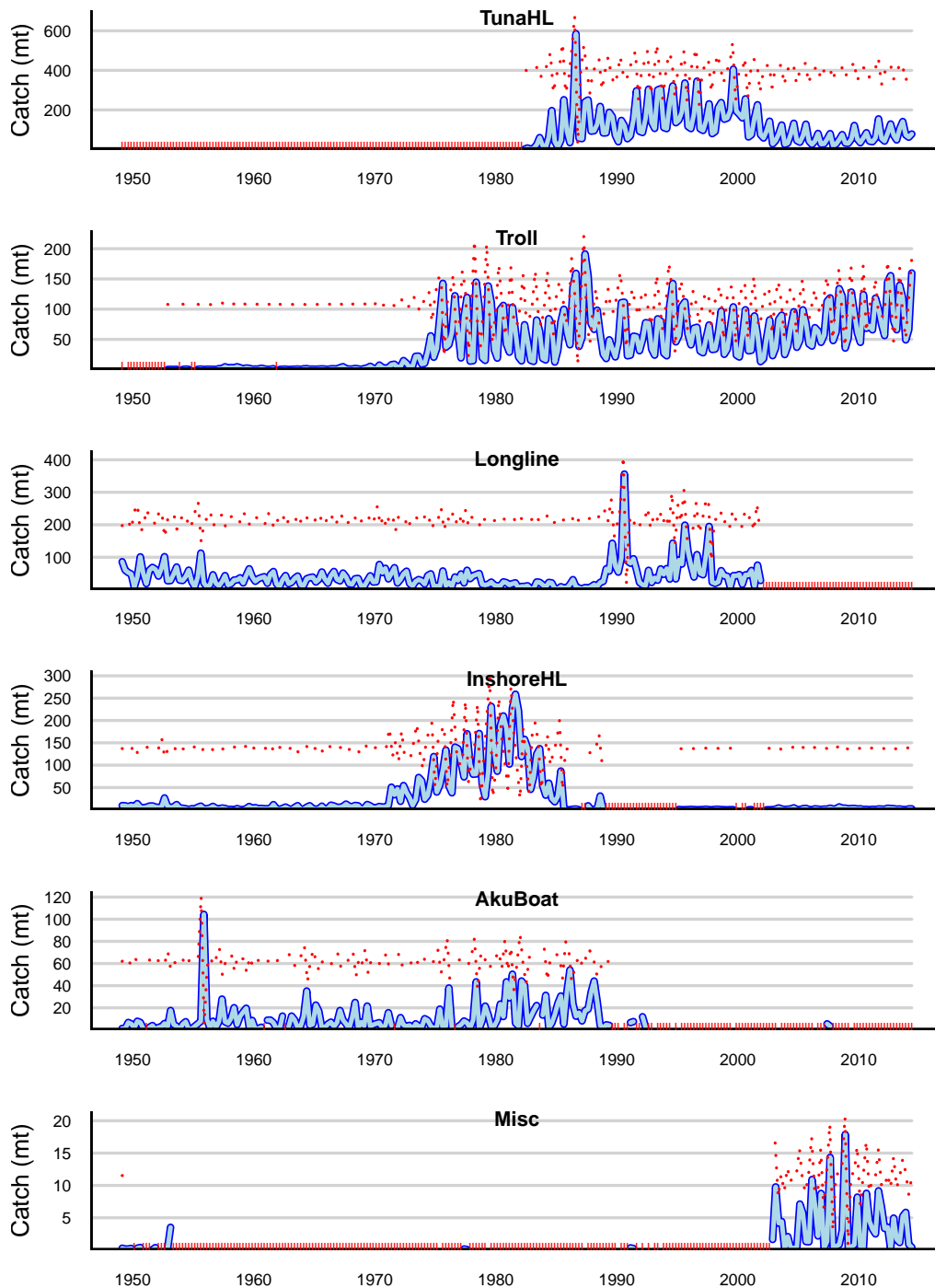
Figure A.1: Yellowfin catch in metric tonnes by principle fisheries operating in the Main Hawaiian Islands from the HDAR data. The dotted red line superimposed on each time series is the difference in catch between successive quarters. The red tick marks on the abscissa indicate quarters where reported catches were zero.
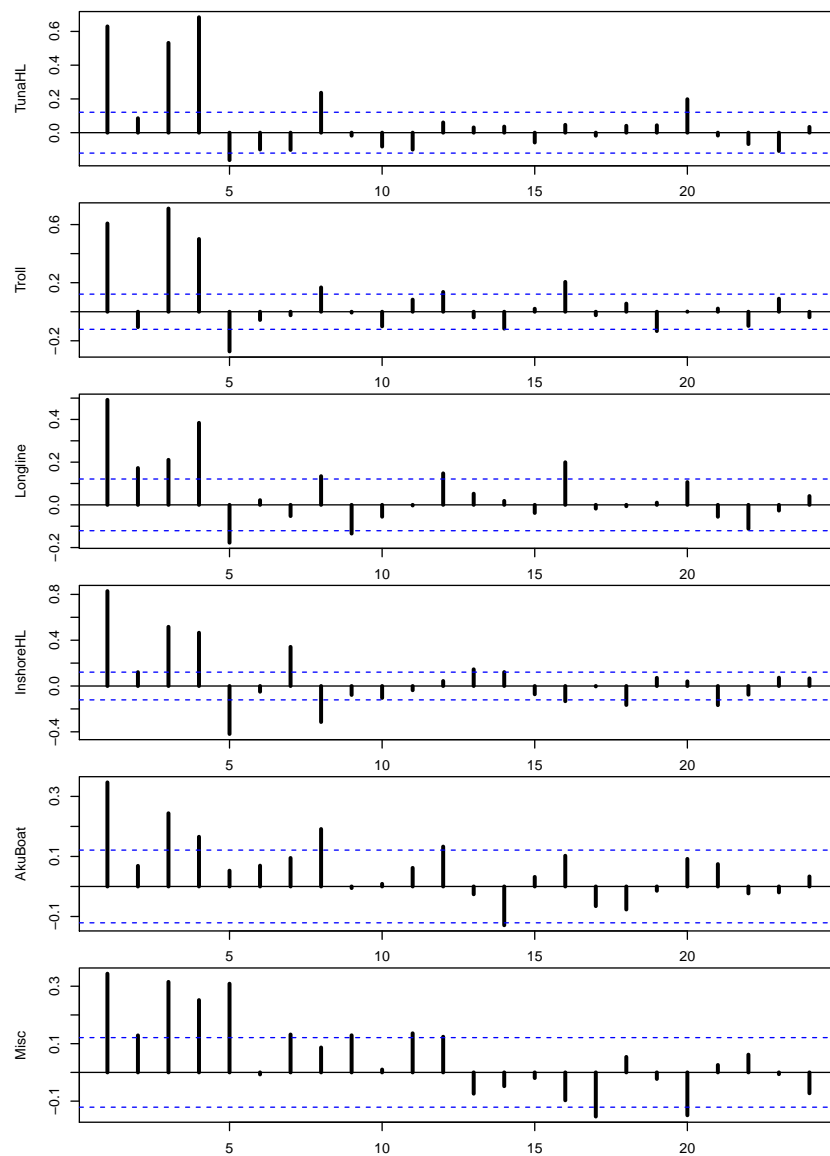
Figure A.2:   Partial autocorrelation coefficients of the HDAR catch time series. The dashed blue lines indicate approximate 95% confidence limits of the correlations.
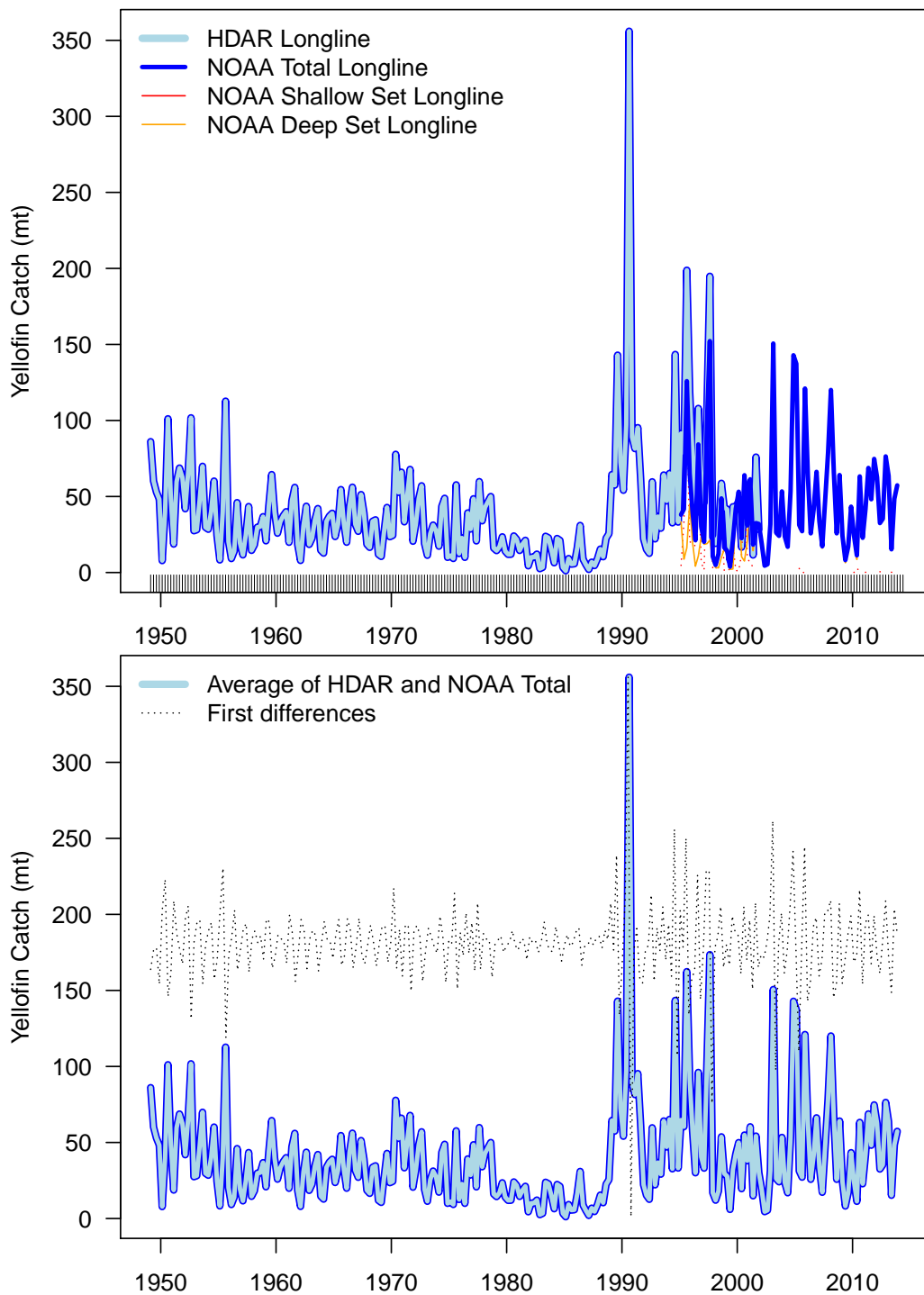
Figure A.3: Comparison between HDAR and NOAA longline time series. The upper panel shows the NOAA deep and shallow set data superimposed on the HDAR data. The lower panel shows the time series produced by a simple average of the HDAR data and the sum of the NOAA deep and shallow catches.
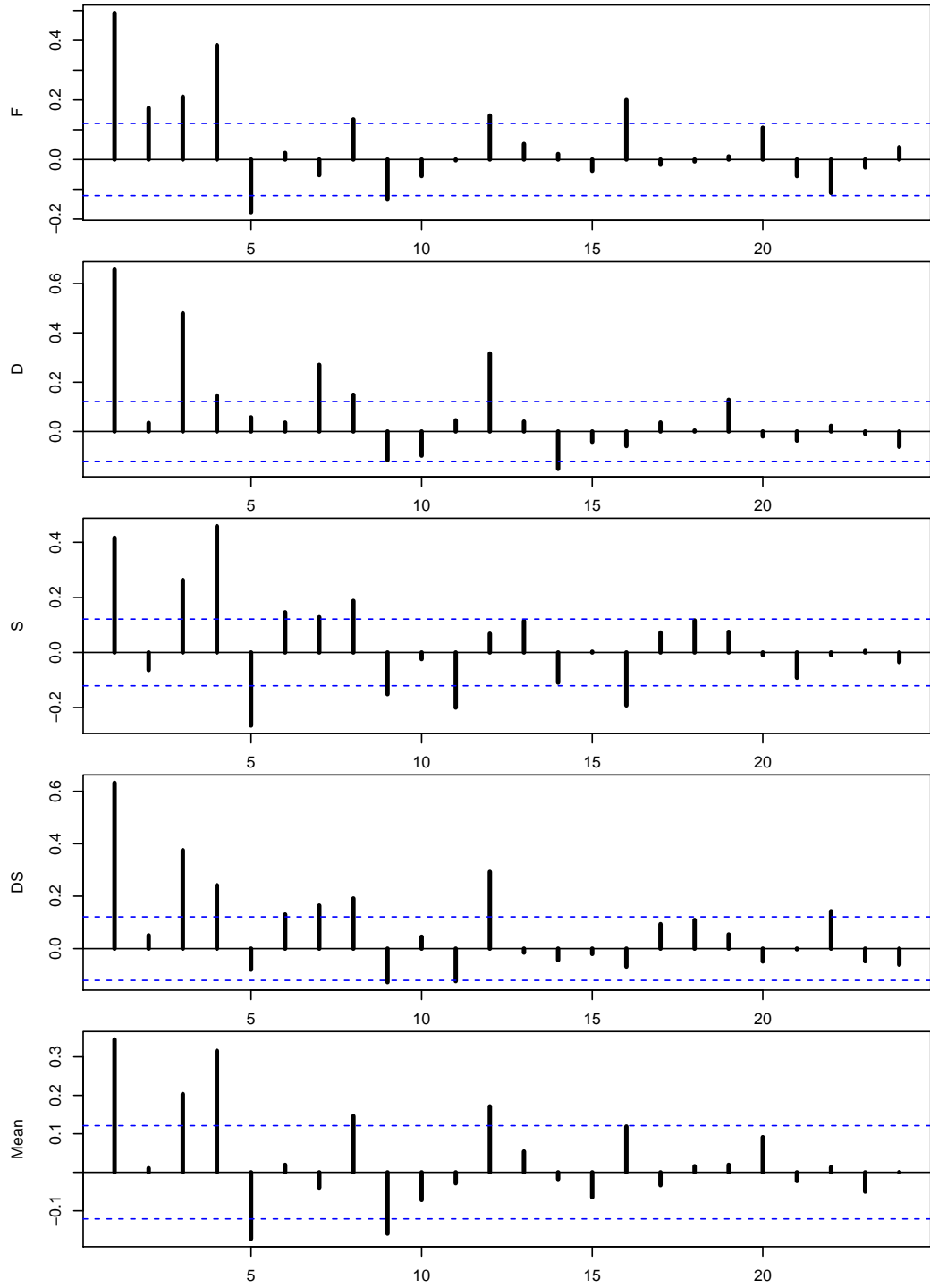
Figure A.4: Partial autocorrelations within the NOAA and combined time series. F refers to the HDAR longline data, D refers to the NOAA deep set data, S to the shallow set data, DS to the combined deep and shallow set data, and Mean to the average of the HDAR and NOAA deep plus shallow.

Figure A.5: Histograms of first differences of the logarithm of catch time series. The blue line is a normal distribution with the same mean and standard deviation as the first differences. The red line is the equivalent standard Cauchy distribution.

Region 2

Region 4

Figure A.6: Quarterly biomass trends estimated in MFCL Regions 2 and 4 from the 2104 WCPFC YFT stock assessment, Davies et al. 2014.
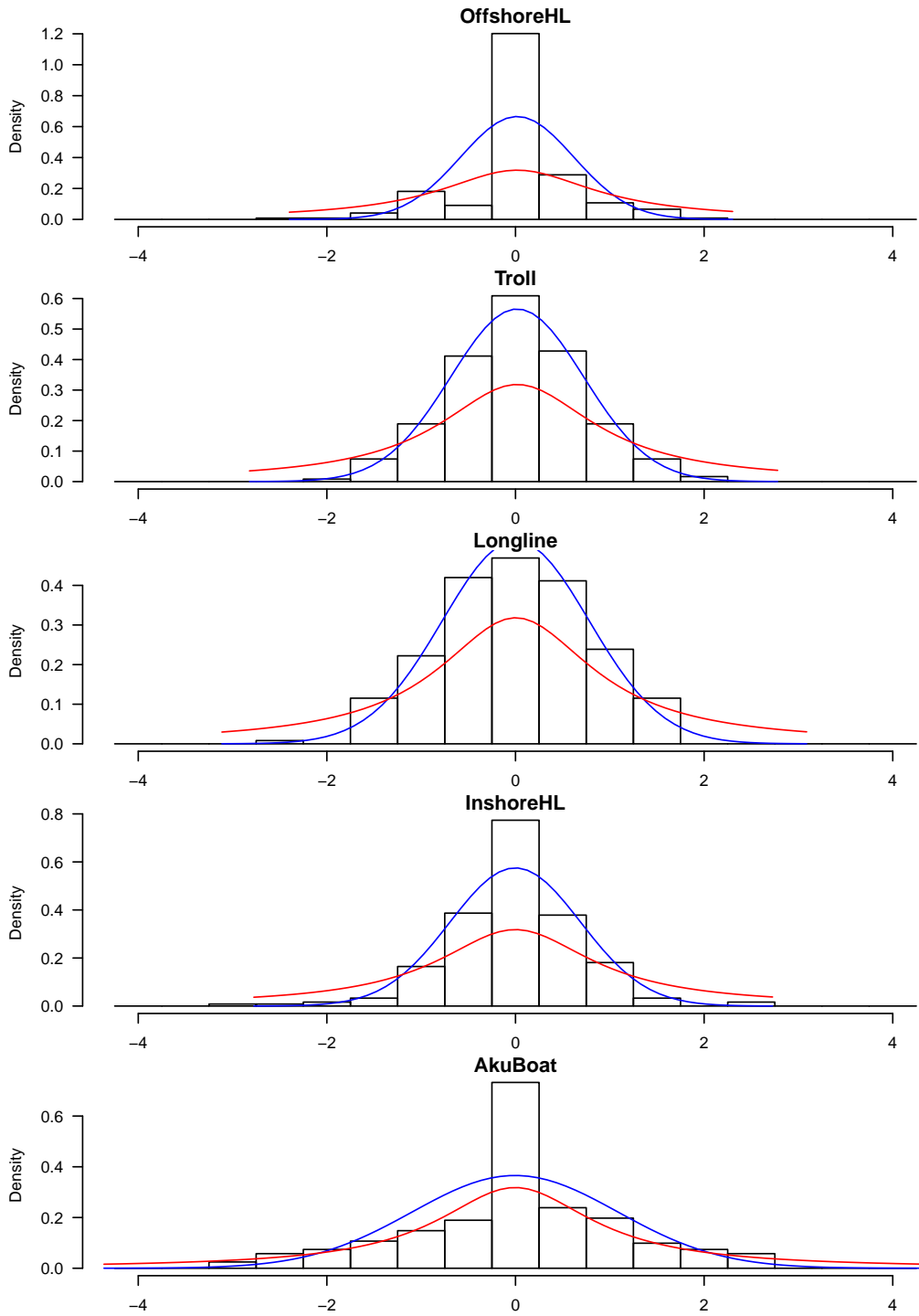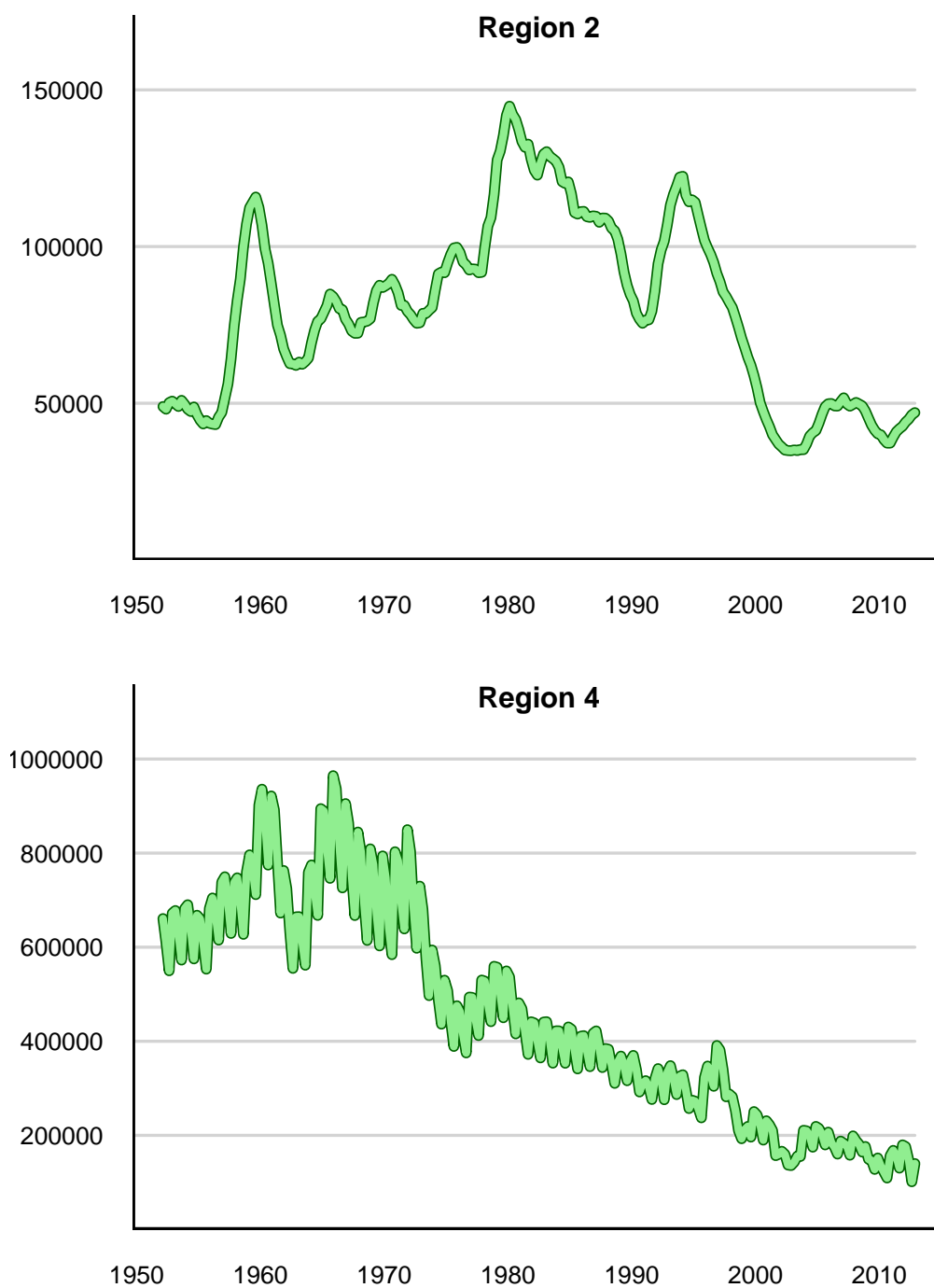
B   Model development

B.1   Two populations with explicit exchange

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:

$$\frac{d}{dt}\left(N_{1,1} + N_{2,1}\right) = \left(N_{1,1} + N_{2,1}\right)\left[r\left(1 - \frac{N_{1,1} + N_{2,1}}{K}\right) - F - T_{12}\right] + T_{21} \quad \text{(B.1)}$$

where r is the 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 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., $\frac{d}{dt}\left(N_{2,2} + N_{1,2}\right)$), but the dynamics of the fish population in region 2 is external to this model. T_{21} can be considered to be a form of population "forcing" by the larger stock in which the MHI population is embedded. T_{21} could be a prediction from other models, such as MULTIFAN-CL or SEAPODYM, in a sort of "off-line" coupling.

The appearance of $N_{1,1} + N_{2,1}$ in the numerator of the logistic term reflects the assumption that the population dynamics of immigrant fish depends on the population dynamics in region 1. This assumption leads to an important non-linearity in the model that predicts potential overwhelming of one stock by the other.

The proportion of "local" fish $p = \frac{N_{1,1}}{N_{1,1}+N_{2,1}}$ is an important constraint, and it is necessary to model the two subpopulations separately. Equation B.1 can be expanded and rearranged to become

$$\begin{aligned}
\frac{d}{dt}\left(N_{1,1} + N_{2,1}\right) &= N_{1,1}\left[r\left(1 - \frac{N_{1,1}}{K}\right) - F - T_{12}\right] \\
&+ N_{2,1}\left[r\left(1 - \frac{N_{2,1}}{K}\right) - F - T_{12}\right] + T_{21} \quad \text{(B.2)} \\
&- 2r\frac{N_{1,1}N_{2,1}}{K}
\end{aligned}$$

The non-linear term, $2r\frac{N_{1,1}N_{2,1}}{K}$, in equation B.2 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 equations for the components of the population inhabiting region 1.

$$
\frac{dN_{1,1}}{dt} = N_{1,1}\left[r\left(1 - \frac{N_{1,1}}{K}\right) - F - T_{12}\right] - (1-q)2r\frac{N_{1,1}N_{2,1}}{K}
$$

$$
\frac{dN_{2,1}}{dt} = N_{2,1}\left[r\left(1 - \frac{N_{2,1}}{K}\right) - F - T_{12}\right] - q2r\frac{N_{1,1}N_{2,1}}{K} + T_{21}
$$

(B.3)

The log transformed equivalent of equation B.3 is

$$
\frac{d\log(N_{1,1})}{dt} = r\left(1 - \frac{N_{1,1}}{K}\right) - F - T_{12} - (1-q)2r\frac{N_{2,1}}{K}
$$

$$
\frac{d\log(N_{2,1})}{dt} = r\left(1 - \frac{N_{2,1}}{K}\right) - F - T_{12} - q2r\frac{N_{1,1}}{K} + \frac{T_{21}}{N_{2,1}}
$$

(B.4)

Transition Equation $T(\alpha_{t-1})$. The state space transition equation for the two components of the MHI population is developed by solving equations B.4 by finite differences with explicit time stepping and adding process error terms.

$$
\log N_{1,1_t} = \log N_{1,1_{t-\Delta t}}
$$
$$
+ \Delta t\left(r\left(1 - \frac{N_{1,1_{t-\Delta t}}}{K}\right) - \sum_{g=1}^{n} F_{g,t-\Delta t} - T_{12} - (1-q)2r\frac{N_{2,1_{t-\Delta t}}}{K}\right) + \eta_t
$$

(B.5)

$$
\log N_{2,1_t} = \log N_{2,1_{t-\Delta t}}
$$
$$
+ \Delta t\left(r\left(1 - \frac{N_{2,1_{t-\Delta t}}}{K}\right) - \sum_{g=1}^{n} F_{g,t-\Delta t} - T_{12} - q2r\frac{N_{1,1_{t-\Delta t}}}{K} + \frac{T_{21_{t-\Delta t}}}{N_{2,1_{t-\Delta t}}}\right) + \eta_t
$$

where $\eta_t \sim N(0, \sigma_N)$ expresses process errors for both populations.

Behavior of model system. Efforts to derive an analytical expression to determine a meaningful equilibrium for equations B.4 have not been completely fruitful (Section C). Figure B.1 shows phase plane diagrams for

several combinations of model parameters assuming constant K and T_{21}. Apparently the system tends towards equilibrium with both $0 < N_{1,1} \leq K$ and $0 < N_{2,1} \leq K$ for some combinations of model parameters. The proportion of "local" fish, $p = \frac{N_{1,1}}{N_{1,1}+N_{2,1}}$, depends on the assumed value of q. For values of $q < 0.5$, the equilibrium value of p is less than 0.9. The functional relationship, if any, between p and q is not clear, but it would appear that q must be be greater than $\simeq 0.5$ to achieve $p = 0.9$.

The evolution of the model system in equation B.5 with time is shown in Figure B.2. The system starts with $N_{1,1} + N_{2,1} = K$ and $p = 0.9$. After a few time steps the population decreases to reflect influence of immigrants and the onset of fishing. Towards the end of the simulation period the proportion local drops below 0.9 in response to the decrease in the size of the immigrant population.

Population Forcing. 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} in equation B.5 is the biomass immigrating into the MHI stock at each time step. The likely candidates for forcing are MFCL regions 2 and 4, both of which overlap the MHI. The estimated biomass trajectories for these two regions are shown in Figure A.6. For testing purposes, I assume that the source of immigrants is MFCL region 2, so that

$$T_{21_t} = T_{21}^* \cdot B_{2,t} \qquad (B.6)$$

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.

Asymptotic biomass. K is the asymptotic biomass in a population growing according to logistic dynamics. It is the population size to which the population tends at equilibrium and is often dubbed "carrying capacity". It is not clear that a coupled logistic model such as equation B.1 has a nontrivial equilibrium. 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 that the equilibrium population size, K, has been constant for over 65 years. An alternative parameterization for K is to assume a random walk, for example,

$$\log K_t = \log K_{t-1} + \omega_t; \quad \omega_t \sim N(0, \sigma_\omega^2) \qquad (B.7)$$

where σ_ω^2 is the variance of the asymptotic population size random walk. This

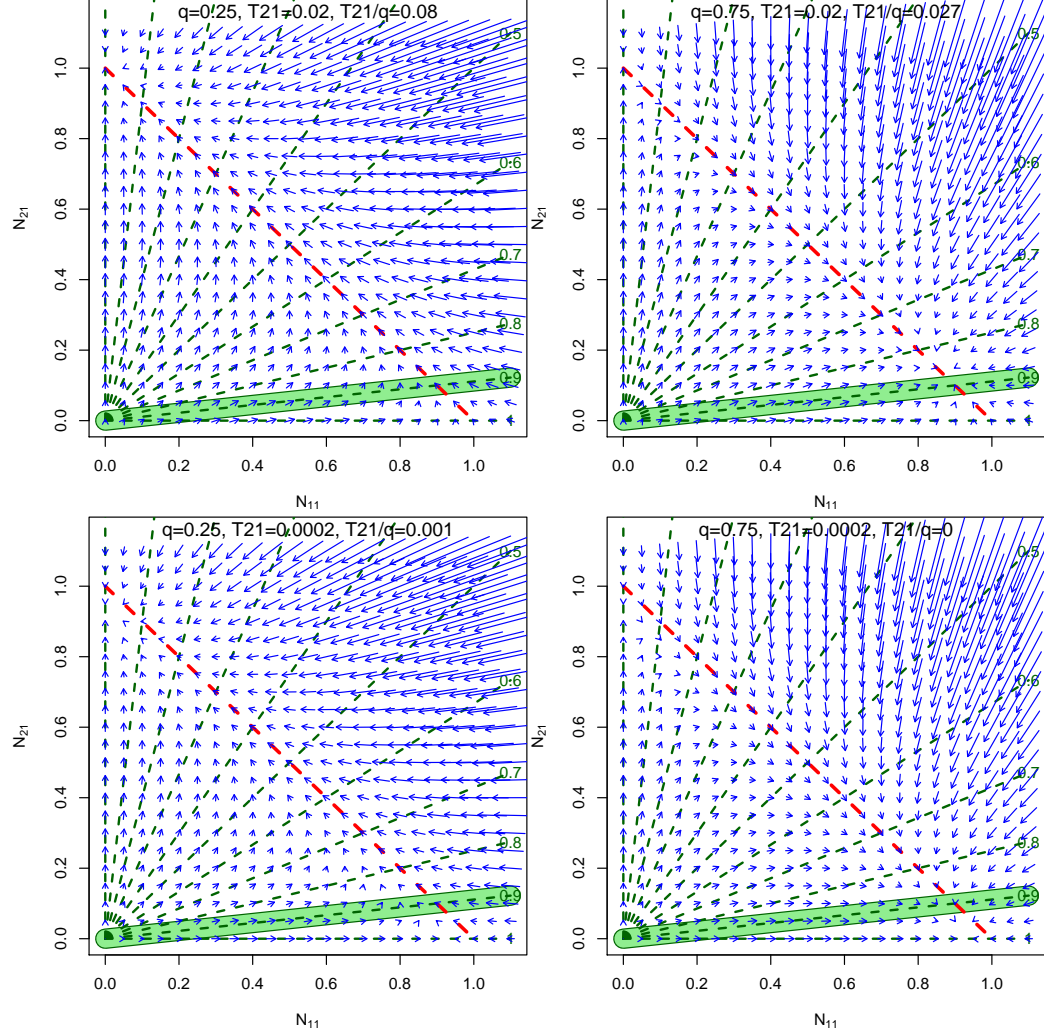Figure B.1: Phase plane for equation B.4. Arrows represent the trajectory of the system for different combinations of $N_{1,1}$ and $N_{2,1}$; dashed red line indicates $K = 1$; dashed green lines indicate different values of p. The thick light green line indicates $p = 0.9$. Values of q and T_{21} are shown in each panel. Other parameter values: $r = 0.3, K = 1.0, F = 0.007, T_{12} = 0.01$.

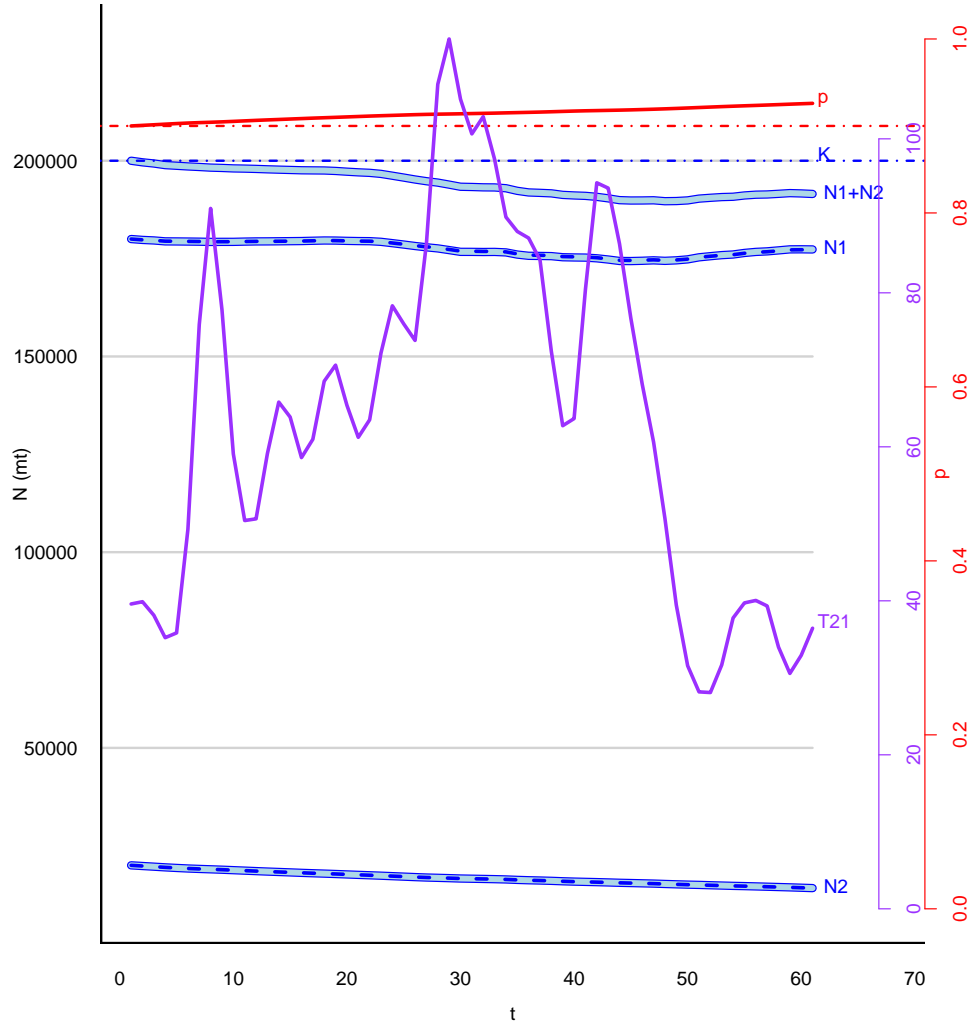Figure B.2: Behavior of equations B.5 over time with forcing from MFCL region 2 biomass (Figure A.6) and fishing mortality by fleet set proportional to observed catch. Light blue lines indicate population components as indicated; dashed blue line indicates K; solid red line indicates the proportion local, p; dashed red line line indicates $p = 0.9$. Other parameter values: $q = 0.54, r = 0.3, K = 200000, T12 = 0.01, T21 = 0.002$.

parameterization was implemented in computer code, but not extensively tested.

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

$$\log C_{g,t} = \log\left(F_{g,t} \cdot \left(\frac{N_{1,1_{t-\Delta t}} + N_{1,1_t}}{2} + \frac{N_{2,1_{t-\Delta t}} + N_{2,1_t}}{2} \right) \right) + \varepsilon_t \quad \text{(B.8)}$$

where the total population in region 1 is the sum of the average population over the time step (Quinn and Deriso, 1999), and ε_t is a robust likelihood given by

$$\varepsilon_t \sim (1 - a_g) * N(0, \sigma_\varepsilon^2) + a_g * C(0, \sigma_\varepsilon^2). \quad \text{(B.9)}$$

C is the standard Cauchy probability density and a_g is a fleet-specifc proportion of "contamination" by the fat-tailed Cauchy distribution. In practice $a_g = 0.07$ $g = 1 \ldots n$.

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, $L(\bar{p})$.

$$L(p) \sim N(L(\bar{p}), \sigma_{L(p)}^2). \quad \text{(B.10)}$$

p is generally assumed to be approximately 0.9. By varying the value of the variance, $\sigma_{L(p)}^2$, p can be made as close to 0.9 as is prudent. This representation of p can be interpreted as a Bayesian prior. In principle, it may be possible to estimate $\bar{p}$ and $\sigma_{L(p)}^2$, but interactions with the parameter q in equation 3.4 should be explored.

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

$$L(\theta, \alpha, x) = \prod_{t=2}^{m} \left[\phi\big(\alpha_t - T(\alpha_{t-1}), \Sigma_\eta\big) \right] \prod_{t=1}^{m} \left[\phi\big(x_t - O(\alpha_t), \Sigma_\varepsilon\big) \right] \quad \text{(B.11)}$$

Here, m is the number of time steps in the catch time series and θ is a vector of model parameters. A complete list of parameters is found in Table B.1. The model is implemented in ADMB-RE (Fournier et al 2012). The actual number of parameters to be estimated depends on the model configuration, specified by phase flags in the input file.

Table B.1: Complete list of variables and parameters for the two population model

Variable	Definition
m	Number of quarterly time steps
n	Number of fishing gears
r	Instantaneous growth rate (y^{-1})
K	Asymptotic biomass (mt)
T_{12}	Emigration rate (y^{-1})
T_{21}^*	Immigration rate (y^{-1})
q	Nonlinear mortality apportionment proportion
σ_N	Population growth SD
σ_F	Fishing mortality random walk SD
σ_Y	Observation error SD
a	Likelihood contribution from fat-tailed distribution
$\bar{p}$	Mean proportion local
$\sigma_{L(p)}$	SD logit transformed $\bar{p}$

B.2   Single population with indexed abundance

The indexed abundance model[1] assumes that the biomass of YFT in the MHI is proportional to the biomass of an "index" population. Let I_t equal the biomass of the index population and Q the constant of proportionality between the MHI population and the index population, thus

$$\varphi_t \sim N(QI_t - N_t, \sigma_Q) \tag{B.12}$$

becomes a component of the total likelihood due to differences between the MHI biomass and the index population biomass.

   Transition Equation $T(\alpha_{t-1})$. The state space transition equation for the single population model is developed by solving 3.5 analytically from one time to the next (see Section C).

$$N_t = \frac{K(r - \bar{F}_t)}{\frac{K(r-\bar{F}_t)}{N_{t-\Delta t}} e^{-\Delta t(r-\bar{F}_t)} - re^{-\Delta t(r-\bar{F}_t)} - r} + \eta_t \tag{B.13}$$

where $\bar{F}_t$ is the total fishing mortality, i. e.,

$$\bar{F}_t = \sum_{g=1}^{n} F_{g,t-\Delta t}.$$

$\eta_t \sim N(0, \sigma_N)$ is the process error expressing variability in population dynamics.

   Observation Equation, $O(\alpha)$. Predicted catch, $\widehat{C}_{g,t}$, for each gear is the product of estimated fishing mortality and the total biomass.

$$\log \widehat{C}_{g,t} = \log \left(F_{g,t} \cdot \left(\frac{N_{t-\Delta t} + N_t}{2} \right) \right) + \varepsilon_t \tag{B.14}$$

where the total biomass is the average biomass over the time step (Quinn and Deriso, 1999), and ε_t is a "zero-inflated" normal likelihood given by

$$\log \varepsilon_t = \begin{cases} C_{g,t} > 0: & (1 - p_0) \cdot \left(\log \frac{1}{\sqrt{2\pi\sigma_Y^2}} - \left(\frac{\log C_{g,t} - \log \widehat{C}_{g,t}}{\sigma_Y} \right)^2 \right) \\ C_{g,t} = 0: & p_0 \cdot \log \frac{1}{\sqrt{2\pi\sigma_Y^2}} \end{cases} \tag{B.15}$$

p_0 is the proportion of observed catch observations equal to 0.0. This proportion may be estimated or fixed at a constant value. For current analysis, it is fixed at $p_0 = 0.15738$ as computed from the data.

---

[1]https://github.com/johnrsibert/XSSA/blob/master/ADMB/issams.tpl

Table B.2: Complete list of variables and parameters for the single population model.

Variable	Definition
m	Number of quarterly time steps
n	Number of fishing gears
r	Instantaneous growth rate (y^{-1})
K	Asymptotic biomass (mt)
Q	Abundance index proportionality constant
p_0	Proportion of zero catch observations; fixed at $p_0 = 0.15738$
σ_N	Population growth SD
σ_F	Fishing mortality random walk SD
σ_Q	Abundance index proportionality SD
σ_Y	Observation error SD

C   Integrating Schaefer Models

The widely used Schaefer (1954) fisheries stock assessment model is a simple extension of the logistic population model with a term added to represent removals from the population due to fishing.

$$\frac{dN}{dt} = rN(1 - \frac{N}{K}) - FN = N(r - F - \frac{r}{K}N) \tag{C.1}$$

where N is the population size, r is the instantaneous growth rate (t^{-1}), K is the asymptotic population size in the same units as N, and F, is the instantaneous rate of removal due to fishing (t^{-1}). Equation (C.1) reduces to the logistic model if F is assumed to be zero. These models are usually integrated numerically with "explict" finite difference methods to compute an approximation of the value of N at some time. Such approximations are often unstable for values of r large relative to the time step used in the finite difference solution. The performance of estimation methods in which logistic models are embedded in a statistical procedure depending on numerical function minimizers, i.e, models build using ADMB, TMB and BUGS, is greatly improved in accuracy, speed and use of computing resources if analytical solutions are used in preference to finite difference approximations.

C.1   Single population

Quinn and Deriso (1999) show the integral of the logistic ODE and point out that it can be obtained using integration by partial fractions. Numerous mathematics tutorials can be found on the World Wide Web use integration of the logistic ODE to illustrate the technique of integration by partial fractions. The same procedure can be used to integrate the Schaefer ODE. Equation (C.1) is rearranged slightly and variables separated to become

$$\frac{K}{N(K(r - F) - rN)}dN = dt. \tag{C.2}$$

The fraction in the left hand side can be factored into two parts,

$$\frac{K}{N(K(r - F) - rN)} = \frac{A}{N} + \frac{B}{(K(r - F) - rN)}. \tag{C.3}$$

A and B are constants that can be found by solving $K = A(K(r - f) - rn)) + BN$ setting $N = K$ and $N = 0$; $A = \frac{1}{r-F}$ and $B = 1 + \frac{F}{r-F}$. The

desired integral becomes

$$\int \frac{K}{N(K(r-F)-rN)}dN = \int dt$$

$$\int \frac{A}{N}dN + \int \frac{B}{K(r-F)-rN}dN = \int dt$$

$$\frac{1}{r-F}\int \frac{1}{N}dN + (1 + \frac{F}{r-F})\int \frac{1}{K(r-F)-rN}dN = \int dt$$

$$\frac{1}{r-F}\log|N| + \frac{1}{r}(1 + \frac{F}{r-F})\log|K(r-f)-rN| + logC = t$$

$$\log|N| - \log|K(r-F)-rN| + logC = t(r-F)$$

$$\frac{|N|}{|K(r-F)-rN|} \cdot C = e^{t(r-F)}$$

$$\frac{|K(r-F)-rN|}{C|N|} = e^{-t(r-F)}$$

where C is the constant of integration. Setting $|N| = N_t$, the population size at time t, yields

$$N_t = \frac{K(r-F)}{Ce^{-t(r-F)} + r} \tag{C.4}$$

A formula suitable for computing population size at successive time steps can be found by setting $N_t = N_{t-\Delta t}$ at time $t = t - \Delta t$ in equation (C.4). The integration constant, C, becomes

$$C = \left(\frac{K(r-f)}{N_{t-\Delta t}} - r\right)e^{(t-\Delta t)(r-F)}, \tag{C.5}$$

and finally

$$N_t = \frac{K(r-F)}{\frac{K(r-F)}{N_{t-\Delta t}}e^{-\Delta t(r-F)} - re^{-\Delta t(r-F)} - r} \tag{C.6}$$

Further simplification of this equation may be possible, but I have not found it. In any case, equation (C.6) is the only general solution of the Schaefer ODE that I have seen, and it appears to work well in numerical applications.

C.2   Two populations with exchange

The motivation for the two popultion Schaefer model with exchange is developed fully in Appendix B.

The basic equations can be written

$$\frac{dN_{1,1}}{dt} = N_{1,1}\left(r - F - T_{12} - 2(1-q)\frac{r}{K}N_{2,1} - \frac{r}{K}N_{1,1}\right) \quad (C.7)$$

$$\frac{dN_{2,1}}{dt} = N_{2,1}\left(r - F - T_{12} - 2q\frac{r}{K}N_{1,1} - \frac{r}{K}N_{2,1}\right) + T_{21} \quad (C.8)$$

where $N_{1,1}$ the biomass of fish originating in region 1 and residing in region 1, and $N_{2,1}$ is the biomass of fish originating in region 2 but residing in region 1. The parameters r, K and F are unchanged from equation (C.1), and T_{12} is the emigration rate from region 1 (t^{-1}), T_{21} is the rate of immigration of biomass from region 2 to region 1 in units of biomass per time, and q; $(0 < q < 1)$ partitions the mortality caused by "competition" between the two subpopulations. Substitute

$$Z_1 = F + T_{12} + 2(1-q)\frac{r}{K}N_{2,1} \quad (C.9)$$

$$Z_2 = F + T_{12} + 2q\frac{r}{K}N_{1,1} \quad (C.10)$$

into equations (C.7) and (C.8) respectively to produce model equations in a similar form to equation (C.1)

$$\frac{dN_{1,1}}{dt} = N_{1,1}(r - Z_1 - \frac{r}{K}N_{1,1}) \quad (C.11)$$

$$\frac{dN_{2,1}}{dt} = N_{2,1}(r - Z_2 - \frac{r}{K}N_{2,1}) + T_{21}. \quad (C.12)$$

Equation (C.11) can be integrated in the same manner as equation (C.1) to yield

$$N_{1,1_t} = \frac{K(r - Z_1)}{C_1 e^{-t(r-Z_1)} + r}. \quad (C.13)$$

All that remains is to find an equivalent integral for equation (C.12) and a means to simultaneously solve for C_1 and C_2.