Feasibility of developing a stock assessment model for 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.

August 19, 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. Management of this fishery is hampered by lack of stock assessment results specific to the MHI. The feasibility of developing a state-space Schafer model with emigration, external forcing, Bayesian constraints, and robust estimation specifically for the MHI was investigated. The model is 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 (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

*sibert@hawaii.edu

support. The documented history of the YFT fishery begins 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, but 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 a stock assessment model for the MHI YFT population that includes emigration and immigration of biomass and a constraint on the proportion of "resident" biomass.

2 Methods

2.1 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 partic-

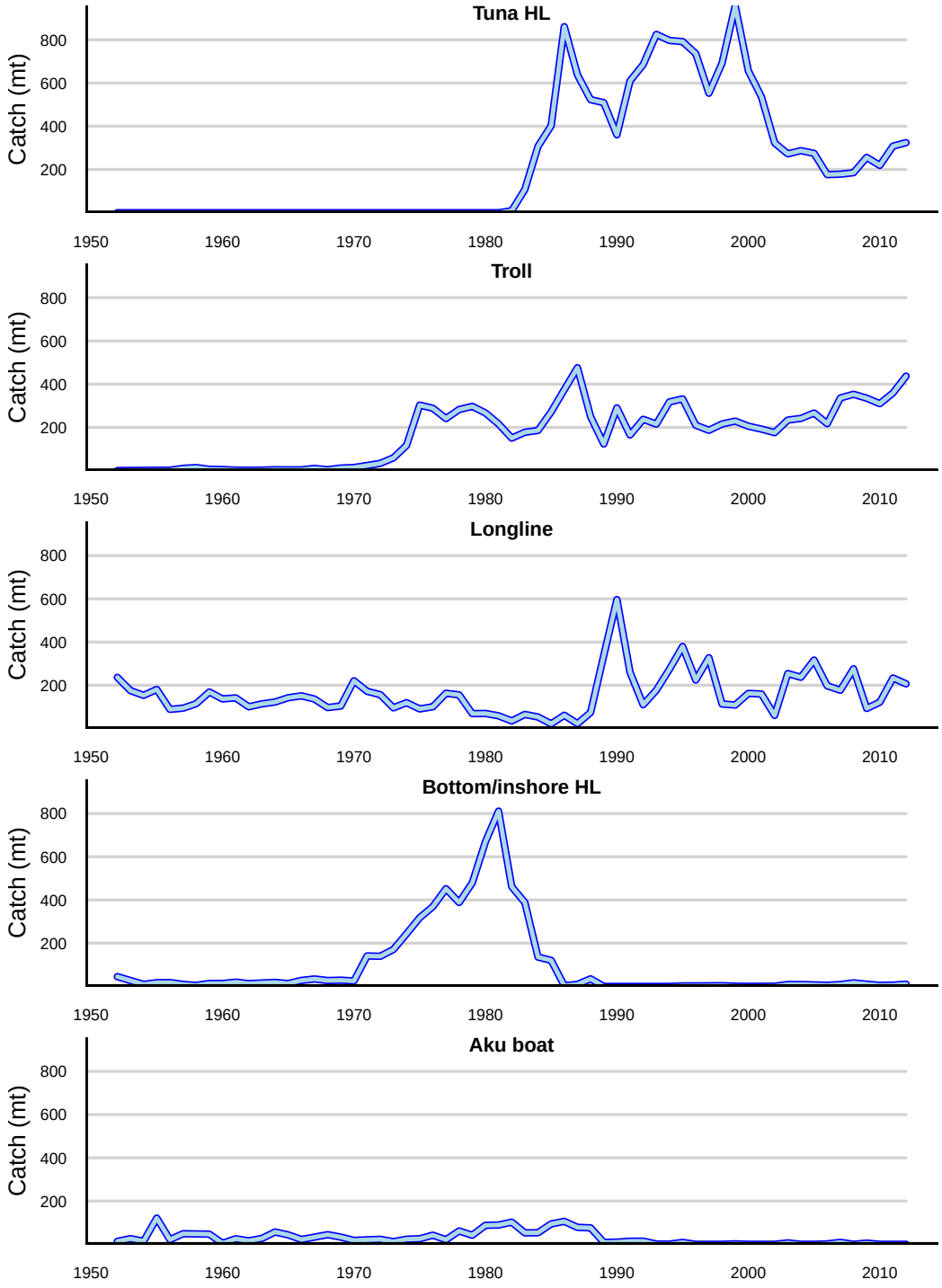

Figure 2.1: Annual yellowfin catch in metric tonnes by fisheries operating in the Main Hawaiian Islands from combined HDAR and NOAA data.

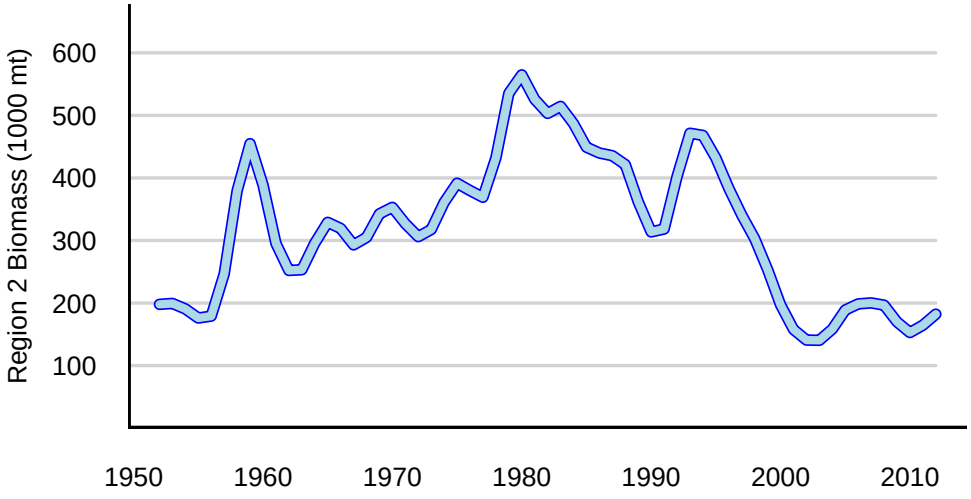

Figure 2.2: Annual estimated biomass trends in MFCL Region 2 from the 2014 WCPFC YFT stock assessment, Davies et al. 2014.

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

A biomass exchange model presupposes a source and sink for biomass exchange. Model estimates of yellowfin biomass by MULTIFAN-CL (MFCL) from the most recent WCPFC stock assessment (Davies et al 2014) provide the most 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.

2.2 Model

An age-aggregated model was developed in preference to a more complex age-structured model for examining the feasibility of a MHI YFT assessment model. 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. 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

$$\alpha_t = T(\alpha_{t-1}) + \eta_t \tag{2.1}$$

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

$$x_t = O(\alpha_t) + \varepsilon_t \tag{2.2}$$

where the function O describes the measurement process with error ε in observing the population.

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 (see Appendix B for details).

$$\begin{aligned} \frac{dN_{1,1}}{dt} &= N_{1,1}\left[r\left(1-\frac{N_{1,1}}{K}\right) - F - T_{12}\right] - (1-q)2r\frac{N_{1,1}N_{2,1}}{K} \\ \frac{dN_{2,1}}{dt} &= N_{2,1}\left[r\left(1-\frac{N_{2,1}}{K}\right) - F - T_{12}\right] - q2r\frac{N_{1,1}N_{2,1}}{K} + T_{21} \end{aligned} \tag{2.3}$$

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 Results

The general behavior of the model system is shown in Figure 3.1, emphasizing the importance of q in maintaining the stability of two subpopulations. For values of $q > 0.5$, the local population, $N_{1,1}$ is stable, $N_{2,1}$ decreases, and the proportion local remains near 0.9. For values of $q < 0.5$, the local population, $N_{1,1}$ decreases, $N_{2,1}$ increases, and the proportion local dips below 0.9. This same dependency on q is also illustrated more generally in Figure B.1.

Initial testing of the estimation procedure gave mixed results. The model is sensitive to initial parameter guesses. Some combinations of initial values cause the numerical estimation procedure to terminate abnormally with the infamous ADMB error message ``Matrix not positive definite in Ln_det_choleski''. Nevertheless, the procedure arrives at a "solution" for some combinations of initial estimates.

Observed and predicted catches are shown in Figure 3.2. The random-walk representation of fishing mortality produces a credible time series predicting catches very accurately for all fleets, except for the Aku Boat fleet.

Predicted biomass trajectories are shown in Figure 3.3. The population appears to begin above equilibrium and to decline to near equilibrium within two years. The size of the $N_{2,1}$ population drops to near zero over the 61 year estimation period. In consequence, the proportion local begins slightly below 0.9, subsequently fluctuates around 0.9 and then increases. The increase in q appears to coincide with the steep decline in the immigration rate of biomass into the MHI (T_{21}) in the late 1990s.

Values of the model parameters just prior to termination of the program are compared to their initial values in Table 3.1. Some model parameters do not change substantially from their initial values during the estimation procedure. The population dynamics parameters, r, K, T_{12}, T^*_{21}, and q, are the most important for stock assessment, but do not change much. Of these, q, the parameter apportioning the non-linear mortality term between the two subpopulations, changes the most.

The variance parameters are of no particular interest and exist only to

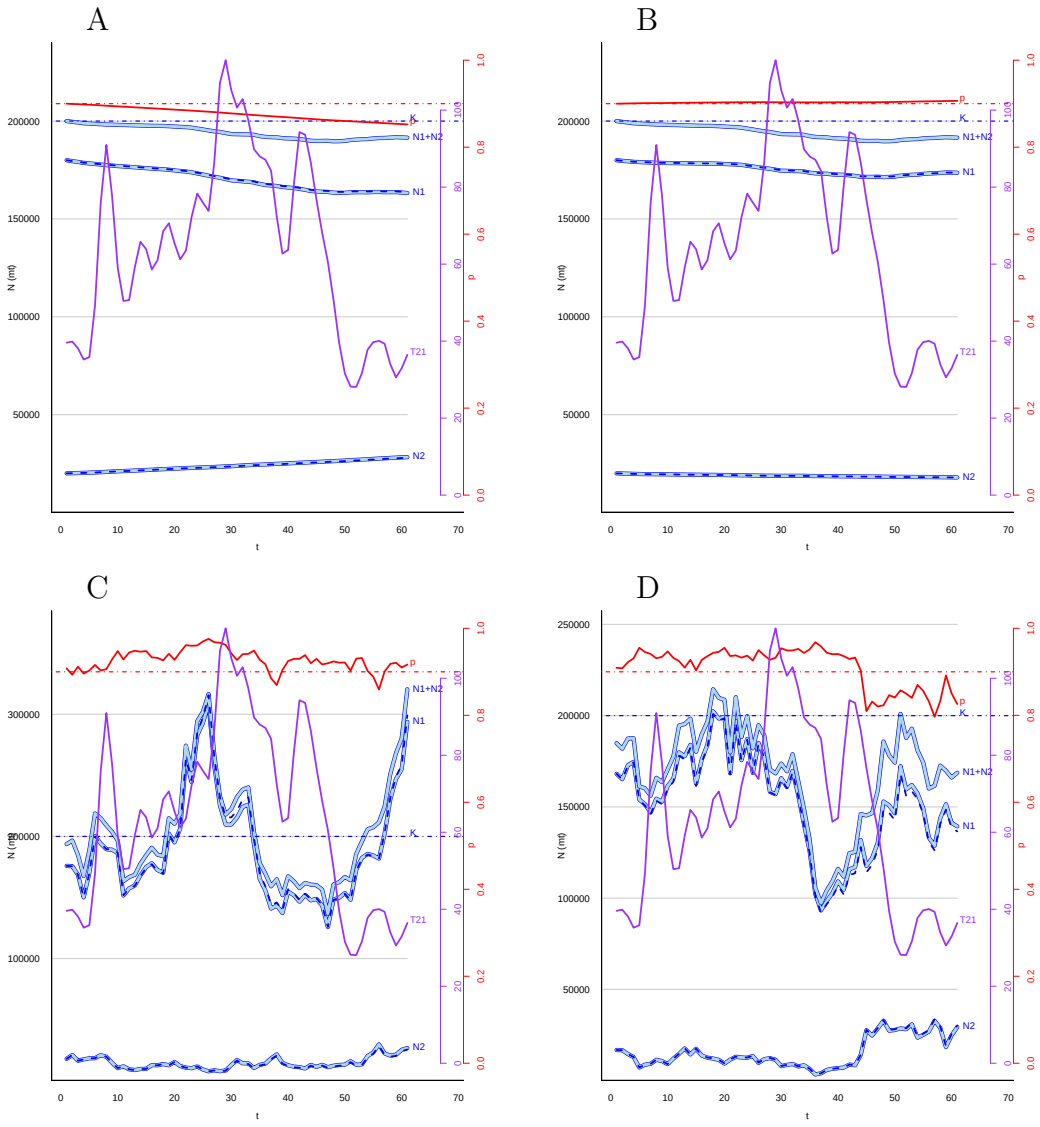

Figure 3.1: Simulated biomass trends with two values of the parameter q and two levels of process error. Light blue lines indicate estimated $N_{1,1}$, $N_{2,1}$, and $N_{1,1} + N_{2,1}$ biomass in metric tonnes. The dashed blue line is the estimated equilibrium biomass, K. Purple line is the estimated immigrant biomass T_{21}, Red line indicates estimated proportion local, $\bar{p}$; dashed red line indicates $\bar{p} = 0.9$. Panels A and C, $q = 0.48$; panels B and D, $q = 0.52$. Panels C and D repeat panels A and B with a random normal errors on $N_{1,1}$ and $N_{2,1}$ with standard deviations of 0.01 and 0.05 respectively.

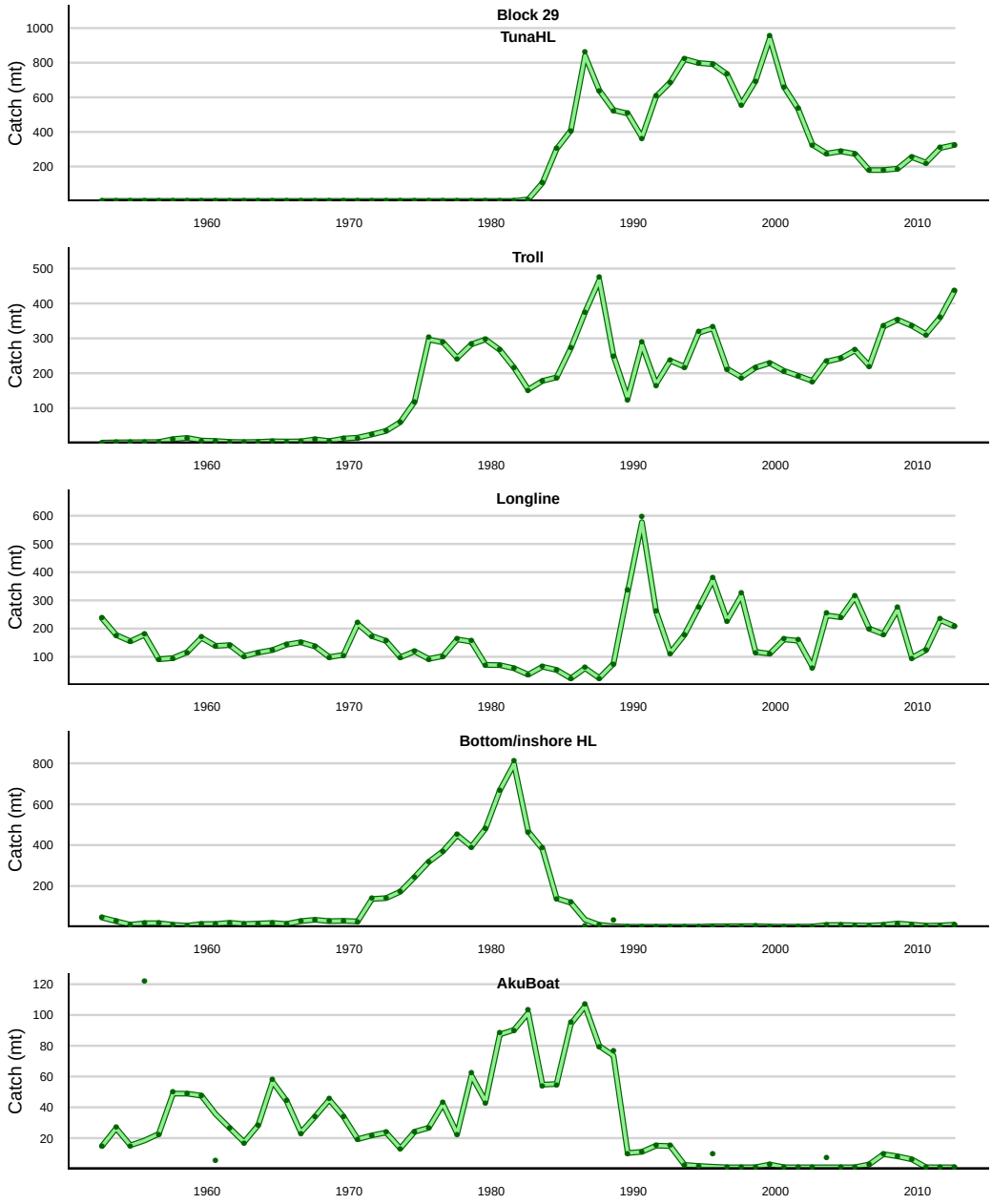

Figure 3.2: Estimated catch by fleet in metric tons. The dark green dots indicate observed catch and light green lines indicate predicted catch,

Table 3.1: Model "estimates". Initial values of model parameters and final values just prior to program exit with ‘‘`Function minimizer: Step size too small -- ialph=1`’’. Dashes (—) indicate that the variable was constrained to be constant at its initial value, i.e., not estimated.

Vatiable	Initial Value	Final Value
r	1.20	1.29
K	150,000	149,000
T_{12}	0.01	0.00999
T^*_{21}	0.002	0.00202
q	0.54	0.525
σ_η	0.1	0.0883
σ_ξ	0.3	0.672
σ_ε	1.0	0.0945
a_1	0.07	—
a_2	0.07	—
a_3	0.07	—
a_4	0.07	—
a_5	0.07	—
$\bar{p}$	0.9	—
$\sigma_{\bar{p}}$	0.8	—

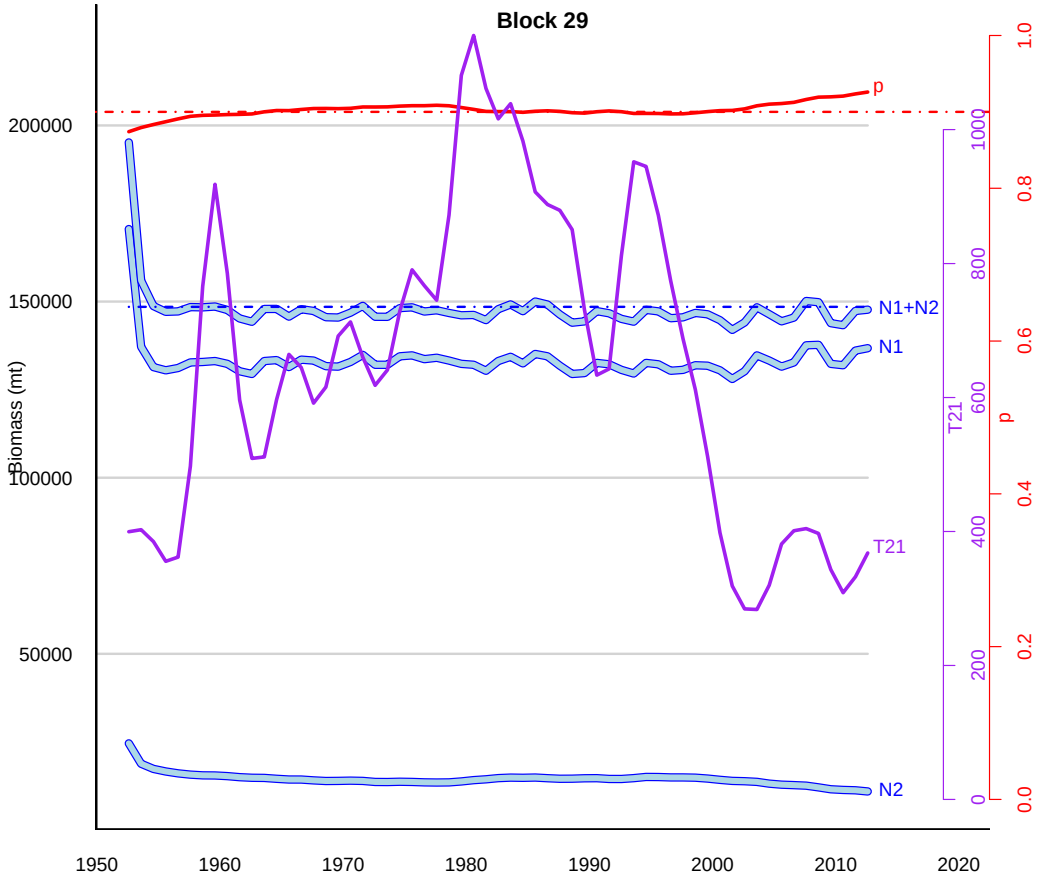

Figure 3.3: Estimated biomass trends. Light blue lines indicate estimated $N_{1,1}$, $N_{2,1}$, and $N_{1,1} + N_{2,1}$ biomass in metric tonnes. The dashed blue line is the estimated equilibrium biomass, K. Purple line is the estimated immigrant biomass T_{21}, Red line indicates estimated proportion local, $\bar{p}$; dashed red line indicates $\bar{p} = 0.9$.

represent variability in population dynamics and errors in observation. These "nuisance" appear to be reasonably sensitive to the data.

4 Discussion

The age-aggregated approach using a logistic model with emigration and immigration shows sufficient promise to continue model development. The general behavior of the model system (Figure 3.1 and B.2) appropriately describes a small persistent population interacting with a much larger one. The efficacy of the random walk representation of fishing mortality is particularly encouraging, and avoids debates about changes in catchability associated with changes in fishing efficiency. The proportion of local fish in the system is sensitive to model parameters and can effectively be constrained to be near the empirically determined value of 0.9. Coupling immigration to the output of some other model is relatively trivial to implement and is sufficiently flexible to use output from alternative models.

Unfortunately, population dynamics parameters, with the possible exception of q, do not appear to be estimatable. Lack of sensitivity could be due to several causes: the model is completely ill-posed, the code is incorrect, there is no signal in data, or some combination of these. Inspection of catch trends (Figure 2.1) does not reveal any obvious signs of stock depletion and recovery. Estimation of population dynamics parameters in a lightly exploited stock is a well-known problem in stock assessment.

Figure B.1 suggests that the model equations (2.3) have non-trivial equilibria for certain parameter values. The relationships between of the parameters q, T_{21}^*, and $\bar{p}$ are not clear. Estimation of fisheries reference point would be simpler if an analytic expression for the equilibrium point were available. Finite difference approximations of logistic population models are notoriously unstable under certain circumstances. An analytical solution of equation 2.3 would be of great help. Further mathematical analysis is called for.

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 to actually take on this challenging project and for their on-going collaboration. Thanks to Mr. David Itano for sharing insights into the small-boat fisheries in Hawaii. 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 log-

book 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., 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.

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.

B Model development

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:

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

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., $\frac{d}{dt}\big(N_{2,2} + N_{1,2}\big)$), but the dynamics of the fish population in region 2 is external to this model. T_{21} can be considered to be 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.

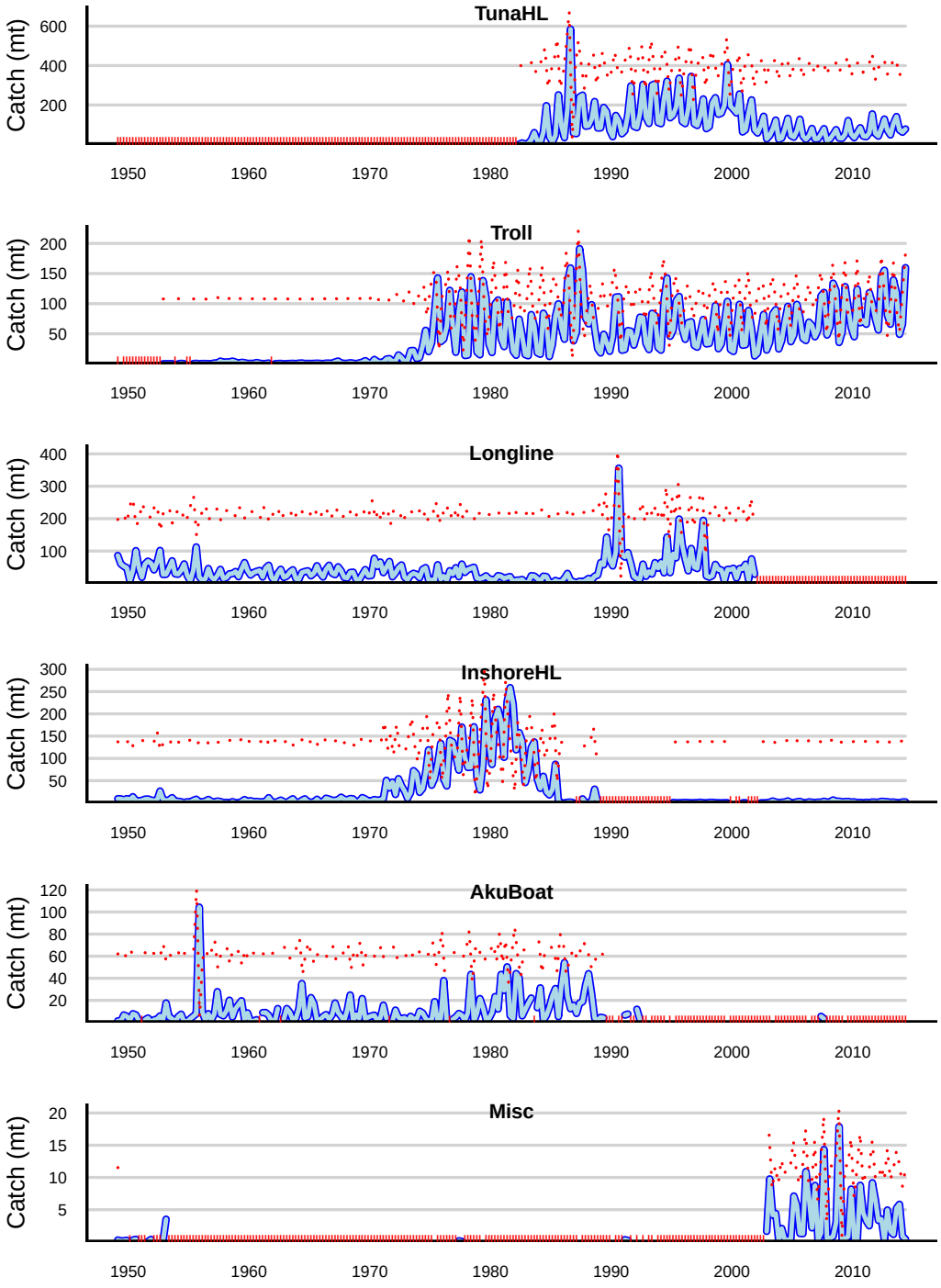

Figure A.1: Yellowfin catch in metric tonnes by principle fisheries operating in the Main Hawaiian Islands from the HDAR data. The dotted red line superimposed on each time series is the difference in catch between successive quarters. The red tick marks on the abscissa indicate quarters where reported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

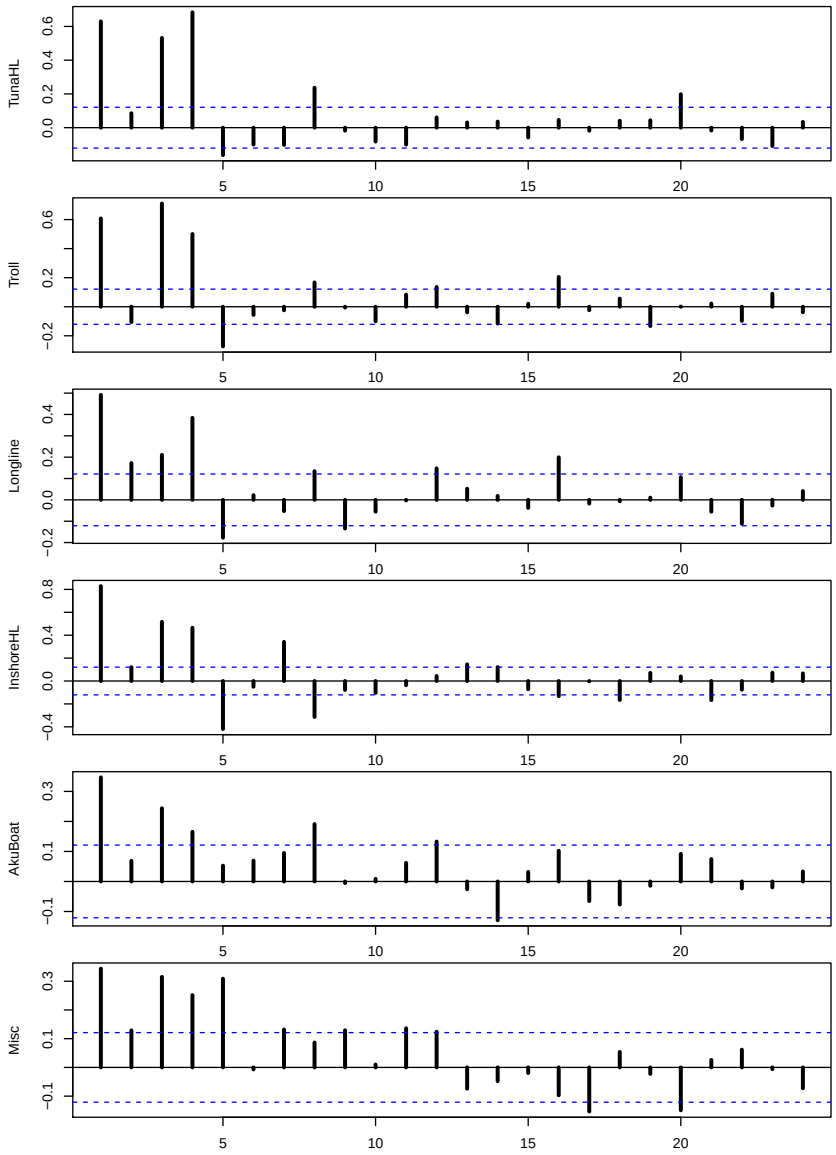

Figure A.2: Partial autocorrelation coefficients of the HDAR catch time series. The dashed blue lines indicate approximate 95% confidence limits of the correlations.

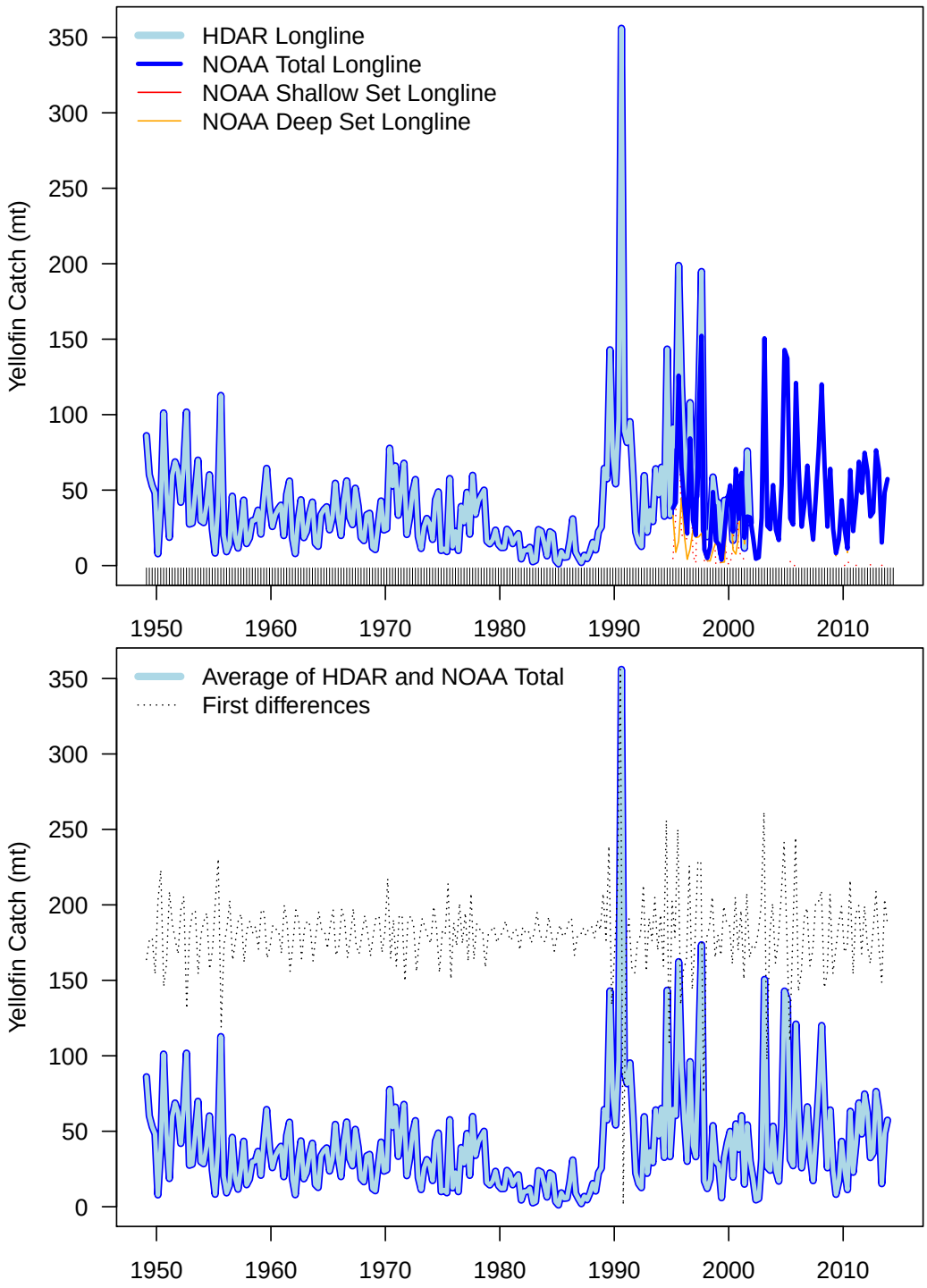

Figure A.3: Comparison between HDAR and NOAA longline time series. The upper panel shows the NOAA deep and shallow set data superimposed on the HDAR data. The lower panel shows the time series produced by a simple average of the HDAR data and the sum of the NOAA deep and shallow catches.

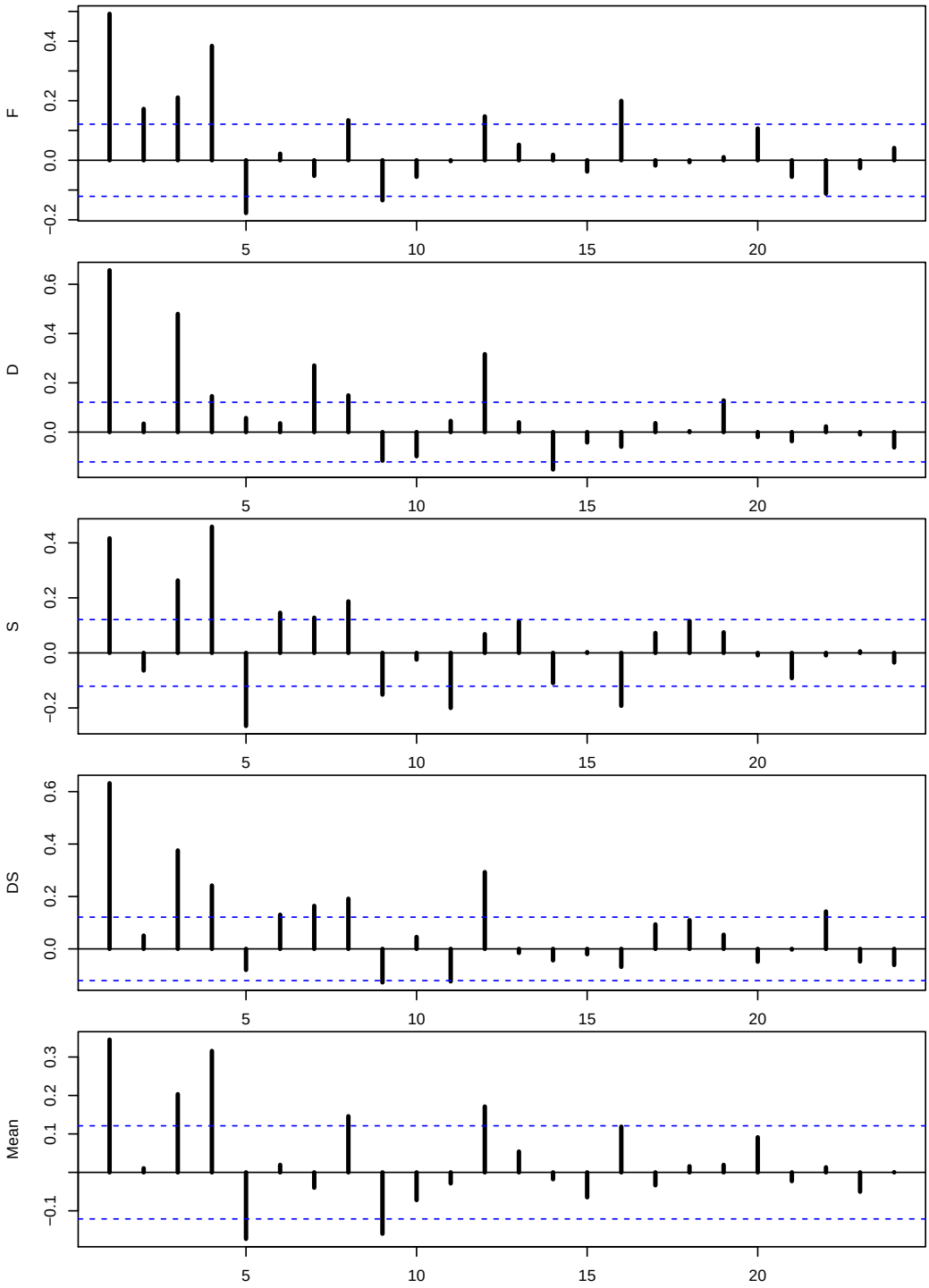

Figure A.4: Partial autocorrelations within the NOAA and combined time series. F refers to the HDAR longline data, D refers to the NOAA deep set data, S to the shallow set data, DS to the combined deep and shallow set data, and Mean to the average of the HDAR and NOAA deep plus shallow.

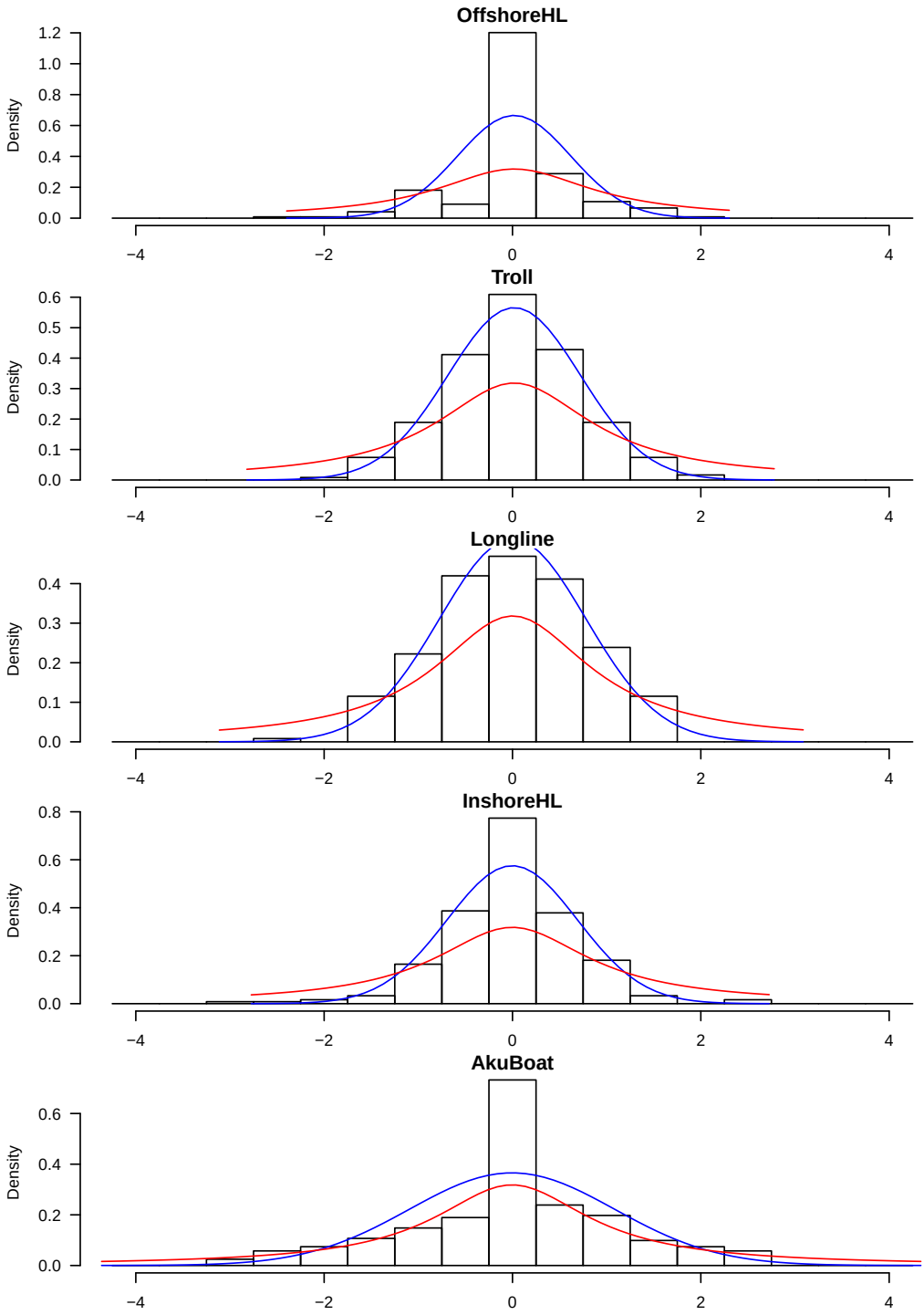

Figure A.5: Histograms of first differences of the logarithm of catch time series. The blue line is a normal distribution with the same mean and standard deviation as the first differences. The red line is the equivalent standard Cauchy distribution.

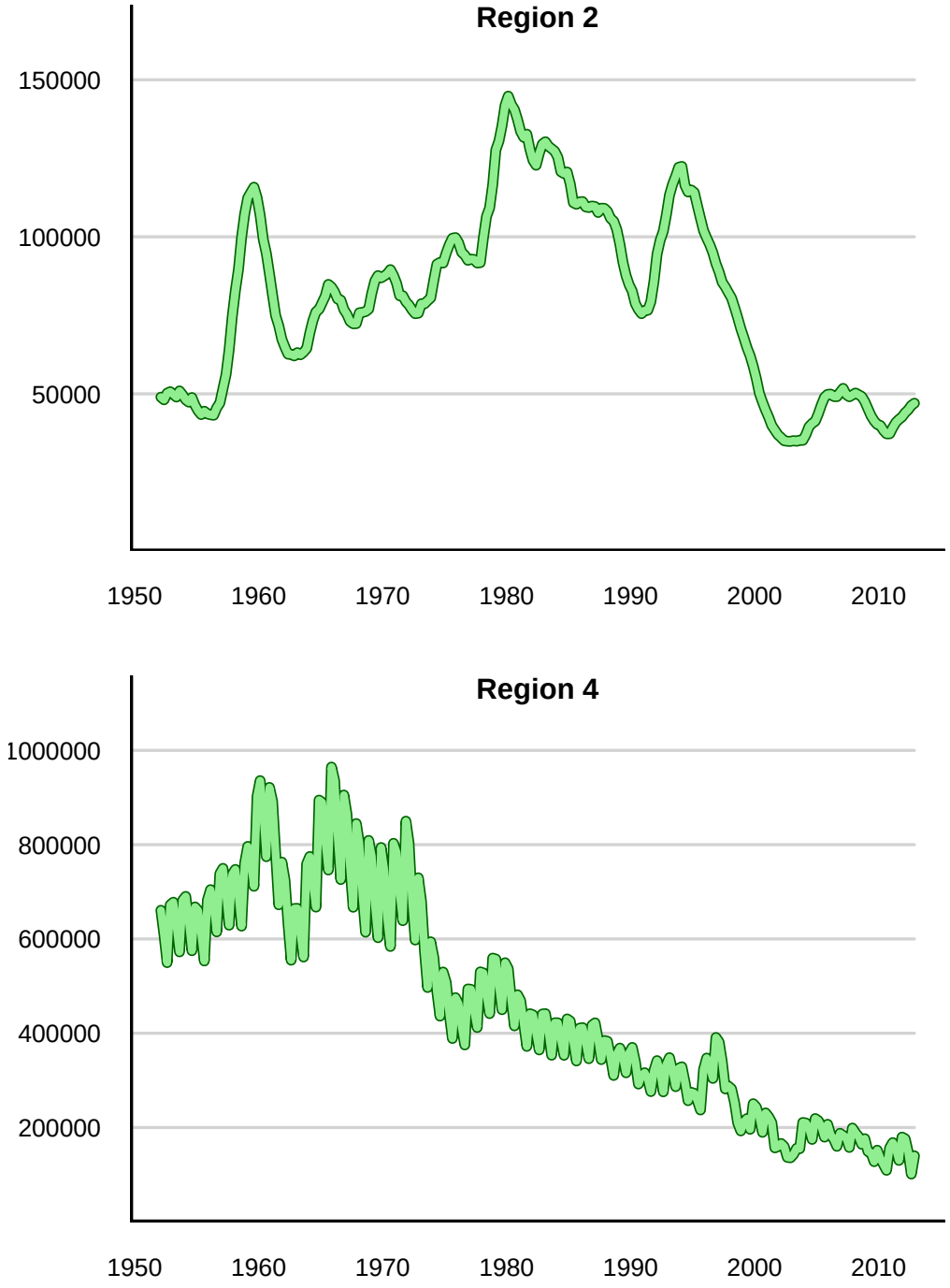

Figure A.6: Quarterly biomass trends estimated in MFCL Regions 2 and 4 from the 2104 WCPFC YFT stock assessment, Davies et al. 2014.

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

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

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.

$$\begin{aligned}\frac{dN_{1,1}}{dt} &= N_{1,1}\left[r\left(1-\frac{N_{1,1}}{K}\right)-F-T_{12}\right]-(1-q)2r\frac{N_{1,1}N_{2,1}}{K} \\ \frac{dN_{2,1}}{dt} &= N_{2,1}\left[r\left(1-\frac{N_{2,1}}{K}\right)-F-T_{12}\right]-q2r\frac{N_{1,1}N_{2,1}}{K}+T_{21}\end{aligned} \tag{B.3}$$

The log transformed equivalent of equation B.3 is

$$\begin{aligned}\frac{d\log(N_{1,1})}{dt} &= r\left(1-\frac{N_{1,1}}{K}\right)-F-T_{12}-(1-q)2r\frac{N_{2,1}}{K} \\ \frac{d\log(N_{2,1})}{dt} &= r\left(1-\frac{N_{2,1}}{K}\right)-F-T_{12}-q2r\frac{N_{1,1}}{K}+\frac{T_{21}}{N_{2,1}}\end{aligned} \tag{B.4}$$

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.

$$\begin{aligned}
\log N_{1,1_t} &= \log N_{1,1_{t-\Delta t}} \\
&+ \Delta t\left(r\left(1-\frac{N_{1,1_{t-\Delta t}}}{K}\right)-\sum_{g=1}^{n}F_{g,t-\Delta t}-T_{12}-(1-q)2r\frac{N_{2,1_{t-\Delta t}}}{K}\right)+\eta_t
\end{aligned} \tag{B.5}$$

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

where $\eta_t \sim N(0,\sigma_\eta)$ 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 fruitful. 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.

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

$$\log F_{g,t} = \log F_{g,t-1} + \xi_t; \quad \xi_t \sim N(0,\sigma_\xi^2) \tag{B.6}$$

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

Population Forcing. Immigration of stock from region 2 to region 1 is a form of forcing whereby events outside of the model influence the

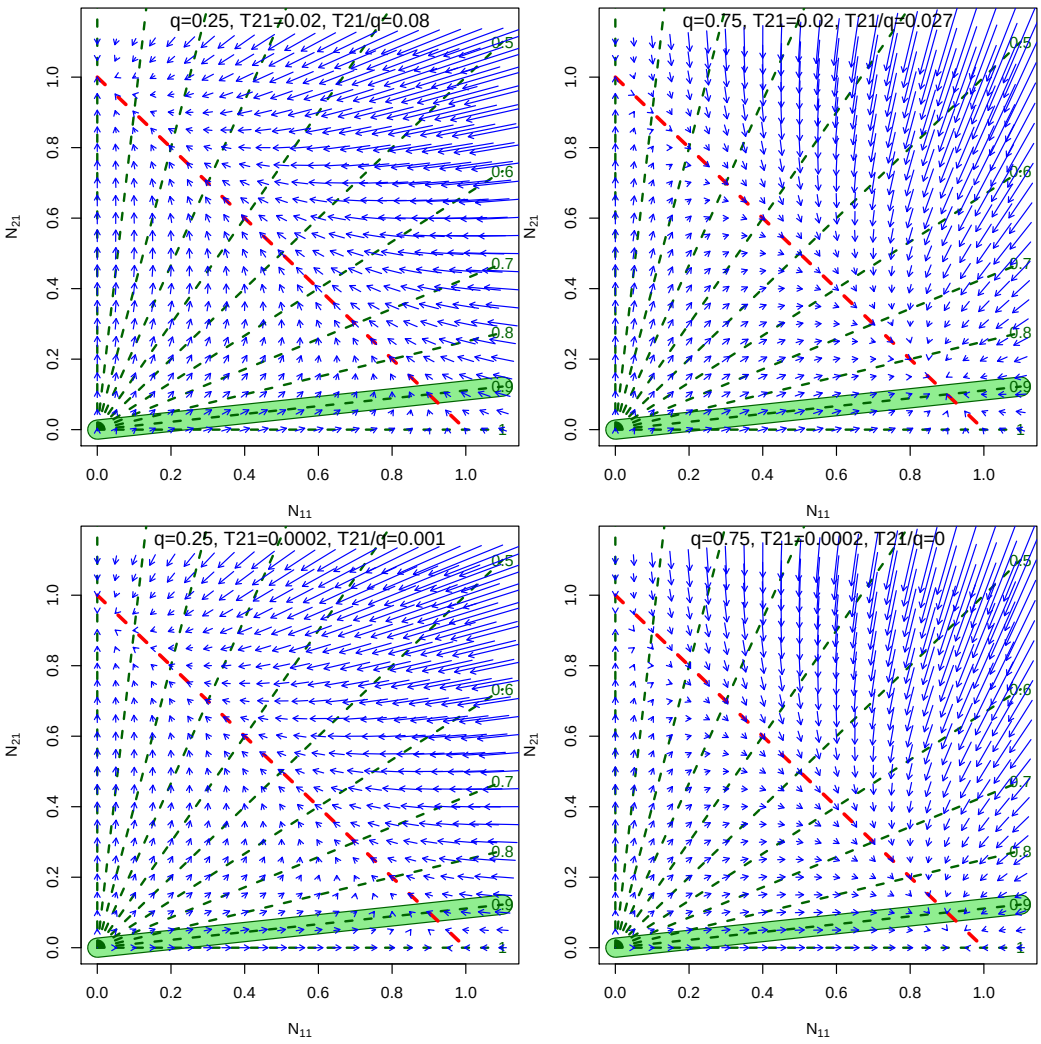

Figure B.1: Phase plane for equation B.4. Arrows represent the trajectory of the system for different combinations of $N_{1,1}$ and $N_{2,1}$; dashed red line indicates $K = 1$; dashed green lines indicate different values of p. The thick light green line indicates $p = 0.9$. Values of q and T_{21} are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

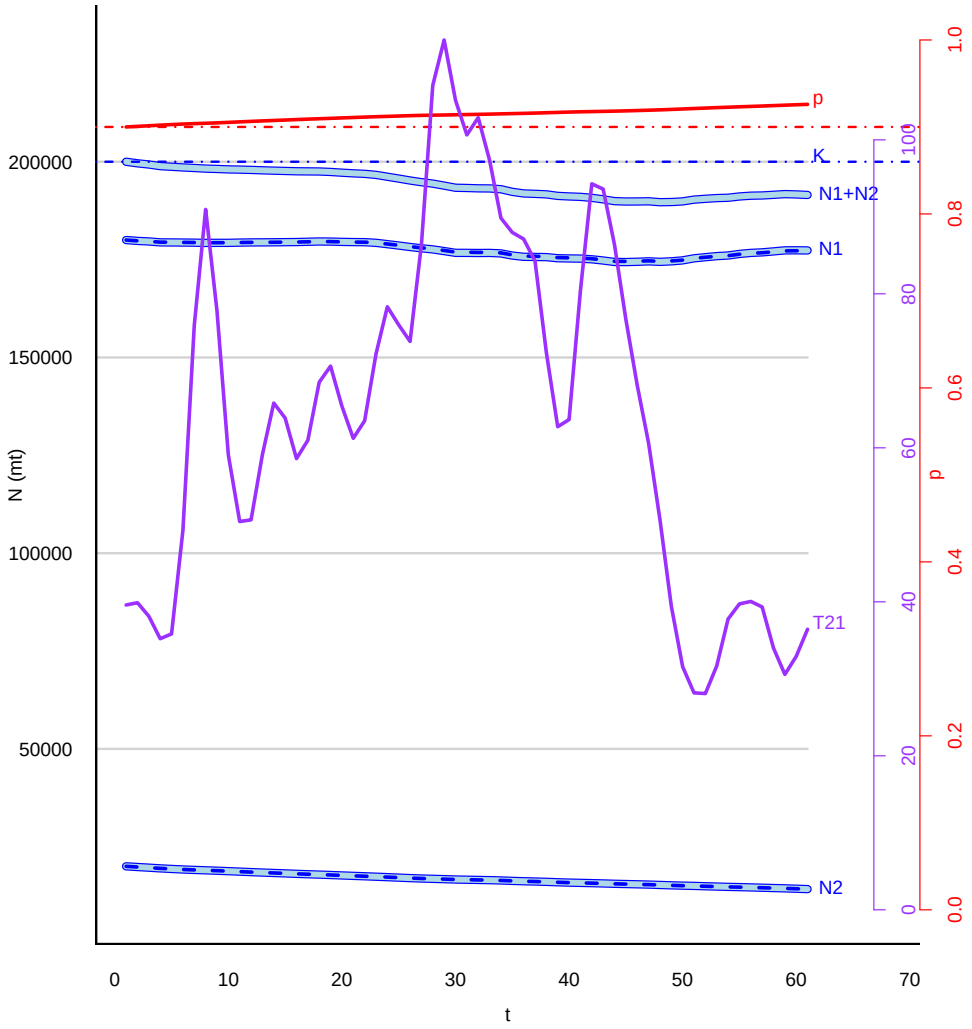

Figure B.2: Behavior of equations B.5 over time with forcing from MFCL region 2 biomass (Figure A.6) and fishing mortality by fleet set proportional to observed catch. Light blue lines indicate population components as indicated; dashed blue line indicates K; solid red line indicates the proportion local,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

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

$$T_{21_t} = T^*_{21} \cdot B_{2,t} \tag{B.7}$$

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 non-trivial 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,

$$\log K_t = \log K_{t-1} + \omega_t; \quad \omega_t \sim N(0, \sigma^2_\omega) \tag{B.8}$$

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

$$\log C_{g,t} = \log\left(F_{g,t} \cdot \left(\frac{N_{1,1_{t-\Delta t}} + N_{1,1_t}}{2} + \frac{N_{2,1_{t-\Delta t}} + N_{2,1_t}}{2}\right)\right) + \varepsilon_t \tag{B.9}$$

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

$$\varepsilon_t \sim (1 - a_g) * N(0, \sigma^2_\varepsilon) + a_g * C(0, \sigma^2_\varepsilon). \tag{B.10}$$

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

$$L(p) \sim N(L(\bar{p}), \sigma^2_{L(p)}). \tag{B.11}$$

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. This representation of p can be interpreted as a Bayesian prior. In principle, it may be possible to estimate $\bar{p}$ and $\sigma^2_{L(p)}$, but interactions with the parameter q in equation 2.3 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.

$$L(\theta, \alpha, x) = \prod_{t=2}^{m} \left[\phi\big(\alpha_t - T(\alpha_{t-1}), \Sigma_\eta\big)\right] \prod_{t=1}^{m} \left[\phi\big(x_t - O(\alpha_t), \Sigma_\varepsilon\big)\right] \tag{B.12}$$

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. All computer code, data files, and draft reports in support of this analysis can be found at Github: `https://github.com/johnrsibert/XSSA.git`.

Table B.1: Complete list of model variables and parameters.

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_η	Population growth SD
σ_ξ	Fishing mortality random walk SD
σ_ε	Observation error SD
$a_g\ g = 1 \ldots n$	Observation error contaminated by fat-tailed distribution
$\bar{p}$	Mean proportion local
$\sigma_{L(p)}$	SD logit transformed $\bar{p}$

C Integrating Schaefer Models

The widely used Schaefer 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 growth rate (t^{-1}), K is the asymptotic population size in the same units as N, and F, is the removals due to fishing (t^{-1}). Equation (C.1) reduces to a 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 large values of r relative to the time step used in the finite difference solutions. The performance of estimation methods in which where logistic models are embedded in a statistical models depending on numerical function minimizers is greatly improved in accuracy, speed and use computing resources if analytical solutions can be used in preference to finite difference approximations.

Numerous mathematics tutorials can be found on the World Wide Web which use integration of the logistic ODE to illustrate the technique of integration by partial fractions. 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 as

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

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

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

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

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

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

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

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

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

Setting $|N|=N_t$, the population size at time t, yields

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

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

$$C=\left(\frac{K(r-f)}{N_{t-\Delta t}}-r\right)e^{(t-\Delta t)(r-F)}, \tag{C.4}$$

and finally

$$N_t=\frac{K(r-F)}{\frac{K(r-F)}{N_{t-\Delta t}}e^{-(r-F)\Delta t}-re^{-(r-F)\Delta t}-r} \tag{C.5}$$

Further simplification of this equation may be possible, but I have not found a way. In any case, equation (C.5) is the only general solution of the Schaefer ODE that I have seen and appears to work well in numerical applications.