Assessing of a portion of the Pacific yellowfin tuna (*Thunnus albacares*) stock: *Ahi* in the Main Hawaiian Islands

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

December 30, 2016

Abstract

Regional tuna fishery management organizations do not provide useful advice to local fishery managers in small island jurisdictions. The State of Hawaii maintains time series of yellowfin tuna catches dating back to 1949, but these data have never been formally applied to evaluating the effects of the yellowfin fishery in the Main Hawaiian Islands on the local stock. I develop a new approach utilizing these data that links the local stock dynamics to the dynamics of the larger Pacific stock. This approach uses a state-space logistic production model linked the larger Pacific stock using an index of abundance. The conclusion is that such a model is feasible, that the local stock is not overfished and that local fisheries are fishing at acceptable levels.

*sibert@hawaii.edu

Introduction

The responsibility to manage fisheries for tunas and tuna-like species usually lies with regional organizations established under international treaties. In the Western Central Pacific Ocean (WCPO), this responsibility devolves to the Western and Central Pacific Fisheries Commission (WCPFC, `https://www.wcpfc.int`). The WCPFC conducts stock assessments for several species of tunas and implements fishery management and conservation measures based on these stock assessments. These WCPFC stock assessments and conservation and management measures may assist the WCPFC in regulating large scale fisheries in the stock habitat, but offer little useful advice to managers of small scale fisheries operating outside of the core areas.

Yellowfin tuna (*Thunnus albacares*, YFT) is an important food resource of many Pacific Island communities and the people of the Hawaiian islands of been fishing for yellowfin for many generations. Recent studies (Rooker et al. 2016) have shown that the Hawaii-based fishery for YFT is supported by local production with little or no subsidies from equatorial production zones or nurseries. The Main Hawaiian Islands (MHI), Figure 1, comprising the 8 largest islands in the Hawaiian Archipelago, has been an important fishing ground throughout this history. Since the recent expansion of the Papahānaumokuākea Marine National Monument in 2016, the MHI is the only legally accessible tuna fishing ground in the State of Hawaii. This change emphasizes the importance of careful management of yellowfin fisheries in the MHI.

The primary purpose of this paper is to explore a means to apply the 65 year time series of data collected from local fisheries to the problem of providing advice to fisheries managers about the effects of local fisheries on the local portion of the larger stock of yellowfin tuna in the Pacific. The general approach is to develop a relatively simple population dynamics model, estimate the model parameters from existing data, and apply the estimated parameters to answer the question of the effects of the fishery. I adopt a state-space variant of the logistic surplus production model with several novel features. Fishing mortality, the proportion of the population harvested each year, is represented as a random walk of multiple gear types. Offline coupling is used to link the local biomass to regional biomass estimates from large scale yellowfin stock assessment models developed for the WCPFC. The logistic parameters of the surplus production model are reparameterized to estimate parameters of direct relevance to fishery management. All state

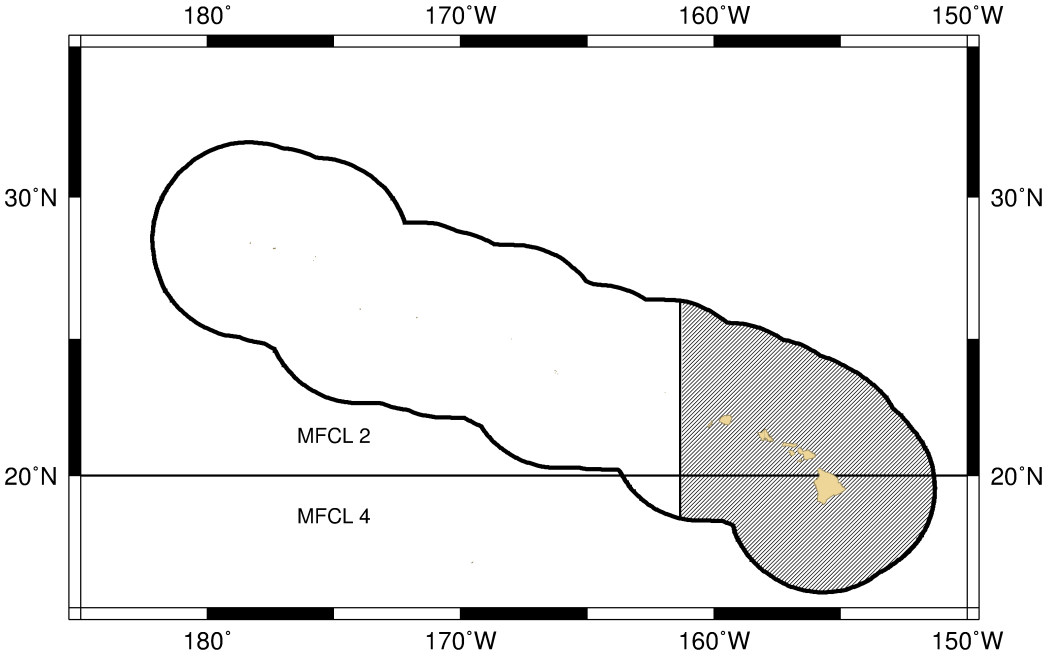

Figure 1: The United States Exclusive Economic Zone around the Hawaiian Archipelago. The Main Hawaiian Islands lie in the shaded area at the extreme east of the EEZ. The line at 20°N latitude is the boundary between WCPFC stock assessment regions 2 and 4. See text for full explanation.

transitions are represented as random effects.

Methods

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 model is

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

where α_t is the state at time t and the function T embodies the dynamics mediating the development of the state at time t from the state at the previous time with random process error, η_t.

Stock dynamics follow the classic Schaefer (1954) differential equation:

$$\frac{dN}{dt} = rN(1 - \frac{N}{K}) - FN \tag{2}$$

where N is the biomass of YFT in the MHI (mt), r is the logistic growth rate (yr^{-1}), K is the asymptotic biomass (mt), and F is the total fishing mortality (yr^{-1}) in the MHI.

The state space transition equation is developed by solving (2) analytically from one time to the next and applying a random error.

$$N_t = \frac{K(r - \bar{F}_t)}{\frac{K(r-\bar{F}_t)}{N_{t-\Delta t}} e^{-\Delta t(r-\bar{F}_t)} - re^{-\Delta t(r-\bar{F}_t)} - r} \cdot e^{\eta_t}; \quad \eta_t \sim N(0, \sigma_N^2) \tag{3}$$

where $\eta_t \sim N(0, \sigma_N)$ is a process error expressing variability in population dynamics and $\bar{F}_t$ is the total fishing mortality exerted by all gears, i. e.,

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

The logarithm of fishing mortality for each gear is assumed to follow a random walk with normal increments, as suggested by Nielsen and Berg (2014).

$$\log F_{g,t} = \log F_{g,t-1} + \xi_t; \quad \xi_t \sim N(0, \sigma_F^2) \tag{4}$$

where ξ_t is a process error expressing the year to year variability in fishing mortality.

An index abundance assumes that the biomass of YFT in the MHI is approximately proportional to some index which is independent of the data analyzed in the model,

$$\log N_t - \log(Q \cdot I_t) = \omega_t; \quad \omega_t \sim N(0, \sigma_I^2) \tag{5}$$

where I_t is the index at time t, Q is the estimated ratio of the MHI population size to the index, and ω_t is a process error representing the difference between the MHI biomass and the abundance index. Conventional fishery independent methods, such as acoustic or trawl surveys, are not available for tunas. Biomass estimates from the most recent WCPFC stock assessment (Davies, et al. 2014) are convenient potential abundance indices. The assessment method used by the WCPFC is spatially structured and estimates biomass in several regions in the Pacific Ocean. The boundary between two of these regions passes directly through the MHI (Figure 1). Biomass estimates from both regions were evaluated as abundance indices.

The logistic parameters r and K are notoriously difficult to estimate in surplus production models, so parameter substitutions were used. Maximum sustainable yeld (MSY, $\widetilde{Y}$) and fishing mortality at MSY ($F_{\widetilde{Y}}$) are commonly used fishery reference points and are simple functions of r and K. Therefore $\widetilde{Y}$ and $F_{\widetilde{Y}}$ were estimated directly and substituted in (3) as $r = 2F_{\widetilde{Y}}$ and $K = \frac{4\widetilde{Y}}{r}$.

Carruthers and McAllister (2011) recommend use of Bayesian priors for the logistic growth parameter, r, in equation (2). They suggest $\tilde{r} = 0.486$ with a standard deviation of $\sigma_r = 0.046$ (a coefficient of variation cv $= \frac{\sigma_r}{r} = 0.094$) for Atlantic YFT. A lognormal prior on r was implemented as

$$\log r - \log \tilde{r} = \rho; \quad \rho \sim N(0, \sigma_r^2) \tag{6}$$

ρ becomes a component of the likelihood (equation 10).

All process errors are assumed to have the same distribution, $N(0, \sigma_P^2)$; σ_N,σ_F, and σ_I are assumed to be equal and estimated as a global process error with variance σ_P^2.

The general form of the state space observation model is

$$x_t = O(\alpha_t) + \varepsilon_t \tag{7}$$

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

Predicted catch, $\widehat{C}_{g,t}$, for each gear is the product of estimated fishing mortality and the biomass,

$$\widehat{C}_{g,t} = F_{g,t} \cdot \left(\frac{N_{t-\Delta t} + N_t}{2} \right) \cdot e^{\varepsilon_t} \tag{8}$$

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

$$\log \varepsilon_t = \begin{cases} C_{g,t} > 0: & (1 - p_0) \cdot \left(\log \frac{1}{\sqrt{2\pi\sigma_Y^2}} - \left(\frac{\log C_{g,t} - \log \widehat{C}_{g,t}}{\sigma_Y} \right)^2 \right) \\ C_{g,t} = 0: & p_0 \cdot \log \frac{1}{\sqrt{2\pi\sigma_Y^2}} \end{cases} \tag{9}$$

where $C_{g,t}$ is the observed catch for gear g at time t, σ_Y is the observation error and p_0 is the proportion of observed catch observations equal to zero.

The model states, N_t and F_{gt}, are assumed to be random effects (Skaug and Fournier, 2006). Model parameters are estimated by maximizing the joint likelihood of the random effects, observations, and prior likelihood.

$$L(\theta, \alpha, x) = \prod_{t=2}^{m} \left[\phi\big(\alpha_t - T(\alpha_{t-1}), \Sigma_\eta\big) \right] \cdot \prod_{t=1}^{m} \left[\phi\big(x_t - O(\alpha_t), \Sigma_\varepsilon\big) \right] \cdot \rho \tag{10}$$

where m is the number of time steps in the catch time series and $\theta = \left\{ \widetilde{Y}, F_{\widetilde{Y}}, Q, \sigma_P, \sigma_y \right\}$ is a vector of model parameters (Table 1). The model is implemented in ADMB-RE (Fournier et al. 2012) and applied to a 61 year ($m = 61$) time series of observed catch by four fishing gears from 1952 through 2012 (Figure A.3). Data preparation is described in Appendix A. Noncommercial fishers may land substantial quantities of YFT, but reliable data from this sector of the fishery are not currently available. The effects of omitting the noncommercial catch is discussed briefly in Appendix D. All computer code and data files discussed in this paper can be found at Github: `https://github.com/johnrsibert/XSSA.git`.

Results

Table 2 compares 6 different models using three index of abundance assumptions and two assumed values of the standard deviation of the r prior, σ_r.

Table 1: Complete list of estimated and computed parameters and model constraints for the state-space surplus production model. There are 5 estimated parameters. Computed variables are functions of estimated parameters. There are four constraints and constants. For a model with 61 time steps and 4 gear types, there are 305 random effects.

Parameter	Definition
	Estimated parameters:
$\widetilde{Y}$	Maximum sustainable yield
$F_{\widetilde{Y}}$	Fishing mortality at maximum sustainable yield
Q	Abundance index proportionality constant
σ_P	Global process error standard deviation, $\sigma_P = \sigma_N = \sigma_F = \sigma_I$
σ_Y	Observation error standard deviation
	Computed variables:
r	Instantaneous growth rate, $r = 2F_{\widetilde{Y}}$
K	Asymptotic population size, $K = \frac{4\widetilde{Y}}{r}$
$\bar{F}_5$	Average of total estimated fishing mortality for the most recent five years.
	Constraints and constants:
p_0	Proportion of zero catch observations, fixed at 0.04918
$\tilde{r}$	An *a priori* assumed value for r fixed at $\tilde{r} = 0.486$ (Carruthers and McAllister, 2011)
σ_r	Assumed standard deviation of r around its prior, fixed at $\sigma_r = 0.8$ or $\sigma_r = 0.2$; cv = 1.64 or cv = 0.411
$\bar{Y}_5$	Average of total observed catch for the most recent five years.

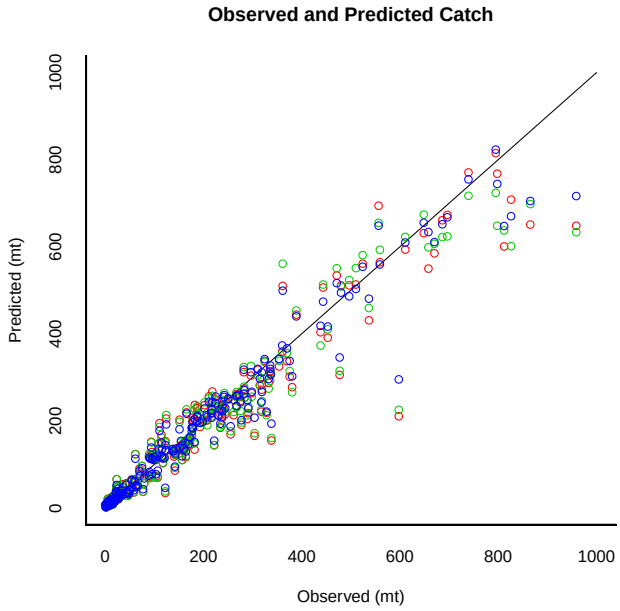

Figure 2: Plot of predicted catch on observed catch for models A, B, and C.

Models with the loose constraint on r, $\sigma_r = 0.8$, and either region 2 or region 4 abundance indices converged to solution, but the model without an index of abundance did not converge. All models with a tighter constraint on r, $\sigma_r = 0.2$ converged to solutions with similar likelihood values and similar parameter estimates, $K \approx 8000$ and $r \approx 0.5$. For convenience, I will refer to models with region 2 indexing and $\sigma_r = 0.8$, region 4 indexing and $\sigma_r = 0.8$, and no indexing and $\sigma_r = 0.2$ as models A, B, and C respectively (Table 2). Model A, indexed to region 2, predicts a large, slowly growing population ($K = 23900; r = 0.232$) whereas model B, indexed to region 4, predicts a small, rapidly growing population ($K = 2130; r = 1.84$). Model C, not indexed, predicts an intermediate situation ($K = 8140; r = 0.497$). In models A and B, the prediction is dominated by the abundance index, and in model C the prediction is dominated by the prior assumption about r.

All models are equally capable of predicting catch, Figure 2. Low catches seem to be predicted accurately, but some high catches are estimated with larger errors, as is expected with log-normal errors. See Figure C.1 for more detail.

Table 2: Results for 6 different models using three abundance indices and two constraints on r. The first section of the table shows model diagnostics. $-\log L$ is the negative logarithm of equation 10; lower values indicate better fits to the data. $|G|_{max}$ is the gradient of the likelihood with respect to the parameters; values of $|G|_{max} > 1e^{-4}$ indicate lack of model convergence. n is the number of parameters estimated. Model Name is the identifier used to discuss the models in the text.

Index Region	2	4	None	2	4	None
Model Name	A	B				C
$-\log L$	199.03	192.65	270.87	198.95	195.04	195.1
n	5	5	4	5	5	4
$\|G\|_{max}$	9.2e-05	4.79e-05	0.138	2.17e-05	5.87e-05	1.32e-05
$\tilde{Y}$	1380	980	5530	1110	940	1010
$F_{\tilde{Y}}$	0.116	0.92	0.436	0.205	0.294	0.249
Q	0.0721	0.000774	—	0.0285	0.00235	—
σ_P	0.233	0.234	0.321	0.245	0.231	0.319
σ_Y	-1.46	-1.45	-1.14	-1.41	-1.46	-1.14
r	0.232	1.84	0.872	0.409	0.588	0.497
K	23900	2130	25400	10900	6400	8140
$\bar{F}_5$	0.0632	1.52	0.749	0.163	0.513	0.149
p_0	0.0492	0.0492	0.0492	0.0492	0.0492	0.0492
$\tilde{r}$	0.486	0.486	0.486	0.486	0.486	0.486
σ_r	0.8	0.8	0.8	0.2	0.2	0.2

Estimated biomass tracks abundance indices for both models A and B, Figure 3, and the abundance indices lie within the process error of the biomass trend. The estimated biomass and abundance index for model A (region 2) reach maxima in the 1980s and subsequently decrease. The estimated biomass and abundance index for model B (region 4) decrease monotonically. The estimated biomass trend in model C is roughly similar to that of model A with a peak in the 1980s and subsequent decrease. The maxima of the biomass trends in all models exceed the estimated equilibrium biomass (K).

The average total YFT catch in the MHI for the most recent 5 years is 818 mt, easily below the estimated $\widetilde{Y}$ for all three models. The average estimated fishing mortality for the most recent 5 years is 0.0632, 1.52 and 0.149 for models A, B and C respectively (Table 2). Models A and C estimate the current fishing mortality to be about half of $F_{\widetilde{Y}}$. Model B (region 4 indexing) estimates the current fishing mortality to be greater than $F_{\widetilde{Y}}$.

Discussion

Failure of the model to converge to a solution without an index of abundance and using the loose prior on r (Table 2) indicates that the data lack sufficient information for this model. The qualitative similarity between the biomass trajectories predicted by models A and C suggests that there is sufficient information in the data to estimate a biomass trend similar to region 2 without an index of abundance, but that additional information is required to scale the trend. An index of abundance or an informative prior constraint on a parameter are alternative assumptions that provide additional information to a model to constrain model behavior. Interpretations of these two assumptions are quite different. Assuming a prior value for r implies *a priori* knowledge of stock productivity, and imposing a low standard deviation around the prior implies a high level of certainty. On the other hand, assuming that local stock size is proportional to stock size in a larger geographic range implies that factors mediating the abundance of the local stock are similar to those mediating the abundance in the larger area.

MFCL Regions 2 and 4 differ from one another in their oceanography, ecology and fisheries (Appendix B). Region 2 is outside the center of abundance for yellowfin in the WCPO, appears to have a smaller stock, and Davies et al. (2014) estimate that region 2 has experienced relatively minor fishery

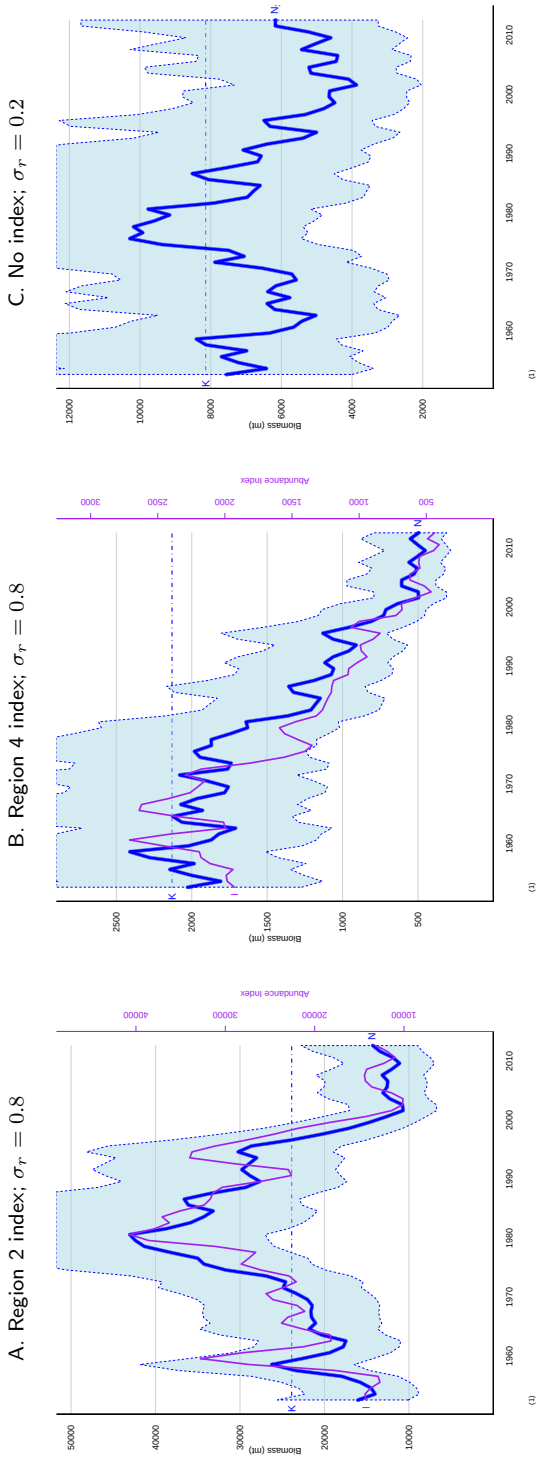

Figure 3: Predicted biomass trends with different indices of abundance and values of σ_r. Blue lines indicate the estimated biomass in metric tons (mt). The purple line indicates the index of abundance. The light blue shaded areas represent the process error as $\pm 2\sigma_P$. The equilibrium biomass K is indicated by the blue dot-dash line. Note that the scale of the ordinate is different in each panel.

impact. In contrast, region 4 hosts some of the largest fisheries in the WCPO, produces a large proportion of the total WCPO yellowfin catch, and Davies et al. (2014) estimate that region 4 has experienced the largest impact of the fishery of any of the MFCL regions. On the basis of low levels of exploitation and the location of the Hawaiian archipelago in the North Pacific Tropical Gyre province, MFCL region 2 would seem *a priori* to be the best choice for an index of abundance.

The random walk representation of fishing mortality enables the model to compute accurate values of $F_{g,t}$. The zero-inflated log normal observation error and the separation of fishing fleets enable the model to interpret low (or zero) catches as changes in $F_{g,t}$ rather than as episodes of low stock size. These features enable the model to estimate catch from the estimated stock extremely accurately.

Current catches ($\bar{Y}_5$) are less than model estimates of maximum sustainable yield ($\widetilde{Y}$) for all model variants. Estimates of current fishing mortality ($\bar{F}_5$) are less than estimates of fishing mortality at maximum sustainable yield ($F_{\widetilde{Y}}$) for all models except model B (region 4 index of abundance). If fishery managers were to adopt $\widetilde{Y}$ and $F_{\widetilde{Y}}$ as management reference points, the conclusion would be that the local stock is not overfished and that overfishing is not occurring.

The model presented here is a promising approach to assessing a local portion of a larger fish stock in a way that provides useful advice to local fisheries managers. The current model is clearly in need of further development. Estimates of noncommercial catch (and possibly their observation error) should be included in the data. The issue of selecting an appropriate index of abundance should be resolved. One hopes that the next round of WCPFC stock assessments will realign the MFCL regions to be more consistent with the Longhurst (1998) ecological provinces so that the entire Hawaiian Archipelago lies within a single region. A region that aligns approximately with the North Pacific Tropical Gyre Province would be the best choice for an index of abundance for the MHI.

Acknowledgements. This work was initiated under a contract from 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 take on this challenging project. Special thanks are due to David Itano for sharing insights into the biology of yellowfin tuna and into small-boat fisheries in Hawaii, to Reginald Kokubun of the State of Hawaii, Division of Aquatic Resources, Department of Land and Natural Resources, for supplying catch report data from the Commercial Marine Landings data base, to Keith Bigelow and Karen Sender of NOAA Pacific Island Fisheries Science Center for supplying logbook reporting data and from the PIFSC data base. Thanks also to John Hampton of the Secretariat of the Pacific Community, Oceanic Fisheries Programme, for making available MULTIFAN-CL output files from the latest Western and Central Pacific Fisheries Commission yellowfin tuna stock assessment, and to Nick Davies for advice on interpreting those files.

References

Carruthers, T. and M. McAllister. 2011. Computing prior probability distributions for the intrinsic rate of increase for Atlantic tuna and billfish using demographic methods. Collect. Vol. Sci. Pap. ICCAT, 66(5): 2202-2205.

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 for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software 27, 233249.

DLNR. 2011. HMRFS Newsletter. Hawaii Department of Land and Natural Resources.https://dlnr.hawaii.gov/dar/fishing/hmrfs/

Longhurst, A. 1998. *Ecological Geography of the Sea.* Academic Press, San Diego. 398pp.

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.

Rooker, J.R., R. J. D. Wells, D. G. Itano, S. R. Thorrold, J. M. Lee. 2016. Natal origin and population connectivity of bigeye and yellowfin tuna in the Pacific

Ocean. Fish. Oceanogr. 25, 277-291.

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

Senina, I., P. Lehodey, B. Calmettes, S. Nicol, S. Caillot, J. Hampton and P. Williams. 2015. SEAPODYM application for yellow tuna in the Pacific Ocean. WCPFC-SC11-2015/EB-IP-01.

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

Wilson, P. T. 2011. *AKU!*. Xlibris, USA. 368pp. ISBN 978-1-4568-5904-6.

A Data preparation

The State of Hawaii Department of Land and Natural Resources Commercial Marine Landings data base (CML) from 1949 through 2014 is the primary source of data used in this analysis. This database documents the continuous 65 year history of commercial fishing in Hawaii.

The CML catch reports were aggregated into 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". Landing in the "Misc" category are highest after year 2000 and less than 2% of the total landings. The catch time series for the CML data are shown in Figure A.1. The "Bottom/inshore HL" and "Tuna HL" are both handline gears and the data for these two gear types were combined at the suggestion of the CML database administrator. The "Aku Boat" gear type refers to storied Japanese-style pole and line fishery that operated in Hawaii through th 1980s (Wilson 2011). The Aku Boats target skipjack tuna (*Katsuwonus pelamis*), but yellowfin were occasionally landed as incidental catch. Some time series contain sustained periods of zero catch which reflect the development and subsequent shift away from a specific gear type. It is assumed that these declines in catches represent "collapse" of a fishery due to social and economic factors rather than to a decline of YFT stocks. Similarly, 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 size.

The Hawaii-based longline fishery began a rapid expansion in the late 1980s, and the United States National Oceanic and Atmospheric Administration (NOAA) began to collect data from the longline fleet under a federally mandated logbook program in 1990. The CML longline data were augmented by NOAA longline log sheet data from 1995 through 2013. NOAA distinguishes deep sets, targeting bigeye tuna (*Thunnus obesus*), and shallow sets, targeting swordfish (*Xiphias gladius*) in the data. The CML data do not distinguish between deep and shallow sets. Since the longline fleet ranges widely in the North Pacific Ocean, only catches reported in the United States EEZ around Hawaii east of 162°W longitude were included in the data. Figure A.2 shows the correspondence between the CML and NOAA time series. The combined deep plus shallow catches from NOAA align fairly well with the overlapping CML data. The simple average of the CML data with the

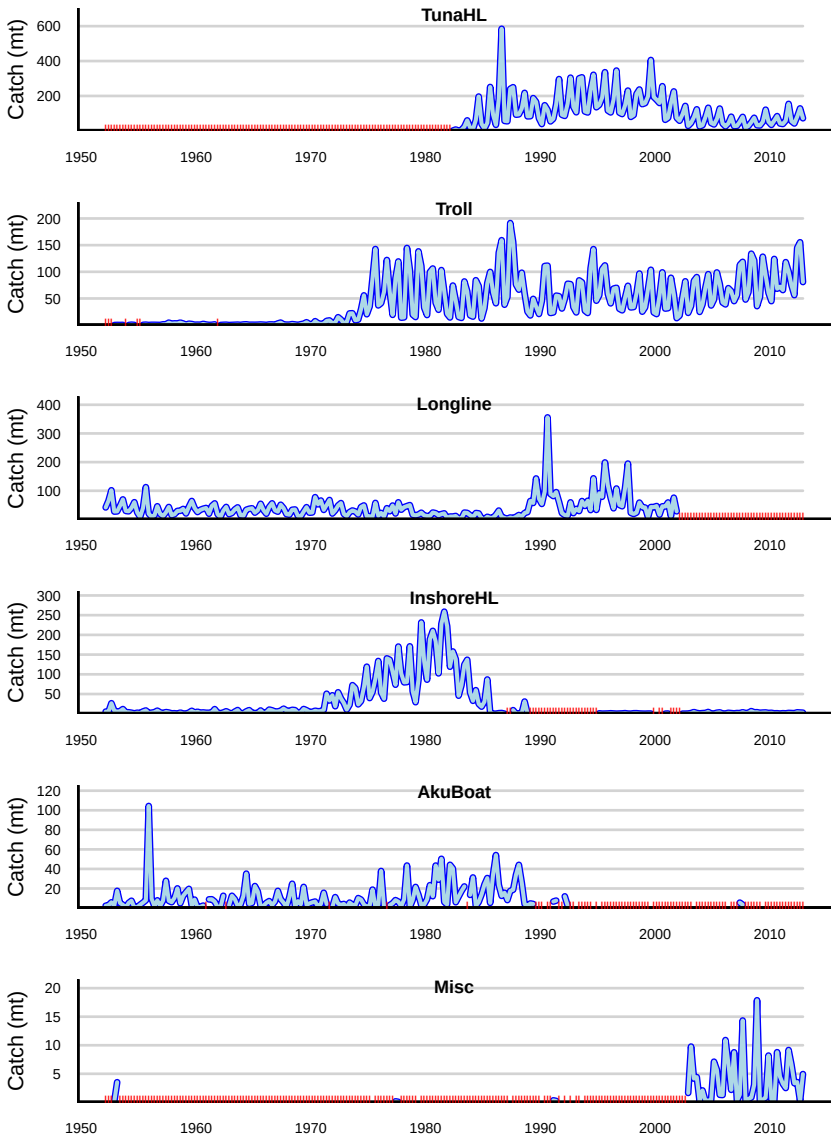

Figure A.1: Yellowfin catch in metric tonnes by principle fisheries operating in the Main Hawaiian Islands from the State of Hawaii Commercial Marine Landings Data. The red tick marks on the abscissa indicate quarters where reported catches were zero.

combined NOAA deep plus shallow data appears to be roughly the same trajectory as the constituent time series. Yellowfin is currently considered an incidental catch in the longline fishery.

Catch data from the CML and NOAA longline data bases were initially aggregated by quarter of the calendar year. All catch time series exhibit annual cycles suggesting strong seasonal signals in the catches by all gears. To avoid the need to estimate autocorrelation matrices for each time series and to minimize the number of zero catch observations, the quarterly time series were aggregated into the annual time series shown in Figure A.3 and used in this analysis.

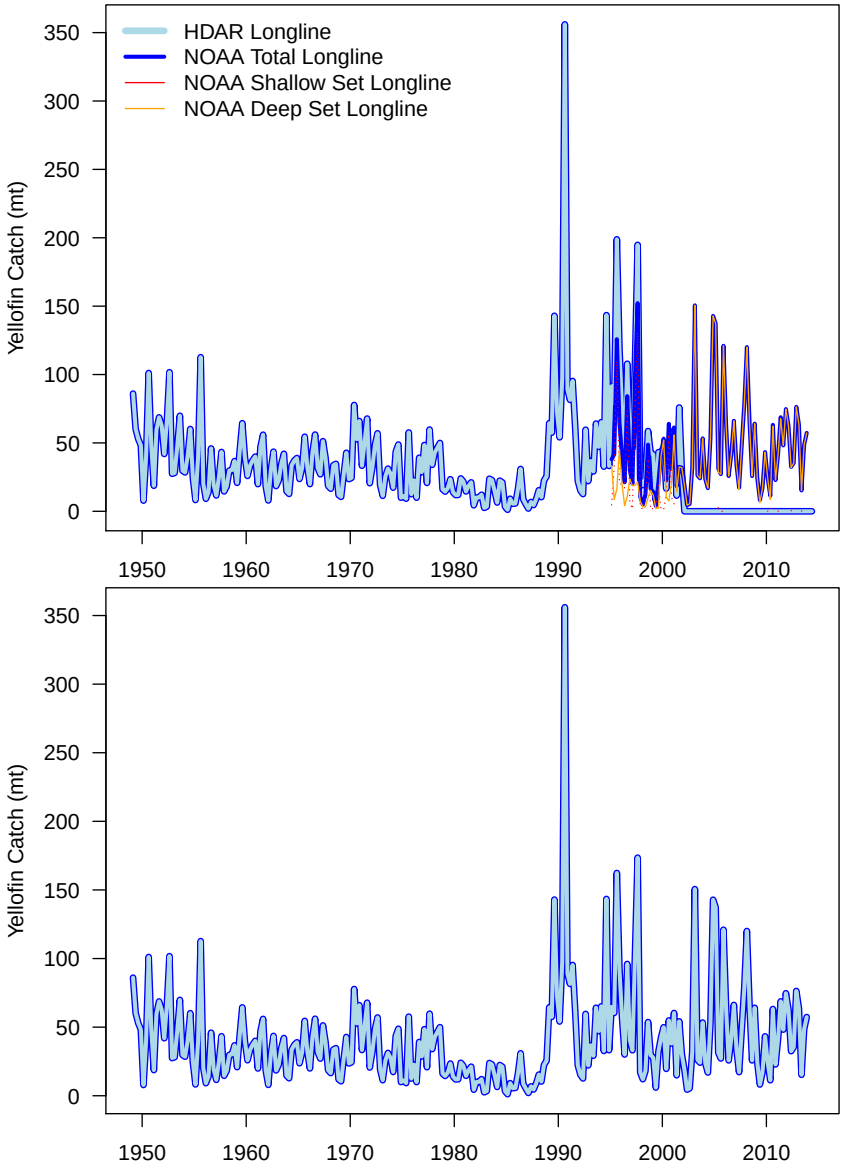

Figure A.2: Comparison between CML and NOAA longline time series. The upper panel shows the NOAA deep and shallow set data superimposed on the HDAR data. The lower panel shows the time series produced by a simple average of the CML data and the sum of the NOAA deep and shallow catches.

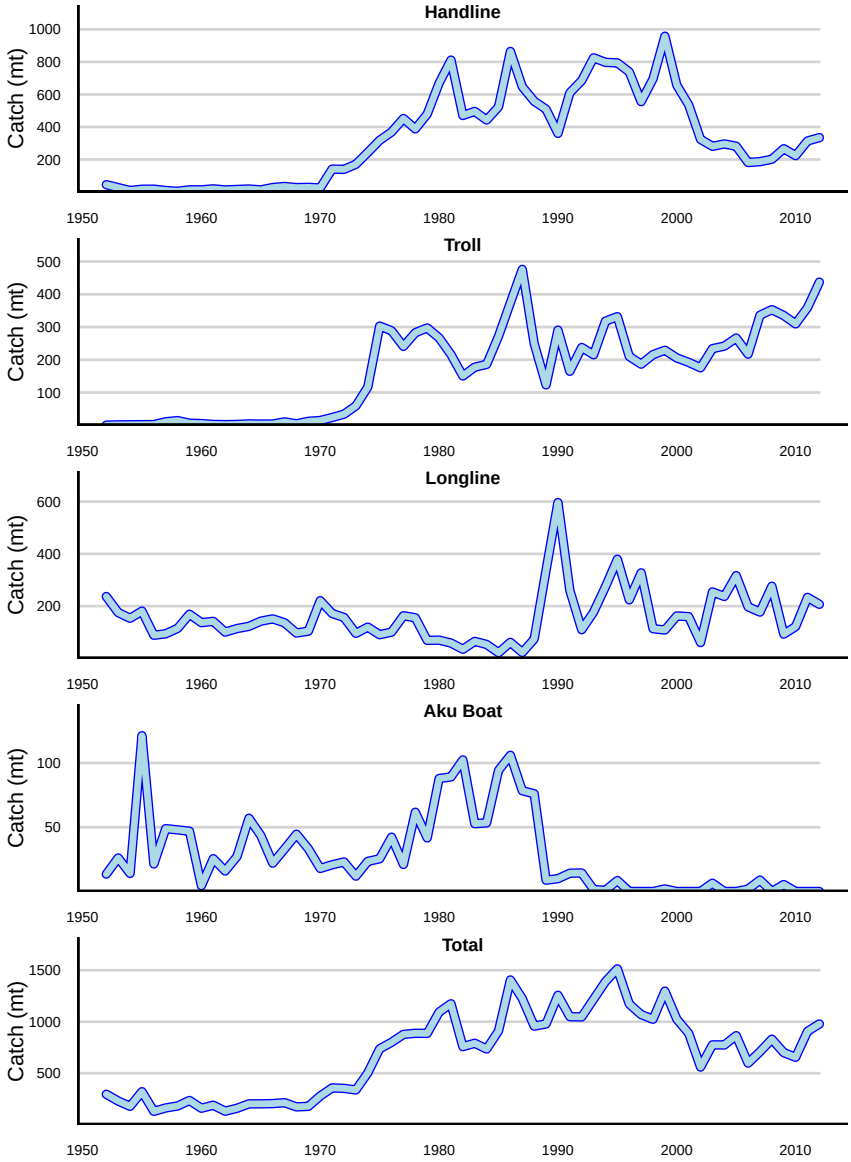

Figure A.3: Four gear type catch time series used in the assessment model.

B Abundance indices and offline coupling

Using the output of one model to constrain a second model is sometimes referred to as "offline" coupling. The output from two quite different models of WCPO yellowfin, SEAPODYM and MULTIFAN-CL, could potentially be used as indices of abundance. Both models are age structured, spatially structured, and have been applied to YFT. SEAPODYM is an ocean basin scale model with 1° spatial resolution, whereas the MULTIFAN-CL yellowfin assessment model is constrained to the western Pacific and partitions the model domain into 9 regions. Figure B.1. The spatial resolution of the SEAPODYM model would make it an ideal candidate for computing an index of abundance, but the model and its application to yellowfin are still under active development (Senina et al, 2015). The 2014 MULTIFAN-CL assessment (Davies et al. 2014) has been officially adopted by the WCPFC as the assessment on which regional conservation and management measures are based.

The Hawaiian archipelago lies in the eastern extension of Longhurst's (1998) North Pacific Tropical Gyre Province (NPTG), Figure B.2. The boundary between MFCL regions 2 and 4 splits both MHI and the NPTG. Region 2 includes the northern extent of the range of YFT, and the MFCL assessment concludes that the YFT biomass is relatively low and that the impact of the WCPO fisheries has been relatively minor. In contrast, region 4 lies in primarily in the Western Pacific Warm Pool Province, includes the core of the YFT range in the Pacific, and supports some of the most intense tuna fisheries in the world. MFCL estimates that the impact of the fisheries on the YFT stock in region 4 to be one among the highest of all MFCL regions, and it is likely that overfishing is occuring. The estimated biomass trends in these two regions are quite different, Figure B.3. Selecting a specific biomass time series for an index of abundance for the MHI stock implicitly assumes that the ecology and productivity of the index population is comparable to the MHI. On the basis of low levels of exploitation and the location of the Hawaiian archipelago in the NPTG, MFCL region 2 would seem *a priori* to be the best choice of a for computing an index of abundance.

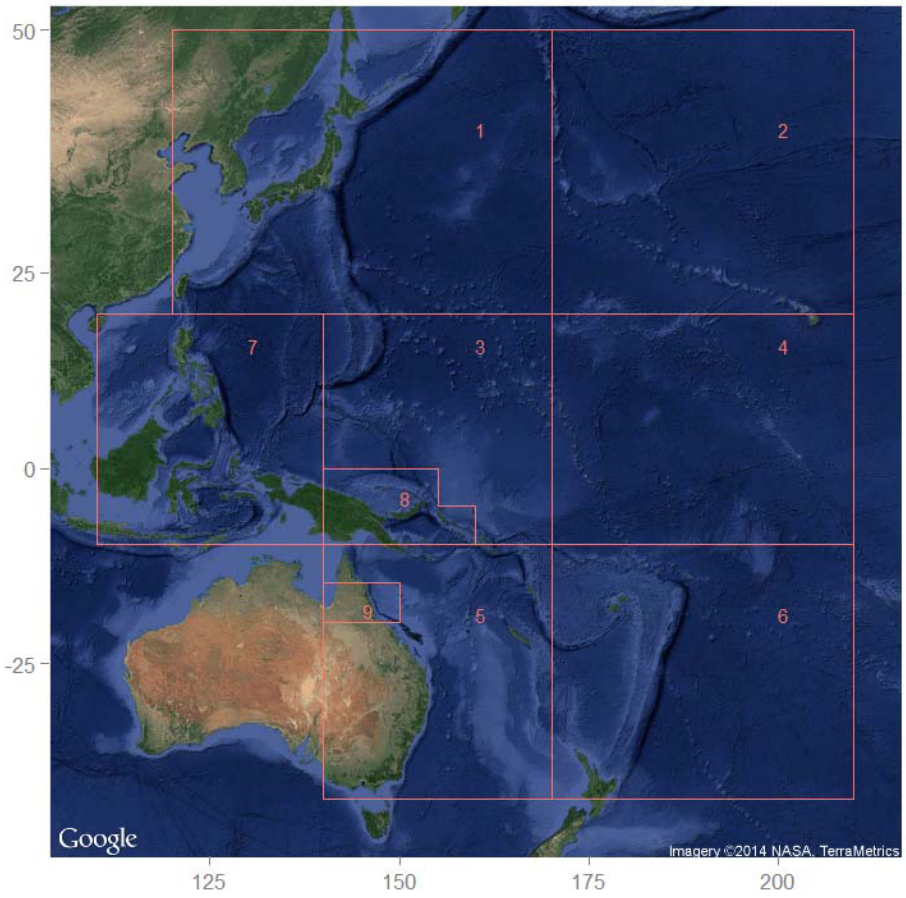

Figure B.1: Regional structure used in the 2014 WCPFC YFT stocks assessment from Davies et al. (2014).

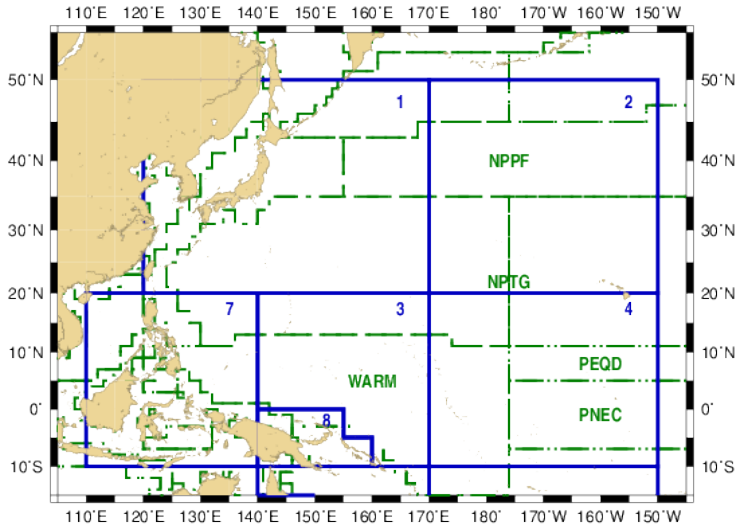

Figure B.2: Map of the central north Pacific Ocean showing selected MULTIFAN-CL stock assessment regions (blue lines and numbers) and Longhurst (1998) ecological provinces (dashed green lines and labels); NPFF, North Pacific Transition Zone Province; NPTG, North Pacific Tropical Gyre Province; WARM, Western Pacific Warm Pool Province; PEQD, Pacific Equatorial Divergence Province; PNEC, North Pacific Equatorial Countercurrent Province.

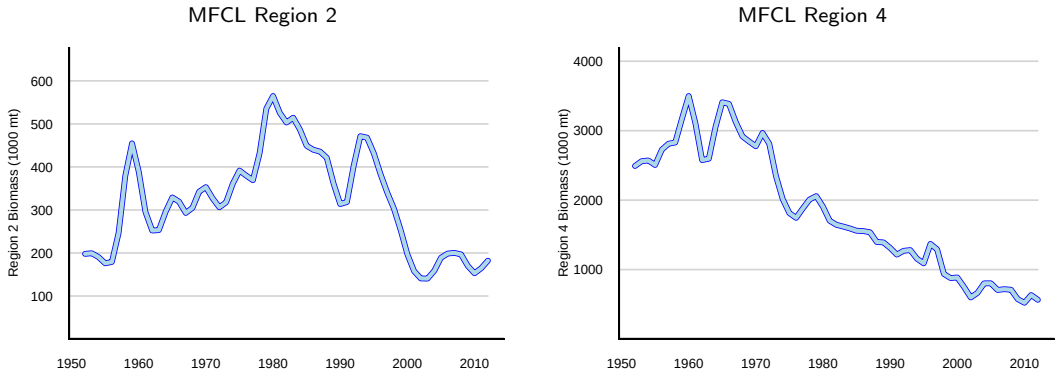

Figure B.3: Estimated biomass trends in MULTIFAN-CL regions 2 and 4 (Davies, et al. 2014).

C Additional Diagnostics

Figure C.1 shows the trends in fishing mortality and catch with different indices of abundance and values of the r prior, σ_r. The solid brown lines in the upper panel show the estimated fishing mortality by fleet. The orange shaded areas represent the process error as $\pm 2\sigma_P$. The solid green lines in the lower panel show the predicted catch; + symbols indicate observed catch. The light green shaded areas represent the observation error as $\pm 2\sigma_Y$. The ordinate in each panel is scaled to the specific gear type.

Figure C.2 places the model results in a fishery management context. The parabolic dashed red lines indicate the theoretical yield from a population with logistic growth. The maximum yield if the population were at equilibrium would occur at the peak of the parabola. There are many possible equilibria, but the maximum equilibrium yield has been labeled "Maximum Sustainable Yield", $\widetilde{Y}$, in the fisheries literature. The parameter $F_{\widetilde{Y}}$ is the fishing mortality that would produce $\widetilde{Y}$ at equilibrium.

Figure C.3 shows the frequency distributions of the parameter estimates obtained by sampling the output of 10^6 Markov Chain Monte Carlo simulations. Some parameters are clearly estimated more accurately than others. The process and observation error estimates have well defined modes and appear to be normally distributed around their point estimates. The distributions of other parameters depends on the model. For model B, indexed to a biomass trajectory with a well defined trend, estimates of $\widetilde{Y}$, $F_{\widetilde{Y}}$, and Q also seem well defined. The corresponding estimates are more ambiguous for model A. Figure C.3 also demonstrates the pervasive (some might say pernicious) effects of using an informative prior on r. Estimates of all parameters that depend on r appear to be normally distributed around the prior value.

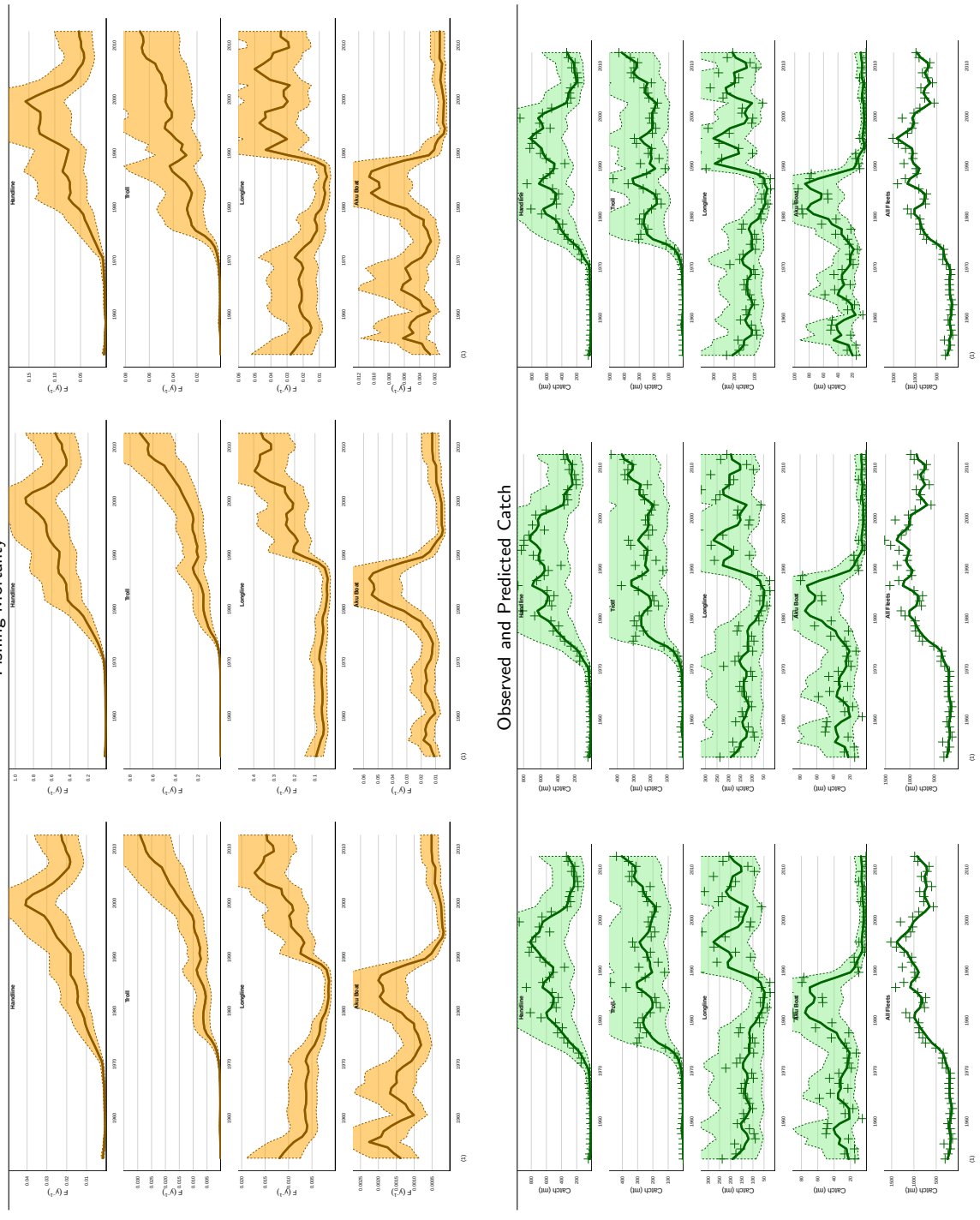

Figure C.1: Trends in estimated fishing mortality (top, orange) and catch (bottom, green).

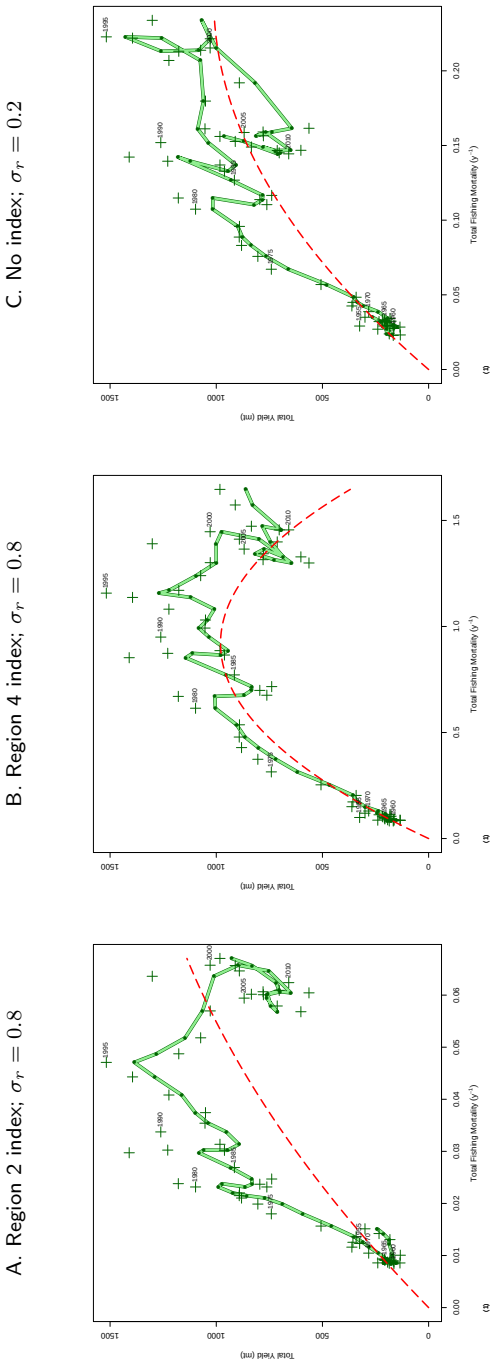

Figure C.2: Estimated production curves showing catch plotted against estimated fishing mortality with different indices of abundance and values of σ_r. The green line and dark green dots are estimated catch. The green + symbols are the observed catch annotated with the year. The dashed red line is the theoretical equilibrium yield. Note that the scale of the abscissa is different in each panel, and the ordinate has been scaled to the observed catch.

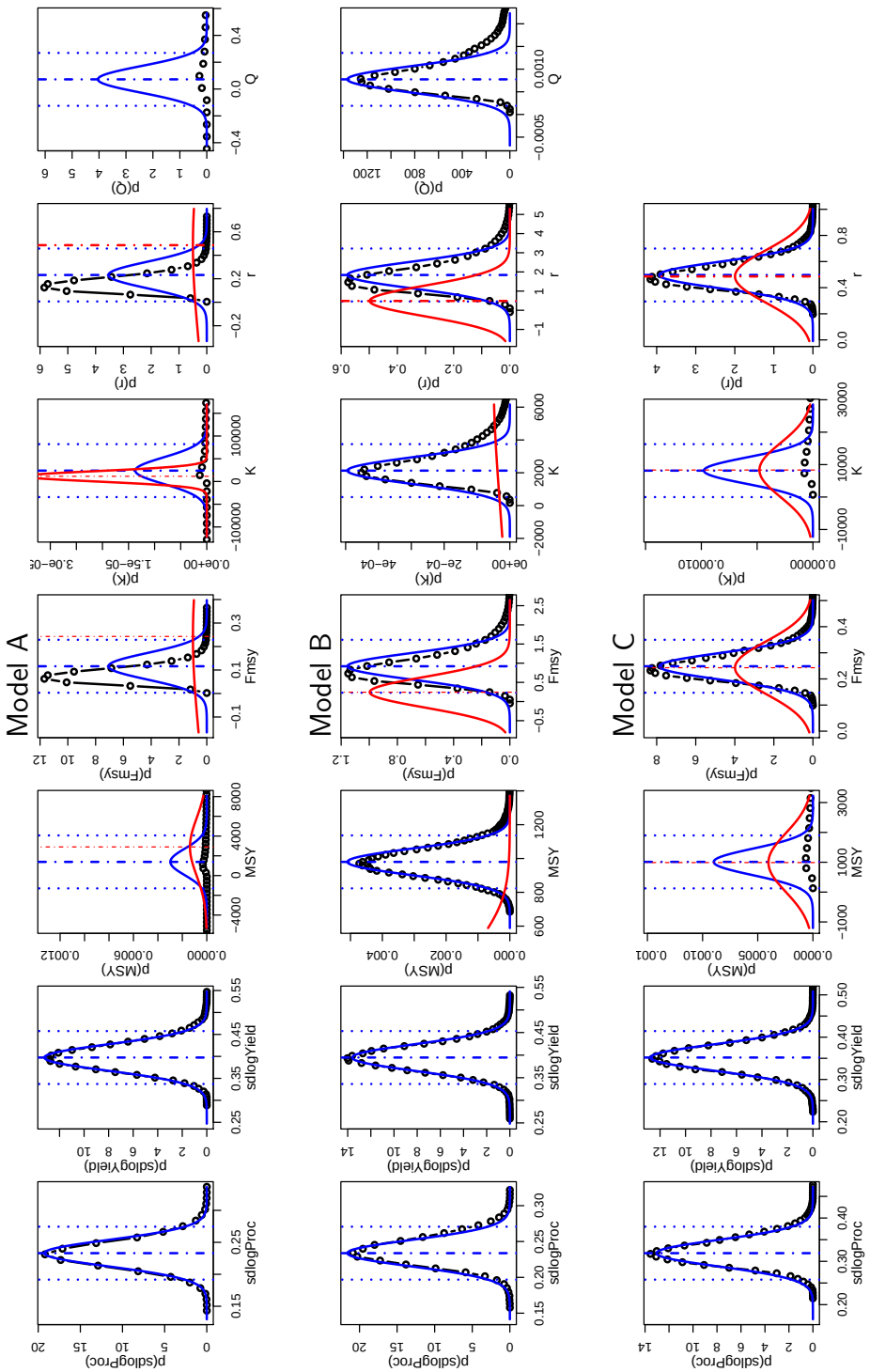

Figure C.3: Posterior distributions of parameter estimates from 10^6 MCMC samples. Black circles and connecting lines are frequencies of the estimates. Blue curves are the theoretical normal distribution of the maximum likelihood estimates computed from the standard deviations estimated from the inverse Hessian matrix at the minimum. Vertical blue dotted lines indicate $\pm 2\sigma$. The solid red curves indicate the effect of the prior on r; the vertical red line is the prior.

D Noncommercial Catch

The Main Hawaiian Islands support a large number of noncommercial fishers. There is no catch reporting system and no mandatory marine fishing license. Accurately estimating the noncommercial catch is therefore difficult. The State of Hawaii Department of Land and Natural Resources and the United States National Oceanic and Atmospheric Administration are collaborating to improve survey methods for the Hawai'i Marine Recreational Fishing Survey (HMRFS). The HMRFS Newsletter (DLNR, 2011) reports that non-commercial catches of YFT range between 2300 and 6900 mt from 2003 through 2010, considerably in excess of the most recent commercial catch. The assessment model developed here has the capability to augment the reported catch with catch by an additional fleet in proportion to the total reported catch and use the augmented catch in the assessment as if it were reported.

Tables D.1 and D.2 summarize the results of attempting to fit the observed catch data augmented by different multiples of the total catch. Four different multipliers were tested, 1, 2, 5, and 10 times total catch. All 8 augmented catch models converge to solutions. The negative log likelihood values are higher than the non-augmented fits as would be expected with more data points. *A priori* one would expect that a fish population capable of producing a higher than reported yield over a 60 year history would have a higher productivity and a higher biomass. The results of the augmented catch parameter estimates are generally consistent with this conclusion. The model compensates for the higher catch by increasing $\widetilde{Y}$ in all models. $F_{\widetilde{Y}}$ increases slightly, but the prior on r constrains the increases. Q increases in the abundance index models. $\bar{F}_5$ decreases substantially in the models without an abundance index.

This simple catch augmentation simulation indicates that the unreported noncommercial catch causes this model to underestimate maximum sustainable yield and the fishing mortality at maximum sustainable yield when applied to the available reported catch data. More realistic simulations should be developed to evaluate this problem.

Table D.1: Results of fitting the model to 4 different multiples of the total catch using MFCL Region 2 as an index of abundance and a loose prior on $r, \sigma_r = 0.8$, comparable to model A in Table 2. See Table 1 for definitions of all variables.

Multiplier	A	1	2	5	10
$-\log L$	199.03	220.1	220.12	220.14	220.14
n	5	5	5	5	5
$\|G\|_{max}$	9.2e-05	2.56e-05	4.44e-05	2.96e-05	5.97e-06
$\widetilde{Y}$	1380	2560	3880	7820	14400
$F_{\widetilde{Y}}$	0.116	0.123	0.123	0.123	0.123
Q	0.0721	0.122	0.185	0.375	0.691
σ_P	0.233	0.234	0.234	0.234	0.234
σ_Y	-1.46	-1.45	-1.45	-1.45	-1.45
r	0.232	0.246	0.246	0.245	0.245
K	23900	41700	63100	128000	235000
$\bar{F}_5$	0.0632	0.0745	0.0739	0.0733	0.0729
p_0	0.0492	0.0492	0.0492	0.0492	0.0492
$\tilde{r}$	0.486	0.486	0.486	0.486	0.486
σ_r	0.8	0.8	0.8	0.8	0.8
$\bar{Y}_5$	818	1630	2450	4890	8960

Table D.2: Results of fitting the model to different multiples of the total catch with no index of abundance and a tighter prior on $r, \sigma_r = 0.2$, comparable to model C in Table 2. See Table 1 for definitions of all variables.

Multiplier	C	1	2	5	10
$-\log L$	195.1	217.51	217.53	217.54	217.54
n	4	4	4	4	4
$\lvert G\rvert_{max}$	1.32e-05	1.62e-05	6.35e-05	4.14e-05	5.7e-05
$\widetilde{Y}$	1010	62100000	1.16e+11	4.41e+10	6.14e+10
$F_{\widetilde{Y}}$	0.249	0.244	0.244	0.244	0.244
σ_P	0.319	0.295	0.295	0.295	0.295
σ_Y	-1.14	-1.22	-1.22	-1.22	-1.22
r	0.497	0.488	0.488	0.488	0.488
K	8140	5.1e+08	9.56e+11	3.62e+11	5.04e+11
$\bar{F}_5$	0.149	3.17e-06	2.50e-09	1.24e-08	1.77e-08
p_0	0.0492	0.0492	0.0492	0.0492	0.0492
$\tilde{r}$	0.486	0.486	0.486	0.486	0.486
σ_r	0.2	0.2	0.2	0.2	0.2
$\bar{Y}_5$	818	1630	2450	4890	8960