Narrative

- Introduction

    – problem statement and motivation

    – introduce reference point and management decision making

    – introduce BH & Derizo

    – State the Bio Model (Pella v. Schaefer)

- Methods

    – Reference Point Derivation

    – Reference Point Inversion (Math) data generation

    – Profile likelihood on q and MLE parameterization

    – State Stats Model

    + distributional results (Appendix B) for reference point and biological inference

- Results

    – bias surface

    – table of estimated equilibrium values at examples

    – yield/SRR curves at example locations

    – Haddon r>-2.575 chaos, page 33

- ?discussion?

- Proposal (<6 pages)

    – challenges of filling out response space with three parameter SRR (numerical methods)

    – embedded simulator space filling (output space-filling)

        * Shepherd-BH, Derizo-BH?

        * Age Structured Models and Data Weighting

    – Age Structured Model and Data Weighting

Nicholas R. Grunloh

- – ?MSEs?

- – timeline (gant chart

- Appendicies

  - – Appendix A: SRR exposition and examples (reference to proposal section)

    - ∗ two v. three parameter

    - ∗ Pella v. Schaefer

    - ∗ ?Derizo Model?

    - ∗ Shepherd v. BH

  - – Appendix B: Distributional Results for SRR parameterizations

1 Introduction

Data for a typical surplus-production model comes in the form of a time series of observations of an index of abundance for some population of interest. The index is often observed alongside a variety of other known quantities, but at a minimum, each observed index will be observed in the presence of some known catch for the period. The index of abundance is assumed to be proportional to biomass with the proportionality constant being a nuisance parameter that is often referred to as the catchability parameter.
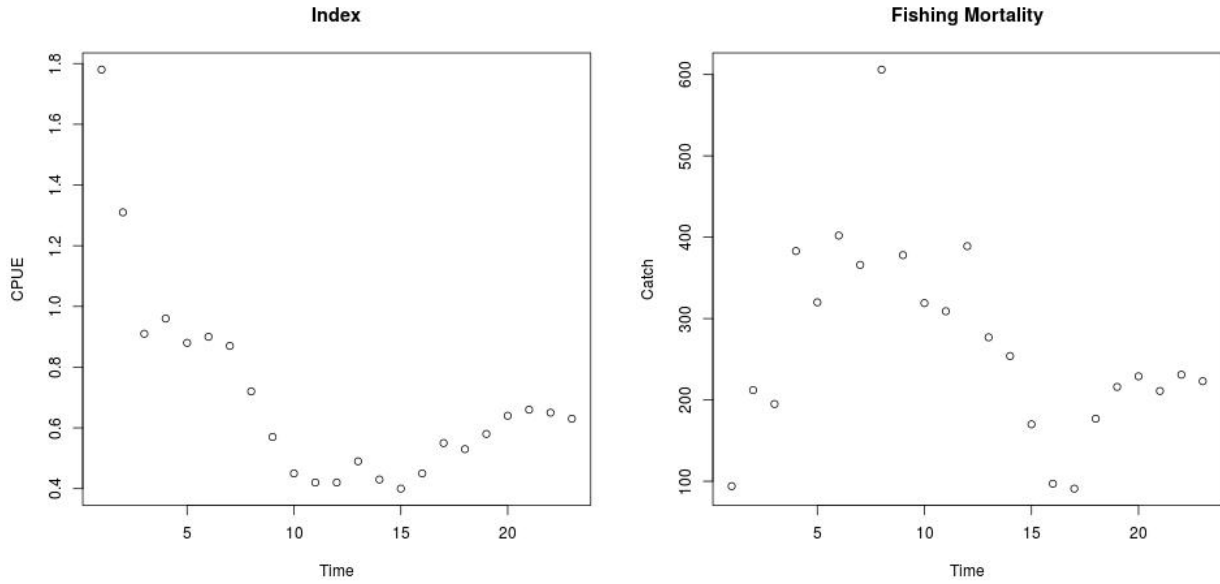

Figure 1: *left*: An observed series of index of abundance data for Namibian Hake from year to year. *right*: Catch data associated the observed Namibian Hake index over the same time period.

The observed indices are assumed to have multiplicative log-normal errors, and thus the following observation model arises naturally,

$$I_t = qB_t e^\epsilon \quad \epsilon \sim N(0, \sigma^2). \tag{1}$$

Above q is the catchability parameter and σ^2 models residual variation. Biologically speaking these parameters are often treated as nuisance parameters with the more biological parameters entering the model thru a process model on biomass.

3

Biomass is assumed to evolve as an ordinary differential equation; in this case I focus on the following form,

$$\frac{dB}{dt} = P(B; \boldsymbol{\theta}) - C. \tag{2}$$

Here biomass is assumed to change in time by two processes, net recruitment into the population, and catches removing biomass from the population.

Firstly, the population grows through a stock recruitment relationship (SRR). Recruitment in this setting is defined as the net biomass increase due to all birth, maturation, and migration processes after accounting for all other naturally occurring sources of mortality other than the recorded fishing from humans. The recruitment function is assumed to be parametric function that relates the current biomass of the population to an aggregate production of biomass.

Secondly, the population decreases as biomass is removed due to catch (C). While catches are observable quantities (cite), the model assumes that catch is proportional to biomass with the proportionality constant representing the fishing rate (F), so that $C = FB$. From a management perspective a major goal of the model is to accurately infer a quantity known as *maximum sustainable yield* (MSE). One could maximize simple yield at a particular moment in time (and only for that moment) by fishing all available biomass in that moment. This strategy is penny-wise but pound-foolish (not to mention ecologically devastating) since it doesn't leave biomass in the population to reproduce for future time periods. We seek to fish in a way that allows (or even encourages) future productivity in the population. This is accomplished by maximizing the equilibrium level of catch (or yield) over time. Equilibrium yield is considered by replacing the steady state biomass ($\bar{B}$) in the assumed form for catch, so that $\bar{C} = F\bar{B}(F)$, where $\bar{}$ indicates a value at steady state. Naturally the steady state biomass is a function of F; we will see a specific example of this in Section (2.1). MSY is found by optimizing $\bar{C}(F)$ with respect to F, and F^* is the fishing rate at MSY. Going forward let * decorate any value derived under the condition of MSY.

The canonical production model in fisheries is the Schaeffer model. The Schaeffer model is formed by choosing P to be logistic growth (cite) parameterized so that $\boldsymbol{\theta} = [r, K]$ and

the family takes the following form,

$$P(B; [r, K]) = rB\left(1 - \frac{B}{K}\right). \tag{3}$$

r is parameter controlling the maximum reproductive rate of the population in the absence of competition for resources (i.e. the slope of SRR at the origin). K is the so called "carrying capacity" of the population. The carrying capacity can be formally stated as steady state biomass in the absence of fishing $\bar{B}(0)$.
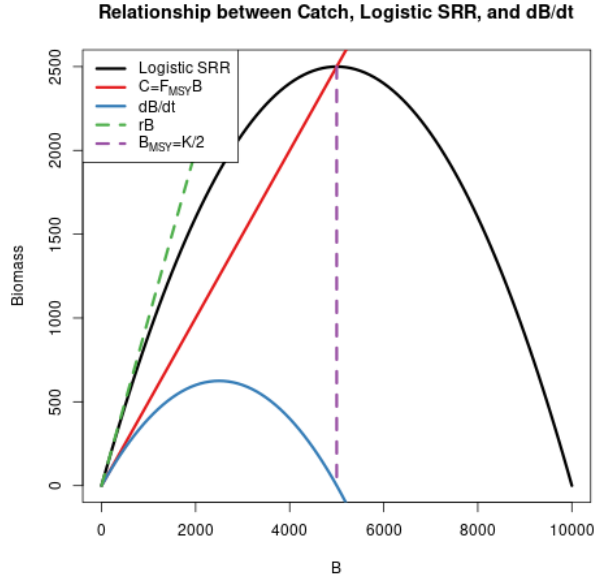

Figure 2:

Logistic SRR produces idealized parabolic recruitment with equilibrium quantities taking very simple forms that can be easily understood from the graphical construction seen in Figure (2). Positive recruitment is observed when $B \in (0, K)$. Due to the second order parabolic shape of the logistic SRR it is straightforward to see that MSY will be maximized by fishing at the peak productivity of the stock. By symmetry it is clear that this peak occurs at $B^* = \frac{K}{2}$. The fishing rate required to position the stock at MSY is $F^* = \frac{r}{2}$, which is half of the stock's maximum reproductive rate at the origin. In the absence of fishing $\frac{dB}{dt}$ would be driven entirely by P, but at MSY $\frac{dB}{dt} = 0$ since the population equilibrates at B^*.

While this idealized form is instructive, and convenient, these simplistic dynamics are also

5

potentially problematic. The symmetry of the logistic functional form is very rigid in that it assumes dynamics in the lower (mate limited) regime of the dynamics ($B \in (0, \frac{K}{2})$) have the same shape as the upper (density limited) regime of the recruitment dynamics ($B \in (\frac{K}{2}, K)$). What in nature ties these phenomena together so that this assumption should be true?

Fisheries are very often managed based upon reference points (RP) which serve as simplified heuristic measures of population behavior. The mathematical form of RPs depends upon the model assumptions primarily thru the SRR (cite). Here the focus is on two RPs that will be denoted by ξ and ζ going forward.

$$\xi = \frac{F^*}{M} \qquad\qquad \zeta = \frac{B^*}{B_0} \qquad\qquad (4)$$

ξ is the optimal (in the MSY sense) fishing rate rescaled so that it is in terms relative to natural mortality. ζ is the biomass at MSY relative to the unfished virgin biomass of the population. In general $\xi \in \mathbb{R}^+$ and $\zeta \in (0, 1)$, however under the under the assumption of logistic recruitment these quantities take the following form,

$$\xi = \frac{r}{2M} \qquad\qquad \zeta = \frac{1}{2} \qquad\qquad (5)$$

so that $(\xi, \zeta) \in \left(\mathbb{R}^+, \frac{1}{2}\right)$.

In practice, at this time, the SRR is typically chosen to depend only on two parameters. Above the Schaeffer model is presented, but other common two parameter choices of the SRR are the Beverton-Holt (BH) and Ricker curves. All of these two parameter SRRs struggle similarly to model the full space of reference points.

Nature does not generate data from a simple model. However, what would happen if nature were even just slightly more complex than the most commonly used fisheries models? I consider data simulated from the three parameter Pella-Tomlinson (PT) SRR model, and subsequently fit these data using the two parameter Schaeffer model to observe the consequences of SRR model-misspecification with special interest on RF management quantities.

The three parameter PT family has a convenient form that includes the logistic SRR as

a special case to form the Schaeffer model. The Pella-Tomlinson SRR is parameterized so that $\boldsymbol{\theta} = [r, K, \gamma]$ and the family takes the following form,

$$P(B; [r, K, \gamma]) = \frac{rB}{\gamma - 1} \left(1 - \frac{B}{K}\right)^{\gamma - 1}. \tag{6}$$

γ is a parameter which breaks PT out of the restrictive symmetry of the logistic curve. In Figure (3) PT recruitment is shown for a range of parameter values so as to demonstrate the various recruitment shapes that can be achieved by PT recruitment.
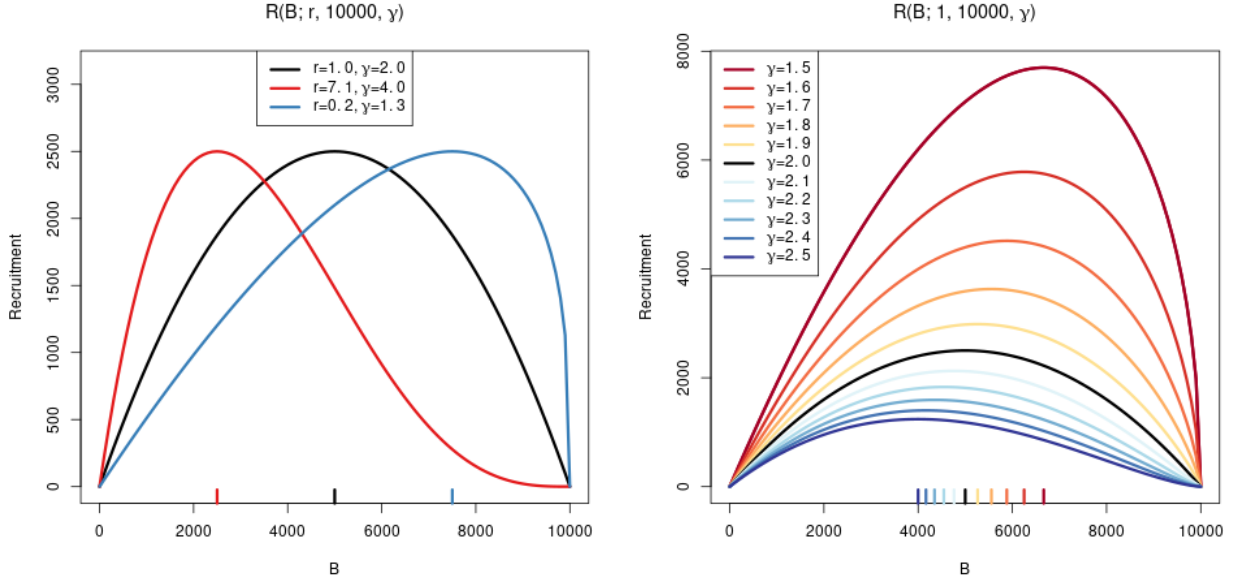

Figure 3:

While the particular form of how γ appears in PT still produces some limitations to the form of the SRR, importantly the introduction of a third parameter allows enough flexibility to fully describe the space of reference points used in management. To see this, the reference points are analytically derived for the PT model in the following section.

- introduce fisheries type data

- ODE modeling formulation

- Schaeffer Model (Logistic production function)

7

- Pella-Tomlinson production function

- Reference Points

- general simulation idea

2 Methods

2.1 PT Reference Points

With $B(t)$ representing biomass at time t, under PT production, the dynamics of biomass are defined by the following ODE,

$$\frac{dB}{dt} = \frac{rB}{\gamma - 1}\left(1 - \frac{B}{K}\right)^{\gamma - 1} - FB. \tag{7}$$

An expression for the equilibrium biomass is attained by setting Eq(7) equal to zero, and rearranging the resulting equation to solve for B. Thinking of the result as a function of F gives,

$$\bar{B}(F) = K\left(1 - \left(\frac{F(\gamma - 1)}{r}\right)^{\frac{1}{(\gamma - 1)}}\right). \tag{8}$$

At this point it is convenient to notice that $\bar{B}(0) = K$. The expression for B^* is given by evaluating Eq(8) at F^*.

To get an expression for F^*, the equilibrium yield is maximized with respect to F,

$$F^* = \underset{F}{\operatorname{argmax}} F\bar{B}(F). \tag{9}$$

In the case of PT production this maximization can be done analytically (however many three parameter production functions do not result in tractable analytical solutions). In this case maximization can proceed by differentiating the equilibrium yield with respect to F as follows,

$$\frac{d\bar{Y}}{dF} = \bar{B}(F) + F\frac{d\bar{B}}{dF} \tag{10}$$

$$\frac{d\bar{B}}{dF} = -\frac{K}{F(\gamma-1)}\left(\frac{F(\gamma-1)}{r}\right)^{\frac{1}{\gamma-1}}. \tag{11}$$

Setting Eq(10) equal to 0, substituting $\bar{B}(F)$ and $\frac{d\bar{B}}{dF}$ by Equations (8) and (11) respectively, and then solving for F produces the following expression for the fishing rate required to produce MSY,

$$F^* = \frac{r}{\gamma-1}\left(\frac{\gamma-1}{\gamma}\right)^{\gamma-1}. \tag{12}$$

Plugging the above expression for F^* back into Eq(8) gives the following expression for biomass at MSY,

$$B^* = \frac{K}{\gamma}. \tag{13}$$

The above derived expressions for $\bar{B}(0)$, B^*, and F^* can then be used to build a specific analytical form for the biological reference points in terms of only biological model parameters.

$$F^* = \frac{r}{(\gamma-1)}\left(\frac{\gamma-1}{\gamma}\right)^{\gamma-1} \qquad\qquad \frac{B^*}{\bar{B}(0)} = \frac{1}{\gamma} \tag{14}$$

2.2   Simulation Study

Indices of abundance are simulated from the three parameter PT production model over a grid of F^* and $\frac{B^*}{\bar{B}(0)}$ values. These PT data are then fit with a two parameter Schaefer model. Let ˜ decorate any quantity that is derived under the restricted two parameter SRR.

Generating simulated indices of abundance from the PT model requires inverting the relationship between $\left(F^*, \frac{B^*}{\bar{B}(0)}\right)$, and (r, γ). It is not generally possible to analytically invert this relationship for many three parameter production functions(cite Derizo paper). Most

9

three parameter production functions lead to RPs that require expensive numerical methods to invert; more over the numerical inversion procedure is often extremely unstable. That said, for the case of PT this relationship is analytically invertible, and leads to the following relationship

$$r = F^* \left(\frac{1 - \frac{B^*}{\bar{B}(0)}}{\frac{B^*}{\bar{B}(0)}} \right) \left(1 - \frac{B^*}{\bar{B}(0)} \right)^{\left(\frac{\frac{B^*}{\bar{B}(0)} - 1}{\frac{B^*}{\bar{B}(0)}} \right)} \qquad \gamma = \frac{1}{\frac{B^*}{\bar{B}(0)}}. \qquad (15)$$

$$r = F^* \left(\frac{\bar{B}(0) - B^*}{B^*} \right) \left(1 - \frac{B^*}{\bar{B}(0)} \right)^{\left(\frac{B^* - \bar{B}(0)}{B^*} \right)} \qquad \gamma = \frac{\bar{B}(0)}{B^*}. \qquad (16)$$

Indices are generated under the following conditions. Data are simulated at each point on the grid $\mathcal{F} \times \mathcal{B}$, with $F^* \in \mathcal{F}$ and $\frac{B^*}{\bar{B}(0)} \in \mathcal{B}$, where $\mathcal{F} = \{0.1, 0.2, ..., 0.7\}$ and $\mathcal{B} = \{0.2, 0.3, ..., 0.6\}$ as seen in Figure (4). These ranges of values for F^* and $\frac{B^*}{\bar{B}(0)}$ are selected to include the majority of most commonly assesed fish species cite. The red X's in Figure (4) show four 4 simulation locations where the Schaeffer model is misspecified to a large degree and will be considered in more detail in Section(). For each $\left(F^*, \frac{B^*}{\bar{B}(0)} \right)$, the associated pair (r, γ) are computed from Eq (15). Since K does not enter the RP calculation its value is fixed arbitrarily at 10000. The value of q is fixed at a typically small value of 0.0005. σ is fixed at the relatively small value of 0.01 to focus specifically on the behavior of population parameters. These parameters fully specify the PT model for the purposes of generating index data for each $\left(F^*, \frac{B^*}{\bar{B}(0)} \right)$ pair.
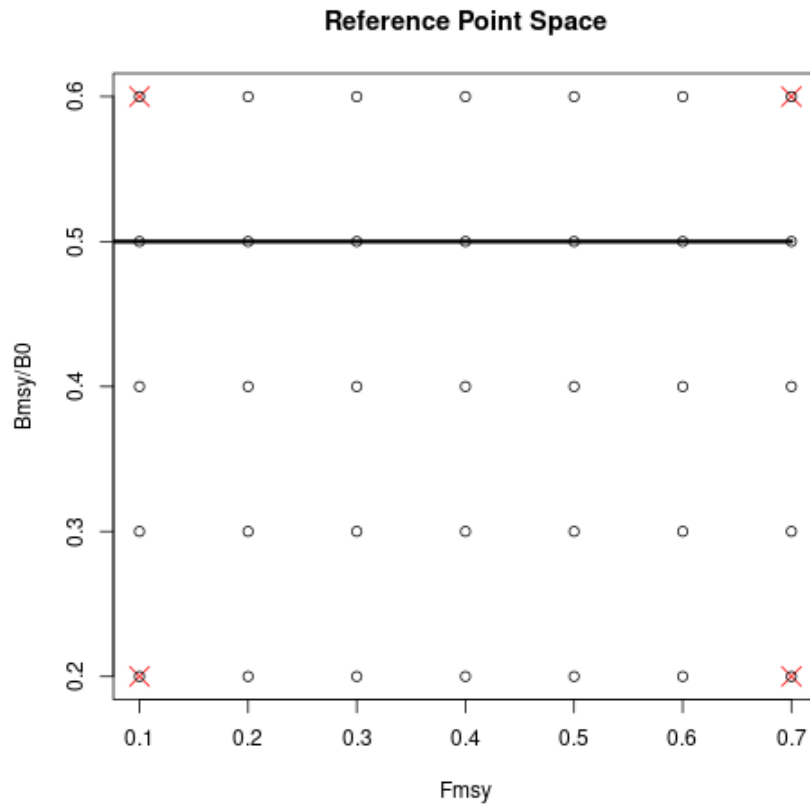
Figure 4: Open circles show the location of the simulation grid $\mathcal{F} \times \mathcal{B}$. The horizontal line shows the constrained space of RPs for the Schaeffer model. The red X's indicated 4 simulation locations where the Schaeffer model is a particularly misspecified.

2.3 Model Fitting

The goal of model fitting is to assess how the biological parameters of the two parameter Schaeffer model behave under MLE inference when fit to PT data. Thus, let I_t be an observation of PT index data at time t. The observation model is log-normal such that,

$$I_t | q, \sigma^2, \boldsymbol{\theta} \sim LN(qB_t(\boldsymbol{\theta}), \sigma^2). \tag{17}$$

For the Schaeffer model $\boldsymbol{\theta} = [r, K]$, and $B_t(\boldsymbol{\theta})$ is defined by the solution of the following ODE

$$\frac{dB}{dt} = rB\left(1 - \frac{B}{K}\right) - FB. \tag{18}$$

The I_t are assumed independent conditional on q, σ^2, r, K and the ODE model for biomass. Thus the log likelihood can be written as

$$\log \mathcal{L}(q, \sigma^2, \boldsymbol{\theta}; I) = -\frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_t \log\left(\frac{I_t}{qB_t(\boldsymbol{\theta})}\right)^2. \tag{19}$$

In this setting, q is fixed at the true value of 0.0005 to focus on the inferential effects of model misspecification on biological parameters. σ^2, r, and K are reparameterized into the log scale as $log(\sigma^2)$, $log(r)$, and $log(K)$ and fit via MLE. σ^2 is allowed to be fit to assess overall model fit. Reparameterization of the parameters into the log scale improves the relability of optimization in addition to facilitating the use of hessian information for parameter estimate standard errors.

Given that the biological parameters enter the likelihood via a nonlinear ODE, and further the parameters themselves are related to each other nonlinearly, the likelihood function can often be difficult to optimize. A hybrid optimization scheme is used to maximize the log likelihood to ensure that a global MLE solution is found. The R package GA (cite GA) is

used to run a genetic algorithm to explore parameter space globally. Optimization occasional jumps into the L-BFGS-B local optimizater to refine optima within a local mode. The scheme functions by searching globally to iteratively improving hot starts for the local optimizer.

In Appendix (3) a profile likelihood method for estimating all of the parameters of the model is derived. The profile likelihood technique greatly improves the reliability of local optimizers when fitting the biological parameters alongside additional nuisance parameters. The catchability parameter q has the effect of rescaling biomass which can often function similarly the role of the carrying capacity parameter K. Thus, the structure of the likelihood may confound q and K, and for some data these parameters may only be weakly identifiable. Posing the model in a Bayesian context provides a convienient mechanism for managing these weak identifiability issues. In a tactful Bayesian formulation q and σ^2 may then be marginalized out of the joint posterior to yeild fast and reliable inference (Cite Maria DeYorio pdf).

2.4 integrating ODEs, Stiffness, and Interpolation.

a preface to regularity issues: identifiability, stiffness, and continuity.

2.4.1 Uniqueness, Continuity, and Identifiability

An important (and often overlooked) implementation detail is the solution to the ODE which defines the progression of biomass through time (See Eq(2)). As a statistical model it is of paramount importance that this ODE not only have a solution, but also that the solution be unique. Of primary concern, uniqueness of the ODE solution is necessary for the identifiability of the statistical model.

If the form of $\frac{dB}{dt}$ is at least Lipschitz continuous, then the Cauchy-Lipschitz-Picard theorem provides local existence and uniqueness of $B(t)$. Recall from Eq(2) that $\frac{dB}{dt}$ is separated into a term for population production, $P(B)$, and a term for removals via catch, FB. For determining Lipschitz continuity of $\frac{dB}{dt}$, the smallest Lipschitz constant of $\frac{dB}{dt}$ will be the sum of the constants for each of the terms $P(B)$ and FB separately. Typically any choice of $P(B)$ will be continuously differentiable, which implies Lipschitz continuity (since the set of continuous differentiable functions is a subset of the set of Lipschitz continuous functions). Thus, the assumed form of $P(B)$ does not typically introduce continuity concerns, unlike

some potential assumptions for catch.

In practice catch is determined by a series of observed, assumed known, catches. Catch observations are typically observed on a quarterly basis, but in practice may not be complete for every quarter of the modeled period. It is overwhelmingly common to discretize the ODE via Euler's method with integration step sizes to match the observation frequency of the modeled data. This is often convenient but can present several issues. This strategy often pushes the assumption of catch continuity under the rug, but for identifiability of the statistical model an implicit assumption of continuity of the catches is required. While mechanistically at the finest scale fishers must only catch discrete packets of biomass (i.e. individual fish), it is sensible to consider aggregate catches at the quarterly (or yearly) scale as accruing in a continuous way. Furthermore any assumption of continuity will be required to be at least Lipschitz continuous for the required regularity of the model.

Here I assume catches accrue linearly between observed catches. This assumption defines the catch function as a piecewise linear function of time, with the smallest Lipschitz constant for the catch term defined by the steepest segment of the catch function. This assumption represents one of the simplest ways of handling catch, while retaining Lipschitz continuity overall. Furthermore linearly interpolated catch is adequately parsimonious for the typical handling of catches.

2.4.2   Integration and Stiffness

2.4.3   Integration, Stiffness, and Chaos

As previously mentioned, the overwhelming majority of implementations of population dynamics models discretize the ODE using Euler's method with the integration step sized fixed so as to match the observation frequency. In this setting we explore model parameterizations that explore the full extent of biologically relevant reference points. This exercise produces some combinations of parameters that result in numerically stiff ODEs.

The concept of stiffness in ODEs is hard to precisely characterize (cite). Hairer and Wanner [5, p. 2] describe stiffness in the following pragmatic sense, "Stiff equations are problems for which explicit methods don't work". It is hard to make this definition more mathematically precise, but this

2.5 Gaussian Process Model

Recall that indices of abundance are simulated from the three parameter PT SRR over an unrestricted grid of ξ and ζ values. After data are generated, γ is then fixed so that the SRR reduces to the special cases previously described, and the restricted model is subsequently fit to the simulated indices.

By working with the biological models parameterized in terms of $\log\left(\tilde{F}^*\right)$ it tends to improve optimization convergence. Furthermore, the normality this induces on the log scale, via the Laplace approximation, yields Log-Normality on $\tilde{F}^*$. Let $\hat{\mu}$ be the maximum likelihood estimate (MLE) of $\log\left(\tilde{F}^*\right)$. Additionally let $\hat{\omega}$ be the inverted Hessian information of the log likelihood evaluated at $\hat{\mu}$.

A GP is a stochastic process generalizing the normal distribution to an infinite dimensional analog. GPs are often specified primarily through the choice of a covariance function which defines the relationship between locations in an index set. Typically the index set is spatial for GPs, and in this setting the model is in reference point space (ξ, ζ). A GP model implies an n dimensional multivariate normal distribution on the observations of the model and the covariance function fills out the covariance matrix for the observations.

Each iteration of the simulation produces a single fitted $\hat{\mu}_i$ at an associate (ξ_i, ζ_i) location

15

with $i \in \{1, ..., n\}$. $\hat{\boldsymbol{\mu}}$ is jointly modeled over the space of reference points as the following GP,

$$
\begin{aligned}
\boldsymbol{x} &= (\xi, \zeta) \\
\hat{\boldsymbol{\mu}} &= \beta_0 + \boldsymbol{\beta}' \boldsymbol{x} + f(\boldsymbol{x}) + \boldsymbol{\epsilon} \\
f(\boldsymbol{x}) &\sim \mathrm{GP}(0, \tau^2 R(\boldsymbol{x}, \boldsymbol{x}')) \\
\epsilon_i &\sim \mathrm{N}(0, \hat{\omega}_i).
\end{aligned}
\tag{20}
$$

$\hat{\omega}_i$ is the observed variance for $\hat{\mu}_i$ from inference in the simulation step. This model allows for the full propagation of inferred information from the simulation step to be propagated into the reference point meta-model.

Here R is the Matern correlation function (?, ?).

$$
\begin{aligned}
R(\boldsymbol{x}, \boldsymbol{x}') &= Matern(\|\boldsymbol{x} - \boldsymbol{x}'\|_{\boldsymbol{\Lambda}}; \nu) \\
\|\boldsymbol{x} - \boldsymbol{x}'\|_{\boldsymbol{\Lambda}} &= \sqrt{(\boldsymbol{x} - \boldsymbol{x}')^{\top} \boldsymbol{\Lambda}^{-1} (\boldsymbol{x} - \boldsymbol{x}')} \\
\boldsymbol{\Lambda} &= \begin{pmatrix} \lambda_\xi & 0 \\ 0 & \lambda_\zeta \end{pmatrix}.
\end{aligned}
\tag{21}
$$

R has an anisotropic separable form of $\|\boldsymbol{x} - \boldsymbol{x}'\|_{\boldsymbol{\Lambda}}$ to allow for differing length scales in the ξ and ζ axes. The flexibility to model correlations separately in the ξ and ζ axes is key due to the differences in the extent of the ξ and ζ domains marginally. λ_ξ and λ_ζ model the length scales for ξ and ζ respectively. Additionally the Matern models the smoothness of the relationship through the smoothness parameter, ν.
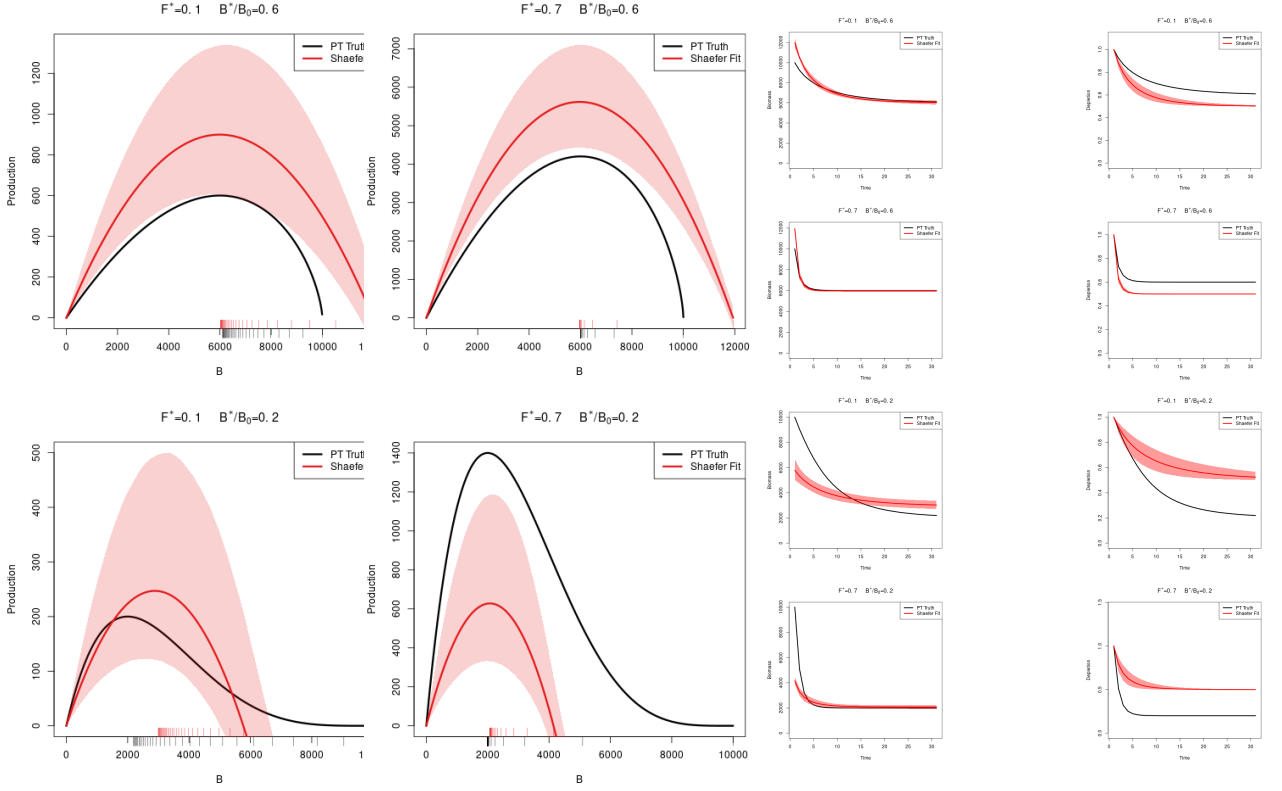
- GP serves as a flexible, stochastic interpolator for understanding the basic behavior of biases induced my limiting the SRR model.

- This model allows for the full propagation of inferred information from the simulation step to be propagated into the spatial meta-model in reference point space.
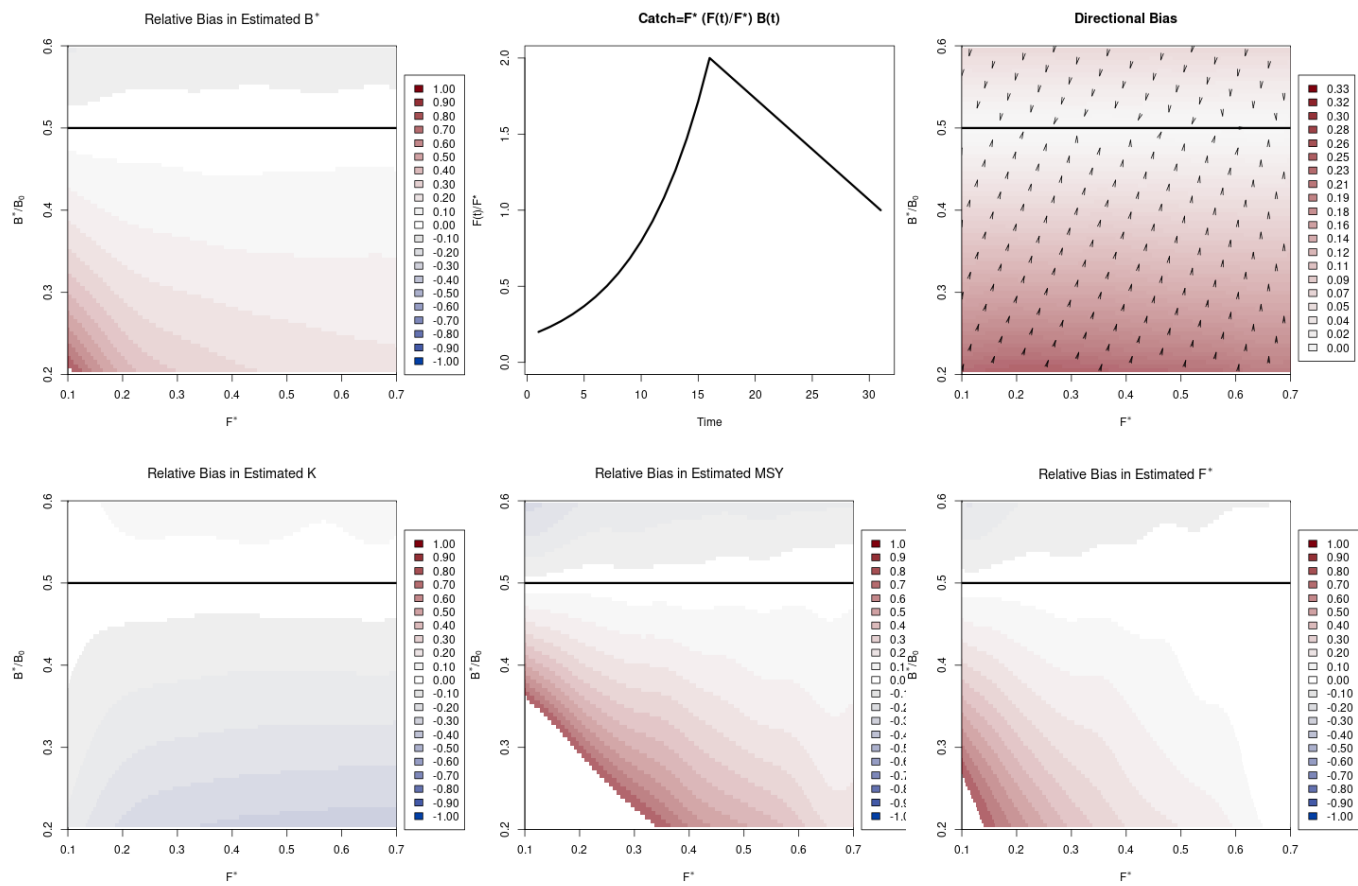
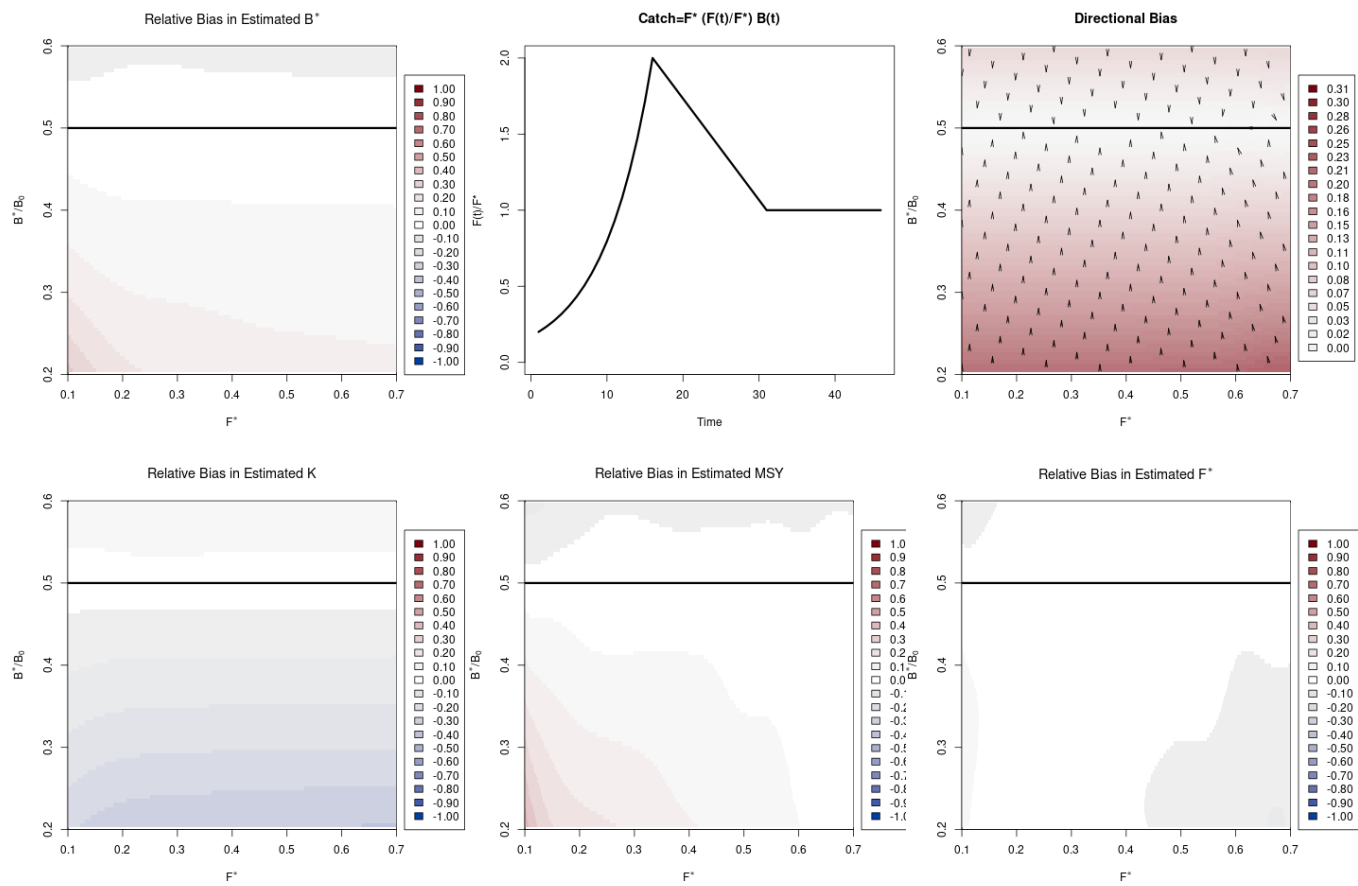- Motivate gaussianity

3   Results
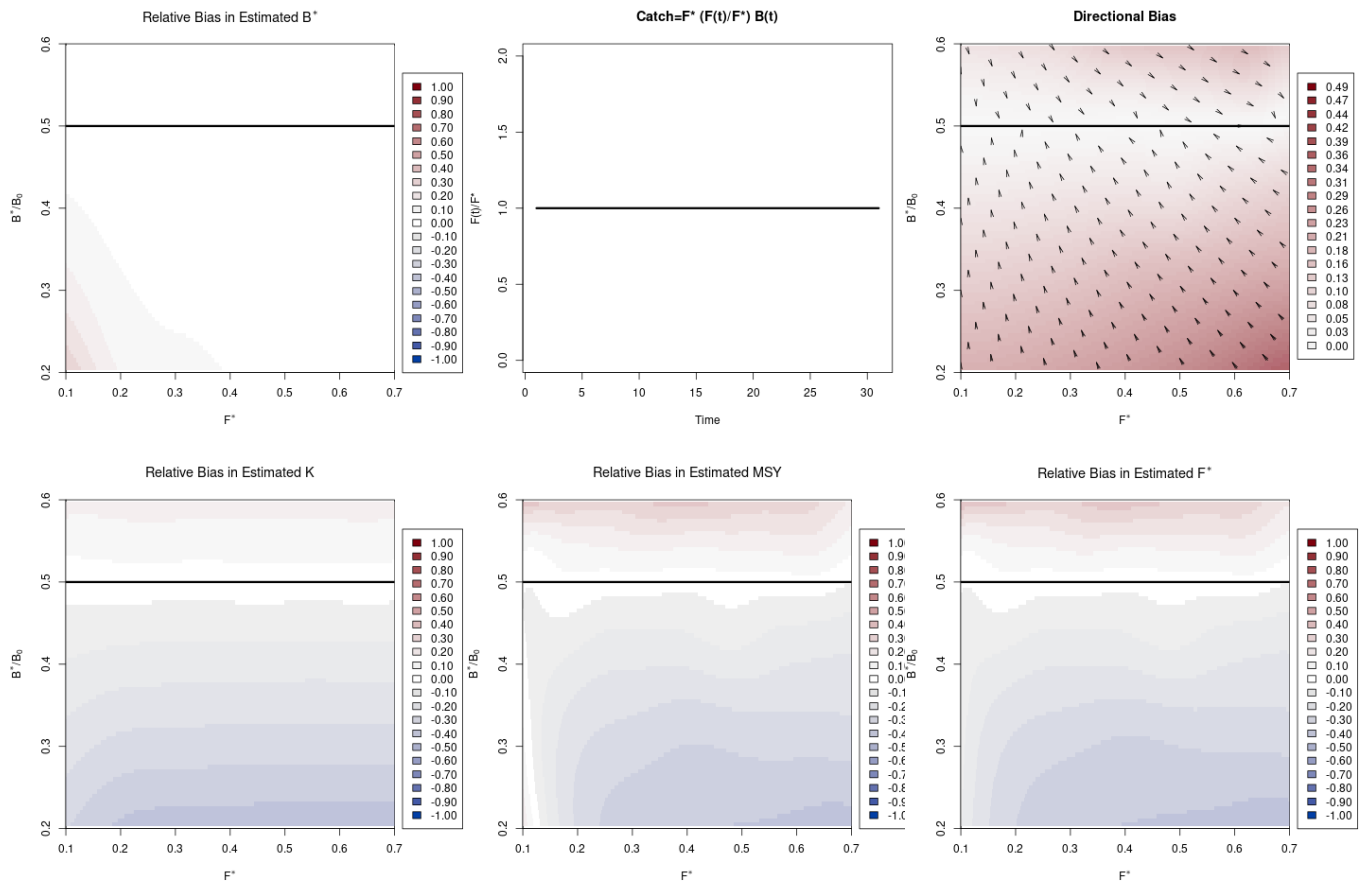
- Biomass story

- Fmsy Story

- cite lack of contrast

- transition to contrast example

(ξ, ζ)	$(\hat{\xi}, \hat{\zeta})$	$\hat{K}$	$\hat{r}$	$\hat{\sigma}$
$(0.5, 0.6)$	$(0.80, 0.5)$	11982.7	0.30	0.048
$(3.5, 0.6)$	$(4.70, 0.5)$	11934.8	1.90	0.031
$(0.5, 0.2)$	$(0.43, 0.5)$	5769.3	0.17	0.282
$(3.5, 0.2)$	$(1.52, 0.5)$	4145.7	0.61	0.184

Relative Bias in Estimated B^*

Catch=F* (F(t)/F*) B(t)

Directional Bias

Relative Bias in Estimated K

Relative Bias in Estimated MSY

Relative Bias in Estimated F^*

Relative Bias in Estimated B^*

Catch=F* (F(t)/F*) B(t)

Directional Bias

Relative Bias in Estimated K

Relative Bias in Estimated MSY

Relative Bias in Estimated F^*

Nicholas R. Grunloh

Directional Bias

Bias in Estimated Optimal Biomass

Figure 5: This is a caption

Bias in Estimated Optimal Fishing

ξ=1 ζ=0. 5
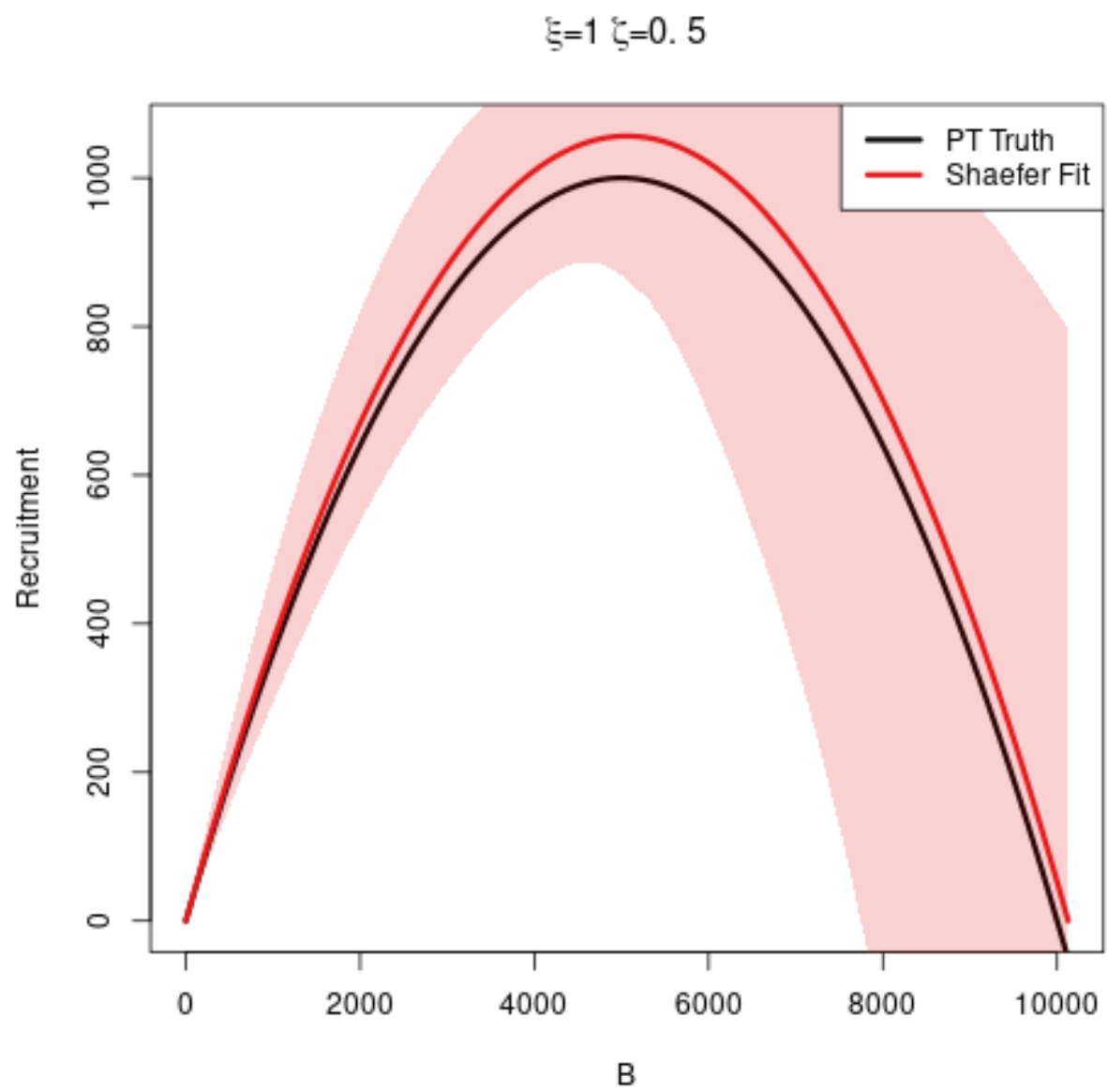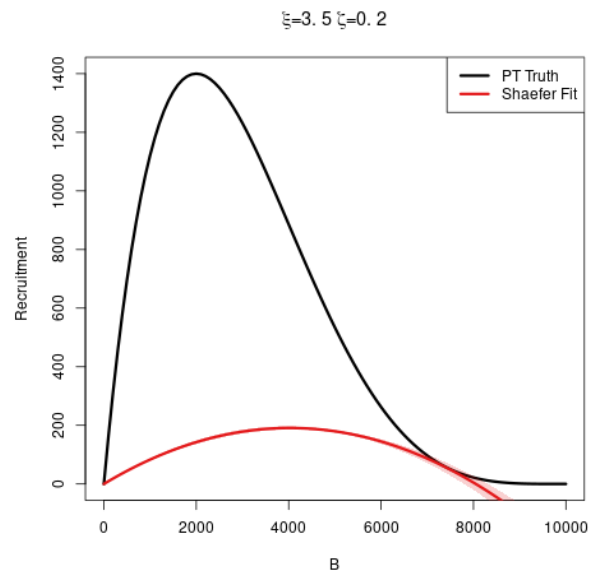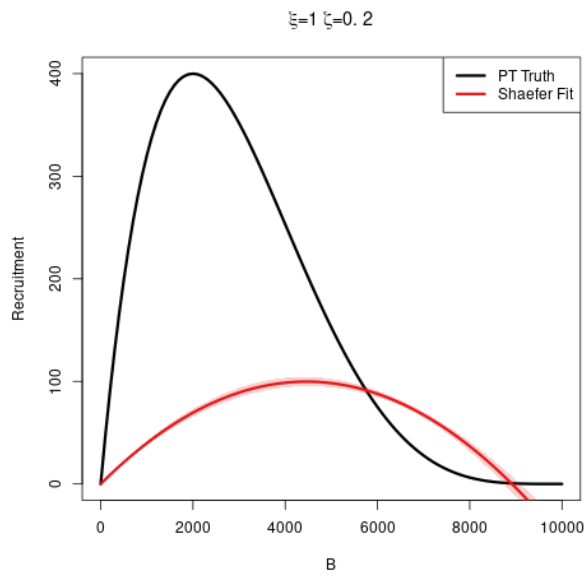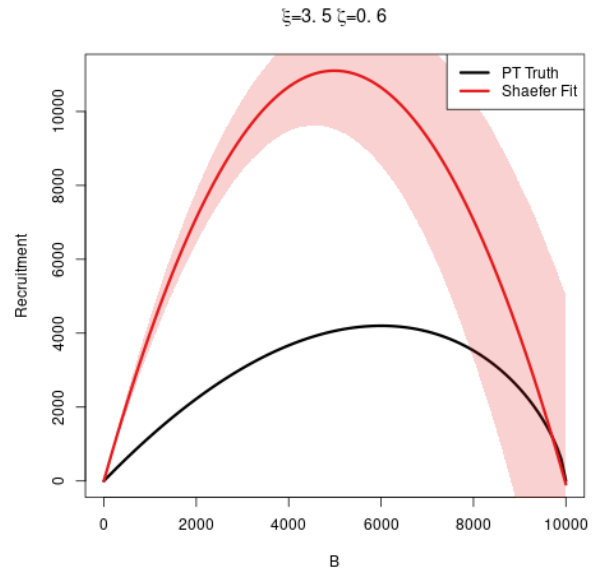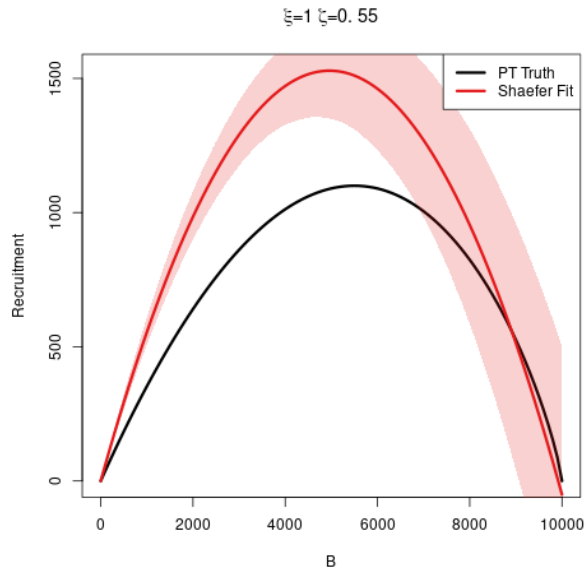
Appendix A: Distributional results for $\tilde{\xi}$ and $\tilde{\zeta}$

Given the Log-Normality of $\tilde{F}^*$ as seen in Eq. (**??**), for fixed M, $\tilde{\xi}$ is clearly just a scaled Log-Normal distribution (Log-Normal parameters are given in terms of the mean and variance on the log scale).

$$\tilde{\xi} = \frac{\tilde{F}^*}{M}$$

$$\tilde{\xi} \ \sim \ \text{LN}\left(\frac{1}{M}e^{\mu+\frac{\sigma^2}{2}}, \frac{1}{M^2}(e^{\sigma^2}-1)e^{2\mu+\sigma^2}\right)$$

Now working with $\tilde{\zeta}$ in terms given by Eq. (**??**) and considering the quantity $\text{logit}(2\tilde{\zeta})$ provides a simplification in terms of $\log\left(\tilde{F}^*\right)$.

$$\tilde{\zeta} = \frac{1}{\tilde{\xi}+2}$$

$$\text{logit}(2\tilde{\zeta}) = \log\left(\frac{\frac{2}{\tilde{\xi}+2}}{1-\frac{2}{\tilde{\xi}+2}}\right)$$

$$= \log\left(2/\tilde{\xi}\right) = \log(2) - \log\left(\tilde{\xi}\right) = \log(2M) - \log\left(\tilde{F}^*\right)$$

The given simplification of $\text{logit}(2\tilde{\zeta})$ reveals the distribution of ζ as a scaled and shifted Logit-Normal distribution.

$$\text{logit}(2\tilde{\zeta}) \ \sim \ \text{N}\left(\log(2M)-\mu, \sigma^2\right)$$

$$2\tilde{\zeta} \ \sim \ \text{logit-N}\left(\log(2M)-\mu, \sigma^2\right)$$

Notice that due to Eq. (**??**) these distribution results hold for any fixed M Beverton Holt model with Log-Normality in $\tilde{F}^*$. These results are not specific to the Laplace approximation setting here. For example under Beverton Holt and fixed M, a Log-Normal prior on $\tilde{F}^*$

Nicholas R. Grunloh

necessarily implies a scaled Log-Normal prior on $\tilde{\xi}$ and a scaled Logit-Normal prior on $\tilde{\zeta}$. Furthermore if M is not fixed, but instead it also follows a Log-Normal distribution, this also implies Log-Normality for $\tilde{\xi}$ and Logit-Normality on $\tilde{\zeta}$, albeit with slightly different parameters.

Appendix B: Profile Likelihood MLE

Given that q has the effect of rescaling the mean function, a naive handling of q has the potential to interfere with the inference on $\boldsymbol{\theta}$. While the parameter q is typically identifiable, it can introduce lesser modes which complicate naive inference.

Below I outline a profile likelihood method for MLE inference on q and σ^2. However if posed in a tactful Bayesian context, q and σ^2 may be marginalized out of the joint posterior to yield a direct sampling scheme for q and σ^2 which factors the posterior into the form $p(q, \sigma^2, \boldsymbol{\theta}|I) = N(\log(q)|\sigma^2, \boldsymbol{\theta}, I)IG(\sigma^2|\boldsymbol{\theta}, I)p(\boldsymbol{\theta}|I)$ (Cite Maria DeYorio pdf??).

The joint likelihood on the log scale can be written as,

$$\log \mathcal{L}(q, \sigma^2, \boldsymbol{\theta}; I) = -\frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_t \log\left(\frac{I_t}{qB_t(\boldsymbol{\theta})}\right)^2. \tag{22}$$

First Eq(22) is maximized with respect to q by partial differentiation of Eq(22) with respect to q,

$$\frac{\partial \log \mathcal{L}}{\partial q} = -\frac{1}{q\sigma^2}\left(\sum_t \log\left(\frac{I_t}{B_t(\boldsymbol{\theta})}\right) - T\log(q)\right) \tag{23}$$

The maximum of the likelihood in the q direction is attained when $\frac{\partial \log \mathcal{L}}{\partial q} = 0$. By setting $\frac{\partial \log \mathcal{L}}{\partial q}$ to 0 and solving for q, the MLE of q in terms of $\boldsymbol{\theta}$ can be written as

$$q(\boldsymbol{\theta}) = e^{\frac{1}{T}\sum_t \log\left(\frac{I_t}{B_t(\boldsymbol{\theta})}\right)} = \left(\prod_t \frac{I_t}{B_t(\boldsymbol{\theta})}\right)^{\frac{1}{T}}. \tag{24}$$

Notice that $\hat{q}(\boldsymbol{\theta})$ is the geometric mean of the empirical scaling factors between the observed index and modeled biomass at each time. This form is emblematic of the interpretation of the q parameter as the proportionality constant between the observed index and the modeled biomass. Additionally notice that $\hat{q}$ is a function of $\boldsymbol{\theta}$, so that achieving the global maximum of the likelihood function still requires maximization over $\boldsymbol{\theta}$. Furthermore, $\hat{q}(\boldsymbol{\theta})$ is only a function of $\boldsymbol{\theta}$ and that σ^2 does not enter the expression. This will be helpful in further maximization of the likelihood with respect to σ^2.

Now to maximize in the σ^2 direction Eq(22) is differentiated with respect to σ^2,

$$\frac{\partial \log \mathcal{L}}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2(\sigma^2)^2} \sum_t \log \left(\frac{I_t}{qB_t(\boldsymbol{\theta})} \right)^2. \tag{25}$$

The maximum of the likelihood in the σ^2 direction is attained when $\frac{\partial \log \mathcal{L}}{\partial \sigma^2} = 0$. Setting $\frac{\partial \log \mathcal{L}}{\partial \sigma^2}$ to 0 and solving for σ^2 produces the following MLE as a function of $\boldsymbol{\theta}$,

$$\sigma^2(\boldsymbol{\theta}) = \frac{1}{T} \sum_t \log \left(\frac{I_t}{q(\boldsymbol{\theta})B_t(\boldsymbol{\theta})} \right)^2 \tag{26}$$

Notice that the conditionally MLE of σ^2 is not only a function of $\boldsymbol{\theta}$ but also a function of q. As previously noted, $q(\boldsymbol{\theta})$ is only a function of $\boldsymbol{\theta}$, and so to achieve a global maximum of the joint likelihood, $\sigma^2(\boldsymbol{\theta})$ is written entirely in terms of $\boldsymbol{\theta}$ by replacing q by $q(\boldsymbol{\theta})$ as seen above.

By combining Eq(24) and Eq(26) the MLEs of q and σ^2 can be written entirely in terms of $\boldsymbol{\theta}$. Furthermore, this realization allows the joint maximization of the likelihood to be reduced to the following profile log-likelihood,

$$\log \mathcal{L}(\boldsymbol{\theta}; I) = -\frac{T}{2} \log \left(\sigma^2(\boldsymbol{\theta}) \right) - \frac{1}{2\sigma^2(\boldsymbol{\theta})} \sum_t \log \left(\frac{I_t}{q(\boldsymbol{\theta})B_t(\boldsymbol{\theta})} \right)^2. \tag{27}$$

This profile log-likelihood is maximized numerically over $\boldsymbol{\theta}$, and the estimates for q and σ^2 are given by evaluating Equations (24) and (26) at $\hat{\boldsymbol{\theta}}$.

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \log \mathcal{L}(\boldsymbol{\theta}; I) \tag{28}$$

$$\hat{\sigma}^2 = \sigma^2(\hat{\boldsymbol{\theta}}) \tag{29}$$

$$\hat{q} = q(\hat{\boldsymbol{\theta}}) \tag{30}$$

This profile formulation via $\hat{q}(\boldsymbol{\theta})$ and $\hat{\sigma}^2(\boldsymbol{\theta})$ reduces the computational complexity of this numerical optimization, while also avoiding the multimodality issues induced by q.