Metamodeling for Bias Estimation of Biological Reference Points.

Nick Grunloh

Introduction

- Hello. My name is Nick Grunloh.
- Talking about: A Metamodeling approach for assessing RP bias in two parameter models of productivity.
- Collaboration with UCSC, SWFSC, and funded: NMFS Sea Grant.

Basic Modeling Structure

- Context: Single Species Surplus-Production Models.
- Production models are an admittadly simple setting, but…
  - have plenty of secrets that we dont tend to talk about.
- Even being simple:
  - they capture many relevant dynamics for management sake
  - and are plenty instructive.
- General Structure:
  - We typically observe an index of abundance
  - Assume the index is proportional to biomass with proportionality constant q.
  - I_t forms a response variable with lognormal residuals.
- Most of the action here comes from the biomass process model.
  - Biomass is modeled as a (typically) nonlinear ODE.
  - Growth via a nonlinear production function, $P(B)$
  - Removals via natural mortality and catch.
    - Here $Z(t)$ lumps together all instantaneous removal rates.
- For management mostly interested in Biological RP Inference.
  - Should say: RPs are functions of productivity parameters in P.
- Commonly RPs are ways of noticing MSY.
  - Here I focus on two:
    - Fmsy: fishing rate to result in MSY (Relative. Fmsy/M)
    - Bmsy: population biomass at MSY (Relative. Bmsy/B0)

RP Constraints

- Conceptually $\frac{F^*}{M}$ and $\frac{B^*}{B_0}$ coexist in an entire 2D space.
  - $\frac{F^*}{M} \in \mathbb{R}^+$ $\frac{B^*}{B_0} \in (0, 1)$
- (Mangel et.al., 2013) Canadian Journal of Fisheries
  - Two parameter Production functions limit the space.
  - In particular the BH model is limited to the 1D curve $\frac{1}{x+2}$
  - **Right:** Plot Relative Bmsy against Relative Fmsy
    * black: posterior samples of the RPs for a 3 parameter Shepherd-like model. (cowcod)
    * red: posterior samples of the RPs for a 2 param BH model.
    * the red posterior is squashed into the curve $\frac{1}{x+2}$
    * Refer to this subspace of RP's as the {BH line}.
  - **Next:** Mangel et. al. suggests looking into 3-parameter curves

Schnute (1985)

- I'm working with a 3 parameter production function as developed in Schnute (1985)
  - The production function is written here.
  - Notice the three parameters α, β, and γ.
  - A number of important 2 parameter special cases
    * The Ricker, The Logistic, and ...
    * Most importantly here the Beverton-Holt when $\gamma = -1$.

- Use the 3 param Schnute model to simulate a species off the BH line
- fit those data under the BH Model.
  - Observe how BH RP estimates are biased relative to true Schnute RPs.
- **Right Panel:** On the right you can see the Schnute production function and how it uses its third γ parameter to get us off of the BH Line via dramatically different productivity behaviors.

Breadcrumb Slide

- Understanding the mapping of the greater conceptual RP space onto these constrained 2 parameter spaces is complicated even in simple cases.
  - Chaos in the Dynamical System -Time Integrator Inaccuracy
  - Model Identifiability -Global Optimization
- Production models are simplified places which are easier to hunt down the many computational issues, and are simple enough to make it possible to understand the mechanisms
- At the link provided here you can see our anlysis of the mechanisms of Bias for the Schaefer Model.

Simulation Design

- Again, we use Schnute to generate data broadly in RP-space.
- In order to do this, one would need to invert the relationship between RPs and productivity parameters.
- Schnute and Richards (1998) show that it's not analytically possible to invert this relationship and numerically inverting is unstable.
- However Schnute and Richards (1998) do provide some results which we have used to generate approximate LHS designs in RP space.
- **Next:** With a design in place Schnute data can be generated for example in the upper right.
- Fitting the BH model against those data will neccessarily land somewhere on the $1/(x+2)$ BH line. (Say the red dot)
- The aim is then to understand the behaviour of these bias arrows broadly in RP space.
  - Will one RP be prioritized over the other?
  - Is there a compromise between RPs?
  - Is the mapping a shortest distance onto the BH line?
- **Next:** Design locations used to train a GP metamodel of $\hat{RP}$.
- Particular BH fits are as helpful as their standard errors, but in repeated sampling metamodeling can discern global RP bias trends.

Catch

- Assume synthetic catch series, To complete the model specification.
- I show Catch in red and the population in black.
- Two cases: Low and High contrast
- In all cases, the intitial biomass is fixed to K.
- Low Contrast:
  - Exponential decay from K to Bmsy
  - low contrast, relatively low information setting
- High Contrast:
  - Population goes through various levels of high and lower fishing.
  - high contrast, relatively high information setting
  - wiggles about until coming to equilibrium

High Contrast

- Here we Visualize RP Biases as a bias field.
- Color indicates the percent error in MLE relative to the Truth
- Arrows indicate the Direction of Bias (from Truth to BH MLE)
- **Next:** For example data is generated off of the line (say here)
- **Next:** and then fit w/ BH. (maps to the red dot)
- We can observe a few things:
  - Overall As Model misspecification increase (far from the line), estimation bias increases.
  - In high contrast example we can see a fairly reasonable mapping.
    * Resembles a shortest distance mapping onto the BH line.
    * Neither RP dominates bias; both RPs fail togehter

Low Contrast

- In the lower information, low constrast, evironment:
  - lower information content in the data
  - we see a higher bias pattern overall
- **Next:** When Model Misspecification is large (in the upper part) we see an interesting limiting behaviour of BH
  - **Next:** the arrows in the upper part of this picture shoot off to the left.
  - Looking at the BH yeild curve fit we can see that we are drastically underestimating steepness.
  - this pushes the BH yeild curve toward the limiting flat symmetric case as relative Fmsy goes to 0.
  - at this level of model misspecification this pattern prioritzes relative Bmsy over relative Fmsy.
- **Next:** For better specified Models: **Next**
  - bias pattern returns to something like shortest distance mapping
  - in this shortest distance mapping neither RP dominates bias;
  - again both RPs compromise to fail togehter (less overall error)
- **Next:** It is interesting to note that for the lower information setting we are only allowed to misspecify model so much before we observe this "Catstrophic" inferential failure.

Conclusion

- We have a rich simulation environment describing global RP bias.
- It's a robust and easily extensible simulator that is hardened against a lot of numerical failings of ODE models.
- We have observed several Mechanisma of inference failure in the prodcution model setting.
- We will use this framework to further extend this analysis to study how patterns may be affected by dynamics of individaul growth and maturity.
- This study reminds us that RP are not observable quantities, but rather modeled quantites that are subject to Model misspecification, uncertainty, and bias.
- In particular the severly constrained setting of using a two parameter production function we are going to pay for our modeling mistakes primarily via estimation bias.
- The role of contrast in our data serves to distribute information amoung our RP estimates, but
- The information content in our data also interacts with how poorly we are misspecifying our models of RPs. So we should be very thoughtful about choosing models of population productivity as misspecified models my produce horribly biased estimate of RPs.