A Metamodel Based Clustering of Production Model Reference Point Estimation.

Nick Grunloh[a], E.J. Dick[b], and Herbie Lee[a]

[a] Department of Statistical Science, Baskin School of Engineering, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA
[b] Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, 110 McAllister Way, Santa Cruz, CA 95060, USA

Abstract

Integrated fisheries models are based upon differential equations which model stock dynamics through time. Fisheries are largely managed based upon quantities derived from the equilibrium equations of these dynamics, known as Reference Points (RP). RP behavior is primarily driven by the functional form of the productivity assumed in the differential equations. Mangel et al. (2013) demonstrate that the most commonly used models of productivity limit the domain of RPs due to a lack of flexibility induced by their two-parameter functional forms. Three-parameter models of production release this theoretical RP limitation (Punt & Cope, 2019). Nonetheless, two-parameter models of productivity are overwhelmingly used in practice. When RP model misspecification of this type is present in population dynamics models, what are the useful limits of statistical inference with respect to estimating these RPs? Here, a simulation environment is designed which explores how misspecified two-parameter production models bias RP inference. Using a Gaussian Process metamodel of the inferred RPs (under two-parameter productivity), the full theoretical space of RP bias behavior is explored. This structured simulation setting allows for clustering of RP failure modes which use the metamodel to predict when a given species is most likely to be subject to catastrophic model failure.

1. Fisheries Production Model

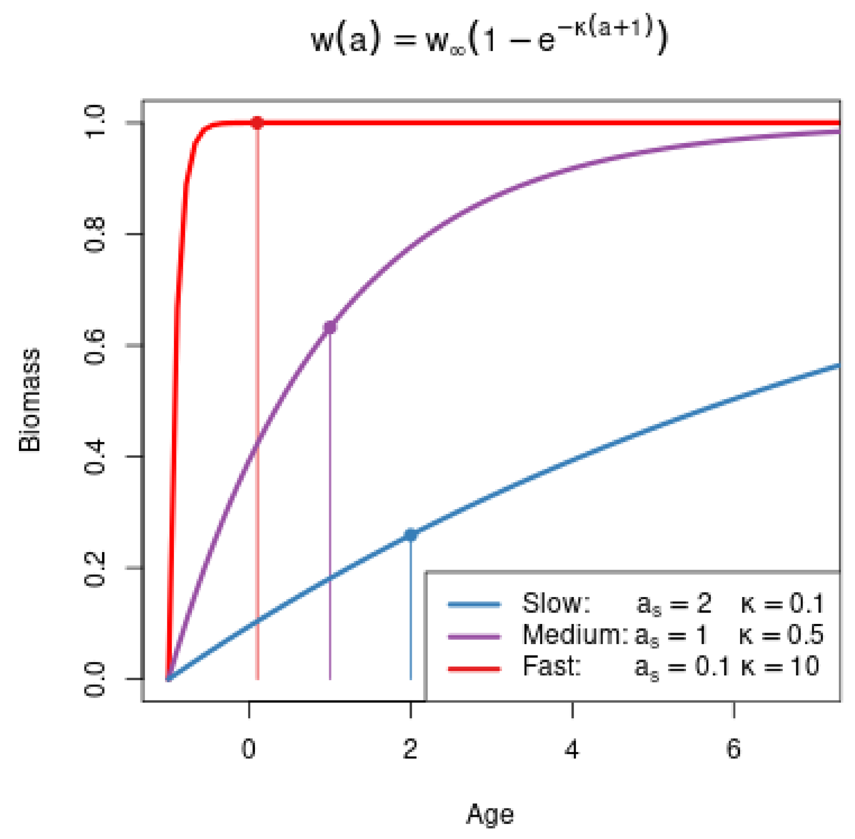
$$I_t = qB_te^{\epsilon} \quad \epsilon \sim N(0,\sigma^2)$$

Population Dynamics:

$$\frac{dB}{dt} = \overbrace{w(a_s)R(B_{t-a_s})}^{\text{Recruitment}} + \overbrace{\kappa\left[w_\infty N - B\right]}^{\text{Net Growth}} - \overbrace{(M+F)B}^{\text{Mortality}}$$

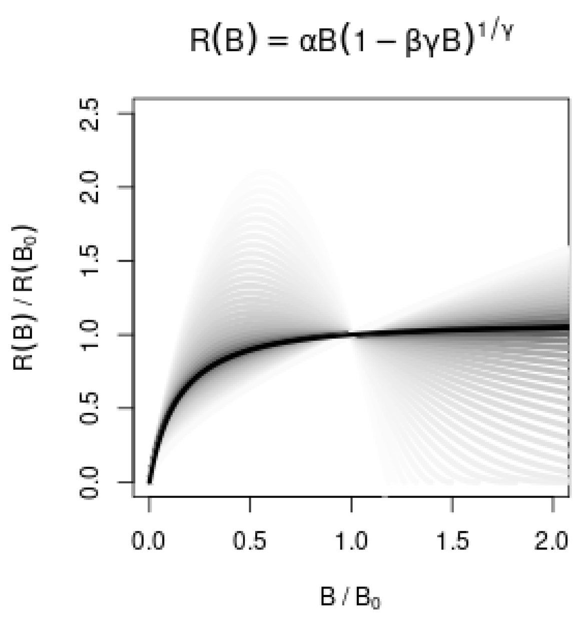
$$\frac{dN}{dt} = R(B_{t-a_s}) - (M+F)N$$

Individuals Grow:



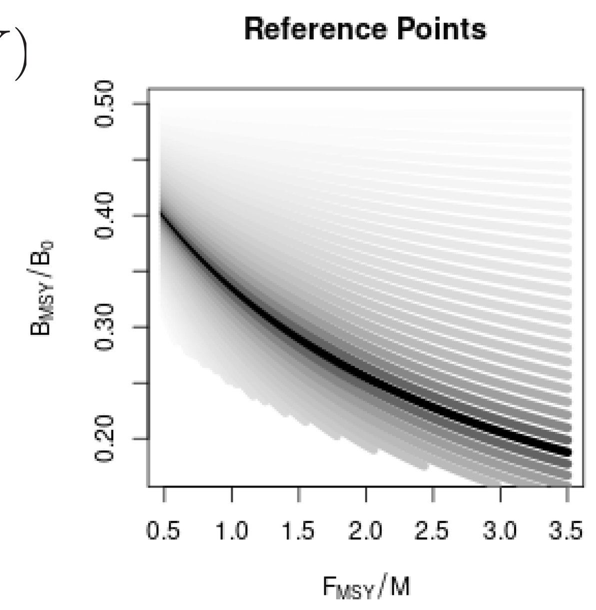
Recruitment Into the Population:

- R is notoriously nonlinear and unknown.
- Most models assume two-parameter $R_{(\alpha,\beta,-1)}$
- Fixing γ limits RPs & freeing γ expands RPs.
- $R_{(\alpha,\beta,\gamma)}$ $\gamma \in (-2,1)$ $\longrightarrow$



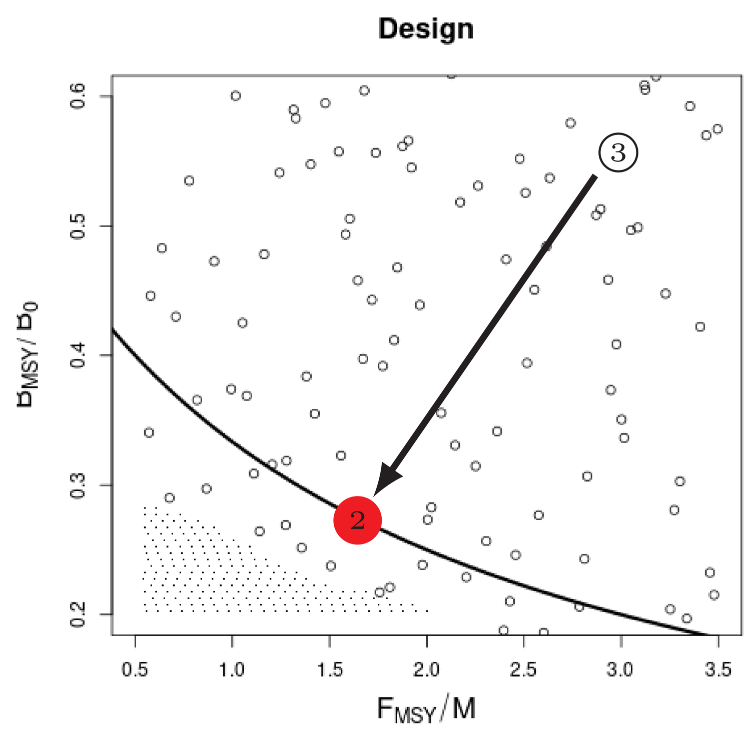
Reference Points: $\frac{F_{MSY}}{M}$, $\frac{B_{MSY}}{B_0}$

- Measures of Maximum Sustainable Yield (MSY)
- Fishing Rate and Biomass Depletion.
- Primarily driven by R and parameters α, β, γ.
- RPs are the foundation of mangement action.



2. Simulation & Metamodeling

Datasets are simulated broadly in RP space and fit with a misspecified two-parameter $R(B)$. Two-parameter $R(B)$ induces bias in RPs by limiting estimates to the curve $\frac{1}{x+2}$. Each dataset produces an example of the bias mapping.

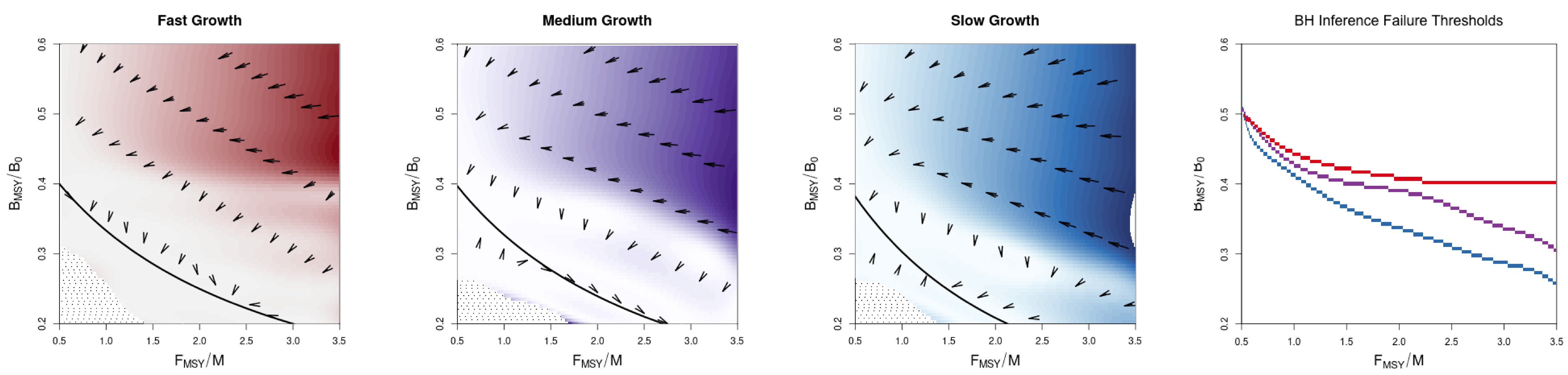


$$\underbrace{\left(\frac{F_{MSY}}{M}, \frac{B_{MSY}}{\bar{B}(0)}\right)}_{\text{Truth}} \overset{\text{GP}}{\mapsto} \underbrace{\left(\frac{\hat{F}_{MSY}}{M}, \frac{1}{\frac{\hat{F}_{MSY}}{M}+2}\right)}_{\text{2-Parameter Estimate}}$$

A Gaussian Process (GP) metamodel models the simulated mapping of RPs under the misspecified two-parameter model.

3. Clustering

Overall RPs are underestimated by the two-parameter model. When two-parameter model misspecification becomes "large" RPs are catastrophically underestimated. This breakpoint is a function of growth.



Since RPs are generally underestimated, a one-sided test is used to idetify model failure. Hypotheses may be refined by specifying P. Catastrophic model failure is identified by taking P to be 0.5.

The GP metamodel predicts the inferred RP mean $\hat{y}(\mathbf{x})$ & variance $\hat{\sigma}^2(\mathbf{x})$ as a function the true RPs,

$$\log(\hat{F}_{MSY}) \sim N\Big(\hat{y}(\mathbf{x}), \hat{\sigma}^2(\mathbf{x})\Big).$$

We want a small percent error in RP estimation:

$$\frac{\frac{F_{MSY}}{M} - \frac{\hat{F}_{MSY}}{M}}{\frac{F_{MSY}}{M}} \leq P$$

$$\hat{F}_{MSY} \geq (1-P)F_{MSY}$$

Declare model failure when:

$$e^{\hat{y}(\mathbf{x})+\sqrt{2\hat{\sigma}^2(\mathbf{x})}\Phi^{-1}\left(\frac{1}{10}-1\right)} < (1-P)F_{MSY}$$

References

Mangel, M., MacCall, A. D., Brodziak, J., Dick, E., Forrest, R. E., Pourzand, R., & Ralston, S. (2013, April). A perspective on steepness, reference points, and stock assessment. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(6), 930–940.

Punt, A. E., & Cope, J. M. (2019, September). Extending integrated stock assessment models to use non-depensatory three-parameter stock-recruitment relationships. *Fisheries Research*, *217*, 46–57.