A Metamodel Based Clustering of Fisheries 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 Model

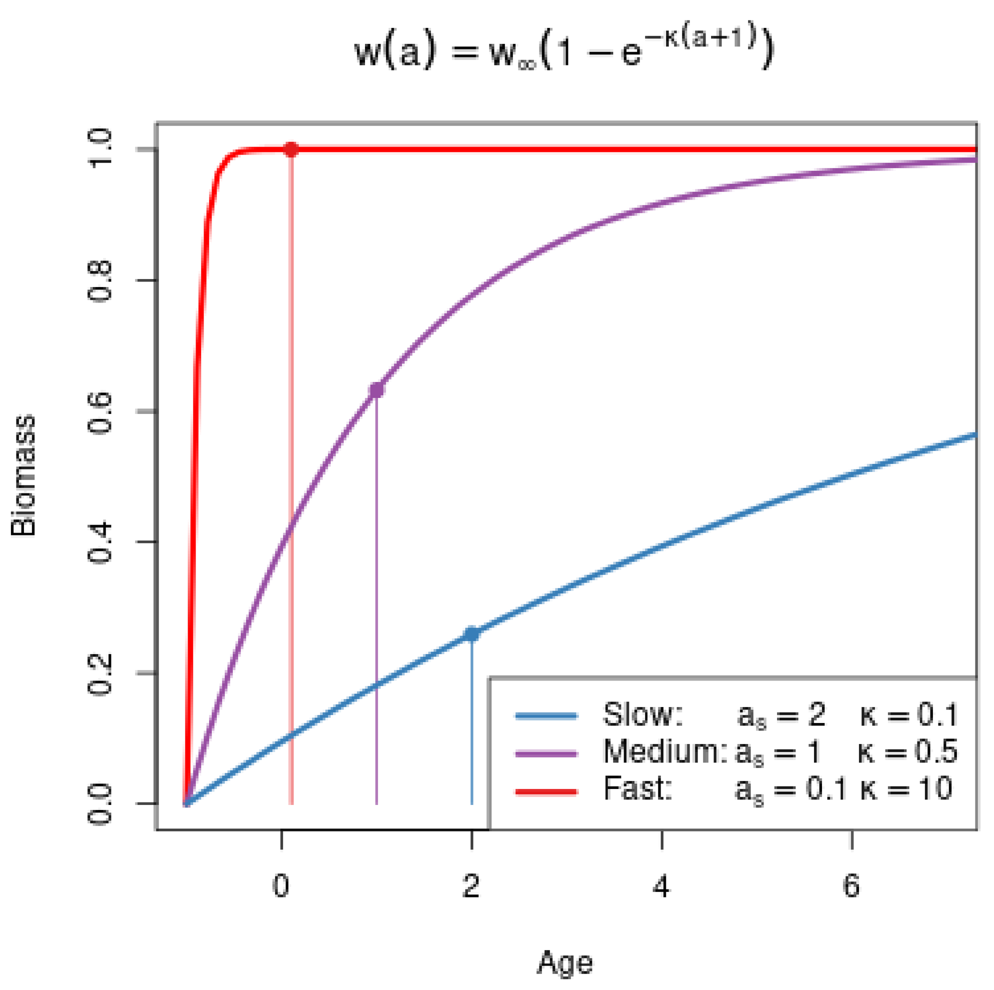
$$I_t = qB_te^{\epsilon} \quad \epsilon \sim N(0, \sigma^2)$$

Population Dynamics:

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

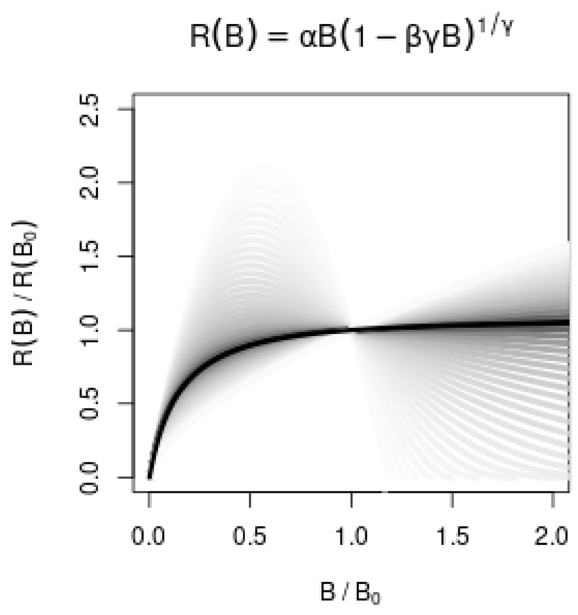
$$\frac{dN}{dt} = R(B;\theta) - (M+F)N$$

Individuals Grow:



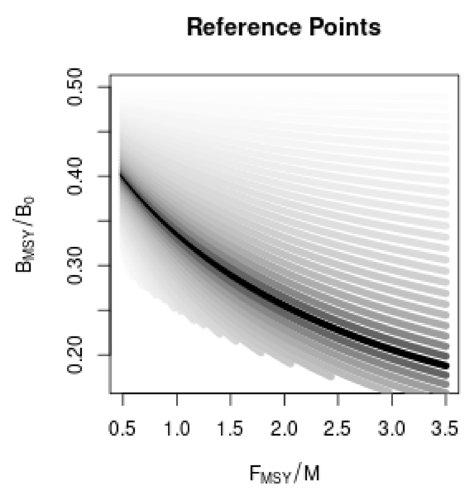
Recruitment Into the Population:

- R is notoriously nonlinear and unknown.
- Three parameter model shown left.
- Most models assume a two parameter form.
- $\gamma = -1$ (black curve) is most common.
- Freeing γ expands RPs.



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

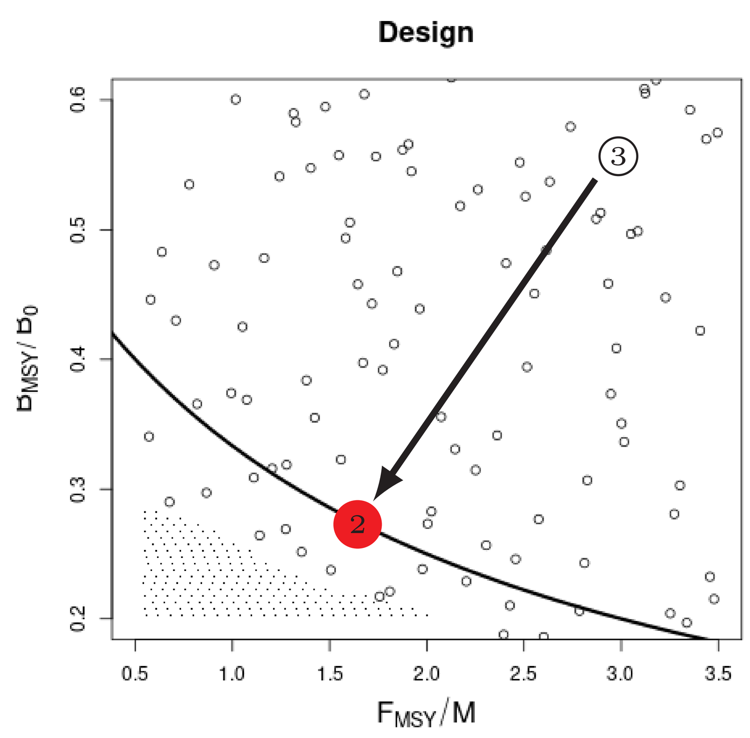
- Equilibrium quantities from dynamics.
- Primarily driven by R.
- Inferred by learning θ as possible.
- RPs are the foundation of mangement action.



2. Simulation & Metamodeling

Datasets are simulated broadly in RP space and fit with misspecified two-parameter $R(B)$. Each dataset produces an example 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. The GP metamodel then predicts the inferred RP mean $\hat{y}(\mathbf{x})$ & variance $\hat{\sigma}^2(\mathbf{x})$ as a function the true RPs,

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

3. Clustering

We want a small percent error in RP estimation:

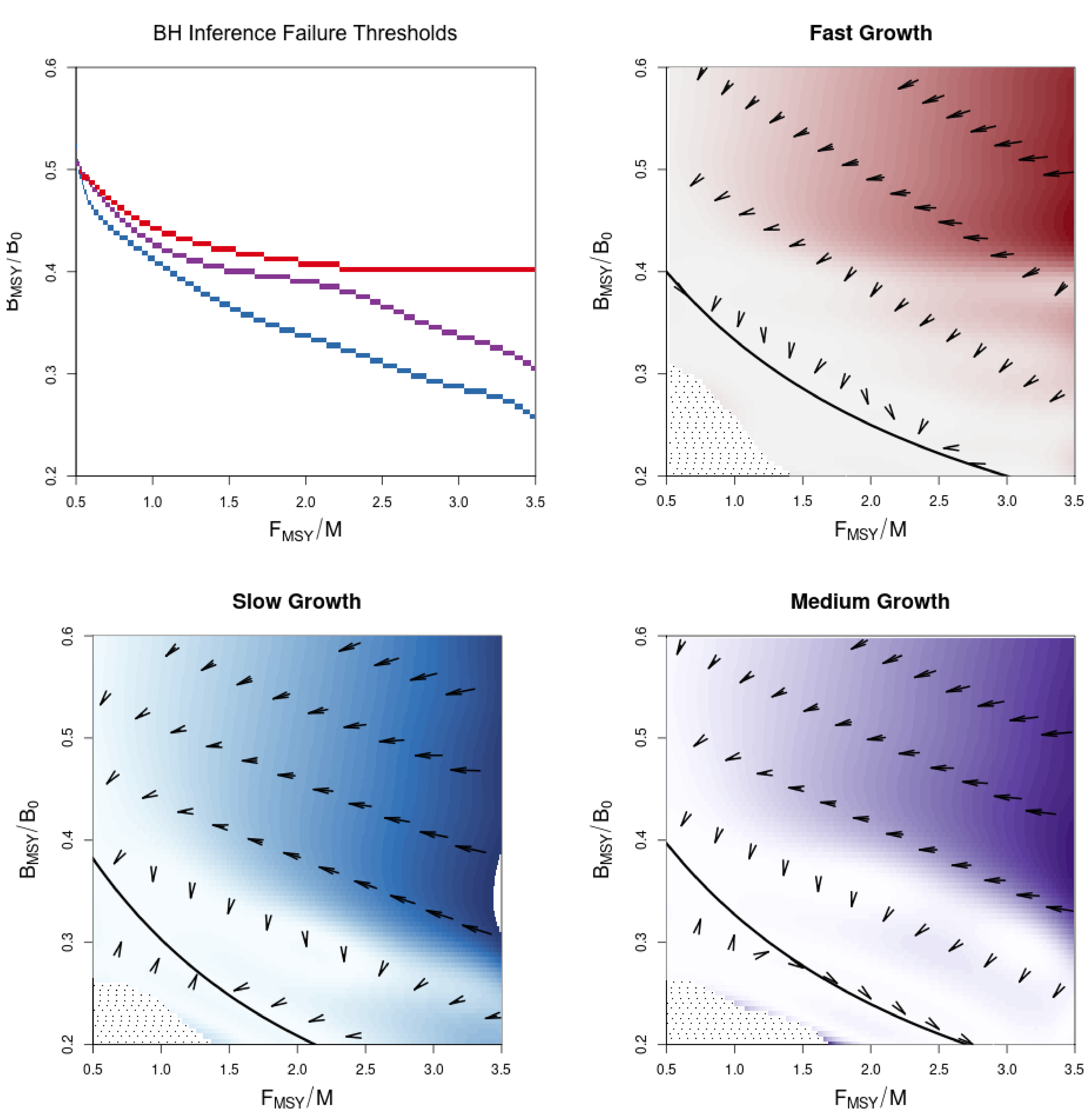
$$\frac{\frac{F_{MSY}}{M} - \frac{\hat{F}_{MSY}}{M}}{\frac{F_{MSY}}{M}} \leq P$$

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

Declare model failure when:

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

Catastrophic model failure is identified by taking P to be 0.5. Hypotheses may be refined by reducing P as relevant.



4. Conclusions

- A rich simulation-based method for describing global RP bias.
- GP metamodel, describes and identifies breakpoints well in this setting.
- Two-parameter model largely under estimates $\frac{F_{MSY}}{M}$.
- As model misspecification becomes "large", RP estimation breaks catastrophically.
- Slower growth exaggerates model misspecification.