

# Improving Catch Estimation Methods in Sparsely Sampled, Mixed Stock Fisheries.

Nicholas Grunloh, E.J. Dick, Don Pearson, John Field, Marc Mangel

2/22/2016

## Abstract

In order to effectively manage exploited populations, accurate estimates of commercial fisheries catches are necessary to inform monitoring and assessment efforts. In California, the high degree of heterogeneity in the species composition of many groundfish fisheries, particularly those targeting rockfish (*Sebastes*), leads to challenges in sampling all market categories, or species, adequately. Limited resources and increasingly complex stratification of the sampling system inevitably leads to gaps in sample data. In the presence of sampling gaps, ad-hoc point estimation is currently obtained according to historically derived “data borrowing” protocols which do not allow for tractable uncertainty estimation. In order to move from the current, but admittedly rigid sampling design, we have continued previous exploratory efforts to develop, and apply, Bayesian hierarchical models of the landing data to estimate species compositions. Furthermore, we introduce a formalized method for discovering consistent “borrowing” strategies across overstratified data. Our results indicate that this approach is likely to be more robust than the current system, particularly in the face of sparse sampling. Additionally, our method should also help inform, and prioritize, future sampling efforts. Perhaps more significantly, this approach provides estimates of uncertainty around species-specific catch estimates.

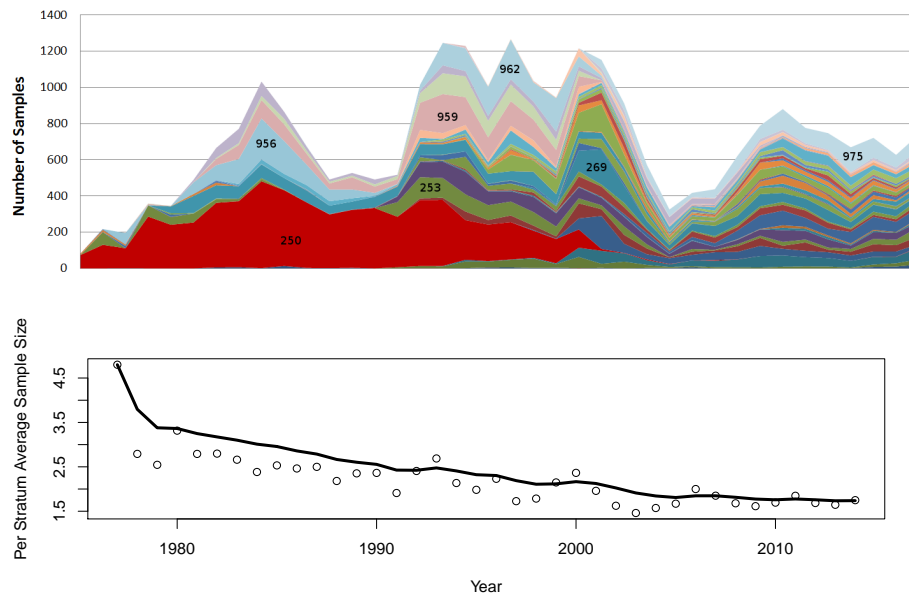
## The Goodness

In the realm of statistical modeling, biologists face a conflated set of expectations. Of primary theoretical concern the role of the biologist it to come to an understanding of biological truth, and rightfully so they have turned to statistical methods to discern truth in the face of uncertainty. However, the modern applied biologist must not only face the challenge of developing a model for understanding the truth structure of a given biological system, but must also be willing to be pragmatic practitioners of statistical methods to produce statistical

models with the goal of producing accurate prediction. Both roles are important, and often these goals may align, however it is important to understand that mathematically these goals are not the same. It is certainly possible that the most effective first step for producing a pragmatic predictor is to model a system with the admission that possible theoretical knowledge may not be well represented in a particular dataset. Approaching problems with this pragmatic perspective allows for the design of modeling approaches which better model observed uncertainty and thus result in models which are capable of making better predictions than an otherwise biologically theoretically sound model. Differences between two such models represent sub-ideal sampling conditions or experimental design, and still the pragmatic applied biologist is often confronted with making decisions in the light of such data.

## Introduction

- Data
- Stratification Pictures



## Methods

Given the complexity of these data, and thus the complexity of the modeling challenge that they present, it is necessary to define a few notational conventions

for representing such a model. In an attempt to conserve symbols, and hopefully reveal structure,  $(\cdot)$  is used as a superscript on variable, or index, symbols to indicate a new symbol of the sort. For example,  $a^{(1)}$  represents the first subclass of a class of  $a$  variables,  $j^{(1)}$  represents a the first of a class of  $j$  indecies, and thus  $a_{j^{(1)}}^{(1)}$  represents the  $j^{(1)th}$  variable of the first subclass of  $a$  variables. Additionally,  $\{\cdot\}_\Omega$  is used as notation for condensing repeated structures within the overall model, where  $\Omega$  represents indexing directions over some set. For example:

$$\left\{ x_{j^{(k)}}^{(k)} \sim N\left(0, v^{(k)}\right) \right\}_{k \in \{1, \dots, m\}} = x_{j^{(1)}}^{(1)} \sim N\left(0, v^{(1)}\right), \dots, x_{j^{(m)}}^{(m)} \sim N\left(0, v^{(m)}\right)$$

Introduce model. Check hyperprior values. maybe use c notation?

---

Species	$j^{(1)} \in \{1, \dots, J^{(1)}\}$
Gear	$j^{(2)} \in \{1, \dots, J^{(2)}\}$
Port	$j^{(3)} \in \{1, \dots, J^{(3)}\}$
Quarter	$j^{(4)} \in \{1, \dots, J^{(4)}\}$
Year	$j^{(5)} \in \{1, \dots, J^{(5)}\}$

---

Ole Poisson paper(cite). Build case for Overdispersion(cite). As  $\psi \rightarrow \infty$ ,  $NB \rightarrow Poisson$

$$y_{ij^{(1)}j^{(2)}j^{(3)}j^{(4)}j^{(5)}} \sim NB\left(\exp(\boldsymbol{\theta}), \exp(\psi)\right)$$

$$\boldsymbol{\theta} = \beta_0 + \sum_{k=1}^5 a_{j^{(k)}}^{(k)} + \sum_{\substack{k \in \{3,4\} \\ l \in \{1,2,3,4,5\} \\ l \neq k}} b_{j^{(k)}j^{(l)}}^{(k,l)}$$

Talk about Model, linear predictor (unpack it piece by piece),

$$\psi \sim N\left(0, 10^4\right)$$

$$\beta_0 \sim N\left(0, 10^4\right)$$

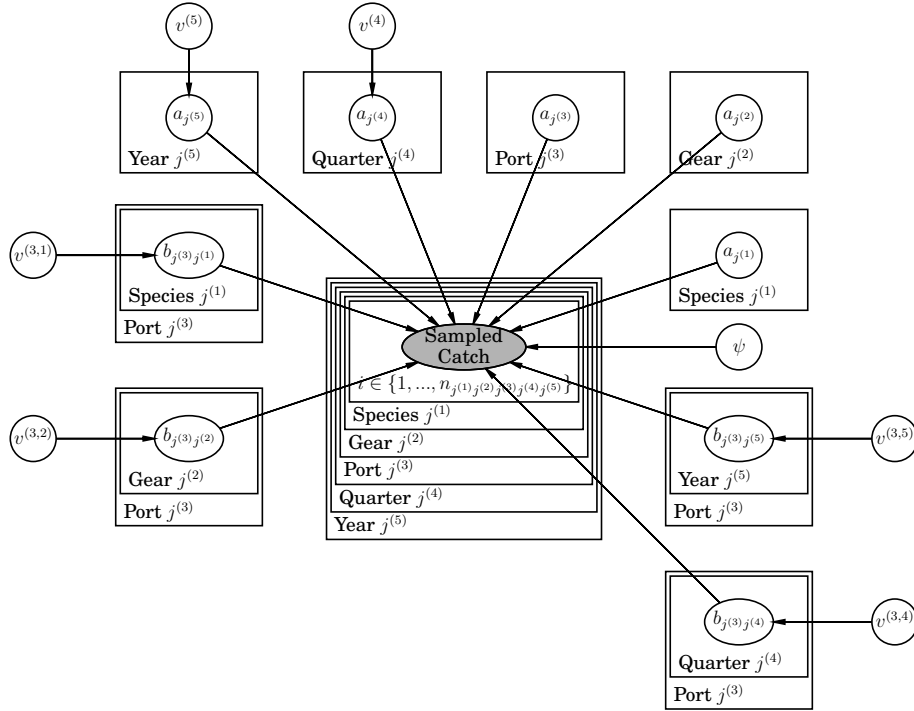
$$\left\{ a_{j^{(k)}}^{(k)} \sim N\left(0, 10^4\right) \right\}_{k=\{1,2,3\}}$$

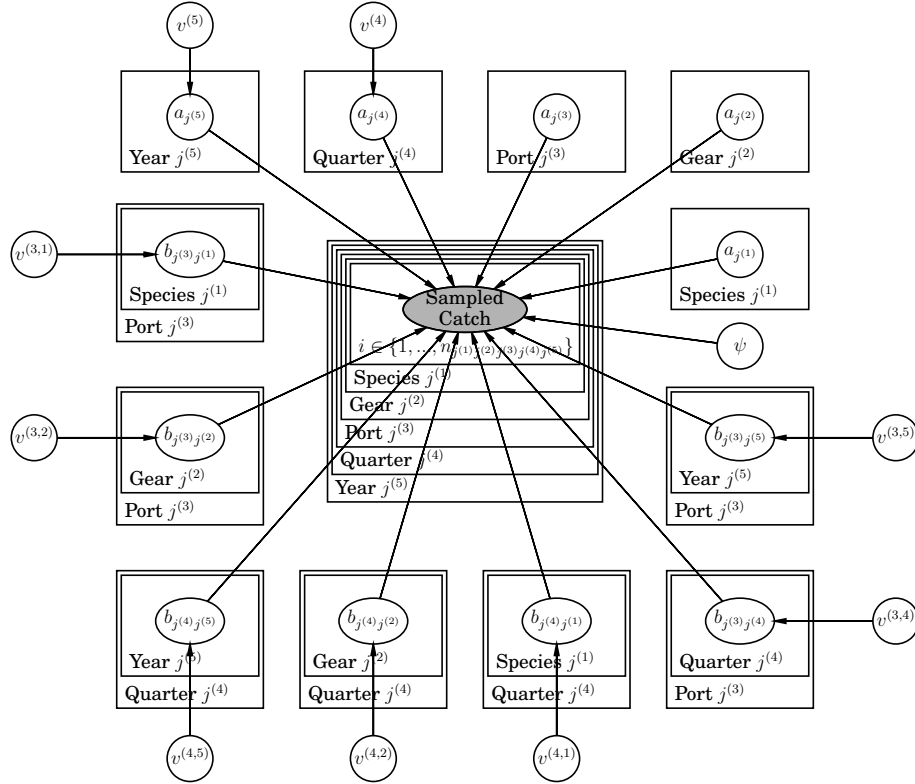
$$\left\{ a_{j^{(k)}}^{(k)} \sim N\left(0, v^{(k)}\right) \right\}_{k=\{4,5\}}$$

$$\left\{ b_{j^{(k)}j^{(l)}}^{(k,l)} \sim N\left(0, v^{(k,l)}\right) \right\}_{\substack{k \in \{3,4\} \\ l \in \{1,2,3,4,5\} \\ l \neq k}}$$

$$v \sim IG(1, 10^5) \quad \forall \quad v$$

Sampled Catch. Better graphical model???





Explain process for arriving at this structure. Rational for including interactions (also latent heirarchy). How exactly does this “share” (which strata, in what sense).

- Model (likelyhood/prior)
  - Overdispersion (Poisson Ole paper)
  - notaion explaintion
  - model
  - How did we arrive at this model (prior likelihood)
  - Heirarchy (sharing)
    - \*  $??v^{-1/2} \sim Unif(0, Big) \quad \forall \quad v??$
    - \*  $??v^{-1/2} \sim Half - Cauchy \quad \forall \quad v??$
  - Graphical Model
  - INLA
- Predictive

- species composition
  - species landing expansion
- Model Selection (Borrowing)
  - Port/??Qtr??
    - \* ??Combinatorics??Abstract
  - WAIC/MSE
    - \* predictive performance
    - \*  $MSE(\hat{\theta}) = Var(\hat{\theta})^2 + Bias(\hat{\theta}, \theta)^2$

## Results

- Bar plot picture
- Predictive Performance
- Sample/sample generating structure availability ??abstract??

## Discussion

- The Good
- The Bad
- The Ugly
- Moving Forward