Improving Catch Estimation Methods in Sparsely Sampled Mixed-Stock Fisheries.

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

Introduction

- I am Nick
- describe the california spp comps. port sampling data for modeling
- describe our modeling efforts for estimating ssp comps.

Request: Diagnostic

- Create fully stratified performace diagnostics based on my tabulated (tables 2 and 3) aggragate performance numbers.
- Quickly refresh on the model in question
- Tools for sorting through all of this information
- Diagnostics for evaluating performance

Beta-Binomial Model

- $y_{ijklm\eta}$: i^{th} sample of the j^{th} species' integer weight, in the k^{th} port, caught with the l^{th} gear, in the η^{th} quarter, of year m, for a particular market category.
- Stratum μ linked to θ and observed cluster size (n)
- Stratum σ^2 is largely a function of μ but with overdispersion ρ
  - $\rho \to 0$: Binomial variance
  - $\rho \to 1$: n times Binomial variance
- Modeling of θ (all predictors are categorical):
  - Intercept
  - Additive offsets for: Species, Port, Gear
  - Consider multiple time models

Diagnostic Files

- Consider a toy example to get our hands dirty with the diagnostic.
- Recall the model (M4).
- Recall we had some model selection criterion. Here I show the DIC and WAIC information criterion not as a diagnostic, but merely to guide our search through the numerous models under consideration.
- Through out this document you will see green underlined tags.
  - Depending on your pdf viewer, these lines may look slightly different and you may get slightly different behavior when clicking.
  - In any case these should be clickable links to various github pages.
  - If you are veiwing this in a browser you may prefer to [ctrl]-click to avoid redirecting away from the presentation tab.
- **click** We'll talk more about them later
  - marginal species directories (BCAC pdf and csv)
  - directories of various levels of stratification (pdfs) species-gear-year
  - stratifcation csvs gearYearSpp68.csv

MAD Diagnostic

- As we add more models there is a lot of information to sort through, consider the MAD diagnostic as a tool for sorting.
  - ℓ_i: the landings in stratum i,
  - $\mathcal{O}_{ij}$: the observed predictive accuracy of species j in stratum i
  - $\aleph$: the nominal level of prediction for a particular model run
- Low MAD scores occur when ℓ_i is low -or- $|\mathcal{O}_{ij} - \aleph|$ is small.
- High MAD scores occur when ℓ_i is large and $|\mathcal{O}_{ij} - \aleph|$ is large.
- ???? example ????
  - High MAD v. Low MAD

Stratum Plots

- Prediction shown at three levels of stratification
  - Disaggregated
  - By species, gear group, and year aggregating across port complexes, and quarters. ("data-rich assessment")
  - By species, and year aggregating across port complexes, gears, and quarters. ("data-moderate/poor assessment")
  - csv versions of these files are in base run directory.
- **click each**

Diagnostic Wrap-up

- Marginal plots organized by species, each marginal stratum summed over everything else.
- Sort species by MAD, explore margins via margin plots
- Explore within margins via previously described stratum plots

Request: Sample Size

- a request for Sample sizes by mcat and time block
- through out the rest of the requests we work with the top 3 landed mcats in 1978-1982
  - 250, 253, 269
- tables show number of port sampling sightings
- other mcats and higher stratifications are provided as supplemental excel files.

MCAT 250 Sample Sizes

- observed species all time in mcat 250
  - note these are not multinomial sample sizes, but rather sighting occurances.
  - multinomial structure fills in zeros for all unsighted species in a particular sample id.
- We'll see that model performance will get some of the common species, while the less common species are very hard to predict.
  - Common: BCAC, CLPR, CNRY, WDOW, YTRK
  - Intermediate: BANK, BLGL, CWCD *often worrisome
  - Uncommon: BRNZ, MXRF

MCAT 253 & 269 Sample Sizes

- MCAT 253:
  - Common: BCAC, CLPR, WDOW
  - Intermediate: SNOS, YTRK
  - Uncommon: BLGL, CWCD
- MCAT 269:
  - Common: WDOW
  - Intermediate: CLPR, YTRK // BCAC, CNRY
  - Uncommon: DBRK, POP

Flatfish and Elasmobranchs

- number of port sampling sightings
- Largest landed Flatfish and Elasmobranchs
  - Sampling Flatfish since 2002
  - Sampling Elasmobranchs since 2009
- **See Flat/Elasmobranch Table**

Request: Redo modeling w/o So-Cal

- Redo modeling in early time block MCAT 250 w/o Southern California
- Here we look at predictions from top species:
  - CLPR, CNRY, WDOW, YTRK

Redo SoCal Summary

- Out of sample predictions do not effect observed strata.
- Small difference just come from slight run-by run variation
- When Sample sizes become very sparce it can cause slight model instability.

Request: Time Model & Prior Sensitivity

- Top landings MCATS in early time period: 250, 253, 269
  - M models
  - Prior models

Time Models

- Bayesian Modeling
  - Heirarchical v. Random Effect Disclaimer
- (M1) Fixed main effect time model
  - No pooling
- (M2) Random main effect time model
  - years/quarter pool separately
- (M3) Random main effects + random interaction
- (M4) Random interactions jointly pooled
- (M5) Random interactions quarterly variances pooling across years
- (M6) Random interactions yearly variances pooling across quarters
- All with default IG prior

TIME MODEL: 250

- M2, M3, M4
- Least MAD worrisome: WDOW, BCAC, CLPR, CNRY
  - BCAC:
    - Most of landings in TWL, in later years, in all qtrs,
    - largest landings in BRG
    - generally good performance
  - WDOW:
    - Most landings in ERK, TWL, 1986, all qtrs
    - very good performance
- Most MAD worrisome: BRNZ, MXRF, BLGL, CWCD, BANK * consistent
  - CWCD:
    - MRO, HKL (some TWL), all years, spring
    - Over fitting: mostly 0s and interval contains 0
  - MXRF:
    - BRG, TWL, 1980, Winter/Spring
    - Very small sample sizes
      - wouldn't be surprised if some of difference are due to model instability
      - large variance: even a little instability could cause some what larger predictive differences.
      - flat likelihood in MXRK axis => small Δ likelihood across wide area.

TIME MODEL: 253

- M4, M5, M6
- Least MAD worrisome: WDOW, BCAC, CLPR, CNRY, BANK * consistent
  - BCAC:
    - TWL
    - predictions are relatively good
    - not a huge difference between models
  - WDOW:
    - TWL
    - slight overfitting, not bad fit
- Most MAD worrisome: DBRK, CWCD, YTRK, BLGL, SNOS * consistent
  - CWCD:
    - OSF/MRO, TWL, 1978/1981, winter/summer
    - Overfitting: Simpler models do better
      - Mostly zeros
  - DBRK:
    - OSF/MNT, TWL, 1980, summer
    - Underfitting: More complex model works better

TIME MODEL: 269

- M4, M5, M6
- Few species, some spp show on best and worst
- Least MAD worrisome: YTRK, BCAC, CLPR, CNRY
  - CNRY:
    * CRS, TWL, 1982, Q3
    * Overfitting: simpler model
  - YTRK:
    * CRS/MRO, TWL, 1982, spring
    * almost no model sensativity
- Most MAD worrisome: WDOW, DBRK, POP
  - BCAC:
    * ERK/MNT, TWL, 1982, Fall/Summer
    * Overfitting
    * almost no model sensativity
  - WDOW:
    * BDG, TWL, 1982, spring
    * underfitting:??

Request: Landings

- Aggregate across MCATs 250, 253, 269 by year and year:gear for each spp.
  - new model runs against calcom in black
- Only show select species realevant for management
- **WDOW**
  - little sensativity to M model
  - low estimates in TWL, 1979, 1980, 1981
- **BCAC**
  - little sensativeity
  - driven by TWL
  - reasonably small differences
- **CLPR**
  - little sensativity
  - driven by TWL (very similar)
  - other gears off a bit (S:G)
- **DBRK & CWCD**
  - little sensativity
  - lots of variance but basically similar
  - we can estimate variance! wouldn't be able to say that from calcom
- **MXRK**
  - suuuper skewed distributions
    - 9000 samples at 0 and some up to 1
    - high variance
  - Bayesian inference estimating higher moments
    - complex posterior as a result of extreme lack of information
    - statistics are breaking down, yet mean not far from calcom.
    - all of the instability is masked in calcom but the model can seee it.
      - we should want to see those failures in the data; the model does.

Priors

- Diffuse priors
- Main effects diffuse Normals
- ρ transformed to be a real number
  - $\text{logit}(\rho) \to (-3.91, 3.91)$
  - $\rho \to (0.02, 0.98)$
- Any heirarchical variances:
  - Default: IG prior
  - Informative: $\sqrt{v} \sim \text{Half-Cauchy}(10^1)$
  - Diffuse: $\sqrt{v} \sim \text{Half-Cauchy}(10^3)$
  - Flat: $\sqrt{v} \sim \text{Unif}(0, 10^4)$