A hierarchical approach to joint estimation of juvenile salmonid abundance and detection efficiency

Rethinking methods for quantifying snorkel survey detection efficiency

March 5, 2020
Oregon Chapter AFS Meeting

B. Staton[1], C. Justice[1], S. White[1], T. Sedell[2], L. Burns[1], M. Kaylor[3]

[1]Fisheries Science, Columbia River Inter-Tribal Fish Commission
[2]Oregon Department of Fish and Wildlife
[3]Department of Fish and Wildlife, Oregon State University

Snorkel Surveys

What are they?

- Basically what they sound like
- Intended to count juvenile salmonids
- Summer base-flow counts
- Data: counts of fish
  - By species, size class

Snorkel Surveys

What are they?

- Basically what they sound like
- Intended to count juvenile salmonids
- Summer base-flow counts
- Data: counts of fish
  - By species, size class
- Counts $\xrightarrow[AREA]{}$ Densities
- Fish vs. _____ relationships
  - Habitat type/quality/restoration
  - Den. dep. growth/mortality
  - Weighted usable habitat metrics
- Substantial observation errors

Snorkel Surveys

What are they?

- Basically what they sound like
- Intended to count juvenile salmonids
- Summer base-flow counts
- Data: counts of fish
  - By species, size class
- Counts $\xrightarrow[AREA]{}$ Densities
- Fish vs. _____ relationships
  - Habitat type/quality/restoration
  - Den. dep. growth/mortality
  - Weighted usable habitat metrics
- **<u>Substantial observation errors</u>**

Snorkel Surveys

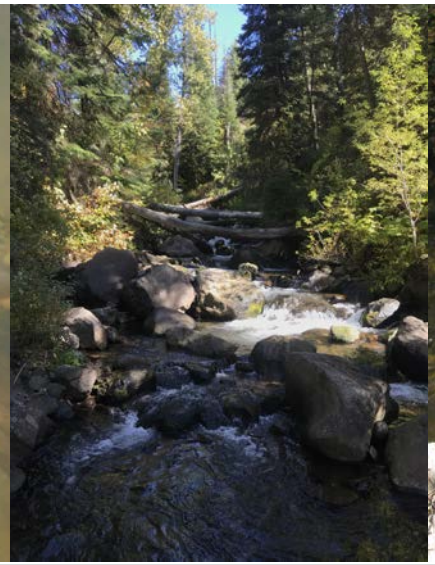
Types of observation errors

(1) Partial Detectability (*individual fish may not be seen*)

(2) Misclassification

- Double, triple, etc. counting
- Incorrect species assignment
- Incorrect size class assignment

(3) Transcription (*counts may be incorrectly recorded*)



Snorkel Surveys

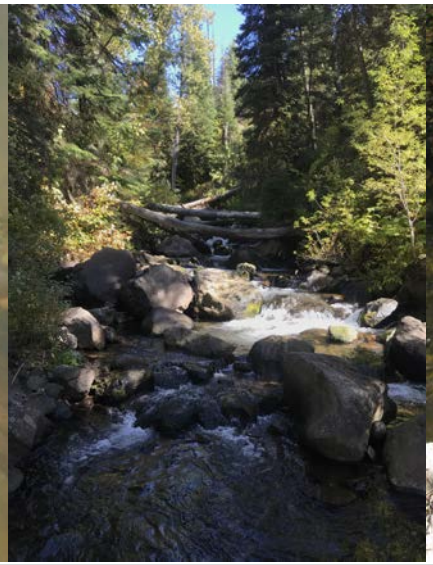
Types of observation errors

(1) <u>Partial Detectability</u> (*individual fish may not be seen*)

(2) Misclassification

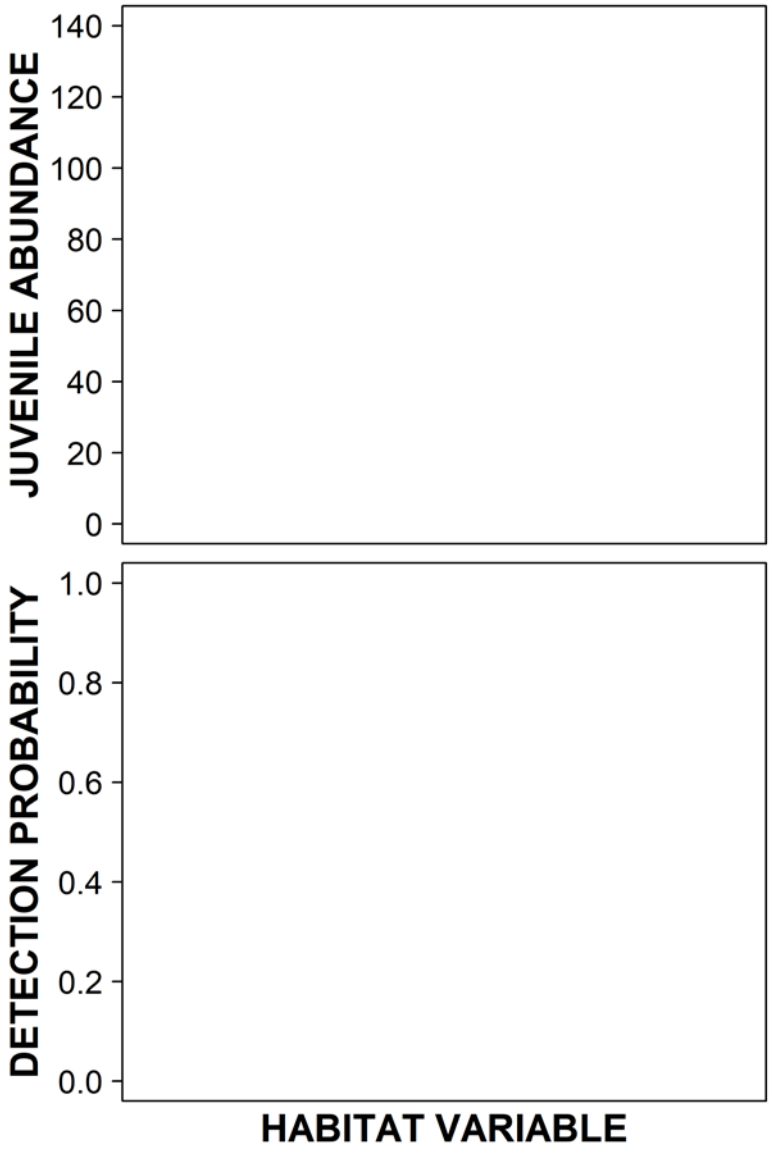
- Double, triple, etc. counting
- Incorrect species assignment
- Incorrect size class assignment

(3) Transcription (*counts may be incorrectly recorded*)



Partial Detectability in Snorkel Surveys

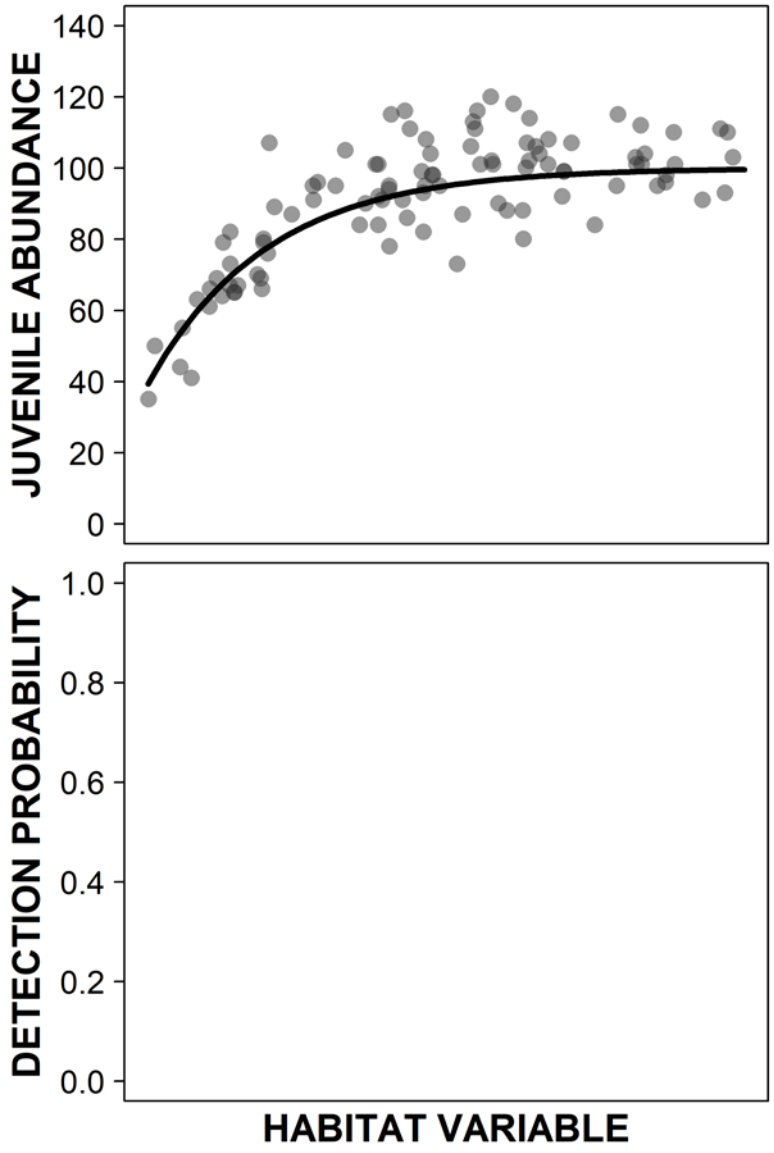
Why does it matter?



Counts are a **biased representation** of abundance

Partial Detectability in Snorkel Surveys

Why does it matter?



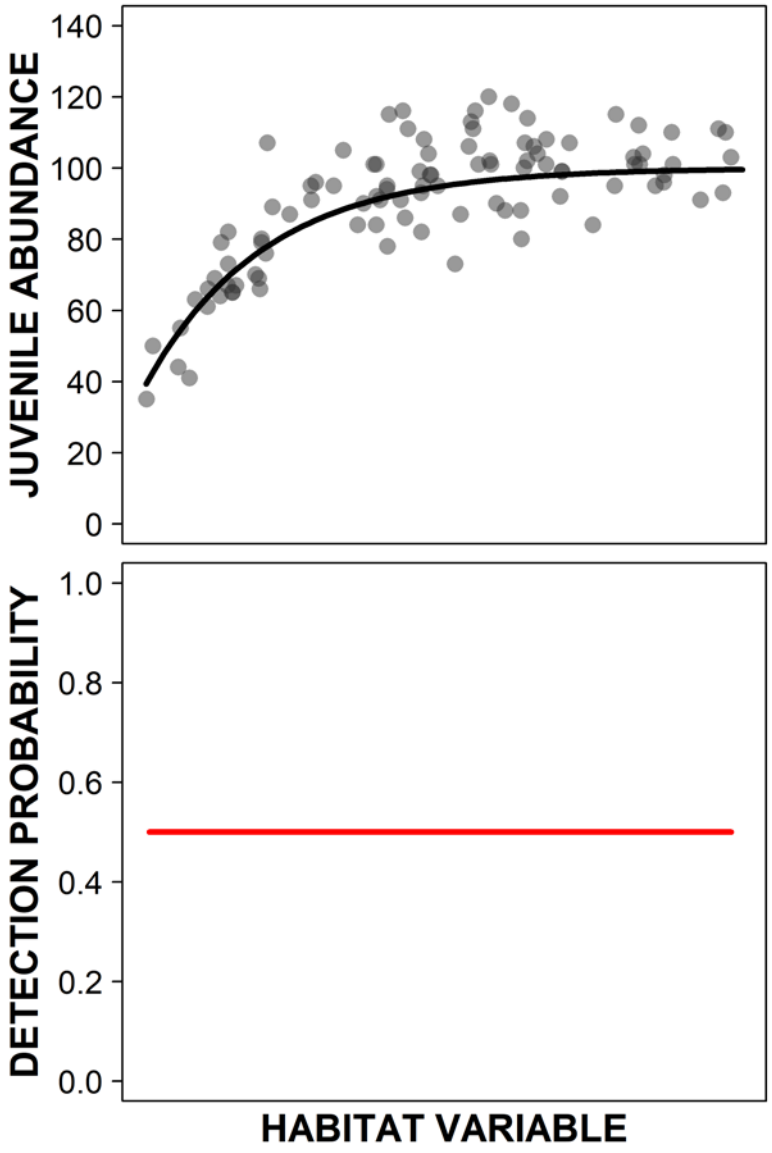
Counts are a **biased representation** of abundance

True Relationship

True

Partial Detectability in Snorkel Surveys

Why does it matter?



Counts are a **biased representation** of abundance

True Relationship

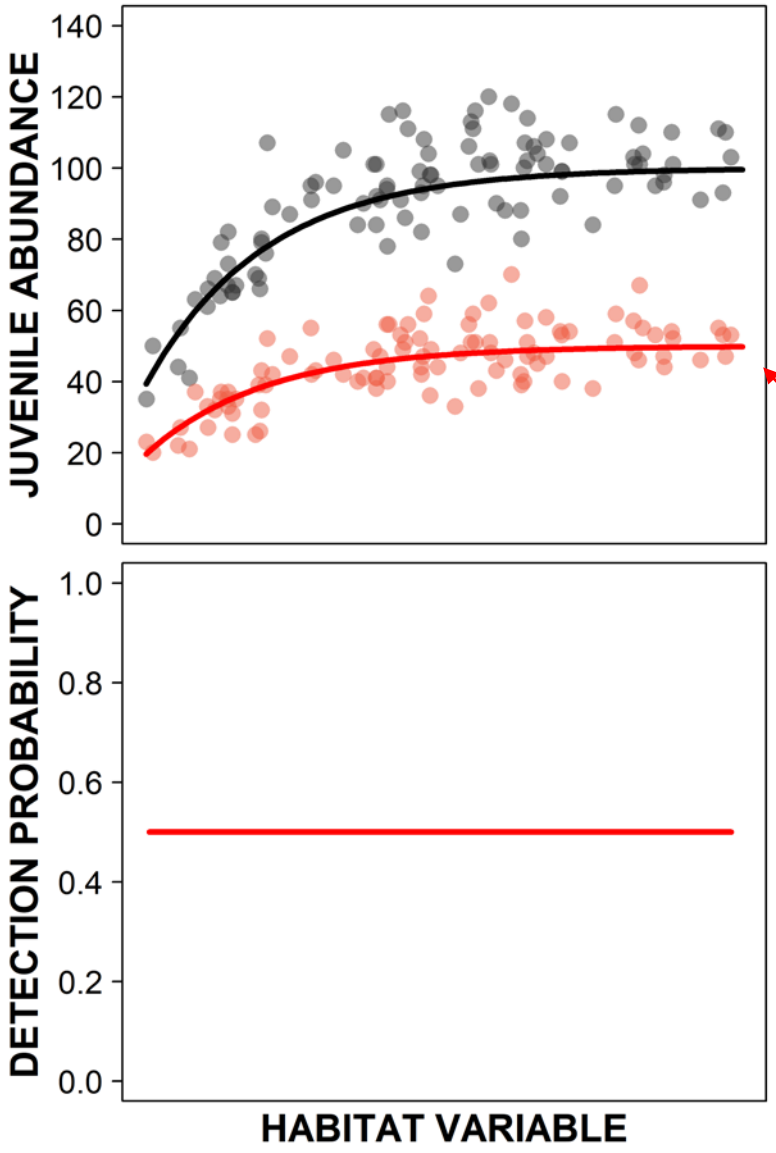
Count half of fish present, regardless of habitat (on average)

True

Observed

Partial Detectability in Snorkel Surveys

Why does it matter?



Counts are a **biased representation** of abundance

True Relationship

Counts are consistently lower than abundance

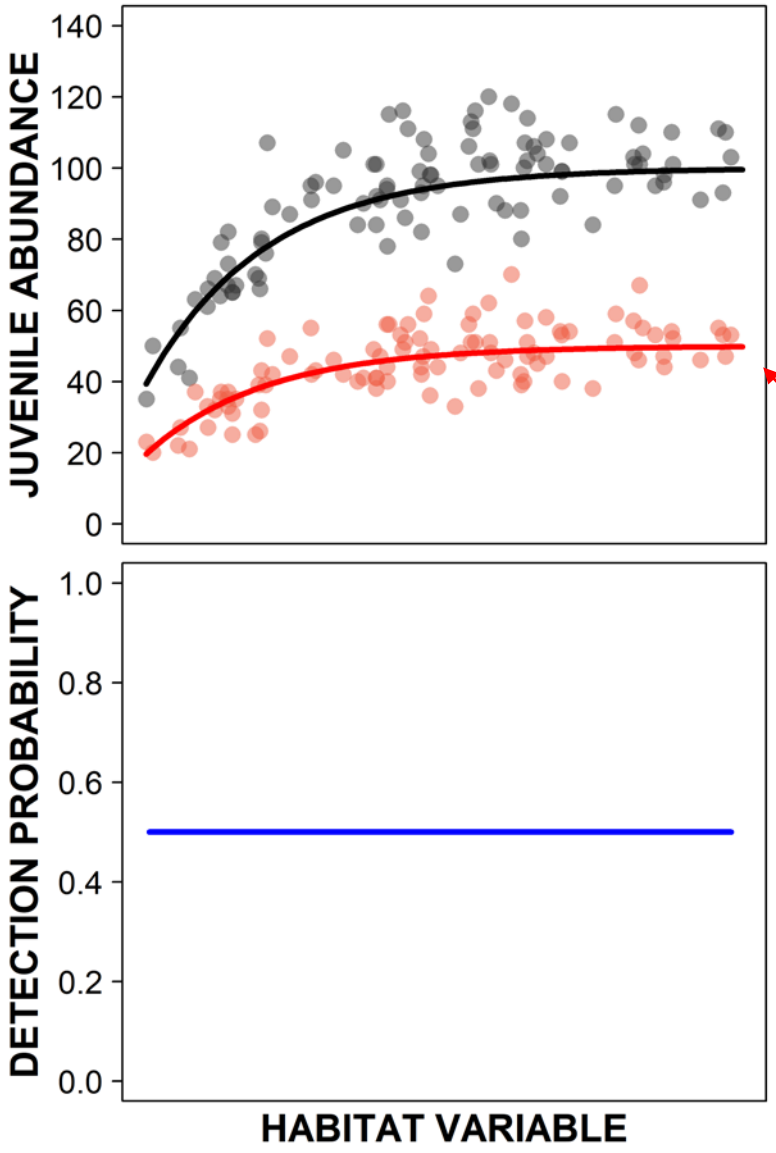
Count half of fish present, regardless of habitat (on average)

True

Observed

Partial Detectability in Snorkel Surveys

Why does it matter?



Counts are a **biased representation** of abundance but can be corrected

True Relationship

Counts are consistently lower than abundance

Correctly assume we count half of fish present

Count half of fish present, regardless of habitat (on average)

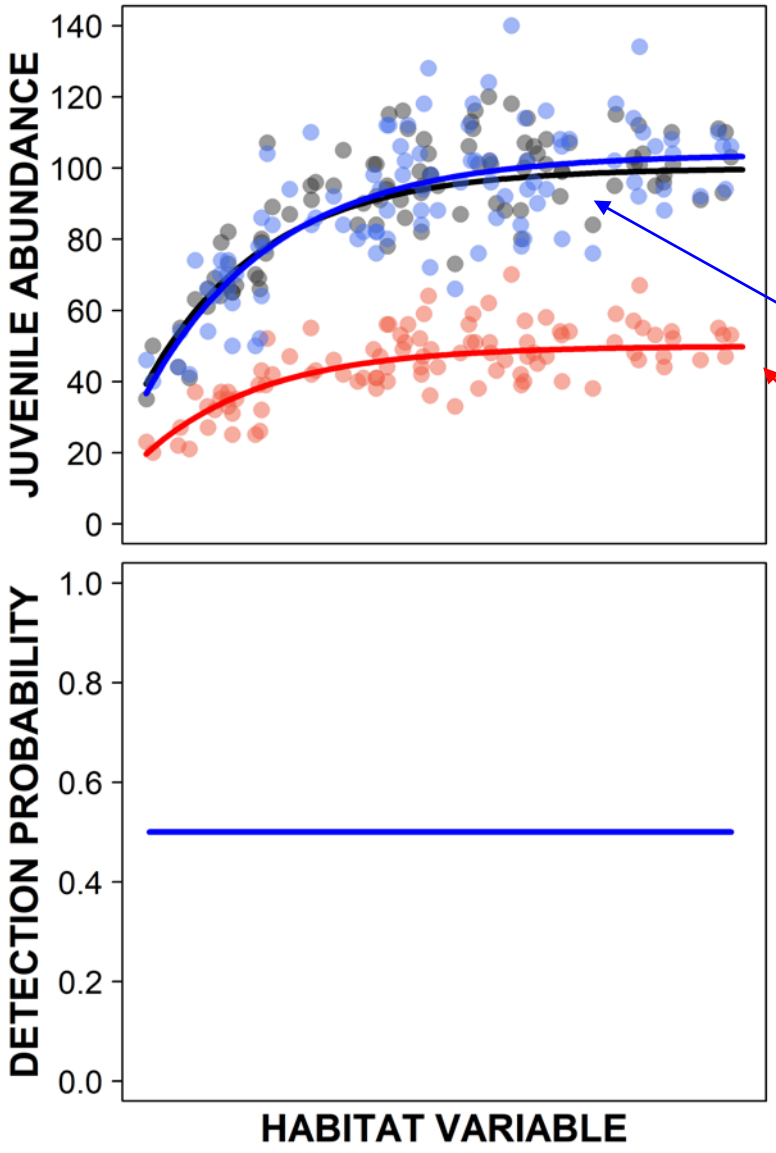
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



Counts are a **biased representation** of abundance but can be corrected

True Relationship

Can still approximate true relationship

Counts are consistently lower than abundance

Correctly assume we count half of fish present

Count half of fish present, regardless of habitat (on average)

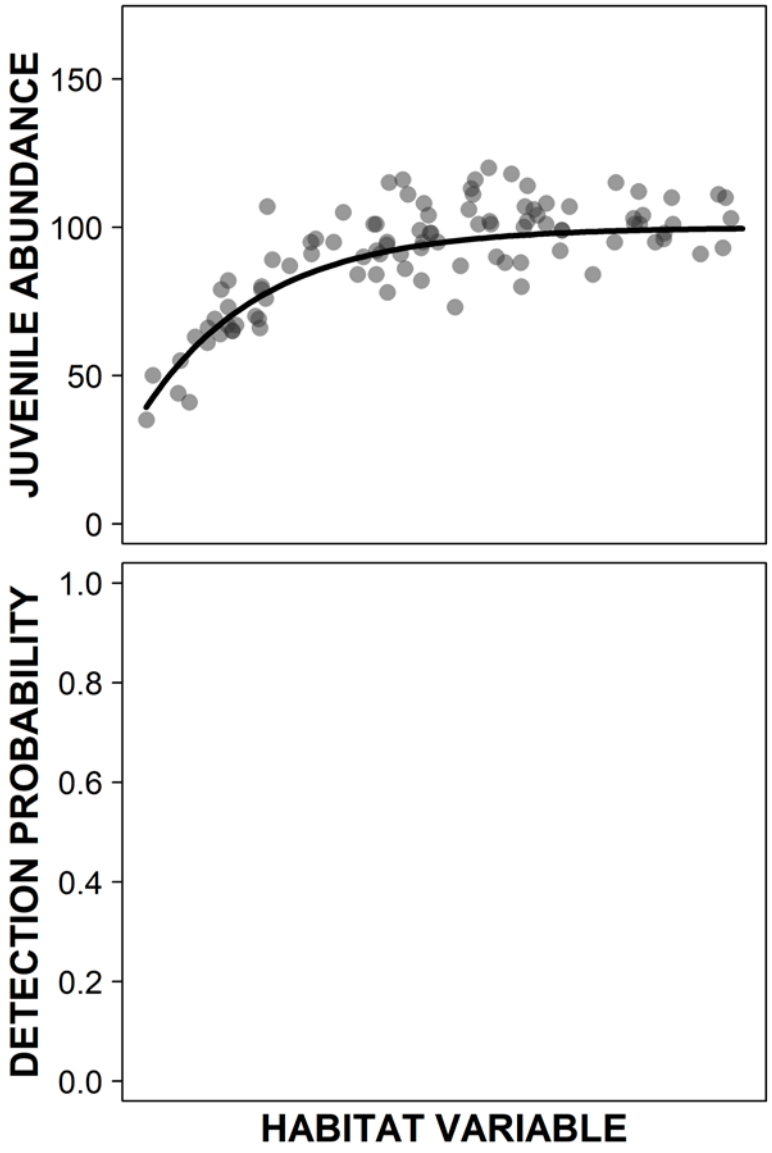
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



If <u>incorrectly</u> accounted for, it can **exacerbate biases**

True Relationship

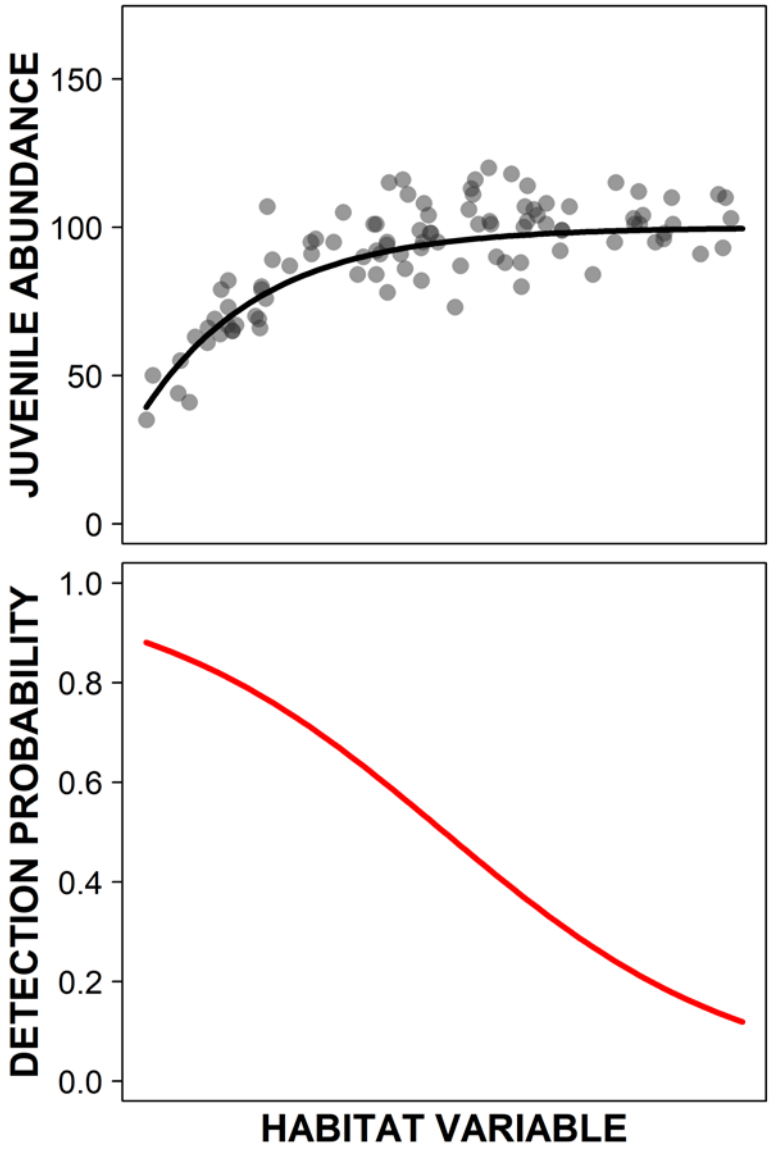
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



If <u>incorrectly</u> accounted for, it can **exacerbate biases**

True Relationship

Fraction of population counted depends on habitat

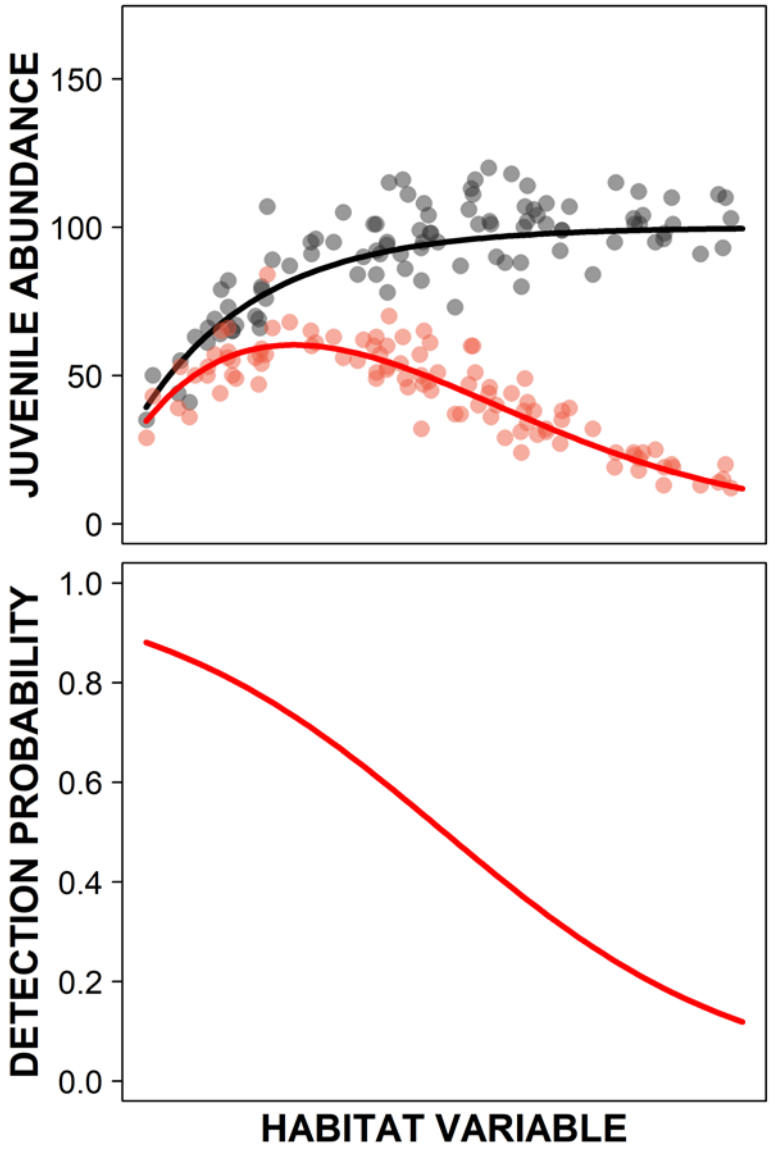
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



If <u>incorrectly</u> accounted for, it can **exacerbate biases**

True Relationship

Raw counts no longer represent the true relationship

Fraction of population counted depends on habitat

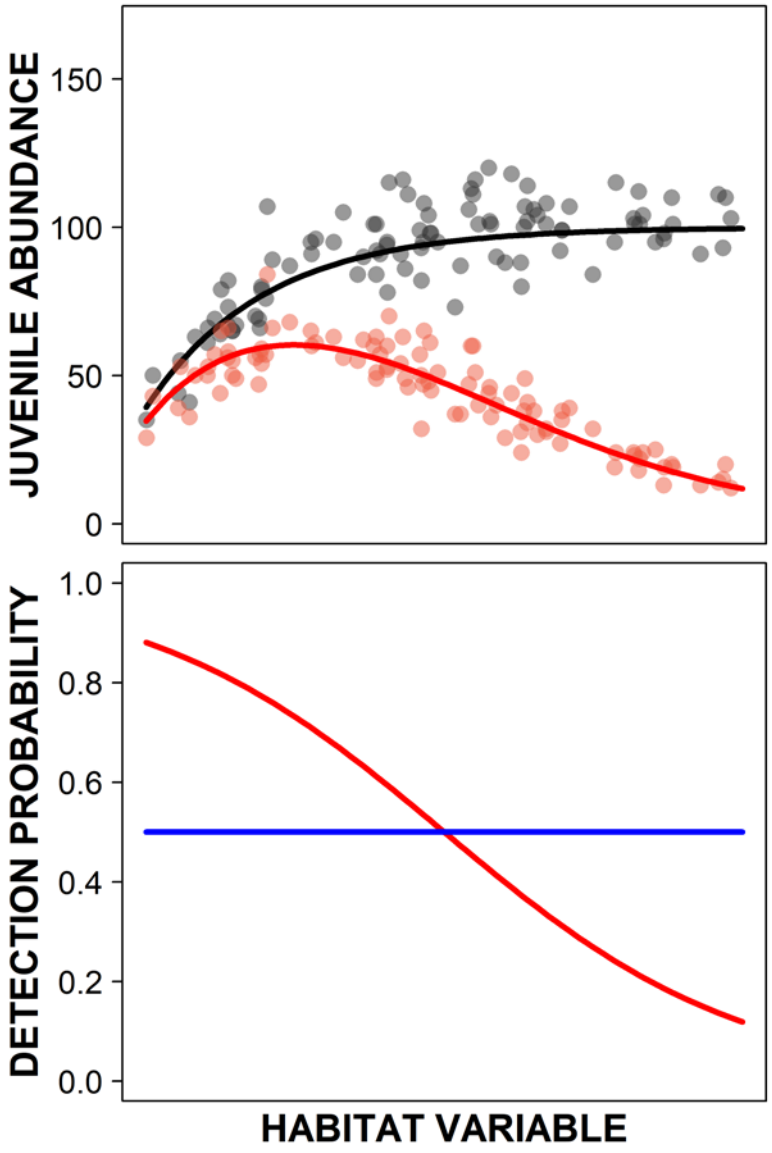
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



If <u>incorrectly</u> accounted for, it can **exacerbate biases**

True Relationship

Raw counts no longer represent the true relationship

Incorrectly assume we count half of fish present

Fraction of population counted depends on habitat

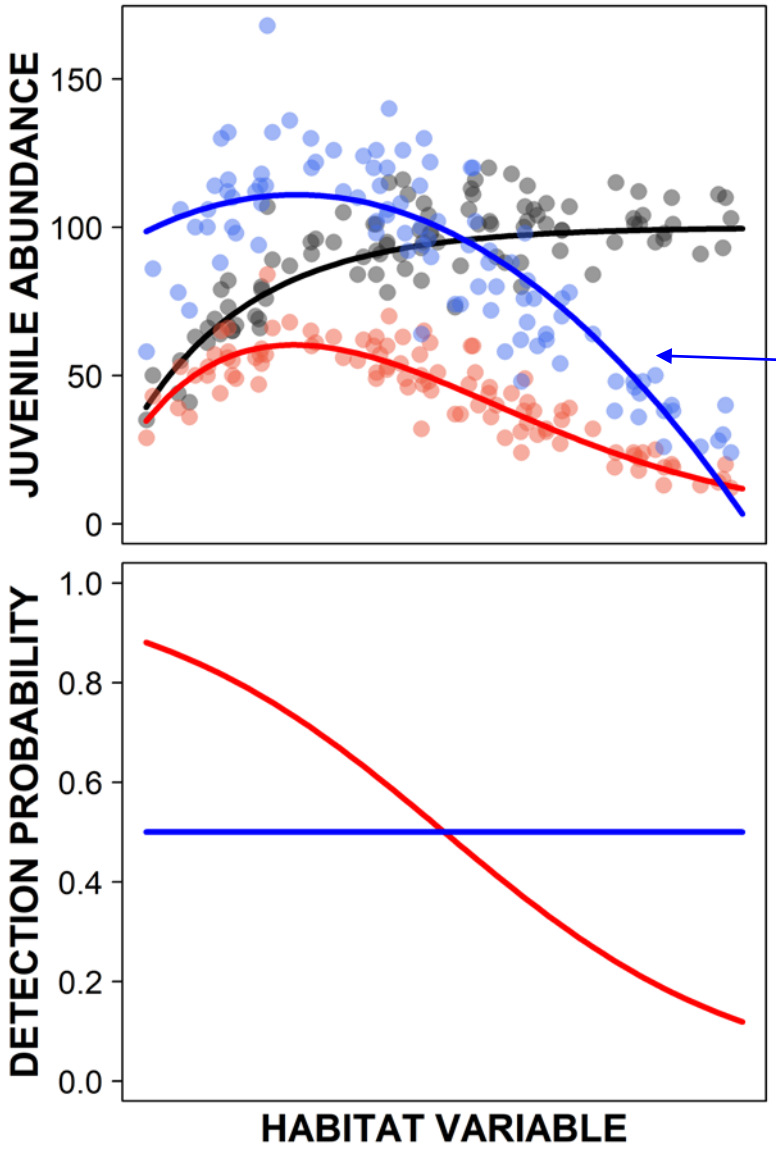
True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

Why does it matter?



If <u>incorrectly</u> accounted for, it can **exacerbate biases**

True Relationship

Do not correctly approximate true relationship

Raw counts no longer represent the true relationship

Incorrectly assume we count half of fish present

Fraction of population counted depends on habitat

True

Observed

Assumed/Estimated

Partial Detectability in Snorkel Surveys

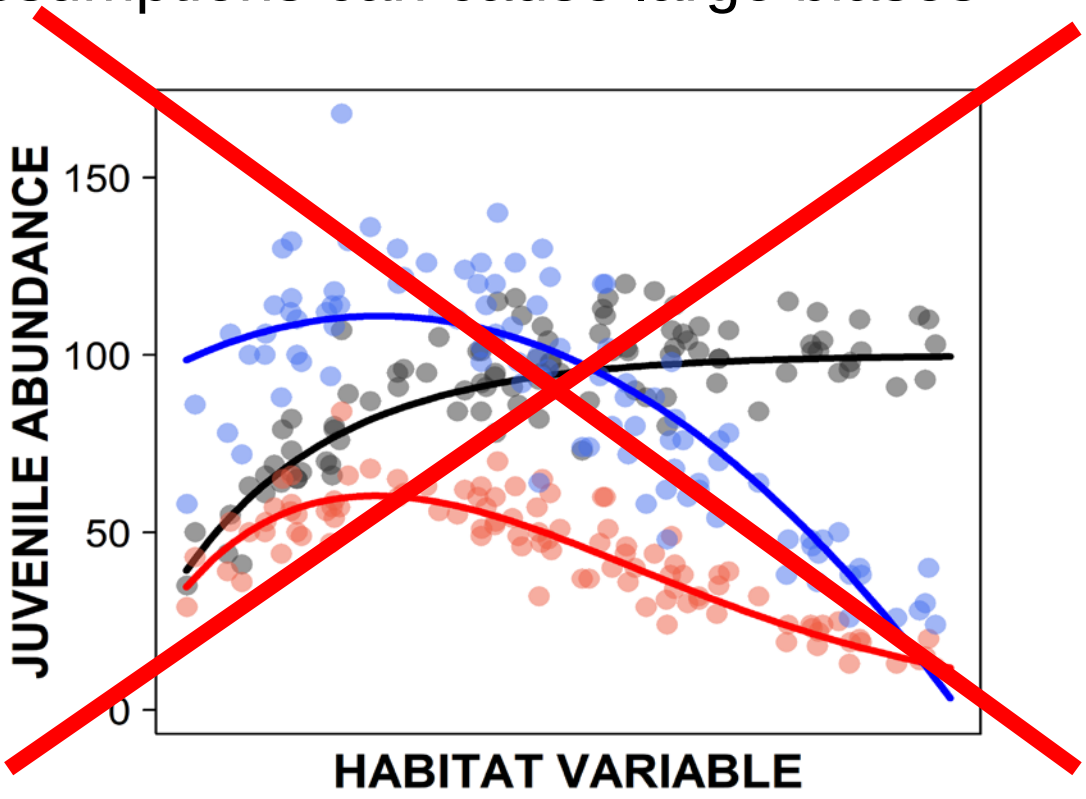
What can we do about it?

- Make assumptions about factors, direction, and magnitude

Partial Detectability in Snorkel Surveys

What can we do about it?

- Make assumptions about factors, direction, and magnitude
  - Bad assumptions can cause large biases



Partial Detectability in Snorkel Surveys

What can we do about it?

- Make assumptions about factors, direction, and magnitude
  - Bad assumptions can cause large biases

- Conduct **directed studies** to quantify it
  - Must have pairs of snorkel counts and abundance

- How to obtain abundance?
  - Depletion estimators (*keep removing fish until you don't catch any more*)
  - Repeated count estimators (*N-mixture models*)
  - Mark-recapture estimators (*based on fraction of population with marks*)

Partial Detectability in Snorkel Surveys

What can we do about it?

- Make assumptions about factors, direction, and magnitude
  - Bad assumptions can cause large biases

- Conduct **directed studies** to quantify it
  - Must have pairs of snorkel counts and abundance

- How to obtain abundance?
  - Depletion estimators
  - Repeated count estimators

  Assumptions difficult to meet, much fish handling, very labor intensive

  - Mark-recapture estimators (*based on fraction of population with marks*)

Partial Detectability in Snorkel Surveys

What can we do about it?

- Make assumptions about factors, direction, and magnitude
  - Bad assumptions can cause large biases

- Conduct **directed studies** to quantify it
  - Must have pairs of snorkel counts and abundance
  - Count/abundance = efficiency -> link to habitat

- How to obtain abundance?
  - Depletion estimators
  - Repeated count estimators

  Assumptions difficult to meet, much fish handling, very labor intensive

  - Mark-recapture estimators

  "Middle ground" on these

Quantifying Partial Detectability

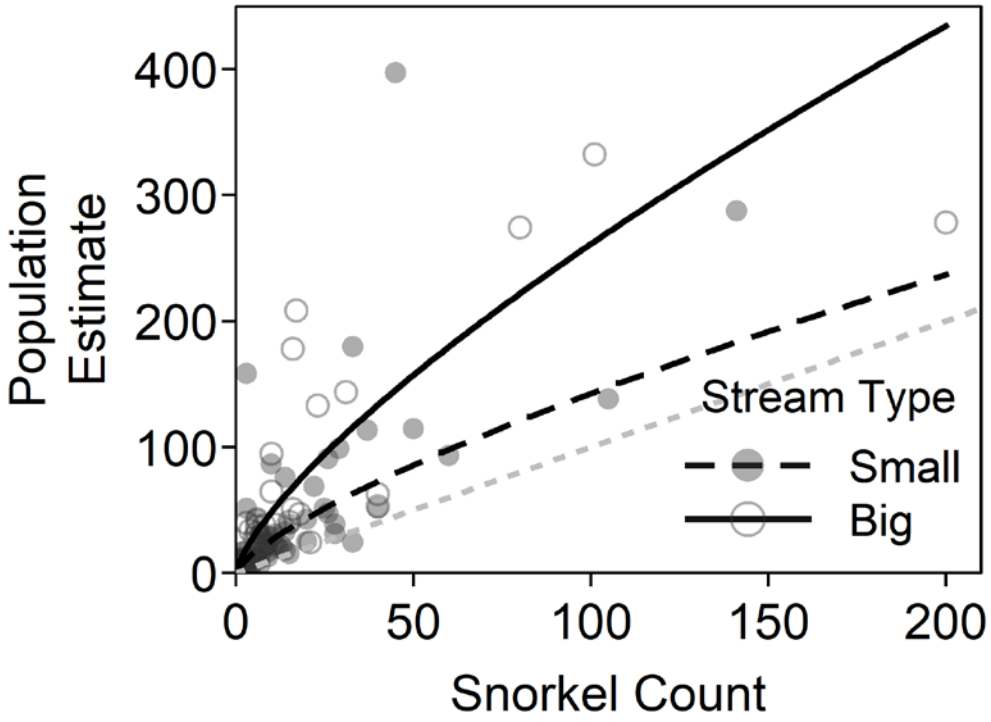
ODFW/CRITFC study

- Summers of 2012/2015; Grande Ronde Basin
- At a wide variety of habitat types:
  (1) Snorkel the channel unit
  (2) Insert block nets
  (3) Capture fish and mark them *(electrofishing + fin clips)*
  (4) Allow population to mix
  (5) Capture fish and count recaptures + novel captures
  (6) Remove block nets
- Obtain a statistical relationship for correcting snorkel counts for partial detectability
- ~100 paired snorkel counts and abundance estimates

Quantifying Partial Detectability

The current approach

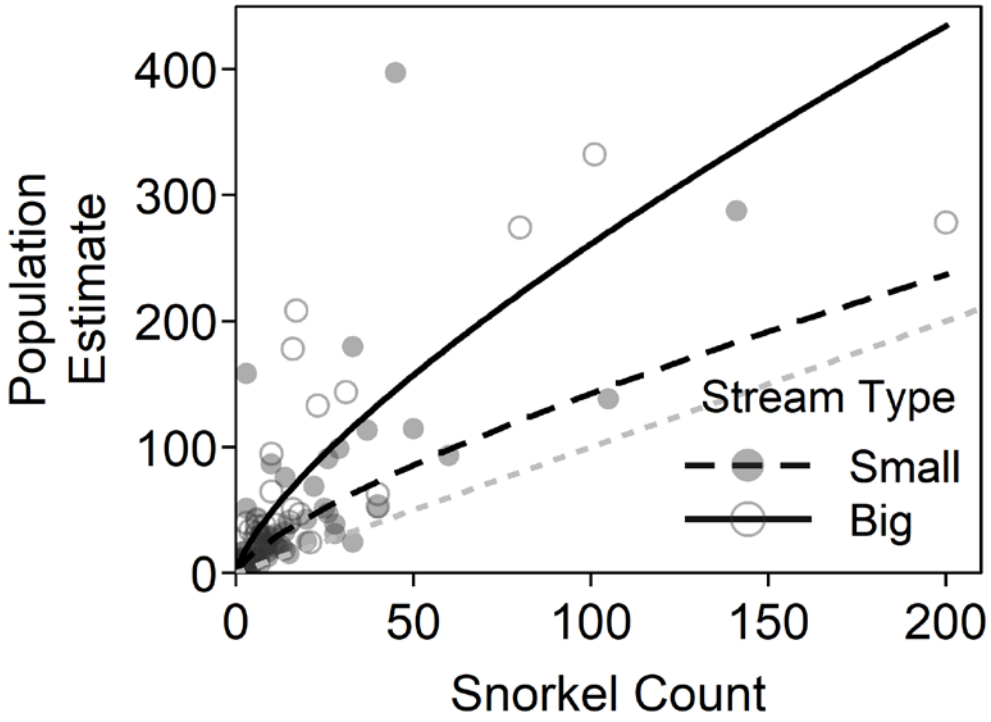
- Regression: predict total salmonid abundance directly from snorkel count
- Same count gives different abundance in small vs. large streams



Quantifying Partial Detectability

***Potential problems** with the current approach*

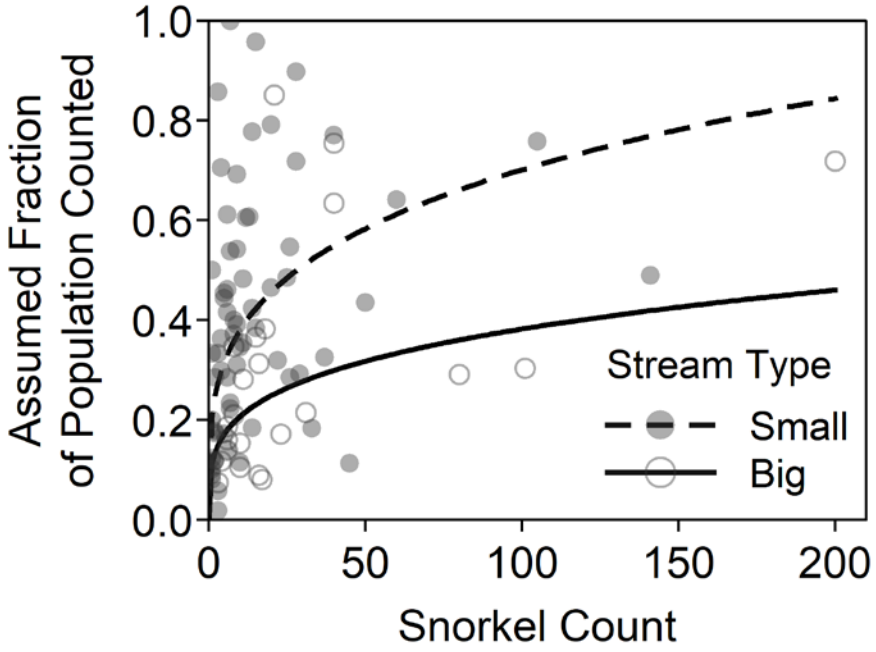
- The population estimate is highly uncertain, ignored
- Assumes stream size (binary) is the only factor impacting detection
- Assumes all salmonids are equally detectable

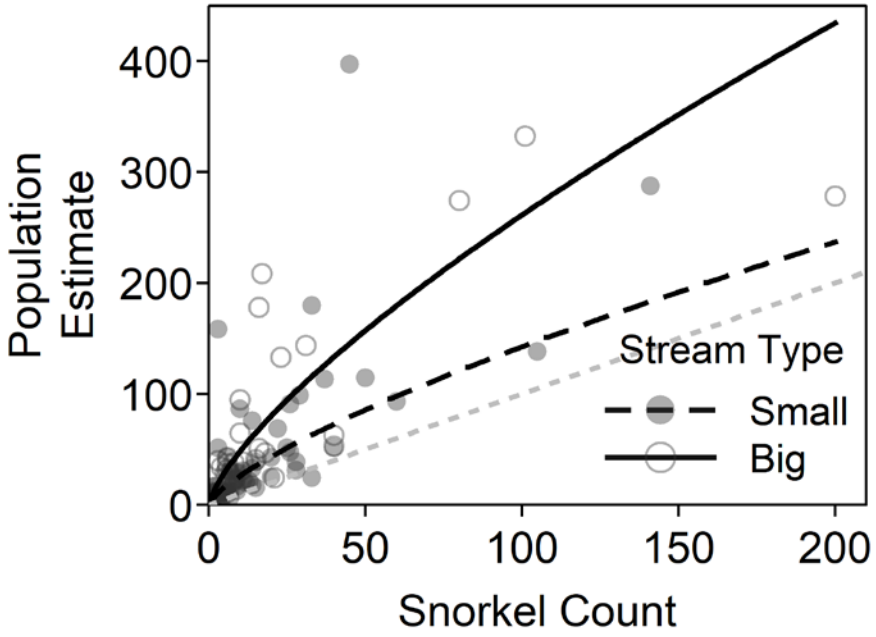


Quantifying Partial Detectability

***Potential problems** with the current approach*

- The population estimate is highly uncertain, ignored
- Assumes stream size (binary) is the only factor impacting detection
- Assumes all salmonids are equally detectable
- Assumes detection probability is a function of count





Rethinking Partial Detectability

As a binomial process

- *N* fish in channel unit *j*
  - Outcome 1: individual *i* is seen (with probability p_j)
  - Outcome 2: individual *i* is not seen (with probability 1 - p_j)
- p_j is a complex function of
  - Habitat conditions, fish behavior, snorkeler aptitude
- If *N* is known, we can estimate *p* as a function of these covariates via logistic regression

Rethinking Partial Detectability

As a binomial process

- *N* fish in channel unit *j*
  - Outcome 1: individual *i* is seen (with probability *p*)
  - Outcome 2: individual *i* is not seen (with probability 1 - *p*)
- p_j is a complex function of
  - Habitat conditions, fish behavior, snorkeler aptitude
- **If *N* is known**, we can estimate *p* as a function of these covariates via logistic regression

PROBLEM

N is not known! Mark-recapture provides an estimate of *N*, with substantial uncertainty due to sampling variability

SOLUTION

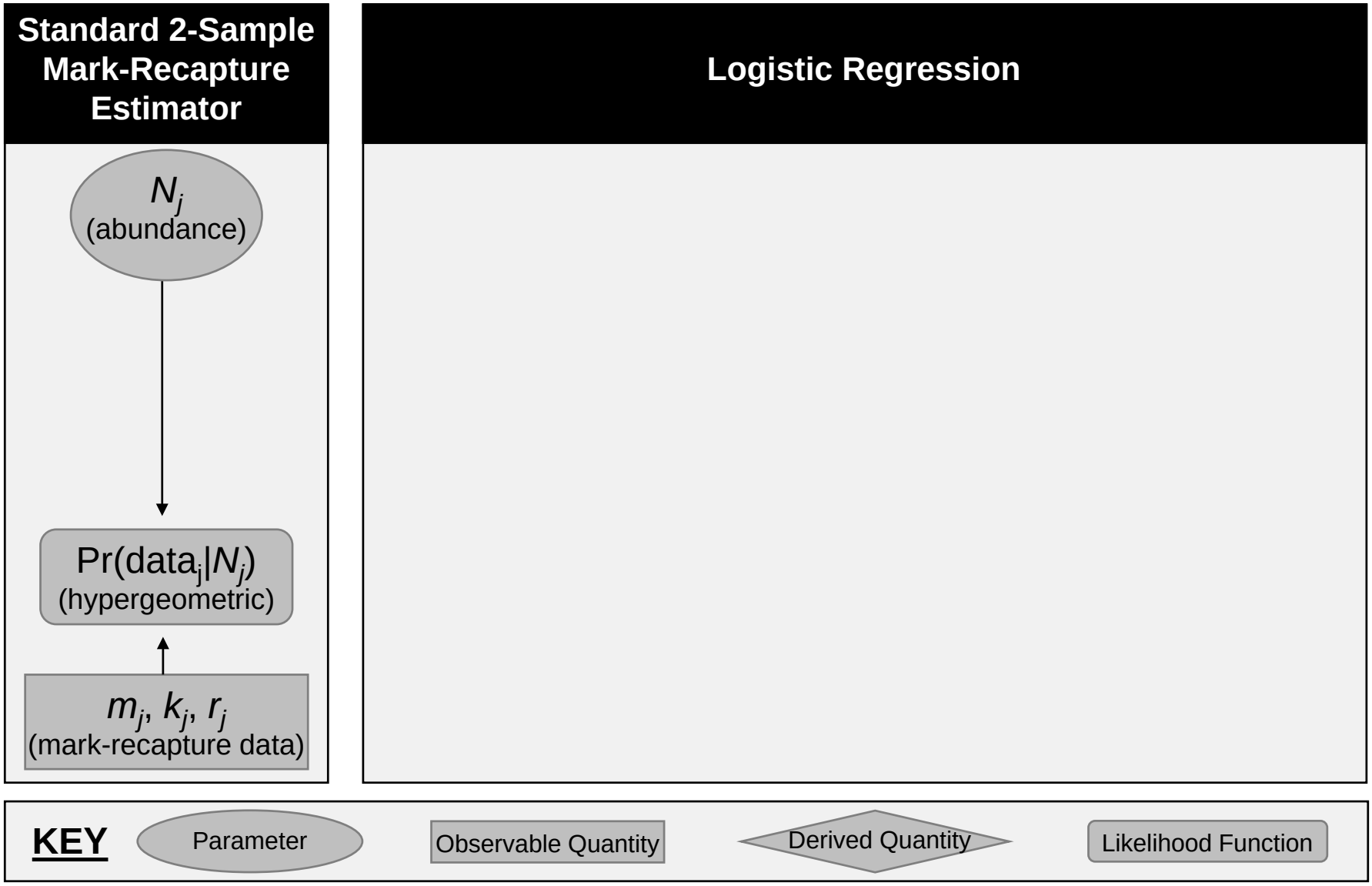
Treat *N* as a latent parameter and apply a joint likelihood to explain both mark-recapture and snorkel data

Estimation of Partial Detectability

Layout of hierarchical model

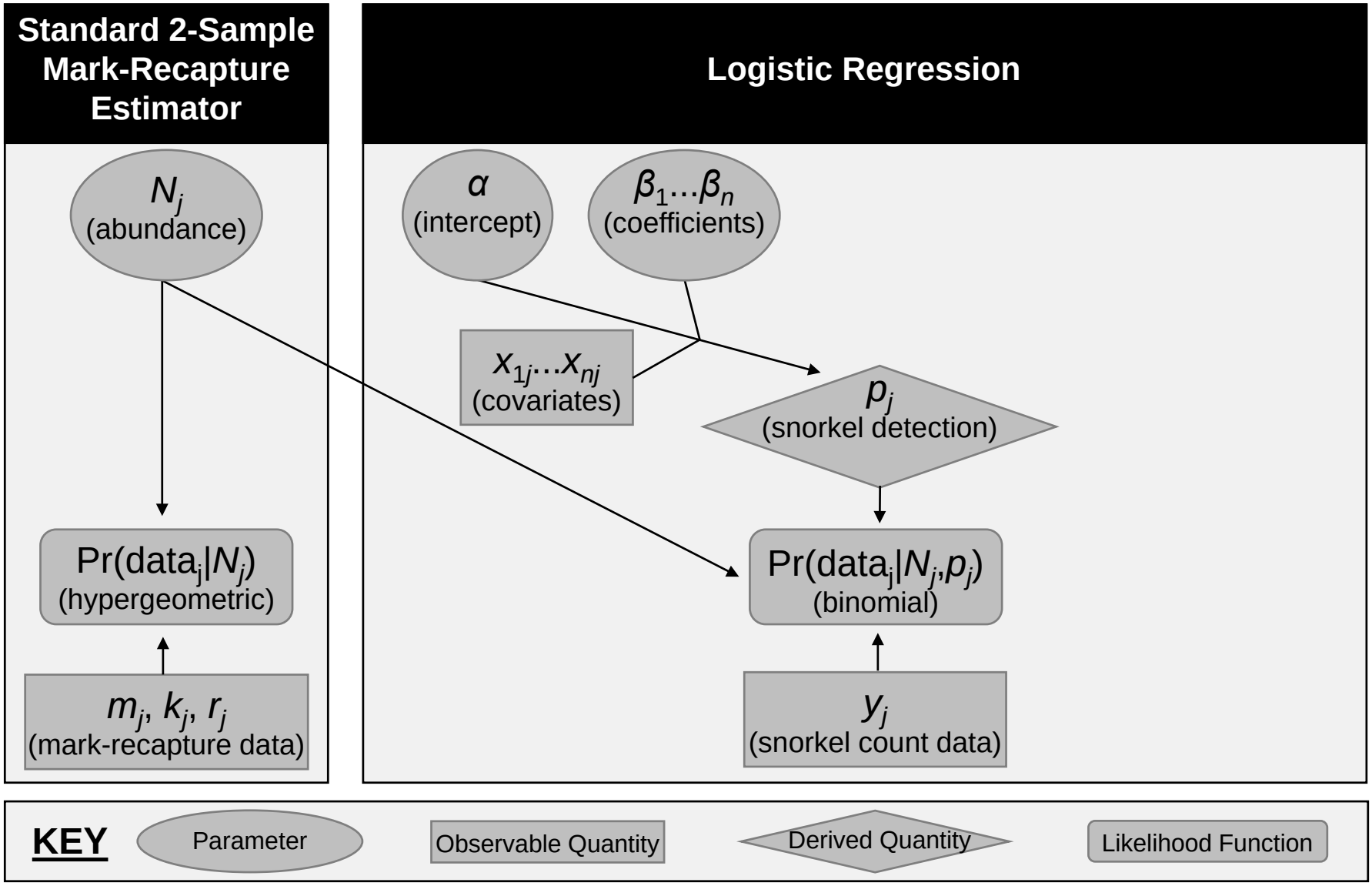
Estimation of Partial Detectability

Layout of hierarchical model



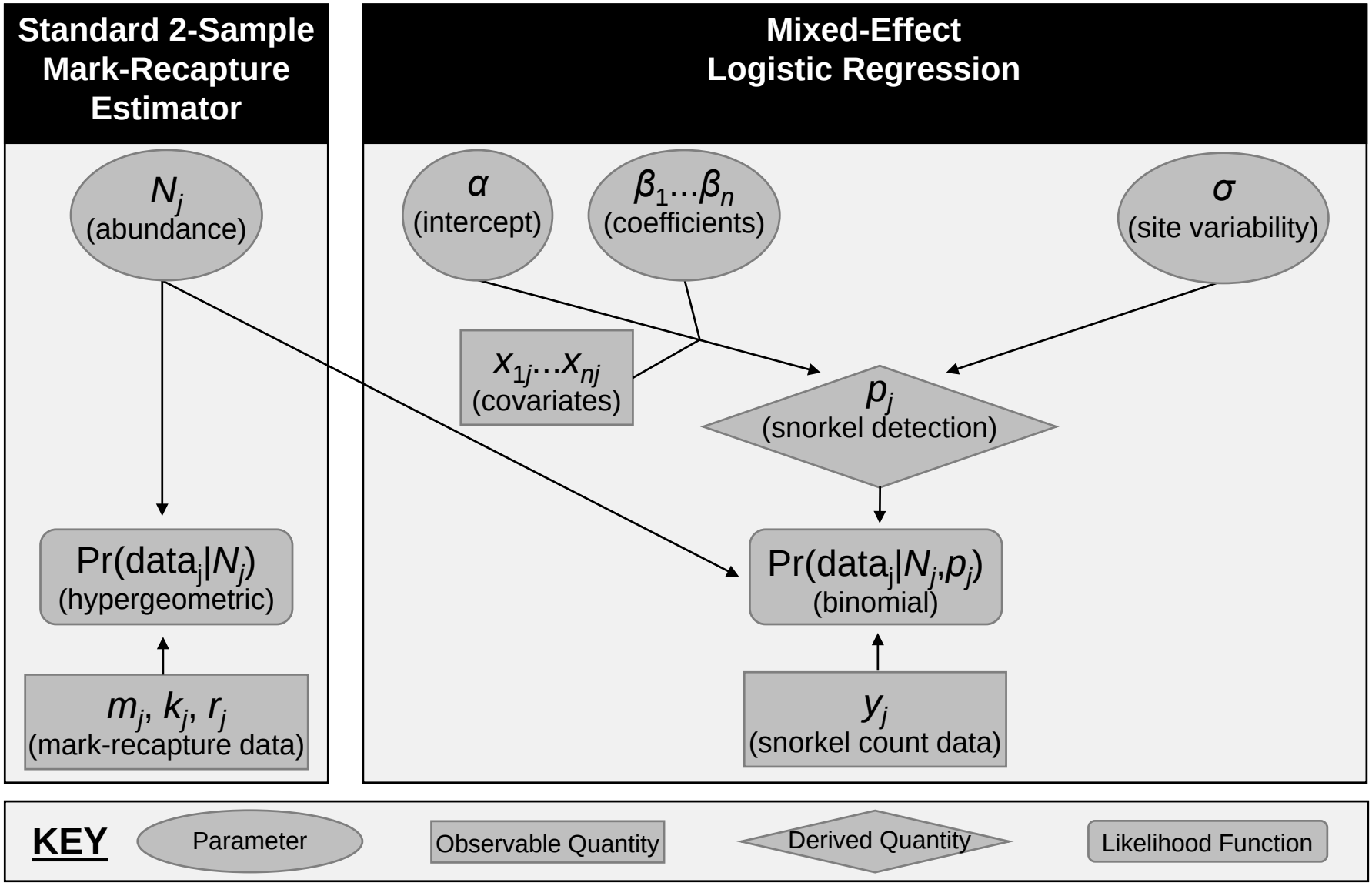
Estimation of Partial Detectability

Layout of hierarchical model



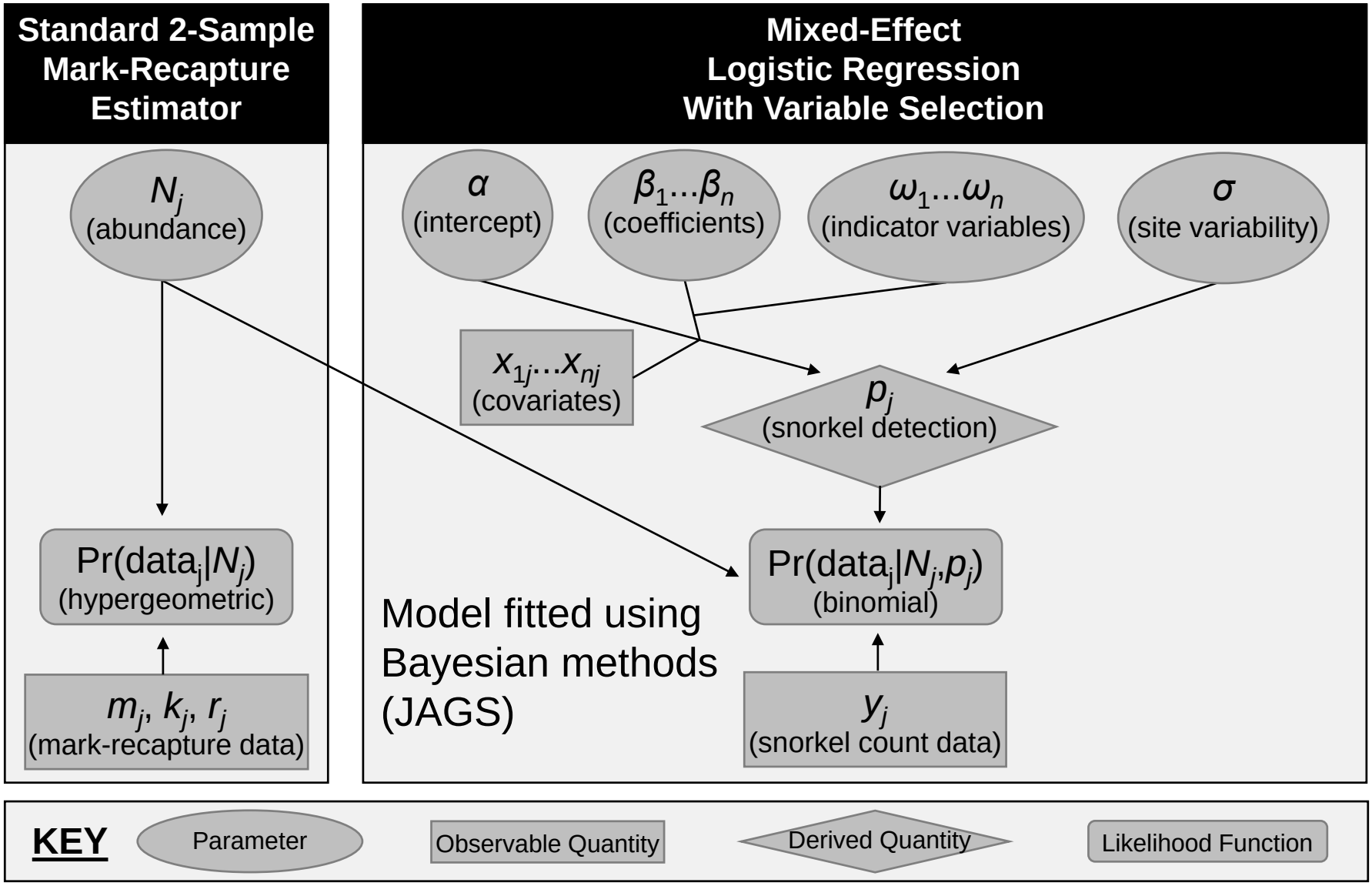
Estimation of Partial Detectability

Layout of hierarchical model



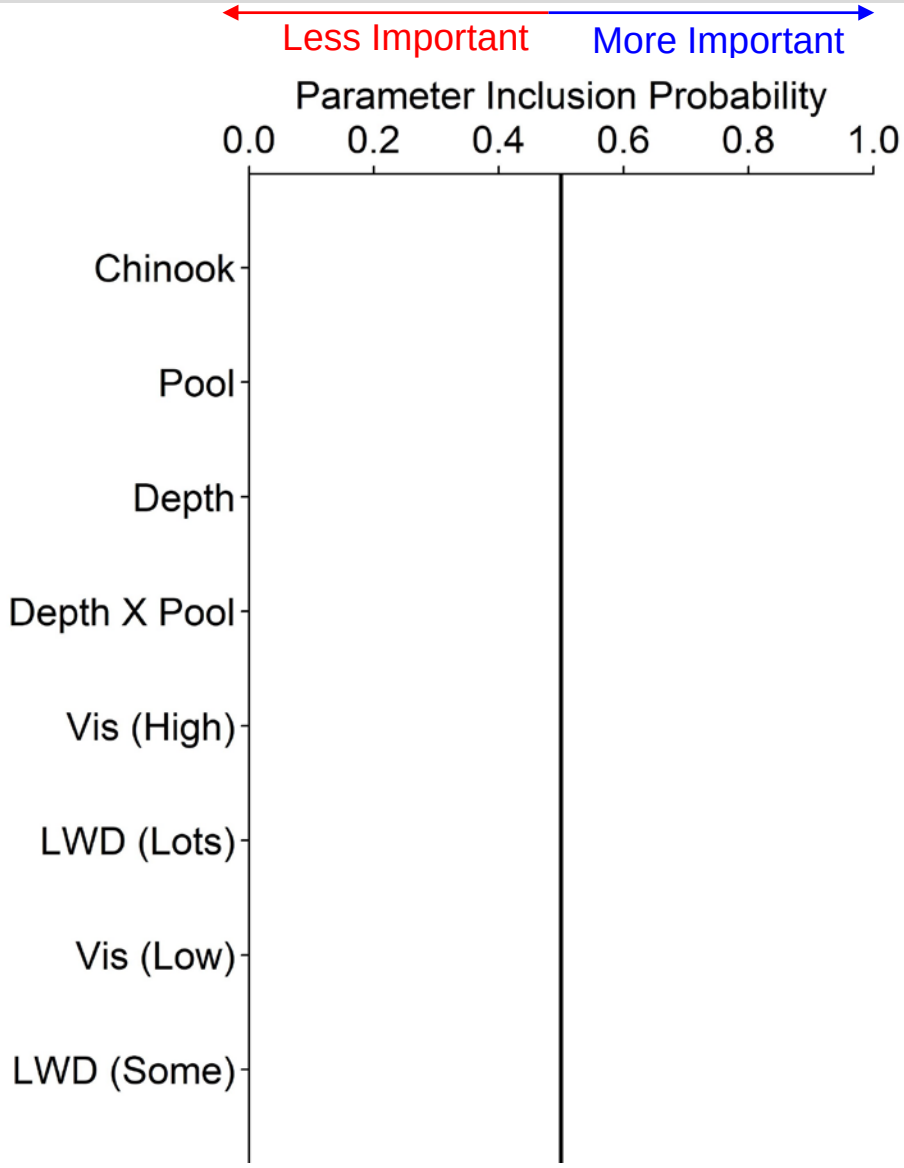
Estimation of Partial Detectability

Layout of hierarchical model



Results

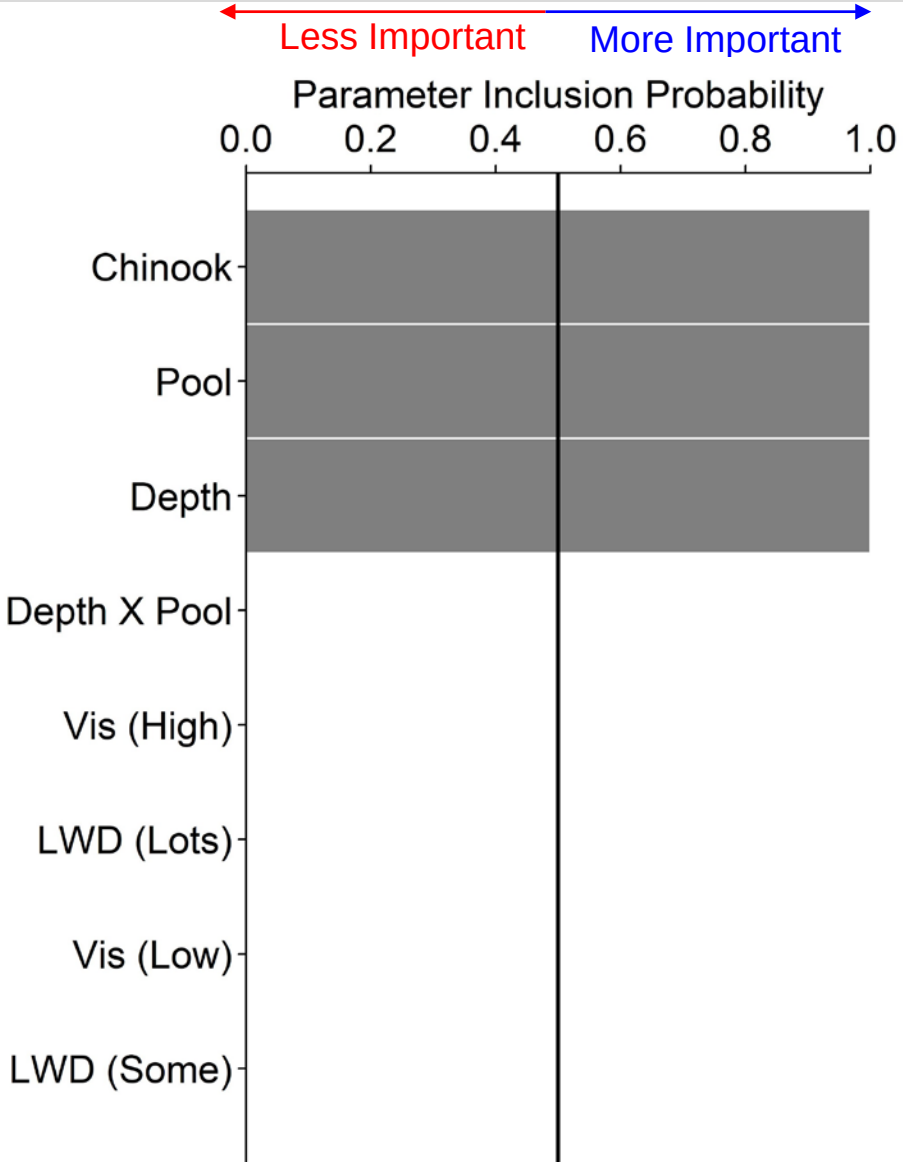
Variable importance: parameter inclusion probability



Results

Variable importance: parameter inclusion probability

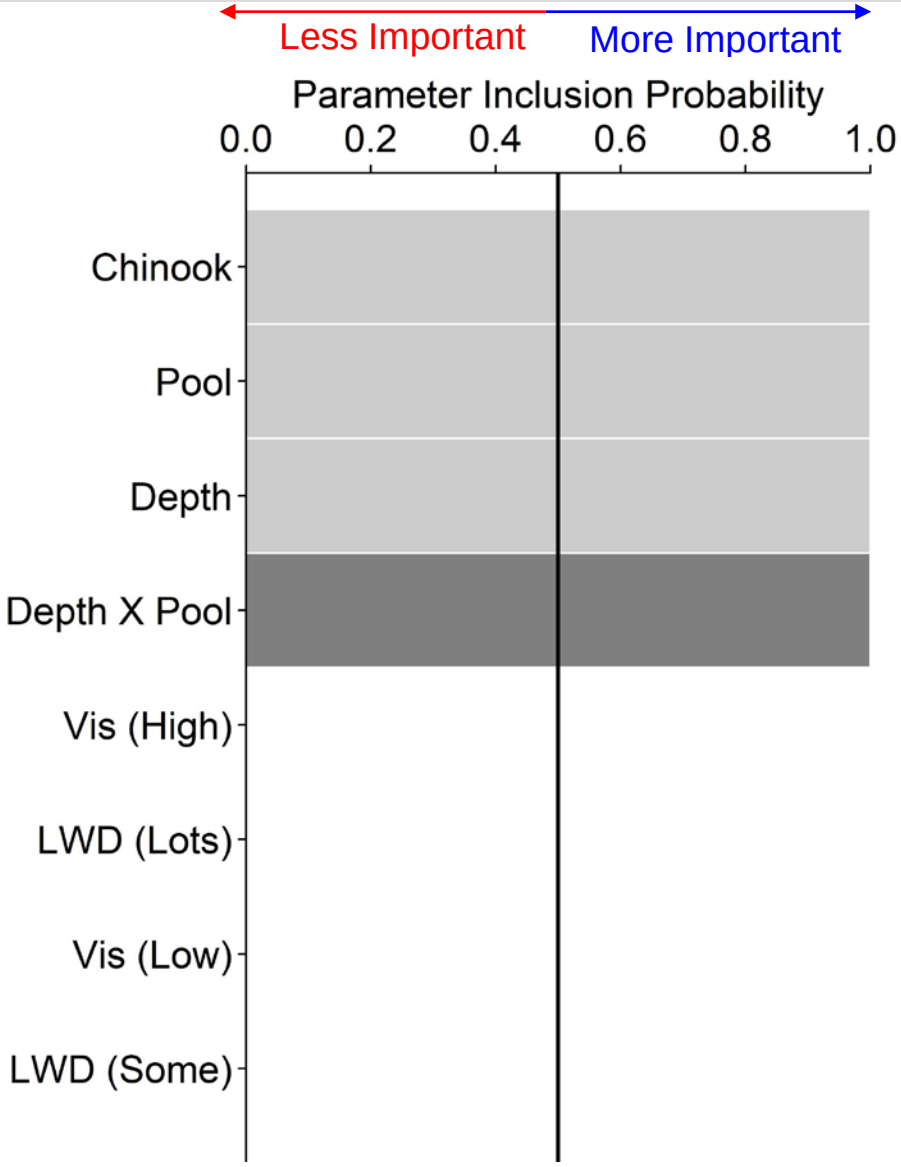
- Strong evidence for *p* varying with
  - Species
  - Unit type
  - Depth



Results

Variable importance: parameter inclusion probability

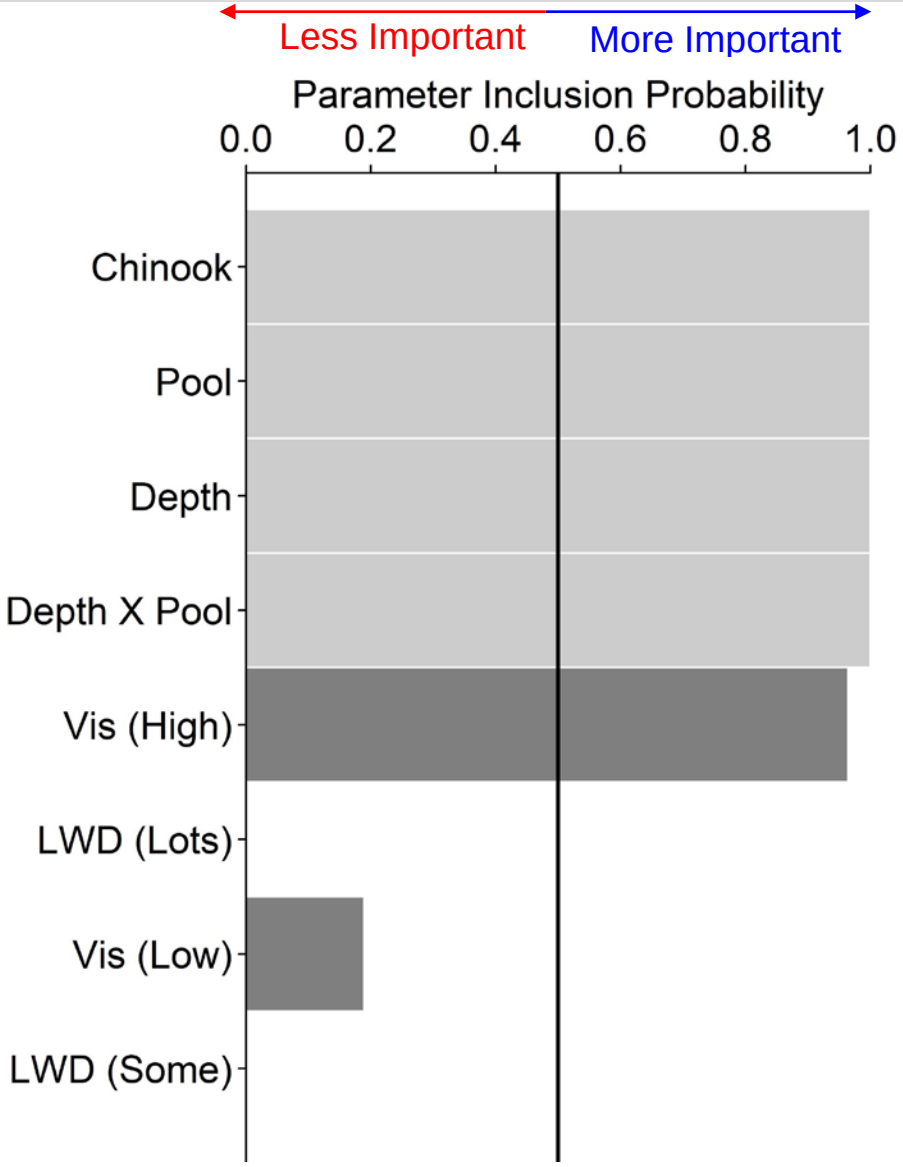
- Strong evidence for *p* varying with
  - Species
  - Unit type
  - Depth
- Effect of depth depends on unit type



Results

Variable importance: parameter inclusion probability

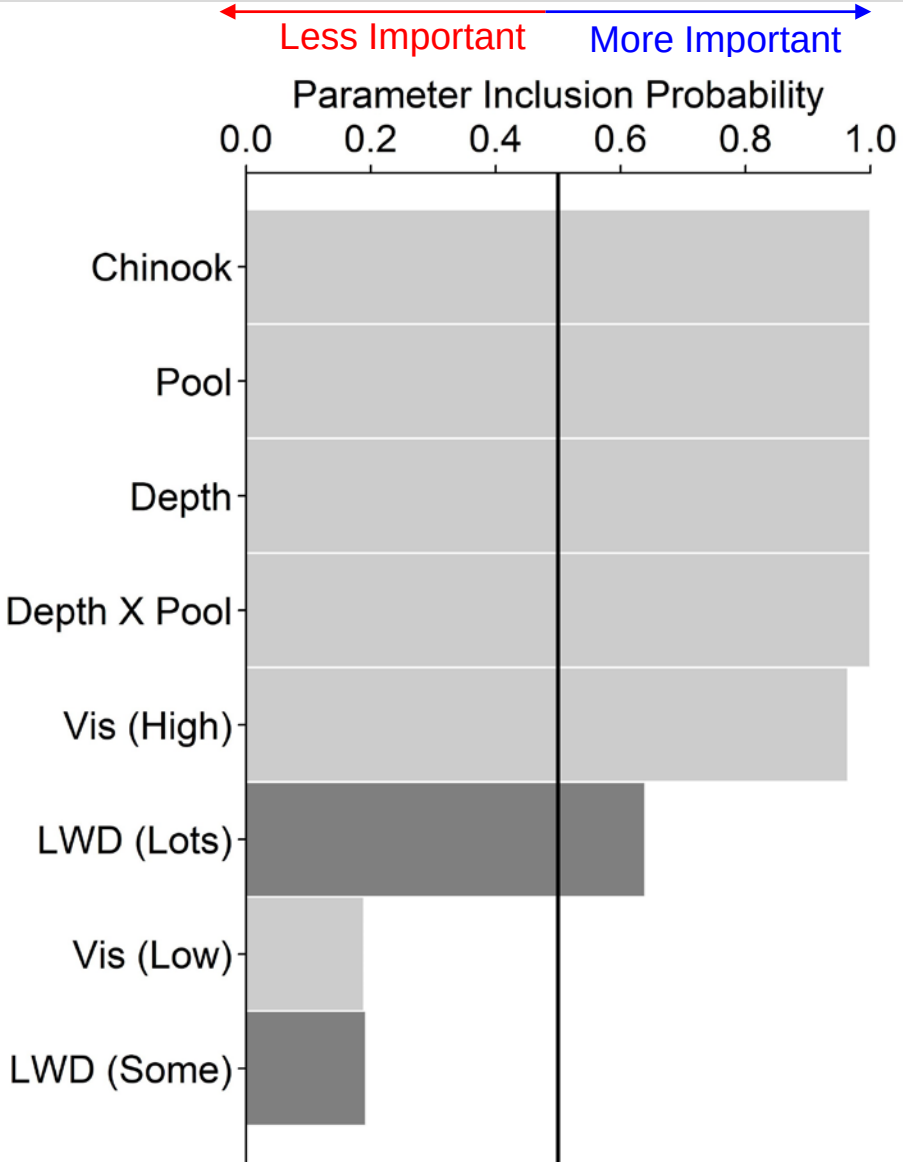
- Strong evidence for *p* varying with
  - Species
  - Unit type
  - Depth
- Effect of depth depends on unit type
- Visibility only important if it is good



Results

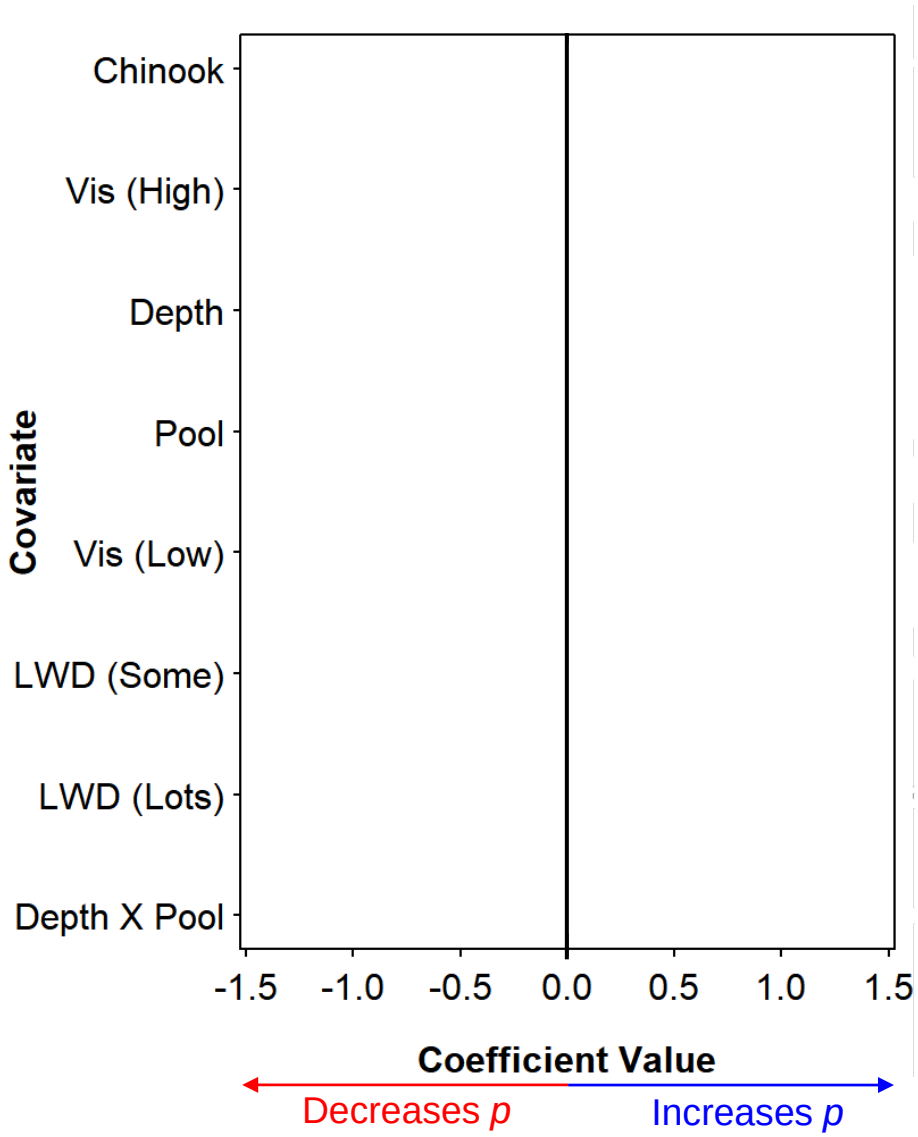
Variable importance: parameter inclusion probability

- Strong evidence for *p* varying with
  - Species
  - Unit type
  - Depth
- Effect of depth depends on unit type
- Visibility only important if it is good
- Wood only important if there is lots of it



Results

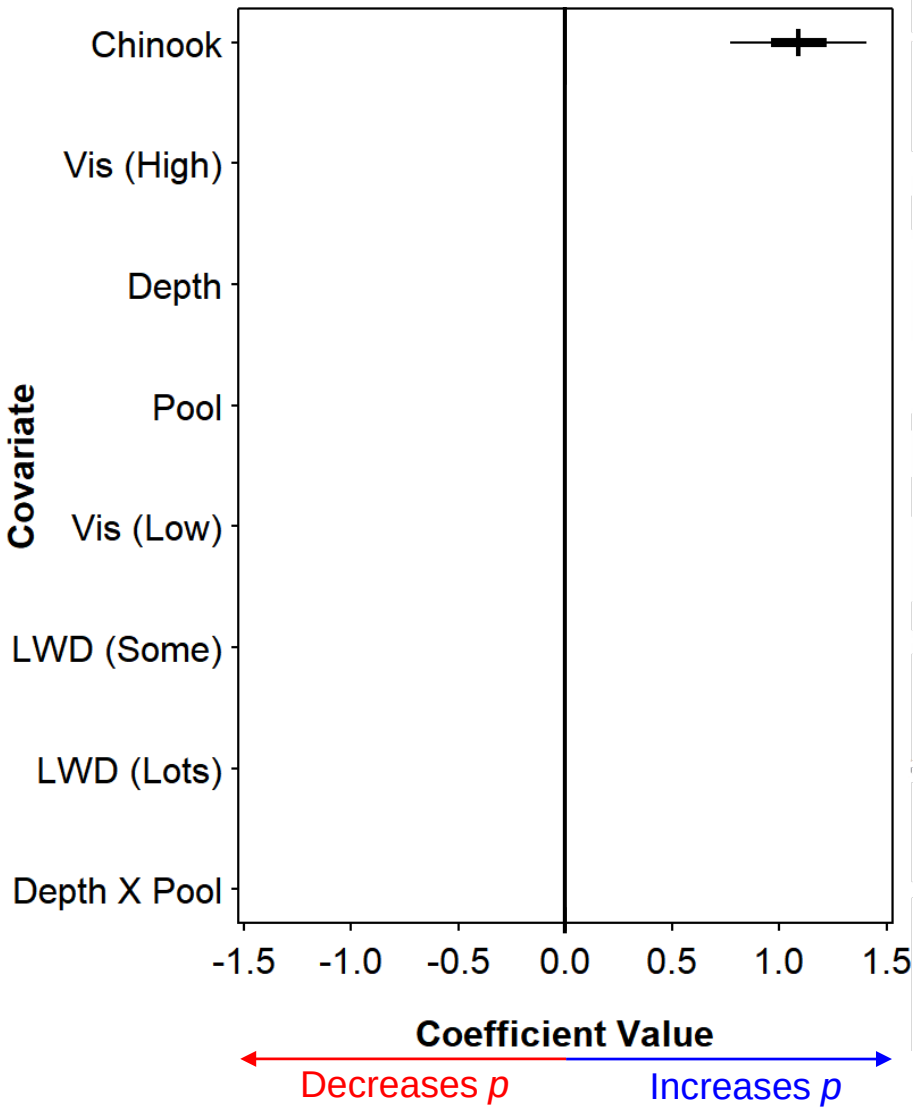
Effect sizes and direction: model-averaged



- Intuitive effect directions

Results

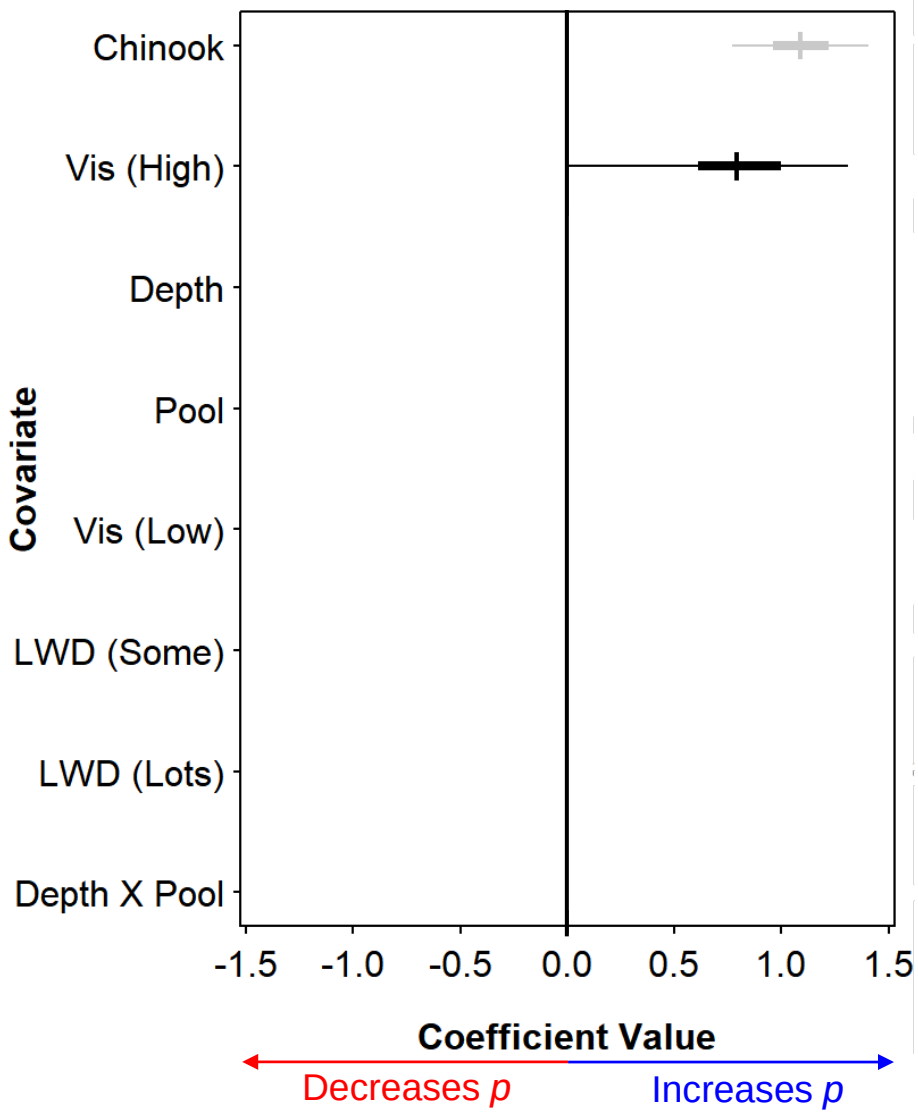
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead

Results

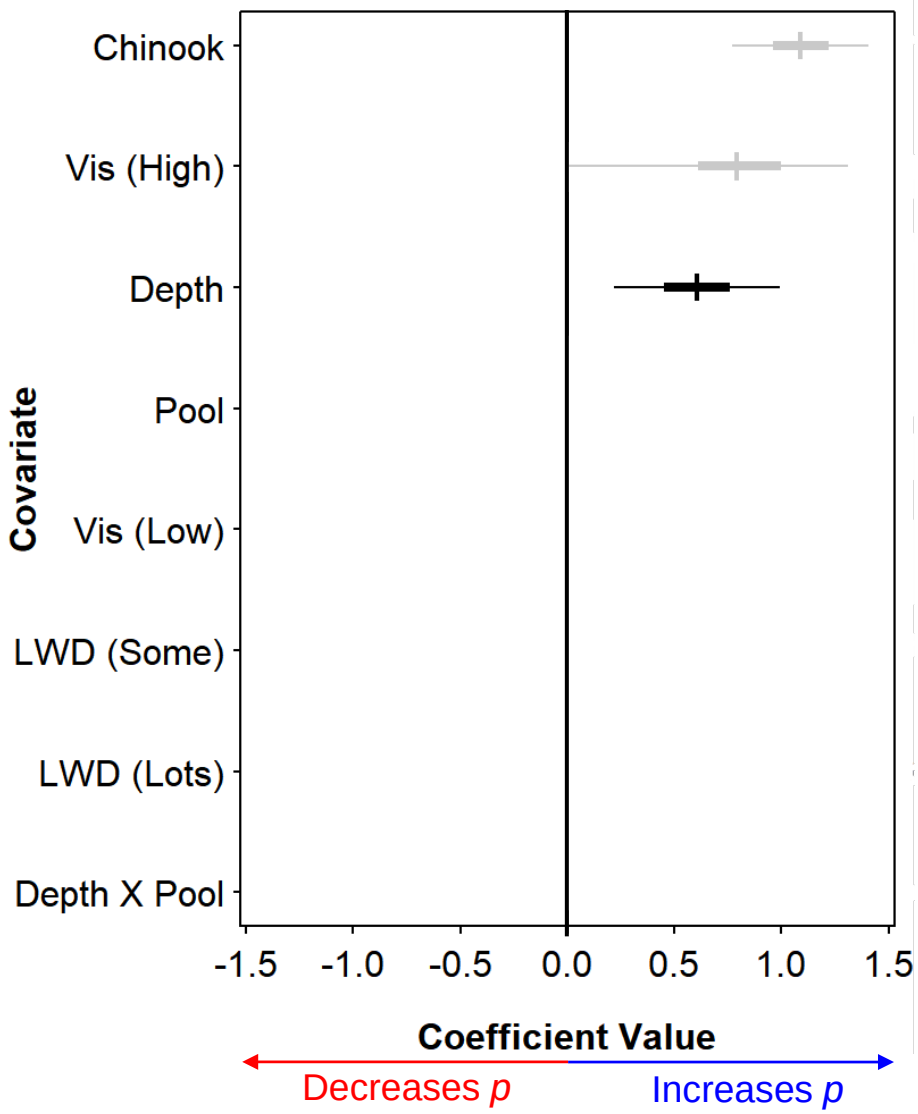
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead
  - High visibility increases detection

Results

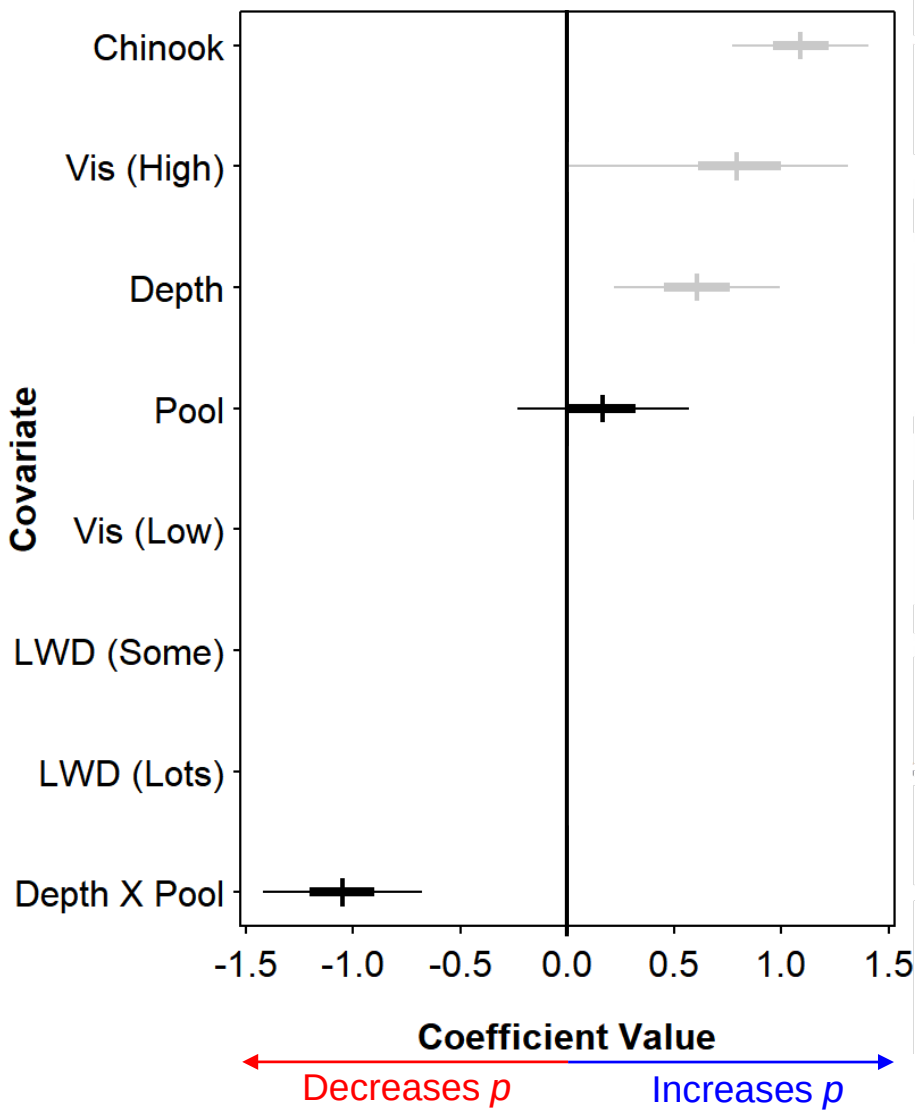
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead
  - High visibility increases detection
  - Depth increases detection in non-pools

Results

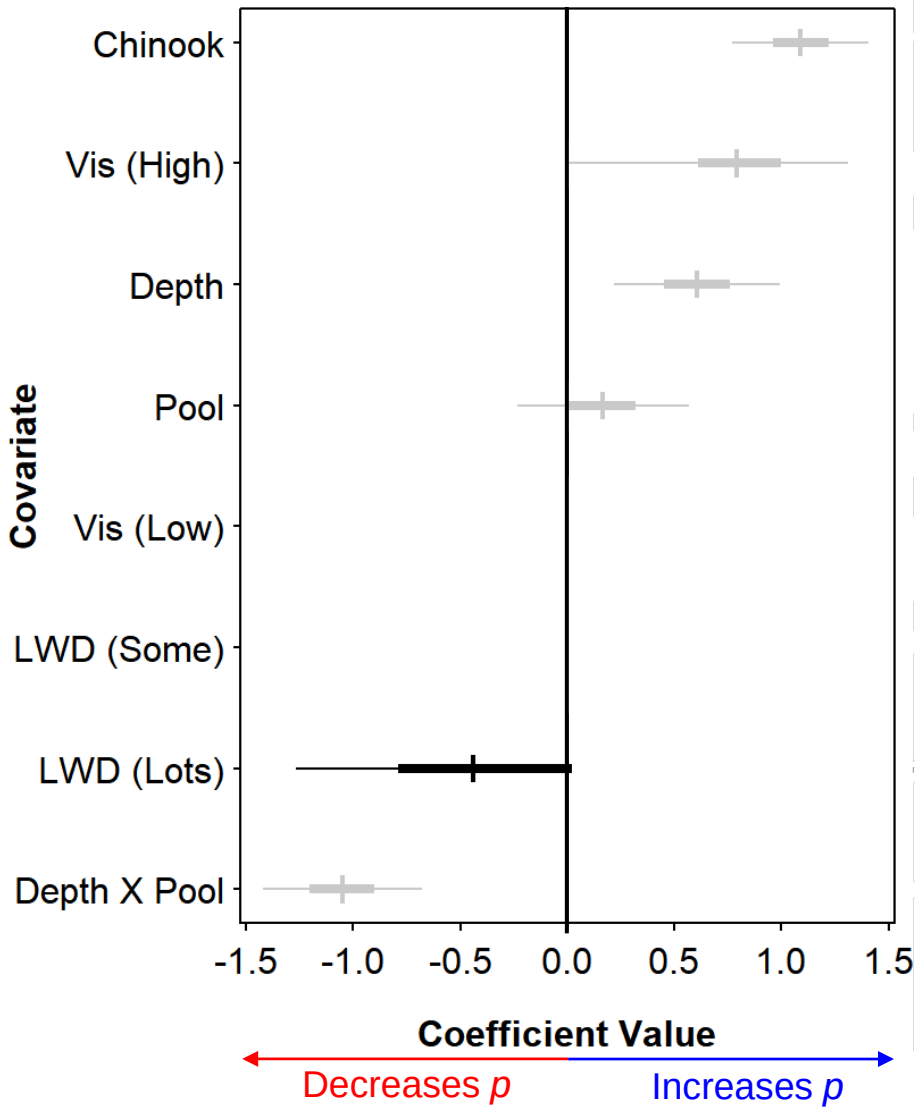
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead
  - High visibility increases detection
  - Depth increases detection in non-pools
  - Depth decreases detection in pools

Results

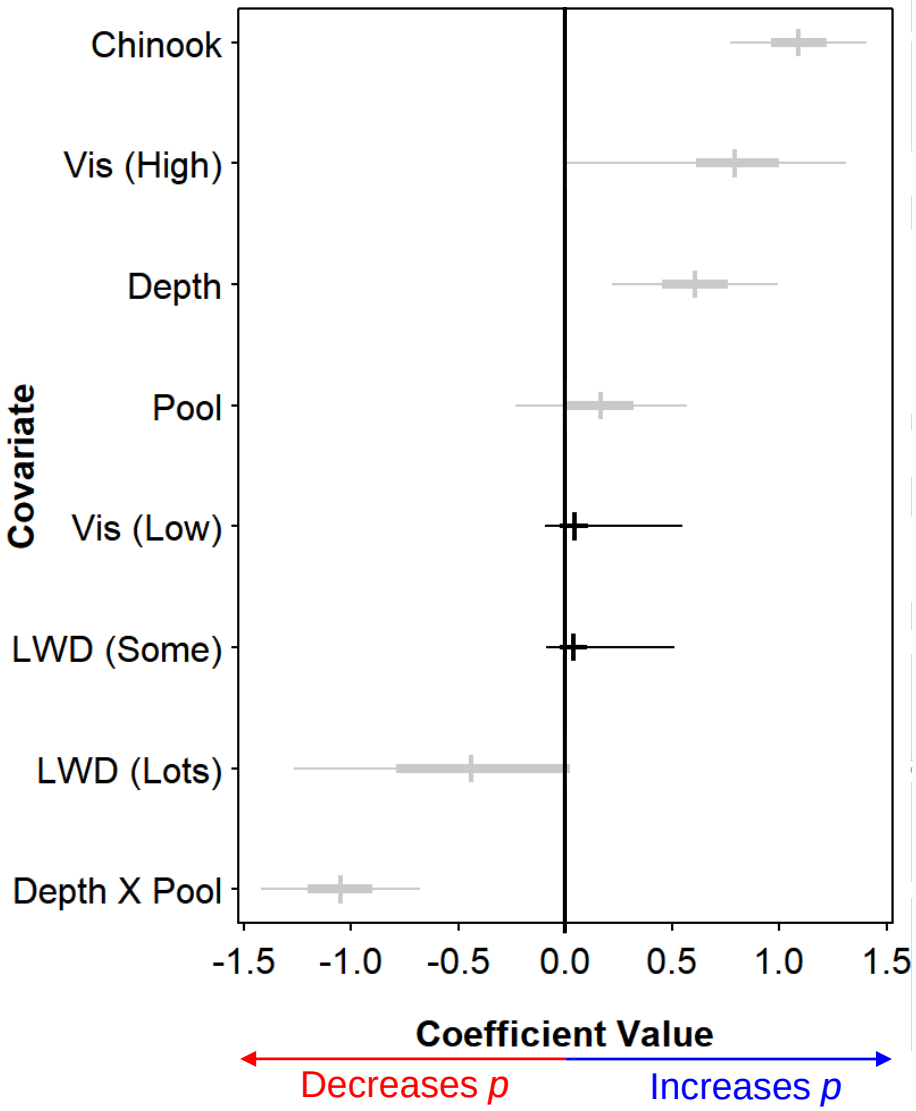
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead
  - High visibility increases detection
  - Depth increases detection in non-pools
  - Depth decreases detection in pools
  - Lots of wood decreases detection

Results

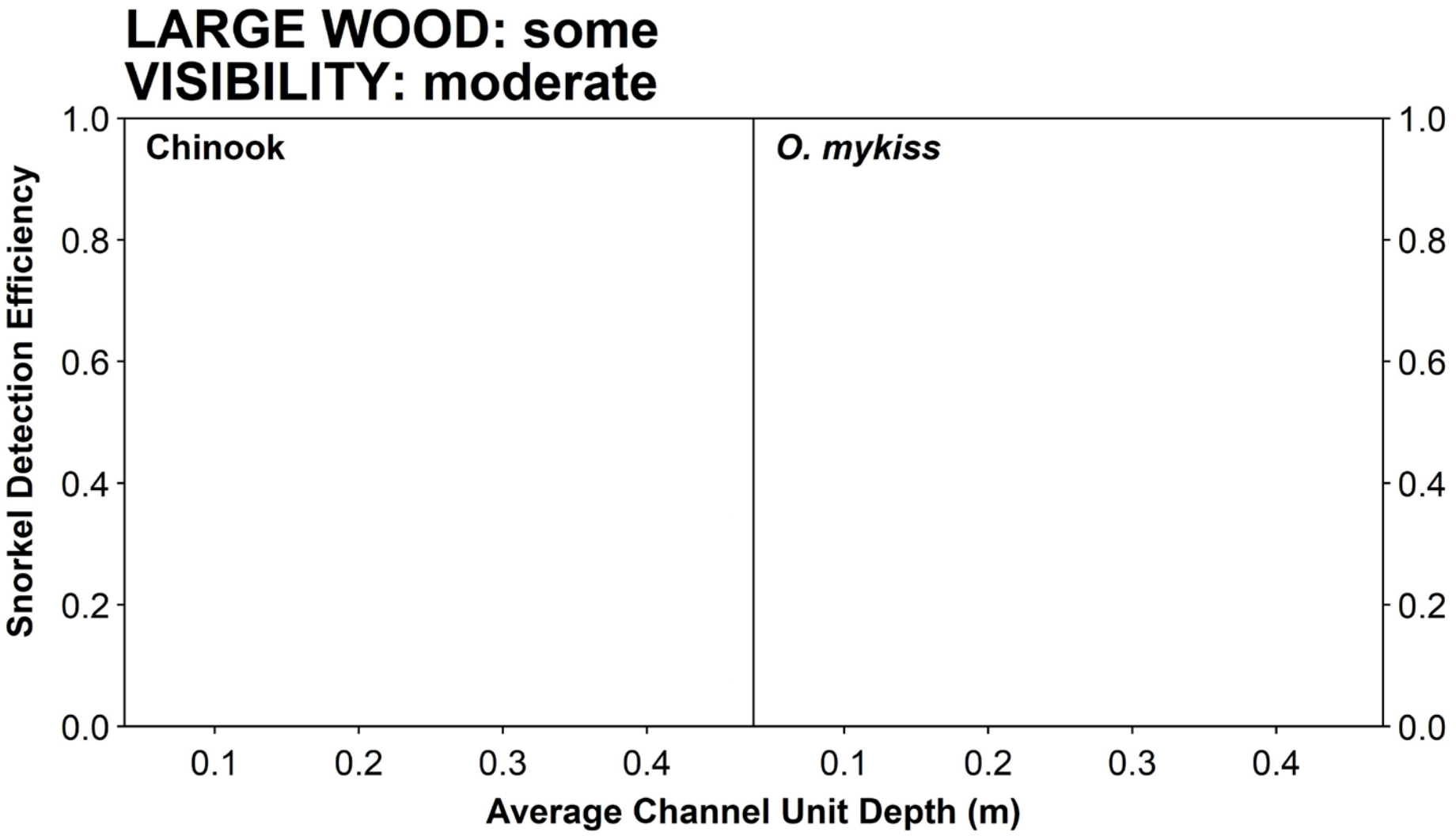
Effect sizes and direction: model-averaged



- Intuitive effect directions
  - Chinook easier to see than steelhead
  - High visibility increases detection
  - Depth increases detection in non-pools
  - Depth decreases detection in pools
  - Lots of wood decreases detection

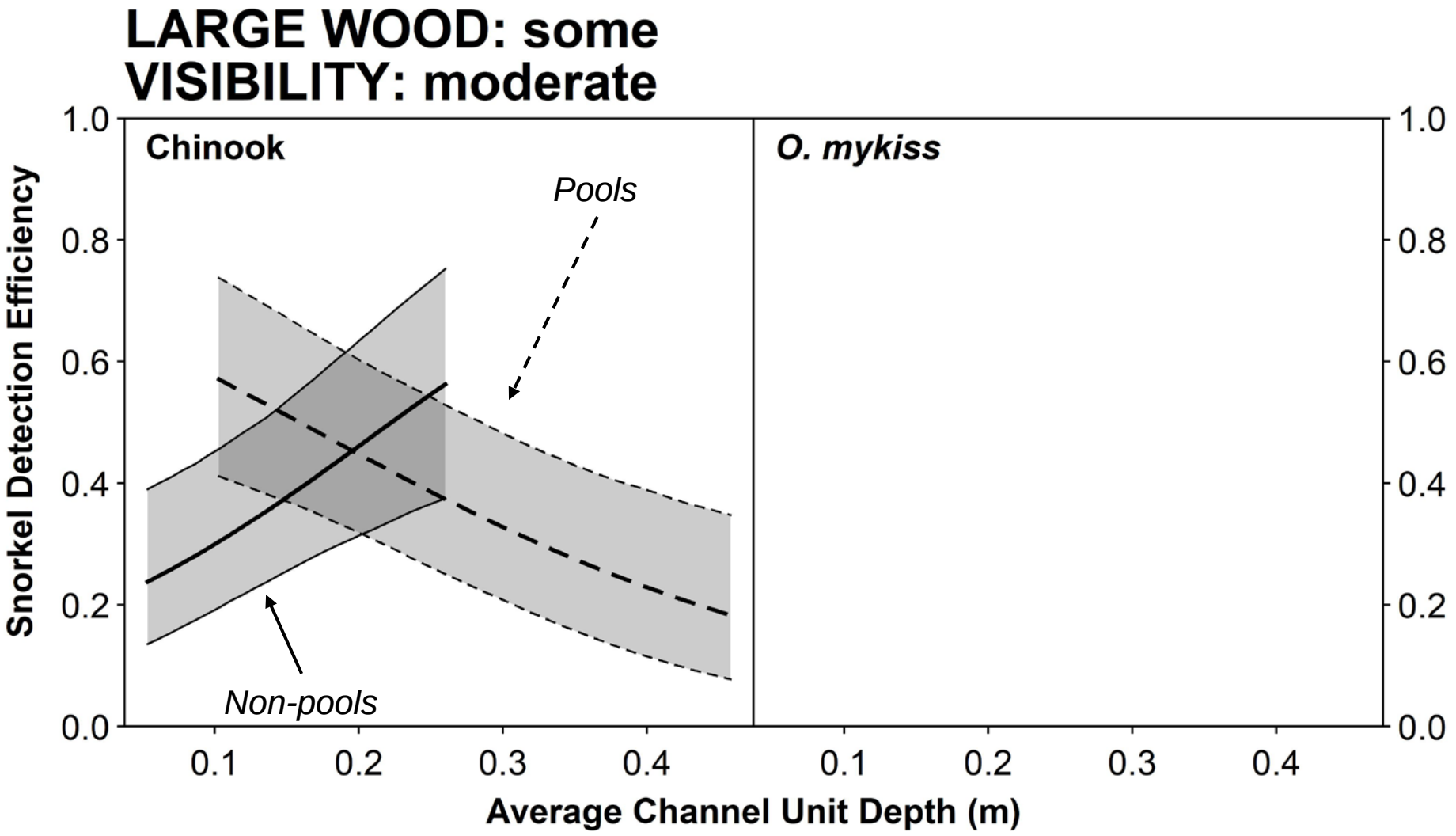
Results

*Effect sizes and direction: **typical** conditions*



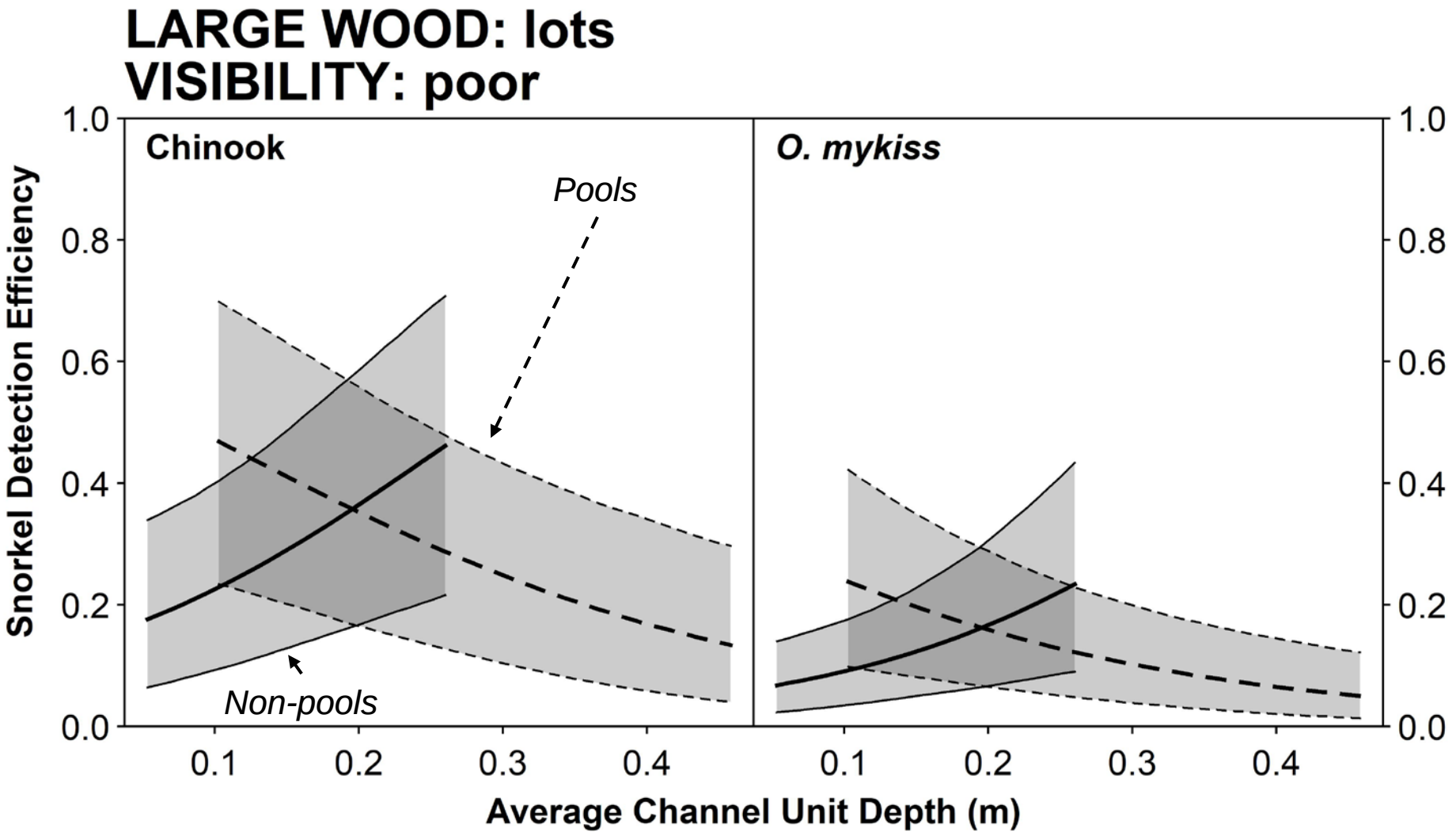
Results

*Effect sizes and direction: **typical** conditions*



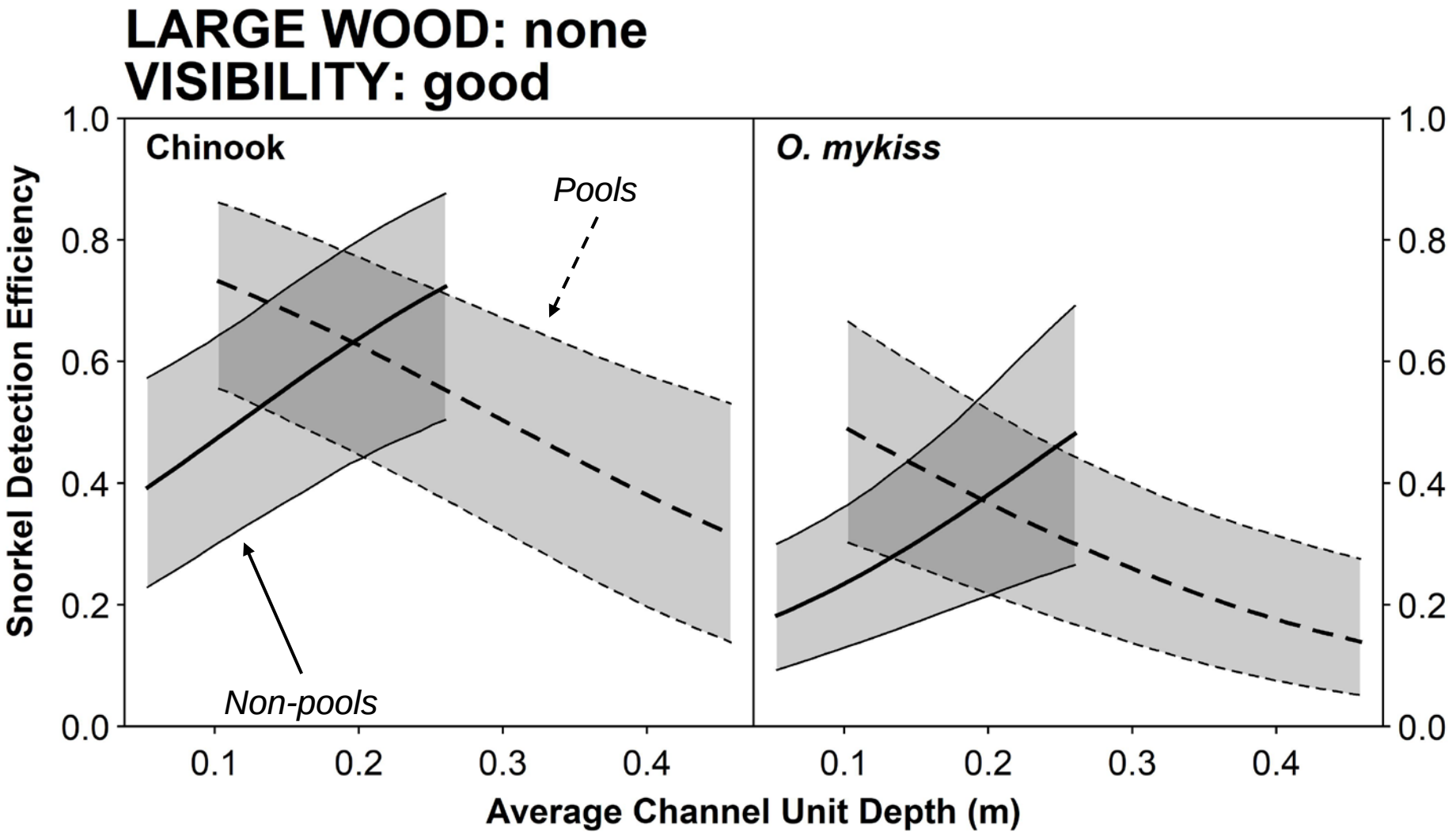
Results

*Effect sizes and direction: **poor** snorkel conditions*



Results

*Effect sizes and direction: **good** snorkel conditions*



Summary

Strengths and Weaknesses

WEAKNESSES

- Lots of random site variability
- Observer-specific effects? Year effects?
- Inference limited to channel units with similar characteristics
- Is count process truly binomial? (*independent, homogenous p*)

STRENGTHS

- Direct estimates of detection probability
- How it changes with local conditions
- Acknowledges uncertainty in mark-recapture sampling
- Output is automatically model-averaged (usage of ω_i terms)

NEXT STEPS: SIMULATION TRIALS

- Sensitive to violated assumptions?
- Necessary sample size?
- Incredibly uninformative mark-recapture data?

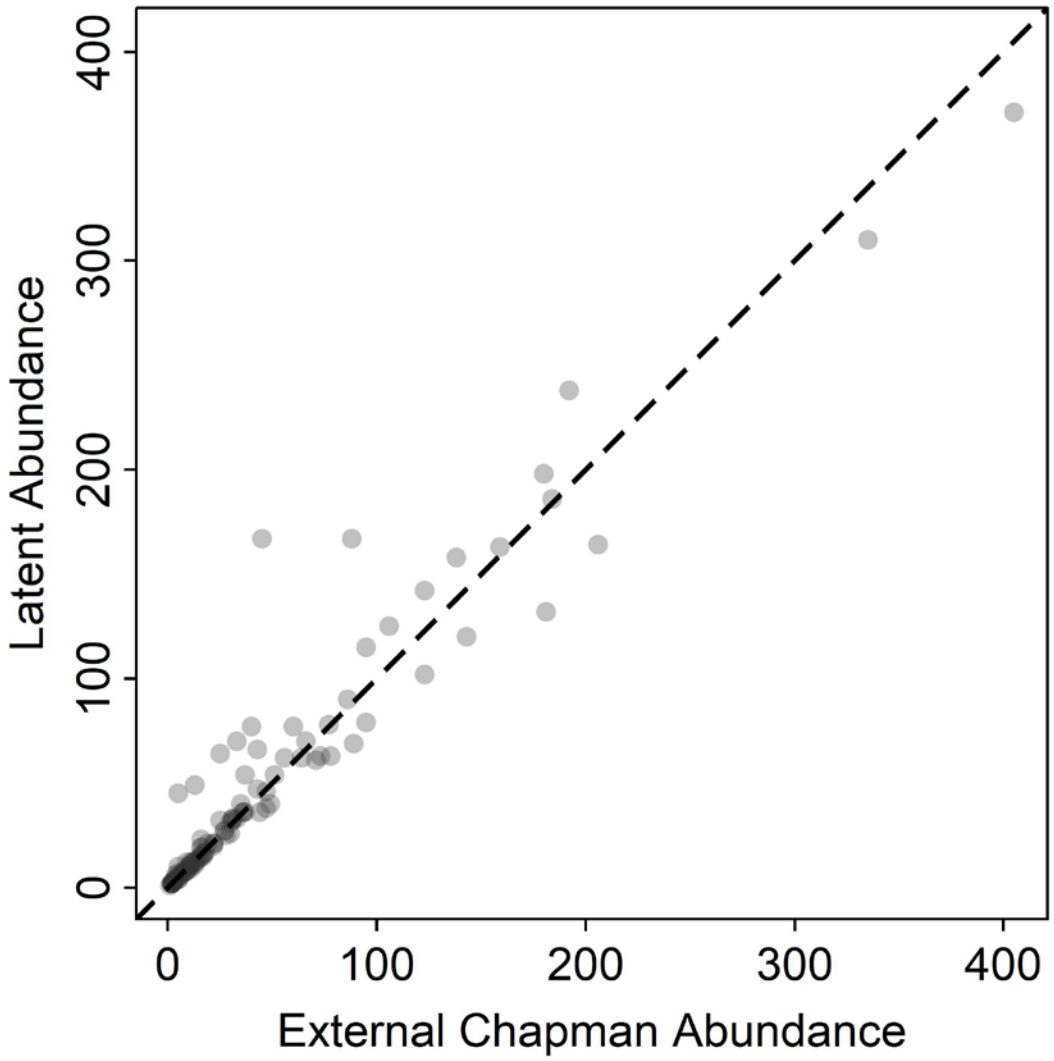
Questions?

Ben Staton
bstaton@critfc.org

BONUS SLIDES

Abundance Estimates

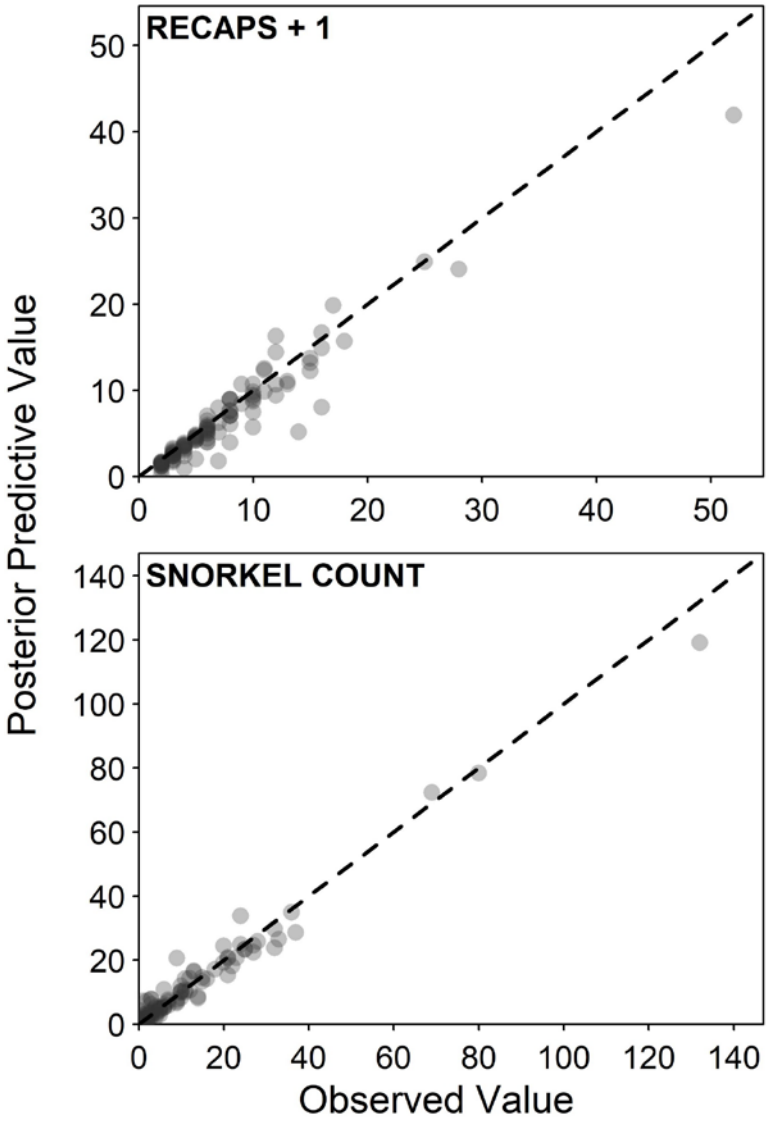
Hierarchical vs. external Chapman estimates

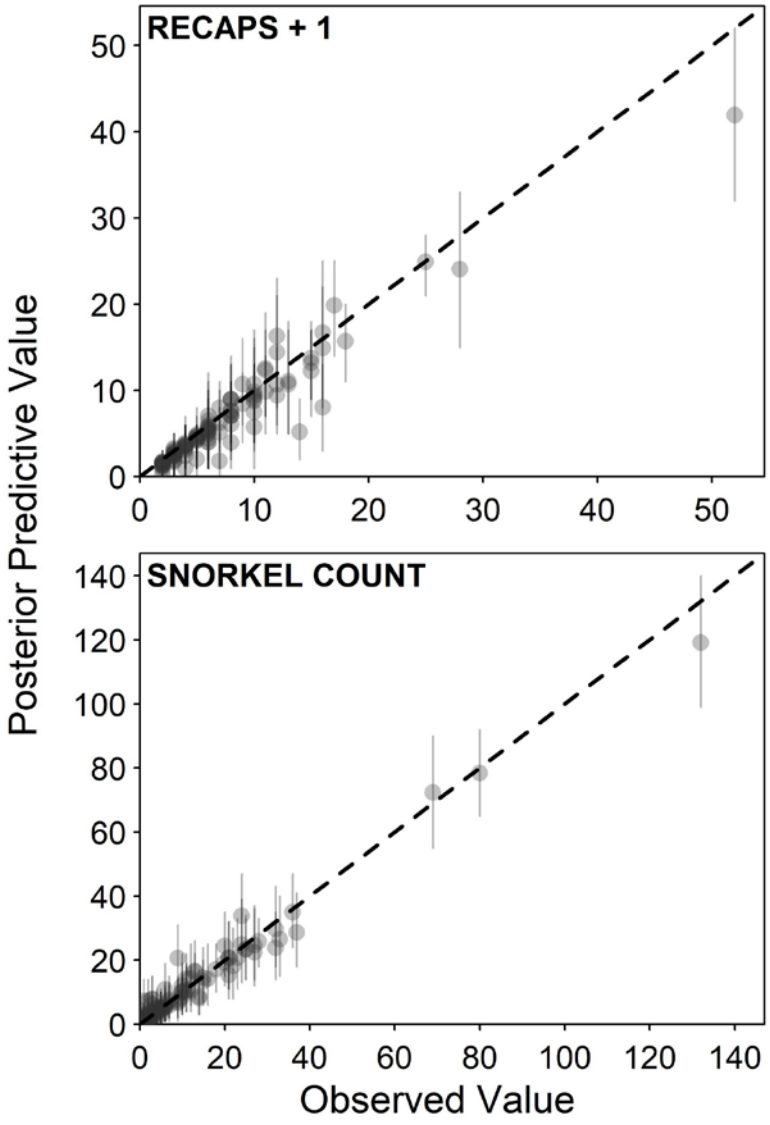


Hierarchical (latent) abundance estimate does not always equal the Chapman estimate

Model must satisfy **both** snorkel and mark-recap data

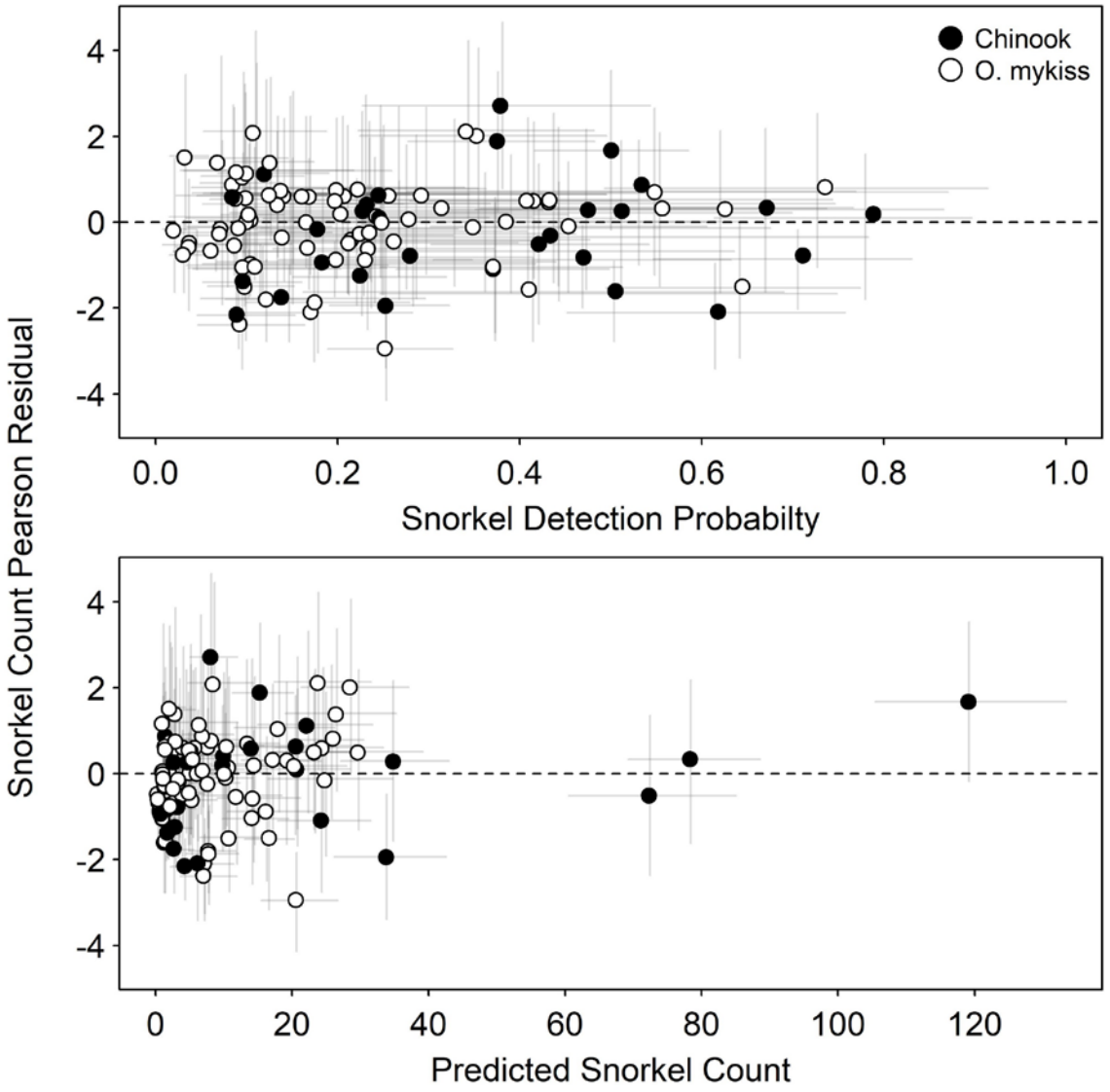
Fit to Data


RECAPS + 1
SNORKEL COUNT
Posterior Predictive Value
Observed Value
0
10
20
30
40
50
0
20
40
60
80
100
120
140


RECAPS + 1
SNORKEL COUNT
Posterior Predictive Value
Observed Value
0
10
20
30
40
50
0
20
40
60
80
100
120
140

Residual Inspection

No residual patterns in snorkel data



Pseudo JAGS Code

```
model {
  # PRIORS: ABUNDANCE
  for (i in 1:nobs) {
    Nu[i] ~ dunif(minN[i], maxN)
    N[i] <- round(Nu[i])
  }

  # PRIORS: LOGIT MODEL COEFFICIENTS
  a ~ dnorm(0, 0.001)
  b ~ dnorm(0, 0.001)
  w ~ dbern(0.5)

  # LIKELIHOOD
  for (i in 1:nobs) {
    # mark-recap data
    recaps1[i] ~ dhyper(marked[i]+1, N[i]-marked[i], K[i]+1, 1)

    # snorkel data
    snorkel[i] ~ dbin(p[i], N[i])
    logit(p[i]) <- a + b * w * cvt[i]
  }
}
```