Factors influencing the efficacy of GSI; examples using Washington, Oregon and Idaho stocks in the GAPS 2.1 database

Workshop: Current and Future Applications of GSI to Ocean Salmon Management

Portland, OR
May 15, 2007

Kenneth I. Warheit

Washington Department of FISH and WILDLIFE

This Talk is NOT . . . .

- Power analysis of the GAPS baseline
- Banks et al. (GAPS consortium)
  - Comprehensive analysis (entire baseline)
  - Mix-stock analysis and individual assignments
  - Multiple methods and procedures

This Talk is . . . .

- Exploring specific issues that will affect our ability to conduct genetic analyses on fishery samples
- Present questions that should be addressed by this workshop

Genetics and Fishery Management

- Population Definition
  - Delineation
  - Aggregation

- Population Assignment
  - Type
    - Mixed-stock analysis
    - Individual assignments
  - Methods
  - Criteria
    - Log-likelihood ratios
    - Posterior Probabilities
    - Cutoffs
  - Unassignment

Genetics and Fishery Management

- Population Definition
  - Delineation
  - Aggregation

- Population Assignment
  - Type
    - Mixed-stock analysis
    - Individual assignments
  - Methods
  - Criteria
    - Log-likelihood ratios
    - Posterior Probabilities
    - Cutoffs
  - Unassignment

Definitions

- Mix-stock Analysis
  - Genetic analysis of fishery samples to determine stock proportions. Individual fish are NOT assigned to stock
- Individual Assignment Analysis
  - Individual fish from a fishery sample are assigned to stock based on some criterion
  - Required if additional data are needed (e.g., cohort)

More Definitions

$$P(stock \mid genotype) = \left(\frac{P(genotype \mid stock) \cdot P(stock)}{P(genotype)} \right)$$

$P(stock \mid genotype)$ = Posterior Probability

$P(genotype \mid stock)$ = Likelihood
(calculated using Rannala and Mountain)

$P(stock)$ = Prior Probability

and More Definitions

(sort of)

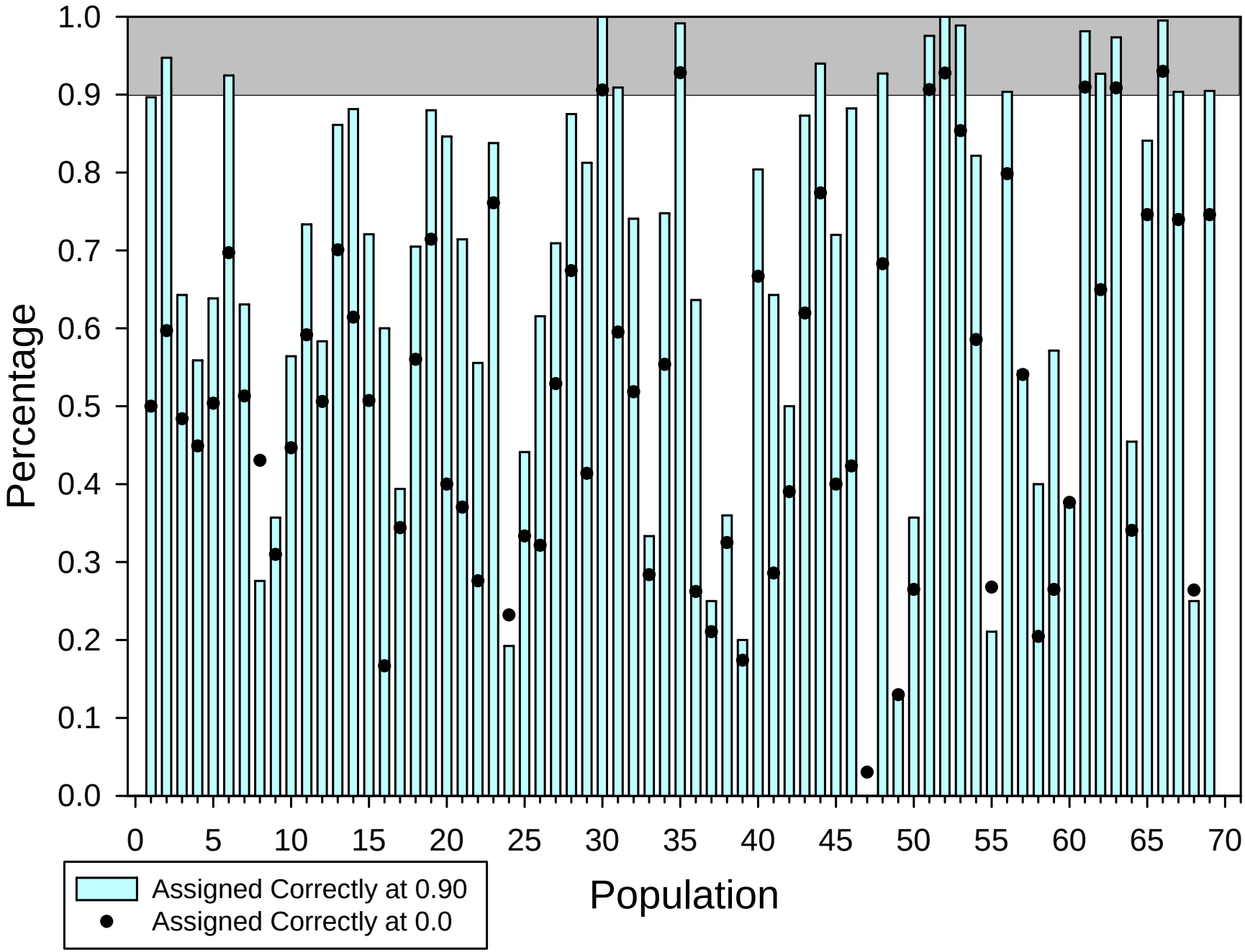
- GAPS Baseline v. 2.1
  - (Genetic Analysis of Pacific Salmonids
  - Coastwide Chinook database
  - 13 microsatellite loci
- Dataset (n=69 populations)
  - Washington
  - Idaho Snake River
  - Oregon Columbia River, Willamette, Coastal
- Jackknife (leave-one-out) analysis

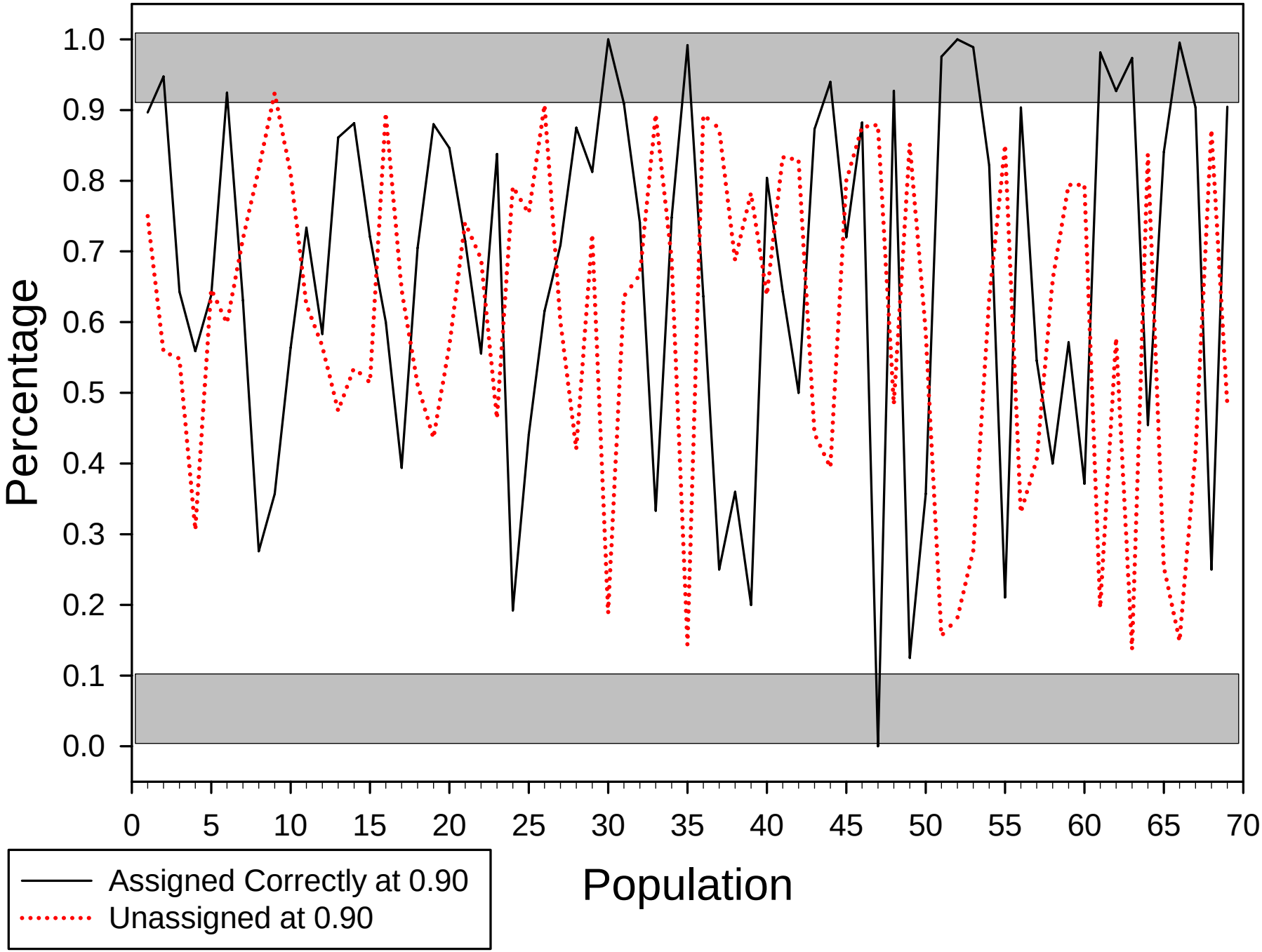
Posterior Probability Cutoffs

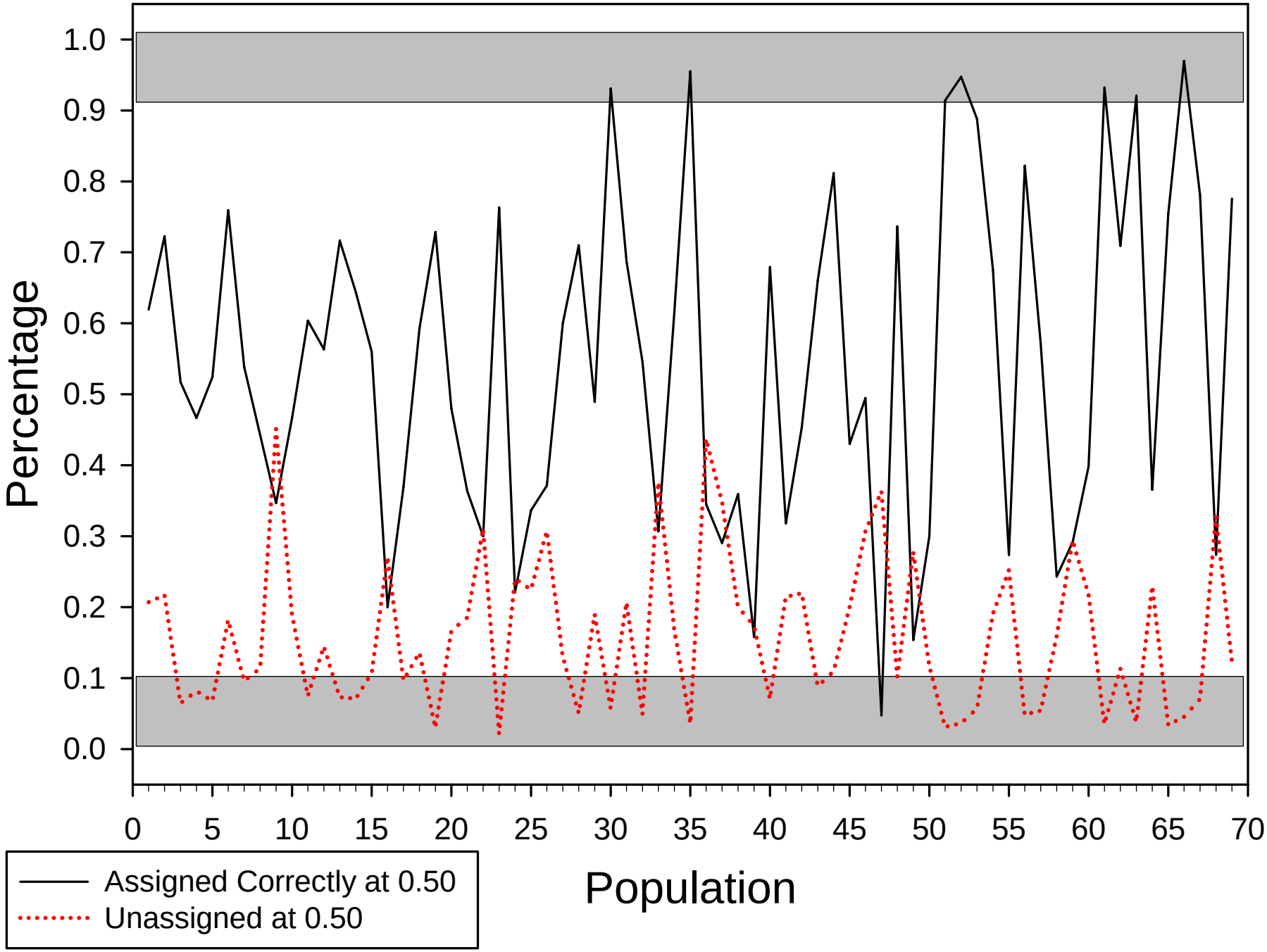
(when do we accept an assignment as being correct)

and

Unassigned Fish

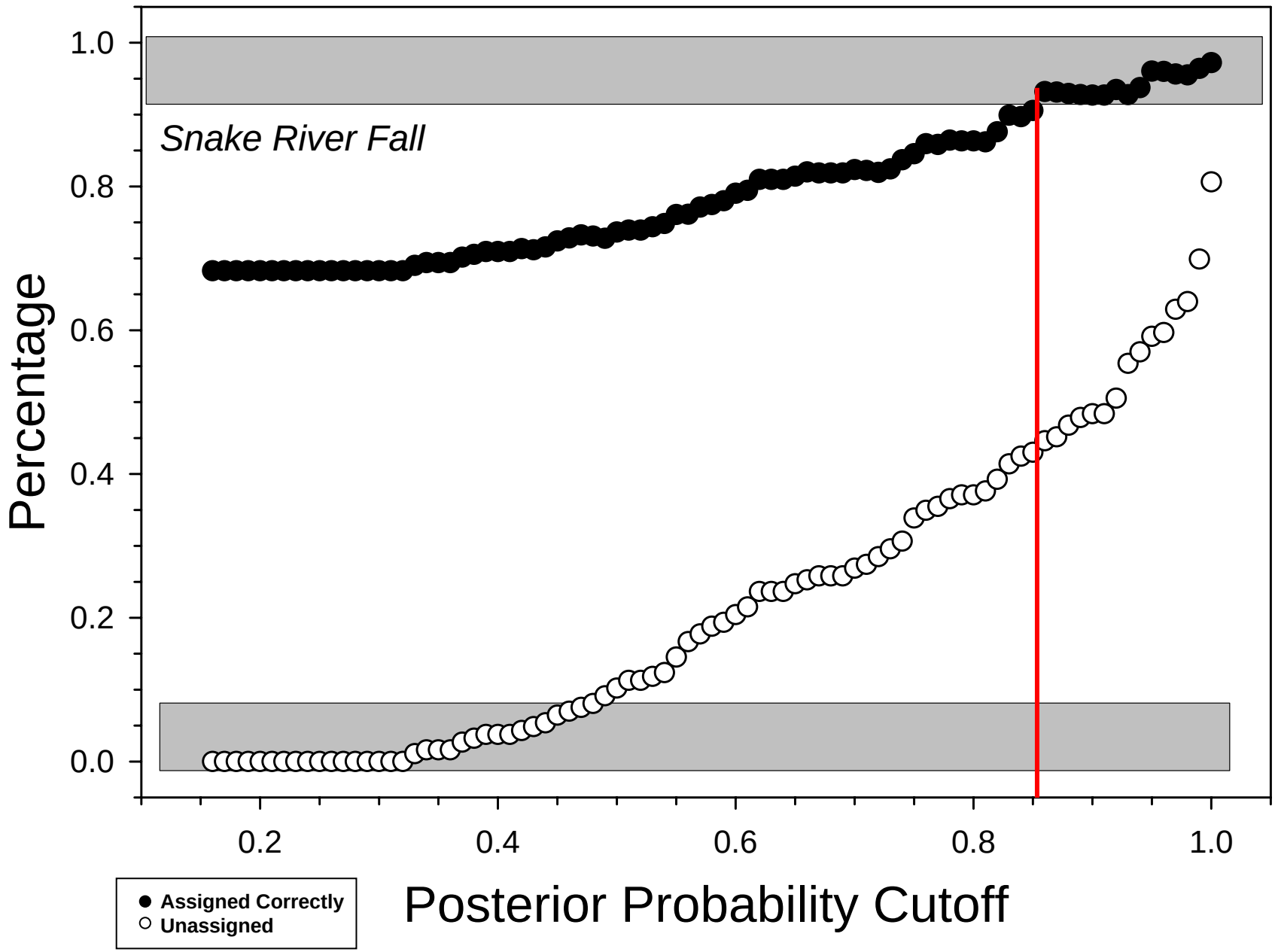

1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Percentage
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Population
Assigned Correctly at 0.90
Assigned Correctly at 0.0

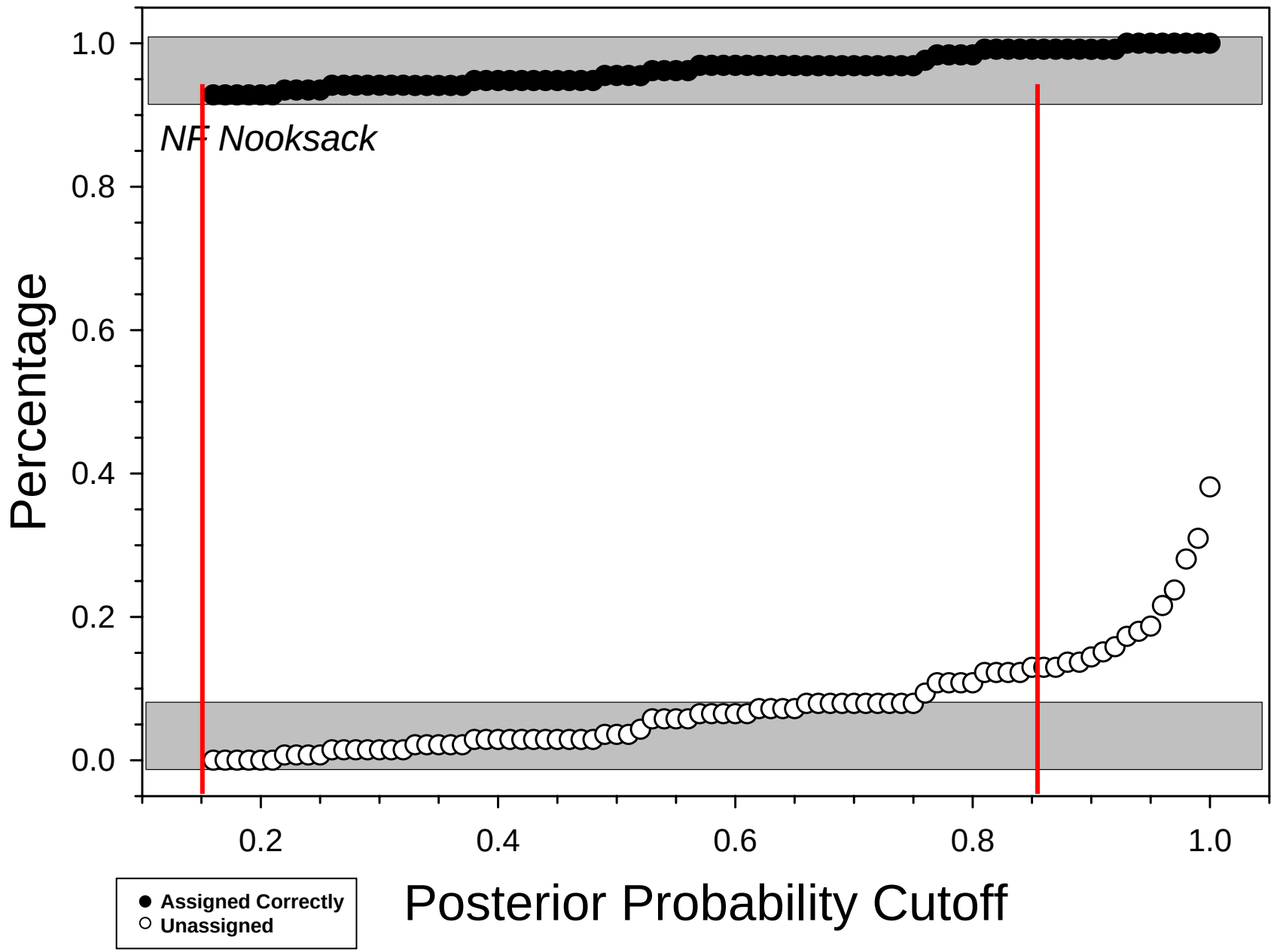

Percentage
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Population
Assigned Correctly at 0.90
Unassigned at 0.90

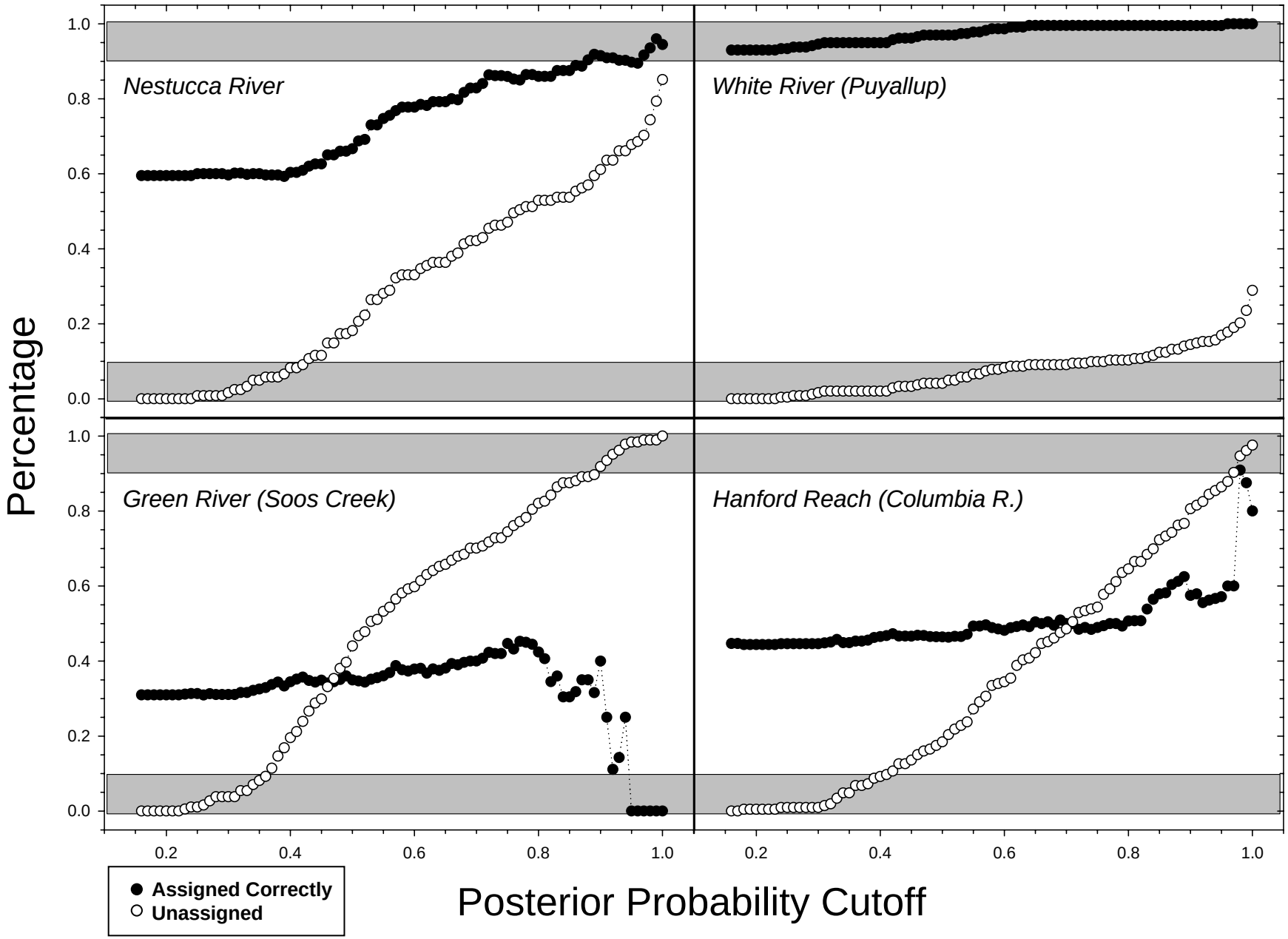

Percentage
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20
25
30
35
40
45
50
55
60
65
70
Population
Assigned Correctly at 0.50
Unassigned at 0.50

Snake River - Fall

Posterior Probability	Percentage	
	Assigned Correctly	Unassigned
0.00	0.68	0.00
0.25	0.68	0.00
0.50	0.74	0.10
0.75	0.85	0.34
0.90	0.93	0.48
0.95	0.96	0.59
1.00	0.97	0.81


Snake River Fall
Percentage
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
Posterior Probability Cutoff
Assigned Correctly
Unassigned

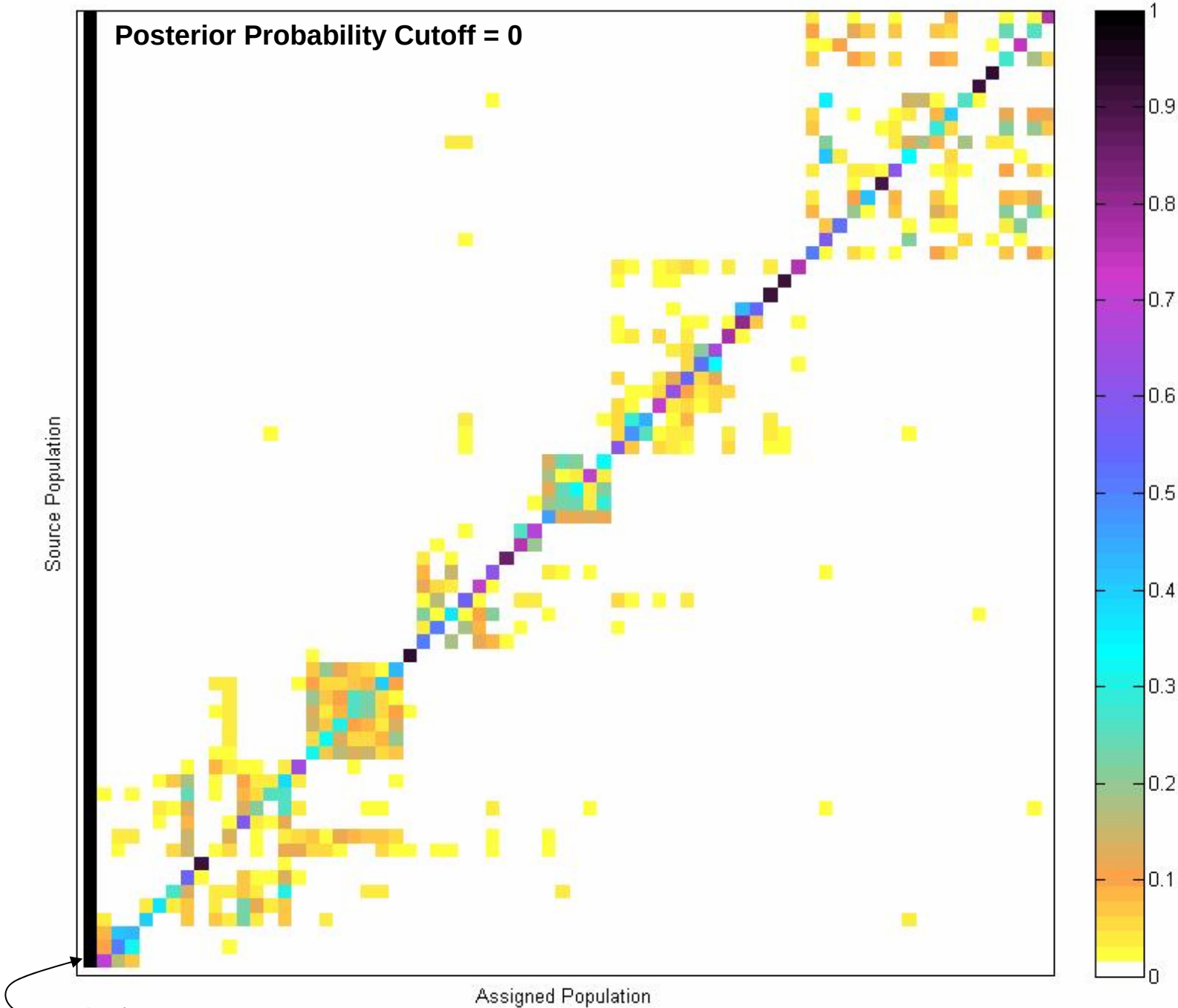

NF Nooksack
Percentage
1.0
0.8
0.6
0.4
0.2
0.0
0.2
0.4
0.6
0.8
1.0
Posterior Probability Cutoff
Assigned Correctly
Unassigned

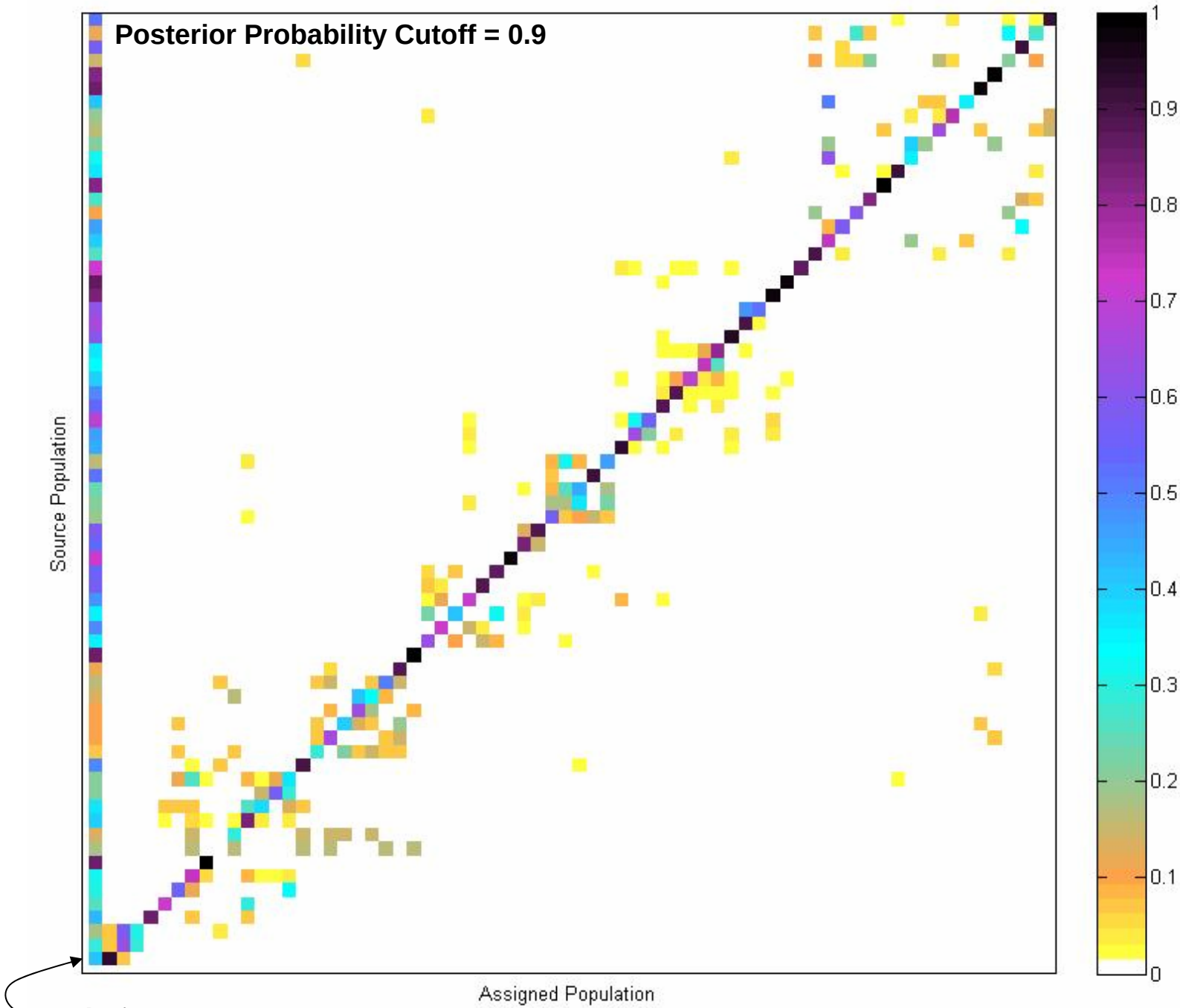

Nestucca River
White River (Puyallup)
Green River (Soos Creek)
Hanford Reach (Columbia R.)
Percentage
Posterior Probability Cutoff
0.0
0.2
0.4
0.6
0.8
1.0
Assigned Correctly
Unassigned

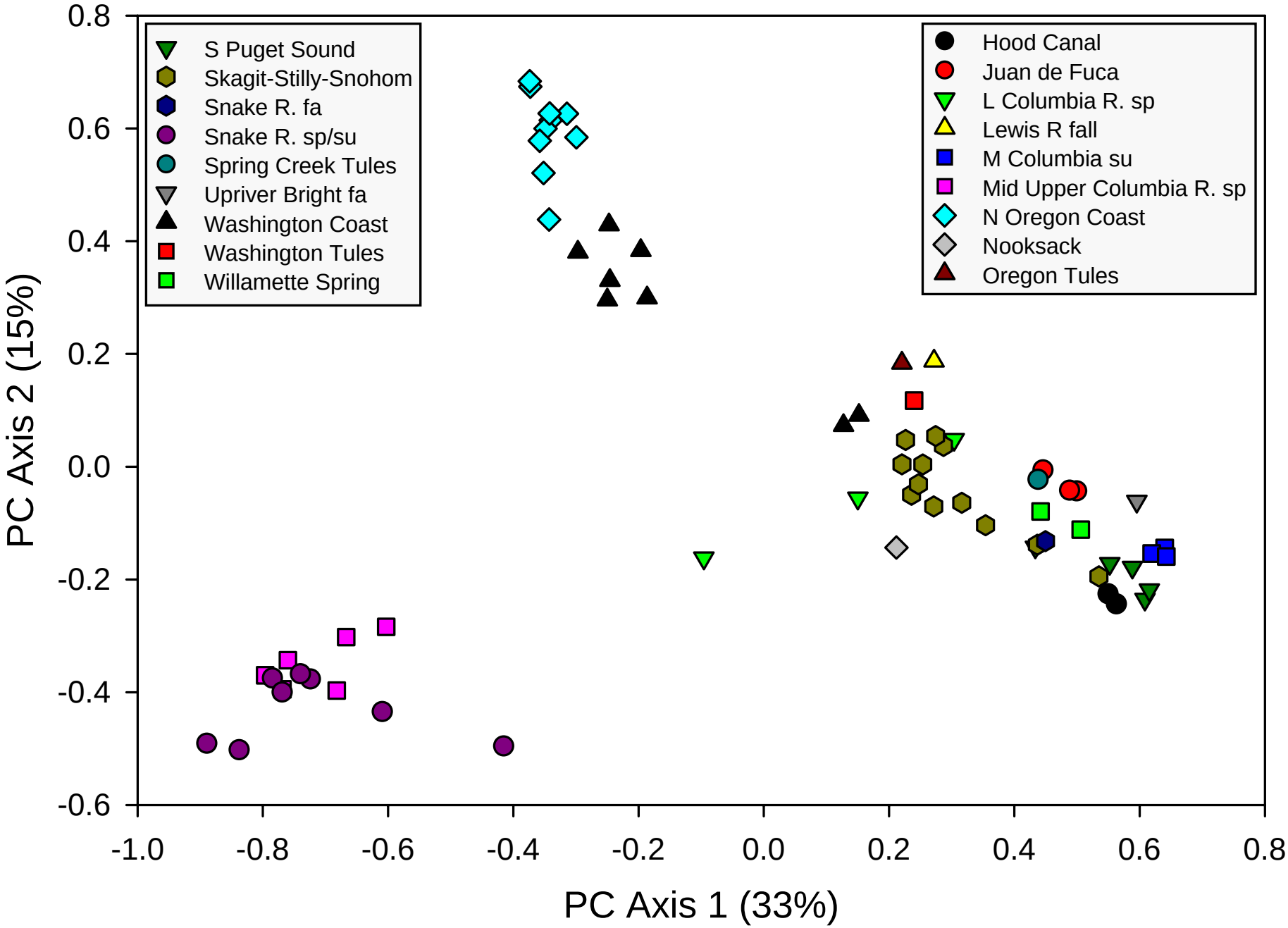
Summary

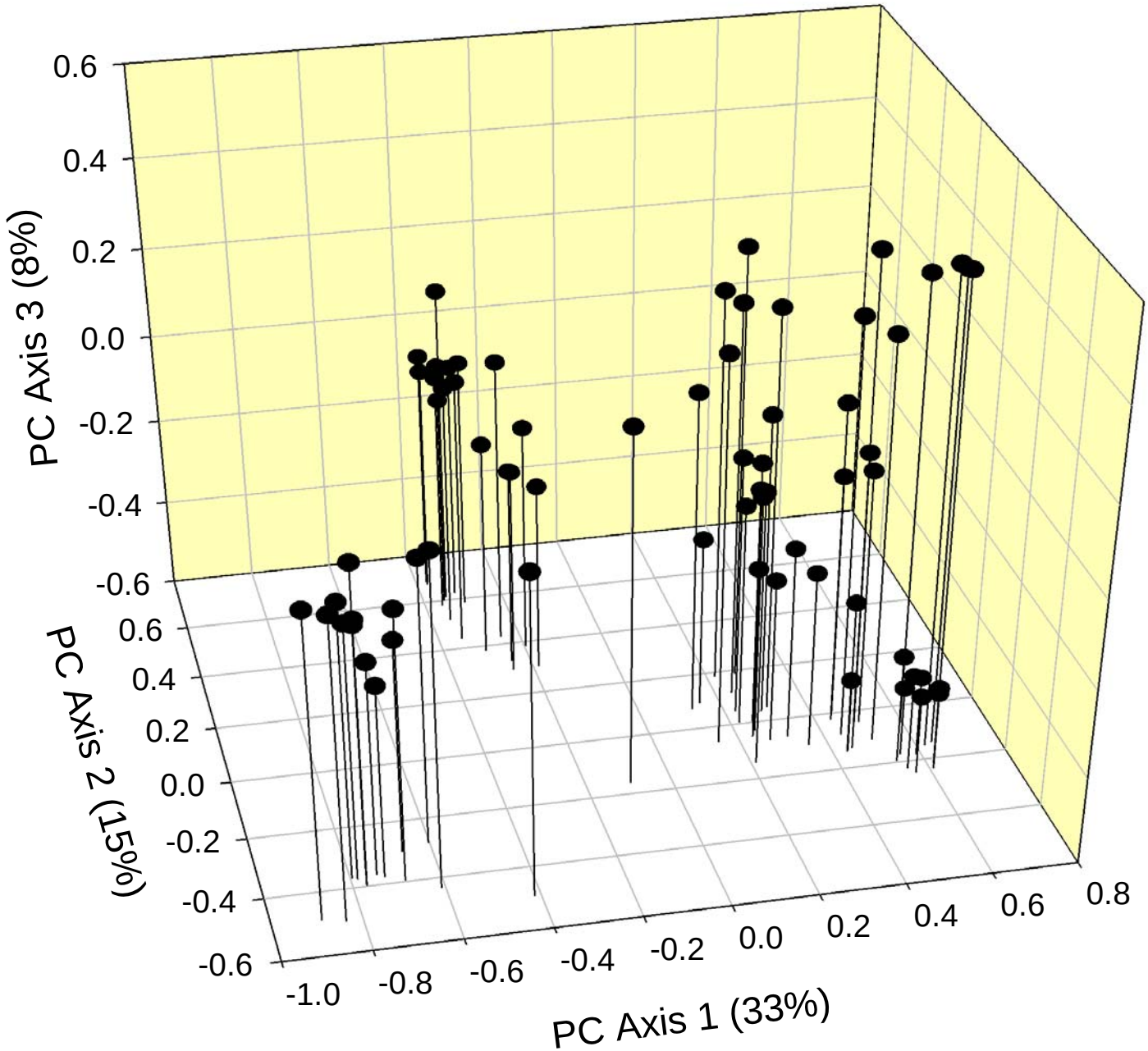
- Assignment error rates not equal
- Increasing stringency (higher posterior probability cutoff) will decrease error
- Increasing stringency will result in more unassigned individuals
- Unassigned rate not equal
- Stock proportions of unassigned fish are not equal to stock proportions of assigned fish

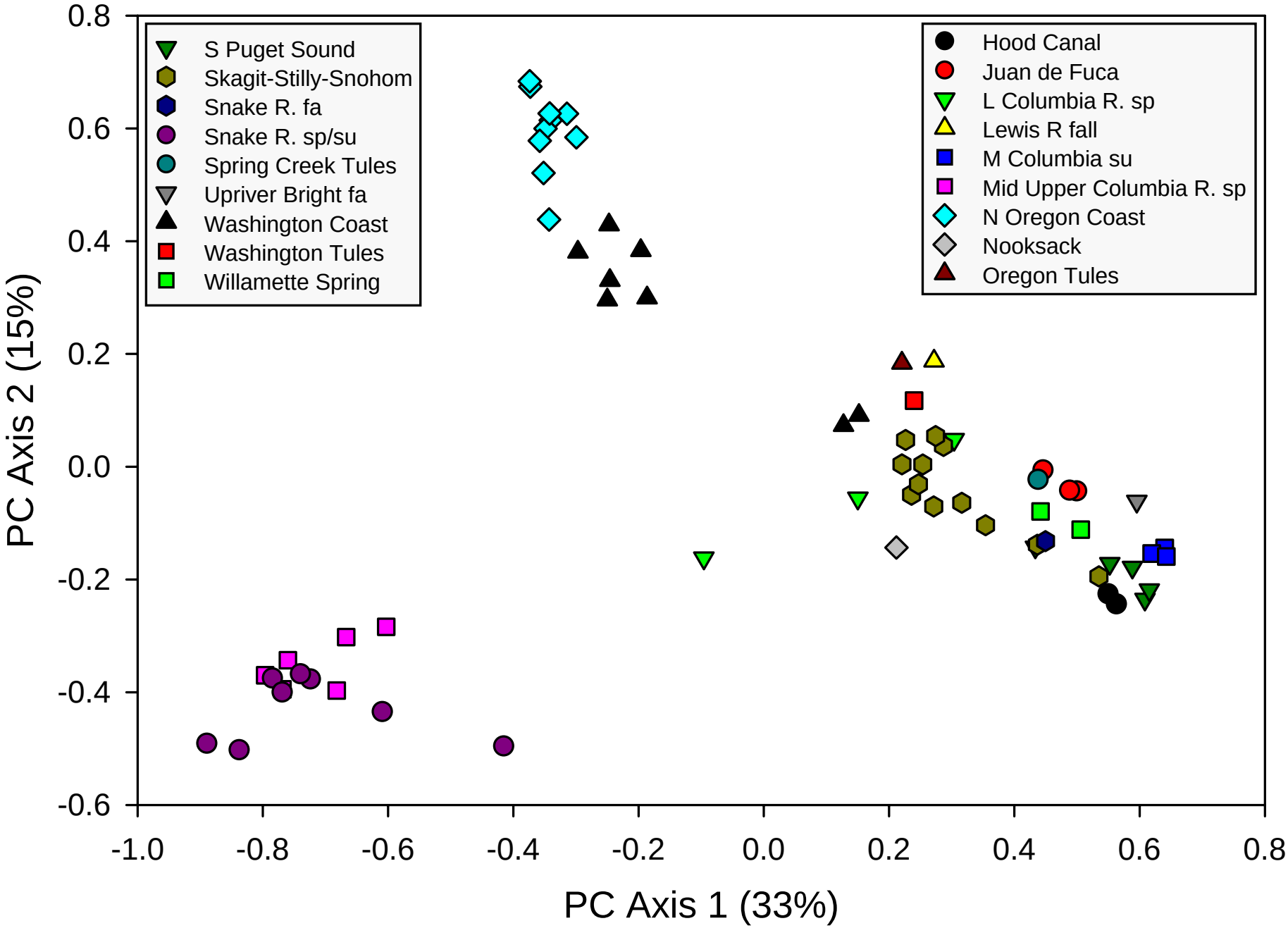
Aggregating Populations


Posterior Probability Cutoff = 0
Source Population
Assigned Population
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Assignment rate

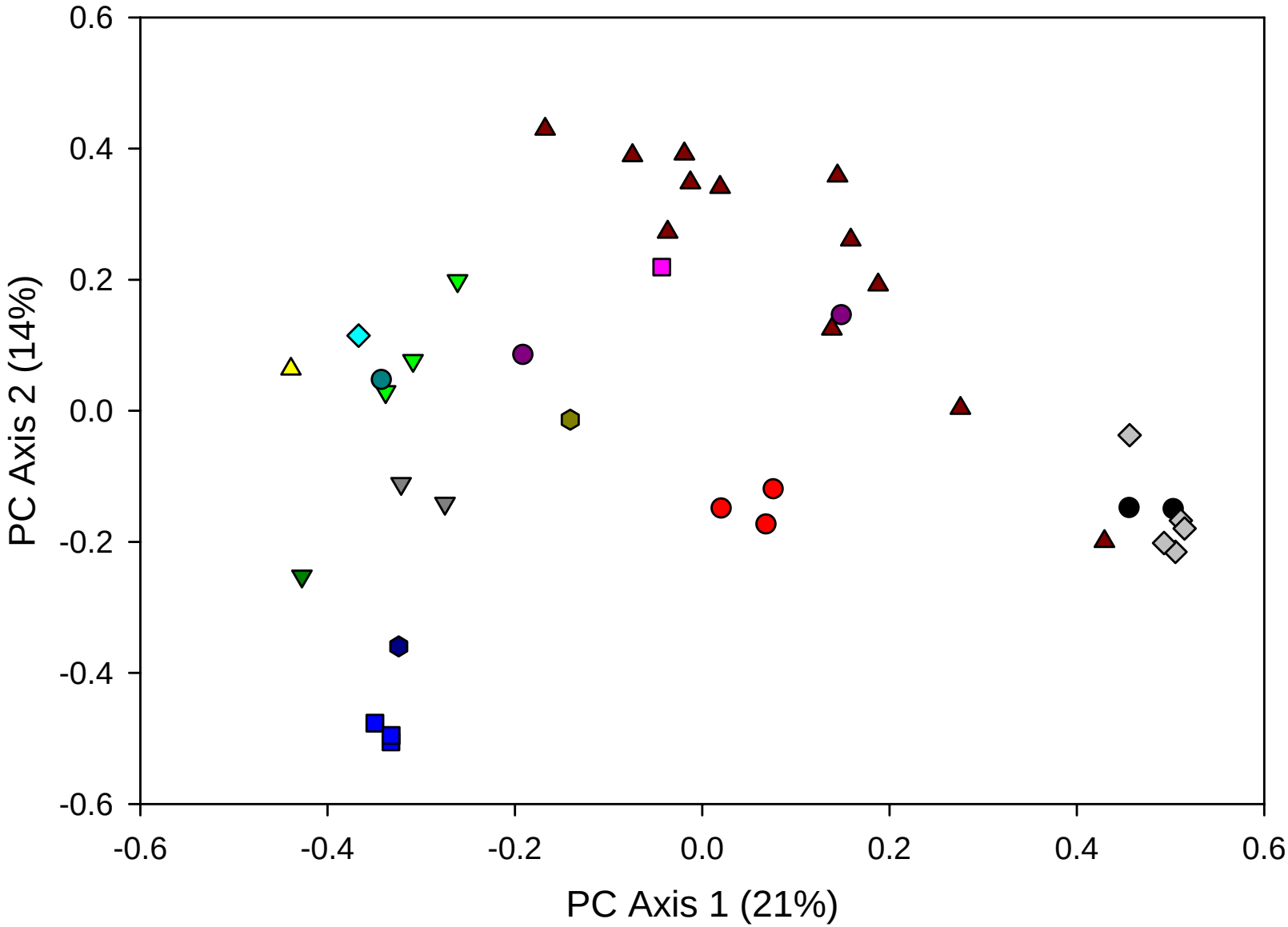

Posterior Probability Cutoff = 0.9
Source Population
Assigned Population
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Assignment rate


S Puget Sound
Skagit-Stilly-Snohom
Snake R. fa
Snake R. sp/su
Spring Creek Tules
Upriver Bright fa
Washington Coast
Washington Tules
Willamette Spring
Hood Canal
Juan de Fuca
L Columbia R. sp
Lewis R fall
M Columbia su
Mid Upper Columbia R. sp
N Oregon Coast
Nooksack
Oregon Tules
PC Axis 2 (15%)
PC Axis 1 (33%)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8

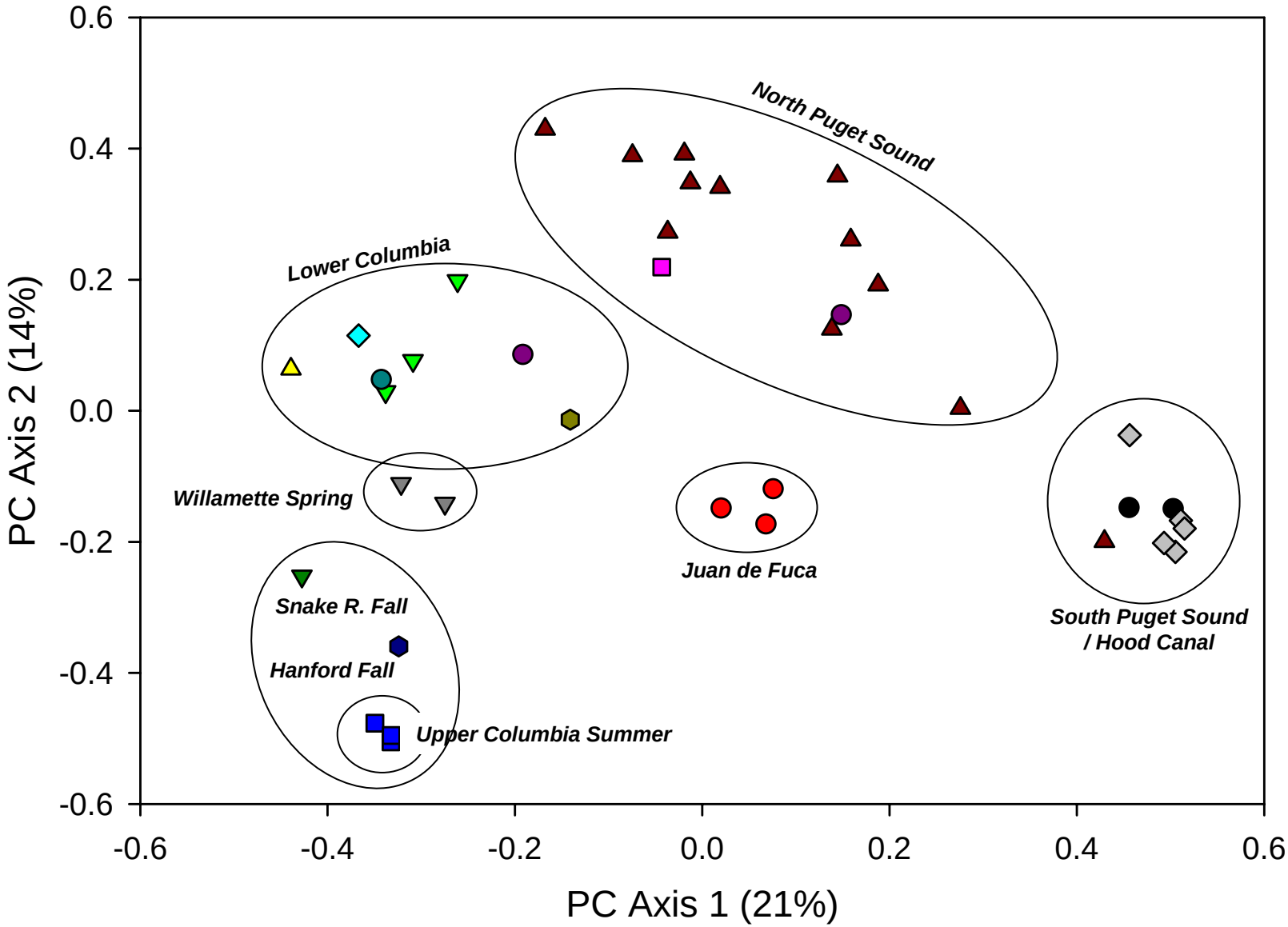

PC Axis 3 (8%)
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
PC Axis 2 (15%)
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
PC Axis 1 (33%)
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8

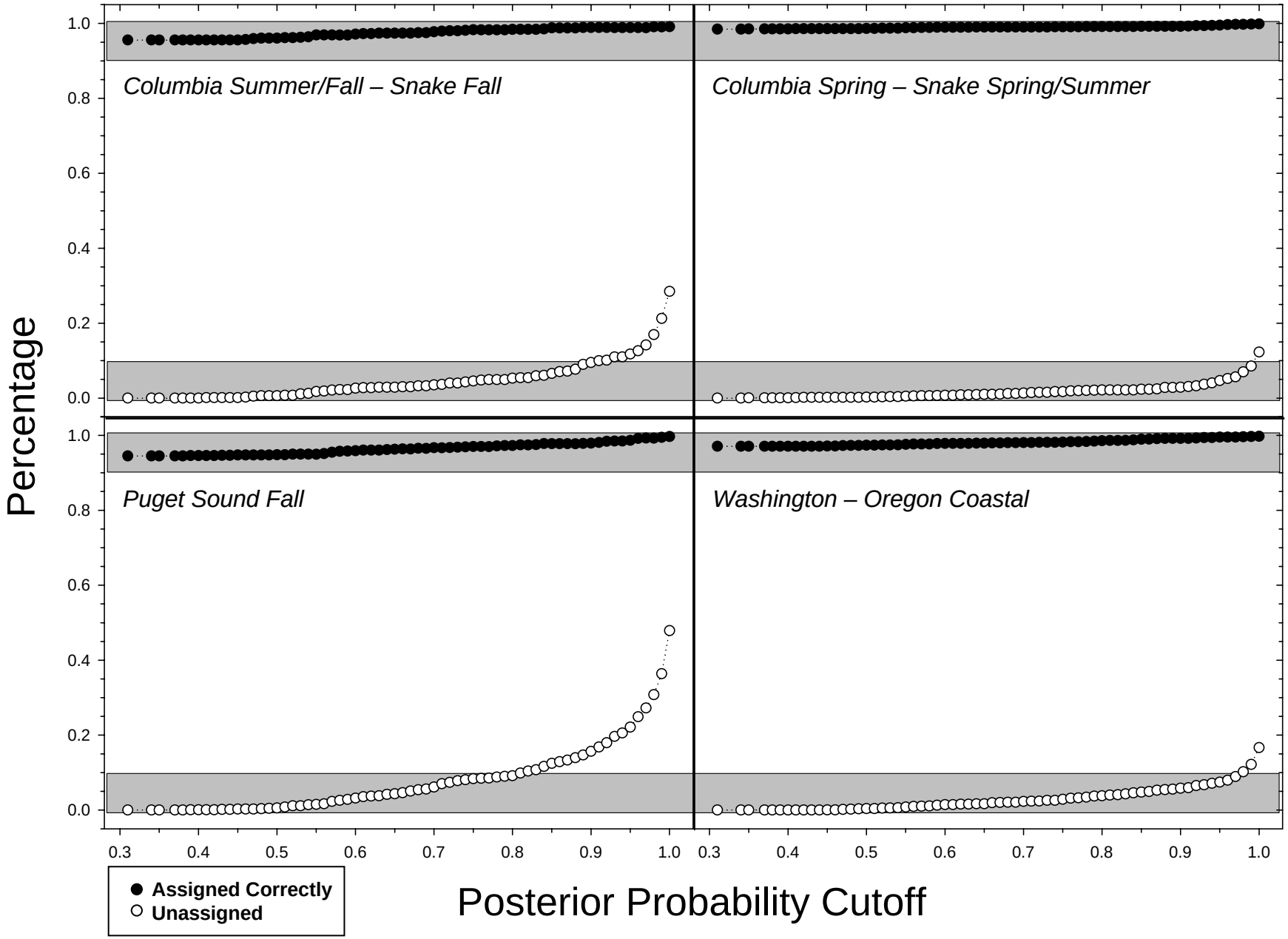

S Puget Sound
Skagit-Stilly-Snohom
Snake R. fa
Snake R. sp/su
Spring Creek Tules
Upriver Bright fa
Washington Coast
Washington Tules
Willamette Spring
Hood Canal
Juan de Fuca
L Columbia R. sp
Lewis R fall
M Columbia su
Mid Upper Columbia R. sp
N Oregon Coast
Nooksack
Oregon Tules
PC Axis 2 (15%)
PC Axis 1 (33%)
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
-0.6
-1.0
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
0.8

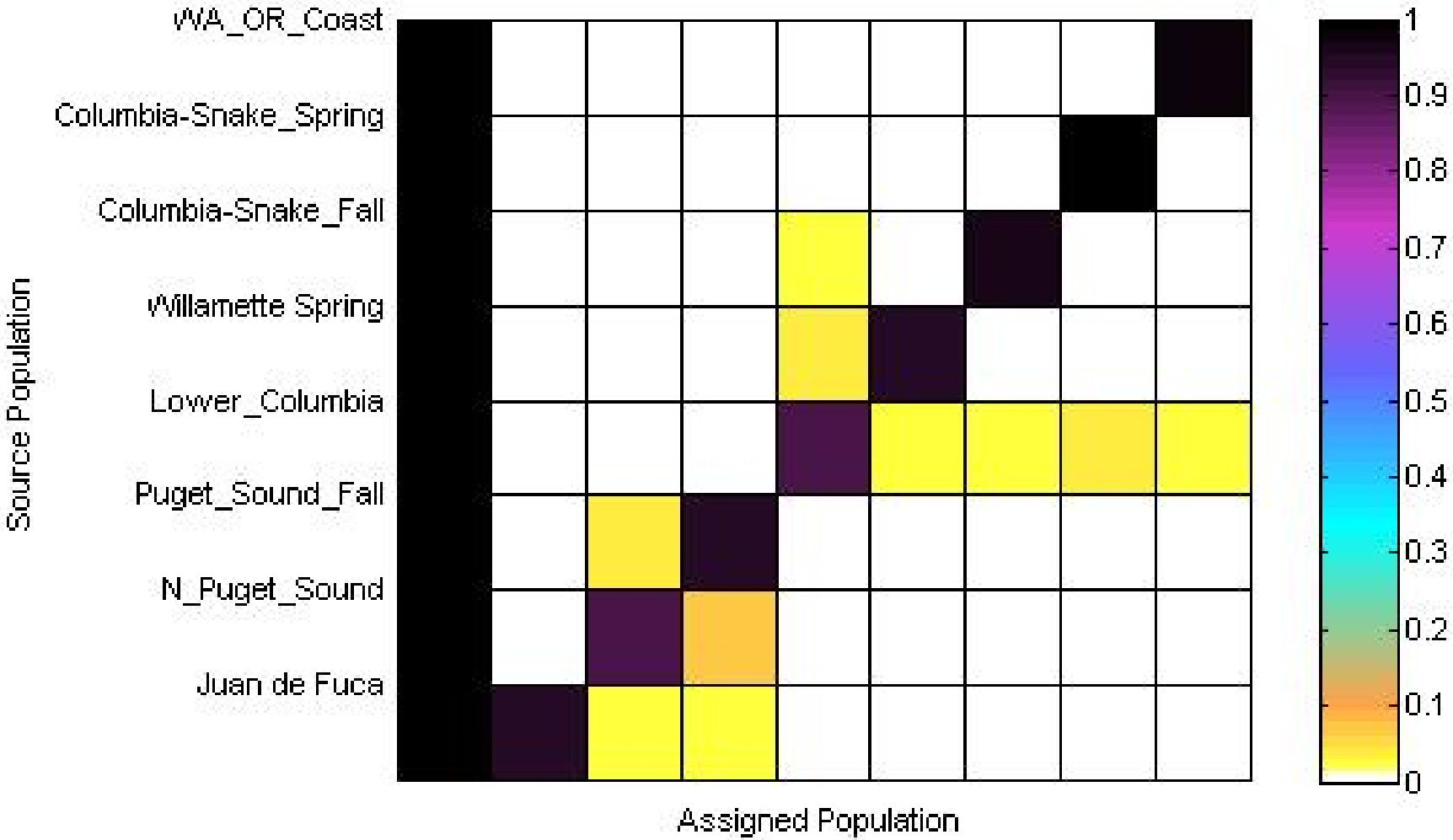
Subgroup Analysis

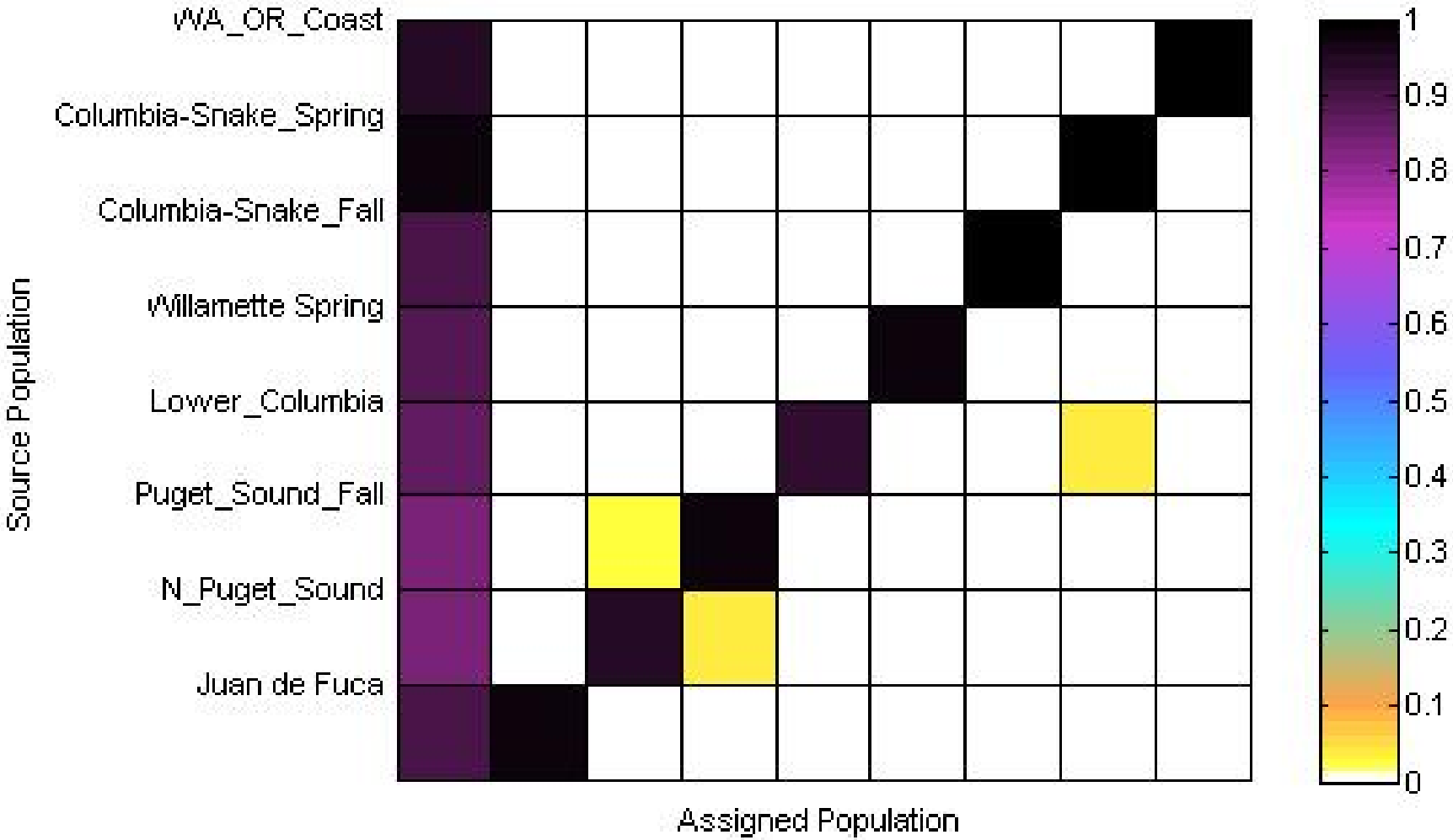


Subgroup Analysis




Percentage
Columbia Summer/Fall – Snake Fall
Columbia Spring – Snake Spring/Summer
Puget Sound Fall
Washington – Oregon Coastal
1.0
0.8
0.6
0.4
0.2
0.0
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Assigned Correctly
Unassigned
Posterior Probability Cutoff


WA_OR_Coast
Columbia-Snake_Spring
Columbia-Snake_Fall
Willamette Spring
Lower_Columbia
Puget_Sound_Fall
N_Puget_Sound
Juan de Fuca
Source Population
Assigned Population
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


WA_OR_Coast
Columbia-Snake_Spring
Columbia-Snake_Fall
Willamette Spring
Lower_Columbia
Puget_Sound_Fall
N_Puget_Sound
Juan de Fuca
Source Population
Assigned Population
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

Summary

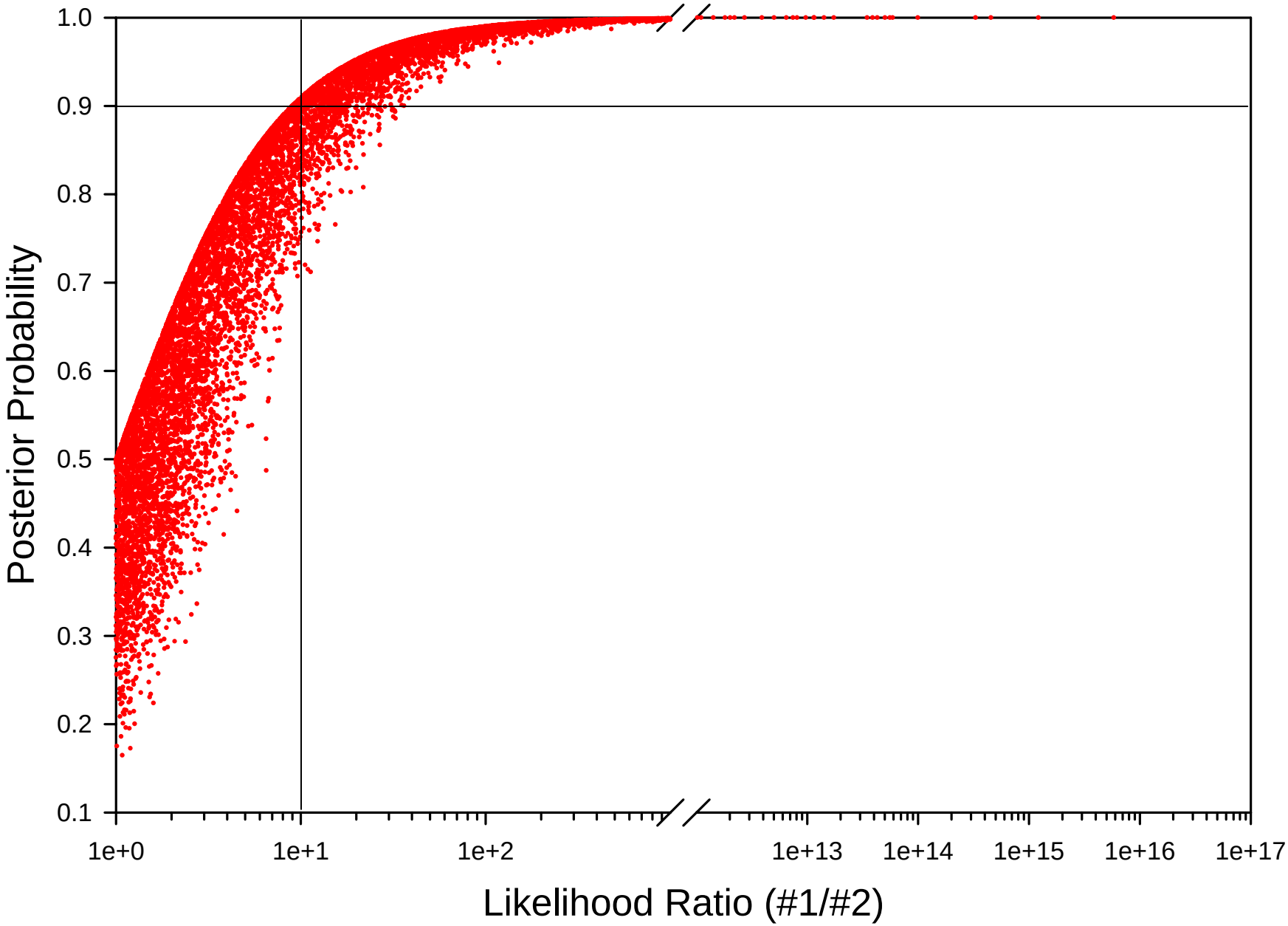
- Stocks are aggregated
- If stocks are aggregated based on geographic proximity and genetic similarity
- Assignment error rate is low
- Unassigned rate is low
- Can use a single, low posterior probability cutoff as assignment criterion

Questions

- Do we need to assign individuals to stocks?
- What method(s)?
- What criteria should we use to define "confidence"?
- How stringent a criterion?
- What do we do with unassigned fish?
- Should stocks be aggregated?
- How should we aggregate stocks?

Acknowledgements

- PSC and Workshop Steering Committee
- WDFW Molecular Genetics Lab, especially Denise Hawkins, Sewall Young
- Craig Busack, Annette Hoffman, Jim Scott, Brodie Cox
- Eric Anderson (NOAA – Santa Cruz)
- GAPS Labs
- Washington State General Fund


1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Posterior Probability
1e+0
1e+1
1e+2
1e+13
1e+14
1e+15
1e+16
1e+17
Likelihood Ratio (#1/#2)