

# Power Analysis to Determine Precision of Genetic Stock Identification of Chinook Smolts at Chandler

Yakima Basin Aquatic Science and Management Conference  
*Ellensburg, WA - June 13, 2007*

Kenneth I. Warheit and Craig Busack



# Summary: GSI Chandler Smolts

---

- *Microsatellite DNA*  
Differentiate Yakima Basin Chinook populations using microsatellite DNA
- *Population of origin of smolts*  
Identify population of origin of smolts passing Chandler
- *Smolt production*  
Use genetic data towards the quantification and comparison of smolt production in upper Yakima versus Naches Rivers
- *1500 smolt samples*  
Analyses based on 1500 smolt samples drawn randomly from a temporally stratified collection

# Questions

---

- Can we reduce sample size from 1500 to a smaller collection, without compromising results?

# The Model

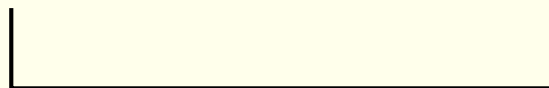
DATE	TOTAL WILD COUNTS	TOTAL WILD PASSAGE	TOTAL DNA SAMPLES	TOTAL DNA SAMPLES ANALYZED
01/22/06	97	647	15	7
01/23/06	153	931	15	10
01/24/06	219	1183	15	13
01/25/06	154	802	15	9
01/26/06	120	597	15	7
01/27/06	22	109	15	1
01/28/06	32	345	15	4
01/29/06	10	52	10	1
01/30/06	17	76	15	1
01/31/06	11	53	15	2
⋮				
06/30/06	729	1884	50	21
07/01/06	195	565	50	6
07/02/06	89	289	5	3
07/03/06	80	273	5	5
07/04/06	281	1398	5	5
07/05/06	165	816	5	5
07/06/06	49	220	5	5
07/07/06	8	33	5	5
07/08/06	292	899	50	19
07/09/06	269	900	5	5
<b>TOTALS</b>	<b>24270</b>	<b>135891</b>	<b>2480</b>	<b>1500</b>

Repeat using N =

1250	500
1000	400
900	300
800	200
700	100
600	

**35,500**    **58,000**

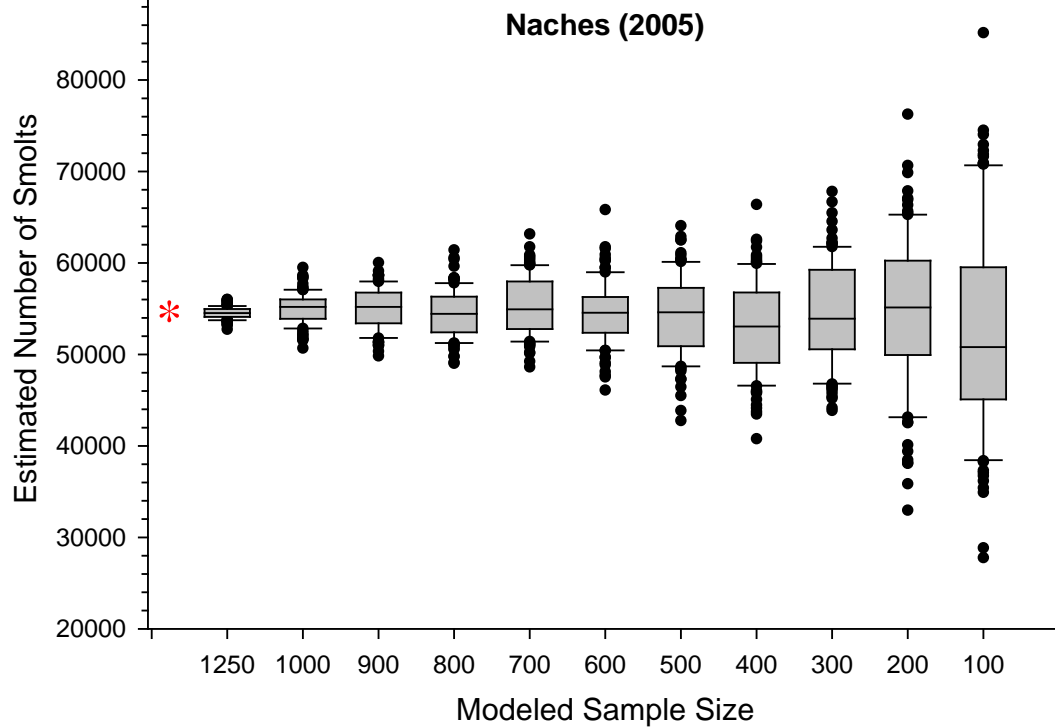
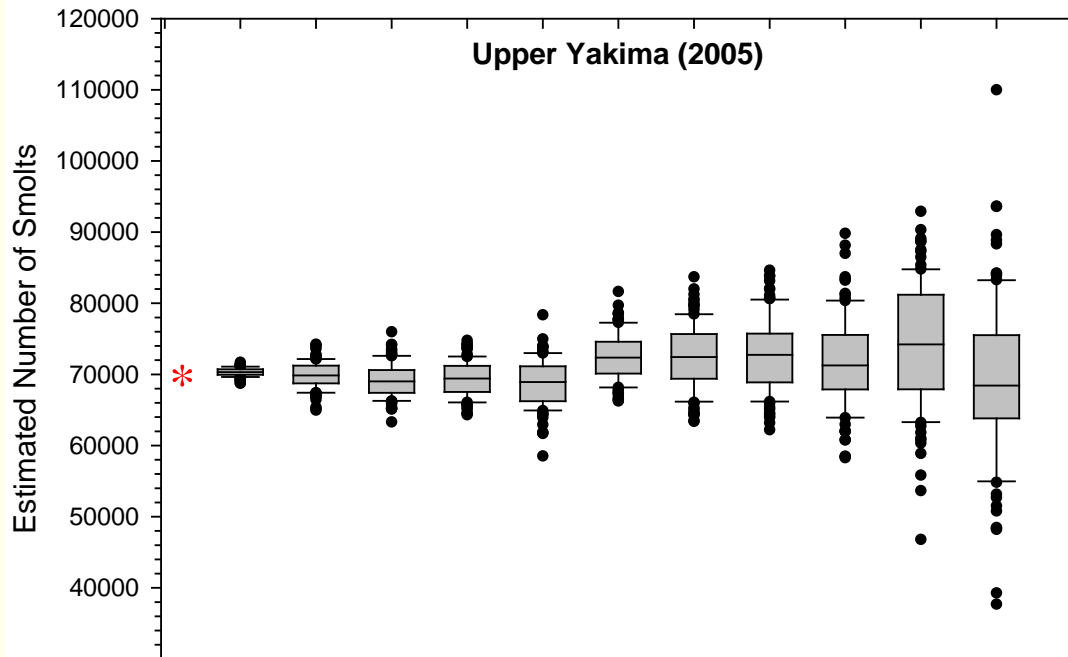
	<u>Naches</u>	<u>uYakimaSp</u>
<b>Jan-Feb</b>	36.1%	42.3%
<b>March</b>	8.4%	71.8%
<b>April</b>	29.2%	27.4%
<b>May</b>	3.7%	4.4%
<b>Jun-July</b>	4.6%	19.3%



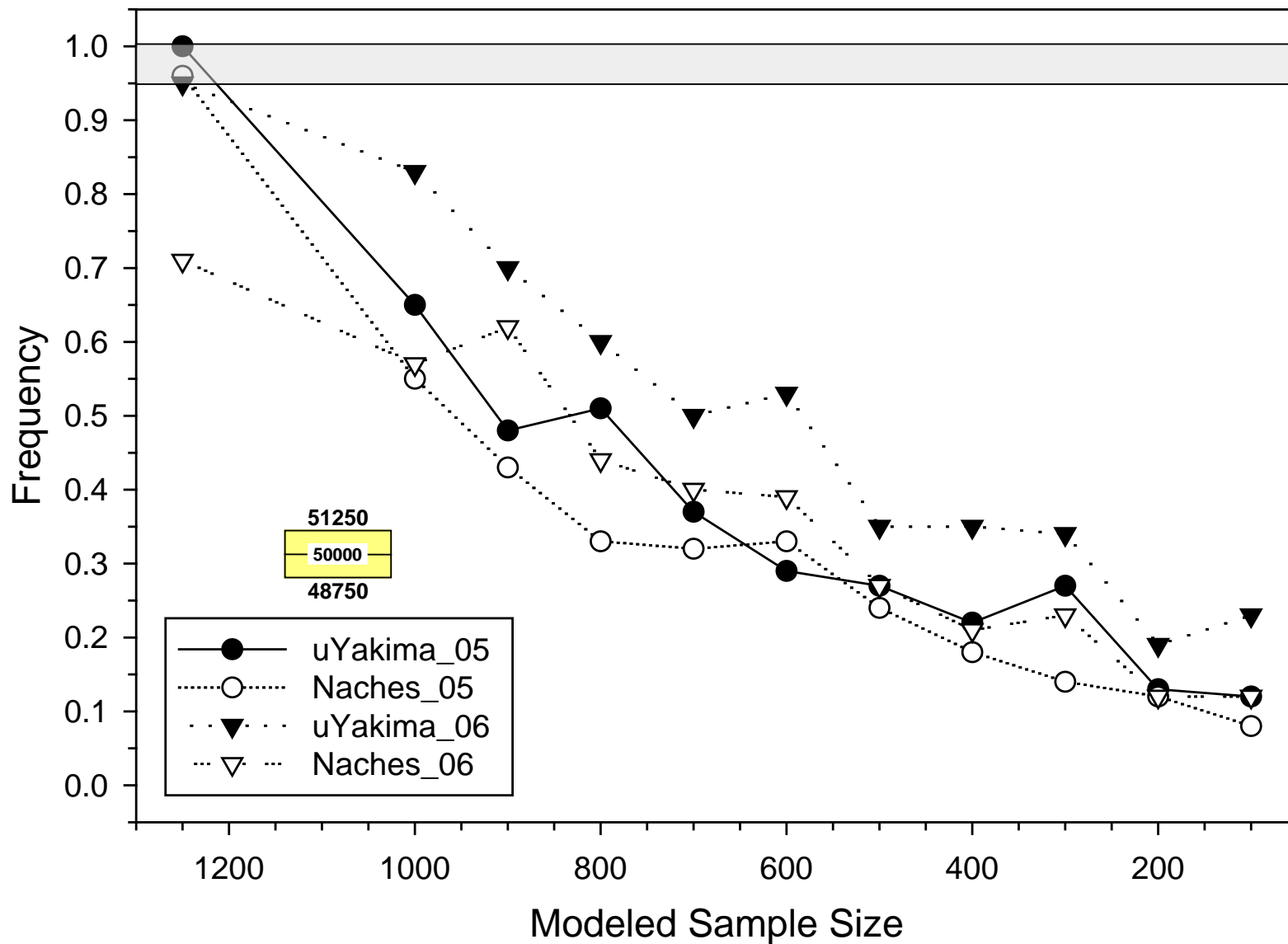
# The Model

---

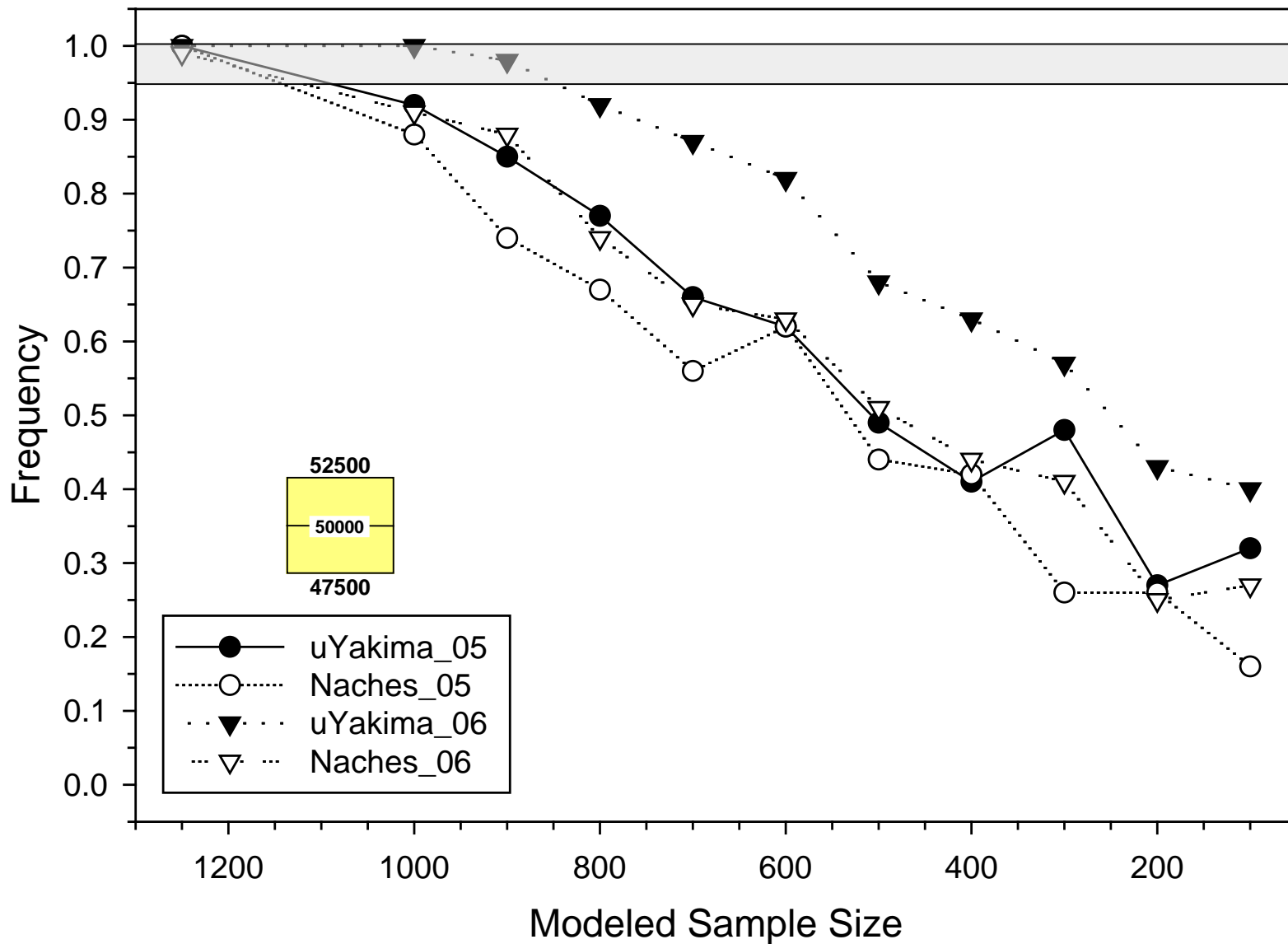
- *Repeat 100 times*  
For each modeled sample size
- *Separate analyses*  
2005 and 2006 data
- *Therefore*  
For each sample size 100 estimated number of smolts for Naches and upper Yakima – for both 2005 and 2006



# Within 5% of 1500 Result ( $\pm 2.5\%$ )

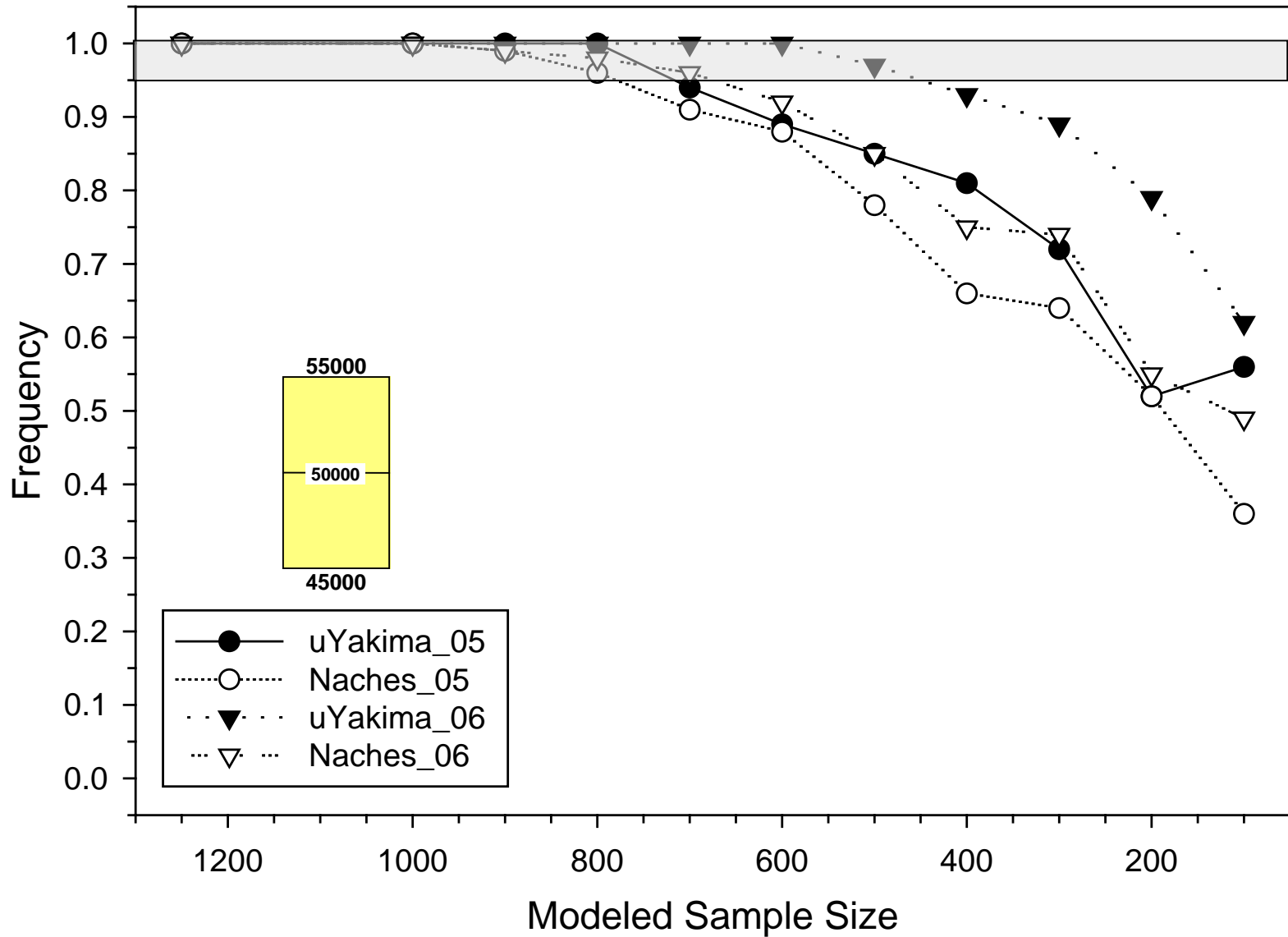


# Within 10% of 1500 Result ( $\pm 5.0\%$ )

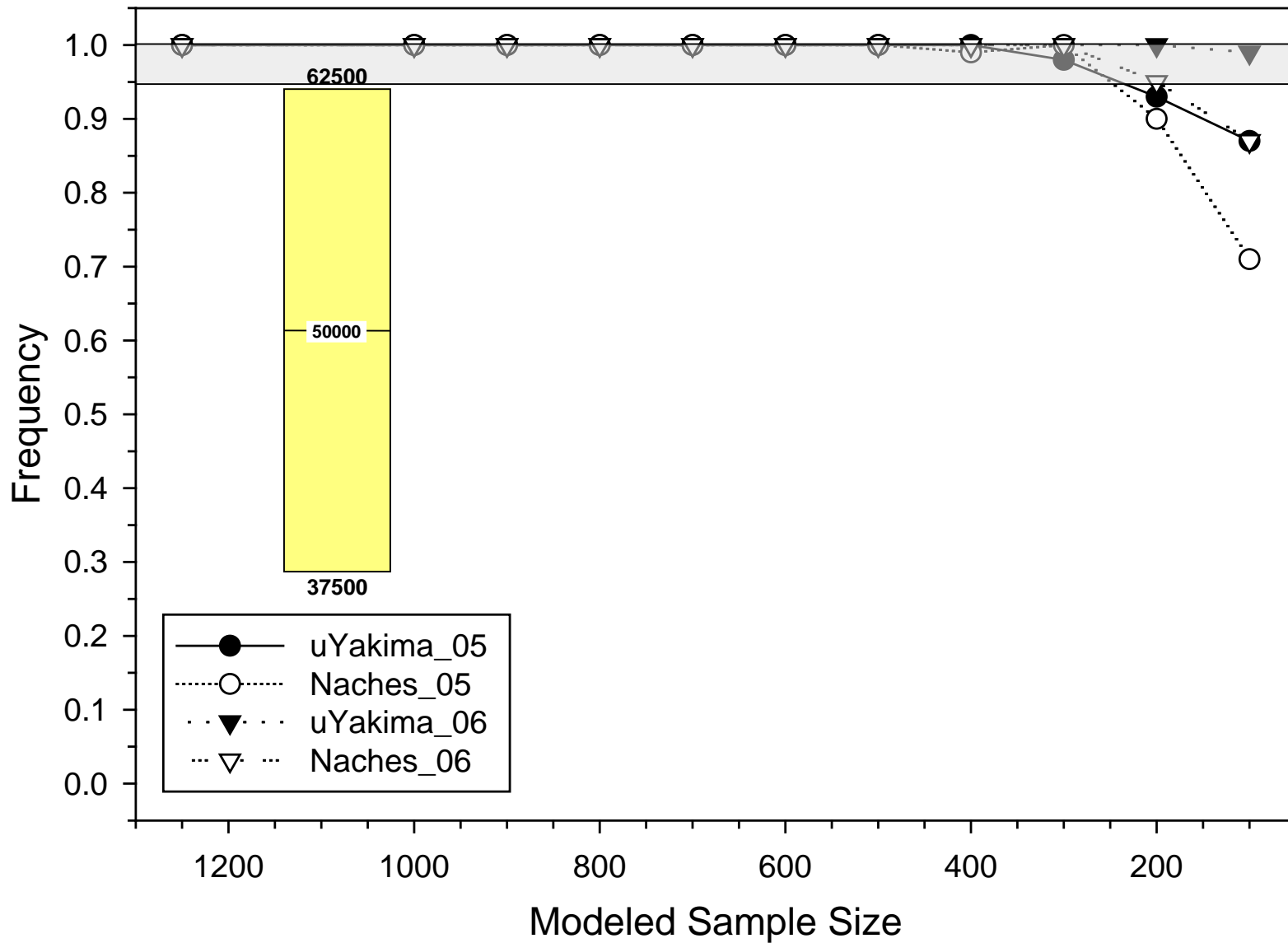




# Within 20% of 1500 Result ( $\pm 10.0\%$ )



# Within 50% of 1500 Result ( $\pm 25.0\%$ )



# Questions

---

- Is 1500 sample appropriate, or should we be sampling more?

2005: 331,821 smolts (0.45%)
2006: 135,891 smolts (1.10%)

# The Model

DATE	TOTAL WILD COUNTS	TOTAL WILD PASSAGE
01/22/06	97	647
01/23/06	153	931
01/24/06	219	1183
01/25/06	154	802
01/26/06	120	597
01/27/06	22	109
01/28/06	32	345
01/29/06	10	52
01/30/06	17	76
01/31/06	11	53
⋮		
06/30/06	729	1884
07/01/06	195	565
07/02/06	89	289
07/03/06	80	273
07/04/06	281	1398
07/05/06	165	816
07/06/06	49	220
07/07/06	8	33
07/08/06	292	899
07/09/06	269	900
<b>TOTALS</b>	<b>24270</b>	<b>135891</b>

All individuals assigned

Randomly select N =

100,000  
50,000  
25,000  
10,000  
5,000  
2,500  
**1,500**

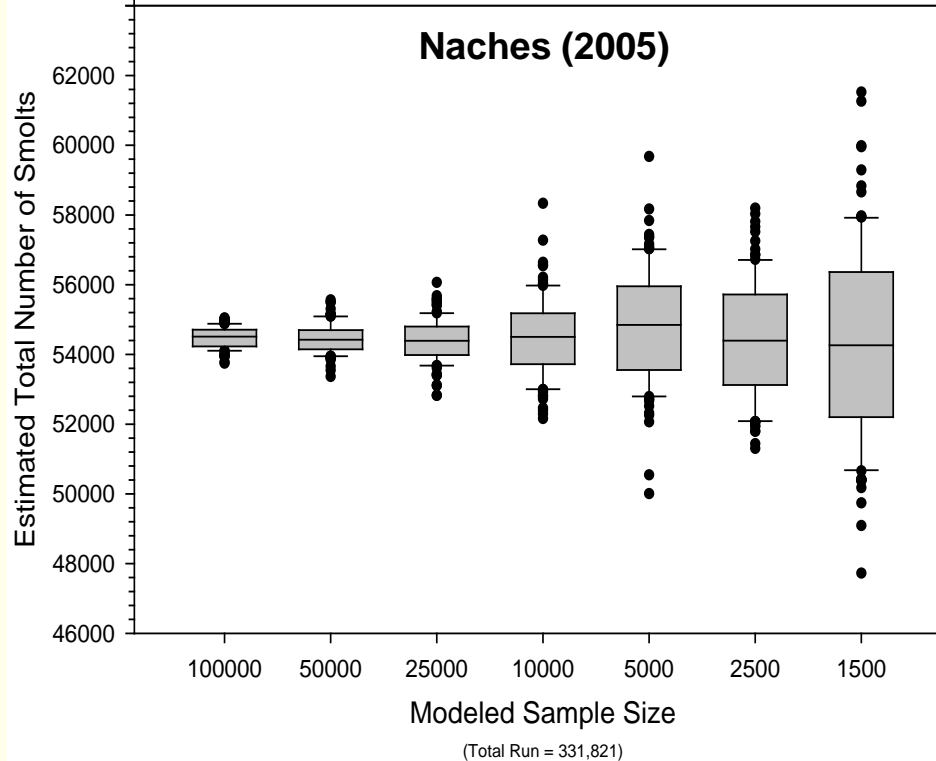
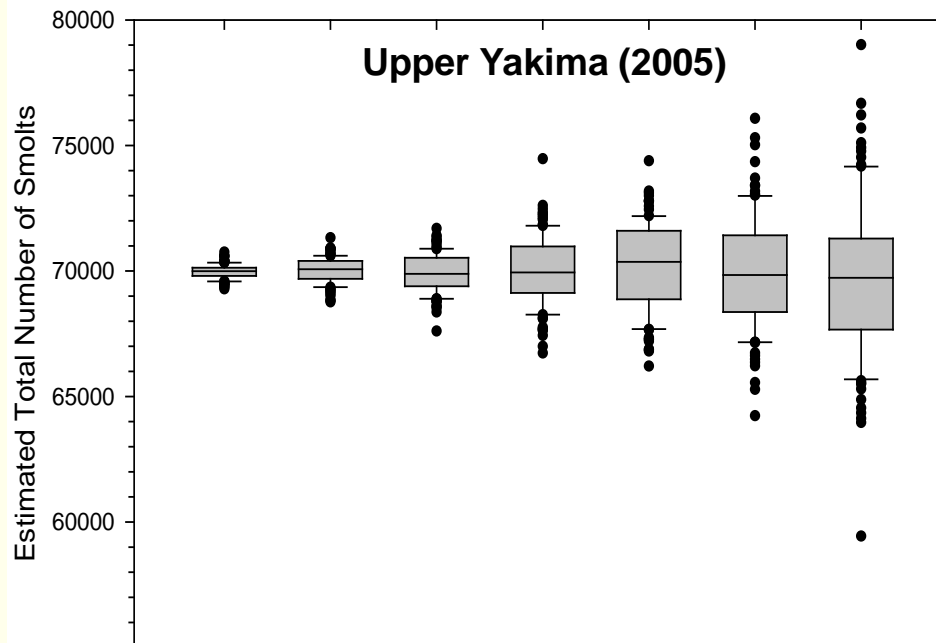
**35,500**    **55,000**

	Naches uYakimaSp	
Jan-Feb	36.1%	42.3%
March	8.4%	71.8%
April	29.2%	27.4%
May	3.7%	4.4%
Jun-July	4.6%	19.3%

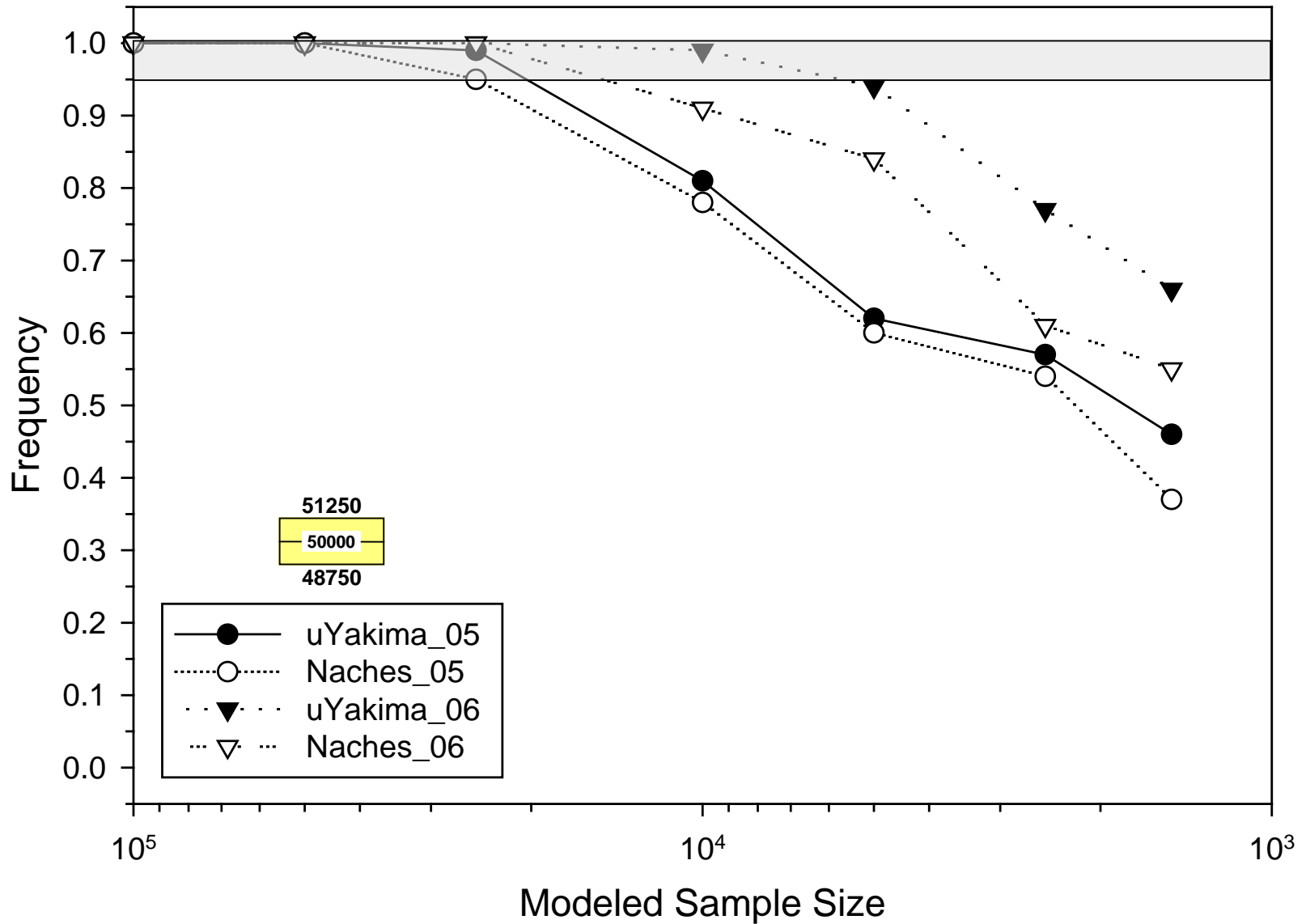
# The Model

---

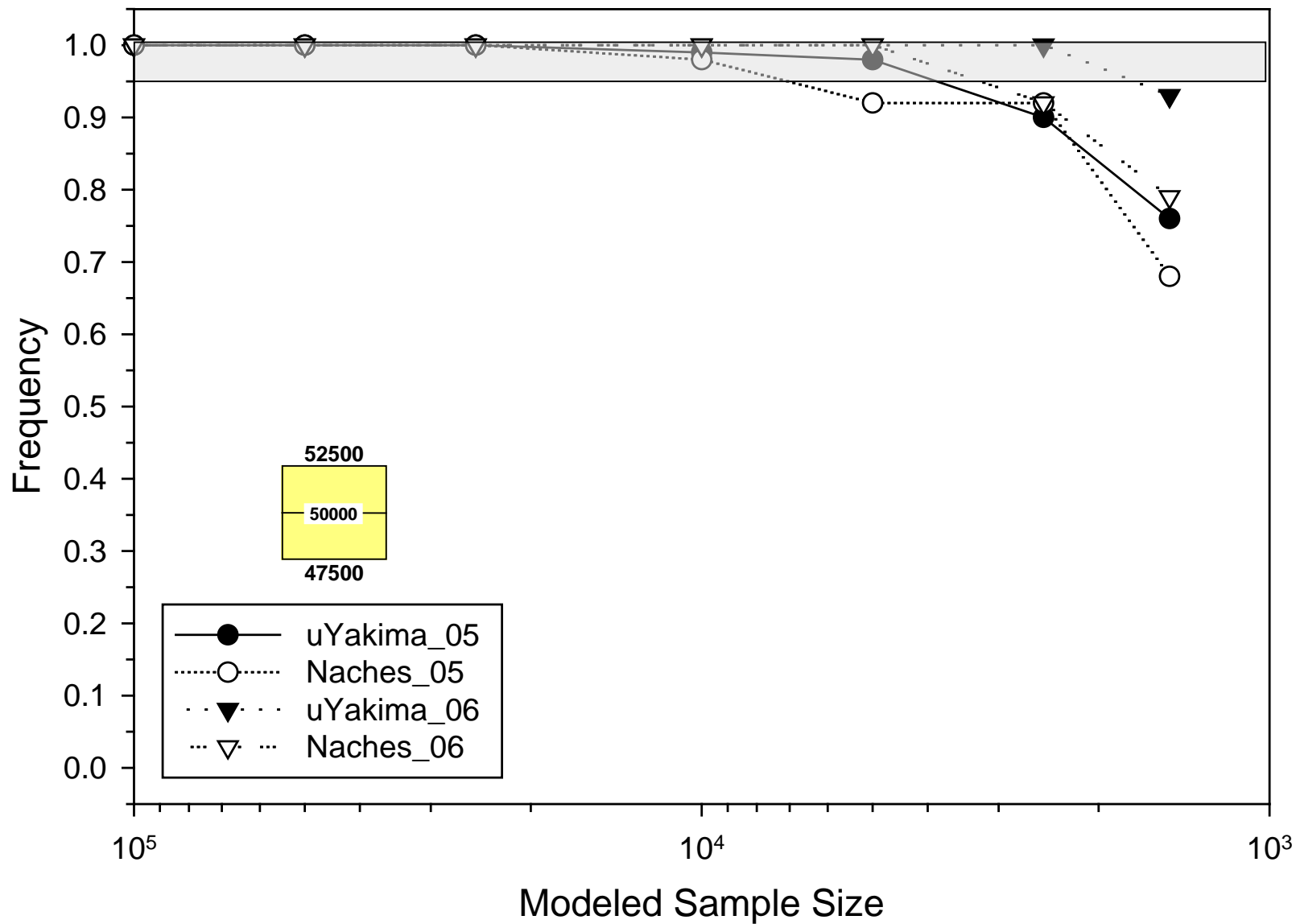
- *Repeat 100 times*  
For each modeled sample size
- *Separate analyses*  
2005 and 2006 data
- *Therefore*  
For each sample size 100 estimated number of smolts for Naches and upper Yakima – for both 2005 and 2006



# Within 5% of 1500 Result ( $\pm 2.5\%$ )

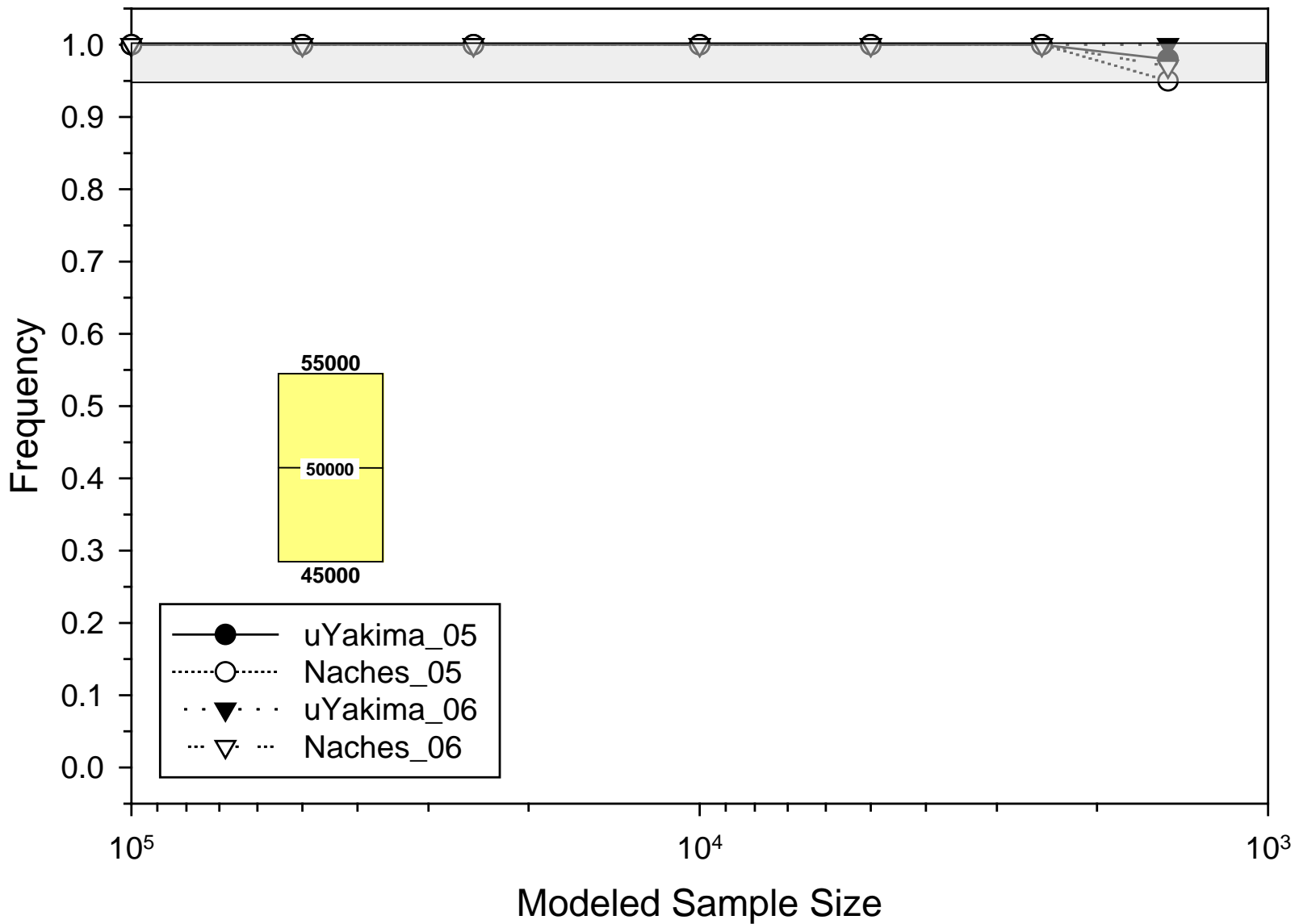


# Within 10% of 1500 Result ( $\pm 5.0\%$ )

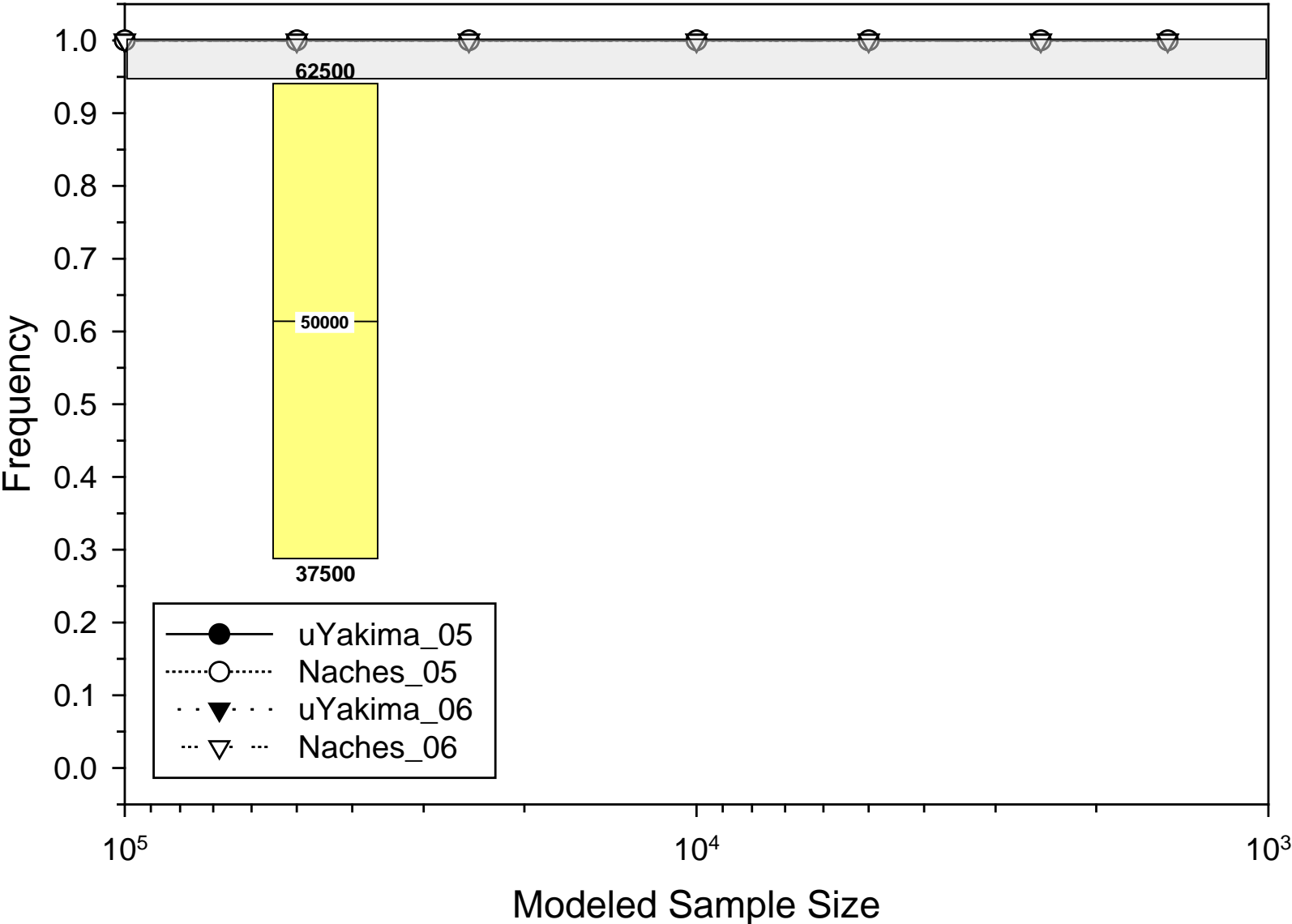




# Within 20% of 1500 Result ( $\pm 10.0\%$ )



# Within 50% of 1500 Result ( $\pm 25.0\%$ )



# Conclusions

---

- Appropriate sample size depends on desired level of precision
- Variability: Temporally (season and year) and population
- Precision (and associated error) is a function of sampling NOT genetics methods
- Other sources of error: GSI, run expansion

# Acknowledgements

---



Yakama Nation

Funding: BPA and WDFW State General Fund

Todd Kassler

Todd Pearsons

WDFW Molecular Genetics Lab