

A New Computer Program to Evaluate Biases in the **Two-Sample Mark-Recapture Abundance Estimator**



Christian Yap, Carl R. Ruetz III, and James McNair Annis Water Resources Institute, Grand Valley State University, Muskegon, Michigan E-mail: <u>yap@mail.gvsu.edu</u>, <u>ruetzc@gvsu.edu</u>, <u>mcnairja@gvsu.edu</u>

ABSTRACT

Two-sample mark-recapture sampling is a common method used to estimate stream fish abundance. The idea is to capture and mark fish in an initial sample. The fish are then released to mix randomly with the whole population. A second sample is obtained, and the number of marked and unmarked fish is recorded. The Chapman estimator uses the number of fish marked in the first sample, the total number of fish captured in the second sample, and the number of recaptured fish to estimate abundance. The assumptions are: 1) the population is closed, meaning no immigration, emigration, births, or deaths, 2) all fish are equally vulnerable to being captured during each sample, meaning marking does not change the behavior of fish, and 3) marks are not lost or overlooked. Violations of these assumptions can happen frequently and examining bias when a combination of assumptions are violated is difficult. To explore how simultaneously violating multiple assumptions of the Chapman estimator affects bias, we developed a computer simulation using Python software that allows the end user to assess bias by simulating a closed or open population, varying fish capture probabilities, and allowing fish to lose marks. The simulations allow the end user to experiment by intentionally violating any combination of model assumptions and determining the effect on estimator bias. This software should be useful to fisheries managers that use the Chapman estimator or instructors that teach the two-pass mark-recapture sampling for abundance estimation.

- **Sub-reach Size:** The percent of the study area being sampled.
- Number of Trials: The total number of times to repeat the simulation.



File name:	Figure_1.png
Save as type:	Portable Network Graphics (*.png)
	Portable Network Graphics (*.png)
	Encapsulated Postscript (*.eps)
 Hide Folders 	Joint Photographic Experts Group (*.jpeg;*.jpg)
	Portable Document Format (*.pdf)
	PGF code for LaTeX (*.pgf)
	Postscript (*.ps)
	Raw RGBA bitmap (*.raw;*.rgba)
	Scalable Vector Graphics (*.svg;*.svgz)
	Tagged Image File Format (*.tif;*.tiff)

File name: Simulation_1 Save as type: CSV Files(*.csv)

INTRODUCTION

- 2-sample mark-recapture is a commonly used method to estimate fish abundance.
- For statistical estimators to be valid, the sampling method must meet certain assumptions.
- Violating model assumptions can result in bias, and assessing the resulting bias is often difficult.
- This computer software was developed to help assess biases when assumptions of the users abundance estimator are violated.



Fig. 2. The blue curve above shows a balanced beta distribution if a user decides to keep migration in and out of the reach equal. Upstream-biased movement is represented by the orange curve.

RESULTS

- Results are available in a variety of formats:
- A Histogram Plot.
- > Textbox showing statistical data.
- \succ Table showing raw data for each trial run.
- \succ CSV, PNG files if user desires to export results



Save as types	corr mest nest)
 Hide Folders 	

Fig. 5. Filetypes available to users to save their simulation results.

Lincoln-Peterson (Chapman's) Estimator Raw Simulation Res	ults		
Lincoln Peterson (Chapman's) Estimator	# of Fishes Caught in Second Catch (N):	100	•
# of Fishes Marked in First Catch (M): 122	# of Fishes Marked in Second Catch (R):	35	•
Results			
10:43:37 - Estimated Fish Population: 419.25			
10:43:47 - Estimated Fish Population: 344.08333333333333			
	Clear Results Screen Estimate Popula	ition	

Fig. 6 Additional tab so that the user can estimate abundance from a sample using the Chapman estimator.



Fig. 1. Main graphical user window to begin use of the software.

METHODS

- Upon start-up of software, a user can manipulate the following variables under the "Raw Simulation" tab":
- **Total Fish Population in Study Reach:** Know fish abundance.
- > Closed or Open Population: With a closed population, there are no births, mortality, or movement in and out of the study reach (Figure 2).
- **Capture Probability:** The chance of capturing a

Fig. 3. A histogram plot to display simulation results.

. A	WRI Mark Recaptur	re Populat	ion Estimator v1.201	19							- 0	×
Lir	Options ncoln-Peterson (Cl	hapman'	s) Estimator R	aw Simul	ation	Results						
M A C P N N	Mean Population estimation: 49.2667 Actual Population Size: 100 Type: Closed Population Capture Probability for First Pass: q = 0.5Capture Probability for Second Pass: q = 0.0 Type: Capture Probability Varying Per Sample. Possible tag loss at 0.0% Subreach Type: No Subreach Parameter Type: Capture Probability Varying Per Sample. Number of Trials: 30 Migration Distance/Rate: None Type: Capture Probability Varying Per Sample.											
Load Simulation Set: 1 ~			~	Load Data		Vi	View Image 🗌 Graph Popu		ulation Clear All Saved Data			
A	ctual Population s	size: 100		^		Capture Prob (Q1)	nitial Subreach (S1	Tagged	Tag Lost	Mortality	Aigration Distance	Car ^
Open Population				1	0.1512	-70	-1	-1.0000	1	20.0	0.3	
Capture Probability for First Pass: q = 0.0Capture Probability for Second Pass: q = 0.0				ture	2	0.0902	-23	-1	-1.0000	1	-27.0	0.2.
Capture Probability Varying Per Sample.				3	0.4750	-90	-1	-1.0000	1	40.0	0.0	
Varied Subreach Size		.0%		4	0.0007	-3	-1	-1.0000	1	-47.0	0.3!	
Balanced: 0.0Migration Distance: 0.5				*	5	0.8367	-38	-1	-1.0000	1	-12.0	0.1(
Tri	als	Circt Da	Cotch Jacond D	Cati A	6	0.5032	-24	-1	-1.0000	1	-26.0	0.5(
1	1 0000	Pirst ras		iss Catt	7	0.9286	-87	-1	-1.0000	1	37.0	0.0
2	1.0000	0	0		8	0.4548	-36	-1	-1.0000	1	-14.0	0.7:
2	1 0000	0	0		9	0.8333	-75	-1	-1.0000	1	-25.0	0.0!
1 J J		Ĩ	Ŭ,	/								

TECHNICAL SPECIFICATIONS

- Executable for Windows Platforms.
- Open-sourced code available on GitHub.
- Multiprocessing to speed up simulations.

FUTURE IMPROVEMENTS

- Add visual simulation feature.
- Refactor code to add in data structures and algorithm to further speed up simulations.

ACKNOWLEDGMENTS

This research was made possible by the Bill & Diana Wipperfurth Student Research Scholarship. Special thanks to Grand Valley State University's Office of Undergraduate Research and Scholarship (OURS) for providing travel funding through the Academic Conference Fund.

REFERENCES

Hunter, J. D. 2007. "Matplotlib: A 2D Graphics Environment",









statistical analysis.

conditions of the most recent simulation as well providing





