

Ocean Ruler – an image-based, AI-driven approach to small-scale fisheries monitoring and catch size estimation

Authors: Jack Elstner¹, Lyall Bellquist^{1,2}, Frank Hurd², Dan Yocum³, Chris Schmuckal⁴, Mary Gleason², Tom Dempsey², Kate Kauer², Alexis Jackson² and Brice Semmens¹

¹Scripps Institution of Oceanography, University of California San Diego

²The Nature Conservancy, Sacramento, CA

³National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, CA

⁴University of California Santa Barbara

Corresponding Author: Jack Elstner (jelstner@ucsd.edu)

Supplementary Information

Evaluating Software Performance Across Users

For field tests where multiple software users produced Ocean Ruler size estimates of the same set of specimens (i.e., abalone, finfish, pen shell), we compared Ocean Ruler performance across different software users. To do this, we again modeled software deviance (D_i), which was standardized prior to model fitting, as a normally distributed response variable with parameters B and σ . The parameter B describes the central tendency of the overall distribution of software deviance values and represents software bias (e.g., the systematic difference, or average offset of software size estimates to hand measurements). The parameter σ , in contrast, represents measurement error associated with software measurements (e.g., the variability observed in software measurements after accounting for bias). For each fishery group, we then fit four candidate models and performed model comparison to assess whether estimates of B and σ are improved when accounting for variance between individual users (j).

$$\begin{aligned} 1) \quad & D_i \sim \text{Normal}(B, \sigma) \\ & B \sim \text{Normal}(0, 2) \\ & \sigma \sim \text{Exponential}(1) \end{aligned}$$

$$\begin{aligned} 2) \quad & D_{i,j} \sim \text{Normal}(B_j, \sigma) \\ & B_j \sim \text{Normal}(0, 2) \\ & \sigma \sim \text{Exponential}(1) \end{aligned}$$

$$\begin{aligned} 3) \quad & D_{i,j} \sim \text{Normal}(B, \sigma_j) \\ & B \sim \text{Normal}(0, 2) \\ & \sigma_j \sim \text{Exponential}(1) \end{aligned}$$

$$\begin{aligned} 4) \quad & D_{i,j} \sim \text{Normal}(B_j, \sigma_j) \\ & B_j \sim \text{Normal}(0, 2) \\ & \sigma_j \sim \text{Exponential}(1) \end{aligned}$$

For the finfish dataset, we fit models 1-4 using software deviance values associated with initial Ocean Ruler estimates, because anglers collected images of specimens, but did not contribute to manually correcting estimates (this task was done by a member of our science staff). Here, software deviance was calculated using only hand measurements taken by a trained scientist, and did not incorporate hand measurements taken by anglers. For the pen shell dataset, we used software deviance values associated with manually corrected Ocean Ruler estimates, because software users corrected initial estimates, but did not photograph specimens themselves. For the abalone dataset, we fit separate models using software deviance values for both initial and manually corrected estimates, because users participated in all facets of software use (e.g., photographed specimen, generated initial estimate, and manually corrected initial estimate if necessary). For all models, we specified weakly informative priors. After model fitting, we performed model selection using Watanabe-Akaike Information Criteria (WAIC), which allowed us to identify candidate models with the greatest predictive capacity based on lowest WAIC scoring.

Results Summary

For all taxa where multiple participants produced software size estimates of the same set of specimens, we did not find evidence for differences in software performance between users. Accounting for differences in B and σ across users did not improve predictive performance and model parsimony, as models without individual user effects consistently produced lower WAIC values. (Table S1). Therefore, we did not take software variability between users into consideration in subsequent analyses.

Table S1 Candidate models fit to determine whether estimates of B and σ are improved when accounting for variability between individual software users (j). Included is a description of each model and calculated Watanabe-Akaike Information Criteria (WAIC) values. Lower WAIC values indicate models with the greatest predictive capacity. “Initial” refers to initial Ocean Ruler software estimates, while “corrected” refers to manually corrected software estimates.

Model	Description	WAIC	
Model 1	Software Bias Fixed Across Users + Measurement Error Fixed Across Users	Abalone (Initial)	1542.1
		Abalone (Corrected)	1541.6
		Finfish (Initial)	854.9
		Pen Shell (Corrected)	3985.9
Model 2	Software Bias Variable Across Users + Measurement Error Fixed Across Users	Abalone (Initial)	3140228.3
		Abalone (Corrected)	2644919.6
		Finfish (Initial)	1259971.5
		Pen Shell (Corrected)	144598583.2
Model 3	Software Bias Fixed Across Users + Measurement Error Variable Across Users	Abalone (Initial)	4653382.5
		Abalone (Corrected)	2706055.2
		Finfish (Initial)	4102549.8
		Pen Shell (Corrected)	713504446.2
Model 4	Software Bias Variable Across Users + Measurement Error Variable Across Users	Abalone (Initial)	5422334.3
		Abalone (Corrected)	3501357.4
		Finfish (Initial)	4279155.7
		Pen Shell (Corrected)	714169674.1