

Ecography

ECOG-05143

Teitelbaum, C. S., Amoroso, C. R., Huang, S., Davies, T. J., Rushmore, J., Drake, J. M., Stephens, P. R., Byers, J. E., Majewska, A. A. and Nunn, C. L. 2020. A comparison of diversity estimators applied to a database of host–parasite associations. – *Ecography* doi: 10.1111/ecog.05143

Supplementary material

Appendix 1

Supplementary figures

Figure A1: Examples of relationships between PSR and a trait given different distributions of error. Red lines show the 1:1 “true” relationship. As error increases, the ability to recover the true relationship declines. Panel headers are the standard deviation of the error term.

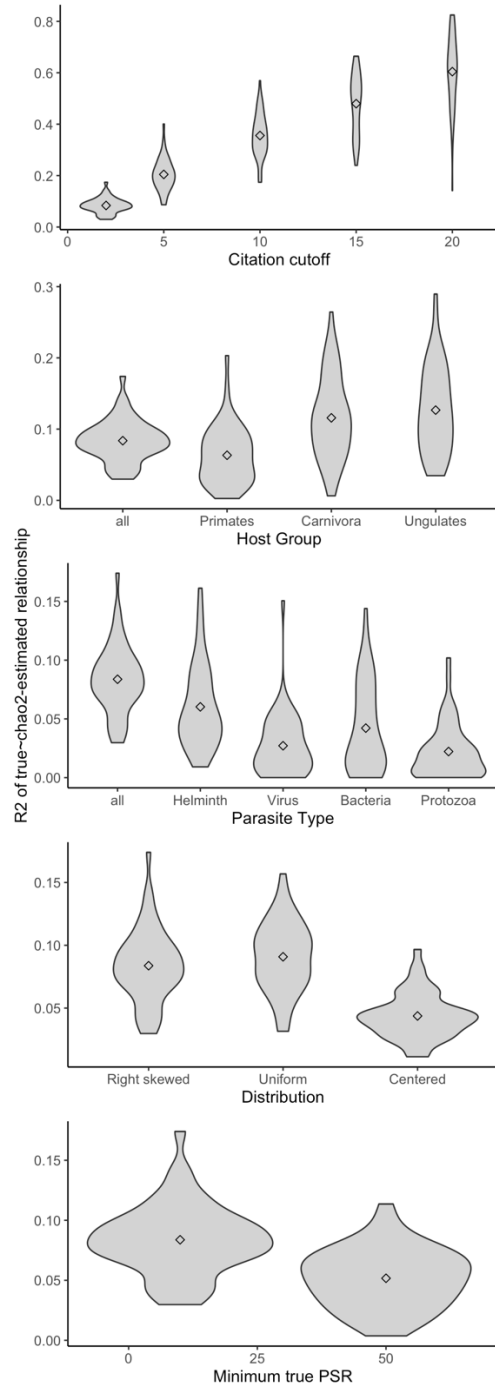


Figure A2: Performance of the Chao2 estimator for subsets of the GMPD. Violin plots show values of 50 simulations; points show means across these 50 simulations. For all parameters not shown in each plot, values of other parameters were held constant at the values shown in Table 1 in the main text. The exception is for parasite groups, where we used a maximum true PSR of 500 (because this value should be lower for individual groups than for all parasites combined). (A) Performance improves when including only better-studied species. (B, C, D) There is little effect of considering single groups of hosts or parasites or a different distribution. (E) Estimates improve with a lower presumed minimum PSR. Note the different y-axis scale in each panel.

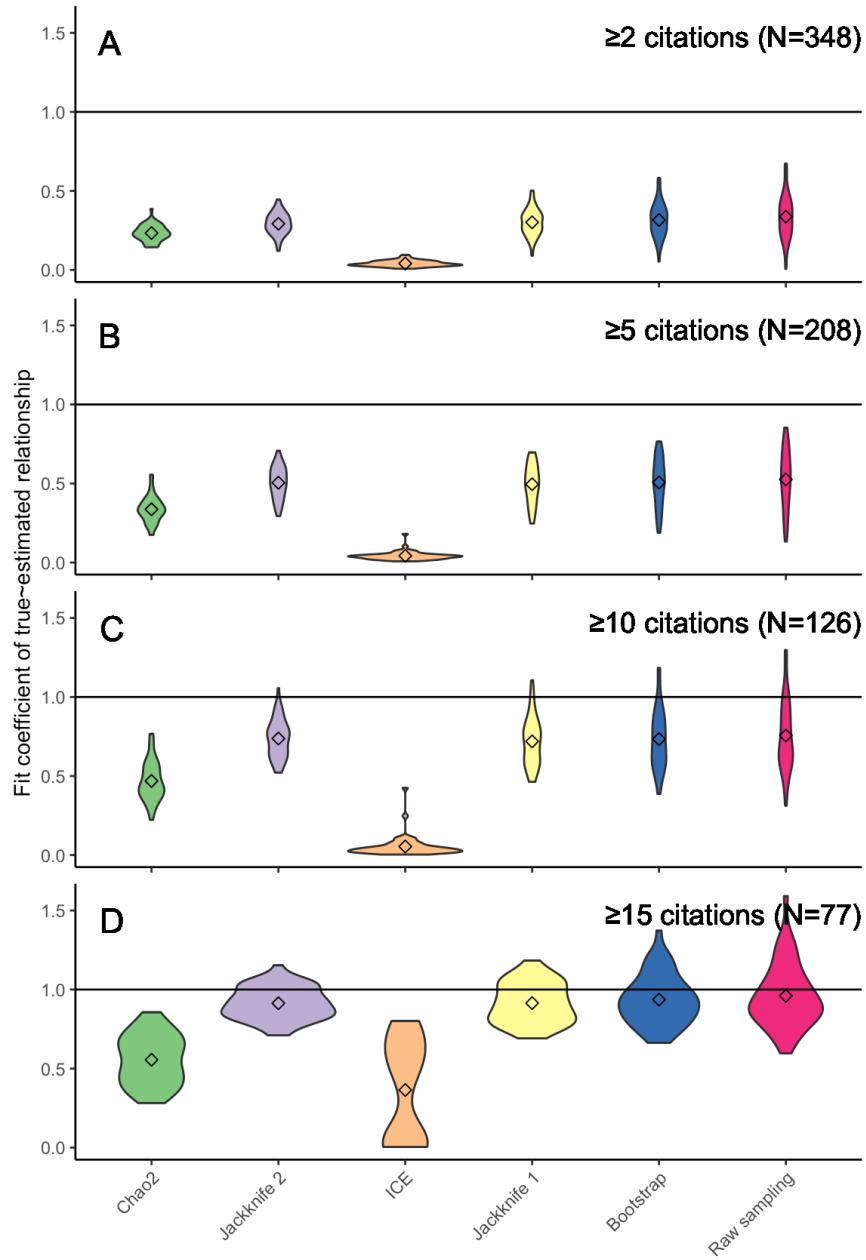


Figure A3: Bias of methods of controlling for sampling effort, measured using the fit coefficient (β) of a linear model predicting true PSR from estimated PSR using each method. Points show means across these 50 simulations. The four panels represent different subsets of hosts, included based on citation cutoffs or a minimum number of citations required for inclusion in the analysis.

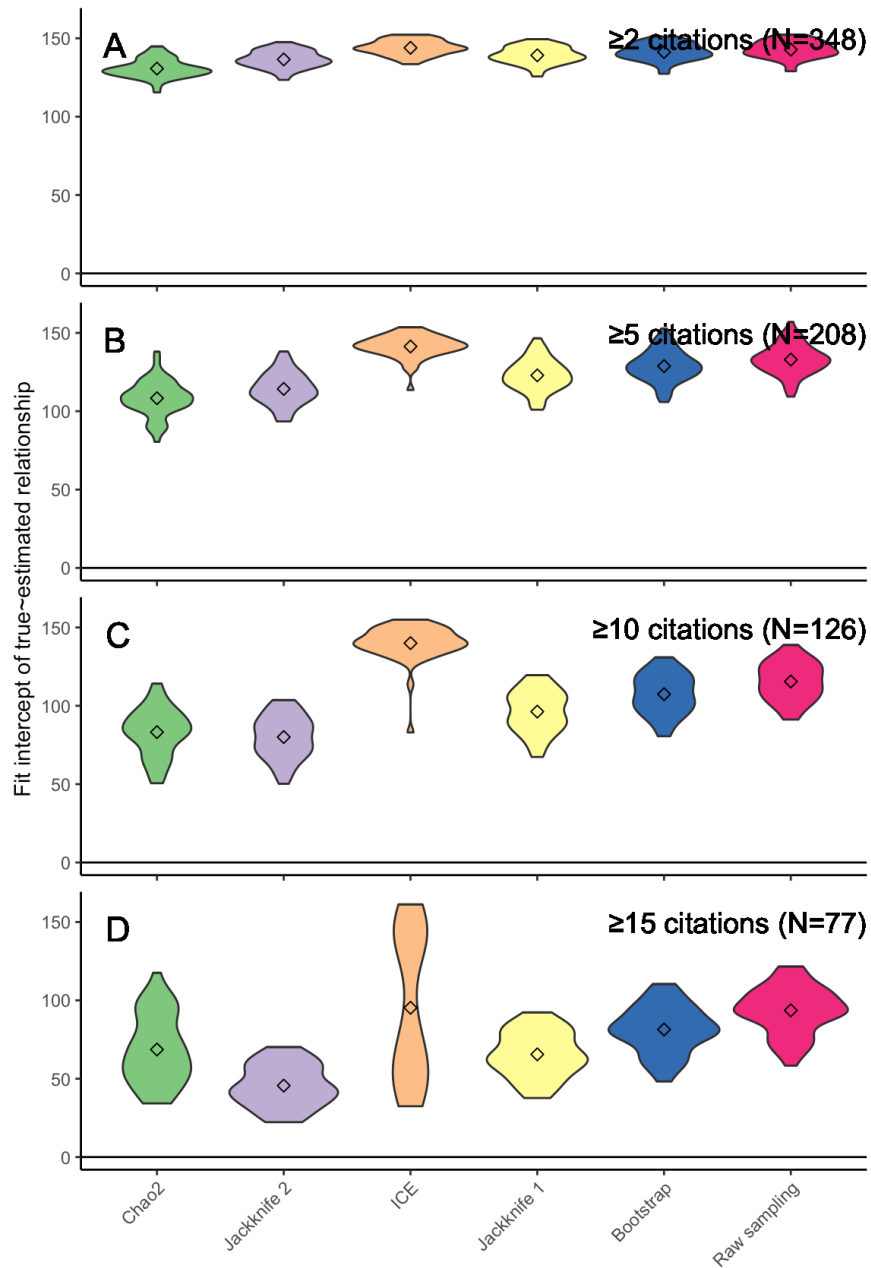


Figure A4: Bias of methods of controlling for sampling effort, measured using the fit intercept of a linear model predicting true PSR from estimated PSR using each method. Points show means across these 50 simulations. The four panels represent different subsets of hosts, included based on citation cutoffs or a minimum number of citations required for inclusion in the analysis.

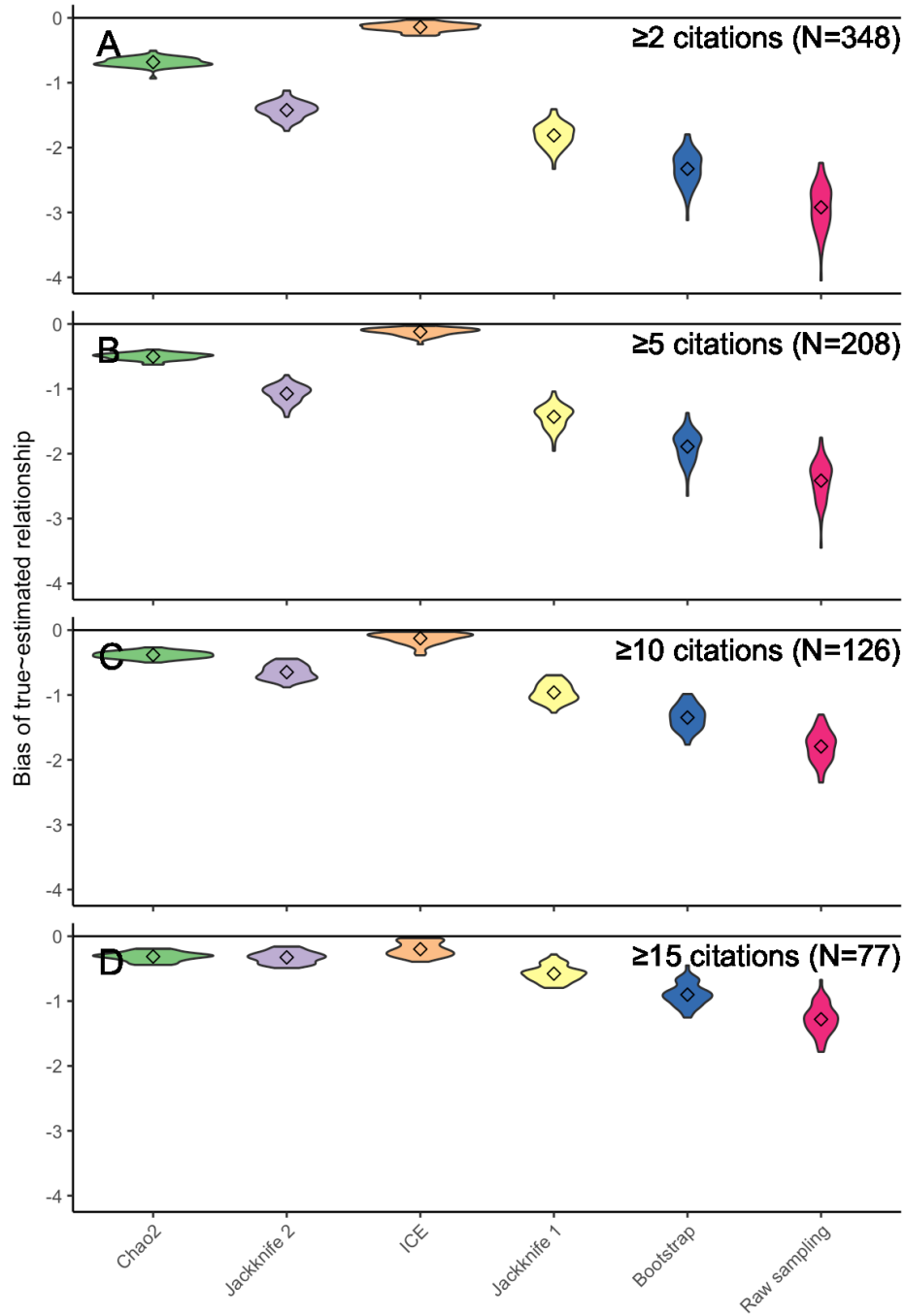


Figure A5: Bias of methods of controlling for sampling effort, measured using the difference in slope (β) between a linear model predicting true PSR from estimated PSR and a linear model with the intercept forced through the origin. Points show means across these 50 simulations. The four panels represent different subsets of hosts, included based on citation cutoffs or a minimum number of citations required for inclusion in the analysis.

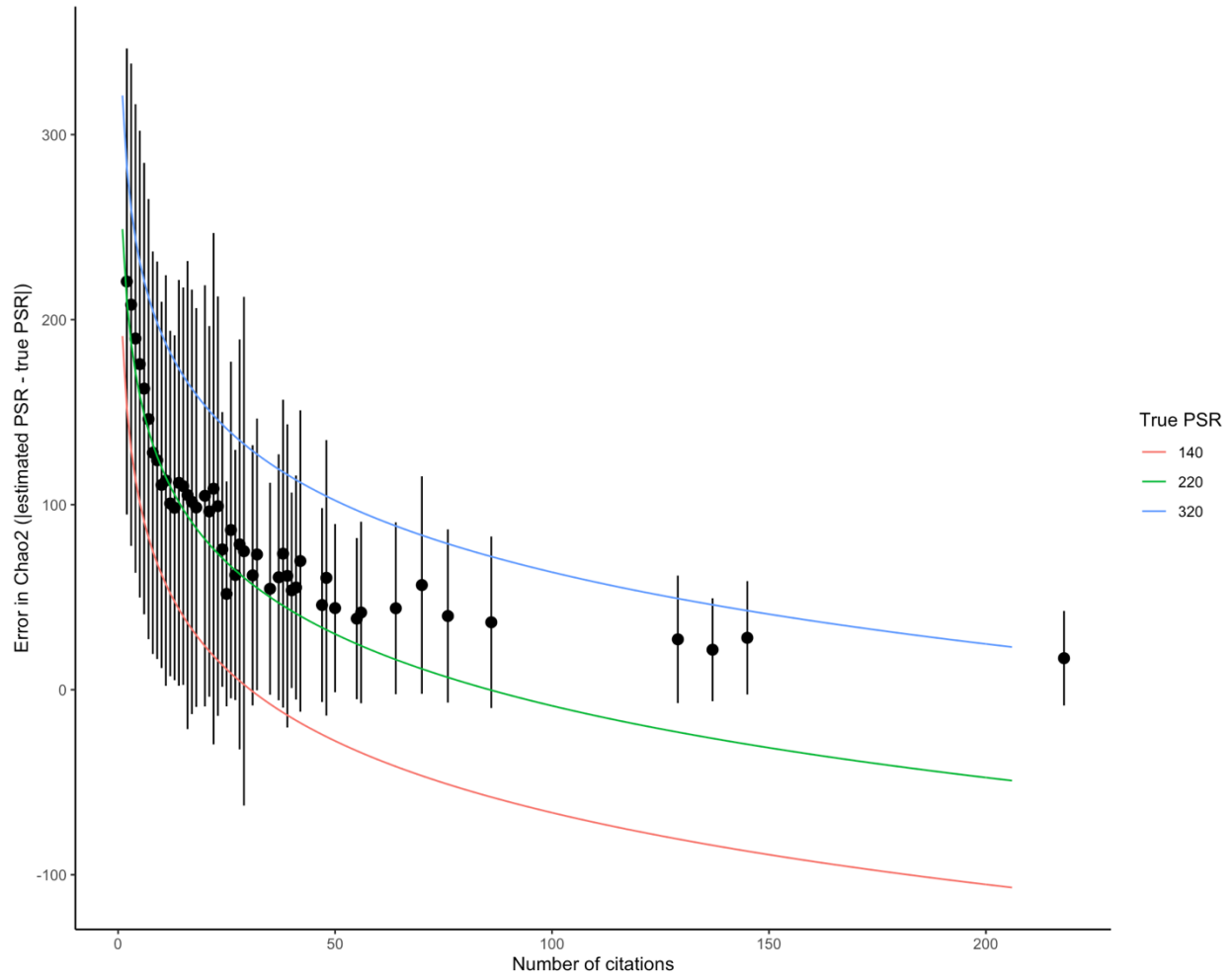


Figure A6: Chao2 is more accurate for better-studied species. “Error in Chao2” is the absolute difference between estimated and true richness, where a value of 0 would be a perfectly accurate estimate. Each point represents the mean and the bar represents the standard deviation of error across the 50 simulated datasets, with parameters in Table 1. Each line shows the modeled linear relationship for the 25%, 50% and 75% quantiles of PSR. Error was positively related to true PSR (Table A2).