



The lack of characterization of inferential operators leaves the robustness of global insect richness estimates undetermined

Simone Fattorini^a and Paulo A. V. Borges^{b,1}

Colwell et al. (1) build on an exceptional DNA-barcoded inventory of tropical insect diversity to estimate global insect richness. Following an approach conceptually similar to that introduced by Erwin (2), Colwell et al.'s analysis proceeds through a sequence of successive inferential steps: 1) Species richness is first estimated for a subfamily of braconid wasps (Microgastrinae) using an estimator that provides a statistical lower bound (CARE-1/ $\hat{N}$). 2) This estimate is subsequently used to infer sampling completeness for the entire insect fauna of the study area (ACG). 3) Local insect richness is then extrapolated to the global fauna through proportional scaling relationships. Therefore, the final estimate is determined not by a single estimator, but by several successive inferential operators that convert an estimate obtained at one stage into an estimate at the next. These operators are motivated by biological arguments, but remain unconstrained by the observations used to construct the estimate. For example, the estimated undersampling ratio for Microgastrinae is applied to all insects, based on the assumption that it approximately reflects sampling completeness across taxa, but the uncertainty associated with this extrapolation is not quantified. Additionally, if undersampling differs systematically among insect taxa, transferring a single ratio to the entire fauna may shift the estimated richness even when the statistical uncertainty of the initial estimator is correctly characterized. Moreover, because richness estimates are highly sensitive to rare species, the applicability of CARE-1 assumptions to highly diverse insect assemblages with many rare species represents an additional source of uncertainty. Similarly, the correlation between cumulative insect and tree richness observed across 12 plots in the exceptionally heterogeneous tropical landscape of Panama does not establish that the global insect-to-tree ratio equals the ACG ratio, yet this ratio is transferred from the study area to the globe through an upscaling step whose associated uncertainty remains unquantified. Quantitative inference requires not only motivated transformations but also explicit characterization of their contribution to the uncertainty of the final result. Since every inferential operator introduces additional uncertainty, the resulting uncertainty propagates

through a sequence of dependent transformations. This propagated uncertainty includes both statistical uncertainty arising from the estimator itself and model uncertainty associated with the inferential operators connecting successive stages of the analysis (3–5). Since each operator conditions all subsequent ones, even modest deviations introduced early in the inferential chain may substantially affect the final estimate. While the confidence intervals for CARE-1/ $\hat{N}$ quantify statistical uncertainty associated with the initial richness estimate, they do not characterize uncertainty arising from the assumptions connecting successive stages. Consequently, the precision of the first step is not automatically inherited by the entire inferential chain, which leaves the robustness of the global estimate of insect richness provided by Colwell et al. (1) quantitatively undetermined. We hope that future studies will explicitly quantify the contribution of each inferential operator to the uncertainty of the final estimate. Until the contribution of each inferential operator has been explicitly quantified, the robustness of the proposed estimate of 14 to 20 million insect species remains quantitatively undetermined.

Data, Materials, and Software Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. P.A.V.B. was supported by FCT-UID/00329/2025–Centre for Ecology, Evolution and Environmental Changes DOI: <https://doi.org/10.54499/UID/00329/2025>.

Author affiliations: ^aDepartment of Life, Health and Environmental Sciences, University of L'Aquila, Coppito (L'Aquila) 67100, Italy; and ^bCentre for Ecology, Evolution and Environmental Changes, Azorean Biodiversity Group, Global Change and Sustainability Institute, School of Agricultural and Environmental Sciences, University of Azores, Angra do Heroísmo, Azores 9700-042, Portugal

Author contributions: S.F. and P.A.V.B. designed research; S.F. and P.A.V.B. performed research; and S.F. and P.A.V.B. wrote the paper.

The authors declare no competing interest.

Copyright © 2026 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: paulo.av.borges@uac.pt.

Published October 5, 2026.

1. R. K. Colwell et al., Constructing a lower-bound estimate of the global number of insect species on a hyperdiverse empirical foundation. *Proc. Natl. Acad. Sci. U.S.A.* **123**, e2524283123 (2026).
2. T. L. Erwin, Tropical forests: Their richness in Coleoptera and other arthropod species. *Coleopt. Bull.* **36**, 74–75 (1982).
3. D. Draper, Assessment and propagation of model uncertainty. *J. R. Stat. Soc. Ser. B* **57**, 45–97 (1995).
4. Joint Committee for Guides in Metrology, *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (JCGM 100:2008)* (JCGM, 2008).
5. A. Saltelli et al., *Global Sensitivity Analysis: The Primer* (Wiley, 2008).