

Closing the gap between statistical and scientific workflows for improved forecasts in ecology

Victor Van der Meersch^{1*}, James Regetz², T. Jonathan Davies^{1,3} & EM Wolkovich¹

May 2, 2025

¹ Forest and Conservation Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

² National Center for Ecological Analysis and Synthesis, 1021 Anacapa St, Santa Barbara, CA 93101, United States

³ Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

* <mailto:victor.vandermeersch@ubc.ca>

For: Workflow for Applied Data Analysis theme issue for *Phil Trans A* as an *Opinion*

Abstract

Concerns about increasing biodiversity loss and climate change have led to greater demands for useful ecological models. Datasets relevant for developing these models have also increased in size and complexity, including in their geographical, temporal and phylogenetic dimensions. New research often suggests that models accounting for these complexities can yield more accurate trends and predictions. We argue, however, that the usual workflows for model fitting in ecology make it difficult to evaluate and compare current models for several reasons. First, the research community is split between two disconnected paradigms: using data to fit simple, trend-like models with few parameters, or developing forecasting models that include complex, mechanistic submodels only indirectly informed by data. Second, in both cases, models tend to be developed not only in isolation of one another, but also without a coherent framework for linking scientific questions and understanding to statistical and modeling decisions, throughout the entire modeling process. We propose a unified, principled workflow for end-to-end empirical model development that bridges the gap between process-based and statistical approaches, integrates sound statistical and scientific practices, and especially relies on data simulation to inform decisions at multiple steps in the process. We argue that this approach, coupled with a shift toward universal training, more open model sharing, and alignment on common datasets, could transform for ecological modeling.

¹ 1 Introduction

- ² Anthropogenic drivers are reshaping natural systems (Díaz *et al.*, 2019). Impacts are projected
³ to increase in coming decades, as climate change accelerates biodiversity loss, altering ecosystem
⁴ services and human well-being (IPBES, 2019). Implementing sustainable policies to mitigate these
⁵ impacts is thus a global priority, but designing the best policies requires estimating and under-
⁶ standing biodiversity and ecosystem trends to date alongside the skill to forecast future dynamics.
⁷ Meeting these policy needs has led often to two separate paths: one focused on estimating trends
⁸ from new global datasets, and another focused on forecasting from generally distinct datasets or

9 mechanistic models based on less data. Newly available large-scale, long-term datasets have pro-
10 vided our first ‘global’ estimates of biodiversity trends (e.g. Loh *et al.*, 2005; Dornelas *et al.*, 2018),
11 but these data—gathered opportunistically from multiple sources—are unbalanced and suffer from
12 large geographic, temporal and taxonomic biases. Models to date have failed to fully address these
13 challenges and, perhaps because of these limitations, are rarely if ever used for forecasting. Instead,
14 forecasting—under different plausible scenarios—has generally relied on entirely different datasets
15 combined with either correlative or process-based models (IPBES, 2019), with process-based mod-
16 els often promoted as the most realistic approach (Urban *et al.*, 2016; Pilowsky *et al.*, 2022) because
17 they focus on mechanistic representations of ecosystem functioning. These approaches have failed
18 to yield clear agreement on current species trends, leading to ongoing debates about the magnitude
19 and even direction (Dornelas *et al.*, 2014; Leung *et al.*, 2020; Buschke *et al.*, 2021; Johnson *et al.*,
20 2024), and producing forecasts that diverge due to high model uncertainty at the ecological level
21 (Cheaib *et al.*, 2012; Thuiller *et al.*, 2019).

22 We argue that current debates and diverging forecasts are driven in large part by the incoherent
23 and disconnected workflows used today in ecology (Loreau *et al.*, 2022; Talis & Lynch, 2023; John-
24 son *et al.*, 2024). Research estimating biodiversity trends has become focused on methodological
25 aspects; the current workflow fails to examine the gap between ideal and available data, and rarely
26 tests for predictive accuracy that could scale up to allow forecasting. At the same time, process-
27 based models developed for forecasting often evolve through the addition of new separate layers
28 or components. These new parts are often disconnected from the original research aim, its data
29 stream, and the previous scientific insights, because current approaches rarely examine the model
30 as a functioning whole and thus ignore major problems (e.g., non-identifiability, discussed below).

31 Workflows that fully integrate all the steps required to build a model from an ecological question,
32 with evaluation of limitations and potential problems before estimating its parameters and making
33 projections, could reduce many of these problems. In particular, we argue that workflows that
34 incorporate data simulation at multiple steps can quickly identify flaws in model structure and
35 constraints in data, and allow us to understand when, where, and why different models diverge
36 (McElreath, 2018; Betancourt, 2020; Gelman *et al.*, 2020; Wolkovich *et al.*, 2024). Towards this
37 aim, we outline the steps of a universal workflow that could harmonize both trend estimation and
38 forecasting.

39 **2 Scientific method and workflows**

40 Quantitative science relies on a model-based framework to confront hypothesis with data (Chamber-
41 lin, 1965). In an idealized scientific method, we would formulate a research question and hypotheses,
42 design an experiment accordingly, build a model, collect data, and use this data to inform our model
43 and differentiate between hypotheses. This method underlies much of the recent pre-registration
44 movement, where hypotheses and methodology are defined prior to data collection (Nosek *et al.*,
45 2018). But this idealized method often does not apply to the reality of ecological research. Many
46 important questions cannot be addressed through controlled experiments and replications. In such
47 cases, we must rely on existing, heterogeneous datasets alongside uncertain and incomplete the-
48 ory to provide a large-scale and long-term perspective (Hilborn & Mangel, 1997). Indeed, most
49 macroecological insights have emerged from exploring patterns in these datasets (exploratory data
50 analysis).

51 This reality should drive researchers to use more robust and coherent methods. But the current
52 workflows combined with the challenges ecologists are facing—both in term of data complexity and

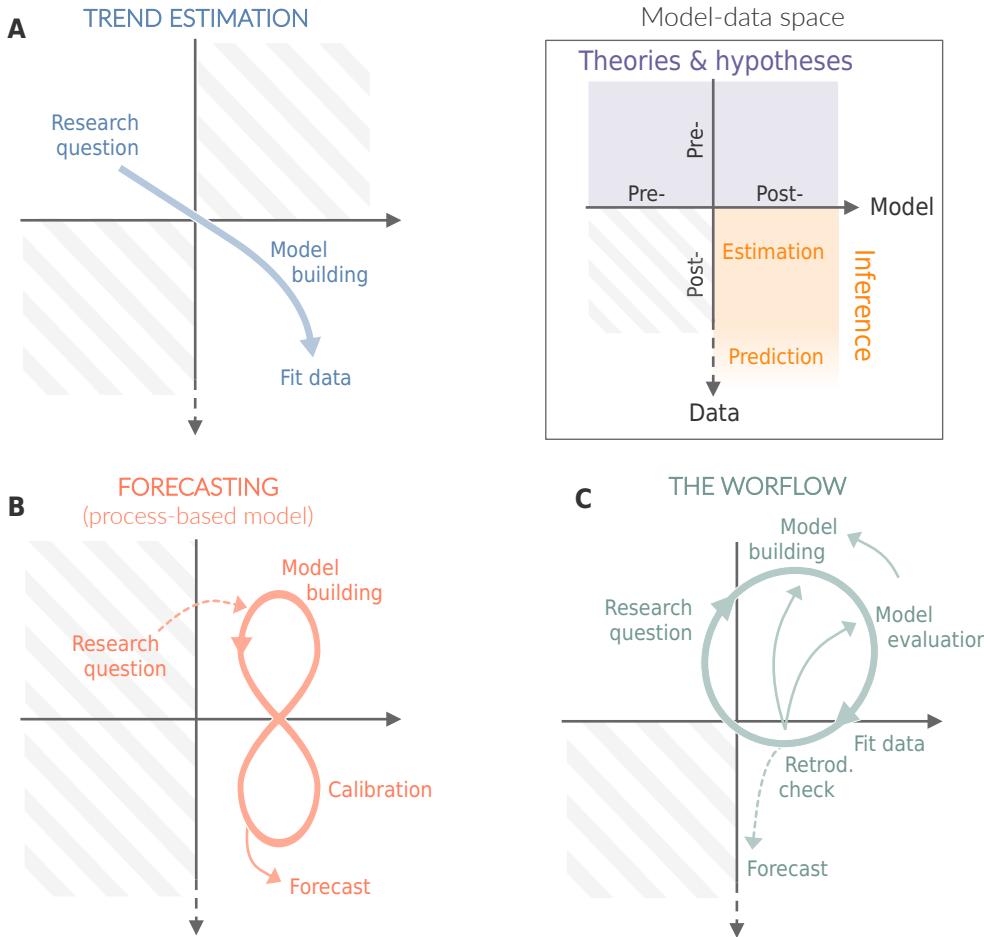


Figure 1: Different trajectories through model-data space during model development. Ideally, this begins in the pre-model/pre-data quadrant, where the research question is defined. The next step moves to the post-model/pre-data quadrant, where a model is build—*independently* of the data. This is where we should spend a good deal of time evaluating model behavior using simulated data. Once we are satisfied with our model, we can move to the post-model/post-data quadrant, where the model is fitted to real data. Then, we perform retrodictive checks to compare model predictions to observations, which likely give us some feedbacks to refine our model.

53 societal needs—instead may lead to persistent problems. Trend estimation has focused mostly on
 54 fitting a model to empirical data in a uni-directional way—without the checks and likely feedbacks
 55 that often highlight uncertainty and related limitations in the model and/or data (figure 1a). For
 56 forecasting, researchers have focused on making predictions with increasingly complex mechanistic
 57 models, frequently obscuring the steps underlying model building and parameterization (figure 1b).
 58 Researchers often calibrate the different parts of these models separately, and fix some parameter
 59 values based on experiments and expert knowledge, to avoid problems when trying to fit the model
 60 as a whole. Addressing these problems while accounting for the realities of working with ecological
 61 data requires a more comprehensive workflow.

62 We argue a workflow that moves along the data-model space in a coherent sequence of steps
 63 (figure 1c) could reduce many of these problems and thus improve ecological science. The first
 64 step of this workflow is to define an explicit research question and formulate hypotheses (step
 65 1, figure 2). This involves making clear assumptions about the most influential drivers, within

the specific context of our study. This step should guide the construction of a narrative model of how we believe the system works, focusing on the mechanisms that could generate the data we observe, including the observational error. From this narrative, we can then develop a mathematical model—an ensemble of equations that encapsulates our knowledge and is designed to answer our research question (step 2, figure 2). The general idea is to start with a relatively simple model that we could refine later (see example workflow we provide). At this stage, prioritizing biologically meaningful parameters is crucial, as it allows us to have a sense of plausible parameter values. This means choosing a model formulation where each parameter corresponds to an interpretable behavior (which sometimes requires considering alternative parameterizations).

With a model in place, the next step focuses on testing and understanding it via data simulation (step 3, figure 2). ‘Fake’ or ‘test’ data are generated directly from the model by fixing parameters to some reasonable range of values (which is straightforward if the parameters are interpretable) and from fake predictor data. We then fit our model to this simulated dataset and evaluate its ability to recover the prescribed values. At this stage, the focus is on understanding the model, so we may need to spend time thinking about whether each parameter is reasonable or not. This is also a step for making sure the model is working as expected (see example workflow).

Once we are confident about our model structure, we can introduce real data as part of an initial model fitting step (step 4, figure 2). This way, we obtain parameter estimates constrained by observations. These parameter estimates lead to the second data simulation step, this time using our fitted model parameters to generate predictions (step 5, figure 2). This—which we call a retrodictive check (Betancourt, 2020, also sometimes called a posterior predictive check, Gelman *et al.*, 2020)—allows model output to be compared to observations. It’s only once all steps have been completed that we can interpret parameter values with respect to our research question. The workflow encourages a focus on the full model, where any parameter (such as a trend estimate) must be carefully interpreted alongside others, as all are fundamental components that shape both inference and forecasting.

Within such a workflow, forecasting emerges as a natural outcome: rather than being a final goal, it only involves jointly modeling new circumstances along with the original data. The adoption of the workflow in macroecological studies—where model building is often informed by patterns in the data—would also make the exploratory analysis more transparent (as an explicit preliminary step for entering the workflow) and would compel researchers to more clearly develop a research question before extensive model fitting.

A key feature of this workflow is the central role of data simulation, which introduces two feedback loops. The first feedback arises when we evaluate the model on simulated data. The failure of the model to recover known parameter values and handle the complexity of the simulated data should prompt reconsidering the model, or even reformulating the research question. Further, this step might reveal that some parameters are highly non-identifiable (meaning the parameter(s) cannot be uniquely estimated), flagging the need to change the model structure—before incorporating empirical observations. The second feedback loop comes from the retrodictive check. Discrepancies here may indicate a missing key driver, and suggest the current model is too simplistic. We can refine the model to integrate the missing process(es) (if we can identify them) and return to the start of the workflow. Insights from the retrodictive check can also lead us to introduce additional complexity when simulating fake data, such as phylogenetic structure or observational biases (e.g. unbalanced data). For example, if a researcher realizes their empirical data is geographically biased, then this bias should be built into the model and thus then into this data simulation step. This iterative evaluation of the model moves beyond a simple reliance on goodness-of-fit metrics.

112 At each iteration, we are able to evaluate the model behavior, both with simulated and real data,
113 taking into account our expert knowledge of the ecological processes.

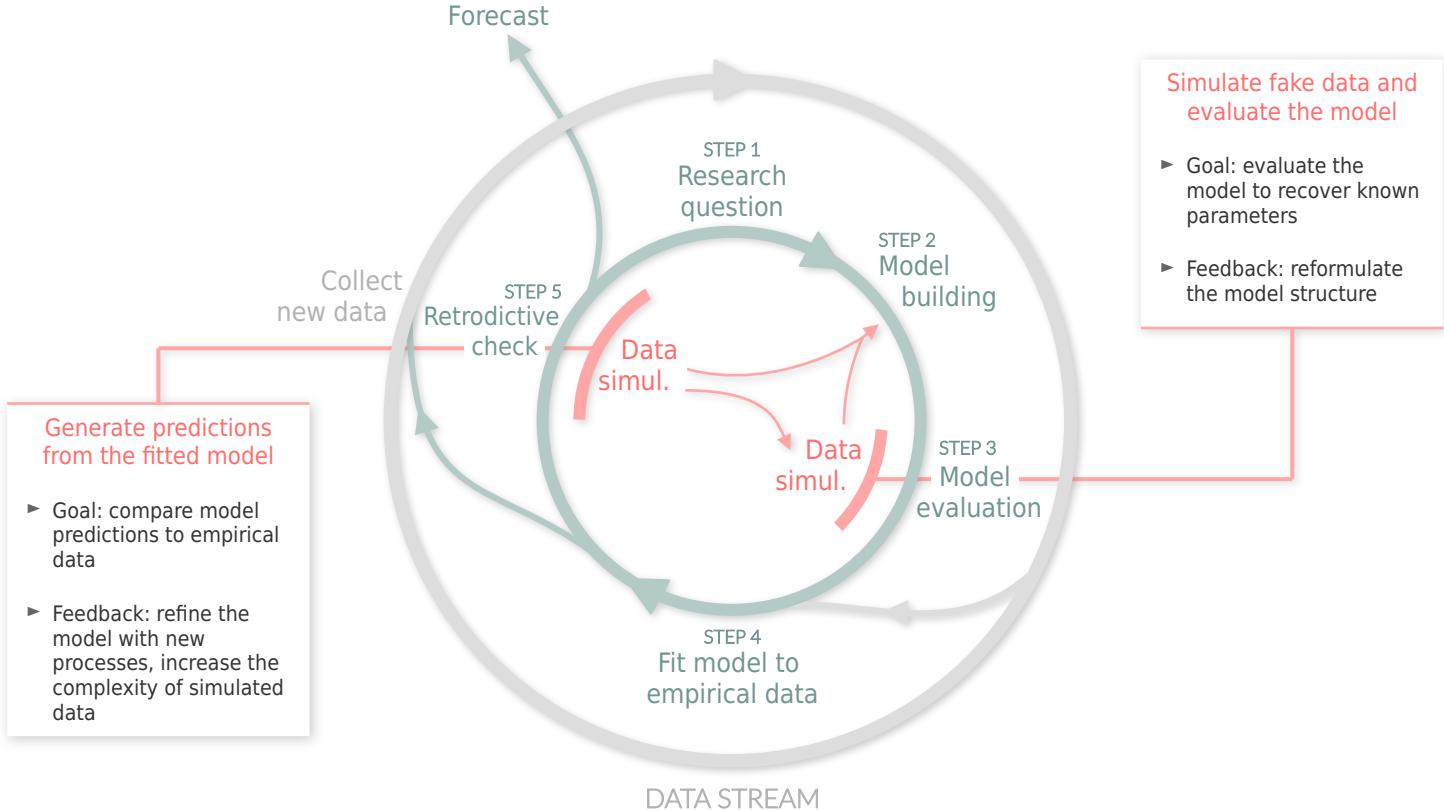


Figure 2: The workflow we propose here (which builds from recent advances in workflows, Betancourt, 2020; Gelman *et al.*, 2020; Schad *et al.*, 2020; Grinsztajn *et al.*, 2021; van de Schoot *et al.*, 2021; Wolkovich *et al.*, 2024) focuses on iterative feedbacks between the research question, model and data. Most research should start at Step 1, with the research question, followed by extensive model building and evaluation using simulated data (Steps 2 and 3) before proceeding to fitting the model on empirical data and examining it through retrodictive checks (Steps 4 and 5). Within this workflow, forecasting is a natural output of the process and not a separate process or one available for only certain modeling approaches. The integration of observations (the data stream) occurs at Step 4—only after the model has been thoroughly evaluated—and can also highlight opportunities to collect new data and enrich the data stream.

114 3 The workflow in practice

115 Across the different fields of ecology—for both parameter estimation and forecasting—a systematic
116 application of a coherent workflow could highlight the best opportunities to reduce uncertainties
117 and yield new scientific insights. This will help refocus the debate on designing new hypothesis,
118 formulating new questions—and guiding efforts to collect new data. Here, we illustrate how such
119 a workflow could lead to significant improvements in two case studies: (i) estimating global bio-
120 diversity trends and (ii) forecasting future species and ecosystem dynamics using process-based
121 models.

122 **3.1 Trends**

123 Ecologists today have amassed data on populations and species across the globe; they have also
124 engaged in an increasing number of debates on regional and global trends over time, with arguments
125 over the magnitude and even direction of population and biodiversity metrics (Dornelas *et al.*, 2014;
126 Leung *et al.*, 2020; Terry *et al.*, 2022; Müller *et al.*, 2024). While shifting estimates are part of the
127 process of science—refining our approaches and thus estimates over time—we believe these debates
128 would be fewer and they would be more rapidly resolved through use of an improved workflow.

129 An improved workflow that required data simulation and retrodictive checks would lead to
130 larger model advances and a greater recognition of uncertainty—thus highlighting likely consistency
131 in estimates across models—that could better aid policy. Using the workflow would make it more
132 obvious that what now appear as major discrepancies are shifts in point estimates that are generally
133 all in the same uncertainty space (Johnson *et al.*, 2024)—and it would challenge modelers to show
134 major predictive advances, which is not currently part of the process. Explanatory power in most
135 models of observational data is usually very low (Low-Décarie *et al.*, 2014; Møller & Jennions, 2002)
136 and thus tests of models’ predictions rarely expected. But the workflow highlights that predictions
137 from the model—what we call retrodictive checks—are part of the process of science, and critical
138 to testing for what may be missing in a model. We expect retrodictive checks on most published
139 trend analyses would highlight major missing components in these models, and drive changes both
140 in the models themselves and in the simulated data to check the models. This step builds somewhat
141 on the skills needed for null models (Gotelli & Ulrich, 2012), but with a shift in focus towards the
142 specific bespoke model at hand. Ecologists have started to use simulated data more to understand
143 potential limitations of their models and data combined, but this is still extremely rare, and efforts
144 to date often treat simulations as separate from the statistical model (Buschke *et al.*, 2021; Dove
145 *et al.*, 2023), short-circuiting their full utility if used in an iterative workflow.

146 Applying the workflow to current trend estimates could importantly highlight the best way to
147 improve data collection for more reliable estimates. Returning to the example of a global estimate of
148 trends in vertebrate populations of species over time (see Box) and applying our proposed workflow
149 would mean more efforts to define the goal and question—is it a simple global estimate? Or a
150 need to also find which species are declining most, including those that may have poor or no data?
151 From there a generative model using simulated data for testing could incorporate many aspects
152 of the populations, and data, that are often only included in ‘null’ or ‘synthetic data generation’
153 currently (Buschke *et al.*, 2021; McRae *et al.*, 2025) but could be built into the models fit to the
154 empirical data. Eventually fitting the empirical data and performing retrodictive checks would
155 likely highlight major missing components of the generative model and, ultimately, this would
156 help inform our global estimates of mean trends (see workflow example). For example, certain
157 populations are recovering for very specific reasons (e.g., elephants in regions where the ivory trade
158 drove declines in the past) that perhaps should be modeled. From this model, what data are most
159 critically needed to address the updated aims would become clearer and could drive new data
160 collection (Toszogyova *et al.*, 2024).

161 **3.2 Forecasting from process-based models**

162 Ecological forecasting is a broad field with a diverse range of methods. Our second case-study
163 focuses on process-based modeling, which is often considered the gold standard for forecasting in
164 ecology (Urban *et al.*, 2016; Pilowsky *et al.*, 2022) and beyond. Newer models, however, generally
165 incorporate greater complexity (and an ever-growing number of parameters), which can make it

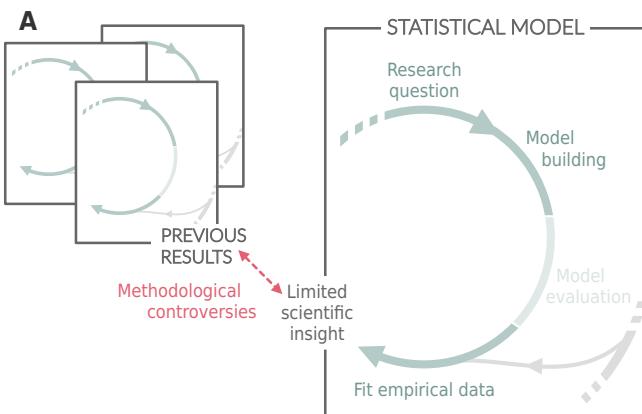
166 difficult to increase scientific understanding (Franklin *et al.*, 2020), and suggest or model potential
167 policies.

168 Increasing model complexity can be beneficial, especially when it reduces uncertainty; however,
169 this is not always the outcome. In process-based modeling for climate forecasts, the uncertainty
170 range on the effect of increasing CO₂ concentration on temperature have remained largely un-
171 changed (Zelinka *et al.*, 2020). This has driven calls for more rigorous and transparent calibration
172 processes (Balaji *et al.*, 2022). Similar concerns arise in ecology, where strong disagreements exist
173 about the effect of climate change on future species distributions (Cheaib *et al.*, 2012) and ecosys-
174 tem dynamics (Lovenduski & Bonan, 2017). These uncertainties have large implications beyond
175 ecology, as they influence simulations of biosphere-atmosphere interactions and, ultimately, future
176 climate projections (Bonan & Doney, 2018; Simpson *et al.*, 2025). Some researchers now advocate
177 for simplifying models, to avoid over-parametrization when the data provide little information to
178 constrain some parameters (Wang *et al.*, 2017; Harrison *et al.*, 2021). If a model becomes too
179 complex, it may become a black box, and understanding the sources of uncertainty and how they
180 propagate through the model may become nearly impossible (see Box). Each additional process
181 and parameter can increase overall uncertainty to the point where model projections lose their
182 usefulness for decision makers (Saltelli *et al.*, 2020).

183 Process-based models used to project species or ecosystem distributions highlight some of these
184 problems. Most of the focus is on the model projections—often represented as maps without
185 uncertainties. Model evaluation generally focuses on how well these projections match observed
186 distributions, sometimes under different climatic conditions to challenge the models (Van der Meer-
187 sch *et al.*, 2025)—most of the time evaluating only one of the many different output variables of
188 the model. The complexity of these models can make it difficult to assess how well the potential
189 parameter space was searched, and whether there was any potential non-identifiability—such work
190 is generally done as a separate effort, exposing potential model problems later on (Van der Meersch
191 & Chuine, 2025), and not required as a preliminary step.

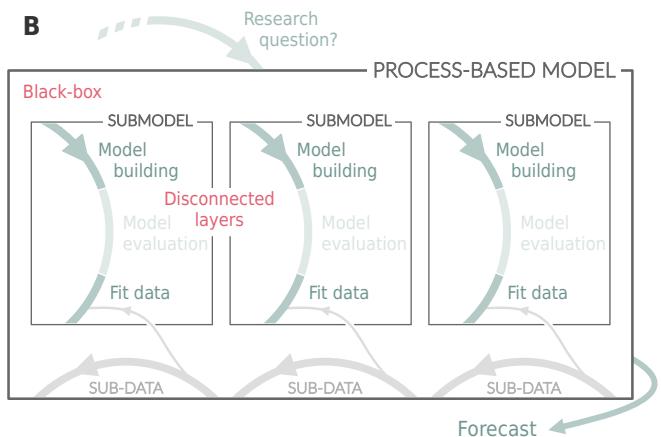
192 Applying the workflow to process-based models would help open the black box. Each successive
193 step of model development in the workflow may highlight current problems and a path to solutions.
194 Incorporating data simulation would introduce a crucial step between model building and data
195 fitting (also called calibration), ensuring a clear delineation between the two, and help expose
196 potential parameter identifiability issues in the model design. Uncovering identifiability issues
197 would likely force researchers to begin with a simpler version of the model, which they could
198 build on iteratively, testing for support—or lack thereof—when adding model complexity. The
199 workflow may also highlight model problems by requiring more explicit model calibration (i.e.
200 data fitting), which is currently hidden within opaque ‘model building’, making it easy to hide
201 non-identifiability in the model. Through the workflow non-identifiable parts of models could be
202 addressed by reformulating the mathematical structure of certain processes, or finding ways to apply
203 additional constraints (e.g., narrower ranges for certain parameters or developing new hypotheses
204 that target the appropriate level of complexity). The resulting process-based models would likely be
205 simpler and thus more tractable for quantifying parameter uncertainty and propagating it through
206 projections.

Trend estimation



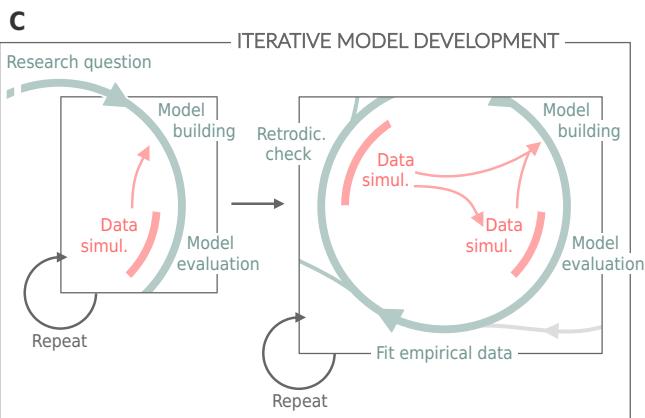
In the current workflow for estimating trends over time, a new model with a new estimate often leads to a paper (see figure A above) because ecologists spend less time interrogating their models with simulated data, or their model performance fit to empirical data. The Living Planet Index (LPI, www.livingplanetindex.org), which aims to include long-term data on vertebrate populations of species across the globe, is emblematic of these conflicting results. With updated data released semi-annually alongside new datasets and new estimates of decline, a growing number of high-profile papers have challenged how strong the evidence is for population decline (Dornelas *et al.*, 2014; Gonzalez *et al.*, 2016; Wagner *et al.*, 2021; Müller *et al.*, 2024), with each paper taking a slightly different analytical approach. For examples using LPI data, Leung *et al.* (2020) published a mixture model that suggested most populations were not significantly declining, followed by other alternative modeling approaches (Buschke *et al.*, 2021; Puurinen *et al.*, 2022) including a recent one suggesting a basic analysis of the dataset should always include three sources of autocorrelation, finding trends that encompassed many previous results (Johnson *et al.*, 2024).

Mechanistic forecasting



Model development is the central step of the process-based workflow, typically requiring several years, yet it often remains opaque for anyone who has not worked the model itself. The step of designing the model—translating knowledge and hypotheses into mathematical equations and parameters—is often blurred with the step of model calibration (or tuning), where parameter values are inferred. Models are often treated as an accumulation of multiple submodels, each governing one or several ecological processes (see figure B above). Rather than being fitted as a whole, submodels are calibrated separately against specific subsets of data, and some parameters are simply prescribed (i.e., fixed to a value found in the literature) or tuned to reproduce certain observations or theory. The way models are currently calibrated is likely not a coincidence, but rather an inappropriate way to accommodate their complexity, where many parameters compensate for one another.

A common workflow to bridge trend estimation and forecasting



A universal workflow offers an opportunity to bridge statistical and process-based frameworks, integrating mechanistic knowledge and leveraging robust statistical approaches (e.g. Rounce *et al.*, 2020). Process-based models would no longer be perceived as deterministic black boxes by other researchers but rather as robust statistical frameworks encapsulating both data structure and mechanistic knowledge. It would be an opportunity to spread the incorporation of mechanistic assumptions beyond the process-based modeling community, potentially improving trend estimates. For both trend estimation and forecasting, the workflow would refocus attention on the research question, highlighting the ecological hypotheses that justify the use and design of the model.

208 **4 Barriers and opportunities**

209 We believe our workflow could help advance ecological science and its applications, but widespread
210 use of it requires overcoming major hurdles that pervade science. The first is pressure to publish
211 quickly, which can be at odds with the reality that good model fitting is inherently iterative and
212 takes time. The second is reluctance to embrace adoption of open science practices that ensure
213 modeling efforts are fully transparent and reproducible. We believe adopting a workflow such as
214 the one we propose may change this reluctance, and aligns with ongoing efforts to place increasing
215 value on research that is carefully developed, openly collaborative (including both data and code),
216 and transparent about areas of uncertainty.

217 **4.1 Adopting the workflow**

218 Advancing ecology to where most researchers use models built more flexibly from ecological theory
219 and insights applied to their ecological systems will not happen rapidly without a major shift in
220 training. Much of ecology still divides the world into training for those who gather data and learn
221 a limited set of pre-built models versus those who develop more complex models. In ecological
222 training today, researchers who conduct field and lab studies often learn a limited set of particular
223 statistical tests matched to particular experiment designs and simple information on their variable
224 types (e.g., categorical x and y leads to using a chi-squared test).

225 When ecologists trained in a limited set of tests require more complex models, they are expected
226 to collaborate with others trained more in model development (though often for highly specific ap-
227 plications, such as wildlife population estimates, where generative models are often predefined).
228 These two groups further differ from process-based modelers, who often train in physical and eco-
229 physiological processes and how to abstract them into mainly deterministic models. Neither group
230 has fully integrated data simulation into their statistical or scientific workflows. Simulation is gen-
231 erally reserved as a form of training needed mainly by those specializing in theoretical ecology, who
232 often solve analytical equations but rarely link to empirical data. While specialization is valuable,
233 we argue the fundamental separation in ecology has overly-siloed these groups and prevented more
234 rapid progress.

235 We suggest unified training in our proposed workflow, or a similar one (Betancourt, 2020;
236 Gelman *et al.*, 2020; Grinsztajn *et al.*, 2021; van de Schoot *et al.*, 2021), would focus on learning
237 to generate questions and then models, and then how to simulate data from them. Through this
238 and the use retrodictive checks, most ecologists would be better equipped to think through what
239 parameters are most critical to their question and/or aim (e.g., management), and also gain a much
240 stronger connection to the level of uncertainty in many of ecological estimates. Empiricists would
241 be more likely to recognize critical gaps in current models fit by those specializing in ecological
242 modeling and help advance those models. Process-based modelers may start a new generation of
243 simpler models that are more tractable to theoretical ecologists, who may see new bridges from
244 their work to empirical data and forecasting.

245 **4.2 A tractable alternative to machine learning**

246 With rapid advances in machine learning, improving current methods to gain greater scientific
247 insights to drive better forecasting seems increasingly important. Machine learning will likely
248 surpass process-based models for forecasting accuracy if the latter lack a robust estimation of their
249 parameters and fall in a complexity trap, at the cost of their interpretability. Similarly, estimates

250 of trends from empirical data using models without clear mechanistic drivers may soon offer fewer
251 advantages over machine learning.

252 Beyond improving model building and evaluation, our proposed workflow also has the po-
253 tential to shift how process-based models are perceived, particularly by those unfamiliar with
254 them. Process-based models could once again be a way to answer a research question—whereas
255 today, model simulations have increasingly become a subject of study on their own. Ideally, ap-
256 plying the workflow would help to move away from the traditional process-based model paradigm,
257 where parameters are typically assigned fixed values without properly accounting for their uncer-
258 tainty. Instead, it would guide a step-by-step model fitting, parameter estimation, and uncertainty
259 quantification—preventing modelers from making biased inferences and unfounded assumptions
260 beyond what the data can support. It would thus define a clear and limited context in which the
261 model should apply, and limit discussion of adding increasing complexity.

262 *Conclusions:* Today model development in ecology is rarely transparent, which limits how easily
263 the research community can understand models, and thus identify potential issues. Instead of broad
264 inclusive conversations about how to improve models to advance our ecological understanding, a
265 significant portion of scientific debate has become mired in methodological considerations. However,
266 we believe our workflow provides a tractable step to fixing this. By focusing on model development
267 more tightly tied to ecological expertise, we argue this workflow should broaden the community
268 that contributes to model development. It may also help resolve apparent conflicts by identifying
269 where divergences in model predictions emerge, whether from differences in assumptions, model
270 structures, or other aspects of the modeling process. As ecologists are increasingly expanding their
271 computational toolkits, many field, and lab, and other forms of ‘empirical’ ecologists have the
272 basic tools to follow this workflow to build models that better represent their ecological domains
273 of interest, and—most importantly—to interrogate those models.

274 **References**

- 275 Balaji, V., Couvreux, F., Deshayes, J., Gautrais, J., Hourdin, F. & Rio, C. (2022) Are general
276 circulation models obsolete? *Proceedings of the National Academy of Sciences* **119**, e2202075119.
- 277 Betancourt, M. (2020) Towards A Principled Bayesian Workflow. https://betanalpha.github.io/assets/case_studies/principled_bayesian_workflow.html.
- 278
- 279 Bonan, G.B. & Doney, S.C. (2018) Climate, ecosystems, and planetary futures: The challenge to
280 predict life in earth system models. *Science* **359**.
- 281 Buschke, F.T., Hagan, J.G., Santini, L. & Coetzee, B.W.T. (2021) Random population fluctuations
282 bias the living planet index. *Nature Ecology & Evolution* **5**, 1145–1152.
- 283 Chamberlin, T.C. (1965) Method of multiple working hypotheses. *Science* **148**, 754.
- 284 Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., François, C., Gritti, E.S.,
285 Legay, M., Pagé, C., Thuiller, W., Viovy, N. & Leadley, P. (2012) Climate change impacts on
286 tree ranges: model intercomparison facilitates understanding and quantification of uncertainty.
287 *Ecology Letters* **15**, 533–544.
- 288 Díaz, S., Settele, J., Brondízio, E.S., Ngo, H.T., Agard, J., Arneth, A., Balvanera, P., Brauman,
289 K.A., Butchart, S.H.M., Chan, K.M.A., Garibaldi, L.A., Ichii, K., Liu, J., Subramanian, S.M.,
290 Midgley, G.F., Miloslavich, P., Molnár, Z., Obura, D., Pfaff, A., Polasky, S., Purvis, A., Razzaque,
291 J., Reyers, B., Chowdhury, R.R., Shin, Y.J., Visseren-Hamakers, I., Willis, K.J. & Zayas, C.N.
292 (2019) Pervasive human-driven decline of life on earth points to the need for transformative
293 change. *Science* **366**.
- 294 Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D., Akhmetzhanova,
295 A.A., Appeltans, W., Arcos, J.M., Arnold, H., Ayyappan, N., Badihi, G., Baird, A.H., Barbosa,
296 M., Barreto, T.E., Bässler, C., Bellgrove, A., Belmaker, J., Benedetti-Cecchi, L., Bett, B.J.,
297 Bjorkman, A.D., Błażewicz, M., Blowes, S.A., Bloch, C.P., Bonebrake, T.C., Boyd, S., Bradford,
298 M., Brooks, A.J., Brown, J.H., Bruelheide, H., Budy, P., Carvalho, F., Castañeda-Moya, E.,
299 Chen, C.A., Chamblee, J.F., Chase, T.J., Siegwart Collier, L., Collinge, S.K., Condit, R., Cooper,
300 E.J., Cornelissen, J.H.C., Cotano, U., Kyle Crow, S., Damasceno, G., Davies, C.H., Davis, R.A.,
301 Day, F.P., Degraer, S., Doherty, T.S., Dunn, T.E., Durigan, G., Duffy, J.E., Edelist, D., Edgar,
302 G.J., Elahi, R., Elmendorf, S.C., Enemar, A., Ernest, S.K.M., Escribano, R., Estiarte, M., Evans,
303 B.S., Fan, T., Turini Farah, F., Loureiro Fernandes, L., Farneda, F.Z., Fidelis, A., Fitt, R., Fosaa,
304 A.M., Daher Correa Franco, G.A., Frank, G.E., Fraser, W.R., García, H., Cazzolla Gatti, R.,
305 Givan, O., Gorgone-Barbosa, E., Gould, W.A., Gries, C., Grossman, G.D., Gutierrez, J.R., Hale,
306 S., Harmon, M.E., Harte, J., Haskins, G., Henshaw, D.L., Hermanutz, L., Hidalgo, P., Higuchi, P.,
307 Hoey, A., Van Hoey, G., Hofgaard, A., Holeck, K., Hollister, R.D., Holmes, R., Hoogenboom, M.,
308 Hsieh, C., Hubbell, S.P., Huettmann, F., Huffard, C.L., Hurlbert, A.H., Macedo Ivanauskas, N.,
309 Janík, D., Jandt, U., Jaźdżewska, A., Johannessen, T., Johnstone, J., Jones, J., Jones, F.A.M.,
310 Kang, J., Kartawijaya, T., Keeley, E.C., Kelt, D.A., Kinnear, R., Klanderud, K., Knutson, H.,
311 Koenig, C.C., Kortz, A.R., Král, K., Kuhnz, L.A., Kuo, C., Kushner, D.J., Laguionie-Marchais,
312 C., Lancaster, L.T., Min Lee, C., Lefcheck, J.S., Lévesque, E., Lightfoot, D., Lloret, F., Lloyd,
313 J.D., López-Baucells, A., Louzao, M., Madin, J.S., Magnússon, B., Malamud, S., Matthews,
314 I., McFarland, K.P., McGill, B., McKnight, D., McLarney, W.O., Meador, J., Meserve, P.L.,

315 Metcalfe, D.J., Meyer, C.F.J., Michelsen, A., Milchakova, N., Moens, T., Moland, E., Moore, J.,
316 Mathias Moreira, C., Müller, J., Murphy, G., Myers-Smith, I.H., Myster, R.W., Naumov, A.,
317 Neat, F., Nelson, J.A., Paul Nelson, M., Newton, S.F., Norden, N., Oliver, J.C., Olsen, E.M.,
318 Onipchenko, V.G., Pabis, K., Pabst, R.J., Paquette, A., Pardede, S., Paterson, D.M., Pélassier,
319 R., Peñuelas, J., Pérez-Matus, A., Pizarro, O., Pomati, F., Post, E., Prins, H.H.T., Priscu, J.C.,
320 Provoost, P., Prudic, K.L., Pulliainen, E., Ramesh, B.R., Mendivil Ramos, O., Rassweiler, A.,
321 Rebelo, J.E., Reed, D.C., Reich, P.B., Remillard, S.M., Richardson, A.J., Richardson, J.P., van
322 Rijn, I., Rocha, R., Rivera-Monroy, V.H., Rixen, C., Robinson, K.P., Ribeiro Rodrigues, R.,
323 de Cerqueira Rossa-Feres, D., Rudstam, L., Ruhl, H., Ruz, C.S., Sampaio, E.M., Rybicki, N.,
324 Rypel, A., Sal, S., Salgado, B., Santos, F.A.M., Savassi-Coutinho, A.P., Scanga, S., Schmidt,
325 J., Schooley, R., Setiawan, F., Shao, K., Shaver, G.R., Sherman, S., Sherry, T.W., Siciński, J.,
326 Sievers, C., da Silva, A.C., Rodrigues da Silva, F., Silveira, F.L., Slingsby, J., Smart, T., Snell,
327 S.J., Soudzilovskaia, N.A., Souza, G.B.G., Maluf Souza, F., Castro Souza, V., Stallings, C.D.,
328 Stanforth, R., Stanley, E.H., Mauro Sterza, J., Stevens, M., Stuart-Smith, R., Rondon Suarez,
329 Y., Supp, S., Yoshio Tamashiro, J., Tarigan, S., Thiede, G.P., Thorn, S., Tolvanen, A., Teresa
330 Zugliani Toniato, M., Totland, Ø., Twilley, R.R., Vaitkus, G., Valdivia, N., Vallejo, M.I., Valone,
331 T.J., Van Colen, C., Vanaverbeke, J., Venturoli, F., Verheyen, H.M., Vianna, M., Vieira, R.P.,
332 Vrška, T., Quang Vu, C., Van Vu, L., Waide, R.B., Waldoch, C., Watts, D., Webb, S., Wesolowski,
333 T., White, E.P., Widdicombe, C.E., Wilgers, D., Williams, R., Williams, S.B., Williamson, M.,
334 Willig, M.R., Willis, T.J., Wipf, S., Woods, K.D., Woehler, E.J., Zawada, K. & Zettler, M.L.
335 (2018) Biotime: A database of biodiversity time series for the anthropocene. *Global Ecology and*
336 *Biogeography* **27**, 760–786.

337 Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C. & Magurran, A.E.
338 (2014) Assemblage time series reveal biodiversity change but not systematic loss. *Science* **344**,
339 296–299.

340 Dove, S., Böhm, M., Freeman, R., McRae, L. & Murrell, D.J. (2023) Quantifying reliability and data
341 deficiency in global vertebrate population trends using the living planet index. *Global Change*
342 *Biology* **29**, 4966–4982.

343 Franklin, O., Harrison, S.P., Dewar, R., Farrior, C.E., Brännström, A., Dieckmann, U., Pietsch, S.,
344 Falster, D., Cramer, W., Loreau, M., Wang, H., Mäkelä, A., Rebel, K.T., Meron, E., Schymanski,
345 S.J., Rovenskaya, E., Stocker, B.D., Zaehle, S., Manzoni, S., van Oijen, M., Wright, I.J., Ciais,
346 P., van Bodegom, P.M., Peñuelas, J., Hofhansl, F., Terrer, C., Soudzilovskaia, N.A., Midgley, G.
347 & Prentice, I.C. (2020) Organizing principles for vegetation dynamics. *Nature Plants* **6**, 444–453.

348 Gelman, A., Vehtari, A., Simpson, D., Margossian, C.C., Carpenter, B., Yao, Y., Kennedy, L.,
349 Gabry, J., Bürkner, P.C. & Modrák, M. (2020) Bayesian workflow.

350 Gonzalez, A., Cardinale, B.J., Allington, G.R., Byrnes, J., Arthur Endsley, K., Brown, D.G.,
351 Hooper, D.U., Isbell, F., O'Connor, M.I. & Loreau, M. (2016) Estimating local biodiversity
352 change: a critique of papers claiming no net loss of local diversity. *Ecology* **97**, 1949–1960.

353 Gotelli, N.J. & Ulrich, W. (2012) Statistical challenges in null model analysis. *Oikos* **121**, 171–180,
354 gotelli, Nicholas J. Ulrich, Werner.

355 Grinsztajn, L., Semenova, E., Margossian, C.C. & Riou, J. (2021) Bayesian workflow for disease
356 transmission modeling in stan. *Statistics in Medicine* **40**, 6209–6234.

- 357 Harrison, S.P., Cramer, W., Franklin, O., Prentice, I.C., Wang, H., Bränström, A., de Boer, H.,
358 Dieckmann, U., Joshi, J., Keenan, T.F., Lavergne, A., Manzoni, S., Mengoli, G., Morfopoulos,
359 C., Peñuelas, J., Pietsch, S., Rebel, K.T., Ryu, Y., Smith, N.G., Stocker, B.D. & Wright, I.J.
360 (2021) Eco-evolutionary optimality as a means to improve vegetation and land-surface models.
361 *New Phytologist* **231**, 2125–2141.
- 362 Hilborn, R. & Mangel, M. (1997) *The Ecological Detective: Confronting Models with Data*. Prince-
363 ton University Press.
- 364 IPBES (2019) Global assessment report on biodiversity and ecosystem services of the intergovern-
365 mental science-policy platform on biodiversity and ecosystem services. Tech. rep.
- 366 Johnson, T.F., Beckerman, A.P., Childs, D.Z., Webb, T.J., Evans, K.L., Griffiths, C.A., Capdevila,
367 P., Clements, C.F., Besson, M., Gregory, R.D., Thomas, G.H., Delmas, E. & Freckleton, R.P.
368 (2024) Revealing uncertainty in the status of biodiversity change. *Nature* **628**, 788–794.
- 369 Leung, B., Hargreaves, A.L., Greenberg, D.A., McGill, B., Dornelas, M. & Freeman, R. (2020)
370 Clustered versus catastrophic global vertebrate declines. *Nature* **588**, 267–271.
- 371 Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. & Randers, J. (2005)
372 The living planet index: using species population time series to track trends in biodiversity.
373 *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 289–295.
- 374 Loreau, M., Cardinale, B.J., Isbell, F., Newbold, T., O'Connor, M.I. & de Mazancourt, C. (2022)
375 Do not downplay biodiversity loss. *Nature* **601**, E27–E28.
- 376 Lovenduski, N.S. & Bonan, G.B. (2017) Reducing uncertainty in projections of terrestrial carbon
377 uptake. *Environmental Research Letters* **12**, 044020.
- 378 Low-Décarie, E., Chivers, C. & Granados, M. (2014) Rising complexity and falling explanatory
379 power in ecology. *Frontiers in Ecology and the Environment* **12**, 412–418.
- 380 McElreath, R. (2018) *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*.
381 Chapman and Hall/CRC.
- 382 McRae, L., Cornford, R., Marconi, V., Puleston, H., Ledger, S.E., Deinet, S., Oppenheimer, P.,
383 Hoffmann, M. & Freeman, R. (2025) The utility of the living planet index as a policy tool and
384 for measuring nature recovery. *Philosophical Transactions B* **380**, 20230207.
- 385 Møller, A. & Jennions, M.D. (2002) How much variance can be explained by ecologists and evolu-
386 tionary biologists? *Oecologia* **132**, 492–500.
- 387 Müller, J., Hothorn, T., Yuan, Y., Seibold, S., Mitesser, O., Rothacher, J., Freund, J., Wild, C.,
388 Wolz, M. & Menzel, A. (2024) Weather explains the decline and rise of insect biomass over 34
389 years. *Nature* **628**, 349–354.
- 390 Nosek, B.A., Ebersole, C.R., DeHaven, A.C. & Mellor, D.T. (2018) The preregistration revolution.
391 *Proceedings of the National Academy of Sciences* **115**, 2600–2606.
- 392 Pilowsky, J.A., Colwell, R.K., Rahbek, C. & Fordham, D.A. (2022) Process-explicit models reveal
393 the structure and dynamics of biodiversity patterns. *Science Advances* **8**, eabj2271.

- 394 Puurtinen, M., Elo, M. & Kotiaho, J.S. (2022) The living planet index does not measure abundance.
395 *Nature* **601**, E14–E15.
- 396 Rounce, D.R., Khurana, T., Short, M.B., Hock, R., Shean, D.E. & Brinkerhoff, D.J. (2020) Quantifying
397 parameter uncertainty in a large-scale glacier evolution model using bayesian inference:
398 application to high mountain asia. *Journal of Glaciology* **66**, 175–187.
- 399 Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W.,
400 Kay, J., Lo Piano, S., Mayo, D., Pielke Jr, R., Portaluri, T., Porter, T.M., Puy, A., Rafols, I.,
401 Ravetz, J.R., Reinert, E., Sarewitz, D., Stark, P.B., Stirling, A., van der Sluijs, J. & Vineis, P.
402 (2020) Five ways to ensure that models serve society: a manifesto. *Nature* **582**, 482–484.
- 403 Schad, D.J., Betancourt, M. & Vasishth, S. (2020) Toward a principled bayesian workflow in cog-
404 nitive science.
- 405 Simpson, I.R., Shaw, T.A., Ceppi, P., Clement, A.C., Fischer, E., Grise, K.M., Pendergrass, A.G.,
406 Screen, J.A., Wills, R.C., Woollings, T. *et al.* (2025) Confronting earth system model trends with
407 observations. *Science advances* **11**, eadt8035.
- 408 Talis, E.J. & Lynch, H.J. (2023) Capturing stochasticity properly is key to understanding the
409 nuances of the living planet index. *Nature Ecology & Evolution* **7**, 1194–1195.
- 410 Terry, J.C.D., O’Sullivan, J.D. & Rossberg, A.G. (2022) No pervasive relationship between species
411 size and local abundance trends. *Nature Ecology & Evolution* **6**, 140–144.
- 412 Thuiller, W., Guéguen, M., Renaud, J., Karger, D.N. & Zimmermann, N.E. (2019) Uncertainty in
413 ensembles of global biodiversity scenarios. *Nature Communications* **10**.
- 414 Toszogyova, A., Smyčka, J. & Storch, D. (2024) Mathematical biases in the calculation of the living
415 planet index lead to overestimation of vertebrate population decline. *Nature Communications* **15**,
416 5295.
- 417 Urban, M.C., Bocedi, G., Hendry, A.P., Mihoub, J.B., Pe’er, G., Singer, A., Bridle, J.R., Crozier,
418 L.G., De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J.J., Holt, R.D., Huth, A., Johst,
419 K., Krug, C.B., Leadley, P.W., Palmer, S.C.F., Pantel, J.H., Schmitz, A., Zollner, P.A. & Travis,
420 J.M.J. (2016) Improving the forecast for biodiversity under climate change. *Science* **353**, aad8466.
- 421 van de Schoot, R., Depaoli, S., King, R., Kramer, B., Maertens, K., Tadesse, M.C., Vannucci, M.,
422 Gelman, A., Veen, D., Willemse, J. & Yau, C. (2021) Bayesian statistics and modelling. *Nature
423 Reviews Methods Primers* **1**.
- 424 Van der Meersch, V., Armstrong, E., Mouillot, F., Duputié, A., Davi, H., Saltré, F. & Chuine, I.
425 (2025) Paleorecords reveal biological mechanisms crucial for reliable species range shift projec-
426 tions amid rapid climate change. *Ecology Letters* **28**.
- 427 Van der Meersch, V. & Chuine, I. (2025) Can inverse calibration help improving process-explicit
428 species distribution models? *Ecological Modelling* **506**, 111132.
- 429 Wagner, D.L., Grames, E.M., Forister, M.L., Berenbaum, M.R. & Stopak, D. (2021) Insect decline
430 in the anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*
431 **118**, e2023989118.

- 432 Wang, H., Prentice, I.C., Keenan, T.F., Davis, T.W., Wright, I.J., Cornwell, W.K., Evans, B.J. &
433 Peng, C. (2017) Towards a universal model for carbon dioxide uptake by plants. *Nature Plants*
434 **3**, 734–741.
- 435 Wolkovich, E., Davies, T.J., Pearse, W.D. & Betancourt, M. (2024) A four-step bayesian workflow
436 for improving ecological science.
- 437 Zelinka, M.D., Myers, T.A., McCoy, D.T., Po-Chedley, S., Caldwell, P.M., Ceppi, P., Klein, S.A. &
438 Taylor, K.E. (2020) Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research
439 Letters* **47**.