Coding for Life: Designing a Platform for Projecting and Protecting Global Biodiversity

MARK C. URBAN, JUSTIN M. J. TRAVIS, DAMARIS ZURELL, PATRICK L. THOMPSON, NICHOLAS W. SYNES, ALICE SCARPA, PEDRO R. PERES-NETO, ANNE-KATHLEEN MALCHOW, PATRICK M. A. JAMES, DOMINIQUE GRAVEL, LUC DE MEESTER, CALUM BROWN, GRETA BOCEDI, CÉCILE H. ALBERT, ANDREW GONZALEZ, AND ANDREW P. HENDRY

Time is running out to limit further devastating losses of biodiversity and nature’s contributions to humans. Addressing this crisis requires accurate predictions about which species and ecosystems are most at risk to ensure efficient use of limited conservation and management resources. We review existing biodiversity projection models and discover problematic gaps. Current models usually cannot easily be reconfigured for other species or systems, omit key biological processes, and cannot accommodate feedbacks with Earth system dynamics. To fill these gaps, we envision an adaptable, accessible, and universal biodiversity modeling platform that can project essential biodiversity variables, explore the implications of divergent socioeconomic scenarios, and compare conservation and management strategies. We design a roadmap for implementing this vision and demonstrate that building this biodiversity forecasting platform is possible and practical.

Keywords: biodiversity, forecasting, global change, modeling, mechanistic

Accelerating threats from climate change, habitat degradation, overexploitation, and species invasions threaten biodiversity worldwide (Ceballos et al. 2015, Urban 2015). These threats are reorganizing biological communities, threatening a million species with extinction, and altering ecosystems through loss of key species and altered nutrient and energy flows (Ceballos et al. 2015, Urban 2015, IPBES 2019). The resultant biodiversity loss and ecosystem collapse are reducing nature’s contributions to human health, wellbeing, and economy (Costanza et al. 2014) and causing a growing sense that humankind has surpassed the planetary boundaries for maintaining life on Earth (Rockström et al. 2009). Therefore, protecting and restoring biodiversity constitutes one of the greatest challenges for science in the twenty-first century.

The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES 2019) recently issued a dire assessment of global biodiversity and the efforts required to protect it. The assessment concluded that efforts to conserve biodiversity are falling far short of international goals and needs. Furthermore, the IPBES expressed low confidence in the current capacity to project biodiversity changes and their responses to divergent future scenarios and mitigation strategies (IPBES 2019). Unlike climate science, biologists have not devoted substantial resources to developing shared and comprehensive modeling frameworks to project future biodiversity change (Urban 2019). Therefore, the world lacks the predictive infrastructure needed to address a rapidly accelerating biodiversity crisis.

An important action urgently needed to prevent further biodiversity loss entails developing accurate models to forecast future biodiversity change, highlight data needs, guide effective conservation strategies, and prioritize conservation of the most threatened species and ecosystems. Owing to the manifold, multiscale, and nonlinear ways humans disrupt nature (Gilman et al. 2010), process-based modeling efforts are particularly needed to unravel the complex feedbacks between threats and biodiversity responses and reveal unrecognized threats to biodiversity. Such knowledge can inform effective conservation strategies and prevent wasting limited resources (Barbier et al. 2018) on otherwise resilient species and ecosystems (Parmesan 2014). Whereas most conservation efforts currently respond to short-term threats (Baillie et al. 2004, Pereira et al. 2013), conservation also needs to focus on evidence-based, proactive measures that prevent biodiversity from becoming critically endangered in the first place.

We review current efforts to model, project, and mitigate biodiversity loss and find critical deficiencies in modeling efforts and forecasting accuracy that increase uncertainty...
and contribute to inaction. To address these shortcomings, we design and present a comprehensive platform for predicting and preventing biodiversity loss, define essential standards, and outline practical recommendations for effective implementation. We argue that investing in a biodiversity projection platform now would facilitate the design of strategies that protect most of the remaining biodiversity and critical ecosystem services despite accelerating threats.

**Current efforts to predict nature**

To understand the current state of biodiversity prediction, we review models readily accessible to scientists, conservation professionals, managers, and policymakers for projecting future biodiversity change under different human disturbance scenarios by gathering information via a systematic keyword literature search (see the supplemental material), expert knowledge, and published reviews (Hoban et al. 2012, Evans et al. 2013, Lurgi et al. 2015, Cabral et al. 2017, Norberg et al. 2019). We define biodiversity projection models as algorithms that project biological responses to external drivers such as land use and greenhouse gas emissions. We include models that predict a wide range of biological responses, including genetics, traits, population abundances, species diversity, and ecosystem properties. These responses are modeled in ways that range from highly complex and mechanistic (e.g., physiological models) to simple and correlational (e.g., species distribution models), which are then altered to explore future outcomes. We exclude system-specific frameworks that inform particular questions or species but cannot be easily modified to address others, but acknowledge their important insights for individual questions.

We found 50 models that met our search criteria (see the supplemental material). Two evaluators independently read pertinent publications and manuals on each model, categorized them along multiple dimensions, and then resolved any conflicts. We evaluated the degree to which models incorporated six key biological processes that enhance realism and predictive accuracy (Gilman et al. 2010, Hoffmann and Sgro 2011, Urban et al. 2016), including physiology, demography, dispersal, species interactions, evolution, and other responses to environmental variation (Urban et al. 2016). We next evaluated if model outputs aligned with the essential biodiversity variables (EBVs) developed to define key measurements needed for global monitoring efforts (Pereira et al. 2013). We also evaluated the degree to which inexperienced users can adapt models to new species, systems, and questions.

We found that most biodiversity models omit key biological processes (figure 1) such as species-specific dispersal, biotic interactions, or adaptations that could mediate biodiversity responses to perturbations such as climate change (Buckley et al. 2010, Gilman et al. 2010, Urban et al. 2016). However, models lacking these key processes are routinely used to inform decision-making. The most popular and accessible approaches apply species distribution models (e.g., Maxent-based approaches; Phillips et al. 2006), to extrapolate correlations between contemporary species distributions and environments to project future responses. Phenomenological models that use algorithms such as Maxent are simpler to fit with existing data and can perform sufficiently well for short time horizons or when little is known about an organism's biology. These simple models have been widely adopted by managers that need readily accessible tools to make conservation decisions. However, models that incorporate even limited biological information generally outperform correlative approaches over longer time horizons and as underlying mechanisms become more diverse or interactive (Zurell et al. 2016). For example, mechanistic models consistently predicted species' range dynamics over longer horizons, whereas statistical models became increasingly inaccurate (Pagel and Schurr 2012).

Despite their many advantages, mechanistic models remain underused. Existing mechanistic models usually focus on one or a few key processes (figure 1), with a few exceptions incorporating multiple biological mechanisms, including sPEGG (Okamoto and Amarasekare 2017), Nemo (Guillaume and Rougemont 2006), and RangeShifter (Bocedi et al. 2014, Bocedi et al. 2021). Many mechanistic models cater to specific taxa (e.g., trees), include few essential processes, or require specialized programming skills to modify them (figure 1). General mechanistic models of biomass and energy have been created that depict ecosystem functioning (Harfoot et al. 2014), but their outputs cannot easily be resolved into finer biodiversity details, such as species abundances, diversity, and interactions.

Most biodiversity models address specific questions for particular species and ecosystems and do not interact with each other, nor are they easily modified to apply to other systems. Such models were not designed to interface with Earth system models. Consequently, existing biodiversity models cannot be applied effectively to investigate large-scale and dynamic interactions among biodiversity and drivers such as climate and land-use change (Clark et al. 2001, Gilman et al. 2010, Hoffmann and Sgro 2011, Urban et al. 2016).

Most models predict only a subset of the EBVs expected to encapsulate the major dimensions of biodiversity change. For instance, genetic models project future genetic variation and adaptations, demographic models project population abundances of single species, and community models project community richness and composition. A few more sophisticated models predict a greater range of dimensions but are often restricted to particular taxa (e.g., trees in TreeMig, Lischke et al. 2006). Understanding how humans shape the many layers of biodiversity currently requires multiple models, each with different data needs, modeling languages, and configurations, and substantial postprocessing of outputs.

If diverse users cannot access, adapt, integrate, and apply models to new problems, then even the best models are unlikely to be adopted widely to promote the best conservation and management solutions. We found that model platforms vary in their accessibility to nonexperts and adaptability for alternative species, ecosystems, and questions.
Figure 1. Current biodiversity projection models and their characteristics. We assessed from left to right how models incorporated six important biological processes, the levels of biodiversity modeled, incorporation of spatial or and temporal components, essential biodiversity indicators returned as outputs, and model generality, modifiability, and open access. More sophisticated incorporation of mechanistic components and greater accessibility indicated by darker shading. Models are ordered from bottom to top on the basis of the number of components incorporated and their sophistication.
example, some species distribution and genetic models can be modified for any species or system and are applied widely. However, potentially more accurate mechanistic models are often specific to particular species or taxonomic groups, and modifying them to apply to new systems or circumstances is usually difficult. Therefore, the current penchant for phenomenological, correlative models likely reflects not just missing biological data for parameterization (Urban et al. 2016) but also the limited availability of flexible and easily modified mechanistic models.

Overall, we find that current biodiversity models generally lack the biological realism, adaptability, interoperability, and integration needed to address the complexities of the biodiversity crisis. We propose one universal modeling platform that would facilitate seamless integration and application to a multitude of systems, species, and uses.

**Toward a universal biodiversity projection platform**

A universal biodiversity projection platform is needed to advance biodiversity understanding, prediction, conservation, management, assessment, and policy solutions (figure 2). Like the trusty Swiss Army knife and its diverse tools, this platform would harmonize existing modeling frameworks and enable projections that are both sophisticated and adaptable to the full range of fundamental and applied biodiversity questions. We envision that such a platform would be a quantum leap forward compared with our current toolbox of individual models. First, one is more likely to use an existing, comprehensive Swiss Army knife rather than cobble together various independent tools. Second, users can select from the most relevant tools to meet individual needs rather than always needing to recreate existing tools. Third, integrating tools into one platform promotes their interactions and feedbacks with each other and with external drivers. Fourth, differences among different modeling types (e.g., correlative and mechanistic) can suggest information about underlying process and inspire more sophisticated approaches. Fifth, combining the available model types into ensembles often increases predictive accuracy. Sixth, by having an open-access platform, a diverse community of developers and users can efficiently contribute to building and integrating models and sharing data, parameterizations, and intellectual developments.

We next define a set of objectives for this platform. A unified biodiversity projection platform should improve projection accuracy and certainty relative to existing approaches; flexibly adapt to any species, system, scale, or region; facilitate model optimization and comparison; prioritize data needs; integrate model validation and monitoring; facilitate transparency and collaboration; and enable cost-effective design and evaluation of management solutions. To support these seven objectives, we delineate 16 design principles found in bold throughout the text and outlined in table 1.

**Improving accuracy and certainty.** A biodiversity platform that integrates diverse modeling types, including statistical and mechanistic biological models and Earth systems models, can improve both accuracy and certainty, by which we mean high precision and confidence in projections. Therefore, our first design principle (table 1) is that biodiversity models should be made realistic by including biological mechanisms and understanding. We define realism as incorporating biological processes into models as opposed to using correlations. Incorporating realism is challenging when faced with model structural uncertainty and when biological parameters are scarce or uncertain (Urban et al. 2016). We aim to advocate for mechanistic representations as much as possible but realize that, at times, correlative approaches will be useful and perhaps the only way forward when we do not yet know how to model key biological processes. The platform could address this issue, however, by combining insights from both statistical and mechanistic approaches (Buckley et al. 2010, Hartig et al. 2011).

A biodiversity projection platform should also enable seamless integration with Earth system drivers, such as climate, land-use change, and socioeconomic models (Clark et al. 2001, Rounsevell et al. 2014, Adam et al. 2015). Few biodiversity models currently account for such drivers even though these drivers interact strongly with biodiversity change (Lovejoy and Nobre 2018, Newbold 2018). Figure 3 illustrates one example for how to couple models of land-use drivers and biodiversity, and how feedbacks between the two can generate substantially different outcomes than when modeled individually. This integrated model of climate change, land use, and a climate-sensitive crop pollinator produced stronger impacts from climate change on biodiversity relative to projections from uncoupled models. As climate change reduced pollinator abundances, crop yields decreased. Lower crop yields increased demand for agricultural land, prompting subsequent conversion of natural lands into agriculture and reduced biodiversity in natural areas. The takeaway from this exercise is that interactions between biodiversity and land-use decisions generate different outcomes than when modeled separately, but these interactions are usually ignored (Albert et al. 2020). For example, although 35% of global food production depends on pollinators, most risk assessments neglect feedbacks between pollinator dynamics and agricultural land-use decisions (Prestele et al. 2021). Even if these specific results require validation with future data, integrating drivers and biodiversity models will be necessary to project their joint dynamics accurately.

An integrated platform also can account for error propagation across all steps of the predictive process so as to represent uncertainties more faithfully (Yates et al. 2018). Otherwise, errors at one stage do not affect or interact with errors at later stages, often providing an overly optimistic and unrealistic interpretation of model certainty (Nicol et al. 2019).

**Enhancing flexibility.** A universal biodiversity projection platform should be flexible enough to model all species, ecosystems, regions, and socioeconomic scenarios. Modularity
Figure 2. A universal biodiversity projection platform would project changes in species abundances, traits, genetics, and associated environmental impacts. The basic model begins with the environment, which varies in space (the bottom orange to yellow heat map) and can change through time naturally but also through human impacts and management actions with bottom layers indicating human impacts (e.g., urbanization) and management strategies (e.g., reserve design). The environmental layers interact with the genome (the blue funnel) to determine traits from the yellow ring (e.g., physiology and phenology) as shown by arrows. These traits in turn combine (the green funnel) to determine demographic inputs (births, immigration) and outputs (deaths, emigration). Each species is embedded in an interaction network (a). Arrows inside the funnel indicate how changes in species abundances feed back to alter genetics and the environment. Managers can design mitigation measures and test them with the modeling framework (e.g., corridors linking green habitats). Essential biodiversity variables are entered from monitoring and recorded in a data cube (light blue). The platform would follow a nested modular design (b), such that users can choose from multiple options that then reveal additional options and ultimately input parameters. An example nested set of options is presented here to inform the death rate of a focal species. In this case, a user selected the green-highlighted nodes to model a species with a death rate that depended on an enemy species. This interaction was also determined by trait variation in the prey species that was underlain by quantitative genetic variation, which has the potential to evolve through a nonzero heritability ($h^2$).
<table>
<thead>
<tr>
<th>Design principle</th>
<th>Design objectives</th>
<th>Description</th>
<th>Enables</th>
<th>Requires</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological realism</td>
<td>Accuracy and certainty, flexibility, optimization and comparison</td>
<td>Incorporates key biological processes that shape how biodiversity responds to environmental variation</td>
<td>Projections that include biological processes, and therefore capture causation, rather than relying on correlations which might be specific to current observations</td>
<td>Mechanistic sub-models that can be used when data is available; should interact with statistical models to provide enhanced flexibility</td>
</tr>
<tr>
<td>Error propagation</td>
<td>Accuracy and certainty, transparency and collaboration, solutions</td>
<td>Errors often interact with other errors during model runs, and therefore, models should propagate errors appropriately, which might not be possible when combining standalone model outputs</td>
<td>An accurate representation of uncertainty</td>
<td>Platform that propagates errors across submodules</td>
</tr>
<tr>
<td>Cointegration with Earth system drivers</td>
<td>Accuracy and certainty, flexibility, transparency and collaboration, solutions</td>
<td>Models feedbacks with major Earth-systems models, including climate, land use, and other ecosystem models</td>
<td>Modeling of feedbacks between biodiversity, Earth-system, and socio-economic models and broad-based optimization and feasibility assessments</td>
<td>Common input and output currencies and other coordinated features (e.g., spatial/temporal scales)</td>
</tr>
<tr>
<td>Nested modularity</td>
<td>Accuracy and certainty, flexibility, transparency and collaboration</td>
<td>Submodels can be easily added, exchanged, expanded, simplified, or removed</td>
<td>Model comparison and ensemble-forecasting</td>
<td>Hierarchical submodels that can be turned on or off according to user needs</td>
</tr>
<tr>
<td>Scalability</td>
<td>Accuracy and certainty, flexibility, validation</td>
<td>Data and processes that are available or operate at different scales can be scaled appropriately to operate within the modeling framework</td>
<td>Input of data and processes of varying resolutions</td>
<td>Procedures to upscale and downscale state variables</td>
</tr>
<tr>
<td>Optimization</td>
<td>Accuracy and certainty, optimization and comparison, validation, solutions</td>
<td>Optimal model structure and parameterization for maximal accuracy and minimal uncertainty based on validation with observed data; it might include procedures to optimize unknown parameters</td>
<td>Finding the model and parameters that produce the most accurate and certain projections</td>
<td>Platform that produces models of varying structure and complexity. Might include adaptive management and artificial intelligence</td>
</tr>
<tr>
<td>Ensemble projections</td>
<td>Accuracy and certainty, validation, transparency and collaboration</td>
<td>By enabling models of varying structure and complexity, a platform can produce multiple models that can be compared and combined to improve accuracy</td>
<td>Ensemble projections that often demonstrate enhanced performance over independent models</td>
<td>Platform that easily produces models of varying structure and complexity</td>
</tr>
<tr>
<td>Simplification</td>
<td>Optimization and comparison, solutions</td>
<td>Model structure is simplified on the basis of user-defined criteria, including performance, parameter or structural sensitivity or uncertainty, and costs of parameter estimation</td>
<td>Sensitivity, cost-benefit, and validation analyses</td>
<td>Multi-model system and techniques to assign costs to additional complexity</td>
</tr>
<tr>
<td>Prioritizing data needs</td>
<td>Supporting and informing data collection</td>
<td>Model sensitivity and uncertainty is used to prioritize which parameters should be collected or improved</td>
<td>Cost-effective data collection</td>
<td>Multiple parameterizations to assess model sensitivity</td>
</tr>
<tr>
<td>Essential biodiversity variables</td>
<td>Optimization and comparison, validation, transparency and collaboration, solutions</td>
<td>A standardized data hypercube of biodiversity state variables developed in tandem with monitoring outputs</td>
<td>Standardized model inputs and outputs and adaptive feedbacks with validation from global monitoring networks</td>
<td>Standardized “data hypercube” of predictions that are designed for inter-model interoperability</td>
</tr>
<tr>
<td>Open access</td>
<td>Transparency and collaboration</td>
<td>Users can run, share, modify, and contribute their own subroutines</td>
<td>Widespread use, coordinated enhancement of modeling effort, efficient development by users, and transparent understanding of model outcomes</td>
<td>Open use standards</td>
</tr>
<tr>
<td>Reproducibility</td>
<td>Accuracy and certainty, optimization and comparison, transparency and collaboration</td>
<td>Version control practices implemented such that the same code can be run and re-run and obtain the same outcomes</td>
<td>Outcomes can be repeated and traced back to model structure</td>
<td>Version control</td>
</tr>
<tr>
<td>Community</td>
<td>Optimization and comparison, transparency and collaboration, solutions</td>
<td>An organizing structure for biodiversity projections and the scientists that contribute them</td>
<td>Collaboration and synthetic understanding of global impacts and intersectoral impacts</td>
<td>Platform and organization that unites research and researchers</td>
</tr>
</tbody>
</table>
addresses this objective by providing users with a toolbox of modeling options to adapt to their individual needs and to facilitate intermodel comparisons (figure 2b; Golding et al. 2018). For instance, one application might combine modules on land use, species interactions, and demography, whereas another application links statistical species distributions to a mechanistic dispersal module to project range dynamics during climate change (Engler et al. 2009). A nested design further enhances flexibility by offering a hierarchy of modular choices (figure 2b). For example, within a biotic interaction module, users might also choose competition, predation, or mutualism. Within each interaction type, users might choose among different ways to model that interaction and whether the environment or genetics affects the interaction. Therefore, a biodiversity platform with nested modularity allows users to combine, exchange, expand, simplify, and exclude available modules and submodules to enable projections finely tuned to particular species, ecosystems, regions, and scenarios.

Biological processes operate at divergent spatial and temporal scales, and data are often available at different temporal and spatial resolutions. A universal biodiversity platform should feature scalability so that it can accommodate these scale mismatches via downscaling or upscaling of data layers.

**Facilitating optimization and comparison.** By providing diverse models, a universal platform facilitates model optimization. An optimal model depends on the question of interest but is generally the one that best predicts out-of-sample observations from different regions or time periods and therefore relies on causal mechanisms rather than correlations (Dietze et al. 2018, Urban 2019). Model projections become more accurate and precise when the modeling process includes dynamic feedbacks among the processes of model development, validation, and revision (Urban et al. 2016, Dietze et al. 2018). Therefore, the initial model reveals data needs, scientists improve estimates of sensitive or poorly defined parameters, models are reparameterized or revised and rerun, and the cycle continues as new observations challenge model outcomes, much the same way as weather forecasting proceeds every day. Although forecasts might initially be highly uncertain, this dynamic modeling feedback can rapidly improve projections.

Ready access to diverse models also can promote a more accurate solution to predicting biodiversity change than relying on a single model alone. By combining projections from multiple models, so-called ensemble model projections have become standard in weather and climate forecasting given their forecasting advantages (Murphy et al. 2004). For example, the United States predicts the track and intensity of hurricanes using an ensemble of 20 model outcomes (Hamill et al. 2012). However, biologists lag behind in adopting ensemble modeling, particularly for process-based models, largely owing to the difficulties in developing multiple models simultaneously (Araújo and New 2007). By allowing many models to be developed at once, a universal platform would facilitate multimodel development and potentially more accurate ensemble projections from divergent model types (Leroux et al. 2017).

Given the high price of collecting or refining model parameters (e.g., through costly experiments and measurements), users often want simpler models that can still generate accurate and certain outcomes. One way to reduce model complexity while retaining predictive capacity is to assess the sensitivity of model outcomes to parameters and remove those that do not enhance accuracy or precision during validation and prioritize those that do (Canessa et al. 2015). Forecasters can then simplify models to facilitate cost-effective projections that provide similarly reliable projections while also gaining insights about the complexity needed to model biodiversity dynamics.

A comprehensive biodiversity projection platform is needed for all these objectives because we cannot optimize, combine, or simplify models to their essential ingredients without first beginning with all the potentially important ingredients.

**Prioritizing data needs.** Despite increasing efforts to collect biodiversity data and make them accessible through synthetic databases (Meyer et al. 2015, Kattge et al. 2020), we still lack...
Figure 3. Uniting biodiversity and Earth system models. We coupled the RangeShifter (Bocedi et al. 2014) and CRAFTY (Murray-Rust et al. 2014) biodiversity and land-use models (a) to represent feedbacks between climate-induced changes in habitat quality, land use, and a simulated pollinator species in the French countryside surrounding Clermont-Ferand, classified by habitat type, including crops that do or do not support the pollinator (b). In panel (c), we demonstrate changes in habitat types from the uncoupled to the coupled model with arrows and icon size proportional to habitat area. The coupled model predicts higher conversion rates of pasture and natural areas to cropland (c) than uncoupled models because fewer pollinators (d) reduce crop yields, increasing demand for agricultural land and decreasing crop supply (e). For details, see the supplemental material.
critical biological data for most species (Urban et al. 2016). A unified biodiversity projection platform can play an important role in prioritizing data collection by providing a means to assess the sensitivity of outcomes to various parameters and structural elements and indicate which information is most needed to improve predictions rapidly. For instance, let’s assume we want to model a species for which we only have good physiological and demographic data but not a good understanding of dispersal distances. We could build a model with a range of dispersal kernels and evaluate how sensitive responses are to this uncertain parameter relative to other unknown parameters. If the responses are highly sensitive to differences in dispersal, then we could advocate for collecting dispersal data. Alternatively, if that is not possible, then model outputs could span the range of possible dispersal kernels. Focusing on collecting information on these sensitive parameters can produce the largest gains in model accuracy and precision while doing so in the most cost-effective and efficient manner. Therefore, model development and biodiversity data collection are best done concurrently whereby models inform what data are most needed, and new data inform model design and implementation.

**Simultaneous validation and monitoring.** EBVs support the development of standardized indicators of biodiversity trends that inform policy objectives such as the Aichi targets (Pereira et al. 2013). EBVs capture the major dimensions of biodiversity change, ranging from genetics to ecosystem properties. Until now, EBVs have been poorly connected to modeling efforts (figure 1), but forecasters need them to validate, refine, and update model structure and parameters, and monitoring networks need to know what variables are most useful in model projections.

We recommend explicitly incorporating the EBV framework within the biodiversity modeling platform to harmonize observations and predictions and form a more coherent system of, and adaptive feedbacks among, biodiversity projection, validation, monitoring, and assessment. A standardized data hypercube of EBVs would form the core outputs of a universal platform (figure 2). This standardized and consistent output matrix with agreed-upon naming conventions would provide ready-made interconnections with external models and promote validation with data streaming in from global monitoring networks (Fer et al. 2021). As monitored input variables change, model predictions would change, enabling real-time assessments of biodiversity change. Coordinating the joint development of measurable biodiversity indicators between projection and monitoring efforts to assess progress toward biodiversity and sustainable development targets would thereby provide early warnings of impending catastrophic changes (Mace et al. 2018).

**Open forecasting.** A biodiversity projection platform should support the creation of open-access, reproducible, and traceable code; promote user contributions; and facilitate an interconnected and diverse community of modelers. In short, we support the call for building the community-based cyberinfrastructure needed for biodiversity science (Fer et al. 2021). Open access ensures that anyone can acquire model code without paying fees or awaiting author permission. Specifically, the platform should adhere to the copyleft license standards, which guarantees that users can run, share, modify, and contribute their code to software. Moreover, these standards require that code cannot be used in proprietary software and must stay under the same license to promote collaboration. By facilitating user-contributed code, the platform could tap into the global expertise, knowledge, and innovation needed to expand the platform; keep it relevant in the face of changing knowledge; and build an interactive community of biodiversity forecasters. This active community of software developers would efficiently distribute the massive workload of this complex global project. For example, the open-access R software environment has become a universal platform for statistical modeling, which has been expanded and kept relevant by a large user community. The LANDIS-II forest landscape model also exemplifies these open-source principles for biology. LANDIS-II comprises a large community of users and developers and includes customizable libraries that allow exploration of climate, land use, and forestry changes (www.landis-ii.org).

The platform should adopt version control practices that require the inclusion of structural metadata and build a repository to maintain future access to the entire version history. This repository ensures reproducibility because analyses can be repeated from the original model version (Golding et al. 2018). Any altered code would receive a unique version number, and each model run would record version numbers for subsequent use.

A universal biodiversity platform should support a globally connected community of biodiversity and Earth system modelers, not unlike that facilitated by the climate and ecosystem modeling communities (Harfoot et al. 2014, Urban 2019). The platform could act as a community portal to capitalize on the extensive but all too often unconnected expertise required to create biodiversity forecasting models. This platform would not only integrate biologists, but also mathematicians, statisticians, computational scientists, software engineers, geographers, and atmospheric scientists, to name a few. The Inter-Sectoral Impact Model Intercomparison Project already brings together scientists to compare diverse impacts from the same climate change and socioeconomic scenarios (Warszawski et al. 2014) but often lacks contributions from biologists.

**Purpose built for solutions.** We suggest that any biodiversity projection platform should be designed with users, including land managers, policymakers, and stakeholders. Therefore, the platform should support an interactive process among stakeholders, modelers, and monitoring networks to codesign analyses that solve real-world problems (Clark et al. 2001, Land et al. 2017). This way the platform can be conceptualized as a modeling environment within a human decision-making process. An important part of this
process will be conveying both outcomes and uncertainties so that policymakers can make decisions that hedge against uncertain and undesirable outcomes.

A universal platform should facilitate global assessments of biodiversity impacts of shared socioeconomic scenarios (Rosa et al. 2017) to enable accurate, targeted, and agile assessments by international agencies (e.g., IPBES, the Intergovernmental Panel on Climate Change) tasked with recommending global political and economic strategies for mitigation of global changes in climate, biodiversity, and ecosystem services. A biodiversity platform should also contribute to cross-sectoral syntheses of global impacts for assessment purposes. Such a platform can also help define new scenarios focused on biodiversity that more strongly link to local social–ecological dynamics (Kok et al. 2017). For example, although replacing natural, nonforested ecosystems with tree monocultures might seem an efficient approach to climate mitigation, it can negatively affect biodiversity (Seddon et al. 2019). By jointly evaluating climate and biodiversity impacts, better nature-based solutions can be found that optimize both climate and biodiversity solutions.

Just as importantly, this platform should inform the design and testing of specific management strategies—for example, by using corridor and reserve design to promote climate change resilience (Albert et al. 2017). The platform should promote joint adaptive modeling and adaptive management (learn by doing while reducing uncertainty), by including an adaptive management cycle, whereby management interventions are designed to maximize model outcomes. These management actions can then be incorporated into model projections to reduce uncertainty and update observations, thus informing subsequent actions (Walters 1986). In addition to facilitating advanced simulations, an integrated biodiversity platform should allow for scenario testing and the identification of optimal management approaches. These optimization approaches should permit users to ask sophisticated questions and to identify solutions that concurrently maximize biodiversity, climate change mitigation, and
socioeconomic benefits (Alagador and Cerdeira 2020). For instance, models might use real estate values to guide reserve design during climate change, thus minimizing both financial and biodiversity losses.

Although we view biodiversity modeling as becoming more mechanistic, artificial intelligence technologies could help design mitigation strategies that optimize management criteria, including socioeconomic outcomes, on the basis of outputs from biodiversity models (figure 4). Artificial intelligence solves problems through adaptive algorithms that optimize target criteria and is increasingly applied to natural resource management and conservation decision-making (Pichoncourt et al. 2012, Chades et al. 2017). Recent advances such as deep reinforcement learning are enhancing its wider application (Silver et al. 2016). For example, artificial intelligence could optimize the spatial design of land-use patterns across complicated socioeconomic landscapes to maximize both benefits and practicality, where complex socioeconomic dynamics (e.g., figure 3) can generate millions of alternative management strategies that surpass human intuition. By incorporating artificial intelligence techniques, the platform can identify globally optimal and feasible management solutions more readily.

Implementing the vision
One of the greatest challenges of the twenty-first century is to make informed predictions that will enable us to design strategies to protect life on Earth, despite historic threats (Mace et al. 2018). Forty years ago, atmospheric scientists also faced a similar task of predicting climate and weather, but lacked a cohesive modeling platform (US National Academy of Sciences 1975). Rising to the challenge, scientists created multiple dynamic and mechanistic climate models, established shared socioeconomic scenarios, and developed a framework to integrate and compare model outcomes. These collaborative modeling platforms enabled more cohesive and evidence-based assessment for climate and enhanced confidence in projections of future climate change to support policy decisions (Edwards 2011). Biology needs such a tool if we hope to bend the curve of biodiversity loss upward in coming years.

Predicting biodiversity is not easy on the best of days, and even the most sophisticated model projections are likely to be frustrated by high uncertainty and ecological surprises (Doak et al. 2008, Berger and Smith 2019). Given the many complexities of biology, prediction might not ever reach the accuracy levels attained for weather or physical particles. However, we argue that substantial gains in predictive accuracy are possible even with modest gains in model development, given the current state of the field. The scientific community has not developed many mechanistic biodiversity predictions, and even fewer have been validated with monitoring data. But those that have been tested demonstrate considerable promise. For instance, mechanistic models have successfully been used to predict population declines, pest population dynamics, species distributions from phenological traits, forest carbon dioxide exchanges, and fire dynamics in nature (Wildl 1999, Brook et al. 2000, Amthor et al. 2001, Chuine and Beaubien 2001, Emmett et al. 2021). These examples suggest that accurate predictions for biodiversity and ecosystems are possible especially when more mechanistic models and data are available. Finally, even if only modest gains are possible, we will still have gained fundamental insights about the limits to predictability in biology.

Although we still too often lack the basic biological data needed to inform biodiversity models, new data efforts are rapidly filling these data gaps. Governments, organizations, and scientists are collecting and compiling these data at an accelerating rate and storing them in repositories, including species distributions (e.g., the Global Biodiversity Information Facility and Ocean Biodiversity Information System), historical abundances (e.g., the Global Population Dynamics Database, Projecting Responses of Ecological Diversity in Changing Terrestrial Systems [PREDICTS] project, and the BioTIME database), and traits (e.g., TraitBank, TRY database). A biodiversity projection platform could begin using these resources immediately and also facilitate the further sharing and integration of data. Even where data gaps continue to exist, models will be crucial in directing efforts toward more efficient data collection (Ficetola et al. 2018). Waiting until we collect all relevant data would prove too late for such models to be useful. Therefore, we need a comprehensive platform both to make use of the increasing big data of biodiversity (Wuest et al. 2020), but also to guide and streamline the monumental effort of collecting relevant data to support model development and parameterization.

A universal modeling platform could develop either by building from basic principles (figure 2) or by tethering together existing models (figure 3). Building a new platform would be desirable from the standpoint of consistency and ensuring rapid operation and integration. However, limited resources for biodiversity science might make this approach impractical. The alternative is to link existing programs, such as those listed in figure 1. This alternative approach would prove efficient from the standpoint of using existing, error-checked models, and we illustrated the feasibility and usefulness of this approach in figure 3. Moreover, multi-model integration could enable substantial gains with relatively little effort by normally ignored model interactions. However, substantial work is needed to ensure that coupled models correctly interpret inputs and outputs from one another and include appropriate linking functions across spatial and temporal scales. Moreover, these Frankenstein models usually run slowly given the computing overhead of cross-program communication and translation. Likely the best course of action is to link existing models now, while working toward recoding models in a common language and framework to speed up future analyses.

Several options exist for integrating models and building a modular system for coupling code into a biodiversity
projection platform. One option is to create a package that acts as a wrapper for submodules written within the popular R programming environment. For instance, the zoon R package allows users to choose species distribution modules from those contributed to an open, version-controlled online repository and then generate reproducible workflows that combine results from the chosen modules (Golding et al. 2018). A complementary approach is to take advantage of software containers, such as the Docker virtualization platform (www.docker.com), that create stand-alone packages that can integrate multiple applications that require different data and computational environments and encapsulate all software dependencies that might otherwise change through time (Huang et al. 2019, White et al. 2019). Already this system has been used to automate ecological forecasting, including processing new data, fitting, calibrating, and running multiple different process-based models, analyzing the outputs, and creating an ensemble forecast. For instance, Docker has been used to create periodically updated and interactive projection platforms for both rodent abundances and forest carbon sinks (Huang et al. 2019, White et al. 2019). Docker also can create interfaces between biodiversity and land use or Earth system models that often operate on different platforms (Robinson et al. 2018, Millington et al. 2021).

Building on the design principles outlined in table 1, the next step is to form a governing board of global scientists, modelers, and biodiversity professionals to coordinate platform development and explore financing options. Once a version is available, the next phase would be to demonstrate its abilities on simulated and real data sets. Simulated data sets with known drivers and computational environments and encapsulate all software dependencies that might otherwise change through time (Huang et al. 2019, White et al. 2019). Docker also can create interfaces between biodiversity and land use or Earth system models that often operate on different platforms (Robinson et al. 2018, Millington et al. 2021).

Conclusions
Most biodiversity forecasters either rely on more generalizable, but less accurate, models or undergo the time-consuming and costly process of developing process-based models specific to particular questions. Consequently, biodiversity science is less efficient, accurate, integrated, and equitable than it could be with a universal platform. We now find ourselves in the middle of the Anthropocene and ill equipped to predict and prevent biodiversity and ecosystem change. However, advances in biology, computer science, artificial intelligence, and computing power now exist to address this challenge quickly. Scientists now have the capacity to recreate the complexities of diverse interacting species within the silicon brain of the computer, replicate it in servers throughout the world, and implement artificial intelligence to find optimal management schemes. Such tools will allow us to decrease uncertainties and develop better evidence-based mitigation and adaptation strategies. With bold innovation coupled with appropriate coordination and support, this grand deficiency in global science can and should be solved this decade.

Acknowledgments
This article originates from a working group supported by the Quebec Centre for Biodiversity Science and the University of Connecticut Center of Biological Risk. National Science Foundation grants no. DEB-1555876 and no. PLR-1417754 and the Arden Chair in Ecology and Evolutionary Biology supported MCU. The German Science Foundation DFG grant no. ZU-361/1–1 supported DZ and AKM.

Supplemental material
Supplemental data are available at BIOSCI online.

References cited


https://academic.oup.com/bioscience


_____. 2019. Projecting biological impacts from climate change like a climate scientist. WIREs Climate Change 10: e585.


