Received: 10 September 2024 | Accepted: 17 July 2025

DOI: 10.1111/2041-210X.70131

BRITISH ECOLOGICAL SOCIETY Methods in Ecology and Evolution

REVIEW

Advancing causal inference in ecology: Pathways for biodiversity change detection and attribution

Franziska Schrodt[1] | Miriam Beck[2] | Joaquim Estopinan[3] | Diana E. Bowler[4] | Colin Fontaine[5] | Pierre Gaüzère[3] | Romain Goury[3] | Matthias Grenié[3] | Inês S. Martins[6] | Naia Morueta-Holme[7] | Luca Santini[8] | Mickael Hedde[9] | Gabrielle Martin[10] | Emmanuelle Porcher[11] | Sara Si-Moussi[3] | Marianne Tzivanopoulos[3] | Grant Vernham[1] | Cyrille Violle[12] | Wilfried Thuiller[3]

[1]School of Geography, University of Nottingham, Nottingham, UK; [2]CESAB, Montpellier, France; [3]University of Grenoble, University of Savoie Mont Blanc, CNRS, LECA, Grenoble, France; [4]UK Centre for Ecology & Hydrology, Wallingford, Oxfordshire, UK; [5]CESCO, UMR 7204, MNHN-CNRS-SU, Paris, France; [6]Department of Biology, Leverhulme Centre for Anthropocene Biodiversity, University of York, York, UK; [7]Section for Biodiversity, Globe Institute, University of Copenhagen, Copenhagen, Denmark; [8]Department of Biology and Biotechnologies 'Charles Darwin', Sapienza University, Rome, Italy; [9]INRAE Occitanie-Montpellier, Montpellier, France; [10]Centre de Recherche sur la Biodiversité et l'Environnement (CRBE) UMR 5300, Université de Toulouse, Toulouse INP, CNRS, IRD, CRBE, Toulouse, France; [11]Muséum National d'Histoire Naturelle, Île-de-France, Paris, France and [12]CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

Correspondence
Franziska Schrodt
Email: f.i.schrodt@gmail.com

Funding information
Danmarks Frie Forskningsfond, Grant/Award Number: 2064-00091B; French Foundation for Research on Biodiversity; HORIZON EUROPE OBSGESSION, Grant/Award Number: 101134954; Royal Society, Grant/Award Number: URF\R1\241013; Villum Fonden, Grant/Award Number: VIL53086; Natural Environment Research Council, Grant/Award Number: 09045

Handling Editor: Phil Bouchet

Abstract

1. Understanding the causes of biodiversity change is essential for addressing environmental challenges. While causal attribution has advanced in other fields, ecologists remain cautious about causal claims or misinterpret predictive models as causal. With growing spatio-temporal data, computational power and cross-disciplinary collaboration, discussions on improving attribution methods in ecology are gaining momentum. However, practical guidance remains limited for non-experts. Here, we identify the challenges and decisions involved in detecting and attributing biodiversity change and provide an overview of suitable methods based on available data and specific research questions.
2. The first challenge we address pertains to biodiversity and driver data. Unlike controlled experimental data in other disciplines, ecological data often stem from monitoring programs or field samplings with varying degrees of rigour, which complicates the analysis due to sampling biases, interacting drivers, measurement error or spatio-temporal variations. We specifically outline how data structure (e.g. structured vs. opportunistic data) and data coverage along the spatial and temporal scale impact detection and attribution.
3. The second challenge involves the ability to detect directional change in the system of interest, which is associated with numerous hurdles. We provide an

overview of the most relevant approaches to deal with sampling variability, gaps and biases in the data, non-linearity in the temporal trends and to identify the most appropriate spatio-temporal resolution.

4. For the third challenge, causal attribution, we focus on data-driven approaches. We review recent frameworks that draw on methodologies from other disciplines, offering analytical roadmaps and step-by-step guidance for causal inference. These include constructing theoretical causal models a priori, full causal models based on data and theory and *posterior* causal interpretation tailored to specific data and research questions.
5. Moving forward, it is essential to foster interdisciplinary collaboration to adapt and refine methodologies from other fields, ensure robust data collection and sharing practices, promote the integration of advanced computational tools and improve the link between data-driven and theory-driven approaches. This approach will enhance our ability to make robust causal inferences; thereby improving our understanding of biodiversity changes and informing effective conservation strategies.

KEYWORDS
attribution, biodiversity change, causal inference, detection, environmental change, global impactsanthropogenic drivers

1 | INTRODUCTION

Better *attribution* and understanding of *causal relationships* (Table 1) in biodiversity change are needed to create realistic and fit-for-purpose *projections* (Table 1) and scenarios. Both are required for various applications, including informing policy (e.g. IPBES, 2016; Nicholson et al., 2019) and increasingly nature markets and economic models (Ferraro et al., 2019; Takahata et al., 2024; West et al., 2023), effectively allocating resources and conservation priorities, promoting accountability and ultimately, improving our understanding of the consequences of environmental change and human actions on ecosystem functioning and human well-being (MacDonald et al., 2019). Methods for causal *driver attribution* (Table 1)—the process of evaluating the relative contributions of multiple potential *causal factors* (drivers) to a detected biodiversity change, with an assignment of statistical confidence to the causal models used to estimate these effects (Gonzalez et al., 2023)—have been proposed in several fields (e.g. physics, climate sciences, economics and epidemiology) (Hegerl & Zwiers, 2011; Pearl, 2009; Reich et al., 2021; Runge, Bathiany, et al., 2019). However, ecologists often remain conservative in making explicit causal statements and tend to rely on a limited set of causal *attribution* methods, such as structural equation models.

Some forms of predictive modelling are also often misinterpreted as identifying causal *drivers*, even though their primary objective is not to establish causality. Predictive models aim to forecast outcomes based on observed patterns, often using machine learning or regression approaches. However, they do not necessarily identify the mechanisms behind observed outcomes. For example, in their study, Arif and MacNeil (2022a) used random forests to predict reef degradation outcomes across Pacific islands based on multiple covariates like sea surface temperature, fishing intensity and proximity to human settlements. While this approach identified areas of high risk and improved the prediction accuracy of future reef decline, it could not determine which stressor caused the decline. For instance, fishing and temperature were both strong predictors, but their relative causal contributions remained unclear. This distinction is crucial as some forms of predictive modelling have been misinterpreted as identifying causal drivers, even though their primary objective is not to establish causality (Arif & MacNeil, 2022a; Stewart et al., 2023). Such misinterpretations can lead to erroneous policy or management decisions. Conversely, causal models are explicitly designed to test or infer the direction and magnitude of causal relationships. These models help disentangle confounding variables and support stronger inference about what factors truly drive change in ecological systems. For example, Stewart et al. (2023) used Structural Equation Models to investigate causal relationships among agricultural expansion, forest fragmentation and mammal diversity decline in sub-Saharan Africa with the goal to quantify direct vs. indirect effects (e.g. does agriculture directly reduce biodiversity, or does it act through forest loss?). They found that agricultural expansion had a strong indirect effect via forest fragmentation, but a weak direct effect. In contrast to the predictive model, this model was explicitly built around causal hypotheses and used latent variables and path coefficients to infer directionality. Nichols and Cooch (2025),

TABLE 1 Key definitions used in this manuscript.

Term	Definition	Category
Detection	The process of demonstrating that a measure of biodiversity has changed relative to a baseline or reference distribution characterizing undisturbed variation (counterfactual state), or an appropriate model-derived null expectation of biodiversity change in the absence of a human driver(s) (from Gonzalez et al., 2023)	General
Attribution	The process of evaluating the relative contributions of multiple potential causal factors (drivers) to a detected biodiversity change, with an assignment of statistical confidence to the causal models used to estimate these effects (from Gonzalez et al., 2023). Could also be defined as/referred to as causal inference	
Causal relationship	X → Y can be considered a causal relationship if there is reason to think that variations induced in X can propagate to subsequent variations in Y (from Grace, 2024)	
Causal effect	For a given causal relationship, a magnitude measure quantifying how variation in the treatment impacts change in the outcome. Under key assumptions in causal frameworks (SCM, potential outcomes), this measure can be considered unbiased. Causal effects can be quantified with different metrics, depending on the study scale, objective and method applied. The most common are the overall average treatment effect (ATE), its restriction to treated units, the average treatment effect on the treated (ATT) and the conditional ATE (CATE), which restricts the ATE to a subset of units that satisfy a condition on their covariates. The CATE refines the causal effect estimation when there is heterogeneity between units that blur the ATE. See Nogueira et al. (2022) Section 5 *Causal effects* for mathematical definitions	
Useful approximation	A simplified method or assumption that allows researchers to estimate causal effects when exact causal relationships are difficult or impossible to determine due to data limitations, unmeasured confounding or complex causal structures. These approximations aim to provide reliable insights while balancing feasibility and accuracy (see Grace, 2024)	
Projection	An estimate or forecast of future ecological trends or status based on analysis of current and past data on either the ecological response itself (e.g. forest cover) or associated drivers acting as proxies, and on boundary conditions that characterize future conditions (e.g. temperature change)	
Predictive modelling	The use of statistical methods and algorithms to forecast future outcomes by identifying patterns in historical data. Importantly, these models do not generally reveal or confirm causal relationships	
Prediction	A probabilistic assessment of future trends or status based on current knowledge. Unlike projection, prediction is mainly influenced by our current knowledge, that is, initial conditions, rather than future boundary conditions	
Causal graph	Also *Directional Acyclic Graph* (DAG) or *path diagram*. A conceptual representation of the causal assumptions of a system in which all relevant variables are included (both observed and unobserved) and arrows indicate the direction of causality (from cause to effect). This allows variables to be identified as confounders, colliders or mediators and helps to select the optimal methods to be used	Causal model
Confounder	A third variable Z that influenced both the response and the driver variable. Its effect may be direct or indirect (X ← Z → Y). Methods should always control for confounding variables to avoid attributing the wrong drivers	
Mediator	A third variable Z, through which a driver variable affects the response and thus determines indirect effects (X → Z → Y). When estimating total effects, methods should not condition on mediator variables Synonymous terms used in other disciplines are 'modifier', 'moderator' or 'modulator'	
Collider	A third variable Z that is affected by both the driver and response variables (X → Z ← Y). They should not be conditioned on when the total effects are estimated	
Driver	Also *exposure variable*, *treatment variable* or *causal factor* (see def. attribution). Process or feature that leads to a change in a response process or feature	
Response	Also *outcome variable*. Variable upon which the *driver* has a causal effect (e.g. a biodiversity metric)	
Adjustment set	Collection of variables that need to be conditioned on (controlled for) to estimate the causal effect of a driver (X) on a response (Y) accurately. These variables help block confounding pathways that could distort the relationship between the driver and the response, ensuring a valid causal estimate. Adjustment sets are not necessarily unique, and multiple sets may be identified depending on the context and the causal structure of the problem	
Estimand	A precise description of the treatment effect that a study aims to quantify, reflecting the specific scientific question of interest	

TABLE 1 (Continued)

Term	Definition	Category
Treatment	Also intervention. A specific condition or intervention applied to a subset of the study subjects or experimental units, such as plots of land, populations or ecosystems, to observe and measure its effects on ecological variables. A treatment might be active (e.g. application of nutrients, change of soil temperature in an experimental plot) or passive (e.g. decadal land use change)	
Unobserved variable	Also *latent variable*, variable that plays a role in the data-generating process but is not represented in the data distribution for various reasons: overlooked, technically difficult to measure, unethical, etc	
Overlap assumption	A requirement of many sensitivity analyses that, for each value of the covariates (i.e. characteristics or predictors), there must be a positive probability of receiving each treatment or being in each comparison group. In other words, every individual in the study population has a non-zero chance of being assigned to any of the treatment groups under consideration	
Sensitivity analysis	Tests how strong the unmeasured confounding would have to be to explain away the association, that is, how strong the unmeasured confounder would have to be associated with the treatment and outcome for the association between treatment and outcome not to be causal (VanderWeele & Ding, 2017)	
Stationarity	A time series is considered stationary if its mean, variance and autocorrelation structure remain constant over time. This implies the absence of a trend and periodic fluctuations	Time series
Contemporaneous effects	The effect(s) in question occur immediately, without any time lag. The causal relationship between two variables is thus 'contemporaneous'	
Cross-sectional data	Observation of subjects at one point or period of time, or for which the analysis has no regard to differences in time among the observations (Nogueira et al., 2022). In ecology, these data are typically spatial data In causal analysis, cross-sectional data are used to examine how differences in one variable (e.g. an exposure or treatment) are associated with differences in another variable (e.g. an outcome) across subjects at the same time. However, since the data are collected at one point in time, establishing a clear cause-and-effect relationship can be challenging	Data type
Structured data	Data collected using a standardized and systematic sampling design to ensure consistency, comparability and statistical rigour. These datasets are designed to minimize bias and maximize reproducibility, allowing for more robust ecological inference but are highly resource-intensive and thus not available in as large quantities as opportunistic data	
Opportunistic data	Data collected without a standardized or systematic sampling design, often as a byproduct of other activities. Unlike structured monitoring programmes, opportunistic data typically lack predefined spatial, temporal or methodological consistency. Opportunistic data are extremely valuable and often available at larger quantities than structured data, but require careful statistical treatment to account for biases and limitations	
Longitudinal/panel data	Observations about several subjects at multiple points or periods of time, indexed in time order and subject. Time series are a particular case considering only one subject (Nogueira et al., 2022). Most typically called temporal or spatio-temporal data in ecology	
Time series data	Observations about a single subject at multiple points or periods of time, indexed in time order (Nogueira et al., 2022)	

however, argue that the divide between prediction and causality may be overstated, suggesting that under certain assumptions and with appropriate tools (e.g. causal forests, do-calculus, instrumental variables), predictive models can indeed uncover quasi-causal insights, especially when longitudinal data or natural experiments are available. However, this requires careful design and cannot be assumed in general predictive workflows.

Implementation of a conservative causal *attribution* approach in ecology presents four significant challenges due to the inherent complexity of ecological systems, data limitations and the interplay of multiple interacting, often unmeasured factors. (1) Detecting directional change: change *detection* in either a *response* variable or its putative *driver* (see definitions in Table 1) is non-trivial, particularly given issues of data scarcity and non-random sampling. Both observational and experimental ecological data frequently suffer from incomplete spatio-temporal coverage, bias in data collection and methodological inconsistencies (Dee et al., 2023). These challenges are further compounded by the scale dependence of ecological processes, where changes observed at one spatial or temporal scale may not generalize to others (Estes et al., 2018). (2) Complexity of mechanistic understanding: the lack of direct measurement of many key *drivers* results in often limited mechanistic understanding and reliance on inferred or proxy variables. The high-dimensional nature of ecological systems means that *drivers* interact in non-linear and often unexpected ways, with biological *responses* exhibiting lags, feedback loops and context-dependent effects (Essl et al., 2015; Myers-Smith et al., 2020). The spatio-temporal variability of both biotic and abiotic *drivers* complicates efforts to isolate the contribution of specific factors, and even controlled experimental setups cannot fully account for all or potentially even the most important sources of bias and error, limiting their ability to establish mechanistic understanding (but see Kimmel et al. (2021) for a framework to identify and address them). (3) Methodological and computational barriers: the mathematical and coding expertise required for formal causal *attribution* approaches is often perceived as a barrier to widespread implementation and can risk increased use of black box approaches. On the contrary, this complexity and messiness of ecological data also results in data often failing strict statistical requirements of causal statistics (see Grace, 2024 for a thorough discussion) (4) Defining baselines in a changing world: determining appropriate baselines for detecting change is a critical but unresolved issue. Many ecological monitoring programmes were initiated only after environmental degradation became apparent, meaning that true pre-impact conditions are rarely available for comparison (Soga & Gaston, 2018). Furthermore, most long-term ecological monitoring efforts commenced after anthropogenic pressures had already altered ecosystems, making it difficult to disentangle human impacts from natural variability (Mihoub et al., 2017).

In contrast, given the increasing availability of 'big data' in ecology, including biodiversity and *driver* data (Bush et al., 2017; Farley et al., 2018; Schrodt et al., 2024; Wüest et al., 2020), enhanced computational power and open sharing of code among disciplines, there has been a resurgence of discussions on improving *attribution* methods in ecology. Several frameworks have been proposed to address some of the above-mentioned issues, also encouraging application of more comprehensive and sophisticated causal *attribution* methods in ecology (Dee et al., 2023; Ferraro et al., 2019; Gonzalez et al., 2023; Grace, 2024; Laubach et al., 2021). Laubach et al. (2021) and Ferraro et al. (2019) draw on methods from other disciplines, especially epidemiology and economics, providing an analytical roadmap for causal inference based on causal diagrams and structural knowledge of the system. Gonzalez et al. (2023), in contrast, propose a framework based on the '*detection* and *attribution*' approach developed in climate sciences (Bindoff et al., 2014). This five-step approach starts with a theoretical causal model guiding data collation, followed by estimating and detecting change in the target biotic variable, and, finally, attributing this change to a *driver* (Gonzalez et al., 2023). Grace (Grace, 2024) on the contrary, proposes an integrative paradigm harnessing mechanistic understanding and fully utilizing evidence across different research studies. Yet, despite the developments of conceptual frameworks for causal inference, there is a lack of practical guidance to overcome the range of problems that ecologists face.

Here, we address key challenges of biodiversity change *detection* and conservative causal *attribution* and propose solutions to overcome barriers in (1) *biodiversity and* driver *data characteristics*, (2) *detection of change within both data types* and (3) *linking* driver *and biodiversity data for causal inference*. Specifically, we provide guidance on the choices to be made at each step of biodiversity change *detection* and *attribution*, offering a guide for selecting suitable methods based on the available data and the research question. To ease cross-referencing and understanding for non-experts, important technical terms are bolded throughout the text and defined in Table 1. For more detailed discussions on terminology, see: on discipline-specific jargon (Siegel & Dee, 2025), on definitions and interpretations of the term 'causal' (Grace, 2024), on use of causal terminology in Remote Sensing (Van Cleemput et al., 2025).

2 | CHALLENGE 1: BIODIVERSITY AND DRIVER DATA

Several properties of data critically affect the ability to detect and attribute changes in biodiversity (Figure 1).

The first general property is the level of structure in data collection and reporting of observations. *Structured data* (Table 1) typically come from standardized sampling protocols (referred to as 'sampling structure' in Figure 1), often with repeat sampling across years at the same location ('temporal structure' in Figure 1, also referred to as '*longitudinal*' or '*panel*' data) and sometimes with a spatial sampling design ('spatial structure' in Figure 1, Table 1) (Kelling et al., 2019). In contrast, unstructured data consist of *opportunistic* observations (Table 1) compiled from uncoordinated data collection of independent surveyors.

The second property is data coverage: biodiversity and *driver* data can range from small to large spatial extents, short to long timeframes and can target anything from a single species or environmental *driver* to comprehensive documentation of all species or multiple socio-economic, environmental and biological *drivers* in a given location (Figure 1). Further challenges emerge since multiple dimensions of biodiversity may change at different scales. For instance, a single metric of biodiversity can increase when considering certain spatial scales, times and taxa, but decline in others (Dornelas et al., 2023).

Here, we highlight three main structure axes and three main coverage axes (Figure 1); though the importance of each will depend on the type of change being detected and the *driver* and *response* in question. Below, we discuss the main challenges associated with data streams that vary with respect to these axes and highlight the key trade-offs that need to be considered for *detection* and *attribution*.

2.1 | Challenge 1.1—Biodiversity data

2.1.1 | Data structure

While *structured data*, arguably the most desirable type of data for *detection* and *attribution*, remain rare, unstructured data are increasingly abundant due to online platforms like iNaturalist and data aggregators such as GBIF (Global Biodiversity information Facility), which involve reporting and aggregation of observations by experts and casual observers. Semi-structured data, such as eBird (Kelling et al., 2019), combine elements of *structured* and unstructured data, allowing variable data collection protocols but with improved documentation and interoperability. Finally, synthesis databases like BioTime (Dornelas et al., 2018), Living Planet Database (Loh et al., 2005) or PREDICTS (Hudson et al., 2014), aggregate numerous

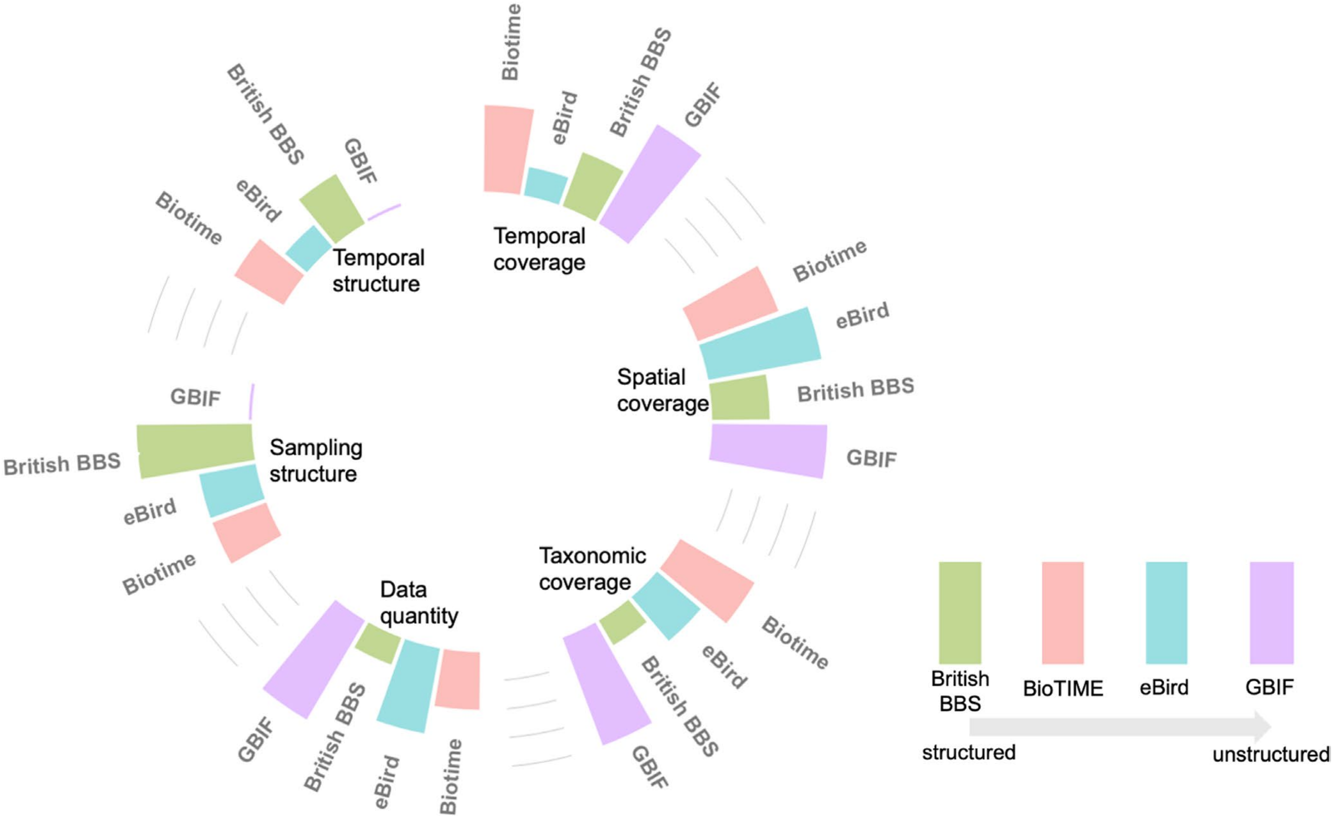

FIGURE 1 Existing biodiversity databases vary along seven major axes of structure and coverage. Biodiversity change detection and attribution data should ideally score highly on all these axes. However, all databases (and sources of data, including from remote sensing) score low to medium on at least one of the axes. We illustrate such trade-offs on the example of four databases that were selected to follow a gradient in data structure (i.e. from a typical highly *structured* Breeding Bird Survey that is found in many European countries and the US (e.g. the British BBS, green) to the unstructured GBIF (purple) database). The height of the bars indicates the qualitative score along each axis for each database (indicated by colours) relative to the other databases based on published data and relative estimates. For example, the *time series* (Table 1) in BioTIME (pink) may span over 100 years and be structured in time (high time coverage and structure) but is less structured in space (low spatial structure) and can be spatially limited to a single lake or forest patch (medium spatial coverage). In contrast, national Breeding Bird Survey (green) gather bird monitoring data with a high spatial structure over multiple years, but are limited in terms of spatial coverage, and citizen science platforms like eBird (blue) provide vast amounts of data (high data quantity) but for one taxon only (low taxonomic coverage) and with little temporal structure. Note that additional aspects that are of relevance, such as the data collection starting year and prior definition of study sites, are not shown in the figure. Semi-quantitative data informing this figure is presented in Table S2.

independent studies, resembling semi-structured data: each study is internally standardized, but observations across studies vary in methodology and spatial–temporal resolution, with some minimum inclusion requirements and no a priori overarching protocol controlling which sites, species and times are sampled.

Structured or semi-structured data have many attractive features for monitoring and attributing trends since sampling variability is minimized or can be modelled with the available metadata (Bayraktarov et al., 2019). Repeated sampling at the same location (e.g. permanent or quasi-permanent plots), common in *structured data*, is ideal for detecting changes over time (Kapfer et al., 2017). However, *structured data* face trade-offs in spatial, temporal or taxonomic coverage (see next section) (Isaac & Pocock, 2015), which means that they are usually restricted in space and time and to select taxa (Estes et al., 2018), limiting their ability to capture ecological processes acting at larger scales (McGill, 2019). Unstructured data, while abundant and broad in coverage, are prone to bias and require complex analytical approaches to isolate true biological signals from sampling variation and bias (Dobson et al., 2020; Isaac et al., 2014; Rapacciuolo et al., 2021).

2.1.2 | Data coverage

Coverage is critical for both *detection* and *attribution* but for different reasons. A broad coverage across a wide range of relevant species, ecosystems and regions or time, enhances the ability to detect change across broad systems and to characterize its mean and variance. For *attribution*, wide spatial and temporal coverage is essential to describe the *response* curve of a species to a *driver*, requiring sampling across a range of *driver* values (Thuiller et al., 2004). Moreover, *responses* to ecological thresholds can only be detected with sufficient sampling at extreme values that surpass the threshold (Spake et al., 2022). Typically, there are trade-offs between the three coverage axes (Figure 1a), such as sampling many species in a limited area over a short period of time versus sampling a single species repeatedly over a large area over a long period of time. The coverage axes (Figure 1a) also trade off with the level of structure. For instance, unstructured data from GBIF have greater spatio-temporal taxonomic coverage than *structured data*, which often cover narrower areas along the three axes. However, developments like the Biodiversity Information Standards (TDWG) ratifying the Humboldt Extension to the Darwin Core (metadata standards for monitoring and survey data, including sampling effort/metadata) and GBIF's adoption of these standards have the potential to reduce these trade-offs, shifting unstructured data towards being semi-structured where possible (Sica et al., 2022).

2.2 | Challenge 1.2—Driver data

Driver data on any natural or anthropogenic feature or process that could influence biodiversity present unique challenges. While we do not aim to cover all the challenges, we here highlight some of the key considerations when dealing with *driver* data for the *detection* and *attribution* of biodiversity change. Similar to biodiversity data, *driver* data are often limited in spatio-temporal structure, resolution and coverage (Joppa et al., 2016), likely leading to spatial and temporal mismatches between biodiversity and *driver* data or between research and policy on *drivers* (Mazor et al., 2018).

2.2.1 | Data structure

Driver data often suffer from a lack of systematic measurement, leading to biases and measurement errors similar to those in biodiversity data. For example, modelled climate data (e.g. CHELSA (Karger et al., 2020)) are based on in situ data from only a limited number of weather stations, often located in built-up areas, resulting in locally unreliable data in some parts of the world, particularly those also affected by a scarcity of biodiversity data (Daly, 2006; Dinku, 2019). Similarly, databases of in situ abiotic characteristics such as soil or geology are often *opportunistic*, resulting in strong spatial and temporal biases (Schrodt et al., 2024). Temporal biases can also arise from changes in protocols, instruments or observers over time (Borges et al., 2024; Somerton et al., 2002). Data FAIRness has not accelerated for *driver* data as fast as it has for biodiversity data, further limiting usability (Schrodt et al., 2024).

Additionally, we often lack variables capturing the actual *drivers* and use proxy variables that may be unsuitable for testing causality, or their value as a proxy may be scale-dependent (e.g. distance to major cities often used as a proxy of human disturbance and exploitation). Consequently, assessing the impacts of some *drivers* (e.g. climate change) is more straightforward than others (e.g. overexploitation). Remote sensing can provide opportunities to collect data on some *drivers* over continuous space and short time periods (Lausch et al., 2019; Pettorelli et al., 2014), but poses their own challenges, including confounding variables (e.g. different sensor types and atmospheric corrections, differing bioclimatic conditions across cloudless views), *structured* or systematic and thus potentially highly problematic measurement errors and sometimes high uncertainty. In addition, the number and range of confounding factors are multiplied at the large spatial scales reached by remote sensing products and random measurement error typically leads to regression dilution bias, whereby the effect of a *driver* on biodiversity would be underestimated. Therefore, their use should be accompanied by consideration of compensatory causal inference techniques to control for these biases and sources of error. For a thorough overview of the premise and challenges of using remotely sensed data for causal *attribution* in ecology see Van Cleemput et al. (2025).

Land cover layers are based on machine learning *predictions* (Table 1) from remote sensing data, validated against a limited set of available ground truthing points, resulting in high uncertainty and substantial variability between different thematic classes. Yet, classification errors from satellite data are rarely accounted for; although recommendations and approaches have been developed to account for measures of uncertainty/precision where available (Simmonds et al., 2022).

2.2.2 | Data coverage

Differences in the evidence base for *driver* impacts partly reflect differences in data availability and coverage. A number of ground-based *driver* databases (e.g. SoilTemp (Lembrechts et al., 2020), LUCAS (d'Andrimont et al., 2020)) provide granular and historical data but often with relatively low and highly biased coverage. Modelled databases (e.g. CHELSA for climate (Karger et al., 2020) or Soil Grids for soil physicochemical properties (Poggio et al., 2021)) are another rich source of *driver* information and now tend to provide uncertainty estimates to help account for the highly unstructured and low coverage input data these models are based on. Thus, although some workarounds exist, accessing highly *structured* in situ or directly remotely sensed (rather than inferred) *driver data* with good spatio-temporal coverage remains challenging. While the situation continues to improve rapidly with new high-resolution, open-access satellite missions being launched and planned, some remotely sensed *driver* data remain inaccessible or expensive to use (Turner et al., 2015) or have too low resolution and grain size for ecological *attribution*. With satellite mission data only recently becoming publicly available and much in situ abiotic or socio-economic data not yet digitized, accessing historical data to match long biodiversity *time series* is challenging. Land use maps, for example, often present issues in temporal and thematic resolution (Daskalova et al., 2020) and lack the nuanced features needed to describe habitat quality or features important to different species (e.g. Lumbierres et al., 2022). In addition, there may be socio-technical barriers to obtaining information on specific *drivers*. For instance, the resolution of human population density data (e.g. CIESIN, 2018) typically matches political boundaries, which correspond to relatively small areas in Europe but large regions in Africa. Finally, important spatial and temporal mismatches remain between biodiversity and driver data, which prevent effective impact attribution. For instance, a widely used 30 m-resolution global deforestation database (Hansen et al., 2013) provides yearly data on forest cover, yet many species monitoring programmes (e.g. bird or mammal surveys) occur less frequently or at coarser resolutions (e.g. national park level). This causes spatial and temporal mismatches, as deforestation trends may show immediate habitat loss, but biodiversity declines often lag (e.g. species persisting in fragmented habitats before population's crash). Similarly, global fisheries data (e.g. Global Fishing Watch, FAO catch data) aggregate fishing effort at coarse spatial scales (~10–100 km) and annual intervals. However, species declines (e.g. sharks, tuna) are driven by localized, seasonally shifting overfishing hotspots (Kroodsma et al., 2018), making it difficult to link broad-scale fishing data to short-term species responses.

However, technological advances combined with state-of-the-art processing and interpolation methods (e.g. use of cloud-based processing solutions such as those offered on the OpenEO platform) are creating new opportunities to overcome *driver* data scarcity (Pettorelli et al., 2014), with improved mapping of land use intensity (Kuemmerle et al., 2013) pollution (Pennock, 2015; Tóth et al., 2016), weather data across Sub-Saharan Africa (Kaspar et al., 2022) and remote sensing data on soils (Lausch et al., 2019), landforms (Lausch et al., 2020) and hydrology (Bauer-Marschallinger et al., 2019). Ecologists are also exploring alternative data sources to complement remotely sensed data, including ground-based imagery (Morueta-Holme et al., 2024), citizen and community science data (Davison et al., 2024; Morueta-Holme et al., 2024) and broader geodiversity data (Schrodt et al., 2024).

3 | CHALLENGE 2: DETECTION OF BIODIVERSITY CHANGE

In the *detection-attribution* framework, *detection* refers to the identification of directional change in the system of interest (Table 1). In biodiversity research, this step involves estimating annual or long-term trends in biodiversity metrics and assessing the degree of support relative to a historical baseline, a counterfactual state or a no-trend scenario. The focus is generally more on estimating the magnitude and rate of change rather than a binary classification of change or no change. Trends may be estimated at the site level, but are more commonly generalized across multiple sites to derive an average rate of change that is representative of a wider area of interest. The method used to compile data significantly influences trend *detection* in biodiversity studies. Top-down approaches, which analyse global datasets using a standardized methodology, may be less sensitive to detecting biodiversity declines compared to bottom-up approaches, which synthesize local or regional studies through meta-analysis (Boënnec et al., 2024). This discrepancy arises because top-down analyses often aggregate coarse-scale data, potentially masking localized declines, whereas bottom-up approaches leverage high-resolution, context-specific data that capture finer-scale biodiversity changes. Furthermore, the choice between *structured* and unstructured *data* introduces distinct but overlapping challenges in change *detection*, influencing both the *sensitivity* and accuracy of inferred trends. Below, we explore some of these issues and present established and emerging strategies for the *detection* of biodiversity change while addressing these issues.

3.1 | Observational processes and sampling variability

Trend *detection* methods for *structured datasets* are relatively well-established (Dornelas et al., 2013; Martins et al., 2023), often using generalized linear models to describe long-term mean trends, generalized additive models for fluctuations over time or less commonly, threshold regression to identify abrupt changes (Chan et al., 2015). However, there is still variation in how well observational processes (i.e. sampling variability) are modelled (Chadwick et al., 2024). Ignoring observational processes is possible if the sampling methodology does not vary over time (i.e. high sampling structure, Figure 1); although the noise caused by sampling effects can still reduce the statistical power to detect a trend. State-space models (SSMs) can effectively minimize the noise caused by sampling variation (Kindsvater et al., 2018). For instance, random walk models in SSMs can account for complex data structures and can be fitted in both frequentist (e.g. Kalman filters, Laplace approximation methods) and Bayesian frameworks (e.g. Metropolis-Hastings samplers; see Auger-Méthé et al., 2021 for a guide to SSMs).

Unstructured data typically require more complex hierarchical models to account for sampling variation over space and time. If sampling variation is not accounted for, biodiversity trend estimates could be driven by changes in sampling rather than true changes in species populations or distributions. Hierarchical models such as occupancy-detection models explicitly model factors affecting *detection* probabilities (see MacKenzie et al., 2017 for a guide to occupancy models) and have been used to characterize trends in a wide range of taxa that lack large-scale *structured datasets* (e.g. most invertebrates; Outhwaite et al., 2020). Other methods adjust for sampling variation in effort over space and time based on the frequency of benchmark species, such as the frequency local scaling method (FRESCALO; Auffret & Svenning, 2022; Eichenberg et al., 2020; Hill, 2012). Using citizen science data, Isaac et al. (2014) found that occupancy-*detection* models and FRESCALO are both efficient methods for detecting temporal trends from noisy ecological data. Other approaches, such as thinned point process models (PPMs), similarly attempt to model both factors affecting species occurrence and the *detection* of those occurrences (Adjei et al., 2023; Bachl et al., 2019), although they have primarily been used for spatial models (see Wiegand and Moloney (2020) for a guide to PPMs) rather than trend *detection* (but see Seaton et al., 2024). Unstructured data can suffer from different types of measurement error. The above-mentioned methods primarily aim to deal with false absences (i.e.

when a species is present at a site but fails to be detected), but methods are emerging to deal with false positives (Kéry & Royle, 2020).

3.2 | Biases and data gaps

As illustrated above, biases and gaps exist in all types of biodiversity and *driver* data. Addressing these spatial, temporal and taxonomic data gaps is critical to draw broad inferences about biodiversity change (e.g. Henriques et al., 2020), and some assessment and reporting frameworks for assessing bias are emerging (e.g. Boyd et al., 2022; ROBITT). Whether data gaps lead to biased trend estimates depends on whether they covary with the underlying distribution of true trends (Bowler et al., 2024). Solutions to potential spatio-temporal biases due to data gaps include subsampling, weighting or imputation techniques (Bowler et al., 2024; Nakagawa & Freckleton, 2011). However, the strength and validity of such approaches are highly dependent on understanding the causes of data gaps and the quality and representativeness of the original data (Bowler et al., 2024; Schrodt et al., 2015; Ten Caten & Dallas, 2023). Many 'generic' imputation techniques assume the data are 'missing at random', an assumption that is rarely met in environmental change studies. For example, plant trait data available in databases are highly biased towards larger specific leaf area and larger seed mass, so bias correction methods need to be able to address this specific type of non-random bias (Johnson et al., 2021; Sandel et al., 2015; Schrodt et al., 2015). Similarly, approaches have been developed to specifically address non-random taxonomic bias (Henriques et al., 2020; McRae et al., 2017). Mismatches between the availability of *driver* data and biodiversity data add to the challenge of data gaps and highlight the importance of considering the coverage of heterogeneous *time series* in all facets of biodiversity and environmental change (Courter et al., 2013; Duchenne et al., 2022) (see section below on process-based modelling for *attribution* of historical trends in biodiversity).

3.3 | Appropriate resolution in time and space

Detecting a signal depends on the extent sampled along the axis of interest relative to the noise and the data resolution (Metcalf et al., 2021; Santini et al., 2017). Larger extents along the temporal or spatial axis facilitate trend detection despite the noise. Since the primary focus is often on identifying signals along environmental or anthropogenic gradients, the specific extent of the environmental gradient (e.g. variation in human impact) is more important than the spatial extent.

Temporally capturing peak periods of environmental change is challenging. For example, determining peak periods of land cover change requires attention to nuanced changes in land use intensity (Daskalova et al., 2020; Mihoub et al., 2017). Remote sensing products with higher spatio-temporal resolution can further increase statistical power and enhance signal detection within a given extent; however, increased resolution can also lead to increased noise, masking the true signal. Conversely, short temporal extents may not capture trends that reverse over longer periods of time, highlighting the need for appropriate temporal sampling resolutions and long-term studies that are aligned with ecological processes (Harte et al., 2015). For example, for birds/mammals, a few sampling events during the breeding season are sufficient to assess changes among years, but taxa with strong seasonal variation in activity require sampling across the activity period (e.g. flight period for butterflies) in order to account for within-year patterns before attempting to detect between-year patterns.

3.4 | Non-linearity

Mostly, *detection* of biodiversity change focuses on linear trends, which, for population or species abundance data, amounts to estimating mean annual population growth rates. Yet, growth rates naturally fluctuate over time (Duchenne et al., 2022; Rigal et al., 2020). For instance, several studies have documented the decline and subsequent increase in freshwater insect populations in Europe over the past 40 years (Haase et al., 2023; Outhwaite et al., 2020) Non-linear patterns can be more difficult to detect (Spake et al., 2022), but can have significant conservation implications. For example, a generally positive linear trend may not always indicate a growing population of no conservation concern if it is associated with a non-linear concave trend ('decelerating increase'), indicating a recent decline in abundance. In addition to non-linearity, non-monotony of trends also implies that the temporal extent of the data, particularly the baseline year, strongly impacts linear trend estimation (Daskalova et al., 2021; Duchenne et al., 2022). Similarly, a narrow temporal window increases the risk of capturing transient dynamics (e.g. Harte et al., 2015). Recently, frameworks have been developed to characterize and classify non-linear trends (Rigal et al., 2020) and abrupt shifts (Pélissié et al., 2024) in *structured* time series using polynomial effects and break detection methods, offering more consistent solutions to characterize the variability in trends. In the former, second-order polynomial functions are fitted to the trends using least squares regression and then summarized using metrics such as direction, acceleration, velocity and change points. In addition to the stable, linear and quadratic options defined in Rigal et al. (2020) and Pélissié et al. (2024) add an alternative by detecting step changes and selecting the best fit among the four models using the Akaike information criteria adapted for small *time series*.

4 | CHALLENGE 3: ATTRIBUTION

Change *attribution* typically refers to the process of identifying and assigning causality to the factors, events or processes ('*drivers*') that lead to an observed biodiversity change or outcome ('*response*'). In other words, a *causal relationship* is a 'situation where two variables are connected through some mechanism or means such that variations in one can propagate to subsequent variations in the other'

(Grace, 2024). Using the classic definition by Pearl (2009), 'a causal relationship exists when variation in an independent variable leads to, or causes, variation in a dependent variable, and this relationship is not spurious or confounded by other variables'.

In most cases, the interest extends to quantifying the magnitude and direction of these *causal relationships*. Causal biodiversity change *attribution* combines insights from change *detection* with data on biological, environmental or anthropogenic pressure variables. Recent studies have introduced ecologists to causal inference methods commonly used in other disciplines (Dee et al., 2023; Ferraro et al., 2019; Laubach et al., 2021). Here, we aim to contextualize causal inference methods within biodiversity change, discuss concepts that influence decisions and outline possible workflows using different data types on *drivers* and biodiversity *responses*.

4.1 | Analytical workflows for causal attribution of biodiversity change

Change *attribution* can be depicted as a multi-step process of data preparation/exploration, building the causal model (i.e. identification of the potential *drivers*, covariates and the biodiversity metrics), followed by quantifying the *causal effect* of *drivers* on the biodiversity metrics, and finally model evaluation and interpretation (Figure 2a). The choice of an *attribution* method depends on the direct *drivers*, potential covariates and biodiversity data characteristics, the question being asked and the prior knowledge of the system. Methods vary along a continuum from more exploratory to knowledge-based methods that may include several of the four steps (see Figure 2b for some examples from published studies). Below and in Box 1 we

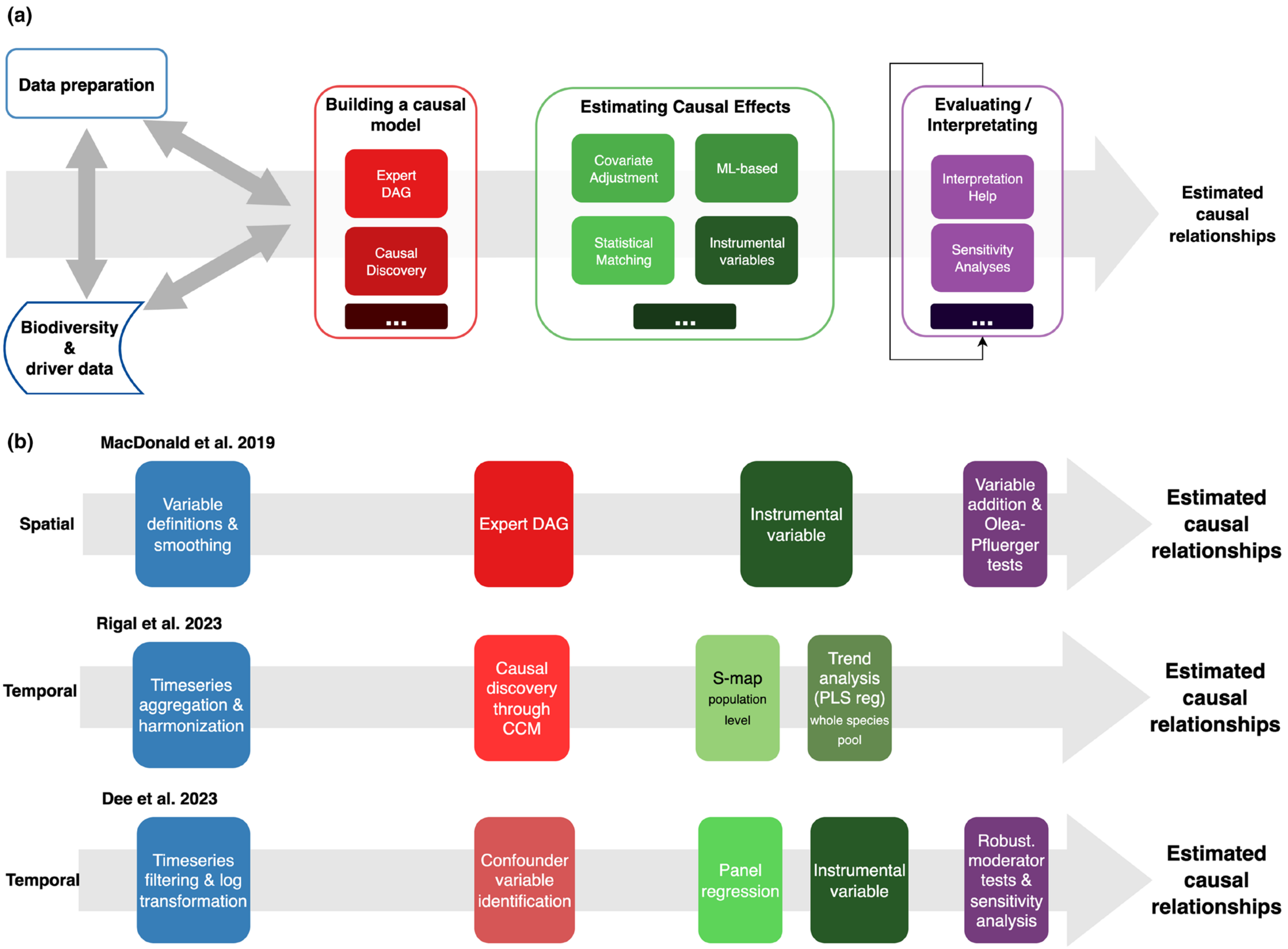

FIGURE 2 (a) Workflow from data to estimated causal relationships, providing key families of methods for each step. Depending on the data and knowledge available, the workflow may start with either data and their preparation (blue boxes) or causal model building (using either an expert DAG (dark red box) or Causal Discovery (light red box)), and several cycles of these steps may be required before continuing to the next one. Causal effect estimation can be performed using one or a combination of several methods (indicated by different shades of green boxes). The final step, evaluation and interpretation (purple boxes) is crucial for interpretation and should ideally be considered prior to causal effect estimation. Please note that the groups of methods shown here are only a few indicative examples, many other methods are available. (b) Key examples of applications of these steps in published ecological studies (Dee et al., 2023; MacDonald et al., 2019; Rigal et al., 2023). Colours correspond to the steps outlined in (a), with different groups of methods indicated by different shading (e.g. dark green box = instrumental variables). Further examples can be found in Figure S3.

BOX 1 Key considerations and steps for causal attribution in ecology

Defining the causal relationship

- Identify the focal relationship of interest (e.g. X → Y).
- Clearly define and state your causal question.

Formulating hypotheses and predictions

- What types of evidence are available?
- How much evidence do we have and how consistent is the evidence?
- How strong is the agreement across different lines of evidence?
- How well does it apply to the context under consideration?
- Are there competing hypotheses, and how do their predictions differ?

Mechanistic understanding and data contributions

- Does the data contribute to understanding causal mechanisms?
- Identify possible mediators and intermediate processes.
- Consider the depth of mechanistic understanding, including biological or social organization levels. For example, if you were modelling a community response, think of the mechanism in terms of population responses, too.

Constructing directed acyclic graphs (DAGs)

- Build expert DAGs and, where appropriate, compare/combine with data-driven causal discovery.
- Clearly communicate which research/source of information you base your causal knowledge and assumptions on.
- Address key challenges:
    - Bidirectional arrows and feedback loops.
    - Temporal and spatial scale dependencies.
    - Stability of observed patterns over time and space.
    - Known and unknown, measured and unmeasured confounders, colliders and mediators

Adjustment sets and method selection

- Determine the appropriate adjustment set (set of variables that should be controlled for based on the DAG for causal inference).
- Choose the right analytical method based on:
    - Nature of treatment variable (binary vs. continuous).
    - Nature of the data (spatial or temporal)
    - Confidence in identifying confounders and colliders.
    - Availability of data on confounders
    - The number of confounders relative to the amount of data
    - Hypothesized functional form of the causal relationship.
    - Availability of instrumental variables.
    - Feasibility of retrieving unbiased estimates.

Interpretation and reporting

- Clearly state assumptions of any methods used for detection and causal attribution and assess their robustness.
- Describe any potential measurement errors or biases and how they were addressed.
- Provide details of sensitivity analyses or robustness checks performed to assess the validity of the results.
- Discuss confounders, colliders and mediators and how they were controlled, including potential sources of bias (e.g. selection bias, reverse causality).
- Consider and discuss alternative explanations or competing causal pathways.
- Do not interpret and report the effects of control variables as causal effects (Table 2 fallacy).
- Discuss the generalizability of the findings to other contexts, populations or settings; conversely, limitations related to the sample, design or scope of the study that may affect the external validity.
- Clearly describe details if useful approximation approaches were applied (Table 1).

outline details of the considerations and choices that need to be made at each of these steps.

4.1.1 | Step 1: Data preparation and exploration

In biodiversity change *attribution*, the task of data preparation and exploration often interacts iteratively with the task of building the causal model until all relevant *causal relationships* are identified (Figure 2). For example, causal discovery methods (see Section 5.2.2) require data that conform to specific formats and assumptions, which necessitates preprocessing before identifying relationships (e.g. dealing with data gaps, see Section 3). The identification of missing relevant variables or covariates may require additional rounds of data collection and preparation, including the identification of proxy variables if measured data for the missing variables are not available. If all relevant variables are known at the start of a project, data collection and preparation can take place in subsequent steps.

4.1.1.1 | *Variable roles within biodiversity change attribution*

Simple models, such as a linear regression between a change in species' population abundance and a change in temperature, often fail to capture the complex causality in biodiversity change due to multiple *drivers* acting simultaneously and influencing each other (e.g. temperature and drought (Bowler et al., 2020; De Palma et al., 2018)). Complex relationships among key variables are common in the *attribution* of biodiversity change, which may explain why *attribution* is rather more challenging in ecology than in climate sciences. In climate *attribution* studies, radiative forcings are modelled as *exogenous* variables (external to the system), which greatly simplifies the inference of *causal effects* (Hannart et al., 2016). In biodiversity science, such exogeneity cannot be assumed because of the multiple cross-interactions between the variables of interest. It is thus key to a robust analysis to carefully identify the role of all variables in a study system before selecting the data and the appropriate model.

Within the framework of structural causal models, variables are classified into three main types: *confounders*, *colliders* and *mediators* (Table 1; Figure 3 (Laubach et al., 2021)). Ecologists have mainly focused on *confounders*, often overlooking the role of *colliders* and *mediators* due to a lack of awareness of their potential importance. Depending on the variable's role, different problems arise when the variable is excluded from the model (e.g. confounding, omitted variable bias) or included in the model (overcontrol, *collider* bias, M-bias, Table 2 fallacy) (Rinella et al., 2020). In ecology, a common modelling approach is to include all potentially relevant variables in a multiple regression model (a 'causal salad' regression) in order to control for potential confounding variables. However, this practice can lead to more biased estimates of variable effects than those resulting from potentially excluding key variables in an a priori selection (Cinelli et al., 2020). Building a DAG (Directed Acyclic Graph) can clarify the roles of different variables and help identify which variables need to be included in a model for causal inference (e.g. Guzman et al., 2024; Figure 3).

Confounders. *Confounders* occur when a third variable (Z) affects both the *driver* (X) and the biodiversity *response* metric (Y) (X ← Z → Y; Figure 3a). For example, a loss of species richness attributed to fragmentation alone could also be driven by habitat loss. Habitat loss confounds the overall effect of fragmentation on species richness ('direct effects' in Figure 3a). Similarly, the long-term recovery of plant and lichen species due to reductions in a specific air pollutant is difficult to attribute with certainty because of simultaneous declines in multiple pollutant intensities (particularly nitrogen and sulphur deposition), driven by concurrent pollution control policies ('indirect effects' in Figure 3a) (Dise et al., 2011). Spatial and temporal analyses can differ in their susceptibility to confounding. Spatial gradients in *drivers* can be useful for *attribution* because they are typically stronger than temporal gradients over sampling periods (Blüthgen et al., 2022; Oedekoven et al., 2017), but they are also more likely to be confounded by other covariates (Viana et al., 2022). Studies using time series can often better handle confounding factors because of weaker coupling of *drivers* and the fact that cause must precede effect.

Mediators. *Mediators* are variables (Z) that lie on the causal path between the *driver* (X) and the biodiversity metric (Y), creating indirect paths from exposure to outcome that break the flow of the information (X → Z → Y). They should not be conditioned when estimating a total effect; that is, data should not be stratified to consider specific/individual values of the *mediator* variable. Including *mediators* in biodiversity change *attribution* within a 'causal salad' regression can lead to overcontrol, blocking indirect pathways. For instance, macroclimate change affecting species may be mediated by microclimate buffering from forest cover (De Frenne et al., 2021), while urban night lighting may mediate the effect of city regulation on insect abundances (Figure 3b). In contrast, if the objective is to estimate the direct effect of city regulation, rather than the total effect of both night lighting and city regulation, indirect pathways should be blocked by conditioning on the *mediator*.

Colliders. *Colliders* (Z), or common effects, are shared effect variables of both the *driver* X and the biodiversity metric Y (X → Z ← Y). When conditioned on, these variables induce bias by opening up non-*causal relationships*, and consequently should not be controlled for, nor should their descendants. Generally, this occurs when a post-*treatment* variable (Table 1) is included in the model. For example, plant abundance may be a *collider* variable when studying the impact of ecosystem disturbance (e.g. a fire) on herbivores, since plant abundance can also be affected by herbivore abundance (Figure 3c).

Unobserved variables. The application of causal discovery algorithms necessitates the formulation of robust assumptions about the data-generation process, which in turn affects the data preparation. These assumptions typically include time series causal *stationarity* (constant mean, variance and autocorrelation structure over time, Table 1) and causal sufficiency (the absence of *unobserved confounders*, Table 1). While some methods can deal with unobserved *confounders* under

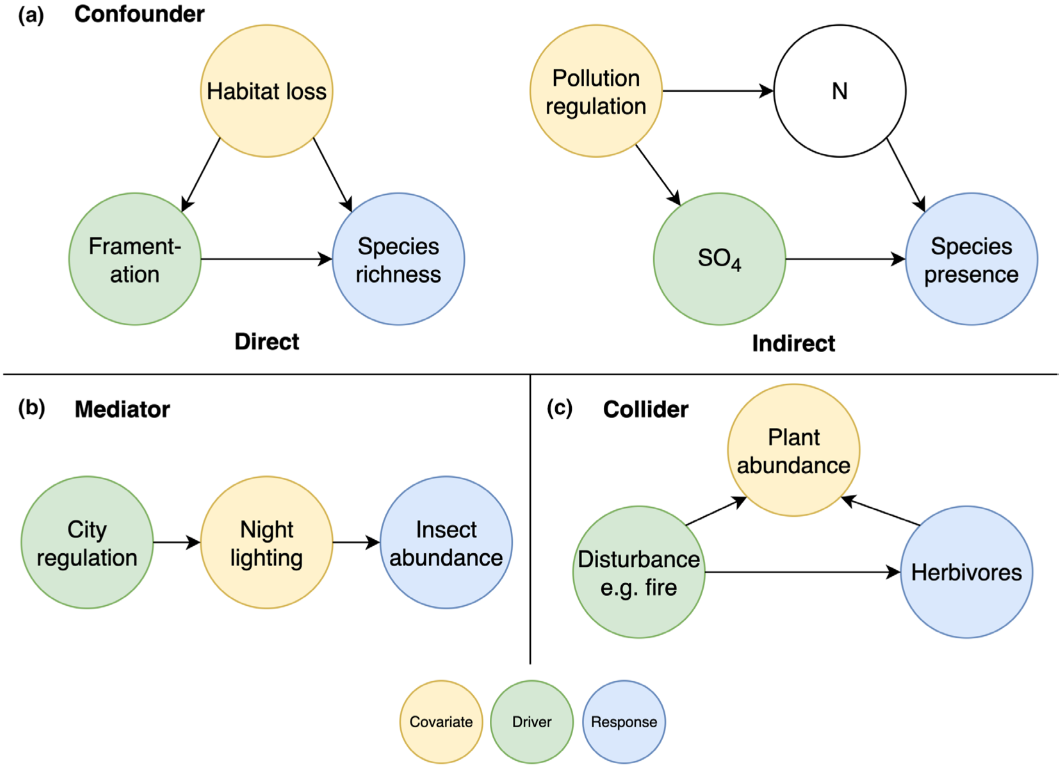

FIGURE 3 Examples illustrating the primary roles of covariates (yellow) interacting with a driver (green, the target of the causal effect estimation) and/or biodiversity response (blue). (a) direct (left) and indirect (right) confounders, (b) mediator and (c) collider covariates. In (a), confounding covariates must be included in the model to correctly estimate the causal effect of the driver. In (b), the mediating covariate could be included in the model to understand a possible pathway through which city regulation acts, but this would split the estimated total effect of city regulation. In (c), the covariate is a collider and should not be included in the model to estimate the effect of the driver.

additional assumptions, identifying the target effect (i.e. the specific *causal relationship* or quantity we aim to identify, measure or infer from our data) requires having access to variables measuring the processes that generate/cause the data. While these processes are rarely fully known or accurately captured with ecological data, we can assume, infer, and estimate them. Several methods have been adopted from other disciplines to do this using ecological data, including design adjustments (nested sampling and statistical designs) and causal diagrams (Byrnes & Dee, 2024).

Lag effects. Finally, lag processes or 'ecological memories' are ubiquitous in ecological systems but can be difficult to detect or incorporate. For example, forests may not show measurable *responses* to droughts until years after the drought event (*driver*) has occurred (Kuhn-Régnier et al., 2021; Pretzsch, 2022). Moreover, time lags can lead to spurious feedback loops when constructing *causal graphs* (see Table 1 and Figure S2, step 2 of the workflow). Similarly, lags can be of a spatial nature, often considered in terms of spatial buffers or landscape-scale effects. For instance, in freshwater systems, *drivers* acting upstream may affect biological *response* measurements downstream, that is, spatially removed from the *driver*. If these lag effects are not incorporated into a causal model, even very strong connections might be missed or at least misestimated. Methods such as signal regression offer efficient ways to estimate how the effects of a variable change with spatial scale or temporal lag and can be fit with commonly used generalized additive models (Wood, 2020).

4.1.2 | Step 2: Building a causal model

Directed Acyclic Graphs (DAGs) provide conceptual representations of causal assumptions of a system (Arif & MacNeil, 2022a; Ferraro et al., 2019; Laubach et al., 2021). Highlighting direct and indirect pathways helps identify variables as *colliders*, *mediators* or observed/unobserved *confounders* (Figure 3), without first making assumptions about the variable distributions (normal, non-normal) or form of the relationships between the links (i.e. linear, non-linear, abrupt). It is recommended to also include *unobserved variables* as incomplete DAGs may lead to inaccurate causal estimates/conclusions. In addition, the scale (temporal, spatial, organismal) at which phenomena are studied needs to be clearly defined, as mechanisms in biodiversity are strongly scale-dependent. For example, climate influences land use at global to regional scales, but locally, microclimate is strongly modulated by specific land uses and configurations, potentially reversing the main direction of causality (Figure S1). Opposing directions of causality can create unwanted feedback cycles in DAGs which are acyclic by construction. Careful consideration of the space and time-scale for which the DAG applies can help clarify the most relevant direction of causality (see Figure S2). DAGs can be built based on prior knowledge of the system ('expert DAG') or data-driven causal discovery algorithms. While expert DAGs require strong prior understanding of the study system, data-driven causal discovery can be applied when the underlying causal structure is not well established. The aim here is to identify potential causal relationships primarily from data; although some degree of theory and expert knowledge are required for robust causal discovery. This is especially important in complex or novel systems where the relationships among variables are not yet clearly understood. Expert DAGs and data-driven discovery are commonly combined.

4.1.2.1 | *Expert DAG*

In biodiversity sciences, there is often at least some basic domain knowledge about how variables relate to each other. The first step is therefore to put hypotheses down on paper in the form of a DAG, representing the variables present in the dataset and any other variables that are thought to be involved in the system, even if they are not initially connected to the graph. Second, a review of the literature should help to identify important missing covariates. Ideally, data matching the newly identified covariates can be collected and aligned with the existing samples to enrich the dataset (see Section 5.1). If not, such covariates remain unobserved and will affect effect estimation. When working within the structural causal model (SCM) framework, the causal Markov assumptions of the DAG need to be checked: are the (conditional) independencies drawn in the DAG verified in the data distribution? In other words, for any given relationship between a *driver* X and a *response* Y through a *mediator* M ($X \rightarrow M \rightarrow Y$), the full mediation model implies that Y and X are conditionally independent given M. Testing this assumption is difficult, especially in complex models. Statistical tests of the conditional independence (e.g. by simply regressing Y and X on M and assessing whether the residuals show non-zero correlation) are accepted as indicative (Textor et al., 2016). If the causal Markov assumptions are met, the DAG data are consistent, and an effect estimation method can be applied. Otherwise, the DAG should be refined to reflect the interdependencies present in the data. See Arif and MacNeil (2022a) for a comprehensive review in an ecological context.

Overall, expert DAGs are intuitive, but are sensitive to prior knowledge, with the risk of, for example, missing relationships or recent discoveries, or being biased when translating findings from the literature to other systems.

4.1.2.2 | *Data-driven causal discovery*

When researchers lack sufficient confidence in their representation of the data-generating process, or when little system knowledge is available, causal discovery or causal search can help identify pathways and directions of causality (Song et al., 2022; Spirtes & Zhang, 2016). Applying a causal discovery algorithm can be challenging. These models often rely on strong assumptions about the data-generating process underlying the dataset in question (causal sufficiency, faithfulness, etc.). The many configurable degrees of some models can reduce confidence when the parameter search is not guided by system knowledge or results in unstable outcomes.

The choice of method depends on the type (*spatial/cross-sectional* or *temporal/longitudinal* (Table 1)) and sample size of the data, assumptions about the data-generating process (*stationarity*, *contemporaneous effects* (Table 1)) and the need to handle *latent* (*unobserved* but inferable) *variables* or *confounders* (Table 1). Conditional independence-based (or constraint-based) methods which start with an undirected graph linking all variables and then remove or redirect links based on conditional independencies in the data include the Peter-Clark (PC) algorithm (Spirtes et al., 1993), its flexible adaptations for time series data (Runge, Nowack, et al., 2019) and the Fast Causal Inference algorithm (FCI (Spirtes et al., 1993)), which relaxes the causal sufficiency assumption.

Score-based methods (e.g. Greedy Equivalence Search (GES) (Chickering, 2003)), Fast Greedy Search (FGS) (J. D. Ramsey, 2015), or the Greedy FCI algorithm (Ogarrio et al., 2016) are computationally more intensive because they add the causal links one at a time to increase a fit score. Another class of models is functional causal models such as the semi-parametric non-linear additive noise models (Hoyer et al., 2008) that identify pairwise dependencies, or the linear LiNGAM method (Shimizu et al., 2006), which exploits the non-Gaussianity of the data. New methods recast causal discovery as an optimization problem, benefiting from recent deep learning advances in neural combinatorial optimization (Pamfil et al., 2020; Zheng et al., 2018; Zhu et al., 2019). For an exhaustive list of available causal discovery methods and software implementations, we refer the interested reader to table 3 of the review by Nogueira et al. (2022) and to Glymour et al. (2019) for a short illustrated review.

Complementary approaches, notably when assuming deterministic data processes and non-linear relationships, are state-space reconstruction approaches. The most widely used of the latter in ecology is convergent cross-mapping (CCM; Sugihara et al., 2012), which is based on empirical dynamic modelling and can help identify the direction of effects (discovery). CCM assumes that in a dynamic system the exposure variable contains information about its causal predictor and thus allows cross-mapping within a reconstructed state-space. A common workflow combines CCM with Smap for *causal effect* estimation, that is, gaining insight on both direction (positive or negative) and strength of the causal association (see Section 5.3). Many developments are available to tackle challenges present in ecological data such as short *time series* (multispatial CCM (Clark et al., 2015); Cross Map Smoothness (Ma et al., 2014); latent CCM (De Brouwer et al., 2021), synchrony (Convergent Cross Sorting (Breston et al., 2021)), bidirectional coupling (Continuity Scaling (Ying et al., 2022)) or *cross-sectional/spatial data* (Gao et al., 2023)).

4.1.3 | Step 3: Causal effect estimation (causal inference)

Once a *causal relationship*, in the form of a *causal graph* or *causal discovery*, has been established, the objective is to estimate the

TABLE 2 Key assumptions and considerations for the groups of causal methods presented in this article.

Method	Assumptions for causal interpretation	Key considerations
Structural equation modelling (SEM)	Correct model specification (all relevant paths and variables included) No omitted confounders Normally distributed errors Predictors don't correlate with error terms of the response Linear relationships (typically) The model must be identifiable (enough data and constraints to estimate all parameters uniquely)	Assumption that the structure reflects causal relationships. Coefficients represent conditional relationships, not just univariate effects. Model misspecification or omitted variables can bias estimates. Requires domain expertise for model building
Matching/weighting	Conditional Independence Assumption (CIA): all confounders observed Common support/overlap No hidden bias or unmeasured confounding	Balances covariates but does not control for unobserved confounders. Diagnostic checks (e.g. balance metrics, sensitivity analyses) are essential
Instrumental variables (IV)	Instrument relevance (strongly correlated with treatment) Instrument exogeneity (no direct effect on outcome, only via treatment) Exclusion restriction	Powerful for unobserved confounding, but finding valid instruments is difficult. Weak instruments bias estimates; validity must be critically evaluated
BACI/difference-in-differences (DiD)	Parallel trends (before-after change of groups would be the same in the absence of the treatment) Proper randomization No differential shocks (other than treatment)	Pre-trend testing and robustness checks are necessary. External validity can be limited
Regression discontinuity design (RDD)	Random effects are not correlated with any covariates Precise knowledge of the cut-off No manipulation around the cut-off Continuity in potential outcomes at cut-off	RDD gives causal effects near the threshold. Requires careful bandwidth selection and validation that units just above and below cut-off are similar
Bayesian networks/DAGs	Correct graph structure Causal sufficiency/conditional independence assumption (no unmeasured confounders) Faithfulness and Markov assumption	Can handle complex systems and sparse data. Can model effects of interventions. Depends on expert input or data-driven discovery with strong assumptions
Covariate adjustment (selection)	Known DAG and consistent with data No unobserved confounders Backdoors are closed	Effects of control variables must not be interpreted as causal effects (avoiding Table 2 fallacy)
Synthetic controls	Good pre-intervention fit No interference across units and consistent treatment definition No anticipation	Having sufficient pre-intervention periods to fit the synthetic control is key to the success of the method. Sensitivity analyses necessary
Fixed effects panel regression	No interference across units and consistent treatment definition Within-units covariate variability Site-specific but constant and/or time-varying but common unobserved confounding allowed	Ensuring that there is enough within-unit variation and that time-varying confounders have been included is key
Metalearners	No unobserved confounders No interference across units and consistent treatment definition Good estimation of nuisance parameters	Cross-fitting of the nuisance (intermediate) parameters and choice of the metalearner suited to dataset

TABLE 2 (Continued)

Method	Assumptions for causal interpretation	Key considerations
Empirical dynamic modelling (EDM)	Deterministic and stationary enough system Relevant dynamics observed: unobserved confounders of low influence Adapted sampling frequency	Suitable for time series without separability of variables. Can't separate causal effects from confounders and mediators
Invariant causal predictors	No unobserved confounders varying by environments Invariance assumption Multiple environments	Can handle non-linearity and unobserved variables shared between environments, needs variation across environments to identify non-causal associations
Granger causality	No unobserved confounders affecting the time series Stationary time series Prediction-based conception of causality	Time series should be pre-processed to be stationary and the sampling adapted to detect lagged dependencies

associated *causal effects*, that is, to quantify the respective magnitudes of *causal effects* acting on a given *response* variable. As described in the glossary, *causal effects* can be quantified using different metrics, depending on the scale, objective and methods of the study. In the following, we will adopt the generic term *causal effect*, as we believe that it is beyond the scope of this perspective to detail the different estimates associated with each method. It should be noted, moreover, that estimates depend on the assumptions made when constructing the graph. Ideally, graphs and estimates should be repeated over a range of likely assumptions to obtain robust results. Apart from repeatedly running the models, one can increase the robustness of the result and facilitate interpretation by applying methods before (e.g. statistical matching, instrumental variables, see below) or after (e.g. *sensitivity* analyses, see next step and Table 1) the causal model. Some of the more commonly used methods are outlined below, with assumptions for causal interpretation and key considerations given in Table 2. For a near-complete list of available methods and software for *causal effect* estimation, see table 12 in Nogueira et al. (2022).

4.1.3.1 | *Covariate adjustment*

Covariate *adjustment* (or selection) is the most widely used method for estimating *causal effects*. It consists of selecting the set of variables in a DAG to serve as model inputs in the next step. The structural causal model (SCM) framework suggests rules for doing this, such as the backdoor criteria (Pearl, 1995; Saavedra et al., 2022). This approach essentially means controlling for *confounders* (closing back doors) but not controlling for *colliders* (not opening the back doors) or for *mediators* (to avoid interrupting the information flow). Controlling for variables affecting the outcome but unlinked to the tested driver is not necessary to reach unbiased effect estimates. However, it can improve precision. Witte and Didelez (2019) review further selection procedures, which differ in their requirement for prior knowledge and/or the aim of the analysis. Once a set of controlling variables has been identified, the *causal effect* can then be estimated using traditional modelling approaches (i.e. regression models) or, for more complex scenarios, machine learning approaches (random forest algorithms, neural networks, etc.). In the case of *unobserved* confounding variables, there are further approaches using variation within site or years (Byrnes & Dee, 2024). The minimal adjustment set depends on the specific focal *causal relationship* of interest, which means that separate models might have to be constructed if there are multiple *causal relationships* of interest (Guzman et al., 2024). *Causal effects* should therefore be estimated one by one, each time using an appropriate adjustment set (Siegel & Dee, 2025). An important corollary is that the effects of the control variables cannot be directly interpreted as *causal effects*, to avoid the Table 2 fallacy (see Westreich & Greenland 2013 for further details).

4.1.3.2 | *Structural equation models*

SEM is a methodology originally adapted from traditional path analysis that provides an explicit connection between data and theoretical ideas, allowing for complex ecological systems and causal networks to be modelled. This includes modelling direct and indirect pathways to the *response*, as well as constructing *latent* and composite terms. SEMs are evaluated using a theory-driven iterative framework that evaluates models through tests of parsimony. Prior to estimation, a conceptual model is constructed from selected variables using ecological theory, expert opinion and existing literature (i.e. the proposed model). Proposed models are then assessed for relative model fit (e.g. comparing the model-implied and observed covariance matrices), and pathways are re-evaluated if model fit is not satisfactory. Essentially, this process ensures all confounding pathways on endogenous variables are included in the model prior to estimation (see Fan et al., 2016; Grace et al., 2010 for reviews).

There are several types of SEMs (e.g. covariance-based, piecewise and Bayesian), each with strengths and limitations. For example, piecewise SEM can accommodate non-linear pathways (e.g. mixed effects) but cannot accommodate *latent* terms (Lefcheck, 2016). Covariance-based SEM generally requires large datasets and may have limited scope for multi-level endogenous categorical variables due to model over-identification issues. These factors should be assessed and selected based on the data structure, target system and research question at hand. As with other regression-oriented techniques, estimates of *causal effects* within SEM may be biased if predictors correlate with the error terms of the *response* variable, which could be addressed with tools such as instrumental variables (see below) where appropriate (Grace, 2021).

4.1.3.3 | *Statistical matching*

Matching methods allow the effects of specific interventions to be isolated by accounting for potential confounding issues due to selection bias (i.e. the fact that interventions are typically non-randomly placed), prior to model application (Stuart, 2010). These methods essentially adjust for differences in covariates between *treatment* levels to create a design that more closely resembles a randomized experimental design. Matching can also be used to improve the rigour of other quasi-experimental approaches, for example, to select appropriate control units for difference-in-difference or BACI analyses. In ecology, such methods are already commonplace for assessing the effectiveness of protected areas (Schleicher et al., 2020). Often, such matching of *treatment* and control sites is done intuitively during study design (Redhead et al., 2022), but statistical matching methods also allow this to be done after data collection and are thus useful when relying on existing datasets/databases (Emmons et al., 2022). This method involves deriving a propensity score (PS), defined as the probability of *treatment* assignment given the observed covariates (Ramsey et al., 2019), which can be used to group similar control and *treatment* observations—even with high-dimensional covariates—and allow unbiased effect estimation under the causal sufficiency assumption. While originally developed for single binary *treatments*, there are extensions to handle continuous (generalized PS or GPS (Zhao et al., 2020)) or multiple *treatments* (McCaffrey et al., 2013). However, propensity scores do not always perform better than multiple regression as a way to control for *confounders*, at least with large datasets (Wilkinson et al., 2022).

4.1.3.4 | *Instrumental variables*

In some cases, the *treatment* itself might be difficult to directly relate to an outcome, either because of possible reverse causation (i.e. if the outcome also affects the *treatment*) or omitted confounding variables (e.g. affecting both the *treatment* and the outcome). In these cases, an instrumental variable (IV) that explains *treatment* assignment but is not associated with other variables might be a better option to test for a *causal effect* of the *treatment* (Grace, 2021; MacDonald et al., 2019). An IV must satisfy two conditions (Kendall, 2015); it must: (i) be a cause of the *treatment* variable of interest (*relevance*) and (ii) have no *causal effect* on the outcome variable other than through the target pathway (*exclusion condition*). The second condition often severely limits the candidate instruments. IVs are then used in a two-step process consisting of, first, predicting unconfounded *treatment* levels and, second, estimating the *treatment* effect on the outcome based on the predicted *treatment*. However, the *causal effect* is only estimated from the subset of *treatment* that is actually affected by the IV (Dee et al., 2023). Once identified, IVs can be integrated in different models, such as regression (Kendall, 2015) or SEM (Grace, 2021). While IVs are appealing, selecting variables that satisfy the above two conditions is challenging, which has resulted in IVs not being used as much as other causal inference approaches.

4.1.3.5 | *Other quasi-experimental methods*

When trying to estimate *causal effects* but facing confounding effects between *treatment* and outcome biases, quasi-experimental or natural experiments offer practical solutions to mimic a randomized experiment (Butsic et al., 2017; Larsen et al., 2019). Before-After-Control-Impact (BACI) designs are appropriate when *longitudinal (time series) data* are available both before and after the *treatment* assignment; for example, the implementation of a specific intervention in conservation science. The difference-in-differences method, a term more common in econometrics, focuses on the before-after change of the difference between control and *treatment* groups within a BACI design. The key assumption is that the before-after change of both groups would be the same in the absence of the *treatment* (the parallel trends assumption). For instance, such an approach has been used to study the effects of protected areas on species trends (Adams et al., 2015; Wauchope et al., 2022). The BACI design can be replicated retrospectively using statistical matching methods. Regression discontinuity (RD) design is a related approach but with different assumptions. RD is used when there is a clear breakpoint, or discontinuity, in *treatment* allocation, either in space or time, that separates observations into control and *treatment* groups (Wuepper & Finger, 2022). In ecology, this design has been used to test the effects of a wildfire by comparing burned and unburned areas (Butsic et al., 2017). The key assumption is that local randomization holds across '*treatment*' groups around the small interval of *treatment* allocation, which means that it most resembles a natural experiment; the *treatment* effect can then be estimated by comparing the *response* immediately below and above the breakpoint. If the effect of interest is measured across sites with repeated measures over time, fixed effects panel regression can be used (Jones & Lewis, 2015; Larsen & Noack, 2017; Ratcliffe et al., 2024), which uses site or time fixed effects to control for *unobserved* confounding effects that are assumed to be site-specific but time-invariant, or time-specific and common to all sites. In ecology, often random effects are included for site or year but these can lead to violations of the assumption that random effects are not correlated with any covariates, preventing unbiased estimates of causal factors (Antonakis et al., 2021). Fixed effects design variants adapted to different assumptions on the nature of observed and *unobserved confounders* are another option (Byrnes & Dee, 2024).

4.1.3.6 | *Machine learning based methods*

In addition to identifying *causal relationships* as seen in Section 5.2, machine learning algorithms can contribute to quantifying unbiased *causal effects*. This is not to be confused with the field of causal machine learning (Kaddour et al., 2022), where causal inference helps solve machine learning problems to achieve better predictive power; this is beyond the scope of this paper.

One family of techniques for estimating conditional *causal effects* (Table 1) is metalearners (Künzel et al., 2019), which decompose the task of estimating *causal effects* into basic algorithms such as random forests or neural networks. In ecology, these are especially useful for estimating spatially-varying effects of causal factors, or even just describing patterns of biodiversity trends at fine spatial resolution (Fink et al., 2023). There are different metalearners, each adapted to specific data properties and assumptions (Caron et al., 2022). Double or debiased machine learning (Chernozhukov et al., 2018) allows for effect estimation in a two-step process based on the Frisch-Waugh-Lovell theorem: the *causal effect* is estimated separately from the nuisance parameters, that is, the influences of the other variables on the outcome (Fink et al., 2023). Under the assumption of strong ignorability (i.e. causal sufficiency and overlap or positivity, meaning the *treatment* assignment probability is strictly between zero and one), causal forests (Hahn et al., 2020; Wager & Athey, 2018) allow obtaining unbiased causal estimates. Finally, the benefits of supervised learning can be combined with the additional information provided by instrumental variables (Hartford et al., 2017).

4.1.3.7 | *Other methods*

The empirical dynamic modelling (EDM) framework is suited for *time series* data when separability of variables is not given (i.e. the influence of individual variables on an outcome can not be clearly distinguished and isolated) and interaction strengths or signs may change over time. Convergent cross-mapping (CCM), which we discuss in Section 5.2 as a tool to identify the existence and direction (i.e. positive or negative) of a causal link between a pair of variables (possible *driver* and *response*), can be combined with Sequential Locally Weighted Global Linear Maps (S-maps) (Deyle et al., 2016; Sugihara, 1994) to estimate the strength of the identified causal relationship. Compared to other similar *prediction* techniques for *time series* (i.e. DLM), ecosystem states are weighted based on their actual similarity, that is, their distance in the reconstructed state-space, rather than their neighbours in time. However, to our knowledge, methods from the EDM framework do not have a proven ability to disentangle *causal effects* from *confounders* and *mediators*, as done in the SCM or potential outcome frameworks under precise hypotheses. They rely on different assumptions, including full observability (i.e. all relevant system variables are measured or recoverable), low-dimensional deterministic dynamics and sufficient data coverage of the system's state space.

Invariant causal *prediction* methods (Peters et al., 2016) exploit the assumption of invariance, that is, the cohesion of the direct *causal effects* on a *response* variable in subpopulations or changing environments, to identify and estimate causal structures. It relies on common concepts and methods, and helps address challenges such as non-linearity (Heinze-Deml et al., 2018), *unobserved* variables (Peters et al., 2016) or sequential data (Pfister et al., 2019).

Where continuous geospatial data are available, for example, through remote sensing, using information theory in combination with causal methods can be useful (Vernham et al., 2023). For instance, temporal pattern analyses were used together with Markov models within a sliding window approach to attribute the impacts of ecological engineering and climate change on carbon uptake (Li et al., 2023).

Bayesian networks offer another promising avenue for ecological applications. Bayesian networks, also known as belief networks or Bayes nets, are DAGs in which nodes represent random variables and edges represent probabilistic dependencies between these variables (Darwiche, 2008). Bayesian networks can simultaneously integrate expert knowledge with statistically significant information learned from data, handle complex systems with many variables and handle missing data, providing a scalable approach to modelling high-dimensional data (Borsuk et al., 2004). Another advantage is that they can model the effects of interventions (e.g. policy changes) on outcomes using 'do-calculus', which allows predicting the impacts of hypothetical changes, answering what if questions and thus supporting counterfactual reasoning. However, Bayes nets can be computationally demanding—as the number of variables and nodes increases, the complexity of computations for exact inference (such as marginalization and updating beliefs) can grow exponentially. A possible solution is either careful pre-selection of variables or application of distributed platforms such as DistriBayes (Ding et al., 2023). A potentially greater barrier that Bayes nets share with many other *attribution* methods is the inability to integrate feedback loops and the reliance on conditional independence assumptions that may not hold in all real-world scenarios.

Finally, some methods are not strictly causal in the statistical sense but rely on predictability to infer or orientate causal links. The Granger causality framework (Granger, 1969) does not directly provide an *attribution* in the mechanistic sense but instead tests a predictive relationship in *time series* data, that is, whether one variable has predictive power over another, not whether X directly causes Y in a causal or mechanistic sense. Its utility thus lies in its ability to highlight potential *causal relationships* that warrant further investigation, often in combination with other methods that are more suited for *attribution*. Its main limitation apart from this is the requirement of separability, that is, that the *driver* variable is independent from the *response* variable. Variants using random forests can deal with non-linear (Chen et al., 2004; Papagiannopoulou et al., 2017) or short multivariate time series (Wismüller et al., 2021), while extensions relying on the time-frequency approach aim to identify periodic coupling (Detto et al., 2012) or anomalous events (Shadaydeh et al., 2019).

4.1.4 | Step 4: Evaluation and interpretation

The *detection* and *attribution* of directional effects requires assumptions. The plausibility and strength of such assumptions should be assessed, where possible, using *sensitivity* techniques. This step is key to providing confidence in the significance of the estimated effects.

4.1.4.1 | *Sensitivity analyses*

Sensitivity analysis is about assessing the robustness of the key assumptions behind the effect estimation methods. It evaluates how strong unconsidered *confounders* would have to be to explain away the association, that is, how strong the confounding effect between the *treatment* and outcome would have to be for the association between *treatment* and outcome to be non-causal (VanderWeele & Ding, 2017). While the *overlap assumption* (Table 1) can be tested from the observed data, causal sufficiency can only be assessed indirectly. Testing how the results would be affected if the assumption of unconfoundedness is violated can provide confidence or necessary caution in the following interpretations. Various methods have been developed since the 1960s (Cornfield et al., 1959), including bounces for the *treatment* effect (Rosenbaum, 2010) implemented in the rbounds R package (Keele, 2022), but the more recent E-values and robustness values, from VanderWeele & Ding (2017) and Cinelli et al. (2020) respectively, represent a major step forward by dropping strong assumptions about the nature of the unmeasured *confounders*, providing a clear reporting format to facilitate communication and facilitating interpretation considering pre-existing domain knowledge (see also Table S1). Ecological applications of sensitivity tests include, for example, the Oster method (Dee et al., 2023) and the Cinelli method (Andraczek et al., 2024).

Refutation methods test the robustness of the estimate by modifying the data and observing how the effect is affected. For example, adding an independent random variable as a common cause or replacing the data set with a randomly selected subset/bootstrapped samples should not significantly change the estimated effect. However, replacing the true *treatment* variable or the true outcome variable with an independent random variable (or placebo *treatment* and dummy outcome) should drive the estimated effect to zero. The R package DoWhy (Blöbaum et al., 2024) allows such tests to be performed with convenient functions and documentation.

Finally, partial identification consists of considering different assumptions and constructing confidence sets instead of precise point estimates based on possibly unverified assumptions (Tamer, 2010). This *sensitivity analysis* can be seen as a step in uncertainty quantification, where the implausible assumptions are relaxed and the causal model estimate is allowed to vary within its logical limits.

4.1.4.2 | *ML interpretability methods*

When data-driven machine learning approaches are used to predict biodiversity patterns, traditional interpretability (explainable AI) methods can help to explain how model outputs depend on inputs (Moraffah et al., 2020). Such methods can disentangle the relative importance or the local contributions of *drivers* and their interactions. Global model-agnostic methods, e.g. Partial Dependence Plots (Friedman, 2001) and Permutation Feature Importance (Breiman, 2001; Fisher et al., 2018), can be used to explain the average behaviour of *drivers* in the machine learning model. Local interpretability tools, including Local Surrogate models (LIME; Ribeiro et al., 2016) and Shapley values (Lundberg & Lee, 2017) can decompose the importance of *drivers* for individual *predictions*. Although derived from predictive tasks, these measures of *driver* importance are often interpreted as practical model proxies for *driver* effects on a given outcome variable. These techniques were indeed not designed for causal inference methods where the objective is to estimate a precise effect in an already identified data-generating process. However, when applied after careful variable selection informed by a causal diagram, confidence in the usefulness of these metrics can be increased. More recently, causal interpretable models have moved towards a new goal in counterfactual explanations which involve asking 'what if' questions ('what if I had taken a different action, would the outcome have been different?') (Guidotti, 2024; Moraffah et al., 2020). As such, counterfactual explanations are particularly valuable as they can provide actionable insights, for example, in a conservation management context. This emerging set of models and interpretation techniques aims to bridge the gap between the importance and partial effect of variables in *prediction* tasks and effect estimation techniques.

5 | DISCUSSION

Rather than providing a fixed template for model building and parameter estimation, we have identified a general workflow using a selection of key methods and their requirements. The final choice for a *detection* and *attribution* approach must be made by the researcher, as it will depend on the characteristics and amount of data available, the objective of the study, the current understanding of the study system, the ecological question being addressed and the modelling expertise available in the team. This selection can and indeed should be a dynamic process, moving back and forth between steps in the workflow and testing different methods until proceeding to the final analysis. Although it may be tempting to use some of the more advanced methods (e.g. causal discovery), they should not be used as a black box; results should not be interpreted blindly. Ultimately, causal inference and *attribution* are about improving understanding of the system, which includes awareness of model assumptions.

5.1 | Causal model/DAG, facilitation beyond the analysis

Irrespective of the chosen methods and recognizing that this is not always strictly necessary, we encourage the construction and communication of a causal model, that is, a DAG. Prior to analysis, causal models help to identify *confounders*, *colliders* and/or *mediators* and thus aid in selecting an appropriate method for robust analyses. A causal model can also aid decisions on data collection, as (confounding) variables may be identified that are worth collecting and/or integrating into monitoring programmes. The sharing of causal models could also support (cross-disciplinary) collaborations when links towards other disciplines become visible (Dolby, 2021). Once published, a formalized DAG makes it easy to follow the authors' assumptions and rationale for including or excluding certain variables. It also facilitates the comparison between studies and verification that scientific (expert, literature) knowledge has been correctly integrated. However, it is important to note that the development of an expert DAG requires extensive knowledge of biological phenomena, while in reality, our understanding is often superficial. This may lead to the omission of important links or *mediators* and thus to erroneous conclusions about causality. Moreover, DAGs may appear statistically sound but be fundamentally incorrect, thereby reinforcing our erroneous beliefs. Measurement errors, for example, might result in misinterpretation and biologically incorrect DAGs. They should therefore be used with care and a full understanding of the limitations and caveats inherent in DAGs.

5.2 | Reporting uncertainty

Due to the complexity of ecological systems, a *detection* and/or *attribution* analysis will always involve some degree of uncertainty. This is either due to noise and bias, especially in observational data (Boyd et al., 2022; Chadwick et al., 2024; Moudrý et al., 2024), due to assumptions associated with model development (e.g. building a flawed DAG that meets modelling assumptions but may be biologically incorrect), or those associated with model selection (Copas & Eguchi, 2020). The same holds for the spatial and temporal scales addressed in the study, as both trends and *drivers* depend on these dimensions (Johnson et al., 2024; Figure S2). Awareness of uncertainty and scale are key to robust interpretation of the results and should, therefore, be clearly stated (Boyd et al., 2022; Gonzalez et al., 2023; Pescott et al., 2022).

5.3 | Bias-variance trade-off

When studying causal inference, one might assume that achieving an unbiased estimator with minimal variance is the ideal objective. This perspective contrasts with machine learning, where model selection is often driven by minimizing mean squared error (MSE) above all else.

However, in practice, causal inference does not strictly require unbiased estimators. In many cases, estimators with some bias but a lower MSE may be preferable. For instance, in matching methods, the choice of the number of matches directly involves a bias–variance trade-off (Stuart, 2010). The concept of a *useful approximation* standard where estimates that are predominantly causal and at least exceed the likely degree of bias are acceptable is promoted by some (see, e.g. Grace, 2024) but controversial. Importantly, they should not distract from the optimal goal of moving towards unbiased estimates (Siegel & Dee, 2025). Thus, causal inference requires a nuanced approach, balancing bias, variance and practical applicability.

5.4 | To attribute or not to attribute

While causal inference methods offer promising avenues for ecological and global change research, they are not always applicable, nor are they often the best possible choice among many possible inference approaches. In addition to the challenges discussed above, the complexity of certain phenomena is a major limitation. In real-world cases, variables may influence each other (undirected) or mechanisms may involve feedback loops (although there are approaches to deal with the latter e.g. Bongers et al., 2016; Wang et al., 2018). These two relatively common cases violate the classical assumptions of DAGs, limiting the application of these methods. The reliance on proxy variables (e.g. use of human population density or travel distance to large cities acting as proxies for direct human disturbance/exploitation) and inappropriate resolutions (i.e. *driver* data not matching ecological *response* scales) further weaken analyses of *causal relationships*. Most importantly, DAGs can become overly complicated, making it impractical to test complex *causal relationships*. Here, again, inspiration ought to be sought from other, equally complicated systems in other disciplines, such as public health and economics that managed to establish a culture of practicing more robust *attribution* than ecology has so far.

Classical predictive approaches, such as model selection based on information criteria, have long been prevalent in ecology and frequently used to infer causality. However, increasing awareness highlights the limitations of such interpretations, as they can be misleading (Hone & Krebs, 2023). Beyond potential biases arising from omitted variables and spurious correlations, a fundamental distinction exists between statistical and causal inference: while statistical inference prioritises predictive accuracy, causal inference seeks to elucidate the underlying mechanisms driving observed changes. As a result, models incorporating all available variables may appear superior in terms of predictive performance yet remain misaligned with the structural relationships underpinning ecological processes (Arif & MacNeil, 2022a). A more rigorous validation framework is therefore essential, enabling model selection not merely based on predictive power but through explicit comparisons between causal inference methods and traditional predictive models applied without causal considerations (see workflow examples in Figure S3). Such comparative analyses underscore the distinct objectives and divergent conclusions that emerge when causal *attribution* methodologies are explicitly incorporated into ecological modelling.

Causal inference, by contrast, is grounded in explicitly stated assumptions derived from prior knowledge or exploratory data analysis. These assumptions frame hypotheses that must be carefully formulated to avoid reinforcing pre-existing cognitive or methodological biases. Poorly constructed hypotheses—whether overly broad, incorrectly specified or based on flawed premises—can obscure true causal mechanisms and skew inference. While causal inference methodologies rigorously define and test hypotheses about causality, they ultimately rely on classical statistical techniques for effect estimation. Consequently, their validity is contingent on both the robustness of the underlying assumptions and the appropriateness of the chosen methodological framework. Different causal inference approaches—such as Granger causality, convergent cross-mapping (CCM) and other *prediction*-based techniques—embed distinct conceptualizations of causality, which may not always align with the specific demands of a given research question (McCracken & Weigel, 2014). Therefore, results must be critically interpreted within the context of the methodological definition of causality employed.

Simpler but well-designed correlative models may thus be a better choice, for instance, as a first step to understand a system when little is already known (Currie, 2019; Nilsen et al., 2020). Such models create a first layer of scientific evidence of causal links in a system, including possible interactions of variables, which can then strengthen causal discovery or inference approaches as the data and knowledge grow. Correlative models may also be easier to interpret while providing similar results. For instance, propensity score matching may produce the same results as multiple regression in many, but not all, circumstances. Local experiments can further support intuitions where feasible. For more complex processes occurring on larger, less manageable scales, mechanistic models can complement correlative understanding by testing ecological processes in silico. Testing multiple causal inference models can offer additional support.

In conclusion, a comprehensive causal understanding of ecological systems requires both rigorous data analysis methodologies and a robust mechanistic framework (Grace, 2024). These components collectively enhance causal *attribution*, where simpler analytical approaches serve as foundational tools for refining complex causal models, revealing previously overlooked interactions and improving inference precision (Goldberg, 2019; Shipley, 2016). By integrating mechanistic insights with advanced statistical techniques, researchers can more effectively disentangle direct and indirect ecological *drivers*, thereby advancing predictive and explanatory capacities in ecological research.

6 | PERSPECTIVE/OUTLOOK

6.1 | A call for cross-disciplinary collaboration and method transfer

Some *attribution* methods, such as SEMs, are already well established in ecology, while others have great potential but are not yet routinely used in ecological studies (e.g. CCM). The focus of causal discovery research on high-dimensional and mixed data types is particularly promising for biodiversity data studies.

As many of these methods originate from other research fields characterized by different types and quality of data, future studies should test and compare these methods using simulated data representing hypothetical mechanisms under different ecological scenarios of change and/or in situ data from well-studied and understood systems. This will allow us to assess their ability to capture *causal relationships*, their robustness to missing variables and noise in the data, and their data requirements in terms of resolution, coverage and sample size. Crucially, we want to encourage curiosity about methods developed in other disciplines, and collaboration between experts in ecological data and systems and experts in *attribution* and *detection* methods.

6.2 | Better (rather than more) data for *detection* and *attribution*

We have firmly entered the era of big data in ecology, as evidenced by the rise of semi-automated methods such as bioacoustics, camera trapping, eDNA and remote sensing, which will undoubtedly lead to even greater availability of ecological data in the future. But, as highlighted above, *driver* data at sufficient spatio-temporal resolution are lacking and are often a bottleneck for *attribution* studies, and biodiversity data gaps continue to limit the *detection* of change. We therefore need better, not necessarily more, data for *detection* and *attribution*. An easy win would be better collection, sharing and inclusion of metadata (e.g. eBird has fostered more detailed metadata collection than many other citizen science programs outside of *structured* schemes), including change over time and the collection of 'co-located data' on *drivers* and biodiversity (including individual attributes such as traits) at the same sampling locations. Constructing a DAG can help to identify limitations of the dataset/analysis and may also indicate where sampling efforts/monitoring could be increased to collect the right variables and improve the model (Arif & MacNeil, 2022a).

6.3 | Linking causal and predictive frameworks

Predictive frameworks such as scenario models and digital twins are often process-based but currently lack the integration of statistical causal frameworks. This is unfortunate for several reasons. First, embedding causal models into predictive frameworks can help enhance our understanding of underlying causal mechanisms and thus reduce uncertainty around *predictions* of future states. Second, incorporating causal models in predictive frameworks helps to understand how different variables influence outcomes, providing a more robust basis for applying possible interventions in specific contexts (e.g. predicting the impacts of policy intervention on future biodiversity gains). By knowing not only what might happen, but also why, optimal strategies can be tailored to address root causes rather than symptoms. Finally, predictive models that incorporate causal frameworks can better anticipate unintended consequences and identify leverage points for interventions before consequences are realized, resulting in more robust and resilient management approaches. Better integration of *attribution* approaches with predictive ones will lead to more accurate *predictions*, better decision-making and a more comprehensive understanding of ecological dynamics and human impacts. A range of causal predictive approaches are emerging, including causal forests and metalearner algorithms (see Section 5.3.6 (Künzel et al., 2019)), that offer promise to tackle these challenges.

AUTHOR CONTRIBUTIONS

Franziska Schrodt and Wilfried Thiller acquired the funding; Franziska Schrodt administered the project; all authors conceived the ideas and designed the methodology; Franziska Schrodt, Miriam

Beck and Joachim Estopian led the writing of the manuscript with Diana E. Bowler, Colin Fontaine, Pierre Gaüzère, Matthias Grenié, Romain Goury, Ines S. Martins, Naia Morueta-Holme, Luca Santini and Wilfried Thiller contributing significantly. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

This research is the product of the IMPACTS working group led by WT and FS, funded by the synthesis centre CESAB of the French Foundation for Research on Biodiversity (FRB; www.fondationbiodiversite.fr). We also acknowledge funding from the Horizon Europe OBSGESSION (No. 101134954) project. FS was supported by NERC grant no. 09045. NMH was supported by research grant no. 2064-00091B from the Independent Research Fund Denmark and grant no. VIL53086 from the VILLUM FOUNDATION. ISM was supported by Royal Society University Research Fellowship grant no. URF\R1\241013.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.70131.

DATA AVAILABILITY STATEMENT

No data or code were used in this paper.

ORCID

Franziska Schrodt https://orcid.org/0000-0001-9053-8872
Miriam Beck https://orcid.org/0000-0001-8179-3820
Joaquim Estopinan https://orcid.org/0000-0001-9766-3887
Diana E. Bowler https://orcid.org/0000-0002-7775-1668
Colin Fontaine https://orcid.org/0000-0001-5367-5675
Pierre Gaüzère https://orcid.org/0000-0003-1259-6131
Romain Goury https://orcid.org/0009-0001-4051-9937
Matthias Grenié https://orcid.org/0000-0002-4659-7522
Inês S. Martins https://orcid.org/0000-0003-4328-7286
Naia Morueta-Holme https://orcid.org/0000-0002-0776-4092
Luca Santini https://orcid.org/0000-0002-5418-3688
Mickael Hedde https://orcid.org/0000-0002-6733-3622
Gabrielle Martin https://orcid.org/0000-0002-9192-4924
Emmanuelle Porcher https://orcid.org/0000-0002-9264-8239
Sara Si-Moussi https://orcid.org/0000-0002-0519-8699
Marianne Tzivanopoulos https://orcid.org/0000-0001-6931-2879
Grant Vernham https://orcid.org/0000-0002-9075-0045
Cyrille Violle https://orcid.org/0000-0002-2471-9226
Wilfried Thuiller https://orcid.org/0000-0002-5388-5274

REFERENCES

Adams, V. M., Setterfield, S. A., Douglas, M. M., Kennard, M. J., & Ferdinands, K. (2015). Measuring benefits of protected area management: Trends across realms and research gaps for freshwater systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1681), 20140274.

Adjei, K. P., Mostert, P., Parada, J. S., Skarstein, E., & O'Hara, R. B. (2023). The point process framework for integrated modelling of biodiversity data. *arXiv [stat.AP]*. arXiv. http://arxiv.org/abs/2311.06755

Andraczek, K., Dee, L. E., Weigelt, A., Hinderling, J., Prati, D., Le Provost, G., Manning, P., Wirth, C., & van der Plas, F. (2024). Weak reciprocal relationships between productivity and plant biodiversity in managed grasslands. *Journal of Ecology*, *112*, 2359–2373.

Antonakis, J., Bastardoz, N., & Rönkkö, M. (2021). On ignoring the random effects assumption in multilevel models: Review, critique, and recommendations. *Organizational Research Methods*, *24*(2), 443–483.

Arif, S., & MacNeil, M. A. (2022a). Predictive models aren't for causal inference. *Ecology Letters*, *25*(8), 1741–1745.

Auffret, A. G., & Svenning, J.-C. (2022). Climate warming has compounded plant responses to habitat conversion in northern Europe. *Nature Communications*, *13*(1), 7818.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., & Thomas, L. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, *91*(4), e01470.

Bachl, F. E., Lindgren, F., Borchers, D. L., & Illian, J. B. (2019). Inlabru: An R package for Bayesian spatial modelling from ecological survey data. *Methods in Ecology and Evolution*, *10*(6), 760–766.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., & Wagner, W. (2019). Toward global soil moisture monitoring with Sentinel-1: Harnessing assets and overcoming obstacles. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(1), 520–539.

Bayraktarov, E., Ehmke, G., O'Connor, J., Burns, E. L., Nguyen, H. A., McRae, L., Possingham, H. P., & Lindenmayer, D. B. (2019). Do big unstructured biodiversity data mean more knowledge? *Frontiers in Ecology and Evolution*, *6*, 239. https://doi.org/10.3389/fevo.2018.00239

Bindoff, N. L., Stott, P. A. A., Achutarao, K. M., Allen, M. R. R., Gillett, N., Gutzler, D., Hansingo, K., Hegerl, G., Hu, Y., Jain, S., Sebbari, R., Zhang, X., Aldrin, M., Sarojini, B. B., Beer, J., Boucher, O., Braconnot, P., Browne, O., Chang, P., … Zhang, R. (2014). Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis*. https://hal.science/hal-03330685

Blöbaum, P., Götz, P., Budhathoki, K., Mastakouri, A. A., & Janzing, D. (2024). DoWhy-GCM: An extension of DoWhy for causal inference in graphical causal models. *Journal of Machine Learning Research*, *25*(147), 1–7.

Blüthgen, N., Staab, M., Achury, R., & Weisser, W. W. (2022). Unravelling insect declines: Can space replace time? *Biology Letters*, *18*(4), 20210666.

Boënnec, M., Dakos, V., & Devictor, V. (2024). Sources of confusion in global biodiversity trends. *Oikos*, *2024*(6), e10732.

Bongers, S., Forré, P., Peters, J., Schölkopf, B., & Mooij, J. M. (2016). Foundations of structural causal models with cycles and latent variables. *arXiv [stat.ME]*. arXiv. http://arxiv.org/abs/1611.06221

Borges, M. F., Brind'Amour, A., Cardador, F., Chaves, C., Ellis, J. R., Gaudêncio, M. J., Guerra, M., Hille Ris Lambers, R., Narberhaus, I., Rumohr, H., Sapp, M., Schratzberger, M., Velasco, F., & Wlodarska-Kowalczuk, M. (2024). *Practical issues affecting the utility of Field survey data for biodiversity monitoring*. International Council for the Exploration of the Sea, ICES CM 2010/Q:10:ICES CM 2010 Q:09.

Borsuk, M. E., Stow, C. A., & Reckhow, K. H. (2004). A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling*, *173*(2–3), 219–239.

Bowler, D. E., Bjorkman, A. D., Dornelas, M., Myers-Smith, I. H., Navarro, L. M., Niamir, A., Supp, S. R., Waldock, C., Winter, M., Vellend, M., Blowes, S. A., Böhning-Gaese, K., Bruelheide, H., Elahi, R., Antão, L.

H., Hines, J., Isbell, F., Jones, H. P., Magurran, A. E., … Bates, A. E. (2020). Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. *People and Nature*, *2*(2), 380–394.

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J. B., & Pocock, M. J. O. (2024). Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews of the Cambridge Philosophical Society*, *100*, 50–67. https://doi.org/10.1111/brv.13127

Boyd, R. J., Powney, G. D., Burns, F., Danet, A., Duchenne, F., Grainger, M. J., Jarvis, S. G., Martin, G., Nilsen, E. B., Porcher, E., Stewart, G. B., Wilson, O. J., & Pescott, O. L. (2022). ROBITT: A tool for assessing the risk-of-bias in studies of temporal trends in ecology. *Methods in Ecology and Evolution*, *13*(7), 1497–1507.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Breston, L., Leonardis, E. J., Quinn, L. K., Tolston, M., Wiles, J., & Chiba, A. A. (2021). Convergent cross sorting for estimating dynamic coupling. *Scientific Reports*, *11*(1), 20374.

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., Martius, C., Zlinszky, A., Calvignac-Spencer, S., Cobbold, C. A., Dawson, T. P., Emerson, B. C., Ferrier, S., Gilbert, M. T. P., Herold, M., Jones, L., Leendertz, F. H., Matthews, L., Millington, J. D. A., … Yu, D. W. (2017). Connecting earth observation to high-throughput biodiversity data. *Nature Ecology & Evolution*, *1*(7), 176.

Butsic, V., Lewis, D. J., Radeloff, V. C., Baumann, M., & Kuemmerle, T. (2017). Quasi-experimental methods enable stronger inferences from observational data in ecology. *Basic and Applied Ecology*, *19*(March), 1–10.

Byrnes, J. E. K., & Dee, L. E. (2024). Causal inference with observational data and unobserved confounding variables. *bioRxiv*. https://doi.org/10.1101/2024.02.26.582072

Caron, A., Baio, G., & Manolopoulou, I. (2022). Estimating individual treatment effects using non-parametric regression models: A review. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *185*(3), 1115–1149.

Chadwick, F. J., Haydon, D. T., Husmeier, D., Ovaskainen, O., & Matthiopoulos, J. (2024). LIES of omission: Complex observation processes in ecology. *Trends in Ecology & Evolution*, *39*(4), 368–380.

Chan, N. H., Yau, C. Y., & Zhang, R.-M. (2015). LASSO estimation of threshold autoregressive models. *Journal of Econometrics*, *189*(2), 285–296.

Chen, Y., Rangarajan, G., Feng, J., & Ding, M. (2004). Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A*, *324*(1), 26–35.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters: Double/debiased machine learning. *The Econometrics Journal*, *21*(1), C1–C68.

Chickering, D. M. (2003). Optimal structure identification with greedy search. *Journal of Machine Learning Research*, *3*(Nov), 507–554.

CIESIN. (2018). *Gridded population of the world, version 4 (GPWv4): Population density, revision 11*. NASA Socioeconomic Data and Applications Center (SEDAC). https://doi.org/10.7927/H49C6VHW

Cinelli, C., Forney, A., & Pearl, J. (2020). A crash course in good and bad controls. *SSRN Electronic Journal*, September. https://doi.org/10.2139/ssrn.3689437

Clark, T., Ye, H., Isbell, F., Deyle, E. R., Cowles, J., Tilman, G. D., & Sugihara, G. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology*, *96*(5), 1174–1181.

Copas, J., & Eguchi, S. (2020). Strong model dependence in statistical analysis: Goodness of fit is not enough for model choice. *Annals of the Institute of Statistical Mathematics*, *72*(2), 329–352.

Cornfield, J., Haenszel, W., Hammond, E. C., Lilienfeld, A. M., Shimkin, M. B., & Wynder, E. L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *Journal of the National Cancer Institute*, *22*(1), 173–203.

Courter, J. R., Johnson, R. J., Stuyck, C. M., Lang, B. A., & Kaiser, E. W. (2013). Weekend bias in citizen science data reporting: Implications for phenology studies. *International Journal of Biometeorology*, *57*(5), 715–720.

Currie, D. J. (2019). Where Newton might have taken ecology. *Global Ecology and Biogeography*, *28*(1), 18–27.

Daly, C. (2006). Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, *26*(6), 707–721.

d'Andrimont, R., Yordanov, M., Martinez-Sanchez, L., Eiselt, B., Palmieri, A., Dominici, P., Gallego, J., Reuter, H. I., Joebges, C., Lemoine, G., & van der Velde, M. (2020). Harmonised LUCAS in-situ land cover and use database for Field surveys from 2006 to 2018 in the European Union. *Scientific Data*, *7*(1), 352.

Darwiche, A. (2008). Chapter 11 Bayesian networks. In *Handbook of knowledge representation* (Vol. *3*, pp. 467–509). Foundations of Artificial Intelligence. Elsevier.

Daskalova, G. N., Myers-Smith, I. H., Bjorkman, A. D., Blowes, S. A., Supp, S. R., Magurran, A. E., & Dornelas, M. (2020). Landscape-scale Forest loss as a catalyst of population and biodiversity change. *Science (New York, N.Y.)*, *368*(6497), 1341–1347.

Daskalova, G. N., Phillimore, A. B., & Myers-Smith, I. H. (2021). Accounting for year effects and sampling error in temporal analyses of invertebrate population and biodiversity change: A comment on Seibold et al. 2019. *Insect Conservation and Diversity*, *14*(1), 149–154.

Davison, C. W., Rahbek, C., & Morueta-Holme, N. (2024). Changes in Danish bird communities over four decades of climate and land-use change. *Oikos*, *2024*, e10697. https://doi.org/10.1111/oik.10697

De Brouwer, E., Arany, A., Simm, J., & Moreau, Y. (2021). *Latent convergent cross mapping*. ICLR. https://openreview.net/pdf

De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., Ashcroft, M. B., Christiansen, D. M., Decocq, G., De Pauw, K., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klinges, D. H., Koelemeijer, I. A., Lembrechts, J. J., Marrec, R., … Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Global Change Biology*, *27*(11), 2279–2297.

De Palma, A., Sanchez-Ortiz, K., Martin, P. A., Chadwick, A., Gilbert, G., Bates, A. E., Börger, L., Contu, S., Hill, S. L. L., & Purvis, A. (2018). Challenges with inferring how land-use affects terrestrial biodiversity: Study design, time, space and synthesis. In *Next generation biomonitoring: Part 1*. Advances in ecological research (Vol. *58*, pp. 163–199). Elsevier.

Dee, L. E., Ferraro, P. J., Severen, C. N., Kimmel, K. A., Borer, E. T., Byrnes, J. E. K., Clark, A. T., Hautier, Y., Hector, A., Raynaud, X., Reich, P. B., Wright, A. J., Arnillas, C. A., Davies, K. F., MacDougall, A., Mori, A. S., Smith, M. D., Adler, P. B., Bakker, J. D., … Loreau, M. (2023). Clarifying the effect of biodiversity on productivity in natural ecosystems with longitudinal data and methods for causal inference. *Nature Communications*, *14*(1), 2607.

Detto, M., Molini, A., Katul, G., Stoy, P., Palmroth, S., & Baldocchi, D. (2012). Causality and persistence in ecological systems: A nonparametric spectral Granger causality approach. *The American Naturalist*, *179*(4), 524–535.

Deyle, E. R., May, R. M., Munch, S. B., & Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real time. *Proceedings. Biological sciences*, *283*(1822), 20152258.

Ding, Y., Zhou, J., Cui, Q., Wang, L., Zhang, M., & Dong, Y. (2023). DistriBayes: A distributed platform for learning, inference and attribution on large scale Bayesian network. In *Proceedings of the sixteenth ACM international conference on web search and data mining*. ACM. https://doi.org/10.1145/3539597.3573028

Dinku, T. (2019). Chapter 7 - challenges with availability and quality of climate data in Africa. In A. M. Melesse, W. Abtew, & G. Senay (Eds.), *Extreme hydrology and climate variability* (pp. 71–80). Elsevier.

Dise, N. B., Ashmore, M., Belyazid, S., Bleeker, A., Bobbink, R., de Vries, W., Erisman, J. W., Spranger, T., Stevens, C. J., & van den Berg, L.

(2011). Nitrogen as a threat to European terrestrial biodiversity. In M. A. Sutton, C. M. Howard, J. W. Erisman, G. Billen, A. Bleeker, P. Grennfelt, H. van Grinsven, & B. Grizzetti (Eds.), *The European nitrogen assessment* (pp. 463–494). Cambridge University Press.

Dobson, A. D. M., Milner-Gulland, E. J., Aebischer, N. J., Beale, C. M., Brozovic, R., Coals, P., Critchlow, R., Dancer, A., Greve, M., Hinsley, A., Ibbett, H., Johnston, A., Kuiper, T., le Comber, S., Mahood, S. P., Moore, J. F., Nilsen, E. B., Pocock, M. J. O., Quinn, A., … Keane, A. (2020). Making messy data work for conservation. *One Earth*, *2*(5), 455–465.

Dolby, G. A. (2021). Towards a unified framework to study causality in earth-life systems. *Molecular Ecology*, *30*(22), 5628–5642.

Dornelas, M., Antão, L. H., Moyes, F., Bates, A. E., Magurran, A. E., Adam, D., Akhmetzhanova, A. A., Appeltans, W., Arcos, J. M., Arnold, H., Ayyappan, N., Badihi, G., Baird, A. H., Barbosa, M., Barreto, T. E., Bässler, C., Bellgrove, A., Belmaker, J., Benedetti-Cecchi, L., … Hickler, T. (2018). BioTIME: A database of biodiversity time series for the Anthropocene. *Global Ecology and Biogeography: A Journal of Macroecology*, *27*(7), 760–786.

Dornelas, M., Chase, J. M., Gotelli, N. J., Magurran, A. E., McGill, B. J., Antão, L. H., Blowes, S. A., Daskalova, G. N., Leung, B., Martins, I. S., Moyes, F., Myers-Smith, I. H., Thomas, C. D., & Vellend, M. (2023). Looking Back on biodiversity change: Lessons for the road ahead. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *378*(1881), 20220199.

Dornelas, M., Magurran, A. E., Buckland, S. T., Chao, A., Chazdon, R. L., Colwell, R. K., Curtis, T., Gaston, K. J., Gotelli, N. J., Kosnik, M. A., McGill, B., McCune, J. L., Morlon, H., Mumby, P. J., Ovreås, L., Studeny, A., & Vellend, M. (2013). Quantifying temporal change in biodiversity: Challenges and opportunities. *Proceedings. Biological sciences*, *280*(1750), 20121931.

Duchenne, F., Porcher, E., Mihoub, J.-B., Loïs, G., & Fontaine, C. (2022). Controversy over the decline of arthropods: A matter of temporal baseline? *Peer Community Journal*, *2*, e33. https://doi.org/10.24072/pcjournal.131

Eichenberg, D., Bowler, D. E., Bonn, A., Bruelheide, H., Grescho, V., Harter, D., Jandt, U., May, R., Winter, M., & Jansen, F. (2020). Widespread decline in central European plant diversity across six decades. *Global Change Biology*, *27*(5), 1097–1110.

Emmons, A. L., Deel, H., & Davis, M. (2022). Soft tissue decomposition in terrestrial ecosystems. In *Manual of forensic* (pp. 41–78). CRC Press. https://doi.org/10.4324/9781003171492-03

Essl, F., Dullinger, S., Rabitsch, W., Hulme, P. E., Pyšek, P., Wilson, J. R., & Richardson, D. M. (2015). Delayed biodiversity change: No time to waste. *Trends in Ecology & Evolution*, *30*(7), 375–378, ISSN 0169-5347. https://doi.org/10.1016/j.tree.2015.05.002

Estes, L., Elsen, P. R., Treuer, T., Ahmed, L., Caylor, K., Chang, J., Choi, J. J., & Ellis, E. C. (2018). The spatial and temporal domains of modern ecology. *Nature Ecology & Evolution*, *2*(5), 819–826.

Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S. R., Park, H., & Shao, C. (2016). Applications of structural equation modeling (SEM) in ecological studies: An updated review. *Ecological Processes*, *5*(1), 1–12.

Farley, S. S., Dawson, A., Goring, S. J., & Williams, J. W. (2018). Situating ecology as a big-data science: Current advances, challenges, and solutions. *BioScience*, *68*(8), 563–576.

Ferraro, P. J., Sanchirico, J. N., & Smith, M. D. (2019). Causal inference in coupled human and natural systems. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(12), 5311–5318.

Fink, D., Johnston, A., Strimas-Mackey, M., Auer, T., Hochachka, W. M., Ligocki, S., Oldham Jaromczyk, L., Robinson, O., Wood, C., Kelling, S., & Rodewald, A. D. (2023). A double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, *14*(9), 2435–2448.

Fisher, A., Rudin, C., & Dominici, F. (2018). All models are wrong, but many are useful: Learning a Variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research: JMLR*, *20*(177), 1–81.

Friedman, J. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, *29*(5), 1189–1232.

Gao, B., Yang, J., Chen, Z., Sugihara, G., Li, M., Stein, A., Kwan, M.-P., & Wang, J. (2023). Causal inference from cross-sectional earth system data with geographical convergent cross mapping. *Nature Communications*, *14*(1), 5875.

Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of causal discovery methods based on graphical models. *Frontiers in Genetics*, *10*(June), 524.

Goldberg, L. R. (2019). *The book of why: The new science of cause and effect: By Judea Pearl and Dana Mackenzie*. Basic Books (2018). isbn: 978-0465097609.

Gonzalez, A., Chase, J. M., & O'Connor, M. I. (2023). A framework for the detection and attribution of biodiversity change. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *378*(1881), 20220182.

Grace, J. B. (2021). Instrumental variable methods in structural equation models. *Methods in Ecology and Evolution*, *12*, 1148–1157.

Grace, J. B. (2024). An integrative paradigm for building causal knowledge. *Ecological Monographs*, *94*(4), e1628. https://doi.org/10.1002/ecm.1628

Grace, J. B., Anderson, T. M., Olff, H., & Scheiner, S. M. (2010). On the specification of structural equation models for ecological systems. *Ecological Monographs*, *80*(1), 67–87.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, *37*(3), 424.

Guidotti, R. (2024). Counterfactual explanations and how to find them: Literature review and benchmarking. *Data Mining and Knowledge Discovery*, *38*(5), 2770–2824.

Guzman, L. M., Elle, E., Morandin, L. A., Cobb, N. S., Chesshire, P. R., McCabe, L. M., Hughes, A., Orr, M., & M'Gonigle, L. K. (2024). Impact of pesticide use on wild bee distributions across the United States. *Nature Sustainability*, *7*, 1324–1334.

Haase, P., Bowler, D. E., Baker, N. J., Bonada, N., Domisch, S., Garcia Marquez, J. R., Heino, J., Hering, D., Jähnig, S. C., Schmidt-Kloiber, A., Stubbington, R., Altermatt, F., Álvarez-Cabria, M., Amatulli, G., Angeler, D. G., Archambaud-Suard, G., Jorrín, I. A., Aspin, T., Azpiroz, I., … Welti, E. A. R. (2023). The recovery of European freshwater biodiversity has come to a halt. *Nature*, *620*(7974), 582–588.

Hahn, P. R., Murray, J. S., & Carvalho, C. M. (2020). Bayesian regression tree models for causal inference: Regularization, confounding, and heterogeneous effects (with discussion). *Bayesian Analysis*, *15*(3), 965–1056.

Hannart, A., Pearl, J., Otto, F. E. L., Naveau, P., & Ghil, M. (2016). Causal counterfactual theory for the attribution of weather and climate-related events. *Bulletin of the American Meteorological Society*, *97*(1), 99–110.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-resolution global maps of 21st-century Forest cover change. *Science (New York, N.Y.)*, *342*(6160), 850–853.

Harte, J., Saleska, S. R., & Levy, C. (2015). Convergent ecosystem responses to 23-year ambient and manipulated warming link advancing snowmelt and shrub encroachment to transient and long-term climate-soil carbon feedback. *Global Change Biology*, *21*(6), 2349–2356.

Hartford, J., Lewis, G., Leyton-Brown, K., & Taddy, M. (2017). Deep IV: A flexible approach for counterfactual prediction. In *International conference on machine learning* (pp. 1414–1423). PMLR.

Hegerl, G., & Zwiers, F. (2011). Use of models in detection and attribution of climate change. *WIREs Climate Change*, *2*(4), 570–591.

Heinze-Deml, C., Peters, J., & Meinshausen, N. (2018). Invariant causal prediction for nonlinear models. *Journal of Causal Inference*, *6*(2), 20170016. https://doi.org/10.1515/jci-2017-0016

Henriques, S., Böhm, M., Collen, B., Luedtke, J., Hoffmann, M., Hilton-Taylor, C., Cardoso, P., Butchart, S. H. M., & Freeman, R. (2020). Accelerating the monitoring of global biodiversity: Revisiting the sampled approach to generating red list indices. *Conservation Letters*, *13*(3), e12703.

Hill, M. O. (2012). Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution*, *3*(1), 195–205.

Hone, J., & Krebs, C. J. (2023). Causality and wildlife management. *The Journal of Wildlife Management*, *87*(5), e22412.

Hoyer, P., Janzing, D., Mooij, J. M., Peters, J., & Schölkopf, B. (2008). Nonlinear causal discovery with additive noise models. *Advances in Neural Information Processing Systems*, *21*, 689–696. https://proceedings.neurips.cc/paper_files/paper/2008/file/f7664060cc52bc6f3d620bcedc94a4b6-Paper.pdf

Hudson, L. N., Newbold, T., Contu, S., Hill, S. L. L., Lysenko, I., De Palma, A., Phillips, H. R. P., Senior, R. A., Bennett, D. J., Booth, H., Choimes, A., Correia, D. L. P., Day, J., Echeverría-Londono, S., Garon, M., Harrison, M. L. K., Ingram, D. J., Jung, M., Kemp, V., … Purvis, A. (2014). The PREDICTS database: A global database of how local terrestrial biodiversity responds to human impacts. *Ecology and Evolution*, *4*(24), 4701–4735.

IPBES. (2016). *The methodological assessment report on scenarios and models of biodiversity and ecosystem services*. V1. IPBES. https://doi.org/10.5281/zenodo.3235429

Isaac, N. J. B., & Pocock, M. J. O. (2015). Bias and information in biological records. *Biological Journal of the Linnean Society*, *115*(3), 522–531.

Isaac, N. J. B., van Strien, A. J., August, T. A., de Zeeuw, M. P., & Roy, D. B. (2014). Statistics for citizen science: Extracting signals of change from Noisy ecological data. *Methods in Ecology and Evolution*, *5*(10), 1052–1060.

Johnson, T. F., Beckerman, A. P., Childs, D. Z., Webb, T. J., Evans, K. L., Griffiths, C. A., Capdevila, P., Clements, C. F., Besson, M., Gregory, R. D., Thomas, G. H., Delmas, E., & Freckleton, R. P. (2024). Revealing uncertainty in the status of biodiversity change. *Nature*, *628*(8009), 788–794.

Johnson, T. F., Isaac, N. J. B., Paviolo, A., & González-Suárez, M. (2021). Handling Missing Values in Trait Data. *Global Ecology and Biogeography: A Journal of Macroecology*, *30*(1), 51–62.

Jones, K. W., & Lewis, D. J. (2015). Estimating the counterfactual impact of conservation programs on land cover outcomes: The role of matching and panel regression techniques. *PLoS One*, *10*(10), e0141380.

Joppa, L. N., O'Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., Watson, J. E. M., Butchart, S. H. M., Virah-Sawmy, M., Halpern, B. S., Ahmed, S. E., Balmford, A., Sutherland, W. J., Harfoot, M., Hilton-Taylor, C., Foden, W., Minin, E. D., Pagad, S., Genovesi, P., … Burgess, N. D. (2016). BIG DATA AND BIODIVERSITY. Filling in biodiversity threat gaps. *Science*, *352*(6284), 416–418.

Kaddour, J., Lynch, A., Liu, Q., Kusner, M. J., & Silva, R. (2022). Causal machine learning: A survey and open problems. *arXiv [cs.LG]*. arXiv. http://arxiv.org/abs/2206.15475

Kapfer, J., Hédl, R., Jurasinski, G., Kopecký, M., Schei, F. H., & Grytnes, J.-A. (2017). Resurveying historical vegetation data—Opportunities and challenges. *Applied Vegetation Science*, *20*(2), 164–171.

Karger, D. N., Schmatz, D. R., Dettling, G., & Zimmermann, N. E. (2020). High-resolution monthly precipitation and temperature time series from 2006 to 2100. *Scientific Data*, *7*(1), 248.

Kaspar, F., Andersson, A., Ziese, M., & Hollmann, R. (2022). Contributions to the improvement of climate data availability and quality for sub-Saharan Africa. *Frontiers in Climate*, *3*, 815043.

Keele, L. J. (2022). *Package 'rbounds'. R*. Springer. https://doi.org/10.1007/978-1-4757-3692-2

Kelling, S., Johnston, A., Bonn, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., Fernandez, M., Hochachka, W. M., Julliard, R., Kraemer, R., & Guralnick, R. (2019). Using Semistructured surveys to improve citizen science data for monitoring biodiversity. *BioScience*, *69*(3), 170–179.

Kendall, B. E. (2015). A statistical symphony: Instrumental variables reveal causality and control measurement error. In *Ecological statistics* (pp. 149–167). Oxford University Press.

Kéry, M., & Royle, J. A. (2020). *Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 2: Dynamic and advanced models*. Academic Press.

Kimmel, K., Dee, L. E., Avolio, M. L., & Ferraro, P. J. (2021). Causal assumptions and causal inference in ecological experiments. *Trends in Ecology & Evolution*, *36*(12), 1141–1152.

Kindsvater, H. K., Dulvy, N. K., Horswill, C., Juan-Jordá, M.-J., Mangel, M., & Matthiopoulos, J. (2018). Overcoming the data crisis in biodiversity conservation. *Trends in Ecology & Evolution*, *33*(9), 676–688.

Kroodsma, D. A., Mayorga, J., Hochberg, T., Miller, N. A., Boerder, K., Ferretti, F., Wilson, A., Bergman, B., White, T. D., Block, B. A., Woods, P., Sullivan, B., Costello, C., & Worm, B. (2018). Tracking the global footprint of fisheries. *Science (New York, N.Y.)*, *359*(6378), 904–908.

Kuemmerle, T., Erb, K., Meyfroidt, P., Müller, D., Verburg, P. H., Estel, S., Haberl, H., Hostert, P., Jepsen, M. R., Kastner, T., Levers, C., Lindner, M., Plutzar, C., Verkerk, P. J., van der Zanden, E. H., & Reenberg, A. (2013). Challenges and opportunities in mapping land use intensity globally. *Current Opinion in Environmental Sustainability*, *5*(5), 484–493.

Kuhn-Régnier, A., Voulgarakis, A., Nowack, P., Forkel, M., Prentice, I. C., & Harrison, S. P. (2021). The importance of antecedent vegetation and drought conditions as global drivers of burnt area. *Biogeosciences*, *18*(12), 3861–3879.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., & Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(10), 4156–4165.

Larsen, A. E., Meng, K., & Kendall, B. E. (2019). Causal analysis in control–impact ecological studies with observational data. *Methods in Ecology and Evolution*, *10*(7), 924–934.

Larsen, A. E., & Noack, F. (2017). Identifying the landscape drivers of agricultural insecticide use leveraging evidence from 100,000 fields. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(21), 5473–5478.

Laubach, Z. M., Murray, E. J., Hoke, K. L., Safran, R. J., & Perng, W. (2021). A Biologist's guide to model selection and causal inference. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1943), 20202815.

Lausch, A., Baade, J., Bannehr, L., Borg, E., Bumberger, J., Chabrilliat, S., Dietrich, P., Gerighausen, H., Glässer, C., Hacker, J., Haase, D., Jagdhuber, T., Jany, S., Jung, A., Karnieli, A., Kraemer, R., Makki, M., Mielke, C., Möller, M., … Schaepman, M. (2019). Linking remote sensing and geodiversity and their traits relevant to biodiversity—Part I: Soil characteristics. *Remote Sensing*, *11*(20), 2356.

Lausch, A., Schaepman, M. E., Skidmore, A. K., Truckenbrodt, S. C., Hacker, J. M., Baade, J., Bannehr, L., Borg, E., Bumberger, J., Dietrich, P., Gläßer, C., Haase, D., Heurich, M., Jagdhuber, T., Jany, S., Krönert, R., Möller, M., Mollenhauer, H., Montzka, C., … Thiel, C. (2020). Linking the remote sensing of geodiversity and traits relevant to biodiversity—Part II: Geomorphology, terrain and surfaces. *Remote Sensing*, *12*(22), 3690.

Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*(5), 573–579.

Lembrechts, J. J., Aalto, J., Ashcroft, M. B., De Frenne, P., Kopecký, M., Lenoir, J., Luoto, M., Maclean, I. M. D., Roupsard, O., Fuentes-Lillo, E., García, R. A., Pellissier, L., Pitteloud, C., Alatalo, J. M., Smith, S. W., Björk, R. G., Muffler, L., Backes, A. R., Cesarz, S., … Nijs, I. (2020). SoilTemp: A global database of near-surface temperature. *Global Change Biology*, *26*(11), 6616–6629.

Li, H., Gao, W., Liu, Y., Yuan, F., Minchao, W., & Meng, L. (2023). Attributing the impacts of ecological engineering and climate change on carbon uptake in northeastern China. *Landscape Ecology*, *38*(12), 3945–3960.

Loh, J., Green, R. E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V., & Randers, J. (2005). The living planet index: Using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*(1454), 289–295.

Lumbierres, M., Dahal, P. R., Di Marco, M., Butchart, S. H. M., Donald, P. F., & Rondinini, C. (2022). Translating habitat class to land cover to map area of habitat of terrestrial vertebrates. *Conservation Biology*, *36*(3), e13851.

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv [cs.AI]*. arXiv. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf

Ma, H., Aihara, K., & Chen, L. (2014). Detecting causality from nonlinear dynamics with short-term time series. *Scientific Reports*, *4*(1), 7464.

MacDonald, A. J., Larsen, A. E., & Plantinga, A. J. (2019). Missing the people for the trees: Identifying coupled natural–human system feedbacks driving the ecology of Lyme disease. *The Journal of Applied Ecology*, *56*(2), 354–364.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L., & Hines, J. E. (2017). *Occupancy estimation and modeling: Inferring patterns and dynamics of species occurrence* (2nd ed.). Academic Press.

Martins, I. S., Schrodt, F., Blowes, S. A., Bates, A. E., Bjorkman, A. D., Brambilla, V., Carvajal-Quintero, J., Chow, C. F. Y., Daskalova, G. N., Edwards, K., Eisenhauer, N., Field, R., Fontrodona-Eslava, A., Henn, J. J., van Klink, R., Madin, J. S., Magurran, A. E., McWilliam, M., Moyes, F., … Dornelas, M. (2023). Widespread shifts in body size within populations and assemblages. *Science (New York, N.Y.)*, *381*(6662), 1067–1071.

Mazor, T., Doropoulos, C., Schwarzmueller, F., Gladish, D. W., Kumaran, N., Merkel, K., Di Marco, M., & Gagic, V. (2018). Global mismatch of policy and research on drivers of biodiversity loss. *Nature Ecology & Evolution*, *2*(7), 1071–1074.

McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, *32*(19), 3388–3414.

McCracken, J. M., & Weigel, R. S. (2014). Convergent cross-mapping and pairwise asymmetric inference. *Physical Review E*, *90*(6), 062903.

McGill, B. J. (2019). The what, how and why of doing macroecology. *Global Ecology and Biogeography: A Journal of Macroecology*, *28*(1), 6–17.

McRae, L., Deinet, S., & Freeman, R. (2017). The diversity-weighted living planet index: Controlling for taxonomic bias in a global biodiversity indicator. *PLoS One*, *12*(1), e0169156.

Metcalf, O. C., Barlow, J., Devenish, C., Marsden, S., Berenguer, E., & Lees, A. C. (2021). Acoustic indices perform better when applied at ecologically meaningful time and frequency scales. *Methods in Ecology and Evolution*, *12*(3), 421–431.

Mihoub, J.-B., Henle, K., Titeux, N., Brotons, L., Brummitt, N. A., & Schmeller, D. S. (2017). Setting temporal baselines for biodiversity: The limits of available monitoring data for capturing the full impact of anthropogenic pressures. *Scientific Reports*, *7*(January), 41591.

Moraffah, R., Karami, M., Guo, R., Raglin, A., & Liu, H. (2020). Causal interpretability for machine learning - problems, methods and evaluation. *SIGKDD Explorations: Newsletter of the Special Interest Group (SIG) on Knowledge Discovery & Data Mining*, *22*(1), 18–33.

Morueta-Holme, N., Iversen, L. L., Corcoran, D., Rahbek, C., & Normand, S. (2024). Unlocking ground-based imagery for habitat mapping. *Trends in Ecology & Evolution*, *39*(4), 349–358.

Moudrý, V., Bazzichetto, M., Remelgado, R., Devillers, R., Lenoir, J., Mateo, R. G., Lembrechts, J. J., Sillero, N., Lecours, V., Cord, A. F., Barták, V., Balej, P., Rocchini, D., Torresani, M., Arenas-Castro, S., Man, M., Prajzlerová, D., Gdulová, K., Prošek, J., … Šímová, P. (2024). Optimising occurrence data in species distribution models: Sample size, positional uncertainty, and sampling bias matter. *Ecography*, *2024*, e07294.

Myers-Smith, I. H., Kerby, J. T., Phoenix, G. K., Bjerke, J. W., Epstein, H. E., Assmann, J. J., John, C., Andreu-Hayles, L., Angers-Blondin, S., Beck, P. S. A., Berner, L. T., Bhatt, U. S., Bjorkman, A. D., Blok, D., Bryn, A., Christiansen, C. T., Cornelissen, J. H. C., Cunliffe, A. M., Elmendorf, S. C., … Wipf, S. (2020). Complexity revealed in the greening of the Arctic. *Nature Climate Change*, *10*, 106–117. https://doi.org/10.1038/s41558-019-0688-1

Nakagawa, S., & Freckleton, R. P. (2011). Model averaging, missing data and multiple imputation: A case study for Behavioural ecology. *Behavioral Ecology and Sociobiology*, *65*(1), 103–116.

Nichols, J. D., & Cooch, E. G. (2025). Predictive models are indeed useful for causal inference. *Ecology*, *106*(1), e4517.

Nicholson, E., Fulton, E. A., Brooks, T. M., Blanchard, R., Leadley, P., Metzger, J. P., Mokany, K., Stevenson, S., Wintle, B. A., Woolley, S. N. C., Barnes, M., Watson, J. E. M., & Ferrier, S. (2019). Scenarios and models to support global conservation targets. *Trends in Ecology & Evolution*, *34*(1), 57–68.

Nilsen, E. B., Bowler, D. E., & Linnell, J. D. C. (2020). Exploratory and confirmatory research in the Open Science era. *The Journal of Applied Ecology*, *57*(4), 842–847.

Nogueira, A. R., Pugnana, A., Ruggieri, S., Pedreschi, D., & Gama, J. (2022). Methods and tools for causal discovery and causal inference. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *12*(2), e1449.

Oedekoven, C. S., Elston, D. A., Harrison, P. J., Brewer, M. J., Buckland, S. T., Johnston, A., Foster, S., & Pearce-Higgins, J. W. (2017). Attributing changes in the distribution of species abundance to weather variables using the example of British breeding birds. *Methods in Ecology and Evolution*, *8*(12), 1690–1702.

Ogarrio, J. M., Spirtes, P., & Ramsey, J. (2016). A hybrid causal search algorithm for latent variable models. In *Conference on probabilistic graphical models* (pp. 368–379). PMLR.

Outhwaite, C. L., Gregory, R. D., Chandler, R. E., Collen, B., & Isaac, N. J. B. (2020). Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature Ecology & Evolution*, *4*(3), 384–392.

Pamfil, R., Sriwattanaworachai, N., Desai, S., Pilgerstorfer, P., Georgatzis, K., Beaumont, P., & Aragam, B. (2020). DYNOTEARS: Structure learning from time-series data. In *International conference on artificial intelligence and statistics* (pp. 1595–1605). PMLR.

Papagiannopoulou, C., Miralles, D. G., Decubber, S., Demuzere, M., Verhoest, N. E. C., Dorigo, W. A., & Waegeman, W. (2017). A nonlinear Granger-causality framework to investigate climate–vegetation dynamics. *Geoscientific Model Development*, *10*(5), 1945–1960.

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669–688.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys*, *3*, 96–146.

Pélissié, M., Devictor, V., & Dakos, V. (2024). A systematic approach for detecting abrupt shifts in ecological Timeseries. *Biological Conservation*, *290*, 110429.

Pennock, D. (2015). *Status of the World's soil resources: Technical summary*. FAO.

Pescott, O. L., Stroh, P. A., Humphrey, T. A., & Walker, K. J. (2022). Simple methods for improving the communication of uncertainty in species' temporal trends. *Ecological Indicators*, *141*, 109117.

Peters, J., Bühlmann, P., & Meinshausen, N. (2016). Causal inference by using invariant prediction: Identification and confidence intervals. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, *78*(5), 947–1012.

Pettorelli, N., Laurance, W. F., O'Brien, T. G., Wegmann, M., Nagendra, H., & Turner, W. (2014). Satellite remote sensing for applied ecologists: Opportunities and challenges. *The Journal of Applied Ecology*, *51*(4), 839–848.

Pfister, N., Bühlmann, P., & Peters, J. (2019). Invariant causal prediction for sequential data. *Journal of the American Statistical Association*, *114*(527), 1264–1276.

Poggio, L., de Sousa, L. M., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: Producing soil information for the globe with quantified spatial uncertainty. *The Soil*, *7*(1), 217–240.

Pretzsch, H. (2022). The emergent past: Past natural and human disturbances of trees can reduce their present resistance to drought stress. *European Journal of Forest Research*, *141*(1), 87–104.

Ramsey, D. S. L., Forsyth, D. M., Wright, E., McKay, M., & Westbrooke, I. (2019). Using propensity scores for causal inference in ecology: Options, considerations, and a case study. *Methods in Ecology and Evolution*, *10*(3), 320–331.

Ramsey, J. D. (2015). Scaling up greedy causal search for continuous variables. *arXiv [cs.AI]*. arXiv. http://arxiv.org/abs/1507.07749

Rapacciuolo, G., Young, A., & Johnson, R. (2021). Deriving indicators of biodiversity change from unstructured community-contributed data. *Oikos*, *130*(8), 1225–1239.

Ratcliffe, H., Kendig, A., Vacek, S., Carlson, D., Ahlering, M., & Dee, L. E. (2024). Extreme precipitation promotes invasion in managed grasslands. *Ecology*, *105*(1), e4190.

Redhead, J. W., Hinsley, S. A., Botham, M. S., Broughton, R. K., Freeman, S. N., Bellamy, P. E., Siriwardena, G., Randle, Z., Nowakowski, M., Heard, M. S., & Pywell, R. F. (2022). The effects of a decade of Agri-environment intervention in a lowland farm landscape on population trends of birds and butterflies. *The Journal of Applied Ecology*, *59*(10), 2486–2496.

Reich, B. J., Yang, S., Guan, Y., Giffin, A. B., Miller, M. J., & Rappold, A. (2021). A review of spatial causal inference methods for environmental and epidemiological applications. *International Statistical Review = Revue Internationale de Statistique*, *89*(3), 605–634.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why should I trust you?: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. ACM. https://doi.org/10.1145/2939672.2939778

Rigal, S., Dakos, V., Alonso, H., Auniņš, A., Benkő, Z., Brotons, L., Chodkiewicz, T., Chylarecki, P., de Carli, E., del Moral, J. C., Domşa, C., Escandell, V., Fontaine, B., Foppen, R., Gregory, R., Harris, S., Herrando, S., Husby, M., Ieronymidou, C., … Devictor, V. (2023). Farmland practices are driving bird population decline across Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(21), e2216573120.

Rigal, S., Devictor, V., & Dakos, V. (2020). A method for classifying and comparing non-linear trajectories of ecological variables. *Ecological Indicators*, *112*(106113), 106113.

Rinella, M. J., Strong, D. J., & Vermeire, L. T. (2020). Omitted variable bias in studies of plant interactions. *Ecology*, *101*(6), e03020.

Rosenbaum, P. R. (2010). *Observational studies*. Springer series in statistics. Springer.

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., … Zscheischler, J. (2019). Inferring causation from time series in earth system sciences. *Nature Communications*, *10*(1), 2553.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., & Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, *5*(11), eaau4996.

Saavedra, S., Bartomeus, I., Godoy, O., Rohr, R. P., & Zu, P. (2022). Towards a system-level causative knowledge of pollinator communities. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *377*(1853), 20210159.

Sandel, B., Gutiérrez, A. G., Reich, P. B., Schrodt, F., Dickie, J., & Kattge, J. (2015). Estimating Themissing species bias in plant trait measurements. *Journal of Vegetation Science*, *26*(5), 828–838.

Santini, L., Belmaker, J., Costello, M. J., Pereira, H. M., Rossberg, A. G., Schipper, A. M., Ceaușu, S., Dornelas, M., Hilbers, J. P., Hortal, J., Huijbregts, M. A. J., Navarro, L. M., Schiffers, K. H., Visconti, P., & Rondinini, C. (2017). Assessing the suitability of diversity metrics to detect biodiversity change. *Biological Conservation*, *213*(September), 341–350.

Schleicher, J., Eklund, J., Barnes, M. D., Geldmann, J., Oldekop, J. A., & Jones, J. P. G. (2020). Statistical matching for conservation science. *Conservation Biology*, *34*(3), 538–549.

Schrodt, F., Kattge, J., Shan, H., Fazayeli, F., Joswig, J., Banerjee, A., Reichstein, M., Bönisch, G., Díaz, S., Dickie, J., Gillison, A., Karpatne, A., Lavorel, S., Leadley, P., Wirth, C. B., Wright, I. J., Wright, S. J., & Reich, P. B. (2015). BHPMF—A hierarchical Bayesian approach to gap-filling and trait prediction for macroecology and functional biogeography: Gap-filling in trait databases. *Global Ecology and Biogeography*, *24*(12), 1510–1521.

Schrodt, F., Vernham, G., Bailey, J., Field, R., Gordon, J. E., Gray, M., Hjort, J., Hoorn, C., Hunter, M. L., Jr., Larwood, J., Lausch, A., Monge-Ganuzas, M., Miller, S., van Ree, D., Seijmonsbergen, A. C., Zarnetske, P. L., & Daniel Kissling, W. (2024). The status and future of essential geodiversity variables. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *382*(2269), 20230052.

Seaton, F. M., Jarvis, S. G., & Henrys, P. A. (2024). Spatio-temporal data integration for species distribution modelling in R-INLA. *Methods in Ecology and Evolution*, *15*(7), 1221–1232.

Shadaydeh, M., Denzler, J., Guanche García, Y., & Mahecha, M. (2019). Time-frequency causal inference uncovers anomalous events in environmental systems. In *Lecture notes in computer science*. Lecture Notes in Computer Science (pp. 499–512). Springer International Publishing.

Shimizu, S., Hoyer, P., Hyvärinen, A., & Kerminen, A. J. (2006). A linear non-Gaussian acyclic model for causal discovery. *Journal of Machine Learning Research: JMLR*, *7*, 2003–2030.

Shipley, B. (2016). *Cause and correlation in biology: A user's guide to path analysis, structural equations and causal inference with R*. Cambridge University Press.

Sica, Y. V., Ingenloff, K., Gan, Y.-M., Kachian, Z., Baskauf, S. J., Wieczorek, J., Zermoglio, P. F., & Stevenson, R. D. (2022). Application of Humboldt extension to real-world cases. *Biodiversity Information Science and Standards*, *6*, e91502.

Siegel, K., & Dee, L. E. (2025). Foundations and future directions for causal inference in ecological research. *Ecology Letters*, *28*(1), e70053.

Simmonds, E. G., Adjei, K. P., Andersen, C. W., Aspheim, J. C. H., Battistin, C., Bulso, N., Christensen, H. M., Cretois, B., Cubero, R., Davidovich, I. A., Dickel, L., Dunn, B., Dunn-Sigouin, E., Dyrstad, K., Einum, S., Giglio, D., Gjerløw, H., Godefroidt, A., González-Gil, R., … O'Hara, R. B. (2022). Insights into the quantification and reporting of model-related uncertainty across different disciplines. *IScience*, *25*(12), 105512.

Soga, M., & Gaston, K. J. (2018). Shifting baseline syndrome: Causes, consequences, and implications. *Frontiers in Ecology and the Environment*, *16*(4), 1540–9295. https://doi.org/10.1002/fee.1794

Somerton, D. A., Otto, R. S., & Syrjala, S. E. (2002). Can changes in tow duration on bottom trawl surveys Lead to changes in CPUE and mean size? *Fisheries Research*, *55*(1), 63–70.

Song, C., Simmons, B. I., Fortin, M.-J., & Gonzalez, A. (2022). Generalism drives abundance: A computational causal discovery approach. *PLoS Computational Biology*, *18*(9), e1010302.

Spake, R., Barajas-Barbosa, M. P., Blowes, S. A., Bowler, D. E., Callaghan, C. T., Garbowski, M., Jurburg, S. D., van Klink, R., Korell, L., Ladouceur, E., Rozzi, R., Viana, D. S., Xu, W.-B., & Chase, J. M. (2022). Detecting thresholds of ecological change in the Anthropocene. *Annual Review of Environment and Resources*, *47*, 797–821.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. 1993rd ed. Lecture notes in statistics 81. MIT Press.

Spirtes, P., & Zhang, K. (2016). Causal discovery and inference: Concepts and recent methodological advances. *Applied Informatics*, *3*(1), 3.

Stewart, P. S., Stephens, P. A., Hill, R. A., Whittingham, M. J., & Dawson, W. (2023). Model selection in occupancy models: Inference versus prediction. *Ecology*, *104*(3), e3942.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1–21.

Sugihara, G. (1994). Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions. Physical Sciences and Engineering*, *348*(1688), 477–495.

Sugihara, G., May, R., Ye, H., Hsieh, C.-H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science (New York, N.Y.)*, *338*(6106), 496–500.

Takahata, K., Suetsugu, H., Fukaya, K., & Shirota, S. (2024). Bayesian state-space synthetic control method for deforestation baseline estimation for Forest carbon credits. *Environmental Data Science*, *3*(January), e6.

Tamer, E. (2010). Partial Identification in Econometrics. *Annual Review of Economics*, *2*(1), 167–195.

Ten Caten, C., & Dallas, T. (2023). Thinning occurrence points does not improve species distribution model performance. *Ecosphere (Washington, DC)*, *14*(12), e4703.

Textor, J., van der Zander, B., Gilthorpe, M. S., Liskiewicz, M., & Ellison, G. T. (2016). Robust causal inference using directed acyclic graphs: The R package 'Dagitty'. *International Journal of Epidemiology*, *45*(6), 1887–1894.

Thuiller, W., Brotons, L., Araújo, M. B., & Lavorel, S. (2004). Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, *27*(2), 165–172.

Tóth, G., Hermann, T., Szatmári, G., & Pásztor, L. (2016). Maps of heavy metals in the soils of the European Union and proposed priority areas for detailed assessment. *The Science of the Total Environment*, *565*, 1054–1062.

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z., Buchanan, G., Dech, S., Dwyer, J., Herold, M., Koh, L. P., Leimgruber, P., Taubenboeck, H., Wegmann, M., Wikelski, M., & Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation*, *182*, 173–176.

Van Cleemput, E., Adler, P. B., Suding, K. N., Rebelo, A. J., Poulter, B., & Dee, L. E. (2025). Scaling-up ecological understanding with remote sensing and causal inference. *Trends in Ecology & Evolution*, *40*(2), 122–135.

VanderWeele, T. J., & Ding, P. (2017). Sensitivity analysis in observational research: Introducing the E-value. *Annals of Internal Medicine*, *167*(4), 268–274.

Vernham, G., Bailey, J. J., Chase, J. M., Hjort, J., Field, R., & Schrodt, F. (2023). Understanding trait diversity: The role of geodiversity. *Trends in Ecology & Evolution*, *38*(8), 736–748.

Viana, D. S., Keil, P., & Jeliazkov, A. (2022). Disentangling spatial and environmental effects: Flexible methods for community ecology and macroecology. *Ecosphere*, *13*(4), e4028.

Wager, S., & Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, *113*(523), 1228–1242.

Wang, Y., Luo, Y., Wang, M., & Miao, H. (2018). Time-invariant biological networks with feedback loops: Structural equation models and structural identifiability. *IET Systems Biology*, *12*(6), 264–272.

Wauchope, H. S., Jones, J. P. G., Geldmann, J., Simmons, B. I., Amano, T., Blanco, D. E., Fuller, R. A., Johnston, A., Langendoen, T., Mundkur, T., Nagy, S., & Sutherland, W. J. (2022). Protected areas have a mixed impact on Waterbirds, but management helps. *Nature*, *605*(7908), 103–107.

West, T. A. P., Wunder, S., Sills, E. O., Börner, J., Rifai, S. W., Neidermeier, A. N., Frey, G. P., & Kontoleon, A. (2023). Action needed to make carbon offsets from forest conservation work for climate change mitigation. *Science*, *381*(6660), 873–877.

Westreich, D., & Greenland, S. (2013). The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American Journal of Epidemiology*, *177*(4), 292–298. https://doi.org/10.1093/aje/kws412

Wiegand, T., & Moloney, K. A. (2020). *Handbook of spatial point-pattern analysis in ecology*. 1st edition. Chapman & Hall/CRC applied environmental statistics. Chapman & Hall/CRC. https://doi.org/10.1201/b16195

Wilkinson, J. D., Mamas, M. A., & Kontopantelis, E. (2022). Logistic regression frequently outperformed propensity score methods, especially for large datasets: A simulation study. *Journal of Clinical Epidemiology*, *152*, 176–184.

Wismüller, A., Dsouza, A. M., Vosoughi, M. A., & Abidin, A. (2021). Large-scale nonlinear Granger causality for inferring directed dependence from short multivariate time-series data. *Scientific Reports*, *11*(1), 7817.

Witte, J., & Didelez, V. (2019). Covariate selection strategies for causal inference: Classification and comparison. *Biometrical Journal*, *61*(5), 1270–1289.

Wood, S. N. (2020). Inference and computation with generalized additive models and their extensions. *TEST*, *29*(2), 307–339.

Wuepper, D., & Finger, R. (2022). Regression discontinuity designs in agricultural and environmental economics. *European Review of Agricultural Economics*, *50*(1), 1–28.

Wüest, R. O., Zimmermann, N. E., Zurell, D., Alexander, J. M., Fritz, S. A., Hof, C., Kreft, H., Normand, S., Cabral, J. S., Szekely, E., Thuiller, W., Wikelski, M., & Karger, D. N. (2020). Macroecology in the age of big data – Where to go from here? *Journal of Biogeography*, *47*(1), 1–12.

Ying, X., Leng, S.-Y., Ma, H.-F., Nie, Q., Lai, Y.-C., & Lin, W. (2022). Continuity scaling: A rigorous framework for detecting and quantifying causality accurately. *Research (Washington, D.C.)*, *2022*, 9870149.

Zhao, S., van Dyk, D. A., & Imai, K. (2020). Propensity score-based methods for causal inference in observational studies with non-binary treatments. *Statistical Methods in Medical Research*, *29*(3), 709–727.

Zheng, X., Aragam, B., Ravikumar, P. K., & Xing, E. P. (2018). DAGs with NO TEARS: Continuous optimization for structure learning. *Advances in Neural Information Processing Systems*, *31*, 9492–9503. https://proceedings.neurips.cc/paper_files/paper/2018/file/e347c51419ffb23ca3fd5050202f9c3d-Paper.pdf

Zhu, S., Ng, I., & Chen, Z. (2019). Causal discovery with reinforcement learning. *arXiv [cs.LG]*. arXiv. http://arxiv.org/abs/1906.04477

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Figure S1. Example illustrating the scale dependence of the direction of causality considering covariates (grey) interacting with a driver (green, the target of the causal effect estimation) and/or biodiversity response (blue).

Figure S2. When building causal models, cycles can prevent the identification of a DAG.

Figure S3. Illustration of the realisation of the attribution workflow steps identified in Figure 1 by example studies in the field of biodiversity/ecology.

Table S1. Non-exhaustive overview of programming resources and technical and non-technical entry points to the attribution methods discussed in Section 3.

Table S2. Semi-quantitative data informing diagrams in Figure 1.

How to cite this article: Schrodt, F., Beck, M., Estopinan, J., Bowler, D. E., Fontaine, C., Gaüzère, P., Goury, R., Grenié, M., Martins, I. S., Morueta-Holme, N., Santini, L., Hedde, M., Martin, G., Porcher, E., Si-Moussi, S., Tzivanopoulos, M., Vernham, G., Violle, C., & Thuiller, W. (2025). Advancing causal inference in ecology: Pathways for biodiversity change detection and attribution. *Methods in Ecology and Evolution*, *16*, 2276–2304. https://doi.org/10.1111/2041-210X.70131