Best practices for moving from correlation to causation in ecological research

Hannah E. Correia[1*], Laura E. Dee[2], Jarrett E. K. Byrnes[3], John R. Fieberg[4], Marie-Josée Fortin[5], Clark Glymour[6], Jakob Runge[7,8], Bill Shipley[9], Ilya Shpitser[10], Katherine J. Siegel[11], George Sugihara[12], Betsy von Holle[13], and Paul J. Ferraro[1,14*]

[1]Department of Environmental Health and Engineering, Johns Hopkins University

[2]Department of Ecology and Evolutionary Biology, University of Colorado Boulder

[3]Department of Biology, University of Massachusetts Boston

[4]Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota

[5]Department of Ecology and Evolutionary Biology, University of Toronto

[6]Department of Philosophy, Carnegie Mellon University

[7]Institute of Data Science, German Aerospace Center (DLR)

[8]Technische Universität Berlin

[9]Université de Sherbrooke

[10]Department of Computer Science, Johns Hopkins University

[11]Department of Geography and Cooperative Institute for Research in Environmental Science, University of Colorado-Boulder

[12]Scripps Institution of Oceanography, University of California San Diego

[13]Department of Biology, George Washington University

[14]Carey Business School, Johns Hopkins University

[*]Correspondence to Hannah E. Correia (hec0003@auburn.edu) or Paul J. Ferraro (pferrar5@jhu.edu)

ABSTRACT

In ecology, causal questions are ubiquitous, yet the literature describing systematic approaches to answering these questions is vast and fragmented across different traditions (e.g., randomization, structural equation modeling, convergent cross mapping). In our Perspective, we connect the causal assumptions, tasks, frameworks, and methods across these traditions, thereby providing a synthesis of the concepts and methodological advances for detecting and quantifying causal relationships in ecological systems. Through a newly developed workflow, we emphasize how ecologists' choices among empirical approaches are guided by the pre-existing knowledge that ecologists have and the causal assumptions that ecologists are willing to make.

1 CAUSALITY IN ECOLOGICAL STUDIES

Ecology is centered around investigating causal relationships between living organisms and their environments. In ecology, as in many other scientific fields, causality is understood as a phenomenon where change in one variable (the "cause") induces change (the "effect") in another variable[1–4]. Thus, a causal relationship between X and Y exists if a perturbation in the cause X produces a change in the responding variable Y[4,5], potentially through the perturbations of intermediary variables[6,7]. This "perturbation-based" definition of causality is the definition most familiar to scientists and philosophers[4,8].

Because of a strong tradition of using manipulative experiments to establish causation, ecology has been shaped by two aphorisms: "correlation does not equal causation" and "causal claims can only be made from experiments." The first aphorism oversimplifies the complexity of causal relationships and has been critiqued in the literature[5,9,10] – correlation does not *always* equal causation, but correlation can suggest a causal relationship (see Section 2). More importantly, the first aphorism does not imply the second: imperfectly designed experimental studies can mistakenly suggest causal relationships where none exist, and causation can, in fact, be established through well-designed observational studies[11–13]. Natural history approaches, for instance, have long been used to establish credible causal claims (e.g., sea otters driving trophic cascades in subtidal communities[14,15]). Recently, interest in observational approaches has grown[16,17] due to the economic, ethical, and logistical challenges of manipulating ecological variables[18] and the limitations of experiments in capturing complex, large-scale causal relationships in nature[19]. Observational data, particularly from multiple locations and time points, are increasingly valued for complementing experiments and supporting more generalizable causal claims[19–21].

To formalize the requirements for making causal claims from experimental and observational data, scholars in various fields have made substantial advances in mathematical and statistical tools over the past 50 years[12,22–28]. Applications of these advances have changed how we think about scientific topics such as environmental and genetic causes of disease[29–31], military veterans' health[32], criminology[33,34], and education[35,36], and have influenced policies on air pollution[37,38] and carcinogens[39]. These same advances are increasingly being proposed by ecologists to investigate causal questions using observational[9,27,40–49] and experimental data[50–52]. Yet the way in which these advances relate to each other is not readily apparent from the published literature. For example, what are the conceptual connections between studies that use experimental designs and studies that use convergent cross mapping algorithms? Published reviews typically focus on one set of approaches at a time (e.g. quasi-experimental designs, structural causal models, dynamical systems)[27,41,44,53,54], which makes it difficult for ecologists to understand how, or if, the seemingly disparate approaches are related.

In this Perspective, we connection the assumptions, tasks, frameworks, and methods across these approaches, thereby providing a synthesis of the concepts and methodological advances for detecting and quantifying causal relationships in ecological systems. When answering a causal question, we must first identify the appropriate causal task: either causal discovery, which focuses on detecting whether causal relationships are likely to exist between variables in a system, or causal inference, which focuses on quantifying the direction and magnitude of causal relationships without bias. To accomplish these tasks, we employ causal frameworks, such as the structural causal model framework[12], the potential outcomes framework[25], or the dynamical systems causality framework[55,56], which formally define causal relationships and specify the assumptions that must be satisfied to accurately detect or quantify causal relationships from data. These frameworks then guide the selection of causal methods, that is, study designs and algorithms, which are used to operationalize these assumptions and establish the conditions necessary to make causal claims. To outline the process of navigating tasks, frameworks, and methods, we created a workflow for answering causal questions in ecological research. To provide further readings and software to implement the ideas in the Perspective, we provide comprehensive Supplemental Information (SI).

Throughout our Perspective, we highlight how well-articulated causal assumptions are the “glue” that unifies the myriad approaches to answering causal questions in ecology. These assumptions facilitate transparent discussions about the adequacy of study designs and algorithms that help scholars move from observations of statistical dependence in data to claims about causal relationships in ecological systems.

2 USING ASSUMPTIONS TO MOVE FROM CORRELATION TO CAUSATION

Data never "speak" by themselves. To derive meaningful causal insights from data, we must rely on well-defined hypotheses, statistical models grounded in ecological theory, and both testable and untestable assumptions[57–59]. The importance of hypotheses, appropriate statistical models, and statistical assumptions is well known in ecology.

Less well known is the importance of causal assumptions that allow researchers to go from making claims about correlations to making claims about causation. Unlike most statistical assumptions, causal assumptions are typically untestable; that is, causal assumptions cannot be verified from data, even unlimited data. For example, experimentalists assume that randomization of a treatment ensures that any differences in outcomes across the randomized groups can only be attributed to either the treatment or sampling variability[50]. Yet, experimentalists cannot verify this assumption. Causal assumptions, when combined with principles of probability theory and statistical dependence, allow us to make causal claims from data. The formalization of these assumptions is one of the most important scientific advances for answering causal research questions[26,28,58]. For more details on the contrast between statistical and causal assumptions, see SI Section 1.

Causal assumptions, in tandem with statistical assumptions about the data structure, establish when statistical dependence can be interpreted as evidence for the perturbation-based notion of causality[12,25,27,60]. In ecology, a commonly used measure of statistical dependence is correlation, which describes the linear similarity between two sets of observations. Consider a scenario in which we seek to determine whether, or by how much, variation in abundance of aphid predators (e.g., ladybird beetles) (X) changes the abundance of aphids (Y). If our knowledge about the probability of aphid abundance changes after learning something about ladybird beetle abundance, then ladybird beetle abundance and aphid abundance are statistically dependent. This dependence forms the starting point for investigating potential causal relationships between two variables.

Statistical dependence is linked to causality through the Common Cause Principle[61], which states that if a statistical dependence exists between two variables X and Y, then at least one of the following is true: X causes Y, Y causes X, or X and Y are both caused by a third variable C (Fig. 1). The presence of correlation can thus be mapped to the potential presence of a causal relationship. The lack of correlation, however, does not necessarily rule out statistical dependence or causality, as correlation is just one possible measure of dependence between two variables.

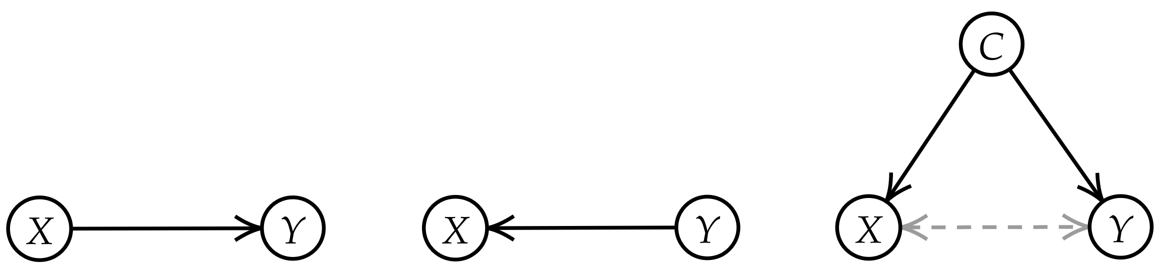


Fig. 1. Statistical dependence implies three possible causal relationships: X causes Y, Y causes X, or X and Y are caused by a common variable C. All three relationships can exist simultaneously in many contexts (indicated by the dashed grey arrows). Causal assumptions aim to eliminate the third possibility because the presence of C introduces additional statistical dependence between X and Y that is not due to any direct causal relationship.

In many causal analyses, eliminating the possibility that a third variable C causes both X and Y is a priority, because we wish to distinguish variables with direct causal links from those that are not causally influencing each other (i.e., we seek to eliminate non-causal, rival explanations for statistical dependencies). For example, in Fig. 2 broad-spectrum pesticide use (C) affects ladybird beetle abundance (X) and earthworm abundance (Y_2). However, ladybird beetle abundance does not influence earthworm abundance, nor vice versa. In this case, any observed statistical dependence between X and Y_2 is entirely attributable to their common cause C.

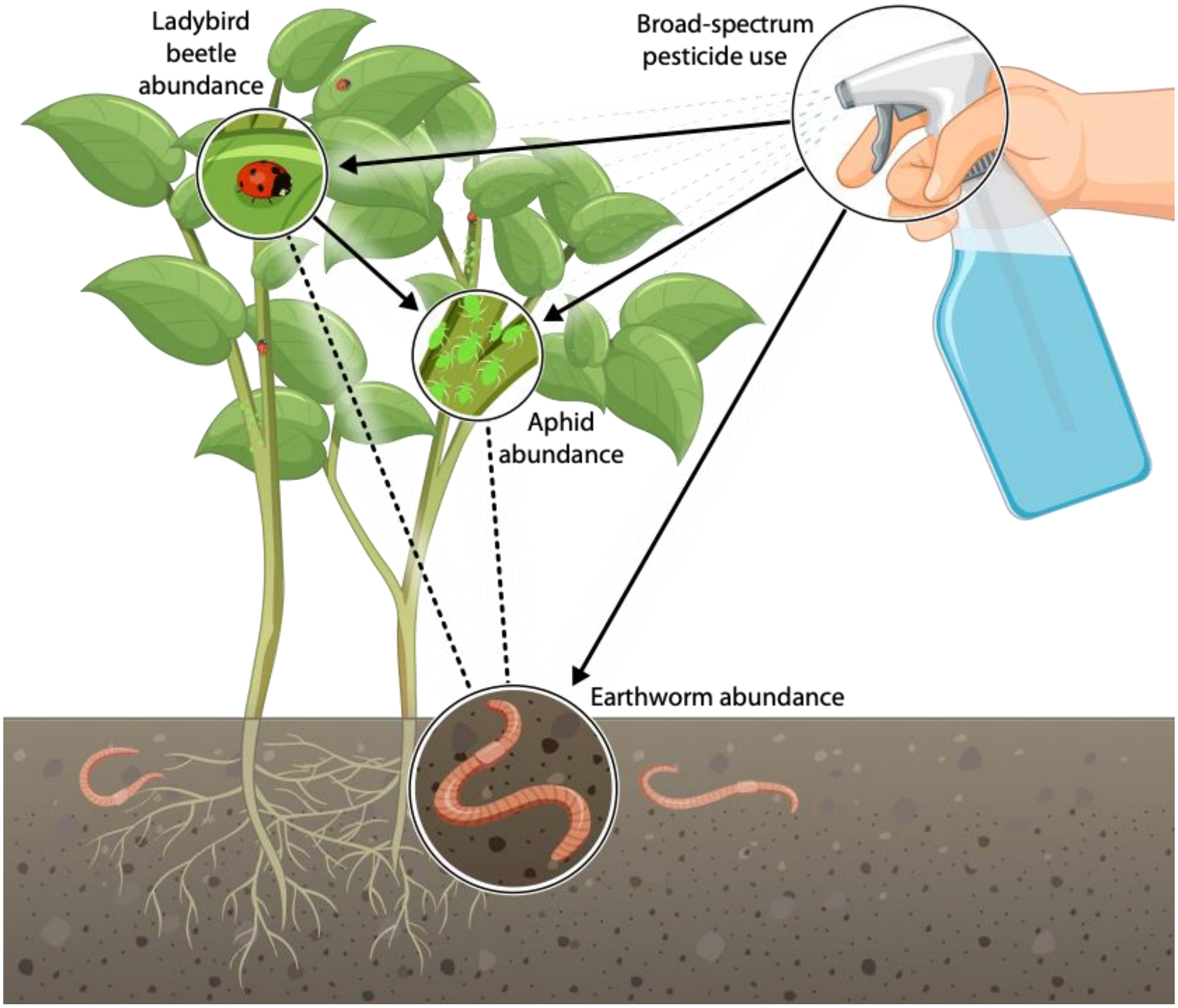


Fig. 2. Illustration of the Common Cause Principle in an ecological system where abundance of ladybird beetles, aphids, and earthworms are statistically dependent but not necessarily causally related. Blue arrows represent directional causal relationships, and red dashed lines represent statistical dependence but not causal relationships.

To eliminate these “common causes” (a.k.a., “confounding variables” or “confounders”), researchers make three assumptions: the Causal Sufficiency Assumption[28], the Causal Markov

Condition[61–63], and the Causal Faithfulness Assumption[28] (Box 1). The combination of these three untestable assumptions allows us to distinguish direct causal relationships between variables from dependence between variables induced by a common cause. By including all common causes in a model describing the relationship between X and Y (A1 in Box 1), we can eliminate the portion of the dependence attributable to shared causes $\boldsymbol{C}$ (A2). We can then interpret the remaining statistical independencies as evidence of no causal relationship between the variables (A3), while any remaining statistical dependence implies the possibility of a direct causal relationship.

For example, if pesticide use (C) is a common cause of both ladybird beetle abundance (X) and aphid abundance (Y), then we should include pesticide use in a model of the relationship between ladybird beetle abundance and aphid abundance (Fig. 2). If pesticide use is the only common cause and, after conditioning on it, ladybird beetle abundance is statistically independent of aphid abundance (i.e., they are conditionally independent), then, under the three causal assumptions, we can infer that no causal relationship between ladybird beetle abundance and aphid abundance exists. Conversely, if ladybird beetle abundance and aphid abundance are *not* independent conditional on pesticide use, then a causal relationship between ladybird beetle abundance and aphid abundance may exist (i.e., a lack of conditional independence simply means we cannot rule out a causal relationship, but it does not provide definitive evidence of causation).

The three causal assumptions required to connect statistical dependence to causal dependence – the Causal Sufficiency Assumption, Causal Markov Assumption, and Causal Faithfulness Assumption – are the foundation upon which causal claims are made from experimental and observational data. These causal assumptions allow us to differentiate the causal dependencies between two variables from the non-causal dependencies created by confounding variables.

Box 1. Three fundamental causal assumptions

For these assumptions, we define two variables X and Y as statistically dependent if the probability that Y takes a specific value given that X has taken a specific value is different from the probability that Y takes a specific value without any information about the value that X has taken. In other words, if X and Y are statistically dependent, knowing something about X changes what is known about the probability of Y.

A1. **Causal Sufficiency**[52] (a.k.a., the "no unmeasured confounding" assumption[55–57]), requires that we observe all variables in a set $\boldsymbol{C}$ that causally influence any pairs of variables X and Y, and we include $\boldsymbol{C}$ in our model that describes the relationship between X and Y, thus ensuring that no confounding variables are unobserved.

A2. The **Causal Markov Condition**[54,58,59] states that if a pair of variables X and Y are statistically dependent solely because both are caused by a common variable C, and if we control for C by including it in our model, then X and Y become conditionally independent given C.

A3. **Causal Faithfulness**[52], stated very loosely, declares that statistical independence (conditional or unconditional) between a pair of variables X and Y indicates the absence of a causal relationship between those variables.

The combination of the **Causal Markov Assumption** (A2) and the **Causal Faithfulness Assumption** (A3) allows us to claim that if two variables, X and Y, are conditionally independent when C is included in the model, then X and Y are not causally related but instead are caused by a third common variable C. The **Causal Sufficiency Assumption** (A1) then ensures that we can distinguish causal relationships from dependence induced by a common cause if we include all possible confounders between variables in a model that describes the relationship between X and Y.

The Causal Markov and Causal Faithfulness assumptions have formal definitions requiring technical notation that are beyond the scope of this article. For a full discussion of these assumptions, we refer the reader to Pearl (2000)[23] and Spirtes and Zhang (2016)[60].

3 SATISFYING CAUSAL ASSUMPTIONS WITH PRE-EXISTING KNOWLEDGE, STUDY DESIGNS, AND ALGORITHMS

Given the restrictive and untestable nature of the three causal assumptions introduced in Section 2, ecologists may wonder whether causal claims can realistically be made from ecological data, since satisfying these assumptions requires building models that account for all confounders. Unlike models built for prediction or description, models built to make causal claims cannot be validated using goodness-of-fit or predictive accuracy metrics, as these metrics assess how well a model describes the observed data but do not evaluate how well the model satisfies the untestable assumptions required for making causal claims[64,65] (for more details, see SI Section 2). In the following subsections, we describe how the foundations for satisfying causal assumptions are provided by pre-existing knowledge, study designs, and algorithms.

3.1 Pre-existing knowledge

To satisfy the three causal assumptions, pre-existing knowledge is essential[59]. We use pre-existing knowledge to hypothesize causal relationships between variables by specifying the outcome(s) of interest (Y) and identifying potential causes (X). We also use pre-existing knowledge to identify potential confounders and determine which confounders can be measured in a study[43,66,67]. The more pre-existing knowledge that we can apply towards satisfying causal assumptions, the more sophisticated the causal questions we can answer.

Pre-existing knowledge can include general and domain-specific ecological theory, subject matter expertise, field experience, and findings from previous studies, including studies that use empirical approaches lacking causal interpretations (see SI Section 2). Because pre-existing knowledge is often complex and wide-ranging, we need succinct and straightforward ways to summarize it. In Section 5, we describe two common tools for organizing our understanding of an ecological system (i.e., our 'mechanistic knowledge'[66]).

3.2 Study designs and algorithms

Pre-existing knowledge is typically not sufficient to satisfy causal assumptions. For instance, even if we can identify all confounders with pre-existing knowledge, we are unlikely to be able to measure them all, which would be necessary to satisfy the Causal Sufficiency Assumption. However, study designs and algorithms provide us with the opportunity to address such challenges by relaxing one or more of the three causal assumptions in Section 2 in favor of equally untestable but (hopefully) more plausible causal assumptions.

Experimental designs, for example, substitute the Causal Sufficiency Assumption with the assumption that treatment randomization eliminates the effects of unmeasured confounding variables[25,68]. Confounders are thus addressed through design rather than measurement. In non-experimental studies, observational designs often relax the Causal Sufficiency Assumption through statistical techniques that define the minimum set of confounding variables that need to be observed to accomplish the desired causal task[57,58,69], or through statistical techniques that allow researchers to pursue alternative research goals that reduce the number of confounders that must be measured (e.g., by defining alternative causal effects[70,71]). These statistical techniques and redefined research goals can also be used with experimental designs that face implementation challenges, such as when the experimental manipulation affects the outcome variable through other pathways (i.e., randomization is a confounder), or when post-randomization observations are missing (i.e., attrition). We provide more details on both experimental and observational study designs and algorithms for each of the causal tasks in Section 8.

4 A WORKFLOW FOR ANSWERING CAUSAL QUESTIONS IN ECOLOGY

We now present a comprehensive workflow that summarizes the key steps for conducting causal analyses (Fig. 3), with examples for each of the steps using two hypothetical ecological studies (Box 2). Our workflow illustrates how to systematically address causal questions in ecology.

The workflow serves as a roadmap, starting with the description of the causal question and ending with the interpretation and validation of the results. Each step in the workflow represents a decision point where we take action to ensure our causal analysis is robust, transparent, and aligned with the assumptions necessary to make causal claims from statistical analyses of ecological data. The workflow is designed to be flexible, allowing ecologists to tailor their approach based on their pre-existing knowledge, data, and methodological preferences.

To illustrate the workflow's application to real-word ecological research, we use two example ecologists, an intertidal ecologist and a tiger ecologist. In Box 2, we summarize how each ecologist navigates the workflow. In Sections 5 through 8, we elaborate on each of the following workflow steps:

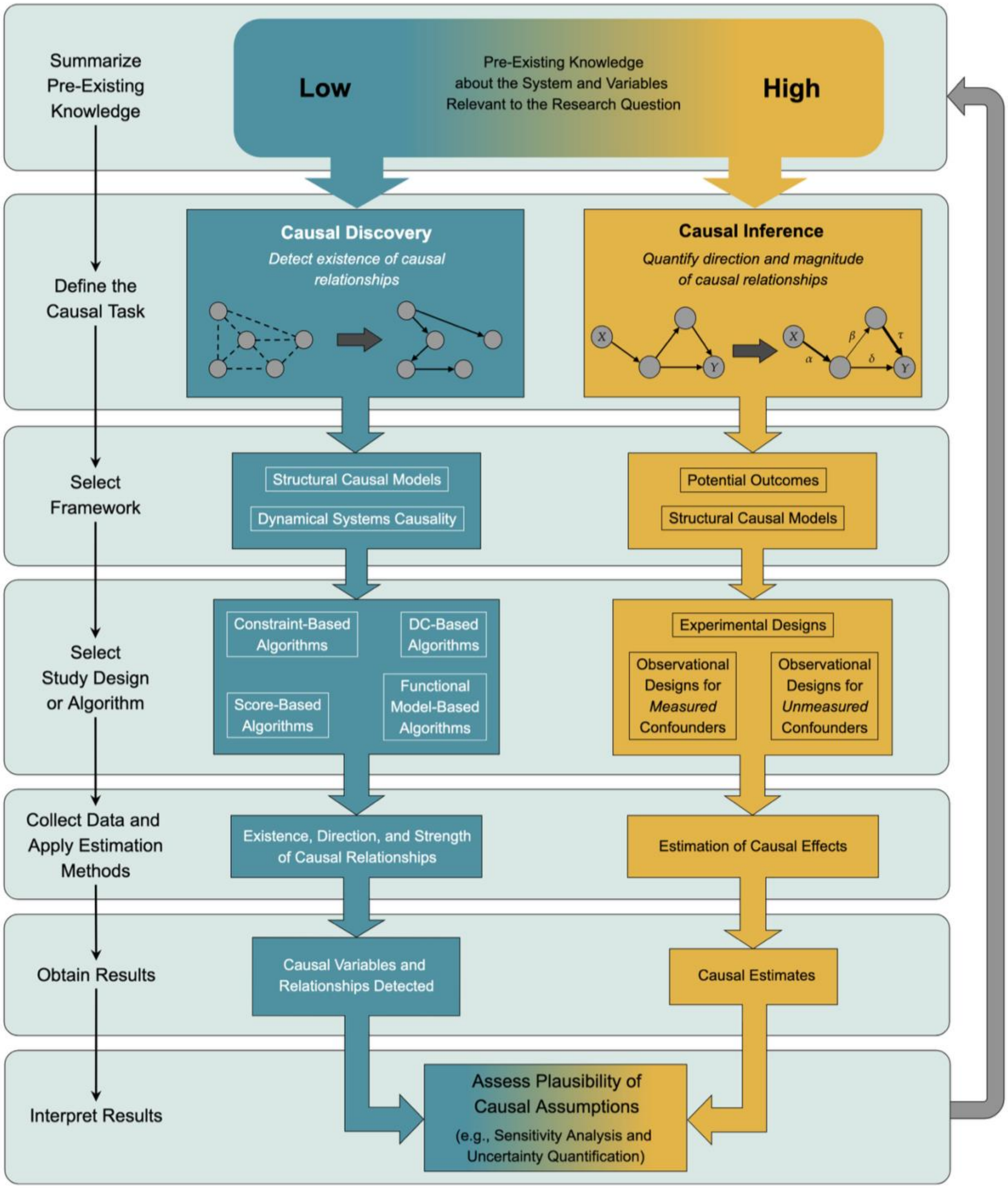
1. **Define the Causal Question and Summarize Pre-Existing Knowledge** (Section 5): We must first define the causal research question with at least one outcome variable (Y) and one or more hypothesized causal variables (X) (see SI Section 2 for differences between causal and non-causal questions in ecology). Then, to identify all confounding variables,

we must assess the corpus of pre-existing knowledge on the causes and outcomes of interest. We can summarize this knowledge using causal diagrams or thought experiments.

2. **Define the Causal Task** (Section 6): When answering causal questions, we use pre-existing knowledge to determine whether to pursue causal discovery or causal inference. Causal inference, which seeks to quantify the magnitudes of causal relationships, is feasible when we have sufficient pre-existing knowledge to be confident of the causal, outcome, and confounding variables and the directions of the causal relationships. If this knowledge is insufficient, we can instead pursue causal discovery, which aims to detect the existence of causal relationships.
3. **Select Framework** (Section 7): To clearly articulate the causal and statistical assumptions that must be satisfied for valid claims in either causal task, we can use one or more causal frameworks. The potential outcomes framework and the structural causal model framework are two common frameworks used for causal inference. For causal discovery, the structural causal model and dynamical systems causality frameworks are frequently used.
4. **Select Study Design or Algorithm**, **Collect Data and Apply Estimation Methods**, **Obtain Results**, and **Interpret Results** (Section 8): For causal inference, study designs can be grouped into three categories: experimental designs, observational designs for measured confounders, and observational designs for unmeasured confounders. Within each of these categories, many methods exist, some of which are described in SI Table S6 (e.g., regression adjustment[72], propensity score matching[45,73], and structural equation modeling[9]). For causal discovery, algorithms are used instead of study designs. These fall

into four categories: constraint-based, score-based, functional model-based, and dynamical systems causality-based. Some of the algorithms are described in SI Table S7 (e.g., convergent cross mapping[27], fast causal inference[28], and greedy equivalency search[74]). Based on the requirements of the study design or algorithm, we then collect data and apply estimation methods to detect causal relationships or quantify causal effect(s). Afterwards, we interrogate the plausibility of the causal and statistical assumptions by identifying potential violations to the assumptions and exploring the implications of those violations for the conclusions.

Although we present the workflow in a linear fashion, researchers will use it iteratively in two ways: (i) the results from one causal analysis will feed into future analyses in the form of pre-existing knowledge[66] (grey arrow in Fig. 3); and (ii) after taking actions at one step, researchers may need to return to previous steps before advancing in the workflow (e.g., reassessing the study design if data collection did not go as planned).



278

279 Fig. 3. A workflow that outlines the key steps and decisions for answering causal questions in

280 ecological research.

Box 2. Ecologists conducting causal analyses using the workflow in Fig. 3.

"Define the Causal Question and Summarize Pre-Existing Knowledge"

An intertidal ecologist seeks to quantify the change in bivalve abundance (Y) caused by floods (X) through changes in nitrogen (M_1) and salinity (M_2) in intertidal zones at the mouth of an estuary. The ecologist summarizes knowledge about all confounders for each of the causal relationships of interest (i.e., floods on bivalves, floods on nitrogen, floods on salinity, nitrogen on bivalves, and salinity on bivalves).	A tiger ecologist seeks to determine the ecological factors ($\boldsymbol{X}$) that encourage more visits or longer time (Y) spent in certain locations by tigers. The ecologist summarizes knowledge about confounders of the causal relationship between ecological factors and tiger occupancy (e.g., geographic and human factors).

"Define the Causal Task"

The intertidal ecologist has robust ecological theory and a significant collection of prior studies that identified the set of all confounding variables that could bias estimation of any one of the causal relationships of interest. Thus, the ecologist pursues causal inference.	The tiger ecologist has theory and field observations to identify certain ecological factors that may influence tiger occupancy, but they do not have sufficient knowledge to identify all human and geographic confounding variables. Thus, the ecologist pursues causal discovery.

"Select Framework"

The intertidal ecologist prefers the structural causal model framework for its structural approach to reasoning about multiple causes jointly.	The tiger ecologist prefers the dynamical systems causality (DC) framework for its focus on complex, evolving systems.

"Select a Study Design or Algorithm"

The intertidal ecologist selects an observational study design in which they measure and condition on all confounding variables.	The tiger ecologist selects a DC-based algorithm.

"Collect Data and Apply Estimation Methods"

The intertidal ecologist collects observational cross-sectional data on <u>all</u> causal, outcome, and confounding variables related to the causal relationships of interest and then fits a structural equation model.	The tiger ecologist collects observational time series data for tiger occurrence, abundance of several prey species, poaching activity, and weather conditions at a series of locations and uses convergent cross mapping (CCM) to detect causal relationships between pairs of variables.

"Obtain and Interpret Results"

The intertidal ecologist obtains estimates of the causal effects of floods on bivalve abundance that arise though the changes in nitrogen and salinity. They perform a causal sensitivity analysis that quantifies how much the estimates change in the presence of an unmeasured confounding variable.	The tiger ecologist obtains a network with detected causal relationships between pairs of variables. They perform a sensitivity analysis that shows how the detected causal relationships change when the CCM hyperparameter settings are changed.

5 SUMMARIZE PRE-EXISTING KNOWLEDGE

One common conceptual tool for summarizing pre-existing knowledge is a causal diagram. Causal diagrams help us organize our pre-existing knowledge by visually mapping the presumed causal relationships among causes (X), their outcomes (Y), and confounding variables ($\boldsymbol{C}$). The most widely-used type of causal diagram is the causal directed acyclic graph (causal DAG), which follows a set of formal rules that define how causal relationships must be encoded[75]. A causal DAG includes the focal variables of a study (i.e., the "cause" and the "outcome" variables), along with all suspected common causes (i.e. confounders) between the focal variables. Directed edges (arrows) between variables indicate that unidirectional causal relationships are presumed to exist, and the absence of an arrow between two variables reflects a strong assumption that a causal relationship does not exist[12]. Causal DAGs, which must include all potential confounders of presumed causal relationships, enable us to identify the confounding variables we need to address with an experimental or statistical technique. Thus, causal DAGs should be constructed at the beginning of a study, before data are collected and the specific study design or algorithm is chosen.

Some ecologists will be familiar with the structural equation model (SEM) diagram[9], which can be interpreted as a causal DAG when its structure represents only unidirectional relationships and explicitly encodes assumptions about causal relationships, including all relevant confounders[76,77]. SEM diagrams also include additional parametric assumptions and are purpose-built for SEM analyses[76], whereas causal DAGs, which require no assumptions about the functional forms of causal relationships between variables, can be used in any type of causal analysis.

Another conceptual tool for summarizing pre-existing knowledge is a thought experiment in which researchers consider how a hypothetical ideal randomized controlled trial (RCT) – often termed a "target trial"[78,79] – would be designed to answer their causal research question[25]. By comparing the ideal (target) trial with the actual data generating process, we can identify discrepancies that may lead to bias through confounding variables that distort the observed relationship between the causal variable and the outcome. Formulating such a target trial forces us to articulate all the key components of an ideal RCT and then systematically determine which of these components may be absent or imperfect in our study. In doing so, it becomes clearer which variables, including potential confounders, should be adjusted for to emulate the conditions of an ideal experiment. Just as drawing causal DAGs helps visualize the network of causal relationships and identify confounders, formulating these thought experiments provides a concrete tool for planning rigorous study designs (i.e., the thought experiment forces us to ask the question, "Where does the variation in the causal variable come from?" a.k.a., "What is the treatment assignment mechanism?"). For resources that describe how to draw causal DAGs or develop thought experiments for studies, see SI Section 3.

6 DEFINE THE CAUSAL TASK – CAUSAL DISCOVERY OR CAUSAL INFERENCE

In deciding the most appropriate causal task for a research question, we must carefully consider the gap between available knowledge and the knowledge that would be required to plausibly satisfy causal assumptions. When pre-existing knowledge is extensive, we may pursue the task of causal inference. When pre-existing knowledge is limited, we may instead pursue causal discovery. Although the dividing lines between these two tasks is not as clearcut as implied in our workflow (i.e., causal research lies on a continuum rather than in one of two camps), the contrast between their goals is illuminating for understanding how each task draws on pre-existing knowledge.

The goal of **causal inference** is to quantify the magnitudes of causal effects, either under a range of typical conditions or under specific interventions (e.g., new management policies, abrupt ecological changes). Causal inference requires substantial pre-existing knowledge about which variables act as causes, outcomes, and confounders, as well as the directions of causal processes ("high" pre-existing knowledge in Fig. 3). Quantifying multiple causal effects within an ecological system is even more challenging because sufficient pre-existing knowledge must exist to satisfy the required causal assumptions for every pair of cause-outcome variables.

When quantifying causal effects, defining the specific effect(s) of interest is important for connecting theoretical quantities to data. Different causal effects require different variations of the causal assumptions[80]. Ecologists are often interested in the average effect of X on Y across all observations, that is, the average change in the outcome Y per unit change in X. However, other effects may also be relevant, such as mediation effects[81] (effects of intermediary variables between a cause and its outcome) or effects for subgroups[82] (e.g., the average effect of X on Y only for observations which experienced specific values of X). Moreover, some causal effects may be preferred because the causal assumptions for these effects can be more plausibly satisfied for a study (e.g., complier average causal effects, local average treatment effects, etc.).

In contrast to causal inference, **causal discovery** aims to detect or "learn" causal relationships among measured variables. Although causal discovery requires causal assumptions, they are less restrictive than they are in causal inference, and thus, less pre-existing knowledge is required ("low" pre-existing knowledge in Fig. 3). While causal discovery methods offer flexibility in investigating causal questions with limited pre-existing knowledge, this advantage comes with the trade-off of potentially less precise or less certain conclusions about causal relationships. Causal discovery is therefore primarily valuable for generating more knowledge to guide subsequent studies.

To detect causal relationships, causal discovery involves defining an initial causal diagram (see Section 5) and refining it with statistical evidence from data. One strategy begins with a causal diagram that assumes causal relationships exist among all variables. Statistical independence tests are then systematically applied to eliminate connections between variables where evidence of a causal relationship is not supported by the data[53]. Another strategy starts with a causal diagram that assumes no causal connections among variables and iteratively adds them where statistical evidence suggests a potential causal relationship[83]. This second strategy is particularly amenable to incorporating pre-existing knowledge by allowing researchers to specify relationships that should be included or excluded from the outset. Both strategies rely on variations of the three causal assumptions introduced in Section 2 and aim to produce a refined causal diagram that reflects only the causal relationships consistent with the observed data and the underlying assumptions.

7 SELECT A CAUSAL FRAMEWORK

Causal frameworks structure how causal assumptions are represented for a given task, ensuring consistency between study design/algorithm, data collection, and estimation procedures.

7.1 Causal frameworks for causal inference

For causal inference, assumptions and estimation procedures are expressed using one of three causal frameworks: the structural causal model (SCM) framework; the Neyman-Rubin causal model, also commonly known as the potential outcomes (PO) framework; and the decision-theoretic framework[22]. While we focus on the SCM and PO frameworks, readers interested in the decision-theoretic framework can refer to Dawid (2000)[22] and Dawid (2012)[84].

The choice of framework is primarily based on researcher preferences, as the PO and SCM frameworks have been shown to be logically and mathematically equivalent[85–87]. The PO framework may appeal to experimentalists because it expresses causal assumptions by approximating the conditions that most accurately represent an idealized "gold standard" randomized controlled experiment. Alternatively, researchers who primarily model ecological systems as collections of simultaneously interacting variables may prefer the SCM framework, which represents systems as causal DAGs. Structural equation modeling, when used to make causal claims under the necessary causal assumptions[9,46], is a subset of the SCM framework[77,88].

Formalizations of the causal assumptions for causal inference as expressed using the PO framework and the SCM framework are described in SI Section 4 and Box S1. Resources for learning more about the core concepts of the PO and SCM frameworks can be found in SI Section 5.

7.2 Causal frameworks for causal discovery

For causal discovery, the assumptions and estimation procedures are expressed using either the SCM framework or the dynamical systems causality (DC) framework[55,56]. Causal discovery using the SCM framework is well-suited for ecological systems with multiple interacting variables, where causal relationships are expected to be stable across observations. SCM-based causal discovery algorithms also allow researchers to incorporate pre-existing knowledge by specifying constraints on potential causal relationships, making them particularly useful for exploratory studies where some causal relationships are known or hypothesized. In contrast, the DC framework may be more suitable for complex dynamic systems where causal effects unfold over time and cannot be represented as static combinations of causative factors. DC-based algorithms typically use time series data to infer causal relationships by testing whether knowledge of once variable's past improves the ability to anticipate changes in another variable. Measures of improvement span changes in predictability or statistical dependence, including those captured by information-theoretic measures[83,89].

SCM-based causal discovery algorithms generally begin with a causal diagram that assumes relationships between all variables in the data, and then they iteratively test for statistical independence between pairs of variables. Edges are removed where statistical independence is found, refining the causal diagram to represent only causal relationships consistent with the statistical independencies reflected in the data[53]. In contrast, DC-based algorithms typically start with no assumed causal relationships among variables, and test whether statistical dependence between each pair of variables in each direction ($X \rightarrow Y$ and $Y \rightarrow X$) are significantly different from white noise or null hypothesis models[83,90]. If the dependence meets the threshold for significance (typically, $\alpha = 0.05$) in only one of the directions, say $X \rightarrow Y$, then asymmetric coupling is detected, indicating a causal information flow from X (the driving system) to Y (the response system). The strength of the causal relationship is then estimated using a distance metric[56,83]. In both the SCM and DC frameworks, multiple causal diagrams can be consistent with the same structure of statistical dependencies in data, but pre-existing knowledge can refine the causal diagrams by constraining what relationships are possible.

Formalizations of the causal assumptions for causal discovery as expressed using the SCM framework and the DC framework are described in SI Section 4 and Box S2. Resources for learning more about the core concepts of the SCM and DC frameworks can be found in SI Section 5.

8 SELECT A STUDY DESIGN OR ALGORITHM, APPLY ESTIMATION METHODS, OBTAIN RESULTS, AND INTERPRET RESULTS

Study designs for causal inference and algorithms for causal discovery provide structured approaches for satisfying or relaxing the untestable causal assumptions through decisions about the data and analysis (i.e., designs and algorithms operationalize causal frameworks). Designs and algorithms also lead us to appropriate methods for estimation and interpretation of the results.

This section provides an overview of key study designs for causal inference and algorithms for causal discovery. The details and applications of each approach are beyond the scope of this Perspective, but in SI Section 6 we provide resources, including guidance on implementation and relevant software packages. While we focus on foundational study designs and algorithms, we summarize in SI Section 7 some advanced methods, including those that integrate machine learning techniques into their estimation procedures, which are rapidly emerging and may offer new opportunities for ecological research.

8.1 Study designs for causal inference

Study designs for causal inference fall into three categories: (1) experimental designs that aim to minimize confounding from both measured and unmeasured variables through manipulation of the causal variable, (2) observational designs that explicitly identify and control for measured confounders, and (3) observational designs that eliminate unmeasured, and potentially unknown, confounding by leveraging external sources of variation (specific designs from these three categories are listed in SI Table S6).

Experimental designs (e.g., randomized controlled trials[50] and factorial designs[91]) are often well-suited for causal inference because they provide a structured approach for directly manipulating the causal variable and defining the temporal order of cause and effect. Through strategies like randomization, we aim to control or eliminate the effects of confounding variables, providing justification for causal claims. However, suboptimal decisions in the design and analysis of experiments can produce invalid causal conclusions[92], and even well-designed experiments may face challenges[93], such as non-compliance or non-random dropout. Moreover, in ecology, experiments may be prohibitively expensive at the scales needed to detect causal effects, or they may distort natural ecological conditions[94], making them impractical or unrepresentative.

When experiments are infeasible, impractical, or unethical, observational designs for measured and unmeasured confounders are available. Advances in causal approaches for observational studies provide statistical techniques to satisfy causal assumptions without experimental manipulation[12,22,25,75,95]. Observational designs for measured confounders (e.g., regression adjustment[72], propensity score matching[73], and structural equation modeling[9]) rely on measuring all confounding variables. When measuring, or even knowing, all relevant confounders is not feasible, we can use observational designs for unmeasured confounders (e.g., before-after-control-impact[96]and multilevel modeling with fixed effects[97]). These designs relax the causal sufficiency assumption of no unmeasured confounders by replacing it with assumptions about the structure of unmeasured confounders, typically informed by pre-existing knowledge. These designs then use statistical techniques to represent the influence of confounders based on their assumed structure, without needing to directly measure the confounders.

Experimental and observational designs can be implemented using either cross-sectional or longitudinal data. However, strong assumptions about temporal ordering (cause must precede its outcome) and stable effects over time are required to quantify causal effects using cross-sectional data. Once data are collected, we can quantify the causal effect of interest using a range of estimation methods ("Collect Data and Apply Estimation Methods" and "Obtain Results" in Fig. 3). Many estimation methods are available to implement a chosen study design, each providing a different statistical approach for estimating the causal effect of interest[98,99].

After estimating a causal effect, we must then interrogate the plausibility of the causal assumptions underlying the study design and explore the implications of violations to these assumptions ("Interpret Results" in Fig. 3). One common approach for assessing the implications of violations is to perform causal sensitivity analyses, which quantify how an estimated effect would change in the presence of unaddressed confounding. Many sensitivity analysis techniques are available for a variety of causal inference methods[100–104], including SEM[105]. An alternative approach to interrogating the plausibility of causal assumptions involves detecting under-adjustment of confounding variables by drawing on pre-existing knowledge to formulate tests of known effects[11,106,107] (e.g., falsification or placebo tests). We must also consider how other forms of bias[108,109], including selection bias[110,111] and measurement bias[112–114], may influence the estimated effects and the robustness of our conclusions.

8.2 Algorithms for causal discovery

Algorithms for causal discovery fall into four categories: DC-based algorithms and three types of SCM-based algorithms, which are called constraint-based, score-based, and functional model-based algorithms (specific algorithms from these four categories are listed in SI Table S6). DC-based methods are suited for dynamic systems and assess causal interactions based on predictability and information flow over time. Constraint-based methods use conditional independence tests to eliminate implausible causal relationships. Score-based methods evaluate possible graphical structures using a scoring criterion that captures how well the graph fits patterns of conditional independencies in the data. Functional model-based methods assume specific functional relationships between variables (e.g., linear or non-linear equations with noise) and infer causal direction by identifying which graph configuration satisfies those assumptions.

Causal discovery algorithms have been developed to accommodate different data structures, with approaches often tailored to either longitudinal data or cross-sectional data. DC-based methods (e.g., Granger causality[60] and convergent cross mapping [CCM][27]) require bivariate or multivariate time-series data (i.e., regularly spaced longitudinal data) to infer causal relationships through changes over time. In contrast, SCM-based algorithms (e.g., Fast Causal Inference [FCI][28], Greedy Equivalency Search [GES][74], and Peter and Clark Momentary Conditional Independence [PCMCI][115]) can be applied to both cross-sectional and longitudinal data, but additional assumptions about temporal ordering (i.e., causes precede their outcomes) must be satisfied when using cross-sectional data. As with causal inference, pre-existing knowledge can enhance results from SCM-based discovery methods by explicitly specifying certain relationships that should or should not be included in the causal diagram.

Once candidate causal diagrams have been obtained ("Collect Data and Apply Estimation Methods" and "Obtain Results" in Fig. 3), we must assess whether the causal assumptions of the chosen discovery algorithm are plausible for the ecological system under study and explore the implications of violations to these assumptions ("Interpret Results" in Fig. 3). To assess the reliability of conclusions drawn from the causal discovery process and to evaluate the robustness of the inferred causal relationships, sensitivity analyses that explore the stability of results across different hyperparameter settings should be undertaken[116].

9 CHALLENGES AND OPPORTUNITIES

Making valid causal claims from ecological data requires moving beyond analyses that use prediction- and association-focused models, which typically fail to represent the true underlying causal structures of ecological systems[64,117,118]. It instead requires satisfying or carefully relaxing the causal assumptions that allow observed statistical dependencies to be interpreted as evidence of causal relationships. While this requirement may seem daunting, especially given the complexity of ecological systems, advances in causal methodologies have demonstrated how the quality and transparency of causal claims can be improved through clearer articulation of the causal assumptions, scrutiny of their plausibility, and attention to potential violations. For example, the intertidal ecologist who uses SEM to estimate causal effects and the tiger ecologist who uses CCM to discover causal relationships (Box 2) would (i) clearly state all the causal assumptions required by their study design or algorithm; (ii) quantify the amount of unmeasured confounding that would be needed to overturn their causal claims, or, for discovery algorithms, report how detected causal relationships change under different hyperparameter settings; and (iii) frankly discuss potential unmeasured confounding variables or other violations to their causal assumptions that could invalidate their conclusions.

By connecting the causal assumptions, tasks, frameworks, and methods that play essential roles in causal research, our workflow (Fig. 3) provides a structured approach for investigating causal questions in ecology. The workflow emphasizes the role of pre-existing knowledge, which helps us to align the causal task with the research objective, clarify assumptions through a causal framework, and select a study design or algorithm that satisfies those assumptions and guides data collection and analysis. Studies that explicitly state and justify the assumptions underlying their causal claims allow subject matter experts to evaluate the credibility of these assumptions and build on them more effectively. Thus, our workflow not only supports ecologists in conducting rigorous and transparent causal analyses, but it also facilitates cogent discussions about the potential for unresolved confounding in prior studies, which can motivate new studies. Through an iterative application of the workflow, we can enhance the accumulation and synthesis of ecological knowledge.

As causal methods evolve, new advances focus on relaxing or probing untestable assumptions in challenging real-world settings, which expand the relevance and applicability of causal methods to the complexities of ecological systems. Ecologists are uniquely positioned not only to benefit from these advances in causal analysis, but also to contribute meaningfully to their development. Ecologists' experience with experimental study designs, multiscale complex systems, and the integration of biotic and abiotic processes offers valuable insights into widespread challenges in causal research, such as spatial interactions, downscaling, and unit-to-unit causation. As causal approaches become more accessible and adaptable, ecologists have an opportunity to refine long-standing questions, generate new theory, and develop credible causal explanations of the natural world.

Acknowledgements

This work emerged partly from discussions at the workshop “Causality in Ecology” in August 21–23, 2023 in Baltimore, MD, USA. We thank Johns Hopkins University for funding and Rachel Pickett, Carter Polston, Kip Hinton, and Shang Jones for assistance in hosting the workshop. We thank Ashley E. Larsen for insightful discussions during the workshop and feedback on drafts of the paper. H.E.C and P.J.F. acknowledge funding support from USDA-NIFA award 2023-67023-39033.

Author Contributions Statement

H.E.C. led the paper. H.E.C, L.E.D and P.J.F co-organized, and P.J.F. funded, the workshop in which J.E.K.B., H.E.C., L.E.D., J.R.F., P.J.F., M-J.F., C.G., B.S., I.S., K.J.S., G.S., and B.vH. contributed to establishing the goals and emphases of the paper. H.E.C., L.E.D and P.J.F initiated the paper concept and framing. H.E.C. and P.J.F. wrote the main text. J.E.K.B., L.E.D., J.R.F., M-J.F., B.S., I.S., K.J.S., G.S., and B.vH. suggested edits to the drafts of the paper. H.E.C. conceived and wrote the Supplemental Information.

REFERENCES

1. Laland, K. N., Sterelny, K., Odling-Smee, J., Hoppitt, W. & Uller, T. Cause and Effect in Biology Revisited: Is Mayr's Proximate-Ultimate Dichotomy Still Useful? *Science* **334**, 1512–1516 (2011).
2. Mayr, E. Cause and Effect in Biology. *Science* **134**, 1501–1506 (1961).
3. Woodward, J. Causation in biology: stability, specificity, and the choice of levels of explanation. *Biol. Philos.* **25**, 287–318 (2010).
4. Ben-Menahem, Y. *Causation in Science*. (Princeton University Press, 2018).
5. Wagner, A. Causality in Complex Systems. *Biol. Philos.* **14**, 83–101 (1999).
6. Poliseli, L., Coutinho, J. G. E., Viana, B., Russo, F. & El-Hani, C. N. Philosophy of science in practice in ecological model building. *Biol. Philos.* **37**, 21 (2022).
7. Raerinne, J. Causal and Mechanistic Explanations in Ecology. *Acta Biotheor.* **59**, 251–271 (2011).
8. Woodward, J. *Making Things Happen: A Theory of Causal Explanation*. (Oxford University Press, 2004). doi:10.1093/0195155270.001.0001.
9. Shipley, B. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference with R*. (Cambridge university press, Cambridge (GB), 2016).
10. Ross, L. N. Causes with material continuity. *Biol. Philos.* **36**, 52 (2021).
11. Rosenbaum, P. R. Known Effects. in *Observational Studies* 136–153 (Springer New York, New York, NY, 1995). doi:10.1007/978-1-4757-2443-1_5.
12. Pearl, J. *Causality: Models, Reasoning, and Inference*. (Cambridge University Press, Cambridge, U.K. ; New York, 2009).

13. Dominici, F., Bargagli-Stoffi, F. J. & Mealli, F. From Controlled to Undisciplined Data: Estimating Causal Effects in the Era of Data Science Using a Potential Outcome Framework. *Harv. Data Sci. Rev.* (2021) doi:10.1162/99608f92.8102afed.

14. Estes, J. A. & Palmisano, J. F. Sea Otters: Their Role in Structuring Nearshore Communities. *Science* **185**, 1058–1060 (1974).

15. Estes, J. E., Smith, N. S. & Palmisano, J. F. Sea Otter Predation and Community Organization in the Western Aleutian Islands, Alaska. *Ecology* **59**, 822–833 (1978).

16. Sagarin, R. & Pauchard, A. Observational approaches in ecology open new ground in a changing world. *Front. Ecol. Environ.* **8**, 379–386 (2010).

17. Sagarin, R. & Pauchard, A. *Observation and Ecology: Broadening the Scope of Science to Understand a Complex World*. (Island Press/Center for Resource Economics, Washington, DC, 2012).

18. Benedetti-Cecchi, L. *et al.* Hybrid datasets: integrating observations with experiments in the era of macroecology and big data. *Ecology* **99**, 2654–2666 (2018).

19. De Boeck, H. J. *et al.* Global Change Experiments: Challenges and Opportunities. *BioScience* **65**, 922–931 (2015).

20. McCleery, R. *et al.* Uniting Experiments and Big Data to advance ecology and conservation. *Trends Ecol. Evol.* **38**, 970–979 (2023).

21. Wootten, T. & Pfister, C. The Motivation for and Context of Experiments in Ecology. in *Experimental ecology: issues and perspectives* (Oxford University Press, 1998).

22. Dawid, P. Causal Inference without Counterfactuals. *J. Am. Stat. Assoc.* **95**, 407–424 (2000).

23. Holland, P. W. Statistics and Causal Inference. *J. Am. Stat. Assoc.* **81**, 945–960 (1986).

24. Imbens, G. W. & Rubin, D. B. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. (Cambridge University Press, 2015). doi:10.1017/CBO9781139025751.

25. Rubin, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701 (1974).

26. Rubin, D. B. Causal Inference Using Potential Outcomes: Design, Modeling, Decisions. *J. Am. Stat. Assoc.* **100**, 322–331 (2005).

27. Sugihara, G. *et al.* Detecting Causality in Complex Ecosystems. *Science* **338**, 496–500 (2012).

28. Spirtes, P., Glymour, C. N. & Scheines, R. *Causation, Prediction, and Search*. (MIT Press, Cambridge, Mass, 2000).

29. D'Onofrio, B. M. *et al.* Causal Inferences Regarding Prenatal Alcohol Exposure and Childhood Externalizing Problems. *Arch. Gen. Psychiatry* **64**, 1296–1304 (2007).

30. Pearce, N., Vandenbroucke, J. P. & Lawlor, D. A. Causal Inference in Environmental Epidemiology: Old and New Approaches. *Epidemiology* **30**, 311–316 (2019).

31. Pingault, J.-B. *et al.* Using genetic data to strengthen causal inference in observational research. *Nat. Rev. Genet.* **19**, 566–580 (2018).

32. White, R. F. *et al.* Recent research on Gulf War illness and other health problems in veterans of the 1991 Gulf War: Effects of toxicant exposures during deployment. *Cortex J. Devoted Study Nerv. Syst. Behav.* **74**, 449–475 (2016).

33. Beck, B., Antonelli, J. & Piñeros, G. Effects of New York City's Neighborhood Policing Policy. *Police Q.* **25**, 470–496 (2022).

34. Wikström, P.-O. H. & Kroneberg, C. Analytic Criminology: Mechanisms and Methods in the Explanation of Crime and its Causes. *Annu. Rev. Criminol.* **5**, 179–203 (2022).

35. Jacob, B. A. & Lefgren, L. Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. *Rev. Econ. Stat.* **86**, 226–244 (2004).

36. Long, B. T. & Kurlaender, M. Do Community Colleges Provide a Viable Pathway to a Baccalaureate Degree? *Educ. Eval. Policy Anal.* **31**, 30–53 (2009).

37. Brewer, D., Dench, D. & Taylor, L. O. Advances in Causal Inference at the Intersection of Air Pollution and Health Outcomes. *Annu. Rev. Resour. Econ.* **15**, 455–469 (2023).

38. National Academies of Sciences, Engineering, and Medicine *et al.* Definition of Causality. in *Advancing the Framework for Assessing Causality of Health and Welfare Effects to Inform National Ambient Air Quality Standard Reviews* (National Academies Press (US), Washington (DC), 2022).

39. International Agency for Research on Cancer. Non-ionizing Radiation, Part 2: Radiofrequency Electromagnetic Fields. in *IARC monographs on the evaluation of carcinogenic risks to humans* vol. 102 (IARC, Lyon, France, 2013).

40. Yuan, A. E. & Shou, W. Data-driven causal analysis of observational biological time series. *eLife* **11**, e72518 (2022).

41. Arif, S. & MacNeil, M. A. Utilizing causal diagrams across quasi-experimental approaches. *Ecosphere* **13**, e4009 (2022).

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

43. Grace, J. B. & Irvine, K. M. Scientist's guide to developing explanatory statistical models using causal analysis principles. *Ecology* **101**, e02962 (2020).

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

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

46. Grace, J. B., Scheiner, S. M. & Schoolmaster, Jr., D. R. Structural equation modeling: building and evaluating causal models. in *Ecological Statistics* (eds. Fox, G. A., Negrete-Yankelevich, S. & Sosa, V. J.) 168–199 (Oxford University PressOxford, 2015). doi:10.1093/acprof:oso/9780199672547.003.0009.

47. Paul, W. L. A causal modelling approach to spatial and temporal confounding in environmental impact studies. *Environmetrics* **22**, 626–638 (2011).

48. Dee, L. E. *et al.* Clarifying the effect of biodiversity on productivity in natural ecosystems with longitudinal data and methods for causal inference. *Nat. Commun.* **14**, 2607 (2023).

49. Siegel, K. J., Larsen, L., Stephens, C., Stewart, W. & Butsic, V. Quantifying drivers of change in social-ecological systems: land management impacts wildfire probability in forests of the western US. *Reg. Environ. Change* **22**, 98 (2022).

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

51. Rubin, D. B. For objective causal inference, design trumps analysis. *Ann. Appl. Stat.* **2**, (2008).

52. Shadish, W. R., Cook, T. D. & Campbell, D. T. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. (Houghton Mifflin, Boston, 2001).

53. Glymour, C., Zhang, K. & Spirtes, P. Review of Causal Discovery Methods Based on Graphical Models. *Front. Genet.* **10**, 524 (2019).

54. Runge, J. *et al.* Inferring causation from time series in Earth system sciences. *Nat. Commun.* **10**, 2553 (2019).

55. Harnack, D., Laminski, E., Schünemann, M. & Pawelzik, K. R. Topological Causality in Dynamical Systems. *Phys. Rev. Lett.* **119**, 098301 (2017).

56. Shi, J., Chen, L. & Aihara, K. Embedding entropy: a nonlinear measure of dynamical causality. *J. R. Soc. Interface* **19**, 20210766 (2022).

57. Hernán, M. A. & Robins, J. M. *Causal Inference: What If*. (CRC Press, Boca Raton, 2025).

58. Pearl, J. Causal inference in statistics: An overview. *Stat. Surv.* **3**, (2009).

59. Robins, J. M. & Wasserman, L. On the Impossibility of Inferring Causation from Association without Background Knowledge. in *Computation, Causation, and Discovery* (eds. Cooper, G. F. & Glymour, C.) (The MIT Press, 1999). doi:10.7551/mitpress/2006.001.0001.

60. Granger, C. W. J. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica* **37**, 424 (1969).

61. Reichenbach, H. *The Direction of Time*. (University of California Press, Berkeley, 1956).

62. Kiiveri, H. T., Speed, T. P. & Carlin, J. B. Recursive causal models. *J. Aust. Math. Soc. Ser. Pure Math. Stat.* **36**, 30–52 (1984).

63. Scheines, R. *An Introduction to Causal Inference*. (1997).

64. Addicott, E. T., Fenichel, E. P., Bradford, M. A., Pinsky, M. L. & Wood, S. A. Toward an improved understanding of causation in the ecological sciences. *Front. Ecol. Environ.* **20**, 474–480 (2022).

65. Arif, S. & MacNeil, M. A. Predictive models aren't for causal inference. *Ecol. Lett.* **25**, 1741–1745 (2022).

66. Grace, J. B. An integrative paradigm for building causal knowledge. *Ecol. Monogr.* **94**, e1628 (2024).

67. Burnham, K. P. & Anderson, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. (Springer, New York, NY, 2010).

68. Greenland, S., Pearl, J. & Robins, J. M. Confounding and Collapsibility in Causal Inference. *Stat. Sci.* **14**, (1999).

69. Shpitser, I., VanderWeele, T. & Robins, J. M. On the validity of covariate adjustment for estimating causal effects. in *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence* 527–536 (AUAI Press, Arlington, Virginia, USA, 2010). doi:10.48550/arXiv.1203.3515.

70. Imbens, G. W. & Angrist, J. D. Identification and Estimation of Local Average Treatment Effects. *Econometrica* **62**, 467 (1994).

71. Rosenbaum, P. R. Choice as an Alternative to Control in Observational Studies. *Stat. Sci.* **14**, (1999).

72. Causal inference using regression on the treatment variable. in *Regression and Other Stories* (eds. Gelman, A., Hill, J. & Vehtari, A.) 363–382 (Cambridge University Press, Cambridge, 2020). doi:10.1017/9781139161879.020.

73. Wiik, E. *et al.* Mechanisms and impacts of an incentive-based conservation program with evidence from a randomized control trial. *Conserv. Biol.* **34**, 1076–1088 (2020).
74. Chickering, D. M. Optimal structure identification with greedy search. *J Mach Learn Res* **3**, 507–554 (2003).
75. Pearl, J. Causal diagrams for empirical research. *Biometrika* **82**, 669–688 (1995).
76. Kunicki, Z. J., Smith, M. L. & Murray, E. J. A Primer on Structural Equation Model Diagrams and Directed Acyclic Graphs: When and How to Use Each in Psychological and Epidemiological Research. *Adv. Methods Pract. Psychol. Sci.* **6**, 251524592311560 (2023).
77. Pearl, J. The causal foundations of structural equation modeling. in *Handbook of structural Equation Modeling* 68–91 (The Guilford Press, New York, NY, US, 2012).
78. Hernán, M. A., Wang, W. & Leaf, D. E. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* **328**, 2446 (2022).
79. Hernán, M. A., Dahabreh, I. J., Dickerman, B. A. & Swanson, S. A. The Target Trial Framework for Causal Inference From Observational Data: Why and When Is It Helpful? *Ann. Intern. Med.* (2025) doi:10.7326/ANNALS-24-01871.
80. Lundberg, I., Johnson, R. & Stewart, B. M. What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory. *Am. Sociol. Rev.* **86**, 532–565 (2021).
81. Correia, H. E., Dee, L. E. & Ferraro, P. J. Designing causal mediation analyses to quantify intermediary processes in ecology. *Biol. Rev.* brv.70011 (2025) doi:10.1111/brv.70011.
82. Spake, R. *et al.* Understanding ‘it depends’ in ecology: a guide to hypothesising, visualising and interpreting statistical interactions. *Biol. Rev.* **98**, 983–1002 (2023).
83. Paluš, M. From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics. *Contemp. Phys.* **48**, 307–348 (2007).

84. Dawid, P. The Decision-Theoretic Approach to Causal Inference. in *Wiley Series in Probability and Statistics* (eds. Berzuini, C., Dawid, P. & Bernardinelli, L.) 25–42 (Wiley, 2012). doi:10.1002/9781119945710.ch4.

85. Ibeling, D. & Icard, T. Comparing Causal Frameworks: Potential Outcomes, Structural Models, Graphs, and Abstractions. in *Advances in Neural Information Processing Systems* (eds. Oh, A. et al.) vol. 36 80130–80141 (Curran Associates, Inc., 2023).

86. Pearl, J. Graphical models, potential outcomes and causal inference: Comment on Linquist and Sobel. *NeuroImage* **58**, 770–771 (2011).

87. Weinberger, N. Comparing Rubin and Pearl's causal modelling frameworks: a commentary on Markus (2021). *Econ. Philos.* **39**, 485–493 (2023).

88. Bollen, K. A. & Pearl, J. Eight Myths About Causality and Structural Equation Models. in *Handbook of Causal Analysis for Social Research* (ed. Morgan, S. L.) 301–328 (Springer Netherlands, Dordrecht, 2013). doi:10.1007/978-94-007-6094-3_15.

89. Wiener, N. *Modern Mathematics for Engineers*. (McGraw-Hill, New York, 1956).

90. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Phys. Nonlinear Phenom.* **58**, 77–94 (1992).

91. Zhao, A. & Ding, P. Regression-based causal inference with factorial experiments: estimands, model specifications and design-based properties. *Biometrika* **109**, 799–815 (2022).

92. Imai, K., King, G. & Stuart, E. A. Misunderstandings Between Experimentalists and Observationalists about Causal Inference. *J. R. Stat. Soc. Ser. A Stat. Soc.* **171**, 481–502 (2008).

93. Bulbulia, J. A. Methods in causal inference. Part 4: confounding in experiments. *Evol. Hum. Sci.* **6**, e43 (2024).

94. Correia, H., Dee, L. & Ferraro, P. Quantifying intermediary processes in ecology using causal mediation analyses. Preprint at https://doi.org/10.32942/X2R628 (2024).

95. Rubin, D. B. Bayesian Inference for Causal Effects: The Role of Randomization. *Ann. Stat.* **6**, 34–58 (1978).

96. Smokorowski, K. E. & Randall, R. G. Cautions on using the Before-After-Control-Impact design in environmental effects monitoring programs. *FACETS* **2**, 212–232 (2017).

97. Gelman, A. & Hill, J. Causal inference using multilevel models. in *Data Analysis Using Regression and Multilevel/Hierarchical Models* 503–512 (Cambridge University Press, 2006). doi:10.1017/CBO9780511790942.

98. Cousineau, M., Verter, V., Murphy, S. A. & Pineau, J. Estimating causal effects with optimization-based methods: A review and empirical comparison. *Eur. J. Oper. Res.* **304**, 367–380 (2023).

99. Igelström, E. *et al.* Causal inference and effect estimation using observational data. *J. Epidemiol. Community Health* **76**, 960 (2022).

100. Huang, M. Y. Sensitivity analysis for the generalization of experimental results. *J. R. Stat. Soc. Ser. A Stat. Soc.* qnae012 (2024) doi:10.1093/jrsssa/qnae012.

101. Rosenbaum, P. R. Sensitivity to Hidden Bias. in *Observational Studies* 105–170 (Springer New York, New York, NY, 1995). doi:10.1007/978-1-4757-2443-1_5.

102. Shen, C., Li, X., Li, L. & Were, M. C. Sensitivity analysis for causal inference using inverse probability weighting. *Biom. J.* **53**, 822–837 (2011).

103. VanderWeele, T. J. & Arah, O. A. Bias Formulas for Sensitivity Analysis of Unmeasured Confounding for General Outcomes, Treatments, and Confounders. *Epidemiology* **22**, 42–52 (2011).

104. Yadlowsky, S., Namkoong, H., Basu, S., Duchi, J. & Tian, L. Bounds on the conditional and average treatment effect with unobserved confounding factors. *Ann. Stat.* **50**, (2022).

105. Sullivan, A. J. & VanderWeele, T. J. Bias and sensitivity analysis for unmeasured confounders in linear structural equation models. Preprint at https://doi.org/10.48550/ARXIV.2103.05775 (2021).

106. Rosenbaum, P. R. The Role of Known Effects in Observational Studies. *Biometrics* **45**, 557 (1989).

107. Rosenbaum, P. R. Sensitivity analyses informed by tests for bias in observational studies. *Biometrics* **79**, 475–487 (2023).

108. Rothman, K. J., Greenland, S. & Lash, T. L. Validity in Epidemiologic Studies. in *Modern epidemiology* 128–147 (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2008).

109. Greenland, S. & Lash, T. L. Bias Analysis. in *Modern epidemiology* 128–147 (Wolters Kluwer Health/Lippincott Williams & Wilkins, Philadelphia, 2008).

110. Bareinboim, E., Tian, J. & Pearl, J. Recovering from Selection Bias in Causal and Statistical Inference. in *Probabilistic and Causal Inference* (eds. Geffner, H., Dechter, R. & Halpern, J. Y.) 433–450 (ACM, New York, NY, USA, 2022). doi:10.1145/3501714.3501740.

111. Hernán, M. A., Hernández-Díaz, S. & Robins, J. M. A Structural Approach to Selection Bias: *Epidemiology* **15**, 615–625 (2004).

112. Imai, K. & Yamamoto, T. Causal Inference with Differential Measurement Error: Nonparametric Identification and Sensitivity Analysis. *Am. J. Polit. Sci.* **54**, 543–560 (2010).

113. Pearl, J. On measurement bias in causal inference. in *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence* 425–432 (AUAI Press, Arlington, Virginia, USA, 2010). doi:10.48550/arXiv.1203.3504.

114. Valeri, L. Measurement Error in Causal Inference. in *Handbook of measurement error* (eds. Yi, G. Y., Delaigle, A. & Gustafson, P.) (CRC Press, Boca Raton, 2022).

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

116. Kummerfeld, E., Williams, L. & Ma, S. Power analysis for causal discovery. *Int. J. Data Sci. Anal.* **17**, 289–304 (2024).

117. Li, J., Liu, L., Le, T. D. & Liu, J. Accurate data-driven prediction does not mean high reproducibility. *Nat. Mach. Intell.* **2**, 13–15 (2020).

118. Tredennick, A. T., Hooker, G., Ellner, S. P. & Adler, P. B. A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology* **102**, e03336 (2021).

Supplementary Information

1. Causal versus statistical assumptions

As noted in the main text, causal and statistical assumptions are both necessary components of deriving valid causal interpretations from observed relationships in data (Stone, 1993). Although the distinctions between these two types of assumptions are not always clear cut in casual research, we find it useful to distinguish them in the following way. Statistical assumptions are formal conditions about the data and model structure that must be satisfied for valid characterizations of relationships between variables from statistical analyses. These assumptions are often testable from data. Causal assumptions are additional conditions that are required to infer causation from statistically dependent relationships and are typically untestable (Hernán et al., 2019). By "untestable", we mean that these assumptions cannot be verified through statistical checks of data, even unlimited data, but instead must be justified using pre-existing knowledge.

Statistical assumptions commonly include assumptions about the probability distribution of random variables or observations, the specifications of relationships between variables, and conditions about data gathering or sampling (see Table S1). For example, they include assumptions about the functional relationships among variables (e.g., linearity, additivity) and about the probability distribution of random errors or observations (e.g., normality, independent and identically distributed random variables, constant variance). Statistical assumptions are encoded in the model structure; thus, they are often not described in applied data analyses.

Unlike causal assumptions (see Section 2 of the main text and Table S1 below), many of the statistical assumptions underlying empirical analyses in ecology are testable – that is, the assumptions can be verified from available data – even if they are often untested by researchers conducting the analyses. There are, however, untestable statistical assumptions that are also necessary for model-based inference, and these assumptions overlap with the causal assumptions described in Section 2 and in Table S1. For example, the basis of the Causal Sufficiency Assumption is a ubiquitous statistical assumption that requires correct specification of the explanatory variables in a model, specifically the inclusion of all confounding variables and the omission of all irrelevant variables. This assumption cannot be directly verified from data (i.e., the assumption is untestable) and must be supported by background knowledge about the system being modeled. Violations to the assumption that explanatory variables have been correctly specified can result in omitted variable bias, overfitting, and simultaneity bias that negatively impact interpretability and generalizability of results.

Other statistical considerations are also important for accurate conclusions from modeled data. These can include: ensuring sufficient statistical power to detect relationships between variables (Kimmel et al., 2023), decreasing measurement error or observational noise to better detect dependent relationships (Brown et al., 1990; Hyslop & Imbens, 2001), appropriately identifying and handling patterns of missingness (Little, 2021), and using robust statistics to accommodate a wider array of probability distributions and modest departures from model assumptions. While these considerations may not be viewed as statistical assumptions *per se*, they play an important role in determining the credibility of quantitative evidence about ecological phenomena.

The statistical and causal assumptions that are fundamental for making causal claims from ecological data are not tied to specific estimation approaches (e.g., frequentist versus Bayesian estimation). Many ecological studies emphasize the mode of estimation (mode of statistical inference) and overlook potential violations to causal and statistical assumptions that must be satisfied for valid inferences, but even minor violations can impair interpretability. Thus, extracting meaningful causal inferences from data in ecology requires both thoughtful construction of models and the scrutiny of the assumptions underlying these models (Burnham & Anderson, 2010).

Table S1. Common statistical and causal assumptions required for valid causal inference from data.

Statistical Assumptions	**Causal Assumptions**
Correct model specification	
- Model(s) include all relevant variables and no irrelevant variables. - Correctly specified functional forms of the relationships among variables (e.g., linearity, additivity).	- No unmeasured or omitted confounding variables (Causal Sufficiency Assumption). - Causal relationships follow the Causal Markov Assumption and Causal Faithfulness Assumption.
Random (unit-level) error conditions	
- Observations are independent and identically distributed (i.i.d.). - Random errors follow a specific probability distribution (e.g., Gaussian). - Random errors have constant variance (homoskedasticity). - Explanatory variables not correlated with random error. - Measurement error in explanatory variables is independent of the true values.	- A unit's treatment does not affect another unit's outcome (i.e., "no interference"). Related to the statistical i.i.d. assumption: i.i.d. can be violated by the presence of interference, which implies a lack of independence across units (see Zhang et al., 2023).
Data-specific criteria	
- For time-series: Stationarity (constant mean and variance over time). - No perfect multicollinearity among explanatory variables.	- No instantaneous causal effects ("no simultaneity"). - Every unit has a non-zero probability of receiving any level of treatment, conditional on covariates (i.e., "positivity" or "overlap").

2. Defining causal and non-causal research questions

Describing and quantifying ecological phenomena often requires a model, which is a mathematical description of how ecologists presume that variables of interest interact with each other. The form of the model is typically determined by the objective of the research question, which we divide into five categories: making causal claims, making associational claims, making predictions, summarizing data through descriptive statistics, and testing logical reasoning of hypotheses via simulations ("Define Research Question" in Figure S1).

Answering the first three types of questions requires statistical inference, which allows ecologists to learn information from observations using probability theory and use that information to make claims about relationships between variables, predict new information, and describe patterns in data (darker-shaded portion of the top box, to the left of the vertical dashed line in Figure S1). When sufficient data are not available or statistical inference is not suitable, mathematical modeling can be used to simulate hypothesized ecological interactions and check for logical fallacies (lighter-shaded portion of the top box, to the right of the vertical dashed line in Figure S1). Associational analyses, predictive models, or simulation-based approaches can also be useful for deriving knowledge that can contribute to future causal research questions (Figure S1 and Figure 3 in main text).

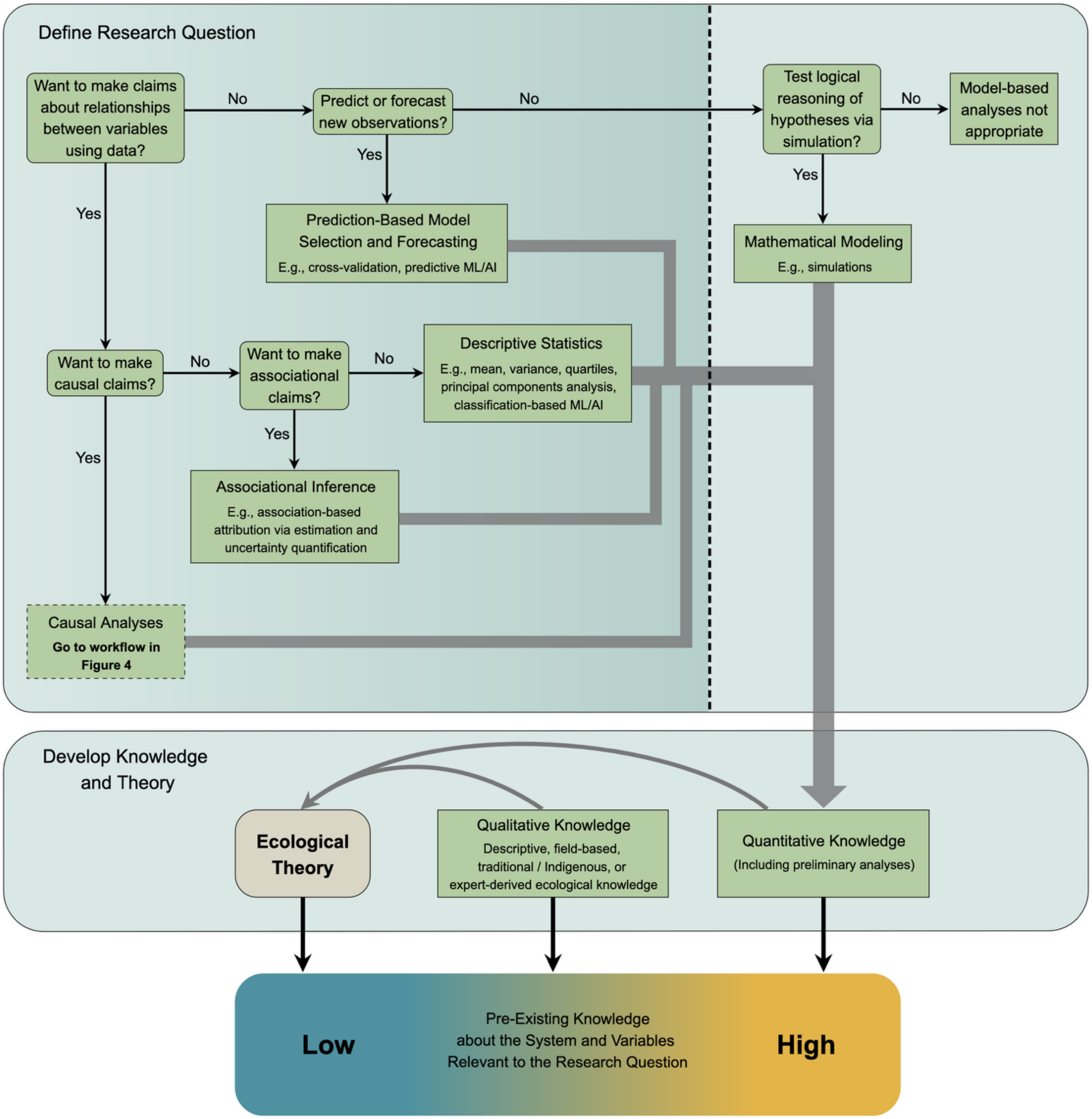


Figure S1. Decision tree for determining the type of analysis most appropriate for the research goal. Prediction-based model selection and forecasting, descriptive statistics, associational inference, and causal analyses use statistical inference, which separates them from approaches like simulation-based mathematical modeling. That separation is represented by the vertical dashed line that separates lighter and darker shaded regions of the top box. The bottom gradient box is also represented in the first box in the workflow of Fig. 3.

A. **<u>Using data to derive claims about relationships between variables</u>**

When causal interpretations of statistical models are desired, causal methodologies, a subset of statistical inference, allow ecologists to make causal claims about relationships between variables from data. However, as we make clear in Section S4, using statistical inference to make causal claims requires that the experimental or nonexperimental data collection and analyses satisfy many conditions (i.e., assumptions). We provide more details on the tasks that can be accomplished through causal studies and specific methods in Section 6.

If causal claims are not desired, ecologists can draw on classical tools from statistical inference (Efron & Hastie, 2016; Holland, 1986; Nakagawa & Cuthill, 2007). These associational studies can also shape the formulation of causal research questions for subsequent studies. Many research questions have causal goals, but researchers will usually cast these questions as associational due to perceived limitations of statistical methodologies or concerns about misuse of their findings (Hernán, 2018; Jones & Schooling, 2018; Kezios & Hayes-Larson, 2018). Researchers also commonly draw causal-sounding conclusions (e.g., using terms like "drives" or "leads to") from predictive or associational analyses (Haber et al., 2022; Han & Guyatt, 2020; Sargeant et al., 2022; Singer, 2022), thus overstating the evidence of causality by implying that the underlying causes have been properly isolated from unrelated or spurious associations (i.e., that alternative explanations for the observed associations have been ruled out). This tendency is now heavily ingrained in the scientific culture of many fields, but we strongly encourage ecologists to principally consider the goals behind their research questions before considering the methods that may be taken to achieve those goals.

Alternatively, ecologists may instead wish to probe data for general patterns among variables by using statistical inference to explore or summarize the data ("Descriptive Statistics" in Figure S1). Approaches used to describe data are often included in studies aiming to make causal or associational claims, but descriptive statistics are not the primary source of evidence for making such claims.

B. **<u>Not deriving claims about relationships among variables from data</u>**

At times, ecologists may want to predict unobserved outcomes from new input data by using training data to optimize parameter estimation such that a set of input features predict output values that most closely match observed data output values in verification data ("Prediction-Based Model Selection and Forecasting" in Figure S1). Predictive studies rely on procedures that emphasize model evaluation and selection through predictive performance, including model averaging that derives inferences from several plausible models (i.e., multi-model inference; Burnham & Anderson, 2010). Results from models selected for high prediction accuracy are often believed to produce more meaningful parameter estimates for inference than models with low prediction accuracy (Harrison et al., 2018), which has spurred the popularity of machine learning approaches touted to provide "data-driven" understandings of complex ecological processes (Christin et al., 2019; Olden et al., 2008). However, prediction models merely need to capture the rudimentary patterns and relationships in the data to produce highly accurate predictions. Thus, models with high prediction accuracy often do not accurately represent the true underlying causal processes of the ecological system from which the data were generated, and thus they are usually not appropriate for making associational or causal claims (Addicott et al., 2022; J. Li et al., 2020; Tredennick et al., 2021).

In other studies, ecologists may wish to simulate hypothesized relationships between variables using mathematical “proof-of-concept” models (sometimes called “mechanistic models”), which play an integral role in translating ecological theories and hypotheses into mathematical language (e.g., the Lokta-Volterra model; Baker et al., 2018; Marquet et al., 2014; Servedio et al., 2014). Numerical analysis of mathematical models allows ecologists to explore and refine hypotheses, examine a model’s internal consistency, and assess how well the model represents theoretical or empirical relationships. Additionally, data collected from experiments and field observations can be used to constrain model parameter values or to compare model output to naturally occurring patterns (Caldararu et al., 2023; Evans et al., 2013; Levins, 1966; Luo et al., 2011; Tredennick et al., 2021), but statistical inference is not the goal of such models.

Although mathematical models, predictive models, associational studies, and descriptive statistics can all contribute to quantitative ecological knowledge and pre-existing knowledge for developing causal research questions (“Develop Knowledge and Theory” in Figure S1), current methodologies for making causal claims from data require principles of probability theory and statistical inference to be combined with the rigorous conditions for experimental and observational data collection and analysis defined by causal assumptions. Some researchers have argued that, under certain conditions, predictive models may also contribute to refining or corroborating causal hypotheses when results from predictive studies align with theoretical expectations (Nichols & Cooch, 2025). While consistent findings from predictive models may contribute to pre-existing or “mechanistic” ecological knowledge (Grace, 2024), particularly when supported by ecological theory and expert understanding, predictive performance alone is insufficient to justify causal claims.

3. Formally summarizing pre-existing knowledge

Establishing the conditions for making valid causal claims from data is achieved by satisfying the causal assumptions that permit us to detect and quantify causal relationships using statistical dependence. A central task for ecologists interested in causal relationships is to carefully consider the study design, the potential variables to be included or not included in the model, and the data collection procedures. One of the fundamental conditions for valid statistical inference and interpretability of results is that the model correctly specifies the true underlying process from which the data were generated. Developing such a correctly specified model requires pre-existing knowledge to identify potentially causative factors and potential pathways of influence through other interacting variables.

The assumptions required for causal analyses highlight how causal tasks (i.e., causal discovery and causal inference) differ from non-causal tasks (e.g., prediction or association). Unlike non-causal analyses, causal tasks depend on pre-existing knowledge to construct and justify models for causal tasks (particularly for causal inference) that satisfy these untestable causal assumptions, rather than selecting the “best” model among several plausible models based on fit metrics that evaluate prediction performance. Even causal discovery is fine-tuned with pre-existing knowledge, guiding algorithms to retain specific plausible relationships specified by the user’s pre-existing knowledge, and its results must be validated through further research.

Proper model specification is crucial for valid causal conclusions (Burnham & Anderson, 2010), thus more attention must be invested in the process of designing studies and building models using pre-existing knowledge to make causal claims from experimental and observational ecological studies. To formalizing pre-existing knowledge in causal analysis, researchers may use two widely used tools: directed acyclic graphs (DAGs) and thought experiments based on ideal randomized controlled trials (RCTs). These tools help define causal relationships and identify confounders that must be addressed to satisfy causal assumptions before any data are analyzed. Table S2 provides a guide to accessible and foundational references for learning how to apply these tools.

Table S2. Key concepts and accessible references for creating and applying causal DAGs and thought experiments of hypothetical ideal RCTs for summarizing pre-existing knowledge.

Concept	**Suggested Readings**
Basics of causal DAGs – What they are, variables to include, why they help in confounder identification	Bulbulia, 2024a; Greenland et al., 1999a; Laubach et al., 2021; Shrier & Platt, 2008
Drawing DAGs in practice – User-friendly guidelines for causal DAGs in experimental and observational settings	Arif & MacNeil, 2022; Textor et al., 2011
Using thought experiments of hypothetical ideal RCTs (i.e., "target trials") – How to use thought experiments to simulate an ideal experiment to find confounders	Greenland, 2003; Hernán et al., 2022, 2025; Hernán & Robins, 2025, pp. 37–40; Morgan & Winship, 2015 (Ch. 1); Rubin, 1974
Distinguishing confounders vs. colliders – Ensuring we do not control for the wrong variables	Arif & Massey, 2023; Bulbulia, 2024a; Greenland, 2003

4. Causal assumptions translated through causal frameworks

Causal inference and causal discovery both rely on untestable assumptions that allow researchers to interpret statistical patterns as evidence of causation. Three foundational assumptions are shared across major causal frameworks: Causal Sufficiency, Causal Markov, and Causal Faithfulness. However, the way these assumptions are expressed, along with the specific terminology and extensions they involve, varies across causal frameworks. In this section, we show how different frameworks formalize these assumptions and illustrate the conceptual bridges between them.

We focus on three widely used causal frameworks: the structural causal model (SCM) framework (Pearl, 2009), the potential outcomes (PO) framework (Rubin, 1974), and the dynamical systems causality (DC) framework (Harnack et al., 2017; J. Shi et al., 2022). A fourth not covered here – the decision-theoretic framework (Dawid, 2000, 2012) – also shares overlapping assumptions. Each framework uses its own notation and formalism to express the causal assumptions and structure causal reasoning. The PO and SCM frameworks are most common for causal inference, while the SCM and DC frameworks are commonly used for causal discovery.

Theoretical work has established formal correspondences among several major causal frameworks. The PO and SCM frameworks have been shown to be theoretically equivalent (Imbens, 2020; Pearl, 2009), with modern formalizations demonstrating that every Rubin Causal Model from the PO framework can be represented as an abstraction of an SCM (Ibeling & Icard, 2023). A measure-theoretic approach has also been proposed to generalize aspects of SCM and PO frameworks and address challenges like cycles, latent variables, and stochastic processes (Park et al., 2023). Causal properties of the decision-theoretic framework can be expressed through extended conditional independence assertions, aligning with the PO and SCM frameworks under specific conditions (Dawid, 2021, 2024; Pearl, 2022). Connections between the SCM and DC frameworks have also been developed, including approaches that extend SCMs to time-dependent settings and systems with feedback loops (Bongers et al., 2018, 2021) and approaches that link Granger causality (a DC-based approach) to SCMs by representing interventions and dynamic feedback processes (White et al., 2011; White & Chalak, 2009). Methods like transfer entropy, which is used in DC-based analyses, have similarly been related back to conditional independence structures central to SCMs (Runge et al., 2012). Commentaries have also highlighted key conceptual differences and areas of overlap between the PO, SCM, and DC frameworks (Lechner, 2010; Markus, 2021). While recent reviews (e.g., Vonk et al., 2023; Yuan & Shou, 2022) have discussed assumptions in causal discovery and causal inference broadly, here we systematically map how core causal assumptions translate across SCM, PO, and DC frameworks for causal inference and causal discovery.

In Box S1, we map the assumptions used for quantifying the average causal effect of X on Y in causal inference via the PO and SCM frameworks onto the three basic causal assumptions. We also summarize two additional assumptions widely used in practice for causal inference. Together, these assumptions allow us to quantify causal effects without bias. For full details of PO assumptions for causal inference, see Hernán & Robins, 2025; for full details of SCM assumptions for causal inference, see Pearl, 2009 or Pearl, 2010.

For causal inference, the inclusion of all relevant confounding variables is necessary to satisfy the causal sufficiency assumption. However, this does not always require directly measuring every confounder. In both frameworks, design-based approaches and statistical techniques can be used to account for unmeasured confounding under certain conditions. Some frameworks, such as SCM, allow for adjustment using variables that are not direct confounders (e.g., descendants of common causes), provided that colliders and other bias-inducing paths are avoided that would otherwise introduce non-causal statistical dependencies (Pearl, 1995; Rohrer, 2018).

In Box S2, we map the assumptions used for causal discovery via the SCM and DC frameworks onto the three basic causal assumptions. We also summarize three additional assumptions commonly required in practice for causal discovery. For full details of SCM assumptions for causal discovery, see Glymour et al., 2019; for full details of DC assumptions for causal discovery, see J. Shi et al., 2022. For relationships between SCM and DC assumptions in causal discovery, see Runge, 2018.

For causal discovery, causal assumptions are used to ensure the reliability of the causal structure inferred from data. SCM-based algorithms primarily rely on the Causal Markov and Causal Faithfulness assumptions, often alongside Causal Sufficiency and additional assumptions like acyclicity and i.i.d. sampling (Glymour et al., 2019). These assumptions can often be relaxed in more advanced approaches. DC-based algorithms often implicitly rely on the causal sufficiency assumption (Paluš, 2007; Runge, 2018), where all common causes are assumed to be measured or contained within the information of the measured variables (i.e., there are no unmeasured confounders, a.k.a., “hidden common causes”), and usually require separability, which is a consequence of the causal faithfulness assumption (Eichler, 2013; Peters et al., 2017; Runge, Nowack, et al., 2019; Spirtes et al., 2000). However, some DC-based causal discovery methods have been developed for non-separable systems (e.g., J. Shi et al., 2022) and for detecting and handling the presence of unmeasured confounders (e.g., Cai et al., 2023).

Together, Boxes S1 and S2 provide a unique synthesis of how the three foundational causal assumptions are formalized and applied across diverse causal frameworks. By explicitly mapping the assumptions of each framework to these shared foundations, the Boxes serve as practical tools for clarifying how these assumptions support valid causal claims across different, and sometimes seemingly disparate, frameworks and causal tasks, thereby clarifying both their common foundations and distinct assumptions.

Box S1. Assumptions for causal inference

Choice of framework

Potential Outcomes (PO) Framework	Structural Causal Models (SCM) Framework
Useful for those familiar with randomized experimental designs. Emphasizes addressing non-causal dependencies (confounding) by leveraging specific experimental designs or imitating such scenarios via statistical techniques.	Useful for those who think about multiple causes jointly ("all-cause models"). Emphasizes defining the minimal set of conditions under which causal effects can be identified and estimated.

Causal assumptions

A1. Causal Sufficiency: All relevant confounders are measured (i.e., no unmeasured common causes).

Potential Outcomes (PO) Framework	Structural Causal Models (SCM) Framework
Terminology: "No unmeasured confounders", "ignorability", or "exchangeability"†. **Key Idea**: Once we adjust for all relevant confounders, the probability of receiving any given exposure level does not depend on any common causes. Therefore, we must measure and adjust for (i.e., include in the model) all variables that influence both the exposure and the outcome (and any intermediary variables; see Correia et al., 2025). Also requires *positivity* – individual units are equally likely to be exposed to a specific value of a causal factor (see below). **References**: Hernán & Robins, 2025; Morgan & Winship, 2015; Rosenbaum & Rubin, 1983	**Terminology**: "All front-door and back-door paths blocked", or "no omitted common causes in the causal DAG". **Key Idea**: All confounders identified by the front-door and back-door criteria (or additional criteria; see Maathuis & Colombo, 2015 and Shpitser & Pearl, 2008) are measured and adjusted for (e.g., included in the model). Also required *consistency* (the statistical property) – with infinite data, the estimated graph will converge to the true causal graph (see Pearl, 2009; Spirtes et al., 2000). **References**: Greenland et al., 1999b; Pearl, 2009
Terminology: "No interference", "no spillover", "no unit-to-unit causation", or "no interactions between units" (see Cox, 1958); part of Stable Unit Treatment Value Assumption (SUTVA) (see Rubin, 1980). **Key Idea:** One unit's exposure does not affect another unit's outcome. Real-world systems often violate this assumption, requiring more complex methods (see Hudgens & Halloran 2008). **References:** Hudgens & Halloran, 2008; Rubin, 1978, 1980	**Terminology:** No spillover is implicitly assumed by SCM notation and causal DAGs. **Key Idea:** In a causal DAG, there are no edges from one unit's exposure to another unit's outcome, i.e., each unit's outcome depends only on its own exposure. Systems that violate this assumption require multi-unit DAGs or specialized methods (see Pearl, 2009). Part of assumption that *units are independent and identically distributed* (i.i.d.) assumption; see Zhang et al., 2023. **References:** Pearl, 2009; Spirtes et al., 2000

A2. Causal Markov Condition: In a system with no cycles or feedback loops, any dependence between two variables that do not directly affect each other must come from a common cause influencing both. Once that common cause is accounted for, the two variables should no longer be dependent.

Potential Outcomes (PO) Framework	**Structural Causal Models (SCM) Framework**
Terminology: "No feedback" or "no cyclic causation" (i.e., simultaneous causation) are implied by the potential outcome notation: the outcome $Y(a)$ is measured after an exposure $A = a$. The exposure and outcome are conditionally independent once we account for all confounding variables. **Key Idea**: Once we measure and adjust for any shared causes, any dependence between two variables that do not share a direct causal relationship should no longer remain. This also requires that the cause precede the effect, ruling out simultaneity. **References**: Hernán & Robins, 2025; Morgan & Winship, 2015; Rubin, 1978	**Terminology**: By definition, causal DAGs are acyclic; therefore, feedback loops or bidirectional arrows (simultaneous causation) are disallowed. Sometimes referred to as *factorization* or the *local Markov property* – each node is conditionally independent of its non-descendants, given its parents. **Key Idea**: Once we condition on the parents (common causes), the dependence between two variables that do not directly affect each other is "blocked". Since arrows in causal DAGs flow in one direction, it is assumed there is no cyclic causation. **References**: Pearl, 2009; Spirtes et al., 2000

A3. Causal Faithfulness: If two variables are statistically independent even after adjusting for confounders, then there is no causal relationship between those variables.

Potential Outcomes (PO) Framework	**Structural Causal Models (SCM) Framework**
Terminology: Implicitly assumed that any true causal effect would manifest as a dependence after all confounders are adjusted for. **Key Idea**: If two variables remain independent after controlling for all relevant confounders, we assume it's not due to a coincidence but instead conclude there is no causal relationship. **References**: Hernán & Robins, 2025; Morgan & Winship, 2015	**Terminology**: Explicitly called *faithfulness* or *stability*, in which the causal DAG encodes all conditional independences. If two variables are independent, there exists no causal path (i.e., no causal relationship) between those variables in the causal DAG. **Key Idea**: If two variables remain independent after conditioning on the variables that block any back-door paths in a causal DAG, we assume this reflects a genuine absence of a causal relationship. **References**: Pearl, 2009; Spirtes et al., 2000; Wermuth & Lauritzen, 1990

Additional assumptions

B1. The exposure is well-defined (i.e., no multiple versions of the treatment, such as different strains of a disease being categorized as a single exposure). That is, there must be no ambiguity about what the cause or exposure is.

Potential Outcomes (PO) Framework	**Structural Causal Models (SCM) Framework**
Terminology: "Causal consistency" (not the same as the statistical property of consistency) or "well-defined treatment"[†]; part of SUTVA (Rubin 1978, 1980). **Key Idea**: No ambiguous exposure or no multiple versions of a single cause. A cause or exposure must be identically represented across all units. **References**: Hernán & Robins, 2025; Rubin, 1978, 1980	**Terminology**: A well-defined or unambiguous exposure is implied by the causal DAGs – the exposure must be unambiguous when declared as node in the causal DAG. **Key Idea**: The causal DAG must represent exactly one well-specified cause or exposure. If we can declare the cause or exposure as one node, we are assuming that it is well-defined. **References**: Pearl, 2009; Spirtes et al., 2000

262

B2. Among units that share the same values for the confounders, there must be some that are exposed and some that are not. In other words, the confounders must not perfectly predict the probability of exposure.[*]

Potential Outcomes (PO) Framework	**Structural Causal Models (SCM) Framework**
Terminology: "Positivity", "overlap", or "common support"[†] **Key Idea**: For any given combination of confounder values, there must be a nonzero chance of receiving each exposure level. **References**: Hernán & Robins, 2025; Morgan & Winship, 2015; Rosenbaum & Rubin, 1983	**Terminology**: All exposure levels are sufficiently represented in the data is implied by representing the exposure as a node in the causal DAG. **Key Idea**: Even if the causal DAG is correctly specified, the data must exhibit variation in exposure for every configuration of confounders. **References**: Pearl, 2009; Spirtes et al., 2000

263

[†]Causal consistency, positivity, and exchangeability make up the 'identifiability conditions' for causal effects. These conditions hold under idealized randomized experiments (see Kimmel et al., 2021).

[*]Positivity is a statistical assumption rather than a purely causal assumption. It requires that our data exhibit variation in exposures across all relevant confounders. See Hernán & Robins, 2025; Morgan & Winship, 2015; Rosenbaum & Rubin, 1983.

264

Box S2. Assumptions for causal discovery

Choice of framework

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Useful for those who think about evolving states of systems over time; focuses on identifying causal relationships for dynamic or complex systems where long time series of observations are available, often under challenging scenarios (e.g., non-separability, high-dimensional nonlinearity).	Useful for those who think about multiple causes jointly (“all-cause models”). Emphasizes defining the minimal set of conditions under which causal effects can be identified and estimated.

Causal assumptions

A1. Causal Sufficiency: All relevant confounders are measured (i.e., no unmeasured common causes).

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: “All variables that drive the system are embedded in the reconstructed state space”, “no missing drivers”, or “intrinsic noise is not attributable to external disturbances or measurement errors”. **Key Idea**: Implicitly assumes the measured variables capture the main dynamic influences. If crucial state variables are omitted, apparent causal links can be spurious. **References**: Ding & Toulis, 2018; Harnack et al., 2017; Orava, 1973; Sun et al., 2015	**Terminology:** “All relevant variables included”, or “no omitted common causes”. **Key Idea**: Discovery algorithms (e.g., PC, FCI) typically assume all major confounders are measured or the algorithm is adjusted to detect them. Also required *consistency* (the statistical property) – with infinite data, the estimated graph will converge to the true causal graph. **References**: Glymour et al., 2019; Peters et al., 2017; Spirtes et al., 2000
Terminology: The observed time series fully capture the dynamics of the unit, with no external influences (i.e., no inter-unit interference). **Key Idea:** The dynamics of each unit are self-contained; the time series used for discovery must reflect the complete internal state of the system. If significant spillover exists, the predictive relationships used to infer causality may be confounded by external influences. **References:** Harnack et al., 2017; Orava, 1973	**Terminology**: “No cross-unit edges” or “independence of units” in causal DAGs. **Key Idea: E**ach unit is independent – one unit’s exposure does not affect another unit’s outcome. Part of the i.i.d. assumption – units are independent and identically distributed (see Zhang et al. 2023). **References**: Glymour et al., 2019; Spirtes et al., 2000

A2. Causal Markov Condition: In a system with no cycles or feedback loops, any dependence between two variables that do not directly affect each other must come from a common cause influencing both. Once that common cause is accounted for, the two variables should no longer be dependent.

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: If two system components do not interact (directly or indirectly), their time series become conditionally independent (or uncorrelated) after controlling for the relevant state variables. **Key Idea**: In time-lagged embedding, if variable A does not help predict B once the relevant lags of B (and possibly other variables) are included, we treat them as causally disconnected. This also requires that the cause precede the outcome, ruling out simultaneity and cyclic causation (see below). **References**: Runge, Bathiany, et al., 2019; Sun et al., 2015	**Terminology**: Sometimes referred to as *factorization* or the *local Markov property* – each variable is conditionally independent of its confounders given its direct causes. **Key Idea**: If two variables are conditionally independent given some conditioning set in the data, they are not connected by any path in the DAG (or are d-separated). Implicitly assumes there is no simultaneity or cyclic causation (see below). **References**: Glymour et al., 2019; Peters et al., 2017; Spirtes et al., 2000

A3. Causal Faithfulness: If two variables are statistically independent even after adjusting for confounders, then there is no causal relationship between those variables.

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: Referred to as *separability* – the influence of measured confounding variables can be eliminated from the information contained in the effect variable's temporal trajectory without changing the direct relationship between the cause and effect; thus, an observed temporal dependence implies the presence of a causal relationship. **Key Idea**: If two variables remain independent after controlling for all relevant confounders, we assume it's not due to a coincidence but instead conclude there is no causal relationship. **References**: Paluš et al., 2018; Runge, Bathiany, et al., 2019; Schreiber, 2000; Sun et al., 2015	**Terminology**: Explicitly called *faithfulness* or *stability*, in which the causal DAG encodes all conditional independences. If two variables are statistically independent, there exists no causal path (i.e., no causal relationship) between those variables in the causal DAG. **Key Idea**: If two variables remain independent after conditioning on the confounders, we assume this reflects a genuine absence of a causal relationship. **References**: Glymour et al., 2019; Peters et al., 2017; Spirtes et al., 2000

Additional assumptions

B1. Cause precedes effect; no simultaneity and no feedback loops.

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: "Temporal ordering", or "one variable's state at the current time t influences the other's state at future time $t + \ell$". **Key Idea**: The future state of a system is conditionally independent of its past states, given its present state (i.e., cause precede effects in time). **References**: Ding & Toulis, 2018; Paluš et al., 2018	**Terminology**: "Acyclic", "no bidirectional edges", or "no feedback loops" implied in the causal DAG. **Key Idea**: Assumes no feedback loops or simultaneous causation exists in the data, since resultant causal DAGs are acyclic. **References**: Peters et al., 2017; Spirtes et al., 2000

B2. Stationarity – the system's behavior doesn't change dramatically over time (i.e., overall distributional patterns such as mean and variance of causes and outcomes remain relatively constant over time).

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: The system's behavior does not change over time. **Key Idea**: Causal relationships remain consistent over time (dependencies should not fundamentally change or vanish). Also requires *ergodicity* – statistical properties (e.g., mean and variance) calculated from time series samples through the ergodic theorem do not change substantially over time. **References**: Harnack et al., 2017; McGoff et al., 2012; J. Shi et al., 2022	**Terminology**: The conditional independencies among variables are consistent over time. **Key Idea**: The influence of a variable's state at a previous time $t - \ell$ on its state at the current time t remains consistent throughout the time series when controlling for the rest of the system's state at the present time t. **References**: McGoff et al., 2012; Peters et al., 2017; Runge, Bathiany, et al., 2019

B3. Sufficient variability within variables in the system so that differences in exposure and outcome can be reliably detected.

Dynamical Systems Causality (DC) Framework	Structural Causal Models (SCM) Framework
Terminology: Time series provide a faithful representation of the system's dynamics. Additionally, many approaches require that states of the system (e.g., from time series data) can be represented as a low-dimensional attractive manifold. **Key Idea**: There must be enough dynamic variation in the observed data to reveal causal influences, and the measured variables must adequately reflect the system's underlying states. **References**: Barański et al., 2020; Deyle & Sugihara, 2011; J. Shi et al., 2022; Takens, 1981	**Terminology**: "Positivity" and "consistency". **Key Idea**: Each variable (cause or outcome) exhibits enough variation to detect dependence (akin to *positivity* in causal inference). Also, each variable must be well-defined, so that distinct real-world processes aren't lumped under one label (*consistency*). **References**: Glymour et al., 2019; Peters et al., 2017

286

5. Core concepts for each causal framework

While assumptions define the foundation for making valid causal claims, each causal framework also introduces a range of concepts and tools that shape how researchers think about variables, causal relationships, and estimation. To help readers navigate these differences, we provide three tables (Tables S3–S5), one for each framework (PO, SCM, and DC, respectively), that highlight foundational concepts across the frameworks, along with seminal and accessible sources for further reading. These tables are designed as navigational tools for readers seeking intuitive or technical entry points into each framework, such as ignorability and causal estimands in the PO framework, d-separation and *do*-calculus in the SCM framework, and state space reconstruction and separability in the DC framework. Familiarity with these concepts is important for understanding how causal inference and causal discovery are framed and implemented within each framework's structure. These frameworks are not mutually exclusive and can be complementary depending on the causal task and data characteristics. Researchers should familiarize themselves with each to determine which assumptions and tools best align with their research goals.

Table S3. Key concepts and recommended references for understanding the potential outcomes (PO) framework.

Concept	**Suggested Readings**
Fundamentals of the PO framework	Holland, 1986; Rubin, 2005; Sobel, 2009
Stable Unit Treatment Value Assumption (SUTVA)	Sobel, 2006; VanderWeele & Hernán, 2013
Ignorability Assumption (Unconfoundedness)	Imbens, 2004; Rosenbaum & Rubin, 1983
Positivity Assumption (Overlap Condition)	Petersen et al., 2012; Westreich & Cole, 2010
Confounding variables to control for in analyses	Gelman et al., 2020; VanderWeele, 2019; VanderWeele & Shpitser, 2011
Causal estimands: average treatment effect (ATE) and others	Heiss, 2024; Imbens, 2004; Imbens & Angrist, 1994; Lipkovich et al., 2020; Wooldridge, 2010 (Ch. 21)
Multiple versions of treatment and interference	Hudgens & Halloran, 2008; Tchetgen Tchetgen & VanderWeele, 2012; VanderWeele & Hernán, 2013

Table S4. Key concepts and recommended references for understanding the structural causal models (SCM) framework.

Concept	Suggested Readings
Fundamentals of the SCM framework	Burnett & Blackwell, 2024; Cheng et al., 2024; Petersen & van der Laan, 2014; Scheines, 1997
Confounding variables to control for in analyses (d-separation; Back-door and Front-door Criteria)	Arif & Massey, 2023; Bulbulia, 2024a; Elwert, 2013; Greenland, 2003; Morgan & Winship, 2015 (Ch. 4 & 10); Pearl, 2010
Graphical rules for causal identification in graphs (*do*-calculus)	Hayduk et al., 2003; Pearl, 2009 (Ch. 1 & 11); Shpitser & Pearl, 2008; Tian & Pearl, 2002
Total and path-specific causal effects	Bulbulia, 2024b; Pearl, 2009 (Ch. 3, 4, 7); VanderWeele, 2015d
Model equivalence and Markov equivalence classes	Andersson et al., 1997; Pearl, 2009 (Ch. 5)
Causal graphs with unmeasured/latent variables	Pearl, 2009 (Ch. 12) ; Richardson & Spirtes, 2002

Table S5. Key concepts and recommended references for understanding the dynamical systems causality (DC) framework.

Concept	Suggested Readings
Fundamentals of the DC framework	Deyle & Sugihara, 2011; Harnack et al., 2017; Runge, 2018; J. Shi et al., 2022; Yuan & Shou, 2022
State space reconstruction (SSR) and attractor manifolds	Cummins et al., 2015; Sauer et al., 1991; Takens, 1981
Causality via predictability	Paluš, 2007; Runge, 2018; Sugihara et al., 2012
Transfer entropy and information-theoretic causality	Schreiber, 2000; Sun et al., 2015; Sun & Bollt, 2014
Separability and causal faithfulness	Eichler, 2013; Peters et al., 2017; Runge, Nowack, et al., 2019
Confounding and hidden variables in time series	De Brouwer et al., 2021; Eichler, 2013; Sun & Bollt, 2014
Limitations in stochastic or weakly coupled systems	Cobey & Baskerville, 2016; McCracken & Weigel, 2014

6. Study designs and algorithms for causal analyses

Selecting a study design or algorithm is a critical step in implementing a causal analysis. Different designs and algorithms offer structured ways to satisfy or relax the untestable causal assumptions and must be chosen in light of the causal task, available data, and pre-existing knowledge. Some approaches are grounded in experimental control, while others rely on statistical adjustments or algorithmic structure learning to address confounding and identify causal relationships.

To help readers explore available options, we provide a series of tables that group study designs and algorithms according to the type of causal task (inference or discovery) and whether they address measured or unmeasured confounding. Table S6 summarizes study designs for causal inference, including experimental designs, observational designs for measured confounders, and observational designs for unmeasured confounders. Table S7 summarizes algorithms for causal discovery, grouped by the causal framework and assumptions each algorithm relies on. These tables are intended to serve as a reference for researchers selecting and comparing appropriate strategies for their study goals, system knowledge, and data constraints. For additional guidance on the selection of specific causal inference study designs and causal discovery algorithms, see the flow chart in Figure 2 in Runge et al., 2023.

Causal inference requires that all confounders be addressed (see Box S1), but this does not necessarily mean every confounder must be explicitly included in a model. Instead, confounding is typically handled using a combination of design-based approaches: directly controlling for measured confounders and employing statistical designs that reduce bias from unmeasured confounders (e.g., experimental randomization or statistical approaches that mimic randomization).

If significant pre-existing knowledge is available and the goal is to obtain system-level understanding (i.e., to model the effects of all causes of an outcome), then SCM-based adjustment methods (e.g., Front-door and Back-door Criteria; see Pearl, 2009 and Arif & MacNeil, 2022) or structural equation modeling (SEM) may be appropriate approaches. While SCM-based adjustment methods typically target specific causal effects, SEM is often used to model entire systems of causal relationships simultaneously. However, this comes with tradeoffs: SEM requires more restrictive assumptions in order to support system-level inferences. These tradeoffs underscore the need to carefully align the use of SEM with the level of pre-existing knowledge and assumptions that can be plausibly justified for the ecological system under study (Grace, 2024; Pearl, 2012; Shipley, 2016). In cases where unobserved variables are present, acyclic directed mixed graphs (ADMGs) can represent the same set of conditional independencies as a DAG. ADMGs also allow for bidirectional (i.e., double-headed) arrows, enabling representation of latent confounding. These graphs rely on an extension of Pearl's *d*-separation criterion, called *m*-separation (for details, see Richardson, 2003 and Drton & Richardson, 2004).

While both SEM and SCM approaches rely on a causal graph to represent assumptions, they differ in how those assumptions are used. SCMs (Pearl, 2009) use the graph to derive conditions under which causal effects can be identified from data, often targeting specific effects of interest via tools such as the Back-door or Front-door criteria. In an SCM approach, the causal graph is used to ask, “Given this DAG, can I even estimate the causal effect of X on Y from observed data, and if so, how?” In contrast, SEMs as used in ecology (Grace et al., 2015; Shipley, 2016) typically assume the full system of causal relationships is known, and use the graph to specify a system of structural equations whose fit can be statistically tested. That is, for SEMs, the causal graph is used to ask, “Assuming this DAG is correct, do the observed data support it, and can I fit a model to estimate the effects I care about?” SEM-based causal inference does not involve formal identification theorems, and estimation is typically linear, even when nonlinear terms are used. Thus, SEMs rely more heavily on model specification and goodness-of-fit, whereas SCMs prioritize identifiability of causal effects under minimal assumptions (Pearl, 1998). SEMs can yield unbiased causal effect estimates if the model includes all relevant confounders and is correctly specified; however, unlike SCM-based methods, they do not provide formal identification criteria to assess whether these conditions are met (Bollen & Pearl, 2013; Markus, 2010; Wang & Sobel, 2013). This distinction highlights that while both approaches can be used for causal modeling, they support different inferential goals and require different standards of justification.

Table S6. Study designs for causal inference, grouped by category. Each study design includes a brief description, key references (including applications in ecology, where available), and links to available software and code. The resources and applications listed are not exhaustive – we prioritized accessible sources and informative, causally focused applications.

Category	Design	Resources and Applications	Software and packages[1]
Experimental designs[2]	**Randomized Controlled Trial (RCT)**: Randomly assign units to treatment or control groups, helping to balance confounders across groups.	Kim & DeVries, 2001; Kimmel et al., 2021; Pynegar et al., 2021; Tilman et al., 2006; Weigel et al., 2021; Wiik et al., 2020	experiment (R package; see https://cran.r-project.org/package=experiment) RCT (R package; see https://cran.r-project.org/package=RCT) ExpAn (Python library; see https://github.com/zalando/expan)
	Factorial Design: Randomly assign units to multiple treatment combinations to test interactions and account for confounding of multiple causal variables.	Dasgupta et al., 2015; Jayewardene, 2009; Kaspari et al., 2012; King & Tschinkel, 2008; Laube & Zotz, 2003; Nicolaisen et al., 2014; Zhao & Ding, 2022	GFD (R package; see https://cran.r-project.org/package=GFD) fullfact (R package; see https://cran.r-project.org/package=fullfact) DoE.base (R package; see https://cran.r-project.org/package=DoE.base) pyDOE2 (Python library; see https://github.com/clicumu/pyDOE2) dexpy (Python library; see https://github.com/statease/dexpy)
	Crossover Trial: Units receive multiple treatments in a random sequence, allowing each unit to serve as its own control and account for confounders that vary between units.	Díaz-Uriarte, 2002; Feinsinger et al., 1991; Fergus et al., 2023; Jaakkola, 2003; Montesanto & Cividini, 2017; Ohrens et al., 2019; Shahn et al., 2023; Treves et al., 2024	crossdes (R package; see https://cran.r-project.org/package=crossdes) CrossCarry (R package; https://cran.r-project.org/package=CrossCarry) Crossover (R package; https://cran.r-project.org/package=Crossover)
	Cluster Randomized Trial: Randomize groups instead of individual units to	Benitez et al., 2023; Branas et al., 2018; Hemming & Taljaard, 2023; Schochet, 2013	cvcrand (R package; see https://cran.r-project.org/package=cvcrand) experiment (R package; see https://cran.r-project.org/package=experiment)

[1] See also https://cran.r-project.org/view=CausalInference
[2] See also https://cran.r-project.org/view=ExperimentalDesign

	account for group-level confounders.		cluster_experiments (Python library; see https://github.com/david26694/cluster-experiments)
Observational designs – controlling measured confounders	**Regression Adjustment**: Include confounders as covariates in the regression model describing the relationship of the causal variable on the outcome.	Fieberg & Ditmer, 2012; Gelman et al., 2020; Moss et al., 2025; Nogueira et al., 2022; Simler-Williamson & Germino, 2022	**R packages:** Base R functions – lm(…), glm(…), etc. – or dedicated regression packages **Python libraries:** statsmodels, linearmodels, etc. *Note: No dedicated packages or libraries – standard regression functions are used when confounders are explicitly specified in models used for causal interpretation.*
	Stratification: Divide units into subgroups, either during study design (e.g., stratified sampling) or during analysis (e.g., subgroup comparisons), based on confounders, then compares those with similar confounders but different exposure levels.	Morgan & Winship, 2014; Oehri et al., 2020; Rosenbaum, 2002	stdReg2 (R package; see https://cran.r-project.org/package=stdReg2) stratamatch (R package; see https://cran.r-project.org/package=stratamatch)
	Inverse Probability Weighting (IPW)[a]: Weight units based on their probability of exposure to create a pseudo-population where confounders are balanced.	Hernán & Robins, 2025 (Ch. 12); Nogueira et al., 2022; West et al., 2022	ipw (R package; see https://cran.r-project.org/package=ipw) twang (R package; see https://cran.r-project.org/package=twang) WeightIt (R package; see https://cran.r-project.org/package=WeightIt) CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)
	Propensity Score Matching (PSM)[a]: Match units with similar probabilities of exposure based on observed confounders (propensity scores) to create treatment	Butsic et al., 2017; Emmons et al., 2024; Nogueira et al., 2022; Pearson et al., 2016; Siegel, Larsen, et al., 2022; Siegel, Macaulay, et al., 2022; Simler-	Matching (R package; see https://cran.r-project.org/package=Matching) MatchIt (R package; see https://cran.r-project.org/package=MatchIt) CausalGPS (R package; see https://cran.r-project.org/package=CausalGPS) and

	and control groups with balanced covariate distributions.	Williamson & Germino, 2022; West et al., 2022; Wiik et al., 2020	pycausalgps (Python library; see https://github.com/NSAPH-Software/pycausalgps) psmpy (Python library; see https://pypi.org/project/psmpy)
	Marginal Structural Modeling (MSM)†: Use weighting to adjust for time-varying confounders when causal variables change over time.	Cole & Hernán, 2008; Hernán & Robins, 2025 (Ch. 12); Lei et al., 2019; Mandujano Reyes et al., 2025; Nandi et al., 2012; VanderWeele et al., 2011	bayesmsm (R package; see https://github.com/Kuan-Liu-Lab/bayesmsm) trajmsm (R package; see https://cran.r-project.org/package=trajmsm)
	Multi-level Modeling with Mixed Effects: Account for confounders from hierarchical data structures by including both fixed and random effects.	Bingenheimer & Raudenbush, 2004; Clough, 2012; Gelman, 2006; Gelman & Hill, 2006	lme4 (R package; see https://cran.r-project.org/package=lme4) brms (R package; see https://cran.r-project.org/package=brms) statsmodels (Python library; see https://www.statsmodels.org/) Bambi (Python library; see https://bambinos.github.io/bambi)
	Structural Causal Model (SCM)-based Back-door Criterion: Use causal diagrams to identify the minimal set of confounders that must be measured to enable unbiased estimation of causal effects.	Arif et al., 2022; Arif & MacNeil, 2022; Paul, 2011; Pearl, 2009; Schoolmaster et al., 2020; Stewart et al., 2023	causaleffect (R package; see https://cran.r-project.org/package=causaleffect) daggity (R package and Web interface; see https://dagitty.net) DoWhy (Python library; see https://py-why.github.io/dowhy)
	Structural Equation Modeling (SEM)[b,c]: Simultaneously quantify multiple causal relationships by including measured confounders as covariates in linear models.	Bollen & Pearl, 2013; Cronin & Schoolmaster, 2018; Grace et al., 2015; Hatami, 2019; Pearl, 1998, 2012; Saavedra et al., 2022	pwSEM[d] (R package; see https://github.com/BillShipley/pwSEM) piecewiseSEM[d] (R package; see https://cran.r-project.org/package=piecewiseSEM) lavaan (R package; see https://cran.r-project.org/package=lavaan)

			semopy (Python library; see https://semopy.com)
Observational designs – controlling unmeasured confounders	**Instrumental Variables (IV)**: Use a variable that influences the causal variable but not the outcome directly, to accounting for unmeasured confounders.	Butsic et al., 2017; Kendall, 2015; Larsen et al., 2019; MacDonald et al., 2019; MacDonald & Mordecai, 2019	ivreg (R package; see https://cran.r-project.org/package=ivreg) AER (R package; see https://cran.r-project.org/package=AER) EconML (Python library; see https://github.com/py-why/econml) CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)
	Before-After-Control-Impact (BACI)[e,f]: Compare changes in the outcome before and after a shift in the causal variable, while using a control group to account for time-varying confounders.	Chevalier et al., 2019; Christie et al., 2019; Comte et al., 2023; Ferraro et al., 2019; Kerr et al., 2019; Paul, 2011; Pitcher et al., 2009; Smokorowski & Randall, 2017; Wauchope et al., 2021	**R packages:** lme4, glmmTMB, or other multi-level modeling packages **Python libraries:** statsmodels, pingouin, or other packages supporting interaction terms in multi-level models *Note: No dedicated packages for BACI designs – analyses typically use mixed-effects models with an interaction term between **Time** (Before vs. After) and **Treatment** (Control vs. Impact) to estimate causal effects.*
	Difference-in-Differences (DiD)[†f]: Compare changes in the outcome over time between units with and without a change in exposure, while accounting for time-invariant confounders.	Butsic et al., 2017; Larsen et al., 2019; Simler-Williamson & Germino, 2022	did (R package; see https://cran.r-project.org/package=did) fixest (R package; see https://cran.r-project.org/package=fixest) CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)
	Regressions Discontinuity Design (RDD)[f]: Compare units just above and below a cutoff, assuming they are similar in all respects except exposure, to remove	Butsic et al., 2017; Cook et al., 2008; Imbens & Lemieux, 2008; Larsen et al., 2019; Noack et al., 2022	rdrobust (R package; see https://cran.r-project.org/package=rdrobust) rddensity (R package; see https://cran.r-project.org/package=rddensity) CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)

	confounding from variables that do not shift abruptly at the threshold.		
	Synthetic Control Methods†: Construct a synthetic control group from a weighted combination of unexposed units to approximate an exposed group with similar distributions of unmeasured confounders.	Abadie et al., 2010; Fick et al., 2021; West et al., 2022; X. Wu et al., 2023	Synth (R package; see https://cran.r-project.org/package=Synth) tidysynth (R package; see https://cran.r-project.org/package=tidysynth) CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)
	Multi-level Modeling with Fixed Effects†: Use only within-unit variation over time to address time-invariant unmeasured confounders.	Byrnes & Dee, 2025; Gelman & Hill, 2006; Simler-Williamson & Germino, 2022	fixest (R package; see https://cran.r-project.org/package=fixest) lfe (R package; see https://cran.r-project.org/package=lfe) plm (R package; see https://cran.r-project.org/package=plm) PyFixest (Python library; see https://github.com/py-econometrics/pyfixest)
	Structural Causal Model (SCM)-based Front-door Criterion: Use causal diagrams to identify sets of measured variables that address the effects of some unmeasured confounders.	Arif et al., 2022; Arif & MacNeil, 2022; Paul, 2011; Pearl, 2009; Stewart et al., 2023	causaleffect (R package; see https://cran.r-project.org/package=causaleffect) daggity (R package and Web interface; see https://dagitty.net) fdtlme (R package; see https://github.com/annaguo-bios/fdtmle) DoWhy (Python library; see https://py-why.github.io/dowhy)
	Interrupted Time Series Analysis†: Leverages a natural or implemented change using repeated outcome measurements before and after the change	Gilmour et al., 2006; Kontopantelis et al., 2015; Lopez Bernal et al., 2016; Wauchope et al., 2021	CausalImpact (R package; see https://github.com/google/CausalImpact) and CausalImpact (Python library; see https://pypi.org/project/causalimpact) segmented (R package; see https://cran.r-project.org/package=segmented)

	to account for pre-existing trends.		CausalPy (Python library; see https://github.com/pymc-labs/CausalPy)

[†]Requires time-series data.

[a]IPW and PSM are both examples of **Covariate Balancing** designs, which balance distribution of confounders across units with different exposure levels.

[b]SEMs can incorporate unobserved constructs (i.e., “latent variables”) which are inferred from measured variables.

[c]To support causal interpretations, SEMs must explicitly invoke untestable causal assumptions (see Bollen & Pearl, 2013; Pearl, 2012) and specify a causal structure via an SEM diagram (see Kunicki et al., 2023).

[d]While some SEM implementations allow some nonlinear specifications (e.g., via generalized additive models), they estimate causal effects using path coefficients or smooth terms derived from model components (Lefcheck, 2016) but do not provide formal identification criteria, as in nonparametric SCMs, to assess whether these effects can be uniquely determined from the data (Wang & Sobel, 2013).

[e]Experimental BACI designs with manipulated treatments are rare. Most BACI studies are observational and may not meet all the assumptions for robust causal inference (see Ferraro et al., 2019; Smokorowski & Randall, 2017; Wauchope et al., 2021).

[f]BACI, DID, and RDD are all examples of **Natural Experiments**, which leverage naturally occurring random variation in the causal variable to mimic randomization and account for unmeasured confounders.

Table S7. Algorithms for causal discovery, grouped by category. Each algorithm includes a brief description, key references (including applications in ecology, where available), and links to available software and code.

Category	Algorithm	Resources and Applications	Software and packages
Constraint-based methods	**PC (Peter and Clark)**: Uses repeated conditional independence tests to infer causal relationships from observed independencies in data, producing a set of causal graphs that represent possible causal relationships consistent with the data.	Bystrova et al., 2024; Chu et al., 2018; Ebert-Uphoff & Deng, 2012; Glymour et al., 2019; Kalisch et al., 2012; J. Li et al., 2015, pp. 9–20; Spirtes et al., 2000	pcalg (R package; see https://cran.r-project.org/package=pcalg) bnlearn (R package; see https://cran.r-project.org/package=bnlearn and https://www.bnlearn.com) Tetrad (GUI, Python library, R package; see https://www.cmu.edu/dietrich/philosophy/tetrad/use-tetrad) causal-learn (Python library; see https://causal-learn.readthedocs.io) pgmpy (Python library; see https://pgmpy.org)
	FCI (Fast Causal Inference): Extends the PC algorithm to detect possible unmeasured confounders, producing a causal graph that reflects uncertainty about edges.	Bystrova et al., 2024; Glymour et al., 2019; Kalisch et al., 2012; La Bastide-van Gemert et al., 2014; Mielke et al., 2022; Nogueira et al., 2022; Shen et al., 2020	pcalg (R package; see https://cran.r-project.org/package=pcalg) Tetrad (GUI, Python library, R package; see https://www.cmu.edu/dietrich/philosophy/tetrad/use-tetrad) causal-learn (Python library; see https://causal-learn.readthedocs.io)
	PCMCI (Peter and Clark Momentary Conditional Independence): A time-series adaptation of PC that improves detection of causal effects in autocorrelated data by iteratively testing for conditional independencies among variables and their lags.	Docquier et al., 2024; Krich et al., 2020; Nogueira et al., 2022; Runge, Nowack, et al., 2019; Tárraga et al., 2024	Tigramite (Python library; see https://github.com/jakobrunge/tigramite) CausalFlow (Python library; see https://github.com/lcastri/causalflow)
Score-based methods	**GES (Greedy Equivalence Search)**: Searches for the best	Gong et al., 2025; La Bastide-van Gemert et al., 2014	pcalg (R package; see https://cran.r-project.org/package=pcalg)

	causal graph by iteratively adding or removing edges based on a scoring criterion, such as the Bayesian Information Criterion (BIC), balancing data fit and simplicity.		Tetrad (GUI, Python library, R package; see https://www.cmu.edu/dietrich/philosophy/tetrad/use-tetrad) pgmpy (Python package; see https://pgmpy.org/) causal-learn (Python library; see https://causal-learn.readthedocs.io)
	GIES (Greedy Interventional Equivalence Search): An extension of GES that incorporates interventional data or assumptions to distinguish between equivalent causal graphs.	Hauser & Bühlmann, 2012; Shah et al., 2023	pcalg (R package; see https://cran.r-project.org/package=pcalg) Causal Discovery Toolbox (Python library; see https://github.com/FenTechSolutions/CausalDiscoveryToolbox) gies (Python library; see https://github.com/juangamella/gies)
	FGES (Fast Greedy Equivalence Search): A variant of GES that uses a parallelized greedy approach to rapidly search for the optimal causal graph, making it suitable for high-dimensional datasets.	Kitson & Constantinou, 2021; Ramsey et al., 2017; Shen et al., 2020	Tetrad (GUI, Python library, R package; see https://www.cmu.edu/dietrich/philosophy/tetrad/use-tetrad)
Functional model-based methods	**LiNGAM (Linear Non-Gaussian Acyclic Model)**: Identifies causal direction among variables by assuming linear relationships and non-Gaussian noise.	Ikeuchi et al., 2023; Kotoku et al., 2020; Kurotani et al., 2024; Shimizu, 2014; Shimizu et al., 2006, 2011	Tetrad (GUI, Python library, R package; see https://www.cmu.edu/dietrich/philosophy/tetrad/use-tetrad) causal-learn (Python library; see https://causal-learn.readthedocs.io) lingam (Python library; see https://github.com/cdt15/lingam)
	ANM (Additive Noise Model): Assumes the outcome variable is an unknown function of the causal variable plus independent additive noise, which enables	Bühlmann et al., 2014; Mooij et al., 2016; Peters et al., 2014; Song et al., 2022	CANM (R package; see https://github.com/Jie-Qiao/CANM) causal-learn (Python library; see https://causal-learn.readthedocs.io)

	identification of causal direction in both linear and nonlinear settings.		Causal Discovery Toolbox (Python library; see https://github.com/FenTechSolutions/CausalDiscoveryToolbox) lingam (Python library; see https://github.com/cdt15/lingam)
	IGCI (Information Geometric Causal Inference): Determines causal direction by analyzing asymmetries in the joint distributions of cause-effect pairs, without inherently controlling for or detecting unmeasured confounders or indirect causal effects.	Janzing et al., 2012; Mooij et al., 2016; Song et al., 2022	CANM (R package; see https://github.com/Jie-Qiao/CANM) Causal Discovery Toolbox (Python library; see https://github.com/FenTechSolutions/CausalDiscoveryToolbox) IGCI (Python library; see https://github.com/amber0309/IGCI)
Dynamical systems causality (DC)-based methods	**Granger Causality (GC)**: Tests whether past values of one time series can predict future values of another, assuming linear relationships in time-series data.	Detto et al., 2012; Granger, 1969; Nogueira et al., 2022; Reygadas et al., 2020; Singh & Borrok, 2019; Yuan & Shou, 2022	NlinTS (R package; see https://cran.r-project.org/package=NlinTS) causal-learn (Python library; see https://causal-learn.readthedocs.io)
	Information Theoretic (IT) Causality: A class of nonparametric and model-based methods that infer direct causal relationships by quantifying how knowledge of one variable reduces uncertainty about the future states of another variable. Includes Transfer Entropy (TE) approaches.	Benocci et al., 2025; Docquier et al., 2024; Hmamouche, 2020; Schreiber, 2000; Sun et al., 2015; Sun & Bollt, 2014; Yang et al., 2018	NlinTS (R package; see https://cran.r-project.org/package=NlinTS) copent (R package; see https://github.com/majianthu/copent) crossmapy (Python library; see https://github.com/PengTao-HUST/crossmapy) IDTxl (Python library; see https://github.com/pwollstadt/IDTxl)
	Convergent Cross Mapping (CCM): Uses state-space reconstruction to infer causal	Chang et al., 2017; Karakoç et al., 2020; Kitayama et al., 2021; Matsuzaki et al., 2018; Nova et al.,	rEDM (R package) and pyEDM (Python library); see

	relationships in nonlinear systems by testing whether past states of the causal variable can reliably predict current states of another variable.	2021; Sugihara et al., 2012; Ushio et al., 2018; J. Wu et al., 2023; Ye et al., 2015; Yuan & Shou, 2022	https://sugiharalab.github.io/EDM_Documentation
	Partial Cross Mapping (PCM): An extension of CCM that adjusts for potential unmeasured confounders to better isolate direct causal relationships.	Leng et al., 2020; Yongmei & Yulian, 2024	MATLAB code (see https://github.com/Partial-Cross-Mapping) crossmapy (Python library; see https://github.com/PengTao-HUST/crossmapy)

7. Advanced methods for causal inference and causal discovery

While many of the fundamental methods for causal discovery and causal inference have existed for several decades, the field of causal inference is continually evolving to incorporate novel statistical techniques and address increasingly complex data scenarios. For example, machine learning (ML) techniques are being integrated into methods for causal discovery and causal inference (Leist et al., 2022). Causal discovery with ML approaches, such as deep causal learning algorithms, use neural approaches to learn causal networks from a combination of empirical data and prior causal knowledge (C. Li et al., 2024; Scherrer et al., 2021; Yu et al., 2019). ML models can also be used in causal inference, provided the model and covariates are specified to accurately represent the underlying causal process (Brand et al., 2023; Hernán & Robins, 2024; Huber, 2023). For example, causal forests estimate causal effects using random forests (Wager & Athey, 2018), while double/debiased ML methods, such as targeted maximum likelihood estimation (TLME) (van der Laan & Rubin, 2006), control for measured confounders using ML models that can capture complex nonlinear and high-dimensional patterns of confounding (Chernozhukov et al., 2018). We summarize some of these advanced methods for both causal discovery and causal inference in Table S8.

It should be noted that not all ML approaches are appropriate for causal analyses (Pichler & Hartig, 2023). ML approaches are merely a class of models that, without pre-existing knowledge and assumptions, are purely intended for predictive tasks and are not appropriate for obtaining causal interpretations (Section S2). Thus, causal ML approaches still require the principles and assumptions linking statistical dependence to causal dependence (Section S4), and careful model building using pre-existing knowledge about all relevant confounding variables is essential for these methods to detect and estimate causal effects without bias (Section S3).

Table S8. Advanced methods for causal discovery and causal inference, grouped by causal task. Each method includes a brief description, key references and links to relevant software and code.

Causal Task	Method	Resources and Applications	Software and packages[3]
Causal discovery	**Deep causal learning**: Uses deep learning models (e.g., neural networks) to detect causal relationships in complex, high-dimensional data, often incorporating pre-existing knowledge to improve accuracy.	C. Li et al., 2024; Luo et al., 2020; Yu et al., 2019; K. Zheng et al., 2024	DAG-GNN (Python code; see https://github.com/fishmoon1234/DAG-GNN) DeFuSE (Python code; see https://github.com/chunlinli/defuse) Dagma (Python library; see https://github.com/kevinsbello/dagma)
	Causal representation learning: Learning disentangled latent representations that correspond to underlying causal variables and capture the structure of the data-generating process.	Ahuja et al., 2023; Brehmer et al., 2022; Scholkopf et al., 2021	Emei (Python library; see https://github.com/FrankTianTT/emei) DRL (Python code; see https://github.com/CausalRL/DRL) gCastle (Python library; see https://pypi.org/project/gcastle)
	Causal reinforcement learning: Incorporates causal assumptions or causal models into reinforcement learning (a machine learning approach where models learn by trying actions and observing which ones produce the best outcomes).	Buesing et al., 2019; Wang et al., 2021; Zeng et al., 2025; Zhu et al., 2020	CARL (Python code; see https://github.com/arquimides/carl) *Note: No dedicated packages or libraries – most implementations of causal reinforcement learning are ad hoc in published papers or preprints.*
	Invariant causal prediction: Identifies causal variables by selecting predictors whose statistical relationships with the outcome remain invariant across environments or experimental settings.	Peters et al., 2016; Pfister et al., 2019	InvariantCausalPrediction (R package; see https://cran.r-project.org/package=InvariantCausalPrediction) causalicp (Python library; see https://github.com/juangamella/icp)

[3]See also https://github.com/rguo12/awesome-causality-algorithms

Causal inference	**Targeted Maximum Likelihood Estimation (TMLE)**: Semi-parametric method that uses machine learning models for flexible outcome and treatment modeling, with a targeted correction step to ensure valid inference.	Luque-Fernandez et al., 2018; Schuler & Rose, 2017; van der Laan & Rubin, 2006	tmle3 (R package; see https://tlverse.org/tmle3) causal-curve (Python library; see https://github.com/ronikobrosly/causal-curve)
	Double/debiased machine learning: Uses machine learning to model outcomes and treatments separately, then combines them to estimate treatment effects while controlling for confounding in high-dimensional settings.	Chernozhukov et al., 2018; Fink et al., 2023; B. Shi et al., 2024	DoubleML (R package; see https://cran.r-project.org/package=DoubleML) EconML (Python library; see https://github.com/py-why/econml)
	Causal forests: Uses ensembles of decision trees to estimate heterogeneous treatment effects while accounting for confounding.	Athey et al., 2019; Athey & Wager, 2019; Fink et al., 2023; Wager & Athey, 2018; Xie et al., 2012; L. Zheng & Yin, 2023	grf (R package; see https://cran.r-project.org/package=grf) EconML (Python library; see https://github.com/py-why/econml)
	Meta-learners for heterogeneous treatment effects (e.g., S-learner, T-learner, X-learner, and R-learner): Use machine learning models to estimate heterogeneous treatment effects by modeling outcomes separately for different treatment levels, with a tradeoff between simple implementation and reduced reliability in inference.	Jiang et al., 2021; Künzel et al., 2019; Nie & Wager, 2021; Salditt et al., 2024	rlearner (R package; see https://github.com/xnie/rlearner) EconML (Python library; see https://github.com/py-why/econml) CausalML (Python library; see https://github.com/uber/causalml) metalearners (Python library; see https://github.com/quantco/metalearners)
	Causal inference using Bayesian machine learning: Estimate treatment effects using Bayesian machine learning models (e.g., Bayesian Additive Regression Trees [BART]) to capture nonlinear	Green & Kern, 2012; Hahn et al., 2020; J. Hill et al., 2020; J. L. Hill, 2011; Zeldow et al., 2019	bartCause (R package; see https://github.com/vdorie/bartCause) BCI Toolbox (Python library; see https://github.com/evans1112/bcitoolbox)

	relationships and quantify uncertainty via posterior distributions.		
	Counterfactual fairness: Defines fairness based on counterfactual comparisons across protected attributes using structural causal models, ensuring outcomes would remain the same in a hypothetical world where protected group membership had been different.	Chiappa, 2019; Nabi & Shpitser, 2018; Y. Wu et al., 2019	EXOC (Python code; see https://github.com/CASE-Lab-UMD/counterfactual_fairness_2025) *Note: No dedicated packages or libraries – most implementations of counterfactual fairness are ad hoc in published papers or preprints.*
	Causal data fusion: Combines data from different sources (e.g., observational and experimental) to estimate causal effects when no single dataset is sufficient, using assumptions encoded in transportability diagrams (causal diagrams that represent differences between data sources).	Bareinboim & Pearl, 2016; Chau et al., 2021; Josey et al., 2022; Pearl & Bareinboim, 2014	*Note: Data fusion methods remain in development, thus general-purpose implementations are not currently widely available. Implementation of some data fusion concepts are available via a GUI at https://causalfusion.net. A Python library called Y_0 (see https://github.com/y0-causal-inference/y0) also implements some data fusion concepts (e.g., parsing transportability graphs).*

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Addicott, E. T., Fenichel, E. P., Bradford, M. A., Pinsky, M. L., & Wood, S. A. (2022). Toward an improved understanding of causation in the ecological sciences. *Frontiers in Ecology and the Environment*, *20*(8), 474–480. https://doi.org/10.1002/fee.2530

Ahuja, K., Mahajan, D., Wang, Y., & Bengio, Y. (2023). Interventional Causal Representation Learning. In A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, & J. Scarlett (Eds.), *Proceedings of the 40th International Conference on Machine Learning* (Vol. 202, pp. 372–407). PMLR. https://proceedings.mlr.press/v202/ahuja23a.html

Andersson, S. A., Madigan, D., & Perlman, M. D. (1997). A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics*, *25*(2). https://doi.org/10.1214/aos/1031833662

Arif, S., Graham, N. A. J., Wilson, S., & MacNeil, M. A. (2022). Causal drivers of climate-mediated coral reef regime shifts. *Ecosphere*, *13*(3), e3956. https://doi.org/10.1002/ecs2.3956

Arif, S., & MacNeil, M. A. (2022). Utilizing causal diagrams across quasi-experimental approaches. *Ecosphere*, *13*(4), e4009. https://doi.org/10.1002/ecs2.4009

Arif, S., & Massey, M. D. B. (2023). Reducing bias in experimental ecology through directed acyclic graphs. *Ecology and Evolution*, *13*(3), e9947. https://doi.org/10.1002/ece3.9947

Athey, S., Tibshirani, J., & Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, *47*(2). https://doi.org/10.1214/18-AOS1709

Athey, S., & Wager, S. (2019). Estimating Treatment Effects with Causal Forests: An Application. *Observational Studies*, *5*(2), 37–51. https://doi.org/10.1353/obs.2019.0001

Baker, R. E., Peña, J.-M., Jayamohan, J., & Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, *14*(5), 20170660. https://doi.org/10.1098/rsbl.2017.0660

Barański, K., Gutman, Y., & Śpiewak, A. (2020). A probabilistic Takens theorem. *Nonlinearity*, *33*(9), 4940–4966. https://doi.org/10.1088/1361-6544/ab8fb8

Bareinboim, E., & Pearl, J. (2016). Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, *113*(27), 7345–7352. https://doi.org/10.1073/pnas.1510507113

Benitez, A., Petersen, M. L., Van Der Laan, M. J., Santos, N., Butrick, E., Walker, D., Ghosh, R., Otieno, P., Waiswa, P., & Balzer, L. B. (2023). Defining and estimating effects in cluster randomized trials: A methods comparison. *Statistics in Medicine*, *42*(19), 3443–3466. https://doi.org/10.1002/sim.9813

Benocci, R., Guagliumi, G., Potenza, A., Zaffaroni-Caorsi, V., Roman, H. E., & Zambon, G. (2025). Application of Transfer Entropy Measure to Characterize Environmental Sounds in Urban and Wild Parks. *Sensors (Basel, Switzerland)*, *25*(4), 1046. https://doi.org/10.3390/s25041046

Bingenheimer, J. B., & Raudenbush, S. W. (2004). Statistical and Substantive Inferences in Public Health: Issues in the Application of Multilevel Models. *Annual Review of Public Health*, *25*(1), 53–77. https://doi.org/10.1146/annurev.publhealth.25.050503.153925

Bollen, K. A., & Pearl, J. (2013). Eight Myths About Causality and Structural Equation Models. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research* (pp. 301–328). Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_15

Bongers, S., Blom, T., & Mooij, J. M. (2018). *Causal Modeling of Dynamical Systems* (Version 4). arXiv. https://doi.org/10.48550/ARXIV.1803.08784

Bongers, S., Forré, P., Peters, J., & Mooij, J. M. (2021). Foundations of structural causal models with cycles and latent variables. *The Annals of Statistics*, *49*(5). https://doi.org/10.1214/21-AOS2064

Branas, C. C., South, E., Kondo, M. C., Hohl, B. C., Bourgois, P., Wiebe, D. J., & MacDonald, J. M. (2018). Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear. *Proceedings of the National Academy of Sciences*, *115*(12), 2946–2951. https://doi.org/10.1073/pnas.1718503115

Brand, J. E., Zhou, X., & Xie, Y. (2023). Recent Developments in Causal Inference and Machine Learning. *Annual Review of Sociology*, *49*(1), 81–110. https://doi.org/10.1146/annurev-soc-030420-015345

Brehmer, J., de Haan, P., Lippe, P., & Cohen, T. S. (2022). Weakly supervised causal representation learning. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), *Advances in Neural Information Processing Systems* (Vol. 35, pp. 38319–38331). Curran Associates, Inc.

https://proceedings.neurips.cc/paper_files/paper/2022/file/fa567e2b2c870f8f09a87b6e73370869-Paper-Conference.pdf
Brown, P. J., Fuller, W. A., American Mathematical Society, Institute of Mathematical Statistics, & Society for Industrial and Applied Mathematics (Eds.). (1990). *Statistical analysis of measurement error models and applications: proceedings of the AMS-IMS-SIAM joint summer research conference held June 10-16, 1989*. American Mathematical Society.
Buesing, L., Weber, T., Zwols, Y., Heess, N., Racaniere, S., Guez, A., & Lespiau, J.-B. (2019). Woulda, Coulda, Shoulda: Counterfactually-Guided Policy Search. *International Conference on Learning Representations*. https://openreview.net/forum?id=BJG0voC9YQ
Bühlmann, P., Peters, J., & Ernest, J. (2014). CAM: Causal additive models, high-dimensional order search and penalized regression. *The Annals of Statistics*, *42*(6). https://doi.org/10.1214/14-AOS1260
Bulbulia, J. A. (2024a). Methods in causal inference. Part 1: causal diagrams and confounding. *Evolutionary Human Sciences*, *6*, e40. https://doi.org/10.1017/ehs.2024.35
Bulbulia, J. A. (2024b). Methods in causal inference. Part 2: Interaction, mediation, and time-varying treatments. *Evolutionary Human Sciences*, *6*, e41. https://doi.org/10.1017/ehs.2024.32
Burnett, J. W., & Blackwell, C. (2024). Graphical causal modelling: an application to identify and estimate cause-and-effect relationships. *Applied Economics*, *56*(33), 3986–4000. https://doi.org/10.1080/00036846.2023.2208856
Burnham, K. P., & Anderson, D. R. (2010). *Model selection and multimodel inference: a practical information-theoretic approach* (2. ed., [4. printing]). Springer.
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*, 1–10. https://doi.org/10.1016/j.baae.2017.01.005
Byrnes, J. E. K., & Dee, L. E. (2025). Causal Inference With Observational Data and Unobserved Confounding Variables. *Ecology Letters*, *28*(1), e70023. https://doi.org/10.1111/ele.70023
Bystrova, D., Assaad, C. K., Si-moussi, S., & Thuiller, W. (2024). *Causal discovery from ecological time-series with one timestamp and multiple observations*. https://doi.org/10.1101/2024.10.10.608447

Cai, M., Wang, Z., Xiao, J., Hu, X., Chen, G., & Yang, C. (2023). XMAP: Cross-population fine-mapping by leveraging genetic diversity and accounting for confounding bias. *Nature Communications*, *14*(1), 6870. https://doi.org/10.1038/s41467-023-42614-7

Caldararu, S., Rolo, V., Stocker, B. D., Gimeno, T. E., & Nair, R. (2023). Ideas and perspectives: Beyond model evaluation – combining experiments and models to advance terrestrial ecosystem science. *Biogeosciences*, *20*(17), 3637–3649. https://doi.org/10.5194/bg-20-3637-2023

Chang, C., Ushio, M., & Hsieh, C. (2017). Empirical dynamic modeling for beginners. *Ecological Research*, *32*(6), 785–796. https://doi.org/10.1007/s11284-017-1469-9

Chau, S. L., Ton, J.-F., González, J., Teh, Y., & Sejdinovic, D. (2021). BayesIMP: Uncertainty Quantification for Causal Data Fusion. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang, & J. W. Vaughan (Eds.), *Advances in Neural Information Processing Systems* (Vol. 34, pp. 3466–3477). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2021/file/1ca5c750a30312d1919ae6a4d636dcc4-Paper.pdf

Cheng, D., Li, J., Liu, L., Liu, J., & Le, T. D. (2024). Data-Driven Causal Effect Estimation Based on Graphical Causal Modelling: A Survey. *ACM Computing Surveys*, *56*(5), 1–37. https://doi.org/10.1145/3636423

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1), C1–C68. https://doi.org/10.1111/ectj.12097

Chevalier, M., Russell, J. C., & Knape, J. (2019). New measures for evaluation of environmental perturbations using Before-After-Control-Impact analyses. *Ecological Applications*, *29*(2), e01838. https://doi.org/10.1002/eap.1838

Chiappa, S. (2019). Path-Specific Counterfactual Fairness. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 7801–7808. https://doi.org/10.1609/aaai.v33i01.33017801

Christie, A. P., Amano, T., Martin, P. A., Shackelford, G. E., Simmons, B. I., & Sutherland, W. J. (2019). Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology*, *56*(12), 2742–2754. https://doi.org/10.1111/1365-2664.13499

Christin, S., Hervet, É., & Lecomte, N. (2019). Applications for deep learning in ecology. *Methods in Ecology and Evolution*, *10*(10), 1632–1644. https://doi.org/10.1111/2041-210X.13256

Chu, T., Danks, D., & Glymour, C. (2018). *Data Driven Methods for Nonlinear Granger Causality: Climate Teleconnection Mechanisms*. https://doi.org/10.1184/R1/6491327.v1

Clough, Y. (2012). A generalized approach to modeling and estimating indirect effects in ecology. *Ecology*, *93*(8), 1809–1815. https://doi.org/10.1890/11-1899.1

Cobey, S., & Baskerville, E. B. (2016). Limits to Causal Inference with State-Space Reconstruction for Infectious Disease. *PLOS ONE*, *11*(12), e0169050. https://doi.org/10.1371/journal.pone.0169050

Cole, S. R., & Hernán, M. A. (2008). Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology*, *168*(6), 656–664. https://doi.org/10.1093/aje/kwn164

Comte, S., Bengsen, A. J., Thomas, E., Bennett, A., Davis, N. E., Brown, D., & Forsyth, D. M. (2023). A Before-After Control-Impact experiment reveals that culling reduces the impacts of invasive deer on endangered peatlands. *Journal of Applied Ecology*, *60*(11), 2340–2350. https://doi.org/10.1111/1365-2664.14498

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management*, *27*(4), 724–750. https://doi.org/10.1002/pam.20375

Correia, H. E., Dee, L. E., & Ferraro, P. J. (2025). Designing causal mediation analyses to quantify intermediary processes in ecology. *Biological Reviews*, brv.70011. https://doi.org/10.1111/brv.70011

Cox, D. R. (1958). *Planning of experiments* (p. 308). Wiley.

Cronin, J. P., & Schoolmaster, D. R. (2018). A causal partition of trait correlations: using graphical models to derive statistical models from theoretical language. *Ecosphere*, *9*(9), e02422. https://doi.org/10.1002/ecs2.2422

Cummins, B., Gedeon, T., & Spendlove, K. (2015). On the Efficacy of State Space Reconstruction Methods in Determining Causality. *SIAM Journal on Applied Dynamical Systems*, *14*(1), 335–381. https://doi.org/10.1137/130946344

Dasgupta, T., Pillai, N. S., & Rubin, D. B. (2015). Causal Inference from 2K Factorial Designs by Using Potential Outcomes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *77*(4), 727–753. https://doi.org/10.1111/rssb.12085

Dawid, P. (2000). Causal Inference without Counterfactuals. *Journal of the American Statistical Association*, *95*(450), 407–424. https://doi.org/10.1080/01621459.2000.10474210

Dawid, P. (2012). The Decision-Theoretic Approach to Causal Inference. In C. Berzuini, P. Dawid, & L. Bernardinelli (Eds.), *Wiley Series in Probability and Statistics* (1st ed., pp. 25–42). Wiley. https://doi.org/10.1002/9781119945710.ch4

Dawid, P. (2021). Decision-theoretic foundations for statistical causality. *Journal of Causal Inference*, *9*(1), 39–77. https://doi.org/10.1515/jci-2020-0008

Dawid, P. (2024). Potential outcomes and decision-theoretic foundations for statistical causality: Response to Richardson and Robins. *Journal of Causal Inference*, *12*(1), 20230058. https://doi.org/10.1515/jci-2023-0058

De Brouwer, E., Arany, A., Simm, J., & Moreau, Y. (2021). Latent Convergent Cross Mapping. *International Conference on Learning Representations*. https://openreview.net/forum?id=4TSiOTkKe5P

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. https://doi.org/10.1086/664628

Deyle, E. R., & Sugihara, G. (2011). Generalized Theorems for Nonlinear State Space Reconstruction. *PLoS ONE*, *6*(3), e18295. https://doi.org/10.1371/journal.pone.0018295

Ding, Y., & Toulis, P. (2018). *Dynamical systems theory for causal inference with application to synthetic control methods* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1808.08778

Díaz-Uriarte, R. (2002). Incorrect analysis of crossover trials in animal behaviour research. *Animal Behaviour*, *63*(4), 815–822. https://doi.org/10.1006/anbe.2001.1950

Docquier, D., Di Capua, G., Donner, R. V., Pires, C. A. L., Simon, A., & Vannitsem, S. (2024). A comparison of two causal methods in the context of climate analyses. *Nonlinear Processes in Geophysics*, *31*(1), 115–136. https://doi.org/10.5194/npg-31-115-2024

Drton, M., & Richardson, T. S. (2004). Iterative conditional fitting for Gaussian ancestral graph models. *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, 130–137. https://doi.org/10.48550/arXiv.1207.4118

Ebert-Uphoff, I., & Deng, Y. (2012). Causal Discovery for Climate Research Using Graphical Models. *Journal of Climate*, *25*(17), 5648–5665. https://doi.org/10.1175/JCLI-D-11-00387.1

Efron, B., & Hastie, T. (2016). *Computer age statistical inference: algorithms, evidence, and data science*. Cambridge University Press.

Eichler, M. (2013). Causal inference with multiple time series: principles and problems. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences*, *371*(1997), 20110613. https://doi.org/10.1098/rsta.2011.0613

Elwert, F. (2013). Graphical Causal Models. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research* (pp. 245–273). Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13

Emmons, S., Woods, T., Cashman, M., Devereux, O., Noe, G., Young, J., Stranko, S., Kilian, J., Hanna, K., & Maloney, K. (2024). Causal inference approaches reveal both positive and negative unintended effects of agricultural and urban management practices on instream biological condition. *Journal of Environmental Management*, *361*, 121234. https://doi.org/10.1016/j.jenvman.2024.121234

Evans, M. R., Grimm, V., Johst, K., Knuuttila, T., De Langhe, R., Lessells, C. M., Merz, M., O’Malley, M. A., Orzack, S. H., Weisberg, M., Wilkinson, D. J., Wolkenhauer, O., & Benton, T. G. (2013). Do simple models lead to generality in ecology? *Trends in Ecology & Evolution*, *28*(10), 578–583. https://doi.org/10.1016/j.tree.2013.05.022

Feinsinger, P., Tiebout, H. M., & Young, B. E. (1991). Do Tropical Bird-Pollinated Plants Exhibit Density-Dependent Interactions? Field Experiments. *Ecology*, *72*(6), 1953–1963. https://doi.org/10.2307/1941550

Fergus, A. R., Hermanstorfer, S. J., & Treves, A. (2023). *Combining two non-lethal methods in crossover design randomized experiments.* https://doi.org/10.31220/agriRxiv.2023.00203

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*, *116*(12), 5311–5318. https://doi.org/10.1073/pnas.1805563115

Fick, S. E., Nauman, T. W., Brungard, C. C., & Duniway, M. C. (2021). Evaluating natural experiments in ecology: using synthetic controls in assessments of remotely sensed land treatments. *Ecological Applications*, *31*(3), e02264. https://doi.org/10.1002/eap.2264

Fieberg, J., & Ditmer, M. (2012). Understanding the causes and consequences of animal movement: a cautionary note on fitting and interpreting regression models with time-dependent covariates. *Methods in Ecology and Evolution*, *3*(6), 983–991. https://doi.org/10.1111/j.2041-210X.2012.00239.x

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. https://doi.org/10.1111/2041-210X.14186

Gelman, A. (2006). Multilevel (Hierarchical) Modeling: What It Can and Cannot Do. *Technometrics*, *48*(3), 432–435. https://doi.org/10.1198/004017005000000661

Gelman, A., & Hill, J. (2006). Causal inference using multilevel models. In *Data Analysis Using Regression and Multilevel/Hierarchical Models* (1st ed., pp. 503–512). Cambridge University Press. https://doi.org/10.1017/CBO9780511790942

Gelman, A., Hill, J., & Vehtari, A. (Eds.). (2020). Causal inference using regression on the treatment variable. In *Regression and Other Stories* (pp. 363–382). Cambridge University Press; Cambridge Core. https://doi.org/10.1017/9781139161879.020

Gilmour, S., Degenhardt, L., Hall, W., & Day, C. (2006). Using intervention time series analyses to assess the effects of imperfectly identifiable natural events: a general method and example. *BMC Medical Research Methodology*, *6*, 16. https://doi.org/10.1186/1471-2288-6-16

Glymour, C., Zhang, K., & Spirtes, P. (2019). Review of Causal Discovery Methods Based on Graphical Models. *Frontiers in Genetics*, *10*, 524. https://doi.org/10.3389/fgene.2019.00524

Gong, K., Chen, Y., Song, X., Fu, Z., & Ding, X. (2025). Causal Inference for Hypertension Prediction With Wearable E lectrocardiogram and P hotoplethysmogram Signals: Feasibility Study. *JMIR Cardio*, *9*, e60238–e60238. https://doi.org/10.2196/60238

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., Scheiner, S. M., & Schoolmaster, Jr., D. R. (2015). Structural equation modeling: building and evaluating causal models. In G. A. Fox, S. Negrete-Yankelevich, & V. J. Sosa (Eds.), *Ecological Statistics* (1st ed., pp. 168–199). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199672547.003.0009

Granger, C. W. J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, *37*(3), 424. https://doi.org/10.2307/1912791

Green, D. P., & Kern, H. L. (2012). Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees. *Public Opinion Quarterly*, *76*(3), 491–511. https://doi.org/10.1093/poq/nfs036

Greenland, S. (2003). Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology (Cambridge, Mass.)*, *14*(3), 300–306.

Greenland, S., Pearl, J., & Robins, J. M. (1999a). Causal diagrams for epidemiologic research. *Epidemiology (Cambridge, Mass.)*, *10*(1), 37–48.

Greenland, S., Pearl, J., & Robins, J. M. (1999b). Confounding and Collapsibility in Causal Inference. *Statistical Science*, *14*(1). https://doi.org/10.1214/ss/1009211805

Haber, N. A., Wieten, S. E., Rohrer, J. M., Arah, O. A., Tennant, P. W. G., Stuart, E. A., Murray, E. J., Pilleron, S., Lam, S. T., Riederer, E., Howcutt, S. J., Simmons, A. E., Leyrat, C., Schoenegger, P., Booman, A., Dufour, M.-S. K., O'Donoghue, A. L., Baglini, R., Do, S., … Fox, M. P. (2022). Causal and Associational Language in Observational Health Research: A Systematic Evaluation. *American Journal of Epidemiology*, *191*(12), 2084–2097. https://doi.org/10.1093/aje/kwac137

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). https://doi.org/10.1214/19-BA1195

Han, M. A., & Guyatt, G. (2020). Systematic survey of the causal language use in systematic reviews of observational studies: a study protocol. *BMJ Open*, *10*(7), e038571. https://doi.org/10.1136/bmjopen-2020-038571

Harnack, D., Laminski, E., Schünemann, M., & Pawelzik, K. R. (2017). Topological Causality in Dynamical Systems. *Physical Review Letters*, *119*(9), 098301. https://doi.org/10.1103/PhysRevLett.119.098301

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Hatami, R. (2019). A Review of the Techniques Used to Control Confounding Bias and How Spatiotemporal Variation Can Be Controlled in Environmental Impact Studies. *Water, Air, & Soil Pollution*, *230*(6), 132. https://doi.org/10.1007/s11270-019-4150-9

Hauser, A., & Bühlmann, P. (2012). Characterization and Greedy Learning of Interventional Markov Equivalence Classes of Directed Acyclic Graphs. *Journal of Machine Learning Research*, *13*(79), 2409–2464.

Hayduk, L., Cummings, G., Stratkotter, R., Nimmo, M., Grygoryev, K., Dosman, D., Gillespie, M., Pazderka-Robinson, H., & Boadu, K. (2003). Pearl's D-Separation: One More Step Into Causal Thinking. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(2), 289–311. https://doi.org/10.1207/S15328007SEM1002_8

Heiss, A. (2024, March 21). Demystifying causal inference estimands: ATE, ATT, and ATU. *Andrew Heiss's Blog*. https://doi.org/10.59350/c9z3a-rcq16

Hemming, K., & Taljaard, M. (2023). Key considerations for designing, conducting and analysing a cluster randomized trial. *International Journal of Epidemiology*, *52*(5), 1648–1658. https://doi.org/10.1093/ije/dyad064

Hernán, M. A. (2018). The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. *American Journal of Public Health*, *108*(5), 616–619. https://doi.org/10.2105/AJPH.2018.304337

Hernán, M. A., Dahabreh, I. J., Dickerman, B. A., & Swanson, S. A. (2025). The Target Trial Framework for Causal Inference From Observational Data: Why and When Is It Helpful? *Annals of Internal Medicine*. https://doi.org/10.7326/ANNALS-24-01871

Hernán, M. A., Hsu, J., & Healy, B. (2019). A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. *CHANCE*, *32*(1), 42–49. https://doi.org/10.1080/09332480.2019.1579578

Hernán, M. A., & Robins, J. M. (2024). Variable selection and high-dimensional data. In *Causal inference: What if* (First edition, pp. 235–246). Taylor and Francis.

Hernán, M. A., & Robins, J. M. (2025). *Causal inference: What if* (First edition). CRC Press.

Hernán, M. A., Wang, W., & Leaf, D. E. (2022). Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA*, *328*(24), 2446. https://doi.org/10.1001/jama.2022.21383

Hill, J. L. (2011). Bayesian Nonparametric Modeling for Causal Inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217–240. https://doi.org/10.1198/jcgs.2010.08162

Hill, J., Linero, A., & Murray, J. (2020). Bayesian Additive Regression Trees: A Review and Look Forward. *Annual Review of Statistics and Its Application*, *7*(1), 251–278. https://doi.org/10.1146/annurev-statistics-031219-041110

Hmamouche, Y. (2020). NlinTS: An R Package For Causality Detection in Time Series. *The R Journal*, *12*(1), 21. https://doi.org/10.32614/RJ-2020-016

Holland, P. W. (1986). Statistics and Causal Inference. *Journal of the American Statistical Association*, *81*(396), 945–960. https://doi.org/10.1080/01621459.1986.10478354

Huber, M. (2023). *Causal analysis: impact evaluation and causal machine learning with applications in R*. The MIT Press.

Hudgens, M. G., & Halloran, M. E. (2008). Toward Causal Inference With Interference. *Journal of the American Statistical Association*, *103*(482), 832–842. https://doi.org/10.1198/016214508000000292

Hyslop, D. R., & Imbens, G. W. (2001). Bias From Classical and Other Forms of Measurement Error. *Journal of Business & Economic Statistics*, *19*(4), 475–481. https://doi.org/10.1198/07350010152596727

Ibeling, D., & Icard, T. (2023). Comparing Causal Frameworks: Potential Outcomes, Structural Models, Graphs, and Abstractions. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, & S. Levine (Eds.), *Advances in Neural Information Processing Systems* (Vol. 36, pp. 80130–80141). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2023/file/fd83f4e0dcaf1c64ea15bbb1695bb40f-Paper-Conference.pdf

Ikeuchi, T., Ide, M., Zeng, Y., Maeda, T. N., & Shimizu, S. (2023). Python package for causal discovery based on LiNGAM. *Journal of Machine Learning Research*, *24*(14), 1–8.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *Review of Economics and Statistics*, *86*(1), 4–29. https://doi.org/10.1162/003465304323023651

Imbens, G. W. (2020). Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics. *Journal of Economic Literature*, *58*(4), 1129–1179. https://doi.org/10.1257/jel.20191597

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467. https://doi.org/10.2307/2951620

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Jaakkola, J. J. K. (2003). Case-crossover design in air pollution epidemiology. *European Respiratory Journal*, *21*(40 suppl), 81s–85s. https://doi.org/10.1183/09031936.03.00402703

Janzing, D., Mooij, J., Zhang, K., Lemeire, J., Zscheischler, J., Daniušis, P., Steudel, B., & Schölkopf, B. (2012). Information-geometric approach to inferring causal directions. *Artificial Intelligence*, *182–183*, 1–31. https://doi.org/10.1016/j.artint.2012.01.002

Jayewardene, D. (2009). A factorial experiment quantifying the influence of parrotfish density and size on algal reduction on Hawaiian coral reefs. *Journal of Experimental Marine Biology and Ecology*, *375*(1–2), 64–69. https://doi.org/10.1016/j.jembe.2009.05.006

Jiang, H., Qi, P., Zhou, J., Zhou, J., & Rao, S. (2021). A Short Survey on Forest Based Heterogeneous Treatment Effect Estimation Methods: Meta-learners and Specific Models. *2021 IEEE International Conference on Big Data (Big Data)*, 3006–3012. https://doi.org/10.1109/BigData52589.2021.9671439

Jones, H. E., & Schooling, C. M. (2018). Let's Require the "T-Word." *American Journal of Public Health*, *108*(5), 624–624. https://doi.org/10.2105/AJPH.2018.304365

Josey, K. P., Yang, F., Ghosh, D., & Raghavan, S. (2022). A calibration approach to transportability and data-fusion with observational data. *Statistics in Medicine*, *41*(23), 4511–4531. https://doi.org/10.1002/sim.9523

Kalisch, M., Mächler, M., Colombo, D., Maathuis, M. H., & Bühlmann, P. (2012). Causal Inference Using Graphical Models with the *R* Package **pcalg**. *Journal of Statistical Software*, *47*(11). https://doi.org/10.18637/jss.v047.i11

Karakoç, C., Clark, A. T., & Chatzinotas, A. (2020). Diversity and coexistence are influenced by time-dependent species interactions in a predator–prey system. *Ecology Letters*, *23*(6), 983–993. https://doi.org/10.1111/ele.13500

Kaspari, M., Donoso, D., Lucas, J. A., Zumbusch, T., & Kay, A. D. (2012). Using nutritional ecology to predict community structure: a field test in Neotropical ants. *Ecosphere*, *3*(11), 1–15. https://doi.org/10.1890/ES12-00136.1

Kendall, B. E. (2015). A statistical symphony: instrumental variables reveal causality and control measurement error. In G. A. Fox, S. Negrete-Yankelevich, & V. J. Sosa (Eds.), *Ecological Statistics* (1st ed., pp. 149–167). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780199672547.003.0008

Kerr, L. A., Kritzer, J. P., & Cadrin, S. X. (2019). Strengths and limitations of before–after–control–impact analysis for testing the effects of marine protected areas on managed populations. *ICES Journal of Marine Science*, *76*(4), 1039–1051. https://doi.org/10.1093/icesjms/fsz014

Kezios, K. L., & Hayes-Larson, E. (2018). A Clarification on Causal Questions: We Ask Them More Often Than We Realize. *American Journal of Public Health*, *108*(8), e4–e4. https://doi.org/10.2105/AJPH.2018.304547

Kim, G. W., & DeVries, D. R. (2001). Adult Fish Predation on Freshwater Limnetic Fish Larvae: A Mesocosm Experiment. *Transactions of the American Fisheries Society*, *130*(2), 189–203. https://doi.org/10.1577/1548-8659(2001)130<0189:AFPOFL>2.0.CO;2

Kimmel, K., Avolio, M. L., & Ferraro, P. J. (2023). Empirical evidence of widespread exaggeration bias and selective reporting in ecology. *Nature Ecology & Evolution*, *7*(9), 1525–1536. https://doi.org/10.1038/s41559-023-02144-3

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. https://doi.org/10.1016/j.tree.2021.08.008

King, J. R., & Tschinkel, W. R. (2008). Experimental evidence that human impacts drive fire ant invasions and ecological change. *Proceedings of the National Academy of Sciences*, *105*(51), 20339–20343. https://doi.org/10.1073/pnas.0809423105

Kitayama, K., Ushio, M., & Aiba, S. (2021). Temperature is a dominant driver of distinct annual seasonality of leaf litter production of equatorial tropical rain forests. *Journal of Ecology*, *109*(2), 727–736. https://doi.org/10.1111/1365-2745.13500

Kitson, N. K., & Constantinou, A. C. (2021). Learning Bayesian networks from demographic and health survey data. *Journal of Biomedical Informatics*, *113*, 103588. https://doi.org/10.1016/j.jbi.2020.103588

Kontopantelis, E., Doran, T., Springate, D. A., Buchan, I., & Reeves, D. (2015). Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. *BMJ*, *350*(jun09 5), h2750–h2750. https://doi.org/10.1136/bmj.h2750

Kotoku, J., Oyama, A., Kitazumi, K., Toki, H., Haga, A., Yamamoto, R., Shinzawa, M., Yamakawa, M., Fukui, S., Yamamoto, K., & Moriyama, T. (2020). Causal relations of health indices inferred statistically using the DirectLiNGAM algorithm from big data of Osaka prefecture health checkups. *PloS One*, *15*(12), e0243229. https://doi.org/10.1371/journal.pone.0243229

Krich, C., Runge, J., Miralles, D. G., Migliavacca, M., Perez-Priego, O., El-Madany, T., Carrara, A., & Mahecha, M. D. (2020). Estimating causal networks in biosphere–atmosphere interaction with the PCMCI approach. *Biogeosciences*, *17*(4), 1033–1061. https://doi.org/10.5194/bg-17-1033-2020

Kunicki, Z. J., Smith, M. L., & Murray, E. J. (2023). A Primer on Structural Equation Model Diagrams and Directed Acyclic Graphs: When and How to Use Each in Psychological and Epidemiological Research. *Advances in Methods and Practices in Psychological Science*, *6*(2), 251524592311560. https://doi.org/10.1177/25152459231156085

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. https://doi.org/10.1073/pnas.1804597116

Kurotani, A., Miyamoto, H., & Kikuchi, J. (2024). Validation of causal inference data using DirectLiNGAM in an environmental small-scale model and calculation settings. *MethodsX*, *12*, 102528. https://doi.org/10.1016/j.mex.2023.102528

La Bastide-van Gemert, S., Stolk, R. P., Van Den Heuvel, E. R., & Fidler, V. (2014). Causal inference algorithms can be useful in life course epidemiology. *Journal of Clinical Epidemiology*, *67*(2), 190–198. https://doi.org/10.1016/j.jclinepi.2013.07.019

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. https://doi.org/10.1111/2041-210X.13190

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. https://doi.org/10.1098/rspb.2020.2815

Laube, S., & Zotz, G. (2003). Which abiotic factors limit vegetative growth in a vascular epiphyte? *Functional Ecology*, *17*(5), 598–604. https://doi.org/10.1046/j.1365-2435.2003.00760.x

Lechner, M. (2010). The Relation of Different Concepts of Causality Used in Time Series and Microeconometrics. *Econometric Reviews*, *30*(1), 109–127. https://doi.org/10.1080/07474938.2011.520571

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. https://doi.org/10.1111/2041-210X.12512

Lei, M.-K., Simons, R. L., Beach, S. R. H., & Philibert, R. A. (2019). Neighborhood Disadvantage and Biological Aging: Using Marginal Structural Models to Assess the

Link Between Neighborhood Census Variables and Epigenetic Aging. *The Journals of Gerontology: Series B*, *74*(7), e50–e59. https://doi.org/10.1093/geronb/gbx015

Leist, A. K., Klee, M., Kim, J. H., Rehkopf, D. H., Bordas, S. P. A., Muniz-Terrera, G., & Wade, S. (2022). Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences. *Science Advances*, *8*(42), eabk1942. https://doi.org/10.1126/sciadv.abk1942

Leng, S., Ma, H., Kurths, J., Lai, Y.-C., Lin, W., Aihara, K., & Chen, L. (2020). Partial cross mapping eliminates indirect causal influences. *Nature Communications*, *11*(1), 2632. https://doi.org/10.1038/s41467-020-16238-0

Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, *54*(4), 421–431. JSTOR.

Li, C., Shen, X., & Pan, W. (2024). Nonlinear Causal Discovery with Confounders. *Journal of the American Statistical Association*, *119*(546), 1205–1214. https://doi.org/10.1080/01621459.2023.2179490

Li, J., Liu, L., & Le, T. D. (2015). *Practical Approaches to Causal Relationship Exploration*. Springer International Publishing. https://doi.org/10.1007/978-3-319-14433-7

Li, J., Liu, L., Le, T. D., & Liu, J. (2020). Accurate data-driven prediction does not mean high reproducibility. *Nature Machine Intelligence*, *2*(1), 13–15. https://doi.org/10.1038/s42256-019-0140-2

Lipkovich, I., Ratitch, B., & Mallinckrodt, C. H. (2020). Causal Inference and Estimands in Clinical Trials. *Statistics in Biopharmaceutical Research*, *12*(1), 54–67. https://doi.org/10.1080/19466315.2019.1697739

Little, R. J. (2021). Missing Data Assumptions. *Annual Review of Statistics and Its Application*, *8*(1), 89–107. https://doi.org/10.1146/annurev-statistics-040720-031104

Lopez Bernal, J., Cummins, S., & Gasparrini, A. (2016). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal of Epidemiology*, dyw098. https://doi.org/10.1093/ije/dyw098

Luo, Y., Ogle, K., Tucker, C., Fei, S., Gao, C., LaDeau, S., Clark, J. S., & Schimel, D. S. (2011). Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications*, *21*(5), 1429–1442. https://doi.org/10.1890/09-1275.1

Luo, Y., Peng, J., & Ma, J. (2020). When causal inference meets deep learning. *Nature Machine Intelligence*, *2*(8), 426–427. https://doi.org/10.1038/s42256-020-0218-x

Luque-Fernandez, M. A., Schomaker, M., Rachet, B., & Schnitzer, M. E. (2018). Targeted maximum likelihood estimation for a binary treatment: A tutorial. *Statistics in Medicine*, *37*(16), 2530–2546. https://doi.org/10.1002/sim.7628

Maathuis, M. H., & Colombo, D. (2015). A generalized back-door criterion. *The Annals of Statistics*, *43*(3). https://doi.org/10.1214/14-AOS1295

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. *Journal of Applied Ecology*, *56*(2), 354–364. https://doi.org/10.1111/1365-2664.13289

MacDonald, A. J., & Mordecai, E. A. (2019). Amazon deforestation drives malaria transmission, and malaria burden reduces forest clearing. *Proceedings of the National Academy of Sciences*, *116*(44), 22212–22218. https://doi.org/10.1073/pnas.1905315116

Mandujano Reyes, J. F., Ma, T. F., McGahan, I. P., Storm, D. J., Walsh, D. P., & Zhu, J. (2025). Spatiotemporal Causal Inference With Mechanistic Ecological Models: Evaluating Targeted Culling on Chronic Wasting Disease Dynamics in Cervids. *Environmetrics*, *36*(2), e2901. https://doi.org/10.1002/env.2901

Markus, K. A. (2010). Structural Equations and Causal Explanations: Some Challenges for Causal SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *17*(4), 654–676. https://doi.org/10.1080/10705511.2010.510068

Markus, K. A. (2021). Causal effects and counterfactual conditionals: contrasting Rubin, Lewis and Pearl. *Economics and Philosophy*, *37*(3), 441–461. https://doi.org/10.1017/S0266267120000437

Marquet, P. A., Allen, A. P., Brown, J. H., Dunne, J. A., Enquist, B. J., Gillooly, J. F., Gowaty, P. A., Green, J. L., Harte, J., Hubbell, S. P., O'Dwyer, J., Okie, J. G., Ostling, A., Ritchie, M., Storch, D., & West, G. B. (2014). On Theory in Ecology. *BioScience*, *64*(8), 701–710. https://doi.org/10.1093/biosci/biu098

Matsuzaki, S. S., Suzuki, K., Kadoya, T., Nakagawa, M., & Takamura, N. (2018). Bottom-up linkages between primary production, zooplankton, and fish in a shallow, hypereutrophic lake. *Ecology*, *99*(9), 2025–2036. https://doi.org/10.1002/ecy.2414

McCracken, J. M., & Weigel, R. S. (2014). Convergent cross-mapping and pairwise asymmetric inference. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, *90*(6), 062903. https://doi.org/10.1103/PhysRevE.90.062903

McGoff, K., Mukherjee, S., & Pillai, N. S. (2012). *Statistical inference for dynamical systems: a review* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.1204.6265

Mielke, K. P., Schipper, A. M., Heskes, T., Zijp, M. C., Posthuma, L., Huijbregts, M. A. J., & Claassen, T. (2022). Discovering Ecological Relationships in Flowing Freshwater Ecosystems. *Frontiers in Ecology and Evolution*, *9*, 782554. https://doi.org/10.3389/fevo.2021.782554

Montesanto, G., & Cividini, S. (2017). A crossover design to assess feeding preferences in terrestrial isopods: A case study in a Mediterranean species. *Biologia*, *72*(2), 194–203. https://doi.org/10.1515/biolog-2017-0020

Mooij, J. M., Peters, J., Janzing, D., Zscheischler, J., & Schölkopf, B. (2016). Distinguishing Cause from Effect Using Observational Data: Methods and Benchmarks. *Journal of Machine Learning Research*, *17*(32), 1–102.

Morgan, S. L., & Winship, C. (2014). Matching Estimators of Causal Effects. In *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (2nd ed., pp. 140–187). Cambridge University Press. https://doi.org/10.1017/CBO9781107587991

Morgan, S. L., & Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (Second edition). Cambridge University press.

Moss, W. E., Binfet, J., Hall, L. E., Allen, S. E., Edwards, W. H., Jennings-Gaines, J. E., & Cross, P. C. (2025). The effectiveness of harvest for limiting wildlife disease: Insights from 20 years of chronic wasting disease in Wyoming. *Ecological Applications*, *35*(1), e3089. https://doi.org/10.1002/eap.3089

Nabi, R., & Shpitser, I. (2018). Fair Inference on Outcomes. *Proceedings of the ... AAAI Conference on Artificial Intelligence. AAAI Conference on Artificial Intelligence*, *2018*, 1931–1940.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, *82*(4), 591–605. https://doi.org/10.1111/j.1469-185X.2007.00027.x

Nandi, A., Glymour, M. M., Kawachi, I., & VanderWeele, T. J. (2012). Using Marginal Structural Models to Estimate the Direct Effect of Adverse Childhood Social Conditions on Onset of Heart Disease, Diabetes, and Stroke. *Epidemiology*, *23*(2), 223–232. https://doi.org/10.1097/EDE.0b013e31824570bd

Nichols, J. D., & Cooch, E. G. (2025). Predictive models are indeed useful for causal inference. *Ecology*, *106*(1), e4517. https://doi.org/10.1002/ecy.4517

Nicolaisen, O., Cuny, M., & Bolla, S. (2014). Factorial experimental designs as tools to optimize rearing conditions of fish larvae. *Aquaculture*, *422–423*, 253–260. https://doi.org/10.1016/j.aquaculture.2013.12.018

Nie, X., & Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, *108*(2), 299–319. https://doi.org/10.1093/biomet/asaa076

Noack, F., Larsen, A., Kamp, J., & Levers, C. (2022). A bird's eye view of farm size and biodiversity: The ecological legacy of the iron curtain. *American Journal of Agricultural Economics*, *104*(4), 1460–1484. https://doi.org/10.1111/ajae.12274

Nogueira, A. R., Pugnana, A., Ruggieri, S., Pedreschi, D., & Gama, J. (2022). Methods and tools for causal discovery and causal inference. *WIREs Data Mining and Knowledge Discovery*, *12*(2), e1449. https://doi.org/10.1002/widm.1449

Nova, N., Deyle, E. R., Shocket, M. S., MacDonald, A. J., Childs, M. L., Rypdal, M., Sugihara, G., & Mordecai, E. A. (2021). Susceptible host availability modulates climate effects on dengue dynamics. *Ecology Letters*, *24*(3), 415–425. https://doi.org/10.1111/ele.13652

Oehri, J., Schmid, B., Schaepman-Strub, G., & Niklaus, P. A. (2020). Terrestrial land-cover type richness is positively linked to landscape-level functioning. *Nature Communications*, *11*(1), 154. https://doi.org/10.1038/s41467-019-14002-7

Ohrens, O., Bonacic, C., & Treves, A. (2019). Non-lethal defense of livestock against predators: flashing lights deter puma attacks in Chile. *Frontiers in Ecology and the Environment*, *17*(1), 32–38. https://doi.org/10.1002/fee.1952

Olden, J. D., Lawler, J. J., & Poff, N. L. (2008). Machine Learning Methods Without Tears: A Primer for Ecologists. *The Quarterly Review of Biology*, *83*(2), 171–193. https://doi.org/10.1086/587826

Orava, P. J. (1973). Causality and state concepts in dynamical systems theory. *International Journal of Systems Science*, *4*(4), 679–690. https://doi.org/10.1080/00207727308920048

Paluš, M. (2007). From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics. *Contemporary Physics*, *48*(6), 307–348. https://doi.org/10.1080/00107510801959206

Paluš, M., Krakovská, A., Jakubík, J., & Chvosteková, M. (2018). Causality, dynamical systems and the arrow of time. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(7), 075307. https://doi.org/10.1063/1.5019944

Park, J., Buchholz, S., Schölkopf, B., & Muandet, K. (2023). A Measure-Theoretic Axiomatisation of Causality. *Thirty-Seventh Conference on Neural Information Processing Systems*. https://openreview.net/forum?id=sPLTQSf6GI

Paul, W. L. (2011). A causal modelling approach to spatial and temporal confounding in environmental impact studies. *Environmetrics*, *22*(5), 626–638. https://doi.org/10.1002/env.1111

Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, *82*(4), 669–688. https://doi.org/10.1093/biomet/82.4.669

Pearl, J. (1998). Graphs, Causality, and Structural Equation Models. *Sociological Methods & Research*, *27*(2), 226–284. https://doi.org/10.1177/0049124198027002004

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed). Cambridge University Press. https://doi.org/10.1017/CBO9780511803161

Pearl, J. (2010). An Introduction to Causal Inference. *The International Journal of Biostatistics*, *6*(2). https://doi.org/10.2202/1557-4679.1203

Pearl, J. (2012). The causal foundations of structural equation modeling. In *Handbook of structural Equation Modeling* (pp. 68–91). The Guilford Press.

Pearl, J. (2022). Causation and decision: On Dawid's "Decision theoretic foundation of statistical causality." *Journal of Causal Inference*, *10*(1), 221–226. https://doi.org/10.1515/jci-2022-0046

Pearl, J., & Bareinboim, E. (2014). External Validity: From Do-Calculus to Transportability Across Populations. *Statistical Science*, *29*(4). https://doi.org/10.1214/14-STS486

Pearson, C. E., Ormerod, S. J., Symondson, W. O. C., & Vaughan, I. P. (2016). Resolving large-scale pressures on species and ecosystems: propensity modelling identifies agricultural effects on streams. *Journal of Applied Ecology*, *53*(2), 408–417. https://doi.org/10.1111/1365-2664.12586

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. https://doi.org/10.1111/rssb.12167

Peters, J., Janzing, D., & Schölkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press.

Peters, J., Mooij, J. M., Janzing, D., & Schölkopf, B. (2014). Causal Discovery with Continuous Additive Noise Models. *Journal of Machine Learning Research*, *15*(58), 2009–2053.

Petersen, M. L., Porter, K. E., Gruber, S., Wang, Y., & van der Laan, M. J. (2012). Diagnosing and responding to violations in the positivity assumption. *Statistical Methods in Medical Research*, *21*(1), 31–54. https://doi.org/10.1177/0962280210386207

Petersen, M. L., & van der Laan, M. J. (2014). Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology (Cambridge, Mass.)*, *25*(3), 418–426. https://doi.org/10.1097/EDE.0000000000000078

Pfister, N., Bühlmann, P., & Peters, J. (2019). Invariant Causal Prediction for Sequential Data. *Journal of the American Statistical Association*, *114*(527), 1264–1276. https://doi.org/10.1080/01621459.2018.1491403

Pichler, M., & Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, *14*(4), 994–1016. https://doi.org/10.1111/2041-210X.14061

Pitcher, C. R., Burridge, C. Y., Wassenberg, T. J., Hill, B. J., & Poiner, I. R. (2009). A large scale BACI experiment to test the effects of prawn trawling on seabed biota in a closed area of the Great Barrier Reef Marine Park, Australia. *Fisheries Research*, *99*(3), 168–183. https://doi.org/10.1016/j.fishres.2009.05.017

Pynegar, E. L., Gibbons, J. M., Asquith, N. M., & Jones, J. P. G. (2021). What role should randomized control trials play in providing the evidence base for conservation? *Oryx*, *55*(2), 235–244. https://doi.org/10.1017/S0030605319000188

Ramsey, J., Glymour, M., Sanchez-Romero, R., & Glymour, C. (2017). A million variables and more: the Fast Greedy Equivalence Search algorithm for learning high-dimensional graphical causal models, with an application to functional magnetic resonance images. *International Journal of Data Science and Analytics*, *3*(2), 121–129. https://doi.org/10.1007/s41060-016-0032-z

Reygadas, Y., Jensen, J. L. R., Moisen, G. G., Currit, N., & Chow, E. T. (2020). Assessing the relationship between vegetation greenness and surface temperature through Granger causality and Impulse-Response coefficients: a case study in Mexico. *International Journal of Remote Sensing*, *41*(10), 3761–3783. https://doi.org/10.1080/01431161.2019.1711241

Richardson, T. (2003). Markov Properties for Acyclic Directed Mixed Graphs. *Scandinavian Journal of Statistics*, *30*(1), 145–157. https://doi.org/10.1111/1467-9469.00323

Richardson, T., & Spirtes, P. (2002). Ancestral graph Markov models. *The Annals of Statistics*, *30*(4). https://doi.org/10.1214/aos/1031689015

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rosenbaum, P. R. (2002). Constructing Matched Sets and Strata. In P. R. Rosenbaum, *Observational Studies* (pp. 295–331). Springer New York. https://doi.org/10.1007/978-1-4757-3692-2_10

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Rubin, D. B. (1978). Bayesian Inference for Causal Effects: The Role of Randomization. *The Annals of Statistics*, *6*(1), 34–58. JSTOR.

Rubin, D. B. (1980). Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment. *Journal of the American Statistical Association*, *75*(371), 591. https://doi.org/10.2307/2287653

Rubin, D. B. (2005). Causal Inference Using Potential Outcomes: Design, Modeling, Decisions. *Journal of the American Statistical Association*, *100*(469), 322–331. https://doi.org/10.1198/016214504000001880

Runge, J. (2018). Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *28*(7), 075310. https://doi.org/10.1063/1.5025050

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. https://doi.org/10.1038/s41467-019-10105-3

Runge, J., Gerhardus, A., Varando, G., Eyring, V., & Camps-Valls, G. (2023). Causal inference for time series. *Nature Reviews Earth & Environment*, *4*(7), 487–505. https://doi.org/10.1038/s43017-023-00431-y

Runge, J., Heitzig, J., Petoukhov, V., & Kurths, J. (2012). Escaping the Curse of Dimensionality in Estimating Multivariate Transfer Entropy. *Physical Review Letters*, *108*(25), 258701. https://doi.org/10.1103/PhysRevLett.108.258701

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. https://doi.org/10.1126/sciadv.aau4996

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 B: Biological Sciences*, *377*(1853), 20210159. https://doi.org/10.1098/rstb.2021.0159

Salditt, M., Eckes, T., & Nestler, S. (2024). A Tutorial Introduction to Heterogeneous Treatment Effect Estimation with Meta-learners. *Administration and Policy in Mental Health and Mental Health Services Research*, *51*(5), 650–673. https://doi.org/10.1007/s10488-023-01303-9

Sargeant, J. M., O’Connor, A. M., Totton, S. C., & Vriezen, E. R. (2022). Watch your language: An exploration of the use of causal wording in veterinary observational research. *Frontiers in Veterinary Science*, *9*. https://www.frontiersin.org/journals/veterinary-science/articles/10.3389/fvets.2022.1004801

Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, *65*(3–4), 579–616. https://doi.org/10.1007/BF01053745

Scheines, R. (1997). *An Introduction to Causal Inference*. Carnegie Mellon University. https://doi.org/10.1184/R1/6490904.V1

Scherrer, N., Bilaniuk, O., Annadani, Y., Goyal, A., Schwab, P., Schölkopf, B., Mozer, M. C., Bengio, Y., Bauer, S., & Ke, N. R. (2021). *Learning Neural Causal Models with Active Interventions* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2109.02429

Schochet, P. Z. (2013). Estimators for Clustered Education RCTs Using the Neyman Model for Causal Inference. *Journal of Educational and Behavioral Statistics*, *38*(3), 219–238. https://doi.org/10.3102/1076998611432176

Scholkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward Causal Representation Learning. *Proceedings of the IEEE*, *109*(5), 612–634. https://doi.org/10.1109/JPROC.2021.3058954

Schoolmaster, D. R., Zirbel, C. R., & Cronin, J. P. (2020). A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations. *Ecology*, *101*(8), e03070. https://doi.org/10.1002/ecy.3070

Schreiber, T. (2000). Measuring Information Transfer. *Physical Review Letters*, *85*(2), 461–464. https://doi.org/10.1103/PhysRevLett.85.461

Schuler, M. S., & Rose, S. (2017). Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *American Journal of Epidemiology*, *185*(1), 65–73. https://doi.org/10.1093/aje/kww165

Servedio, M. R., Brandvain, Y., Dhole, S., Fitzpatrick, C. L., Goldberg, E. E., Stern, C. A., Van Cleve, J., & Yeh, D. J. (2014). Not Just a Theory—The Utility of Mathematical Models in Evolutionary Biology. *PLoS Biology*, *12*(12), e1002017. https://doi.org/10.1371/journal.pbio.1002017

Shah, A., Ramanathan, A., Hayot-Sasson, V., & Stevens, R. (2023). Causal Discovery and Optimal Experimental Design for Genome-Scale Biological Network Recovery. *ArXiv*, arXiv:2304.03210v1.

Shahn, Z., Hernán, M. A., & Robins, J. M. (2023). A Formal Causal Interpretation of the Case-Crossover Design. *Biometrics*, *79*(2), 1330–1343. https://doi.org/10.1111/biom.13749

Shen, X., Ma, S., Vemuri, P., Simon, G., the Alzheimer's Disease Neuroimaging Initiative, Weiner, M. W., Aisen, P., Petersen, R., Jack, C. R., Saykin, A. J., Jagust, W., Trojanowki, J. Q., Toga, A. W., Beckett, L., Green, R. C., Morris, J., Shaw, L. M., Khachaturian, Z., Sorensen, G., … Fargher, K. (2020). Challenges and Opportunities with Causal Discovery Algorithms: Application to Alzheimer's Pathophysiology. *Scientific Reports*, *10*(1), 2975. https://doi.org/10.1038/s41598-020-59669-x

Shi, B., Mao, X., Yang, M., & Li, B. (2024). What, Why, and How: An Empiricist's Guide to Double/Debiased Machine Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4677153

Shi, J., Chen, L., & Aihara, K. (2022). Embedding entropy: a nonlinear measure of dynamical causality. *Journal of The Royal Society Interface*, *19*(188), 20210766. https://doi.org/10.1098/rsif.2021.0766

Shimizu, S. (2014). Lingam: Non-Gaussian Methods for Estimating Causal Structures. *Behaviormetrika*, *41*(1), 65–98. https://doi.org/10.2333/bhmk.41.65

Shimizu, S., Hoyer, P. O., Hyvärinen, A., & Kerminen, A. (2006). A Linear Non-Gaussian Acyclic Model for Causal Discovery. *Journal of Machine Learning Research*, *7*(72), 2003–2030.

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvärinen, A., Kawahara, Y., Washio, T., Hoyer, P. O., & Bollen, K. (2011). DirectLiNGAM: A Direct Method for Learning a Linear Non-Gaussian Structural Equation Model. *Journal of Machine Learning Research*, *12*(33), 1225–1248.

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.

Shpitser, I., & Pearl, J. (2008). Complete Identification Methods for the Causal Hierarchy. *Journal of Machine Learning Research*, *9*(64), 1941–1979.

Shrier, I., & Platt, R. W. (2008). Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology*, *8*, 70. https://doi.org/10.1186/1471-2288-8-70

Siegel, K. J., Larsen, L., Stephens, C., Stewart, W., & Butsic, V. (2022). Quantifying drivers of change in social-ecological systems: land management impacts wildfire probability in forests of the western US. *Regional Environmental Change*, *22*(3), 98. https://doi.org/10.1007/s10113-022-01950-y

Siegel, K. J., Macaulay, L., Shapero, M., Becchetti, T., Larson, S., Mashiri, F. E., Waks, L., Larsen, L., & Butsic, V. (2022). Impacts of livestock grazing on the probability of burning in wildfires vary by region and vegetation type in California. *Journal of Environmental Management*, *322*, 116092. https://doi.org/10.1016/j.jenvman.2022.116092

Simler-Williamson, A. B., & Germino, M. J. (2022). Statistical considerations of nonrandom treatment applications reveal region-wide benefits of widespread post-fire restoration action. *Nature Communications*, *13*(1), 3472. https://doi.org/10.1038/s41467-022-31102-z

Singer, R. S. (2022). Continued abuse of causal inference in studies of antimicrobial resistance: revisiting the confusion between ecological correlation and causation. *Journal of Global Antimicrobial Resistance*, *30*, 485–486. https://doi.org/10.1016/j.jgar.2022.05.007

Singh, N. K., & Borrok, D. M. (2019). A Granger causality analysis of groundwater patterns over a half-century. *Scientific Reports*, *9*(1), 12828. https://doi.org/10.1038/s41598-019-49278-8

Smokorowski, K. E., & Randall, R. G. (2017). Cautions on using the Before-After-Control-Impact design in environmental effects monitoring programs. *FACETS*, *2*(1), 212–232. https://doi.org/10.1139/facets-2016-0058

Sobel, M. E. (2006). What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference. *Journal of the American Statistical Association*, *101*(476), 1398–1407. https://doi.org/10.1198/016214506000000636

Sobel, M. E. (2009). Causal Inference in Randomized and Non-Randomized Studies: The Definition, Identification, and Estimation of Causal Parameters. In R. Millsap & A. Maydeu-Olivares, *The SAGE Handbook of Quantitative Methods in Psychology* (pp. 3–22). SAGE Publications Ltd. https://doi.org/10.4135/9780857020994.n1

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. https://doi.org/10.1371/journal.pcbi.1010302

Spirtes, P., Glymour, C. N., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed). MIT Press.

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. https://doi.org/10.1002/ecy.3942

Stone, R. (1993). The Assumptions on Which Causal Inferences Rest. *Journal of the Royal Statistical Society. Series B, Methodological*, *55*(2), 455–466.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting Causality in Complex Ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/science.1227079

Sun, J., & Bollt, E. M. (2014). Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings. *Physica D: Nonlinear Phenomena*, *267*, 49–57. https://doi.org/10.1016/j.physd.2013.07.001

Sun, J., Taylor, D., & Bollt, E. M. (2015). Causal Network Inference by Optimal Causation Entropy. *SIAM Journal on Applied Dynamical Systems*, *14*(1), 73–106. https://doi.org/10.1137/140956166

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L.-S. Young (Eds.), *Dynamical Systems and Turbulence, Warwick 1980* (Vol. 898, pp. 366–381). Springer Berlin Heidelberg. https://doi.org/10.1007/BFb0091924

Tárraga, J. M., Sevillano-Marco, E., Muñoz-Marí, J., Piles, M., Sitokonstantinou, V., Ronco, M., Miranda, M. T., Cerdà, J., & Camps-Valls, G. (2024). Causal discovery reveals complex patterns of drought-induced displacement. *iScience*, *27*(9), 110628. https://doi.org/10.1016/j.isci.2024.110628

Tchetgen Tchetgen, E. J., & VanderWeele, T. J. (2012). On causal inference in the presence of interference. *Statistical Methods in Medical Research*, *21*(1), 55–75. https://doi.org/10.1177/0962280210386779

Textor, J., Hardt, J., & Knüppel, S. (2011). DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology (Cambridge, Mass.)*, *22*(5), 745. https://doi.org/10.1097/EDE.0b013e318225c2be

Tian, J., & Pearl, J. (2002). A general identification condition for causal effects. *Eighteenth National Conference on Artificial Intelligence*, 567–573. https://doi.org/10.5555/777092

Tilman, D., Reich, P. B., & Knops, J. M. H. (2006). Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature*, *441*(7093), 629–632. https://doi.org/10.1038/nature04742

Tredennick, A. T., Hooker, G., Ellner, S. P., & Adler, P. B. (2021). A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, *102*(6), e03336. https://doi.org/10.1002/ecy.3336

Treves, A., Fergus, A. R., Hermanstorfer, S. J., Louchouarn, N. X., Ohrens, O., & Pineda-Guerrero, A. (2024). Gold-standard experiments to deter predators from attacking farm animals. *Animal Frontiers*, *14*(1), 40–52. https://doi.org/10.1093/af/vfad072

Ushio, M., Hsieh, C., Masuda, R., Deyle, E. R., Ye, H., Chang, C.-W., Sugihara, G., & Kondoh, M. (2018). Fluctuating interaction network and time-varying stability of a natural fish community. *Nature*, *554*(7692), 360–363. https://doi.org/10.1038/nature25504

van der Laan, M. J., & Rubin, D. (2006). Targeted Maximum Likelihood Learning. *The International Journal of Biostatistics*, *2*(1). https://doi.org/10.2202/1557-4679.1043

VanderWeele, T. J. (2015). *Explanation in causal inference: methods for mediation and interaction*.

VanderWeele, T. J. (2019). Principles of confounder selection. *European Journal of Epidemiology*, *34*(3), 211–219. https://doi.org/10.1007/s10654-019-00494-6

VanderWeele, T. J., Hawkley, L. C., Thisted, R. A., & Cacioppo, J. T. (2011). A marginal structural model analysis for loneliness: Implications for intervention trials and clinical practice. *Journal of Consulting and Clinical Psychology*, *79*(2), 225–235. https://doi.org/10.1037/a0022610

VanderWeele, T. J., & Hernán, M. A. (2013). Causal Inference Under Multiple Versions of Treatment. *Journal of Causal Inference*, *1*(1), 1–20. https://doi.org/10.1515/jci-2012-0002

VanderWeele, T. J., & Shpitser, I. (2011). A New Criterion for Confounder Selection. *Biometrics*, *67*(4), 1406–1413. https://doi.org/10.1111/j.1541-0420.2011.01619.x

Vonk, M. C., Malekovic, N., Bäck, T., & Kononova, A. V. (2023). Disentangling causality: assumptions in causal discovery and inference. *Artificial Intelligence Review*, *56*(9), 10613–10649. https://doi.org/10.1007/s10462-023-10411-9

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. https://doi.org/10.1080/01621459.2017.1319839

Wang, X., Du, Y., Zhu, S., Ke, L., Chen, Z., Hao, J., & Wang, J. (2021). Ordering-Based Causal Discovery with Reinforcement Learning. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, 3566–3573. https://doi.org/10.24963/ijcai.2021/491

Wang, X., & Sobel, M. E. (2013). New Perspectives on Causal Mediation Analysis. In S. L. Morgan (Ed.), *Handbook of Causal Analysis for Social Research* (pp. 215–242). Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_12

Wauchope, H. S., Amano, T., Geldmann, J., Johnston, A., Simmons, B. I., Sutherland, W. J., & Jones, J. P. G. (2021). Evaluating Impact Using Time-Series Data. *Trends in Ecology & Evolution*, *36*(3), 196–205. https://doi.org/10.1016/j.tree.2020.11.001

Weigel, C., Harden, S., Masuda, Y. J., Ranjan, P., Wardropper, C. B., Ferraro, P. J., Prokopy, L., & Reddy, S. (2021). Using a randomized controlled trial to develop conservation strategies on rented farmlands. *Conservation Letters*, *14*(4), e12803. https://doi.org/10.1111/conl.12803

Wermuth, N., & Lauritzen, S. L. (1990). On Substantive Research Hypotheses, Conditional Independence Graphs and Graphical Chain Models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *52*(1), 21–50. https://doi.org/10.1111/j.2517-6161.1990.tb01771.x

West, T. A. P., Caviglia-Harris, J. L., Martins, F. S. R. V., Silva, D. E., & Börner, J. (2022). Potential conservation gains from improved protected area management in the Brazilian

Amazon. *Biological Conservation*, *269*, 109526. https://doi.org/10.1016/j.biocon.2022.109526

Westreich, D., & Cole, S. R. (2010). Invited commentary: positivity in practice. *American Journal of Epidemiology*, *171*(6), 674–677; discussion 678-681. https://doi.org/10.1093/aje/kwp436

White, H., & Chalak, K. (2009). Settable Systems: An Extension of Pearl's Causal Model with Optimization, Equilibrium, and Learning. *Journal of Machine Learning Research*, *10*(61), 1759–1799.

White, H., Chalak, K., & Lu, X. (2011). Linking Granger Causality and the Pearl Causal Model with Settable Systems. In F. Popescu & I. Guyon (Eds.), *Proceedings of the Neural Information Processing Systems Mini-Symposium on Causality in Time Series* (Vol. 12, pp. 1–29). PMLR. https://proceedings.mlr.press/v12/white11.html

Wiik, E., Jones, J. P. G., Pynegar, E., Bottazzi, P., Asquith, N., Gibbons, J., & Kontoleon, A. (2020). Mechanisms and impacts of an incentive-based conservation program with evidence from a randomized control trial. *Conservation Biology*, *34*(5), 1076–1088. https://doi.org/10.1111/cobi.13508

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed). MIT Press.

Wu, J., Zhou, Y., Wang, H., Wang, X., & Wang, J. (2023). Assessing the Causal Effects of Climate Change on Vegetation Dynamics in Northeast China Using Convergence Cross-Mapping. *IEEE Access*, *11*, 115367–115379. https://doi.org/10.1109/ACCESS.2023.3325485

Wu, X., Sverdrup, E., Mastrandrea, M. D., Wara, M. W., & Wager, S. (2023). Low-intensity fires mitigate the risk of high-intensity wildfires in California's forests. *Science Advances*, *9*(45), eadi4123. https://doi.org/10.1126/sciadv.adi4123

Wu, Y., Zhang, L., & Wu, X. (2019). Counterfactual fairness: unidentification, bound and algorithm. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, 1438–1444.

Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating Heterogeneous Treatment Effects with Observational Data. *Sociological Methodology*, *42*(1), 314–347. https://doi.org/10.1177/0081175012452652

Yang, A. C., Peng, C.-K., & Huang, N. E. (2018). Causal decomposition in the mutual causation system. *Nature Communications*, *9*(1), 3378. https://doi.org/10.1038/s41467-018-05845-7

Ye, H., Deyle, E. R., Gilarranz, L. J., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, *5*(1), 14750. https://doi.org/10.1038/srep14750

Yongmei, D., & Yulian, L. (2024). Causal Linkage Effect on Chinese Industries via Partial Cross Mapping Under the Background of COVID-19. *Computational Economics*, *63*(3), 1071–1094. https://doi.org/10.1007/s10614-023-10408-0

Yu, Y., Chen, J., Gao, T., & Yu, M. (2019). DAG-GNN: DAG Structure Learning with Graph Neural Networks. In K. Chaudhuri & R. Salakhutdinov (Eds.), *Proceedings of the 36th International Conference on Machine Learning* (Vol. 97, pp. 7154–7163). PMLR. https://proceedings.mlr.press/v97/yu19a.html

Yuan, A. E., & Shou, W. (2022). Data-driven causal analysis of observational biological time series. *eLife*, *11*, e72518. https://doi.org/10.7554/eLife.72518

Zeldow, B., Lo Re, V., & Roy, J. (2019). A semiparametric modeling approach using Bayesian Additive Regression Trees with an application to evaluate heterogeneous treatment effects. *The Annals of Applied Statistics*, *13*(3), 1989–2010. https://doi.org/10.1214/19-AOAS1266

Zeng, Y., Cai, R., Sun, F., Huang, L., & Hao, Z. (2025). A Survey on Causal Reinforcement Learning. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(4), 5942–5962. https://doi.org/10.1109/TNNLS.2024.3403001

Zhang, C., Mohan, K., & Pearl, J. (2023). Causal Inference under Interference and Model Uncertainty. *2nd Conference on Causal Learning and Reasoning*. https://openreview.net/forum?id=TYKk9SWhke0

Zhao, A., & Ding, P. (2022). Regression-based causal inference with factorial experiments: estimands, model specifications and design-based properties. *Biometrika*, *109*(3), 799–815. https://doi.org/10.1093/biomet/asab051

Zheng, K., Yu, S., & Chen, B. (2024). CI-GNN: A Granger causality-inspired graph neural network for interpretable brain network-based psychiatric diagnosis. *Neural Networks*, *172*, 106147. https://doi.org/10.1016/j.neunet.2024.106147

Zheng, L., & Yin, W. (2023). Estimating and evaluating treatment effect heterogeneity: A causal forests approach. *Research & Politics*, *10*(1), 20531680231153080. https://doi.org/10.1177/20531680231153080

Zhu, S., Ng, I., & Chen, Z. (2020). Causal Discovery with Reinforcement Learning. *International Conference on Learning Representations*. https://openreview.net/forum?id=S1g2skStPB