Applying the structural causal model (SCM) framework for observational causal inference in ecology

Article *in* Ecological Monographs · September 2022

DOI: 10.1002/ecm.1554

2 authors:

Suchinta Arif
Dalhousie University

Aaron Macneil
Dalhousie University

Arif Suchinta (Orcid ID: 0000-0001-8381-3071)

Applying the structural causal model (SCM) framework for observational causal inference in ecology

Suchinta Arif and M Aaron MacNeil

Ocean Frontier Institute, Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada, B3H 4R2

Corresponding author: Suchinta Arif. E-mail: suchinta.arif@gmail.com

Open Research Statement: Code for simulations and statistical analysis (Arif, 2022) is available on Figshare at https://doi.org/10.6084/m9.figshare.19541059.v1.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ecm.1554

Abstract:

Ecologists are often interested in answering causal questions from observational data but generally lack the training to appropriately infer causation. When applying statistical analysis (e.g., generalized linear model) on observational data, common statistical adjustments can often lead to biased estimates between variables of interest due to processes such as confounding, overcontrol, and collider bias. To overcome these limitations, we overview the structural causal model (SCM), an emerging causal inference framework that can be used to determine cause and effect relationships from observational data. The SCM framework uses directed acyclic graphs (DAGs) to visualize a researchers' assumptions about the causal structure of a system or process under study. Following this, a DAG-based graphical rule known as the backdoor criterion can be applied to determine statistical adjustments (or lack thereof) required to determine causal relationships from observational data. In the presence of unobserved confounding variables, an additional rule called the frontdoor criterion can be employed to determine causal effects. Here, we use simulated ecological examples to review how the backdoor and frontdoor criteria can return accurate causal estimates between variables of interest, as well as how biases can arise when they are not employed. We further provide an overview of studies that have applied the SCM framework in ecology. SCM and its application of DAGs have been broadly employed in other disciplines to make valid causal inference from observational data. Their use in ecology holds tremendous potential for quantifying causal relationships and investigating a range of ecological questions without randomized experiments.

Key Words: backdoor criterion, causal inference, directed acyclic graphs (DAGs), frontdoor criterion, observational study, statistical ecology, structural causal model (SCM)

Introduction:

Observational studies in ecology rely on data that have not been experimentally manipulated and are commonly used to understand ecological patterns and processes seen in nature (Carmel et al. 2013). Observational approaches are increasing in relevance due to the emergence of large-scale ecological questions that are not easily manipulated or controlled, such as invasive species and the consequences of climate change. New advances in technology, such as remote-sensing, environmental genetics, and animal-borne sensors, as well as increased availability of data online and from citizen science, have enhanced opportunities to answer previously intractable ecological questions using observational data (Sagarin and Pauchard 2010).

Many observational studies in ecology are aimed at answering causal questions, such as the impact of marine protected areas on fishing communities (Mascia et al. 2010) or the effect of forest fragmentation on species richness (Sam et al. 2014). However, causal inference – the leveraging of theory and deep knowledge to estimate the impact of events, choices or other factors on a given outcome of interest (Cunningham 2021) – is rare. Yet without the consideration of causal relationships, statistical analysis can frequently lead to biased estimates (i.e., estimates that differ from the true parameter being estimated) that undermine ecological inferences by providing non-causal correlations among variables of interest (e.g., see Appendix S1). This is the basis of the often-repeated phrase "correlation does not imply causation" (F.A.D.

1900). We believe that increasing the use of causal inference methods in observational ecology will reduce bias throughout the discipline and lead to more accurate assessments across a range of ecological questions, especially when experimental approaches are unfeasible.

Structural causal modelling (SCM, Pearl 2009) is an emerging causal inference framework, which unifies the strong features of structural equation modeling (SEM; Wright 1921, Shipley 2016) and Rubin's potential outcome framework (PO; Rubin 2005) among others, to create a powerful theory of causation and framework for causal inference. Importantly, this framework can be used to determine cause and effect relationships from observational data, without needing to set up randomized control experiments (Pearl 2009). SCM has been widely employed across other disciplines, including econometrics (Imbens 2020), epidemiology (Pearce and Lawlor, 2016), paediatrics (Williams et al. 2018) and psychology (Rohrer, 2018), as well as a few ecological studies (Cronin and Schoolmaster 2018; Schoolmaster et al. 2020; Schoolmaster et al. 2022; Arif et al. 2022; Arif and MacNeil 2022). It holds tremendous potential for increasing the use of causal inference across observational ecological studies.

Under the SCM framework, the derivation of causal effects rests on a set of causal assumptions about the data generating process (e.g., X effects Y and not the other way around). These causal assumptions are visualized using directed acyclic graphs (DAGs), which represent a researchers' assumptions about the causal structure of a system or process under study (Pearl 2009; Morgan and Winship 2014). Given a DAG, a graphical rule known as the backdoor criterion determines the sufficient sets of variables for adjustment required to determine causal effects from observational data. When the backdoor criterion cannot be employed – due to the presence of an unobserved confounding variable – a second graphical rule called the frontdoor criterion can be employed. Using simulated ecological examples with specified (i.e., known) causal effects, we define these criteria and review how they can be employed to determine causal effects between variables of interest.

To date, the few ecological studies that have employed the SCM framework have identified key causal relationships across study systems (Cronin and Schoolmaster 2018; Schoolmaster et al. 2020; Arif et al. 2022), outlined steps required for observational causal inference (Cronin and Schoolmaster 2018; Arif et al. 2022), clarified SCM theory (Schoolmaster et al. 2022), and highlighted the utility of SCM for experimental and quasi-experimental approaches (Schoolmaster et al. 2020; Schoolmaster et al. 2022; Arif and MacNeil 2022). However, these studies can be niche topics and theoretically complex. Here, we provide an easily accessible overview of the SCM framework, highlighting two key tools – the backdoor and frontdoor criteria – that can be used for causal inference across observational ecological studies.

Directed Acyclic Graphs (DAGs):

DAGs are used to represent causal relationships within a given system. A DAG consists of a set of nodes (variables) that are connected to each other by edges (arrows). These arrows represent causal relationships between variables, pointing from cause to effect, with causes preceding their effects. For example, the DAG in Fig 1 shows that X directly effects Y (X → Y), W directly affects both X (W → X) and Y (W → Y), and W indirectly effects Y through X (W → X → Y). It is important to note that the arrows between nodes (variables) represent hypothesized causal relationships (i.e., a lack of causal relationship can be found following a SCM analysis). On the other hand, a lack of arrow between two nodes assumes no causal relationship between variables, representing strong *a priori* causal assumptions. Therefore, missing arrows encode causal assumptions, whereas arrows between nodes represent the possibility of an effect (Elwert, 2013).

A key characteristic of DAGs is that they must be *acyclic*, meaning that they cannot contain bi-directional relationships (i.e. arrows need to be unidirectional) or a feedback loop where a variable causes itself (Glymour and Greenland 2008, Elwert 2013). This limits the application of DAGs to ecological systems that do not contain bi-directionality and or feed-back loops. However, one way to resolve this issue is to articulate the temporal sequence of events more finely (Greenland et al. 1999). For example, if temperature at time one ($Temp_{t1}$) effects ice cover, which then influences temperature at time two ($Temp_{t2}$), $Temp_{t1}$ and $Temp_{t2}$ can be represented as separate nodes within a DAG, without violating acyclic requirements. For interested readers, Schoolmaster et al. (2020) provide a published ecological DAG that incorporates the temporal sequence of events (see their Appendix S2).

DAGs are also non-parametric, meaning that they do not make any assumptions about the stochastic nature of variables or their observation, or the functional form of direct effects (e.g., linear, nonlinear, stepwise) and their effect size (Glymour and Greenland 2008). In this sense, a DAG is qualitative: X -> Y only communicates that X causally affects Y in some way, without specifying any other restrictions. This non-parametric nature of DAGs makes them compatible with a wide range of ecological systems.

Using DAGs under the SCM Framework:

DAGs are central to the SCM framework as they are used to visualize and quantify causal relationships from observational data (Pearl 2009). Fig 2 summarizes the SCM framework which includes creating a DAG (step 1), testing a DAG to ensure DAG-data consistency (step 2), applying either the backdoor or frontdoor criterion (step 3), choosing an appropriate statistical model (step 4), and making inference by quantifying a causal effect (step 5). As we walk through our review, we will follow the workflow in Fig 2 using simulated ecological examples interspersed with relevant theory and background information.

Step 1: Creating a DAG:

DAGs represent a researcher's causal assumptions about the data generating process of a system of process under study (Pearl 2009, Morgan and Winship 2014). As such, researchers should ensure that their DAG represents the complete causal structure of the system or process, including all relevant measured and unmeasured variables, as well as all common causes of any pair of variables included in the DAG (Sprites et al. 2001, Glymour and Greenland 2008). DAGs should also be rigorously justified based on domain knowledge, theory, and research. A combination of background information including experimental data, past literature, and domain knowledge can be used to create DAGs of ecological systems. For example, Ethier and Nudds (2017) gathered information from published literature and local stakeholder knowledge to create

DAGs depicting factors affecting population dynamics of bobolink (*Dolichonyx oryzivorus*). In another study, Cronin and Schoolmaster (2018) synthesized past literature to create a DAG representing the causes of trait covariation. Expert opinion can also be elicited to generate DAGs. To ensure credibility and transparency, researchers should apply formal methods for surveying experts, which has been developed within the ecological literature (e.g., Choy et al. 2009, Kuhnert et al. 2010, Martin et al. 2012, Drescher et al. 2013), including for the development of causal diagrams (e.g., Marcot et al. 2006, McNay et al. 2006). For example, Marcot et al. (2006) show how to use expert review to create their DAG on the probability of capture of northern flying squirrels.

As a general ecological example, Fig 3 presents a DAG adapted from Adams et al. (2015), showing how different factors are expected to influence forest species abundance across a hypothetical region (Step 1, Fig 2). Here, protected areas are shown to effect forest species abundance through three intermediate processes: fire, poaching and logging (Adams et al. 2015). Other variables including distance to roads and cities, slope, and elevation effect both protected areas placement (protected areas are often placed in high and far places; Joppa et al. 2009) and forest species abundance through their effects on fire, poaching and/or logging (Adams et al. 2015). We have created a simulated dataset, matching the causal structure of this DAG (Arif, 2022). We will use this DAG and simulated dataset to work through the rest of the SCM workflow (Steps 2-5, Fig 2). Specifically, we will aim to answer how protected areas, fire, logging and poaching each effect forest species abundance. Because our simulated data was created with specified (i.e., known) causal effects, we can use it to show how the SCM framework can return accurate causal estimates.

Step 2: Test DAG-Data Consistency:

Once a DAG has been created, it can be tested against observational data to check for DAG-data consistency. Simply put, a DAG often asserts multiple independencies that should hold in the observational data, given that both the DAG and observational data are representative of the data generating process. Given a DAG, a pair of variables can be independent of each other (e.g., X is independent of Y) if there are no paths (i.e., a sequence of nodes and arrows) connecting them. As well, a pair of variables can be conditionally independent. Conditional independencies emerge from **d-separation** (dependency separation; Pearl 1988), a graphical rule for deciding whether a variable X is independent of another variable Y, given a set of variable(s), Z in a path.

> **d-separation** (Pearl, 1988): A set of variables, Z, is said to block (or d-separate) a path from one variable to another if either
>
> (i) the path contains at least one arrow-emitting variable that is in Z, or
>
> (ii) the path contains at least one collider variable (variable with two incoming arrows) that is outside Z and does not cause any variables in Z
>
> If all paths between X and Y are blocked (or d-separated) by Z, then X and Y are independent given Z, written $X \perp Y | Z$. For a more detailed discussion of d-separation, readers can reference Shipley (2000) and Shipley (2016) which discuss d-separation within an ecological context.

DAG-data consistency requires that all implied independencies for a given DAG (including conditional independencies based on d-separation rules) are consistent with the observational dataset. For example, in a simplified DAG, X → Z → Y, X is independent of Y, given Z (an arrow-emitting variable that d-separates the path from X to Y). Therefore, the associated observational data should show that X is independent of Y when Z is adjusted for. Often a DAG will hold many independencies and these independencies can be tested against a dataset to ensure DAG-data consistency. If all implied independencies within a DAG coincide with the dataset, then this supports DAG accuracy. However, if at least one implied independency is refuted (i.e., does not match the data), then the DAG is not consistent with the data and would need to be adjusted.

For our DAG (Fig 3), there are 28 independencies that can be tested against our simulated data to ensure DAG-data consistency (Appendix S2: Section S1). In an observational study, we would test these independencies against observational data. Here, we proceed by testing DAG-data consistency using our simulated dataset, to walk readers through the process. Specifically, we use the R package 'dagitty', which provides a user-friendly way to evaluate whether a DAG is consistent with a dataset, even when DAGs become increasingly complex and include many variables (Textor et al., 2016). Dagitty uses a formal test of zero correlation to test whether each identified independency of a specified DAG is consistent with a given dataset (see Textor et al. 2016 for details). Using dagitty, we tested DAG-data consistency and found that all 28 independencies were consistent with our simulated dataset (Step 2, Fig 2; see Arif, 2022 for R code). This is expected as our simulated data was created to match the causal structure of our DAG.

In real world applications, a DAG may require a series of adjustments until DAG-data consistency is reached. As an ecological example, Schoolmaster et al. (2020) provide a real-world example of a DAG used to understand the relationship between tree species composition and canopy cover. Their initial DAG failed DAG-data consistency and was subsequently updated using a combination of domain knowledge and results from failed independence tests (Schoolmaster et al. 2020). Anken et al. (2021) further provide general examples and guidelines on updating DAGs based on DAG-data consistency, using the R package 'dagitty'. Importantly, they note that this process should be handled with care and always supported by domain knowledge. Failed independence tests are not necessarily proof that a DAG is incorrect; they can also indicate problems with the data (e.g., if the collected data does not represent the data-generating process). Ultimately, there should be a firm theoretical basis for creating and revising DAGs.

Once a DAG has been sufficiently justified and tested and updated based on DAG-data consistency, the backdoor (or frontdoor) criterion can be employed (Step 3, Fig 2). Before moving on to application of backdoor and frontdoor criteria, we briefly review why they can be applied to DAGs to determine causal effects from observational data.

DAGs for causal effects

Causal effects describe to what extent a predictor variable X (i.e., the cause) influences a response variable Y (i.e., an effect). The SCM framework uses counterfactual reasoning to determine the causal effect of X on Y (Pearl 2009). A counterfactual represents the potential outcome that would be realized if a predictor variable X was set to a different value, i.e., X=x. Specifically, a counterfactual for response variable Y is noted as Yx(u), which represents the value of (outcome) Y, had (predictor) X been x in unit (or situation) U = u (Rubin 2005, Morgan and Winship 2014). This counterfactual Yx(u) is represented by the equation:

$$Y_x(\mathrm{u}) \triangleq Y_{Mx}(\mathrm{u}) \qquad [1]$$

Under the SCM framework, a DAG represents a structural model, M. In equation 1, M_x stands for a modified version of a model M, where X is intervened upon (*i.e.,* "if X had been x", X=x). Graphically M_x is represented by a modified DAG, where the arrows pointing into X are eliminated. Equation 1 states that the counterfactual Yx(u) is the solution for Y in the modified model M_x (see Galles and Pearl 1998 for axiom of Eq 1).

This definition of counterfactuals can be used to predict the effect of interventions from observational data alone. Under the SCM framework, interventions are denoted by what's known as the do-operator, written do() (Pearl 1995, 2009). For example, the query Q = P(y|do(x)) asks what the distribution of Y would be, if X is set to a particular value of x (*i.e.,* the causal effect of X on Y). Related to Eq 1, this can be defined as

$$P(y|do(x)) \triangleq P_{Mx}(y) \; [2]$$

showing that the distribution of outcome Y (if X is set to a particular value of x) is equal to the distribution of Y in the modified model M_x (Pearl 1995, 2009).

Given that we do not have post-interventional data (following the distribution of M_x), the question becomes whether the query Q = P(y|do(x)) can be estimable from observational data (following the distribution of M) and the set of causal assumptions represented by its associated DAG. When a query includes a do-expression, an algebraic procedure known as do-calculus (Pearl 1995) can be used to *equate* post-interventional distributions (those represented in M_x) to pre-interventional (or observational) distributions (those represented in M). To identify an interventional query, *e.g.,* Q = P(y|do(x), the inference rules of do-calculus (outlined in Pearl 1995) need to be repeatedly applied until an expression is obtained that no longer contains a do-operator. If this can be done, then the post interventional query is estimable from observational data. While the application of do-calculus makes for challenging reading, based on its derived inference rules, Pearl created the *backdoor criterion* and the *frontdoor criterion*, which are two DAG-based graphical rules that can be applied to estimate interventional queries from observational data (i.e., the causal effect of X on Y), without the need for do-calculus operations.

Step 3 (Option 1): Apply Backdoor Criterion:

The backdoor criterion (Pearl 1993, Pearl 2009) is used to identify a set of variables, Z, that when adjusted for, allows the post-interventional query Q = P(y|do(x)) to be accurately estimated from observational data. The backdoor criterion states that a set of variables, Z, is sufficient for estimating the causal effect of X on Y under two conditions:

1. The variables in Z block all *backdoor paths* from X to Y. A *path* within a DAG is any sequence of arrows and nodes connecting two variables of interest, X and Y, regardless of direction. A *backdoor path* is a path between X and Y with an arrow pointing from Y and an arrow pointing into X. Backdoor paths create bias by providing one or more indirect, non-causal pathways through which information can leak from one variable through another, leading to spurious correlation. To block a backdoor path from X to Y, the backdoor path from X to Y must be d-separated. Again, the rules for d-separation are:

   d-separation (Pearl 1988): A set of variables, Z, is said to block (or d-separate) a path from one variable to another if either

   (i) the path contains at least one arrow-emitting variable that is in Z, or

   (ii) the path contains at least one collider variable (variable with two incoming arrows) that is outside Z and does not cause any variables in Z

2. No element of Z is a descendant of (*i.e.,* caused by) X.

When applied, the backdoor criterion blocks all non-causal pathways between a predictor and response variable of interest, while leaving all causal paths open. As such, the application of backdoor criterion eliminates common statistical biases that can otherwise plague observational studies, including confounding, overcontrol, and collider bias. Appendix S1 defines each of these biases and shows how the backdoor criterion removes each of them. The main takeaway is that given a DAG, the application of the backdoor criterion will avoid all three biases, allowing for causal estimates to be made.

Given our DAG (Fig 3), we can use the backdoor criterion to determine the sufficient set for adjustment required to answer our causal questions (Step 3, Fig 2). For example, if we want to quantify the causal effect of protected area on forest species abundance, there are nine backdoor paths that need to be blocked (i.e., d-separated):

1. Forest Species Abundance → Carbon Sequestration ← Logging ← Elevation → Protected Area

2. Forest Species Abundance → Carbon Sequestration ← Logging ← Slope → Protected Area

3. Forest Species Abundance ← Fire ← Distance to Roads and Cities → Logging ← Elevation → Protected Area

4. Forest Species Abundance ← Fire ← Distance to Roads and Cities → Logging ← Slope → Protected Area

5. Forest Species Abundance ← Logging ← Elevation → Protected Area

6. Forest Species Abundance ← Logging ← Slope → Protected Area

7. Forest Species Abundance ← Poaching ← Distance to Roads and Cities → Protected Area

8. Forest Species Abundance ← Logging ← Distance to Roads and Cities → Protected Area

9. Forest Species Abundance ← Fire ← Distance to Roads and Cities → Protected Area7.

The first four backdoor paths are already blocked because we have not adjusted for a collider variable (i.e., a variable with two incoming arrows: → X ←) in each of these four paths. Specifically, carbon sequestration acts as a collider variable in backdoor paths 1 and 2, and logging acts as a collider in backdoor paths 3 and 4. The remaining backdoor paths do not contain collider variables and must be blocked by adjusting for an arrow-emitting variable that isn't a descendent of (i.e., caused by) protected area, our predictor variable. As such, path 5 can be blocked by adjusting for elevation, path 6 can be blocked by adjusting for slope, and paths 7-9 can all be blocked by adjusting for distance to roads and cities. Collectively, the causal effect of protected area on forest species richness, given this DAG can be quantified by adjusting for slope, elevation and distance to roads and cities.

Given that application of the backdoor criterion can rapidly become difficult to keep track of for increasingly complex DAGs, researchers are encouraged to draw out their DAG on www.daggity.net (instructions within site), which will apply the backdoor criterion and generate the minimal sufficient adjustment set(s) required to determine causal effects, given a specified DAG and causal question. As an example, readers can visit dagitty.net/m18S_bV to work with our protected area DAG. Using this website (see Appendix S2: Section S2 for quick steps), to determine the causal effect of fire on forest species abundance, we can adjust for either (distance to roads and cities and protected area) or (logging and poaching). To determine the causal effect of poaching on forest species abundance we can adjust for either (distance to roads and cities and protected area) or (fire and logging). Last, to determine the causal effect of logging on forest species richness we can adjust for either (distance to roads and cities and protected area) or (fire and poaching). When there are multiple options for a sufficient adjustment set based on the backdoor criterion, researchers can choose a set based on data availability and measurement error. If known, it is best to select the set where variables are measured most accurately.

We note that given our DAG and linear simulated data, causal effects between variables of interest could also be determined using alternative methods such as SEM. However, a strength of the backdoor criterion is that it can allow causal estimation without requiring the availability of all variables in a DAG (Pearl 2009). For example, the effect of protected area on forest species abundance requires observational data on only variables for protected area, forest species abundance, slope, elevation, and distance to roads and cities. By only including variables necessary for answering specific causal queries, this can further enhance estimation accuracy by reducing researchers' reliance on noisy and irrelevant data (MacDonald 2004). In addition, the application of the backdoor criterion does not require lengthy algebraic manipulations, isn't computationally taxing and is compatible across linear and non-parametric statistical approaches (Pearl 2009). Ultimately, it provides ecologists with a widely applicable method for covariate selection across observational studies.

Step 4: Choose a Statistical Model

Once the backdoor criterion is used to determine the sufficient set(s) for adjustment, researchers must decide on an appropriate statistical model to carry out their causal analysis. Since our simulated data was created with a linear causal structure, we have chosen linear regression models for analysis (Step 4, Fig 2). However, it is up to each researcher to decide what form of analysis will best suit their data. As DAGs are non-parametric, they make no assumptions about the distribution of variables (e.g., normal) or the functional form of effects (e.g., linear, nonlinear, stepwise), making them compatible with a wide range of statistical methods. DAGs are also compatible with both frequentist and Bayesian statistical approaches since they are used to determine the sufficient set(s) for adjustment, and not the analysis itself. Statistical models developed under the SCM framework are still beholden to the same issues of sample size and measurement error in terms of the precision of resulting estimates; however, they are based on causal reasoning.

Step 5: Causal Effect

Fig 4 shows that when the backdoor criterion was used to determine the sufficient set for adjustment, our linear regression models were able to correctly estimate the causal effect between selected predictor variables and forest species abundance, our response of interest (Step 5, Fig 2; see Arif, 2022 for R code). This is achieved because the backdoor criterion blocks all non-causal pathways (i.e., backdoor paths) between our predictor and response variable of interest, while leaving all causal paths open. By adjusting for specific variables (if necessary) to answer specific causal questions, the backdoor criterion can guide causal inference in observational settings.

Importantly, in performing a causal analysis we are not trying to find a 'best model' of the data according to criteria of model fit such as AIC, which seek to find the model with the greatest predictive support, regardless of potential biases present in estimated effect sizes (Burnham and Anderson 2002, Grace and Irvine 2020). For example, in Figure 5 we include a 'causal salad' model (McElreath 2020) typical of ecological observational studies (including our own past work), whereby all available variables thought to affect a response are thrown into one statistical model and subsequently interpreted, without directly addressing the causal structure of the system. In our simulated example, the causal salad model (Fig 4) is strongly favored over all other models by AIC, yet it provides an entirely inaccurate picture of the causal structure in the system. Under this approach, we obtain inaccurate estimates of our predictor variables of interest (Fig 4). For example, the estimated effect of protected area on forest species abundance is negligible due to overcontrol bias (see Appendix S1) occurring from the inclusion of fire, poaching and logging, which are intermediate variables between the predictor and response variable of interest. Effect sizes for fire, poaching and logging are also biased due to the inappropriate inclusion of carbon sequestration, which is not a predictor variable but is instead influenced by our response variable of interest. Collectively, these results demonstrate the general principle that the models used for causal inference must be carefully built to consider relevant causal relationships within a system prior to analysis. It also directly undermines 'variance explained' as a modelling objective or arbiter of truth – without causal thinking to support modelling decisions, it is easy to add variables that seem to represent a better model according to a range of widely-used statistical criteria. In this, the backdoor criterion can play a critical role in model development that stands apart from typical model-selection methods, by determining the sufficient set(s) for adjustment required for causal inference.

The Front Door Criterion:

The DAG-based approach to causal models up to this point has assumed we have observational data on all variables needed to satisfy the backdoor criterion. However, in some circumstances, there may be a known but unobserved variable that confounds our results, preventing application of the backdoor criterion for determining causal effects. For example, if we want to determine the causal effect of X on Y for the DAG in Fig 5, the backdoor criterion instructs us to adjust for U. However, U is unobserved, and therefore cannot be used as a covariate in our final model. In such cases, an approach called the frontdoor criterion can be employed for causal inference (Pearl 1995, Pearl 2009). To quantify the effect of X on Y in the presence of unobserved confounders, a variable Z satisfies the frontdoor criterion if:

1. Z blocks all directed paths from X to Y
2. There are no unblocked paths from X to Z
3. X blocks all backdoor paths from Z to Y

Once a Z variable is identified, the causal effect of X on Y can be determined by first employing the backdoor criterion to separately determine the effect of X on Z and Z on Y (Fig 5). The product of these two causal effects (i.e., point estimates) then becomes the effect of X on Y (Pearl 1995; 2009). Below we show how to apply the front door criterion to determine the effect of sharks on rays based on a hypothetical ecological example.

Step 1: Create a DAG

The DAG in Fig 6a asserts that sharks effect rays, which in turn effect bivalves, through a top-down trophic cascade which has previously been supported (Myers et al. 2007, Buam and Worm 2009) and refuted (Grubbs et al. 2016) in the literature. In our hypothetical scenario, we also assert that fishing pressure effects both sharks and bivalves, but not rays. Here, observational data on fishing pressure isn't available, making it an unobserved variable. Like our prior example, we created a simulated dataset (with known causal effects) matching our DAG (see Arif, 2022 for R code) to demonstrate the use of the frontdoor criterion. Specifically, we will show how to employ the frontdoor criterion to return the causal effect of sharks on bivalves, which we have set to 0.02.

Step 2: Test DAG-data Consistency

Given the DAG in Fig 6a, there are two independencies that can be tested based on d-separation rules: 1) fishing pressure is independent of rays, given sharks and 2) sharks are independent of bivalves, given fishing pressure and rays. However, testing either independency requires observational data on fishing pressure (our unobserved variable). Therefore, due to our unobserved confounding variable, DAG-data consistency cannot be tested based on d-separation rules in this case. However, we can still apply the frontdoor criterion for causal estimates with our asserted DAG (unchecked for DAG-data consistency).

Step 3 (Option 2): Apply Frontdoor Criterion

The frontdoor criterion can be employed to find the effect of sharks on bivalves. Rays satisfy the frontdoor criterion since (1) they block all directed paths from sharks to bivalves, (2) there are no unblocked backdoor paths from sharks to rays, and (3) all backdoor paths from rays to bivalves are blocked by sharks (see rules for frontdoor criterion above). To determine the effect of sharks on bivalves, we first need to apply the backdoor criterion to determine the effect of sharks on rays (which can be estimated without any adjustments), and the effect of rays on bivalves (which can be estimated by adjusting for sharks). Both sub-models can employ the backdoor criterion without needing to adjust for fishing pressure (our unobserved variable). The causal effect of sharks on bivalves can then be estimated by multiplying the effect of sharks on rays by the effect of rays on bivalves.

Step 4: Choose a Statistical Model

We use linear regression models because our simulated data was created using linear relationships.

Step 5: Causal Effect

Fig 6 shows that when the frontdoor criterion is employed, we were able to accurately determine the causal effect of sharks on bivalves (see Arif, 2022 for R code). Specifically, the product of the effect of sharks on rays (Fig 6b) and the effect of rays on bivalves (Fig 6c) gave us an accurate causal estimate of sharks on bivalves (0.02), without having to adjust for fishing pressure, our unobserved confounding variable. In contrast, a model with just rays regressed on sharks gives a misleading estimate of 0.99. Here, the correlation between sharks and rays is spurious due to the confounding effect of our unobserved fishing pressure variable.

The front door criterion is not as widely applicable to ecological data as the backdoor criterion, given that it requires a specific causal structure, specified by its three rules (see above). However, in cases where these rules are met, the frontdoor criterion can provide causal estimates, regardless of the strength of unobserved confounding. As well, it can be employed in the presence of multiple unobserved confounding variables.

Examples of SCM in Ecology

Although currently underutilized, the SCM framework and its application of DAGs has been used to understand the causal structure of ecological systems. Here, we provide an overview of three recent applications of Pearl's SCM framework in ecology.

What causes climate-induced coral reef regime shifts?

Climate-induced bleaching events can sometimes lead to coral reef regime shifts, whereby the benthic composition of a coral reef ecosystem abruptly transitions from one dominated by coral to one dominated by macroalgae (Bellwood et al. 2004). However, not all coral reefs regime shift following a bleaching event, and it is expected that certain factors may influence the likelihood of climate-induced regime shifts. Although past correlative studies have found correlations between key predictor variables (e.g., depth, structural complexity; Graham et al. 2015) and regime shift trajectory, these studies were not grounded in causal inference. For example, a literature review of observational coral reef regime shift studies showed that no studies to date employed causal inference methods, though they often used causal language to communicate their results (e.g., "the effect of"; Arif et al. 2022). Instead, these studies either used a causal salad approach or did not include any covariate adjustment, without communicating the overall causal structure of the system under study (Arif et al. 2022).

To overcome these limitations, Arif et al. (2022) applied the SCM framework to understand how different factors influenced regime shift trajectory following the 1998 bleaching event across Seychelles coral reefs. They created a DAG depicting the causal structure of how factors are expected to influence regime shift trajectory in Seychelles, based on expert opinion and scientific literature (Fig 7). Given their DAG, they applied the backdoor criterion to determine if, and to what extent, different factors influenced regime shift trajectory in this region. As expected, they found that reduced depth and structural complexity, and high nutrient levels increased the likelihood of regime shifting. Importantly, Arif et al. (2022) found additional factors that were not evident from a past correlative studies using the same dataset and a causal salad approach (e.g., Graham et al. 2015). Additional insights included the positive effect of pre-disturbance macroalgae cover, branching coral and wave exposure of regime shift occurrence (Arif et al. 2022). These results highlight that when dealing with observational data, different statistical adjustments can lead to different conclusions about a study system. Given this, Arif et al. (2022) recommend applying DAGs and the backdoor criterion for model selection across observational coral reef studies.

What causes species-level trait covariation?

Ecological theory suggests that there may be several causes of species-level trait covariation including size, pace of life, evolutionary history, and ecological condition (Cronin and Schoolmaster 2018). Although numerous studies have attempted to quantify the causal effect of these factors on trait covariation, these studies do not consider the causal structure driving trait variation, which in turn can lead to inappropriate statistical adjustments and biased estimates. To resolve this, Cronin and Schoolmaster 2018 synthesized relevant literature and domain knowledge to create a DAG representing the causes of species-level trait covariation that can be applied to across multiple kingdoms.

As their Fig 8 DAG suggests, size and pace-of-life may be two direct causes of trait covariation, and their influence on traits are confounded by evolutionary history and ecological conditions. To determine how size and pace of life effect trait covariation, they first had to accurately quantify their causal effect on each trait, as this information was subsequently used to determine their influence on trait covariation. One way to do this is to employ the backdoor criterion. For example, to determine the effect of size on a trait (e.g. Trait 1 in Fig 8), the backdoor criterion instructs us to adjust for either pace of life or evolutionary constraints and ecological condition to remove the confound of evolutionary history and ecological condition. In contrast, previous studies have estimated the effect of either size or pace of life on traits without first controlling for these confounding variables (e.g. Brown et al. 2004; Johnson et al. 2012). Another widely accepted approach has been to first account for evolutionary constraints and then analyze the residuals (e.g. Bielby et al. 2007, Huang et al. 2013). However, Cronin and Schoolmaster 2018 show that these approaches lead to erroneous estimates about the causes of trait covariation. They also showed that methods including principle component analysis (PCA) and exploratory factor analysis (EFA) are not able to partition trait covariance when the direct causes (size and pace of life) are correlated due to shared drivers (evolutionary history and ecological conditions). This is concerning as several high-profile studies have used these techniques to reach their conclusions (e.g. Wright et al. 2004 concluded from a PCA that size is the only causes of lead trait covariance). Taken together, a well-considered DAG guides ecologists on the sufficient set(s) for adjustment required to quantify the causes of trait-covariation and further highlights the utility of Pearl's SCM framework for observational causal inference.

Is biodiversity a cause of ecosystem functioning?

A central goal of ecology is to understand the causes of ecosystem functioning (Mittelbach 2012); however, correctly identifying these causes has been difficult because there are numerous hypothesized drivers that are often interrelated. A widespread belief among ecologists is that biodiversity is a prominent cause of ecosystem functioning (Tilman et al. 2014). Hundreds of papers have published Biodiversity-Ecosystem Function (BEF) correlations across various ecological systems, with conflicting theories and conclusions (Schoolmaster et al. 2020). To better understand whether biodiversity causally effects ecosystem functioning, Schoolmaster et al. 2020 created a DAG by synthesizing BEF literature and logic (Fig 9a). Their DAG deviates from the standard model whereby species richness is assumed to effect ecosystem functioning through functional trait diversity (Loreau 2001), and instead posits that species composition effect both species richness and functional trait diversity, with functional trait diversity driving ecosystem functioning (Fig 9a).

Given their DAG (Fig 9a), the backdoor criterion states that functional trait distribution and the environment needs to be adjusted for to determine the causal effect (or lack thereof) of biodiversity on ecosystem function. Using simulated and empirical data, Schoolmaster et al. 2020 show that when this is done, there is no causal relationship found between biodiversity and ecosystem functioning. Instead, they argue that previous observational studies that have found an association between biodiversity and ecosystem function arise from model misspecification (i.e., having an incomplete or incorrect set of predictors). For example, given their DAG, confounding bias from failing to condition on environmental factors can lead to spurious (i.e., non-causal) associations between biodiversity and ecosystem functioning. Given their DAG, Schoolmaster et al. 2020 conclude that BEF correlations are non-causal associations. Instead, their model suggests that it is species composition and not biodiversity that drives ecosystem functioning.

Recently, a comment on Schoolmaster et al. (2020) was published by Grace et al. (2021), criticizing their DAG and conclusions, asserting that biodiversity causally effects ecosystem functioning. They provide an alternative DAG, which maintains that biodiversity can causally effect ecosystem functioning indirectly through its effect on trait diversity (i.e., 'distinct functional trait'; Fig 9b). This aligns with the standard model (Loreau, 2001) on BEF correlations being causal. Schoolmaster et al. 2022 responded with a comprehensive reply, addressing critiques of their DAG, clarifying the SCM framework, and showing that the standard model and past interpretations of BEF experiments are not supported by causal analyses. Interestingly, Schoolmaster et al. 2022 note that the simulations provided by Grace et al. (2021)

do not represent the standard model DAG they defend, but instead map onto the DAG presented by Schoolmaster et al. (2020).

Although the issue of BEF correlation versus causation has yet to be resolved, there now exist two contradictory DAGs that can be used to focus critical debate and deepen our understanding of this potential process. As noted by Grace et al. (2021), DAGs allow researchers to state their causal assumptions explicitly and transparently. Ultimately, this allows other researchers to examine those causal assumptions and subsequent interpretations critically, as was done by Grace et al. (2021) and Schoolmaster et al. 2022. Ultimately, communicating and critiquing researchers' causal assumptions through DAGs may lead to a deeper understanding of BEF correlations, as well as for other ecological phenomena.

Additional Considerations

Inaccurate or Unknown Causal Structure

One of the potential limitations of DAGs is that they may not accurately represent the true causal nature of an ecological system. Simply put, inaccurate DAGs will lead to inaccurate causal inference. This can arise when using incorrect theory and background information, or by creating DAGs based on available data, rather than incorporating all relevant variables (such as omitted or unobserved variables). However, as a researcher's causal assumptions are explicitly stated through graphical representation, DAGs allow reviewers to explicitly critique and correct potential problems with far more transparency than is typical (Pearl 2009, Pearl 2010). Further, the ability to test DAG-data consistency via d-separation rules facilitates more reliable conclusions (Textor et al. 2016).

We believe that SCM should be used whenever researchers have causal objectives and sufficient background knowledge to create and justify the assertions made in their DAG. If, however, the causal structure between the predictor and response variables of interest are not fully known, but there exists enough background knowledge and support to create several plausible DAGs (each of which support DAG-data consistency), it may be advantageous to present all DAGs as plausible alternatives, reflecting this epistemic uncertainty. This should provide more accurate estimates, especially when predictor variables have the same covariate adjustments across a range of plausible DAGs. We emphasize that since several DAGs can pass DAG-data consistency, it is always imperative to first justify a DAG (or set of DAGs) based on theory, instead of relying solely on DAG-data consistency.

Application within quasi-experimental and experimental approaches

In recent years, ecologists have promoted the use of quasi-experimental methods for causal inference, including propensity score matching, before-after-control-impact (BACI) studies, regression discontinuity design, and instrumental variables (Butsic et al. 2017, Larsen et al. 2019). Here, DAGs and the principles of the SCM framework (e.g., the backdoor criterion) can be used to create more robust study designs as well as explicitly communicate assumptions required for quasi-experimental approaches (Arif and MacNeil 2022). For example, propensity score matching is employed to remove confounding bias associated with ecological observational studies (e.g., Ramsey et al. 2019). However, although past ecological studies assumed confounding variables that enter a propensity score analysis, it is unclear how these variables relate to one another and within the broader causal structure of a study system. Without this knowledge, it is unclear whether there are unmeasured variables that need to be included in the propensity score (leading to confounding bias) or whether the inclusion of selected variables may lead to other forms of bias (e.g., overcontrol and collider bias; Shrier 2009, Sjolander 2009, Mansournia et al. 2013). As noted by Pearl, for a propensity score analysis to be valid, the selected variables that enter a propensity score must satisfy the backdoor criterion to remove bias (Pearl 2009). In other words, the variables that enter a propensity score should be the sufficient set for adjustment based on the backdoor criterion. For an overview of how the SCM framework can guide quasi-experimental study designs (including propensity score and other matching methods, BACI studies, regression discontinuity designs and instrumental variables), we refer readers to Arif and MacNeil (2022). By utilizing DAGs and the principles of the SCM framework, ecologists can design more robust quasi-experimental approaches, while explicitly communicating their causal assumptions to their audience.

DAGs and the SCM framework can also guide causal inference in experimental studies. Like observational studies, experimental studies rely on causal assumptions that must be ensured by the researcher (Kimmel et al. 2021). Here, DAGs can be used to understand if data collected from an experimental set up (e.g., natural experiment or randomized control trail (RCT)) can be used for causal inference or if there are sources of bias that need to be accounted for (e.g., Williams et al. 2018; Schoolmaster et al. 2020; Schoolmaster et al. 2022). For example, Williams et al. (2018) overview a RCT investigating the effect of an intervention promoting breastfeeding on cognitive development during childhood. A DAG of this study clarifies that only using data from individuals who attend a follow-up session can lead to collider bias because both the intervention and outcome can affect the likelihood of individuals following up; therefore, follow-up data should not be distinctly analyzed (Williams et al. 2018). As an ecological example, Schoolmaster et al. (2020) use their biodiversity-ecosystem function (BEF) DAG to argue that BEF experiments do not directly manipulate biodiversity, but rather manipulate community structure, failing to isolate for the biodiversity effect.

Conclusions

Ecology has relied on observational data from its inception (Elton 1927), yet use of causal logic has typically been limited to controlled randomized experiments. Our ongoing reliance on observational data to understand fundamental questions in ecology requires the increased use of valid causal inference methodologies. Here we have introduced Pearl's SCM framework, which allows causal inference to be made in a wide range of observational contexts. The SCM framework uses DAGs to visualize the hypothesized causal structure of a system or process under study, allowing researchers to explicitly communicate their causal assumptions. Once a DAG has been built that is sufficient to characterize a system or process under study, the backdoor or frontdoor criterion can be employed to guide appropriate statistical adjustments required for causal inference. Doing so can improve conclusions made from observation-based research and will ultimately increase the depth and pace of ecological research.

Author Contributions:

S.A. and A.M. conceived the idea; S.A. led the writing of the manuscript. Both authors contributed to drafts and gave final approval for publication.

Literature Cited:

Adams, Vanessa M., Samantha A. Setterfield, Michael M. Douglas, Mark J. Kennard, and Keith Ferdinands. 2015. "Measuring Benefits of Protected Area Management: Trends across Realms and Research Gaps for Freshwater Systems." *Philosophical Transactions of the Royal Society B: Biological Sciences* 370 (1681): 20140274. https://doi.org/10.1098/rstb.2014.0274.

Ankan, Ankur, Inge M. N. Wortel, and Johannes Textor. 2021. "Testing Graphical Causal Models Using the R Package 'Dagitty.'" *Current Protocols* 1 (2): e45. https://doi.org/10.1002/cpz1.45.

Arif, S. 2022. "Ecological Monographs R Code." Figshare, online resource. https://doi.org/10.6084/m9.figshare.19541059.v1

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

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

Bareinboim, Elias, and Judea Pearl. 2016. "Causal Inference and the Data-Fusion Problem." *Proceedings of the National Academy of Sciences* 113 (27): 7345–52. https://doi.org/10.1073/pnas.1510507113.

Baum, Julia K., and Boris Worm. 2009. "Cascading Top-down Effects of Changing Oceanic Predator Abundances." *Journal of Animal Ecology* 78 (4): 699–714. https://doi.org/10.1111/j.1365-2656.2009.01531.x.

Bellwood, D. R., T. P. Hughes, C. Folke, and M. Nyström. 2004. "Confronting the Coral Reef Crisis." *Nature* 429 (6994): 827–33. https://doi.org/10.1038/nature02691.

Bielby, J., G. M. Mace, O. R. P. Bininda-Emonds, M. Cardillo, J. L. Gittleman, K. E. Jones, C. D. L. Orme, and A. Purvis. 2007. "The Fast-Slow Continuum in Mammalian Life History: An Empirical Reevaluation." *The American Naturalist* 169 (6): 748–57. https://doi.org/10.1086/516847.

Brown, James H., James F. Gillooly, Andrew P. Allen, Van M. Savage, and Geoffrey B. West. 2004. "Toward a Metabolic Theory of Ecology." *Ecology* 85 (7): 1771–89. https://doi.org/10.1890/03-9000.

Burnham, Kenneth P., and David R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media.

Butsic, Van, David J. Lewis, Volker C. Radeloff, Matthias Baumann, and Tobias Kuemmerle. 2017. "Quasi-Experimental Methods Enable Stronger Inferences from Observational Data in Ecology." *Basic and Applied Ecology* 19 (March): 1–10. https://doi.org/10.1016/j.baae.2017.01.005.

Carmel, Yohay, Rafi Kent, Avi Bar-Massada, Lior Blank, Jonathan Liberzon, Oded Nezer, Gill Sapir, and Roy Federman. 2013. "Trends in Ecological Research during the Last Three Decades – A Systematic Review." *PLOS ONE* 8 (4): e59813. https://doi.org/10.1371/journal.pone.0059813.

Choy, Samantha Low, Rebecca O'Leary, and Kerrie Mengersen. 2009. "Elicitation by Design in Ecology: Using Expert Opinion to Inform Priors for Bayesian Statistical Models." *Ecology* 90 (1): 265–77.

Cronin, James Patrick, and Donald R. Schoolmaster Jr. 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.

Cunningham, Scott. 2021. *Causal Inference: The Mixtape*. *Causal Inference*. Yale University Press. https://doi.org/10.12987/9780300255881.

D., F. A. 1900. "The Grammar of Science." *Nature* 62 (1594): 49–50. https://doi.org/10.1038/062049a0.

Drescher, M., A. H. Perera, C. J. Johnson, L. J. Buse, C. A. Drew, and M. A. Burgman. 2013. “Toward Rigorous Use of Expert Knowledge in Ecological Research.” *Ecosphere* 4 (7): art83. https://doi.org/10.1890/ES12-00415.1.

Elton, C.S. 1927. Animal Ecology. University of Chicago Press, Chicago.

Elwert, Felix. 2013. “Graphical Causal Models.” In *Handbook of Causal Analysis for Social Research*, edited by Stephen L. Morgan, 245–73. Handbooks of Sociology and Social Research. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-6094-3_13.

Elwert, Felix, and Christopher Winship. 2014. “Endogenous Selection Bias: The Problem of Conditioning on a Collider Variable.” *Annual Review of Sociology* 40 (July): 31–53. https://doi.org/10.1146/annurev-soc-071913-043455.

Ethier, Danielle, and Thomas Nudds. 2017. “Complexity of Factors Affecting Bobolink Population Dynamics Communicated with Directed Acyclic Graphs: Complexity of Factors Affecting Bobolink.” *Wildlife Society Bulletin* 41 (1): 4-16. https://doi.org/10.1002/wsb.739.

Galles, D., and J. Pearl. 1998. “An axiomatic characterization of causal counterfactuals.” *Foundations of Science* 3(1):151-182.

Grace, James, and Kathryn Irvine. 2019. "Scientist's Guide to Developing Explanatory Statistical Models Using Causal Analysis Principles." *Ecology* 101 (December): e02962. https://doi.org/10.1002/ecy.2962.

Grace, James B., Michel Loreau, and Bernhard Schmid. 2022. "A Graphical Causal Model for Resolving Species Identity Effects and Biodiversity–Ecosystem Function Correlations: Comment." *Ecology* 103 (2): e03378. https://doi.org/10.1002/ecy.3378.

Graham, Nicholas A. J., Simon Jennings, M. Aaron MacNeil, David Mouillot, and Shaun K. Wilson. 2015. "Predicting Climate-Driven Regime Shifts versus Rebound Potential in Coral Reefs." *Nature* 518 (7537): 94–97. https://doi.org/10.1038/nature14140.

Grubbs, R. Dean, John K. Carlson, Jason G. Romine, Tobey H. Curtis, W. David McElroy, Camilla T. McCandless, Charles F. Cotton, and John A. Musick. 2016. "Critical Assessment and Ramifications of a Purported Marine Trophic Cascade." *Scientific Reports* 6 (1): 20970. https://doi.org/10.1038/srep20970.

Greenland, S., J. Pearl, and J. M. Robins. 1999. "Causal Diagrams for Epidemiologic Research." *Epidemiology (Cambridge, Mass.)* 10 (1): 37–48.

Glymour, M., Greenland, S. Causal Diagrams. Modern Epidemiology pp 183-212 (Lippincott Williams & Wilkins, Philadelphia, 2008).

Huang, Zheng Y. X., Willem F. de Boer, Frank van Langevelde, Valerie Olson, Tim M. Blackburn, and Herbert H. T. Prins. 2013. “Species’ Life-History Traits Explain Interspecific Variation in Reservoir Competence: A Possible Mechanism Underlying the Dilution Effect.” *PLOS ONE* 8 (1): e54341. https://doi.org/10.1371/journal.pone.0054341.

Imbens, Guido W. 2020. “Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics.” *Journal of Economic Literature* 58 (4): 1129–79. https://doi.org/10.1257/jel.20191597.

Joppa, Lucas N., and Alexander Pfaff. 2009. “High and Far: Biases in the Location of Protected Areas.” *PLOS ONE* 4 (12): e8273. https://doi.org/10.1371/journal.pone.0008273.

Johnson, Pieter T. J., Jason R. Rohr, Jason T. Hoverman, Esra Kellermanns, Jay Bowerman, and Kevin B. Lunde. 2012. “Living Fast and Dying of Infection: Host Life History Drives Interspecific Variation in Infection and Disease Risk.” *Ecology Letters* 15 (3): 235–42. https://doi.org/10.1111/j.1461-0248.2011.01730.x.

Kuhnert, Petra M., Tara G. Martin, and Shane P. Griffiths. 2010. “A Guide to Eliciting and Using Expert Knowledge in Bayesian Ecological Models.” *Ecology Letters* 13 (7): 900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x.

Larsen, Ashley E., Kyle Meng, and Bruce E. Kendall. 2019. “Causal Analysis in Control–Impact Ecological Studies with Observational Data.” *Methods in Ecology and Evolution* 10 (7): 924–34. https://doi.org/10.1111/2041-210X.13190.

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, et al. 2001. “Biodiversity and Ecosystem Functioning: Current Knowledge and Future Challenges.” *Science* 294 (5543): 804–8. https://doi.org/10.1126/science.1064088.

MacDonald, R. 2004. “The specific analysis of structural equation models.” *Multivariate Behavioral Research* 39: 687-713.

Mansournia, Mohammad A., Miguel A. Hernán, and Sander Greenland. 2013. “Matched Designs and Causal Diagrams.” *International Journal of Epidemiology* 42 (3): 860–69. https://doi.org/10.1093/ije/dyt083.

Marcot, Bruce G, J Douglas Steventon, Glenn D Sutherland, and Robert K McCann. 2006. “Guidelines for Developing and Updating Bayesian Belief Networks Applied to Ecological Modeling and Conservation.” *Canadian Journal of Forest Research* 36 (12): 3063–74. https://doi.org/10.1139/x06-135.

Martin, Tara G., Mark A. Burgman, Fiona Fidler, Petra M. Kuhnert, Samantha Low-Choy, Marissa McBride, and Kerrie Mengersen. 2012. “Eliciting Expert Knowledge in Conservation

Science." *Conservation Biology: The Journal of the Society for Conservation Biology* 26 (1): 29–38. https://doi.org/10.1111/j.1523-1739.2011.01806.x.

Mascia, Michael, C. Claus, and Robin Naidoo. 2010. "Impacts of Marine Protected Areas on Fishing Communities." *Conservation Biology : The Journal of the Society for Conservation Biology* 24 (October): 1424–29. https://doi.org/10.1111/j.1523-1739.2010.01523.x.

McElreath, Richard. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*. 2 edition. Boca Raton: Chapman and Hall/CRC.

McNay, R. S., B. G. Marcot, V. Brumovsky, and R. Ellis. 2006. "A Bayesian Approach to Evaluating Habitat for Woodland Caribou in North-Central British Columbia." *Canadian Journal of Forest Research* 36: 3117-3133.

Mittelbach, G. 2012. Community ecology. Sinauer Associates, Sunderland, Maryland, USA.

Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd ed. Analytical Methods for Social Research. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781107587991.

Myers, Ransom A., Julia K. Baum, Travis D. Shepherd, Sean P. Powers, and Charles H. Peterson. 2007. "Cascading Effects of the Loss of Apex Predatory Sharks from a Coastal Ocean." *Science* 315 (5820): 1846–50. https://doi.org/10.1126/science.1138657.

Pearce, Neil, and Debbie A Lawlor. 2016. “Causal Inference—so Much More than Statistics.” *International Journal of Epidemiology* 45 (6): 1895–1903. https://doi.org/10.1093/ije/dyw328.

Pearl, Judea. 1993. “[Bayesian Analysis in Expert Systems]: Comment: Graphical Models, Causality and Intervention.” *Statistical Science* 8 (3): 266–69.

Pearl, Judea. 1995. “Causal Diagrams for Empirical Research.” *Biometrika* 82 (4): 669–88. https://doi.org/10.2307/2337329.

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Pearl, J. ([2000] 2009). Causality: Models, Reasoning and Inference. 2nd ed. Cambridge, UK: Cambridge University Press.

Ramsey, David S. L., David. M. Forsyth, Elaine Wright, Meredith McKay, and Ian Westbrooke. 2019. “Using Propensity Scores for Causal Inference in Ecology: Options, Considerations, and a Case Study.” *Methods in Ecology and Evolution* 10 (3): 320–31. https://doi.org/10.1111/2041-210X.13111.

Rohrer, Julia 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.

Rubin, D. 2005. Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association* 100: 322–331.

Sagarin, Raphael, and Aníbal Pauchard. 2010. "Observational Approaches in Ecology Open New Ground in a Changing World." *Frontiers in Ecology and the Environment* 8 (7): 379–86. https://doi.org/10.1890/090001.

Sam, Katerina, Bonny Koane, Samuel Jeppy, and Vojtech Novotny. 2014. "Effect of Forest Fragmentation on Bird Species Richness in Papua New Guinea." *Journal of Field Ornithology* 85 (June). https://doi.org/10.1111/jofo.12057.

Schoolmaster, Donald R., Chad R. Zirbel, and James Patrick Cronin. 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.

Schoolmaster Jr., Donald R., Chad R. Zirbel, and James Patrick Cronin. 2022. "A Graphical Causal Model for Resolving Species Identity Effects and Biodiversity–Ecosystem Function Correlations: Reply." *Ecology* 103 (2): e03593. https://doi.org/10.1002/ecy.3593.

Shipley, Bill. 2000. “A New Inferential Test for Path Models Based on Directed Acyclic Graphs.” *Structural Equation Modeling: A Multidisciplinary Journal* 7 (2): 206–18. https://doi.org/10.1207/S15328007SEM0702_4.

Shipley, B. 2016. Cause and correlation in biology: A user’s guide to path analysis, structural equations and causal inference with R. Second edition. Cambridge University Press, Cambridge, UK.

Shrier, I. 2009. “Letter to the editor.” *Statistics in Medicine* 27: 2740–2741.

Sjölander, Arvid. 2009. “Letter to the Editor: Propensity Scores and M-Structures.” *Statistics in Medicine* 28 (April): 1416–20. https://doi.org/10.1002/sim.3532.

Spirtes, Peter, Clark Glymour, and Richard Scheines. 2001. *Causation, Prediction, and Search*. 2nd ed. Adaptive Computation and Machine Learning Series. Cambridge, MA, USA: A Bradford Book.

Textor, Johannes, Benito van der Zander, Mark S. Gilthorpe, Maciej Liskiewicz, and George Th Ellison. 2016. “Robust Causal Inference Using Directed Acyclic Graphs: The R Package ‘Dagitty.’” *International Journal of Epidemiology* 45 (6): 1887–94. https://doi.org/10.1093/ije/dyw341.

Tilman, D., F. Isbell, and J. M. Cowles. 2014. “Biodiversity and ecosystem functioning.” *Annual Review of Ecology, Evolution & Systematics* 45:471–493.

Williams, Thomas C., Cathrine C. Bach, Niels B. Matthiesen, Tine B. Henriksen, and Luigi Gagliardi. 2018. “Directed Acyclic Graphs: A Tool for Causal Studies in Paediatrics.” *Pediatric Research* 84 (4): 487–93. https://doi.org/10.1038/s41390-018-0071-3.

Wright, S. 1921. “Correlation and causation.” *Journal of Agricultural Research* 20: 557–585.

Wright, Ian J., Peter B. Reich, Mark Westoby, David D. Ackerly, Zdravko Baruch, Frans Bongers, Jeannine Cavender-Bares, et al. 2004. “The Worldwide Leaf Economics Spectrum.” *Nature* 428 (6985): 821–27. https://doi.org/10.1038/nature02403.

Figure Legends

Figure 1. A directed acyclic graph (DAG) representing the causal structure between three variables, X, Y and W.

Figure 2. A workflow for going from DAGs to causal inference under the SCM framework.

Fig 3. A DAG representing how different factors may influence forest species abundance.

Figure 4. Results from linear regression models that employed the backdoor criterion to determine the causal effect of different predictor variables on forest species abundance, using our simulated dataset with specified (i.e., known) causal effects. Predictor, response, and control variables are highlighted in green, blue, and black, respectively; omitted variables are shaded in grey. We chose generalized linear regression as our statistical models; for example, the protected area model is represented by the linear regression equation: $Forest\ Species\ Abundance_i = \alpha + \beta 1 Protected\ Area_i + \beta 2 Slope_i + \beta 3 Elevation_i + \beta 4 Distance\ to\ Roads\ and\ Cities_i + \varepsilon_i$. The known and estimated causal effects, along with AIC values are noted for each model. Lastly, the results from a causal salad model (where all variables are placed under one model) are shown as a contrast, with estimated effects for each included variable noted in red.

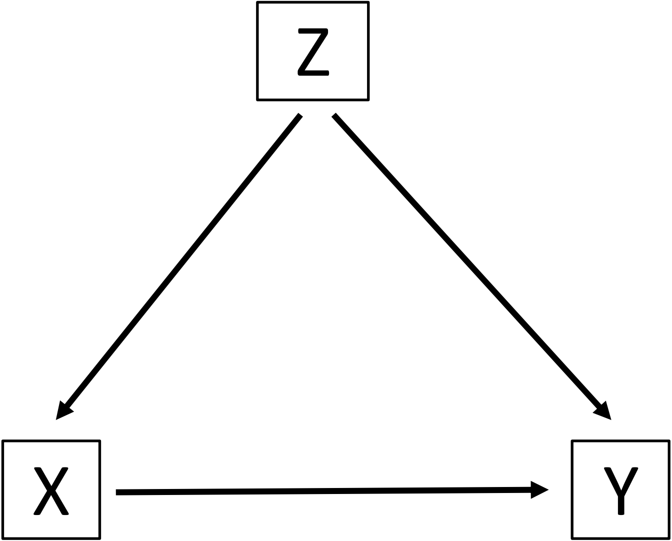
Figure 5. A DAG where the effect of X on Y cannot be estimated (due to an unobserved confounding variable U) without the use of the front door criterion.

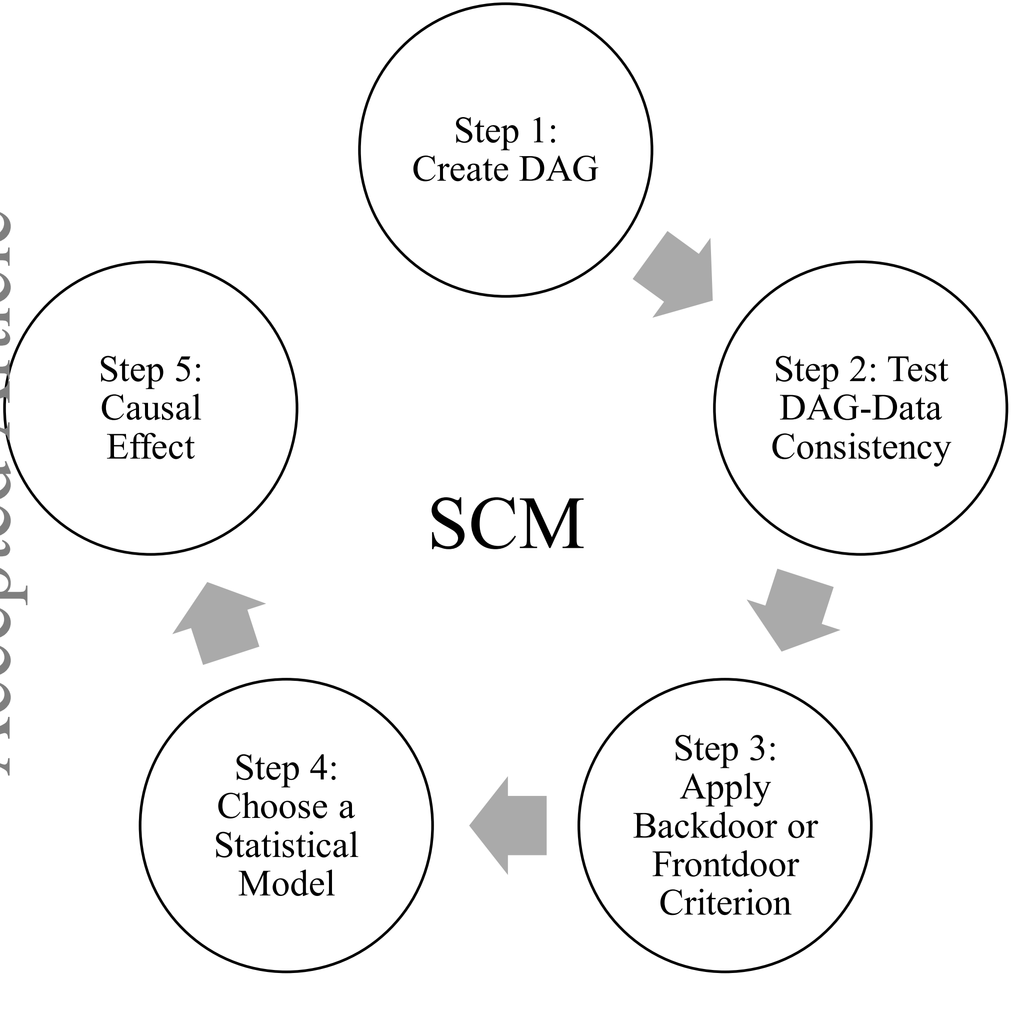
Figure 6. Employing the frontdoor criterion. (a) A DAG representing the causal structure between sharks and bivalves. Here, fishing pressure is an unobserved variable, and the frontdoor criterion needs to be employed to determine the effect of sharks on bivalves. (b, c) Employing the frontdoor criterion to determine the effect of sharks on bivalves from our simulated shark-bivalve dataset. Linear regression models were used to first determine the effect of sharks on rays and the effect of rays on bivalves, using the backdoor criterion to determine the sufficient set for adjustment. The product of these two effects gives us the effect of sharks on bivalves. Known causal effects (from our simulated data) between variables of are noted for comparison.

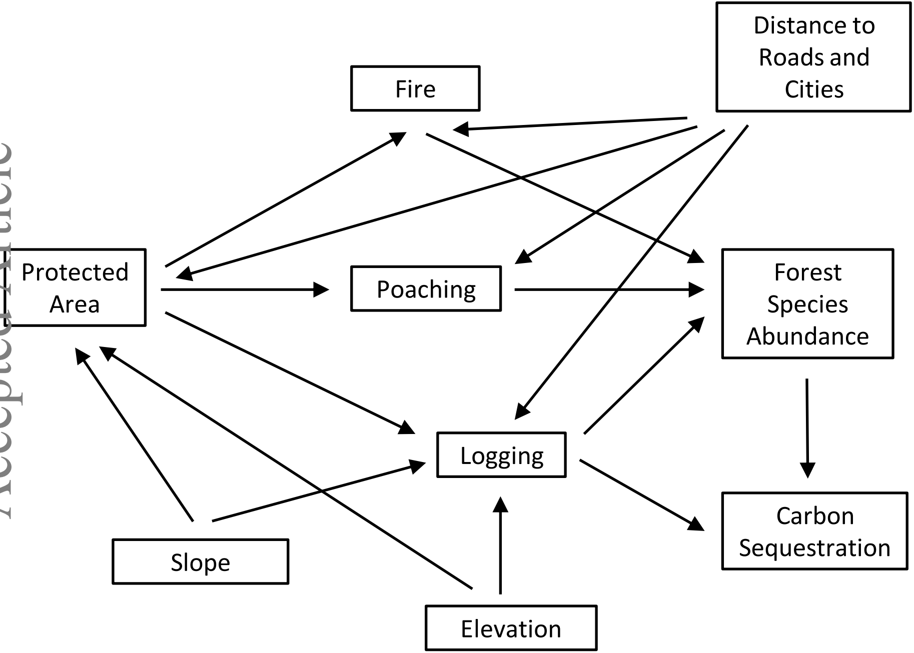
Figure 7. A DAG representing how different factors may influence coral reef regime shifts following a climate-induced bleaching event across Seychelles, from Arif et al. (2022).

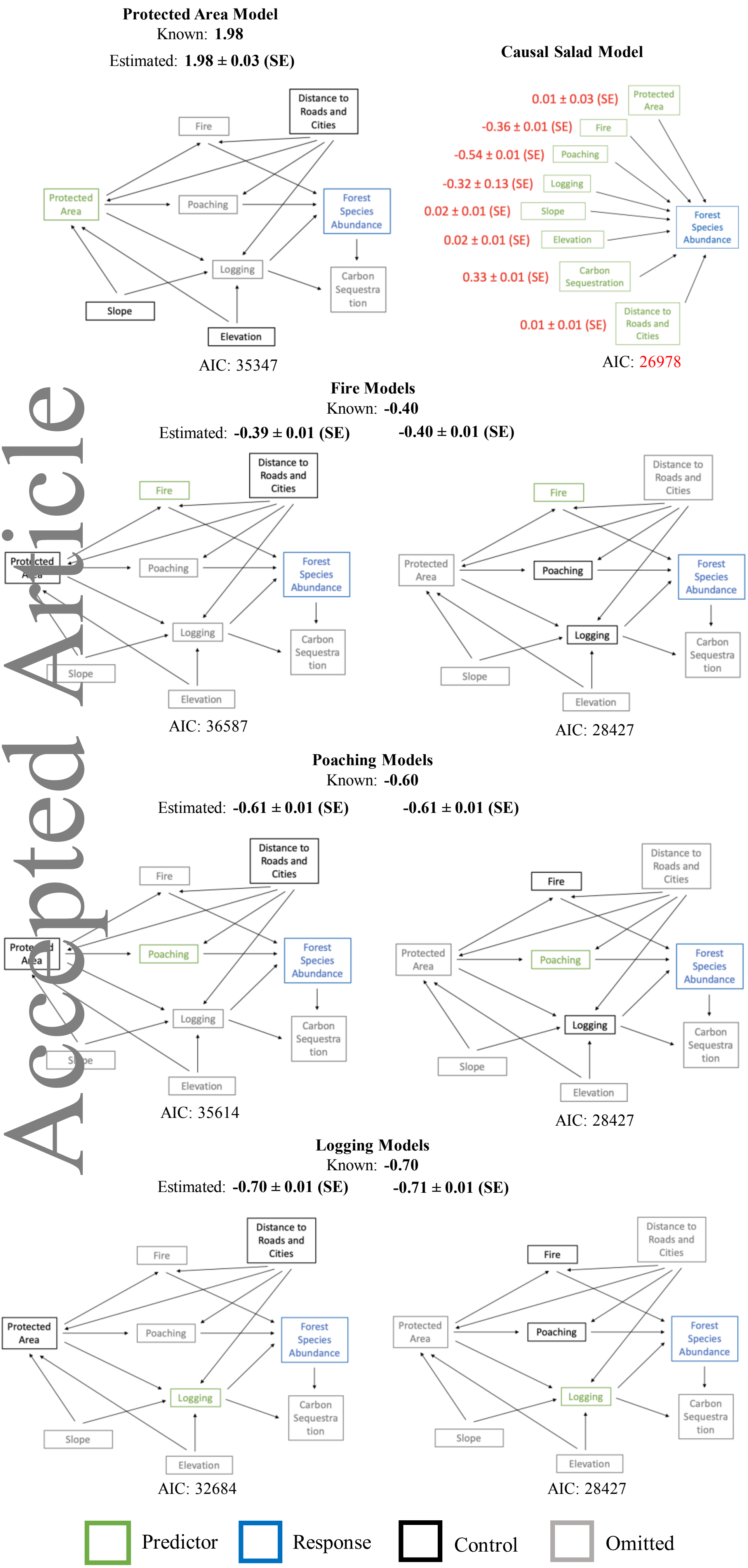
Figure 8. A DAG representing how different factors influence species-level trait covariation, from Cronin and Schoolmaster 2018.

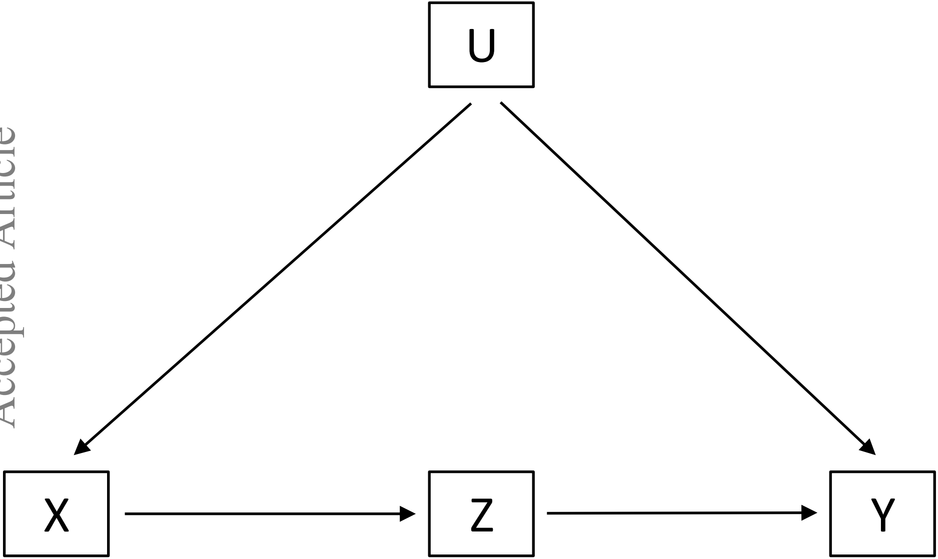
Figure 9. Two DAGs representing the causal relationship between biodiversity and ecosystem function. The DAG in (a) is from Schoolmaster et al. 2020 and the DAG in (b) is from Grace et al. 2021.

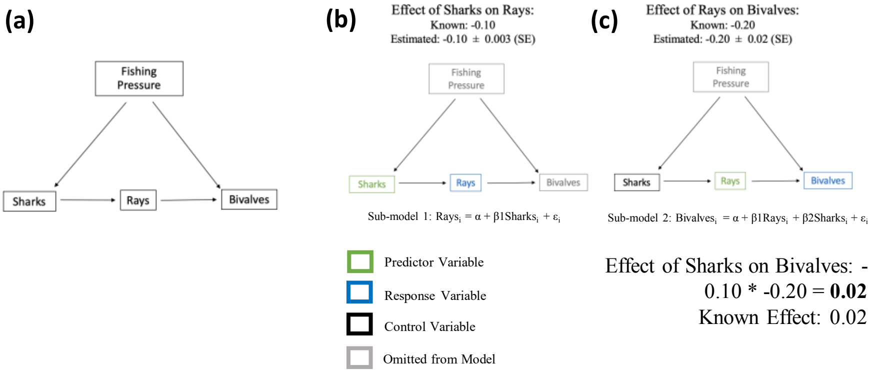

Z
X
Y

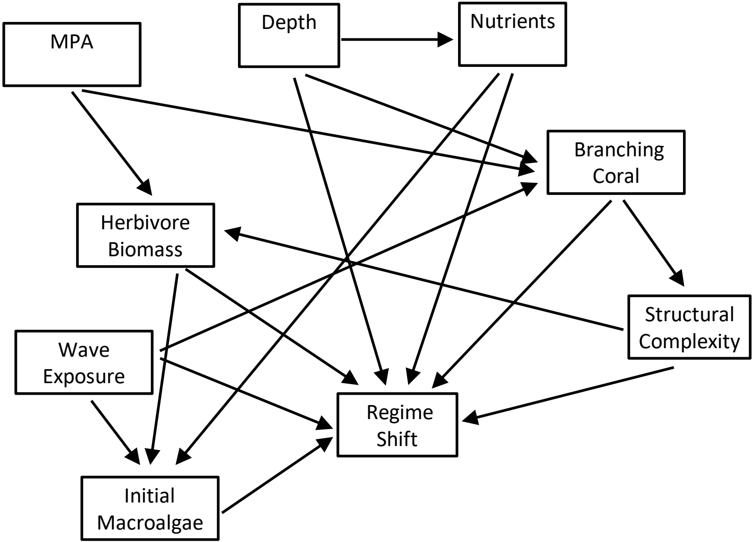

Step 1: Create DAG
Step 2: Test DAG-Data Consistency
Step 3: Apply Backdoor or Frontdoor Criterion
Step 4: Choose a Statistical Model
Step 5: Causal Effect
SCM

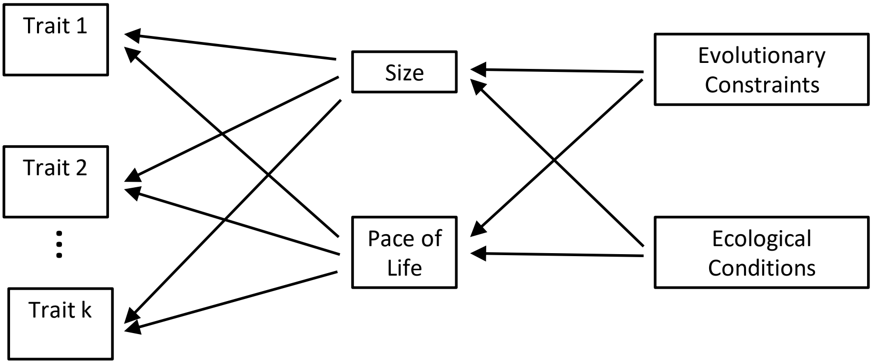

Fire
Distance to Roads and Cities
Protected Area
Poaching
Forest Species Abundance
Logging
Carbon Sequestration
Slope
Elevation

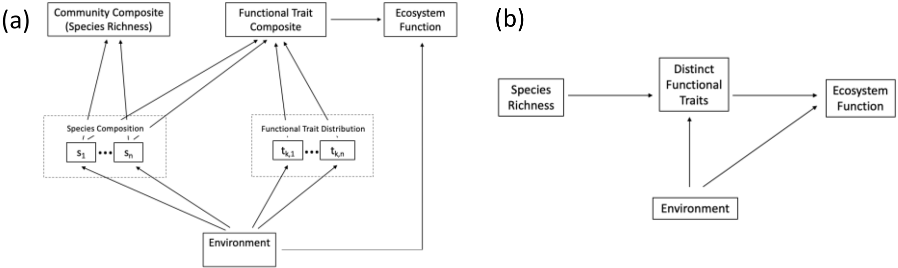

Protected Area Model
Known: 1.98
Estimated: 1.98 ± 0.03 (SE)
Distance to Roads and Cities
Fire
Protected Area
Poaching
Forest Species Abundance
Logging
Carbon Sequestration
Slope
Elevation
AIC: 35347
Causal Salad Model
0.01 ± 0.03 (SE) Protected Area
-0.36 ± 0.01 (SE) Fire
-0.54 ± 0.01 (SE) Poaching
-0.32 ± 0.13 (SE) Logging
0.02 ± 0.01 (SE) Slope
0.02 ± 0.01 (SE) Elevation
0.33 ± 0.01 (SE) Carbon Sequestration
0.01 ± 0.01 (SE) Distance to Roads and Cities
Forest Species Abundance
AIC: 26978
Fire Models
Known: -0.40
Estimated: -0.39 ± 0.01 (SE)
-0.40 ± 0.01 (SE)
AIC: 36587
AIC: 28427
Poaching Models
Known: -0.60
Estimated: -0.61 ± 0.01 (SE)
-0.61 ± 0.01 (SE)
AIC: 35614
AIC: 28427
Logging Models
Known: -0.70
Estimated: -0.70 ± 0.01 (SE)
-0.71 ± 0.01 (SE)
AIC: 32684
AIC: 28427
Predictor
Response
Control
Omitted


U
X
Z
Y


(a)
Fishing Pressure
Sharks
Rays
Bivalves
(b)
Effect of Sharks on Rays:
Known: -0.10
Estimated: -0.10 ± 0.003 (SE)
Fishing Pressure
Sharks
Rays
Bivalves
Sub-model 1: $Rays_i = \alpha + \beta 1 Sharks_i + \varepsilon_i$
(c)
Effect of Rays on Bivalves:
Known: -0.20
Estimated: -0.20 ± 0.02 (SE)
Fishing Pressure
Sharks
Rays
Bivalves
Sub-model 2: $Bivalves_i = \alpha + \beta 1 Rays_i + \beta 2 Sharks_i + \varepsilon_i$
Predictor Variable
Response Variable
Control Variable
Omitted from Model
Effect of Sharks on Bivalves: -0.10 * -0.20 = **0.02**
Known Effect: 0.02


MPA
Depth
Nutrients
Branching Coral
Herbivore Biomass
Structural Complexity
Wave Exposure
Regime Shift
Initial Macroalgae


Trait 1
Trait 2
⋮
Trait k
Size
Pace of Life
Evolutionary Constraints
Ecological Conditions


(a)
Community Composite
(Species Richness)
Functional Trait
Composite
Ecosystem
Function
Species Composition
s_1 ••• s_n
Functional Trait Distribution
$t_{k,1}$ ••• $t_{k,n}$
Environment
(b)
Species
Richness
Distinct
Functional
Traits
Ecosystem
Function
Environment