The benefits of DAGs without causality

Deon & Albert

## Introduction

The field of ecology is currently embroiled in a methodological debate over how best to infer causality from observational data. While all agree that experimental data remains the gold standard for inference (when carried out appropriately), observational data are more often the norm in our field due to the usual and obvious reasons we are all familiar with (logistics, ethics, scales etc.). Given science is predominantly interested in making causal statements using such data, the resulting discourse has created various schools of thought, with each advocating for a particular method (Arif & MacNeil, Hones & Krebs, Nichols et al). This has resulted in discussions which argue over definitions of causality, required assumptions made to infer causality, and flaws (perceived or real) in competing methods. The impetus for these discussions has been the influx of DAGs and the seemingly inextricably linked associated criterion (e.g. backdoor) into ecology, which to a large extent has challenged how much of previous research has been done, principally through making a clear conceptual distinction between predictive models and causal models, even when these may appear statistically identical (e.g. ).

Within this opinion piece we have no interest in arguing in favour of one causal method over another. Instead, we intend to highlight the benefits of DAGs *without* causal inference. We do so because we believe there is clear pedagogical and scientific value in creating a DAG, which is not tied to then making causal statements.

## Confounding in ecology

Confounding remains one of the most significant issues in ecological research (though this is true for all scientific domains). Within this paper, we define confounding as being causal relationships which are part of a larger data generating process (e.g. , and , ), but are not the relationship of focal interest (e.g. ). That is, confounded relationships are those causal relationships present within the system can be related to both and in some, even distant, manner but which are not under current investigation.

Many researchers either rely on intuition (e.g. Roos et al., 2019) or statistical tests (e.g. tests of collinearity or variance inflation factors) to determine whether a variable confounds an observed relationship. This may lead to practices such as when variables and are correlated, then one is *excluded* from the model (traditionally, the one which is “less interesting scientifically”), or if researchers believe that variable is a “confounder” then it is *included* in the model (though in practice decisions of how to deal with suspect confounds may be more ad hoc). If translating these heuristics into DAGs, then the collinearity scenario would suggest researchers consider the relationship is and decide to exclude (assuming is focal) . While the latter form of thinking implicitly assumes a structure like , where is included. In both cases alternative relationships, including , are unintentionally not considered or not known.

Going through such a process, intuiting confounding or using tests, gives researchers a sense of statistical security. If the final model has none of these issues, then it we may assume that we are permitted a degree of freedom to discuss causality (even if using code words and phrases such as “the effect of x on y”, etc REF - Sharif?). In essence, these heuristic steps give us comfort that confounding has been dealt with but, we argue, that this is ultimately a false sense of security.

Collinearity is not a test of confounding and the consequences of collinearity (or multicollinearity) are generally benign and can be ignored is most cases (Bolker’s references that I sent to Xavier).

Converse to the above, mostly *ad hoc,* approach for identifying and dealing with confounding, DAGs give a clear definition of *different types* of confounding and rules for dealing with them: is a fork and must be included; is a mediator and the inclusion of depends on if the researcher is interested in direct or indirect relationships; and is a collider and must not be included. It is worth noting the value of this, without needing to go further into the realms of causality. Simply being able to think about your system in this light, describe how you believe a relationship may be confounded, be able to convey that efficiently to someone else, and identify solutions is a dramatic improvement from where the field is currently.

Both Nichols and Hones argue that predictive models are useful for understanding causality (e.g. Hones says validated predictions are excellent evidence of causation) and we agree but with the important caveat that the model does not reach “strong” predictions through confounding. Indeed, for many people when first introduced to DAGs, this can be a source of confusion. Historically, students are taught that “correlation does not equal causation” but more subtly are taught that confounding largely means “spurious correlations” - e.g. where two randomly chosen variables happen to be correlated with each other. In reality, few researchers would ever choose to include entirely random variables in their analysis.

HOW PAPERS LOOSELY DEFINE PAPER - LEADING TO RABBIT HOLE

## Transparency

In their rejection of DAGs (or structural causal models which are models based on a DAG, SCMs), Nichols et al argue that predictive models, based on specific causal hypotheses are valid for causal inference. We are in agreement, but argue that DAGs can form a valuable part of the H-D approach, where they may act as a complement to formulating hypotheses and deriving predictions. Indeed, we believe that DAGs are just the visual representation of the data generating process, from which the hypotheses is generated, in which case creating the DAG should be relatively painless.

Even when researchers do not explicitly create a DAG, they will never-the-less have a DAG in mind. It’s just that, rather than drawing the DAG, this will exist in their mind and be conveyed through text. The core information will be there (minus the understanding that DAGs give with regards to confounding) but the effort of translating it is delegated to the reader.

Indeed this delegation of translation makes the H-D approach open to abuse. Exciting results based on confounding can be exploited by researchers and journals.

The core problem in the H-D approach, without the inclusion of DAGs, is that it requires that the reader place trust in the author that they have given the hypotheses careful consideration. DAGs make the end results of this process evidently clear and transparent (without requiring that the author subscribe to SCM). This also has the additional benefit of enabling any readers who do subscribe to SCMs to evaluate it under backdoor criterion etc.

Axioms [Expand]

Nichols one and done paper - placing study into broader context - surely DAGs help with that?

DAGs and AIC - do not equate

How to make AIC useful for DAGs

### Scenario 1 - H-P approach works

# A tibble: 7 × 2  
 formula `Number of times ranked best (lowest AIC)`  
 <chr> <int>  
1 y ~ x + z 87  
2 y ~ x + z + v 13  
3 y ~ x 0  
4 y ~ z 0  
5 y ~ v 0  
6 y ~ x + v 0  
7 y ~ z + v 0

# A tibble: 4 × 4  
 formula `Median x effect` `True x effect` `Median AIC rank`  
 <chr> <dbl> <dbl> <dbl>  
1 y ~ x + v 0.826 0 4  
2 y ~ x 0.825 0 3  
3 y ~ x + z + v 1.98 0 2  
4 y ~ x + z 1.98 0 1

### Scenario 2 - H-P approach fails

# A tibble: 7 × 2  
 formula `Number of times ranked best (lowest AIC)`  
 <chr> <int>  
1 y ~ x + z 84  
2 y ~ x + z + v 16  
3 y ~ x 0  
4 y ~ z 0  
5 y ~ v 0  
6 y ~ x + v 0  
7 y ~ z + v 0

# A tibble: 4 × 4  
 formula `Median x effect` `True x effect` `Median AIC rank`  
 <chr> <dbl> <dbl> <dbl>  
1 y ~ x + v 0.00343 0 7  
2 y ~ x 0.00252 0 5  
3 y ~ x + z + v 1.02 0 2  
4 y ~ x + z 1.02 0 1

Though the complication is that should not be interpreted but will very often be interpreted (Greenland Table 2 fallacy paper).

## The band wagon

A central issue in the updake of DAGs may be related to the view that these represent the final culmination in scientific advancement and that all that has come before has been for naught; “we have *finally* developed the silver bullet”. [Expand]

Indeed, it is fair to place currently in fashion methods under special scrutiny to ensure we don’t all jump on a doomed band wagon. [Expand]

SCMs can fail in multiple ways. They may be token or low effort check box exercises, done simply to get the badge. Additionally, they still require trust that the authors have given it due diligence. [Expand]