This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

Causal Diagrams for Civil and Structural Engineers

M.Z. Naser[1,2]
[1]School of Civil & Environmental Engineering and Earth Sciences (SCEEES), Clemson University, USA
[2]Artificial Intelligence Research Institute for Science and Engineering (AIRISE), Clemson University, USA
E-mail: mznaser@clemson.edu, Website: www.mznaser.com

Abstract

Causal diagrams are logic and graphical tools that depict assumptions about presumed causal relations. Such diagrams have proven effective in tackling a variety of problems in social sciences and epidemiology research yet remain foreign to civil engineers. Unlike the traditional means of examining relationships via multivariable regression, causal diagrams can identify the presence of confounders, colliders, and mediators. Thus, this paper hopes to introduce the big ideas behind causal diagrams (specifically, directed acyclic graphs (DAGs)) and how to create and apply such diagrams to several civil engineering problems. Findings from the presented case studies indicate that civil engineers can successfully use causal diagrams to improve their understanding of complex causation relations, thereby accelerating research and practical efforts.

Keywords: Causality; Causal diagrams; Machine learning; Structural engineering.

1.0 Introduction

A primary objective of causality is to establish causal relationships, identifying how a presumed cause leads to and activates the appearance of an effect [1]. More often than not, an effect is generated due to a number of causes. These causes could be independent or dependent on one another. A particular cause may precede other causes, while some causes might happen concurrently with others. Multiple causes can also share a common cause (confounder). A challenge then becomes to recognize a suitable means to tie causes to effects in a proper manner that preserves the reality of the phenomenon at hand. This complexity necessitates various levels of approaches, especially in fields where the interplay of multiple causes may directly impact effects, such as structural engineering.

From the lens of structural engineering, engineers have leveraged first principles to derive meaningful expressions that epitomize causal relationships. For example, the moment capacity of a W-shaped steel beam comprises the product of the plastic modulus and yield strength of structural steel. This illustrates a direct application of causal understanding in engineering design. However, when first principles are insufficient due to the complexity of interactions or are not readily applicable, the engineers revert to statistical approaches – namely linear and multilinear regression. This transition from theoretical to empirical modeling is evident in adopting linear and multilinear regression to formulate engineering solutions. For example, the formula for evaluating the fire resistance of RC columns as adopted in the AS3600 code was also attained via multilinear regression [2]. Other formulae in structural engineering are also built on linear regression, such as those adopted in ASCE 29 [3].

In some instances, linear-like regression may not be able to realize proper formulae, and hence, the use of nonlinear regression analysis can be adopted. A common example of purely-based nonlinear regression is the formula created by Orangun et al. [4] to calculate bar development length in steel reinforcement embedded in concrete. Another example is the shear strength model for reinforced concrete beams as endorsed by the ACI committee 446 (Fracture Mechanics of Concrete) [5], as well as for FRP-reinforced concrete beams [6]. The use of nonlinear regression is quite common in seismic applications of structural design [7,8], as well as in examining size effects [9]. Bazant and Yu [10] note that while problems within the same domain may require rigorous statistical investigation (i.e., size effect) [10], some expressions (such as that for estimating the elastic modulus of concrete in terms of the compressive strength, etc.) can be arrived at empirically. A number of expressions were also arrived at empirically, such as that often used to evaluate the fire resistance of concrete-filled hollow steel columns [11] and in timber design [12].

The recent rise of nonparametric regression (such as symbolic regression) has also proved to be useful in deriving engineering formulae. These techniques offer a way to derive engineering formulas without the constraints of traditional regression assumptions (linearity, uniformity, homoscedasticity, etc.), providing a fresh perspective on modeling complex phenomena. While symbolic regression can also yield formulae, the arrangement of the dependent variables may not mirror that derived via traditional analysis, nor that to be physically plausible. Surprisingly, these models can still predict civil engineering phenomena with ease and accuracy that exceed those of their traditional counterparts [13].

A look into the above variants of regression signifies that such variants are easy to use, well accepted, and, most of all, interpretable. This interpretability arises on two fronts: 1) the fact that regression yields a formula that shows the relation between dependent variables, and 2) quantifies the contribution of each dependent variable upon the response of interest. Attaining a formula allows engineers to simply substitute this formula without the hassle of software or complex procedures.

From an engineering sense, regression is often implicitly used to state that the dependent variables 1) predict or cause the phenomenon, and 2) this prediction and causation are in the form of the formula on hand. In reality, this may, and is likely, not be precise. For example, while linear regression can successfully create formulae to predict phenomena, the same formulas and models cannot be declared simply as causal.

This misconception could arise from pre-specifying the arrangement and form of the relationships adopted in such a formula. For example, the shear strength contribution of a typical concrete beam is given by the following formula:

$$V_c = 0.66\rho^{1/3}\sqrt{f'_c}b_w d \quad \text{Eq. 1}$$

Where ρ, b_w and d represent the ratio of flexural reinforcement, the width of the web, and the depth of the beam, respectively. The origin of formula comes from a comprehensive statistical analysis aimed at creating a formula to predict the shear strength of concrete [14]. The logic behind this formula matches that of our domain knowledge. For example, an increase in any of the above-listed dependent variables increases the shear strength.

Regression analysis serves as a precursor to the application of machine learning (ML) in structural engineering and, hence, bridges traditional statistical modeling with contemporary computational approaches [15]. Yet, such models may not be as transparent (i.e., do not convert into a formula) [16]. The interpretation of such models is tedious and often incomprehensible unless supplemented with explainability measures. These measures can articulate the importance of model input, the relationship between each input due to arriving at correct and accurate predictions, etc. [17,18]. This becomes of the utmost importance when creating and developing ML models [19]. This may also grow from identifying relationships to possibly establishing causal mechanisms [20].

Unlike regression and machine learning, graphs can also be used to describe relationships between dependent and independent variables. The utilization of graphs presents an alternative method to explore and establish relationships, including causal connections, within engineering contexts. In such an exercise, the graphs can represent the relationships in a parametric or nonparametric manner as opposed to linear/nonlinear relations. More interestingly, given the graphic nature of graphs, these can also describe clear causal relations as well.

The historical development of path diagrams and subsequent analytical methods stresses the potential of graphical representations to reinforce causal understanding. One of the earliest mentions of the use of graphs to display relationships is attributed to Sewall Wright [21], who established, mathematically, that a variable that resides in a graph causes another variable and not the other way around. Those graphs were called the "path diagrams" – a key part of the path analysis method. Using such graphs, Wright could articulate and defend causal assumptions mathematically and based on scientific grounds. Path analysis grows into factor analysis, and then structural equation modeling and much more in-depth reviews on the history and development of these methods can be found elsewhere [22–24]. Both of these methods remain closely tied to causal principles and with very limited use to structural engineers. As such, this paper hopes to showcase the potential of causal graphs in structural engineering.

2.0 An overview of regression

A user of regression (say, an engineer) compiles a set of parameters identified empirically or from physics/domain knowledge to be tied to a phenomenon (i.e., temperature rise, deformation, stress, etc.). The hope is to create a tool (or a model or a formula) to predict or estimate the response from the compiled parameters. The creation of a tool must satisfy the requirement of the method used to create the tool.

For example, if linear regression is used to create a model, then this model must conform to the assumptions used in developing linear models. One of the first assumptions an engineer must follow is to set the predictors (X_1, X_2, X_3,.., X_n) and response (Y). Given the linear nature of the linear regression, these predictors can only be linearly tied to the response. In this particular example, the regression is multilinear (see Eq. 2). Frequently, in multilinear regression analysis, each relation of these independent variables is weighted solely to ensure the dependency of the predicted variable.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \epsilon \quad \text{Eq. 2}$$

where β_0; is the y-intercept, β_n; is the regression coefficient of each independent variable X_n, and ϵ is the error.

There are more mechanical assumptions that an engineer must also satisfy when conducting linear and multilinear regression. These include:

- The presence of a linear relationship between the predictors and response.
- Any linear combination of the predictors must have a normal distribution.
- There exists no, or very little, correlation between the predictors.
- The independent of observations are statistically independent.
- Satisfying homoscedasticity (equal or similar variances in error among different groups).

Unfortunately, many of the above assumptions are left unchecked or implicitly assumed to be satisfied (please refer to Bazant et al. [25] for their analysis of data distribution in the shear database by ACI Committee 445 [25]) – as if they are not, then the application of traditional analysis is unlikely to be applied, and engineers may not realize a working model.

By doing so, the engineer assigns that the predictors can, in fact, be thought of to predict, estimate, or cause the response, with little regard to how true, such an assignment is, or to the fact that the predictors are independent of each other – and most importantly to the sufficiency the predictors provide (i.e., there are no other predictors that exist).

The above discussion on linear regression can be extended to nonlinear and symbolic regression (with suitable adjustments for the mechanical assumptions used in each method). The same can be carried over to machine learning, which has the least number of mechanical assumptions to satisfy.

3.0 Descriptive, predictive, and causal engineering queries

In general, queries pertaining to structural engineering problems can be roughly categorized into two classes, namely, 1) descriptive and 2) predictive. In hindsight, a third class also exists, referred to as causal, yet is rarely articulated as one. All of these are discussed herein.

The first type of query answers and describes engineering observations and their statistical relations. Engineers use such queries to establish a foundation for creating and tuning practical solutions (including formulae, methods, and theories). For example, this type of query can describe observations (i.e., on average, structures in coastal areas experience more corrosion issues than that inland) or compare the outcomes of experiments (e.g., all things being equal, larger columns have higher fire resistance ratings than smaller columns). The descriptive nature of these queries implies that descriptive questions cannot support counterfactuals (viz., they do not ask how a response would differ if some predictors were different).

On the other hand, predictive queries use data and observations to predict the response of interest given a set of *predefined* predictors. While descriptive queries are capable of addressing past and present observations, the second type of query can forecast future observations (whether at the individual or population level). The fact that these queries are built on predefined predicates implies that they, just like the first type of queries, cannot be used to identify or answer counterfactuals nor state the mechanisms for how and why the response occurs.

Finally, causal queries ask how a change(s) in structural response results from altering predictors. For example, a causal question can be: how would the midspan deflection curve change if a beam depth is increased by 20%? Another question could be: all things being equal, does increasing a column's section or improving the material's strength lead to a much improved fire resistance? Answering the above questions can simply be done via conducting an experiment on a new deeper beam and by testing two identical columns (with the expectation of one being with a larger cross section and the other being fabricated with strong grade materials). Such tests can be costly yet doable. At the very least, cost-effective numerical models could be developed to answer such questions. However, such models would need to be validated against some ground truth, and if we lack a physical test, then the validity of such models may not truly be established.

In some instances, if causal questions were raised at a system or community level, then physical testing nor numerical models can be used to answer such questions. Herein is where causal models can come in handy.

4.0 Causal diagrams

As can be seen, one of the key drawbacks of statistical models is that they exemplify parametric assumptions that are not known to be correct and may well be incorrect. Another drawback of common statistical models is their incapability of capturing all types of assumptions (e.g., size and scaling effects, fire exposures other than standard fire (ASTM E119), etc.).

An engineer must adopt causal assumptions to transcend beyond mere correlations and associations and realize causal relations. Such assumptions can be described via causal diagrams. These diagrams are graphical constructs that display the relationship between variables and responses in a web-like manner. This paper will focus on one type of causal diagram called a Directed Acyclic Graph (DAG).

A DAG consists of nodes (variables) and arrows (directed edges) that tie the variables. The directed edges are *directed* because they follow time to order, and they are *acyclic* because they do not create any loops. DAGs visually state and encode causal relations and assumptions between the variables and the response, allowing for causal or counterfactual interpretation. A sequence of any unbroken route along or against arrows from variable to variable is called a path (see Fig. 1). A causal path is one that starts with the directed route of arrows leading tail to head from one variable to another (see A → C, also B→D→C).

This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

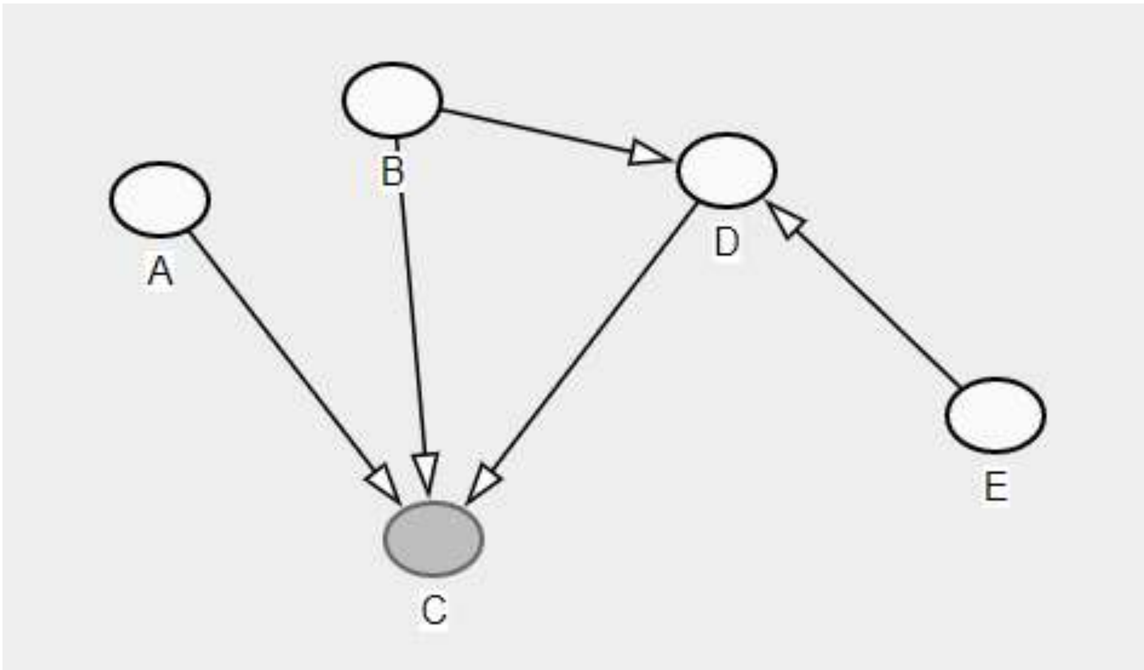

Fig. 1 A sample of a DAG

An intermediate variable along such a path that is not considered the change of interest nor the response is called a *mediator*. Three other types of intermediate variables are of interest, namely *confounders*, *colliders*, and *moderators* (see Table 1 for simplified definitions and refer to the selected references for a more detailed and in-depth discussion on various epidemiological and philosophical aspects). An *instrument* is an external variable (not intermediate) – such as E. As one can see, any predictor in a regression analysis could be any of the above variables. In some instances, a variable can be one type of intermediate variable in one analysis and another in a different analysis – like D. With special treatment in each case, a traditional statistical and engineering analysis is likely to be faulty as it will be prone to bias and misinterpretation.

Table 1 Key concepts

Term	Simplified definition s	Remarks & additional views on definitions
Confounder	A variable that has a causal effect on both the variables and/or variables and response, yet it is not affected by them (such as B in Fig. 1).	- Determining whether a variable is a confounder or not and whether it should be controlled requires logic (i.e., first principles, domain knowledge, etc.), as statistical analysis cannot determine this. - According to Cox and Wermuth [26], " ... *variables, U, whose omission seriously distorts the dependence of interest, but which were not observed, perhaps because their existence and nature were not appreciated*". - According to Yao et al. [27], "*Confounders are the variables that affect both the treatment assignment and the outcome*." - Please refer to more detailed definitions in [28], as well as a historical perspective confounding in [29].

This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

Collider	A variable that is a common effect of two other variables.	- A poor identification of a collider can generate distorted associations (such as D in terms of B and E in Fig. 1). - According to Elwert and Winship [30], " ... *a so-called collider variable, i.e., a variable that is itself caused by two other variables, one that is (or is associated with) the treatment and another that is (or is associated with) the outcome*." - According to Huntington-Klein [31], a collider is "*A variable is a collider along a path if both arrows on either side of it point at it.*".
Mediator	A variable that explains how the independent variable influences the dependent variable .	- An example of a mediator can be seen as B in Fig. 1. - According to MacKinnon et al. [32], "*A mediator is a variable that is in a causal sequence between two variables, whereas a moderator is not part of a causal sequence between the two variables*." - According to Imai et al. [33], *"... the intuitive notion about mediation held by applied researchers that the treatment indirectly influences the outcome through the mediator*." - According to Pearl [34], "*.., the mediator, whose role in transmitting the effect of T [treatment] on Y [outcome] we wish to assess*."
Moderator	A variable that affects the strength and direction of a causal relationship.	- According to Huntington-Klein [31], "*Moderators are variables that don't necessarily cause another variable (although they might do that too). Instead, they modify the effect of one variable on another*." - According to Bauman et al. [35], "*Moderator—an interaction variable that affects the direction, strength, or both of the relationship between an intervention and mediator …*" - According to Judd [36], "*A moderator variable is a variable, which is thought to temper or modulate the magnitude of the effect of an independent variable on a dependent one*."
Instrument	A variable that is associated with the exposure but is not related to the response except through its relationship with a variable.	- According to Morgan and Winship [37], "*An exogenous source of variation that determines Y only by way of the causal variable D.*" - According to Guo and Small [38], "*A valid instrumental variable is a variable that is independent of unmeasured confounders and affects the treatment but does not have a direct effect on the outcome beyond its effect on the treatment*." - According to Pearl [39], "*The traditional definition qualifies a variable Z as an instrumental (relative to the pair (X, Y)) if (i) Z is independent of all variables (including error terms) that have an influence on Y that is not mediated by X and (ii) Z is not independent of X.*"

These are some examples of causal diagrams as applied to the following hypothetical structural engineering problems. Please note that none of the presented examples have a working mathematical/statistical/physical model, nor can they be solved by collecting and analyzing data alone. The discussion will be confined to 3- and 4-variable problems as a more in-depth discussion on higher order causal diagrams will be conducted in a future study. Interested readers are invited to review the following articles [40,41] and reviews [42,43] on causal diagrams.

A case for a confounder in flood-prone zone due to strict building codes

Say that a dense zoning area (Z), such as a metropolitan in a flood-prone region, witnesses a limited number of structural failures (F) despite having a large number of structures (N) as compared to other zones within this region. A complimentary DAG for this case would be that shown in Fig. 2a. This observation could be attributed to several additional elements that are not immediately apparent but are crucial for comprehensive understanding.

Firstly, the strict codal provisions enforced in the zoning area of interest play a pivotal role. These regulations likely mandate rigorous construction methods and standards that account for flood risks. Such measures significantly reduce the vulnerability of structures to flooding and other related hazards, contributing to the observed reduction in structural failures. Moreover, the enforcement of strict codal provisions might also encourage the adoption of innovative construction techniques, such as elevated foundations, drainage systems, etc. These innovations further enhance the resilience of structures and substantially reduce the risk of structural damage during flood events within these densely populated zones. As mentioned earlier, confounders are best identified via logic rather than statistical analysis – which would return empty if applied to the above.

A case for a collider in the rise of structural issues

Say that structural issues (S), like cracking, continue to rise in a given bridge population in one state. Such issues started to grow in the aftermath of lack of funding (L) as well as poor maintenance practices (P) adopted in that state. This can be depicted in Fig. 2b. As one can see, the two problems of lack of funding and poor maintenance collied to amplify the structural issues in bridges.

More specifically, the lack of funding is a critical issue that adversely impacts the ability to allocate adequate resources for bridge maintenance/repair. One of the common consequences of this financial burden is a delayed or completely forgone maintenance routine, which is essential for identifying and addressing structural issues (S) before they escalate into significant structural problems. The situation is further complicated by the historical context of many bridges. Older bridges, designed and built according to past standards, may not be equipped to handle current demands in terms of load and resilience to various stressors. This discrepancy between design expectations and present-day realities necessitates regular maintenance and, in many cases, comprehensive upgrades or replacements, which are significantly hindered by funding and maintenance challenges.

Similarly, poor maintenance practices (P) further compound this issue since even when funds are available, ineffective maintenance or outdated repair techniques can lead to inadequate restoration efforts. Thus, minor wear and tear, which could have been managed with regular maintenance, can grow into critical structural issues over time. Simply, the interaction between lack of funding and

This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

poor maintenance practices creates a feedback loop that significantly amplifies the risk of structural failures. This often leads to a reactive rather than proactive maintenance approach, where actions are taken only after significant damage has occurred, further escalating the costs and complexities of repair.

A case for a mediator (and instrument) due to a seismic event

Say that an earthquake occurs. The magnitude of the earthquake (*E*) is the independent variable, the structural load (*M*) experienced by the building acts as the mediator variable, and the extent of building damage (*D*) is the dependent variable. Additionally, the location of the building - whether it is situated in a seismically active region or not - can be considered one form of an instrument variable (*I*) that influences the relationship between the earthquake magnitude (*E*) and the resulting structural load (*M*). In this scenario, the causal chain unfolds as: The earthquake magnitude (*E*) experienced by the building is influenced by the seismic activity of the region (*I*) in which it is located. This earthquake magnitude (*E*) then translates into the structural load (*M*) experienced by the building's components, with the level of load being the key mechanism through which the earthquake's energy is transferred. Ultimately, this induced structural load (*M*) leads to the observed extent of building damage (*D*). In this scenario, the seismic activity of the region (*I*) acts as an instrument, as it affects the magnitude of the earthquake (*E*) but does not directly influence the building's integrity or its eventual damage (*D*) – but rather, its effect is mediated through the earthquake magnitude (*E*) and the resulting structural load (*M*). (see Fig. 2c).

This understanding is crucial for structural design, guiding decisions on where and how to build resilient structures. In the context of the presented mediator analysis, understanding the role of seismic loads in transmitting earthquake energy to structural damage allows engineers to design better interventions that can be more effectively targeted to reduce the vulnerability of structures to earthquake-induced damage.

A case for a moderator in structural fire engineering

In this instance, the magnitude of deformation (*Y*) in a given structure can be amplified at elevated temperatures (*T*) while being under the same level of loading (*G*) as that seen at ambient temperatures. Such larger deformations may not be observed under ambient conditions (see Fig. 2d). In this example, the more intense the fire, the more intense the temperature and the more severe the fire-induced degradation in materials takes place, thus leading to larger deformations.

To reiterate, the core of this case study lies in understanding how elevated temperatures (*T*) impact the mechanical properties of construction materials. As temperatures rise, materials such as steel and concrete can lose stiffness and strength, contributing to increased deformation under load. As such, there is a clear association between the intensity of the fire with the temperature increase within the structure. As the fire becomes more intense, temperatures rise accordingly, leading to more severe degradation of material properties, resulting in larger deformations under the same load conditions (that would not cause such extensive deformation at ambient temperatures). This scenario highlights the significance of fire engineering in assessing and mitigating the risks associated with fire-induced deformations.

This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

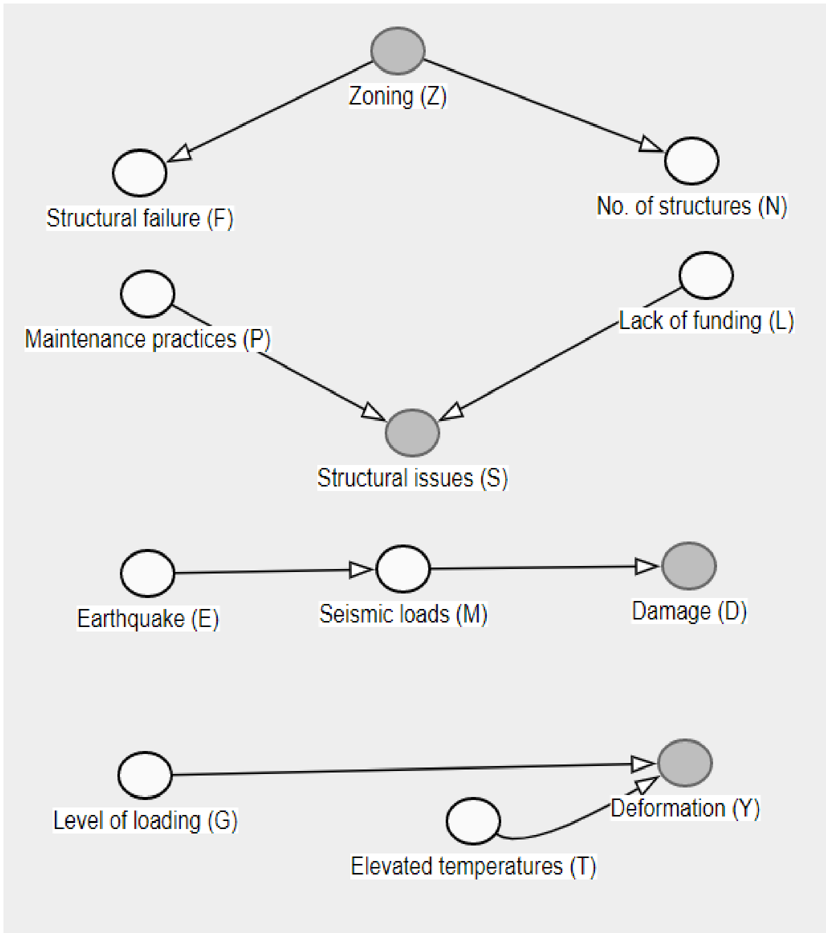

Fig. 2 Causal diagrams [Note: Top to bottom: a-d. Also, note that the slight curve in d distinguishes the mediator effect. Please note the existence of one than one way to plot a mediator (for example, via a straight edge that points at the edge between G and Y, as noted by Reviewer no. 1)

For this introductory short communication, a number of complex notions were not thoroughly presented. For example, the notion of independent and identically distributed (i.i.d.) data, complex case studies, and other causal principles, such as counterfactual reasoning, confounding bias, and selection bias, were only briefly touched upon for the sake of brevity and focus. Methods to address these issues, such as propensity score matching, etc., are complex and require a detailed exposition beyond the scope of this communication. The goal herein is to lay a foundational understanding, emphasizing how causal diagrams can be practically applied in structural engineering without overwhelming readers with the full depth of causal inference theory. The author invites interested researchers and future works to expand on these initial concepts by offering more detailed case studies and exploring advanced causal inference techniques. The author would also like to stress that structural engineers would indeed benefit from the various perspectives on the principles of causal models stemming from parallel domains such as medicine, psychology, economics, philosophy, etc.

This is a preprint draft. The published article can be found at: https://doi.org/10.1016/j.istruc.2024.106398.

Please cite this paper as:
Naser M.Z. (2024). "Causal diagrams for civil and structural engineers." *Structures*. https://doi.org/10.1016/j.istruc.2024.106398.

5.0 Conclusions

This paper presents a case for causal diagrams. Such diagrams can be of aid – especially in problems that lack physical representation or domain knowledge. A key motivation of this paper is to present such diagrams and how they can be implemented in structural engineering case studies. To go beyond regression and improve our predictive capacity, adopting causal principles and diagrams can be of merit.

Data Availability

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of Interest

The author declares no conflict of interest.

Acknowledgment

The author is thankful for www.dagitty.net, which allowed the creation of the presented DAGs. In addition, the author would like to thank the editor and reviewers for their support and for improving this work, especially with regard to the clarity of the presented examples.

References

[1] J. Pearl, D. Makenzie, The Book of Why_ The New Science of Cause and Effect-Basic Books, 2018.

[2] C. Wade, G. Cowles, R. Potter, P. Sanders, Concrete Blade Columns in Fire wade - Google Scholar, in: Concr. 97 Conf., Adelaide, Australia, 1997. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C41&q=Concrete+Blade+Columns+in+Fire+wade&btnG= (accessed November 6, 2022).

[3] ASCE, Standard calculation methods for structural fire protection, ASCE Stand. (2007). https://doi.org/10.1061/9780784408742.

[4] C.O. Orangun, J.O. Jirsa, J.E. Breen, REEVALUATION OF TEST DATA ON DEVELOPMENT LENGTH AND SPLICES., J Am Concr Inst. (1977). https://doi.org/10.14359/10993.

[5] Z.P. Bažant, Q. Yu, Minimizing statistical bias to identify size effect from beam shear database, ACI Struct. J. (2008). https://doi.org/10.14359/20096.

[6] A.H. Ali, H.M. Mohamed, C.E. Chalioris, A. Deifalla, Evaluating the shear design equations of FRP-reinforced concrete beams without shear reinforcement, Eng. Struct. (2021). https://doi.org/10.1016/j.engstruct.2021.112017.

[7] G.D. Hatzigeorgiou, Behavior factors for nonlinear structures subjected to multiple near-fault earthquakes, Comput. Struct. (2010). https://doi.org/10.1016/j.compstruc.2009.11.006.

[8] K.T. Farrow, Y.C. Kurama, SDOF Demand Index Relationships for Performance-Based

Seismic Design, Earthq. Spectra. (2003). https://doi.org/10.1193/1.1622955.

[9] Z.P. Bažant, Scaling of Structural Strength, 2005. https://doi.org/10.1016/B978-0-7506-6849-1.X5000-7.

[10] Z.P. Bažant, Q. Yu, Designing Against Size Effect on Shear Strength of Reinforced Concrete Beams Without Stirrups: I. Formulation, J. Struct. Eng. (2005). https://doi.org/10.1061/(asce)0733-9445(2005)131:12(1877).

[11] T.T. Lie, Fire Resistance of Circular Steel Columns Filled with Bar-Reinforced Concrete, J. Struct. Eng. (1994). https://doi.org/10.1061/(asce)0733-9445(1994)120:5(1489).

[12] American Wood Council, NATIONAL DESIGN SPECIFICATION® for Wood Construction , 2018. www.awc.org. (accessed November 7, 2022).

[13] M.Z. Naser, Machine Learning for Civil and Environmental Engineers: A Practical Approach to Data-Driven Analysis, Explainability, and Causality, Wiley, New Jersey, 2023.

[14] D.A. Kuchma, S. Wei, D.H. Sanders, A. Belarbi, L.C. Novak, Development of the one-way shear design provisions of ACI 318-19 for reinforced concrete, ACI Struct. J. (2019). https://doi.org/10.14359/51716739.

[15] H.-T.T. Thai, Machine learning for structural engineering: A state-of-the-art review, Elsevier, 2022. https://doi.org/10.1016/j.istruc.2022.02.003.

[16] A. Behnood, E.M. Golafshani, Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves, J. Clean. Prod. 202 (2018) 54–64. https://doi.org/10.1016/J.JCLEPRO.2018.08.065.

[17] S.M. Lundberg, S.I. Lee, A unified approach to interpreting model predictions, in: Adv. Neural Inf. Process. Syst., 2017.

[18] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you?" Explaining the predictions of any classifier, in: Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016. https://doi.org/10.1145/2939672.2939778.

[19] X. Chen, J. Abualdenien, M. Mahan Singh, A. Borrmann, P. Geyer, Introducing causal inference in the energy-efficient building design process, Energy Build. (2022). https://doi.org/10.1016/j.enbuild.2022.112583.

[20] M.Z. Naser, B. Ross, J. Ogle, V. Kodur, R. Hawileh, J. Abdalla, H.-T. Thai, Can AI Chatbots Pass the Fundamentals of Engineering (FE) and Principles and Practice of Engineering (PE) Structural Exams?, (2023). https://arxiv.org/abs/2303.18149v2 (accessed July 15, 2023).

[21] S. Wright, Correlation and causation, (1921). https://naldc.nal.usda.gov/catalog/IND43966364 (accessed November 4, 2022).

[22] F.K. Stage, H.C. Carter, A. Nora, Path Analysis: An Introduction and Analysis of a Decade of Research, J. Educ. Res. (2004). https://doi.org/10.3200/JOER.98.1.5-13.

[23] G. Gudjonsson, An easy guide to factor analysis, Pers. Individ. Dif. (1994). https://doi.org/10.1016/0191-8869(94)90040-x.

[24] J.B. Ullman, Structural equation modeling: Reviewing the basics and moving forward, J. Pers. Assess. (2006). https://doi.org/10.1207/s15327752jpa8701_03.

[25] Z.P. Bažant, H.T. Nguyen, A.A. Dönmez, Scaling in size, time and risk—The problem of huge extrapolations and remedy by asymptotic matching, J. Mech. Phys. Solids. 170 (2023) 105094. https://doi.org/10.1016/J.JMPS.2022.105094.

[26] D.R. Cox, N. Wermuth, Causality: A statistical view, Int. Stat. Rev. 72 (2004) 285–305. https://doi.org/10.1111/j.1751-5823.2004.tb00237.x.

[27] L. Yao, Z. Chu, S. Li, Y. Li, J. Gao, A. Zhang, A Survey on Causal Inference, ACM Trans. Knowl. Discov. Data. 15 (2021) 1–46. https://doi.org/10.1145/3444944.

[28] T.J. Vander Weele, I. Shpitser, On the definition of a confounder, Ann. Stat. (2013). https://doi.org/10.1214/12-AOS1058.

[29] A. Morabia, History of the modern epidemiological concept of confounding, J. Epidemiol. Community Health. (2011). https://doi.org/10.1136/jech.2010.112565.

[30] F. Elwert, C. Winship, Endogenous selection bias: The problem of conditioning on a collider variable, Annu. Rev. Sociol. (2014). https://doi.org/10.1146/annurev-soc-071913-043455.

[31] N. Huntington-Klein, The Effect : An Introduction to Research Design and Causality, Chapman and Hall/CRC, Boca Raton, 2021. https://doi.org/10.1201/9781003226055.

[32] D.P. MacKinnon, A.J. Fairchild, M.S. Fritz, Mediation analysis, Annu. Rev. Psychol. (2007). https://doi.org/10.1146/annurev.psych.58.110405.085542.

[33] K. Imai, L. Keele, D. Tingley, A General Approach to Causal Mediation Analysis, Psychol. Methods. (2010). https://doi.org/10.1037/a0020761.

[34] J. Pearl, Interpretation and identification of causal mediation, Psychol. Methods. (2014). https://doi.org/10.1037/a0036434.

[35] A.E. Bauman, J.F. Sallis, D.A. Dzewaltowski, N. Owen, Toward a better understanding of the influences on physical activity: The role of determinants, correlates, causal variables, mediators, moderators, and confounders, Am. J. Prev. Med. (2002). https://doi.org/10.1016/S0749-3797(02)00469-5.

[36] C.M. Judd, Moderator Variable: Methodology, in: Int. Encycl. Soc. Behav. Sci. Second Ed., 2015. https://doi.org/10.1016/B978-0-08-097086-8.44038-9.

[37] S.L. Morgan, C. Winship, Counterfactuals and Causal Inference, 2014. https://doi.org/10.1017/cbo9781107587991.

[38] Z. Guo, D.S. Small, Control function instrumental variable estimation of nonlinear causal

effect models, J. Mach. Learn. Res. (2016).

[39] J. Pearl, Causality, Cambridge University Press, Cambridge, 2009.

[40] S. Greenland, J. Pearl, J.M. Robins, Causal diagrams for epidemiologic research, Epidemiology. (1999). https://doi.org/10.1097/00001648-199901000-00008.

[41] J. Pearl, Causal Diagrams and the Identification of Causal Effects, in: Causality, 2013. https://doi.org/10.1017/cbo9780511803161.005.

[42] A.R. Nogueira, A. Pugnana, S. Ruggieri, D. Pedreschi, J. Gama, Methods and tools for causal discovery and causal inference, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 12 (2022) e1449. https://doi.org/10.1002/WIDM.1449.

[43] M.J. Vowels, N.C. Camgoz, R. Bowden, D’ya like DAGs? A Survey on Structure Learning and Causal Discovery, (2021). https://doi.org/10.48550/arxiv.2103.02582.