Canadian Science Publishing
NRC Research Press journals | Publishing Services

Environmental Reviews
Dossiers environnement

# Cause-effect analysis for sustainable development policy



| | |
|---|---|
| Journal: | *Environmental Reviews* |
| Manuscript ID | er-2016-0109.R2 |
| Manuscript Type: | Review |
| Date Submitted by the Author: | 24-Feb-2017 |
| Complete List of Authors: | Cucurachi, Stefano; University of California, Santa Barbara, Bren School of Environmental Science & Management<br>Suh, Sangwon; University of California, Santa Barbara, Bren School of Environmental Science & Management |
| Keyword: | Sustainable Development Goals, Causality, Cause-effect mechanisms, Quantitative Sustainability Assessment, Sustainability policy |
| | |

# Cause-effect analysis for sustainable development policy

Stefano Cucurachi [a], Sangwon Suh [a,*]

[a] Bren School of Environmental Science and Management, University of California, Santa Barbara, California 93106, United States

[*]corresponding author: e-mail: suh@bren.ucsb.edu, phone: (805) 893-7185, fax: (805) 893-7612

Word count: 15400

## Abstract

The sustainable development goals (SDGs) launched by the United Nations (UN) set a new direction for development covering the environmental, economic and social pillars. Given the complex and interdependent nature of the socio-economic and environmental systems, however, understanding the cause-effect relationships between policy actions and their outcomes on SDGs remains as a challenge. We provide a systematic review of cause-effect analysis literature in the context of quantitative sustainability assessment. The cause-effect analysis literature in both social and natural sciences has significantly gained its breadth and depth, and some of the pioneering applications have begun to address sustainability challenges. We focus on randomized experiment studies, natural experiments, observational studies, and time-series methods, and the applicability of these approaches to quantitative sustainability assessment with respect to the plausibility of the assumptions, limitations and the data requirements. Despite the promising developments, however, we find that quantifying the sustainability consequences of a policy action, and providing unequivocal policy recommendations is still a challenge. We recognize some of the key data requirements and assumptions necessary to design formal experiments as the bottleneck for conducting scientifically defensible cause-effect analysis in the context of quantitative sustainability assessment. Our study calls for the need of multi-disciplinary effort to develop an operational framework for quantifying the sustainability consequences of policy actions. In the meantime, continued efforts need to be made to advance other modeling platforms such as mechanistic models and simulation tools. We highlighted the importance of understanding and properly communicating the uncertainties associated with such models, regular monitoring and feedback on the consequences of policy actions to the modelers and decision-makers, and the use of what-if scenarios in the absence of well-formulated cause-effect analysis.

**Keywords**

Sustainable development goals; causality; cause-effect mechanisms; quantitative sustainability assessment; sustainability policy

## 1 Introduction

The Sustainable Development Goals (SDGs, hereafter) launched on January 1, 2016 include 17 goals, 169 targets and 303 indicators (United Nations 2014, Malik et al. 2015), which will help frame the agendas and policies of the United Nations' member states through 2030 (Hák et al. 2016). These goals are not only comprehensive, covering the economic, social and environmental dimensions of sustainability, but also highly interconnected (International Council for Science 2015), making it essential to understand synergies, trade-offs and conflicts between them in order to support decisions (Schindler and Hilborn 2015). Without such understanding, a policy to improve on one goal could conflict with another goal. For example, policies targeting at improving energy provision could conflict with another goal on climate-change mitigation, or those aiming at the protection of marine ecosystem could clash with the provision of sustainable food for all (Laurenti and Sinha 2015).

Various tools and metrics have supported sustainable development decisions, which we collectively refer to quantitative sustainability assessments (QSAs) in this review. Examples of QSAs include, but not limited to, life cycle assessment (LCA) (Guinée 2002, ISO 2006, Hellweg and Mila i Canals 2014), various footprinting approaches (Wiedmann and Minx 2007, Peters 2010, Hoekstra and Mekonnen 2012, Mancini et al. 2015, Michalsky and Hooda 2015), assessment of planetary boundaries (Rockström et al. 2009, Hughes et al. 2013, Whiteman et al. 2013, Steffen and Richardson 2015), environmental input-output models (Huppes et al. 2006, Tukker et al. 2006, Suh 2009, Hertwich 2010, Lenzen et al. 2012, Hertwich et al. 2014), ecosystem valuation approaches (Groot et al. 2010, Costanza et al. 2014), and material flow analysis (MFA) (Matthews et al. 2000, Brunner and Rechberger 2004, Haberl et al. 2007, Fischer-Kowalski and Swilling 2011), among others [see e.g. (Ness et al. 2007)]. In particular, so called, consequential LCA (CLCA) aims at quantifying the consequences that a certain action or a policy decision has on the environment and natural resources (Brander et al. 2008, Creutzig et al. 2012, Zamagni et al. 2012, Plevin and Delucchi 2014, Suh and Yang 2014).

The complexity and the interconnected nature of the socio-economic and environmental systems, however, poses a challenge to QSA practitioners in modeling the consequences of a policy action in the context of sustainable development (Cucurachi and Suh 2015). Furthermore, recent developments in economics, ecosystem science, and systems biology on causality research have yet to be embraced by QSA approaches.[1] Over the past decades, the causality literature has evolved to address various conceptual and technical issues such as endogeneity (Antonakis et al. 2014, Kreuzer 2016) and reverse causality [see e.g. (Mei-chu 1987, Chong and Calderon 2000, Barsky and Kilian 2004, Chaumont et al. 2012)] in parsing out causal relationships from complex phenomena. For example, Angrist and Krueger (1992) test the effect of children's age when starting school on their eventual years of schooling completed and on educational attainment. Using instrumental variables, the authors conclude that the effect of the starting age on educational attainment is modest. Instrumental variables have also been used to test the effects of education on health (Cutler et al. 2008, Grossman 2008, Conti et al. 2010, Cutler and Lleras-Muney 2010, Heckman et al. 2014), education on well-being (Oreopoulos and Salvanes 2011, Oreopoulos and Petronijevic 2013), and social connections on well-being (Kahneman and Krueger 2006, Fowler et al. 2008). However, few of such techniques have been applied to QSAs.

This review aims at surveying the techniques of cause-effect analysis in the context of QSAs. For each method to infer causality (cause-effect analysis technique in the remainder of the text), we present and review relevant applications in the field of sustainability that show how cause-effect analysis techniques can allow QSAs to increase the value of information they provide to decision makers. Our survey of

[1] For example, Weinzettel et al. (2013) use multi-linear regression and concluded that affluence drives the global displacement of land use, thus being the main cause of biodiversity loss globally. The study does provide a strong correlation between affluence and biodiversity loss but does not univocally allow interpretation of the results as a causal relationship. Likewise, Suweis et al. (2013, 2015) use a stochastic logistic model to assess whether the population growth of a nation is driven (i.e. caused) by either local availability of water resources used or by import of water resources from neighboring countries. As acknowledged by the authors, both studies do not consider a number of "other environmental, cultural, and health-related factors", thus limiting the interpretability of the result as a causal relationship. Some of these problems have been widely discussed and well understood in the causality literature (Aldrich, 1995; Rimer, 1998; Simon and Iwasaki, 1988).

causality literature was drawn from peer-reviewed articles on theory and methods, causality handbooks, and case studies applying the techniques. Based on the literature surveyed, we classify the analytical approaches to cause-effect analysis techniques. Each class of techniques was, in turn, searched on the ISI Web of Science and on Google Scholar in combination with the keywords 'sustain*', 'environ*', 'emissions', 'pollut*', 'econ*', 'CO2', and 'GDP'.

The remainder of the review is organized as follows: the next section presents a short chronology of causality theory; in section 3, we start from the ideal approach to causality provided by Rubin's causal model, and then we analyze the techniques that are based on observational (i.e. non-experimental) data; in section 4 we discuss the applicability of cause-effect analysis techniques to QSA; finally section 5 discusses outlook to close this review.

## 2 A brief chronology

Causality has interested philosophers and scientists since the time of Aristotle (see *Physics II 3 and Metaphysics V 2*). For millennia, however, causal problems have often rested in the realm of philosophical delight rather than inspiring scientific research.

Pearl (2000a) notes that the questions on causality did not enter into formal scientific discourses for a good part of the 19$^{th}$ century. In the dawn of the 20$^{th}$ century, Hume (1902 sec. VII) formally defined a cause as "an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second. Or, specifically, where, if the first object had not been, the second never had existed". A similar idea of cause was also at the basis of the experimental work of Mill (1856). However, Russel (1912) stated that causal relationships and physical equations are incompatible, describing causality as "a word relic" and excluding the existence of causality from mathematics and physics. In 1911, Pearson still described causality as "another fetish amidst the inscrutable arcana of even modern science" (Pearson 1911). Interestingly, a mechanistic view of causality also existed in the early 20$^{th}$ century philosophy. For example, Laplace thought that cause and effect can be understood perfectly given enough knowledge and data: "We may regard the present state of the universe as the effect of the past and

the cause of the future. An intellect which at any given moment knew all of the forces that animate nature and the mutual positions of the beings that compose it, if this intellect were vast enough to submit the data to analysis, could condense into a single formula the movement of the greatest bodies of the universe and that of the lightest atom; for such an intellect nothing could be uncertain and the future just like the past would be present before its eyes" (Laplace 1902).

In the 1950s, further formalizations of probabilistic causality appeared in the philosophical literature (Salmon 1980). Good (1963) and Suppes (1970) attempted to identify the tendency of an event to cause another by (1) constructing causal relations on the basis of probabilistic relations between events, (2) employing the statistical relevance as the basic concept, and (3) assuming temporal precedence of causes [see Russo and Williamson (2007) for a detailed account of probabilistic causality and of assumptions and axioms]. Probabilistic causality "places emphasis upon the mechanisms of causality, primarily uses concepts of process and interaction, and appeals to laws of nature" (Russo 2009).

In the 70s causality still remained as "one of the most important, most subtle, and most neglected of all the problems of Statistic" (Dawid 1979). It is only with the pioneering work of Rubin on a formal potential outcome/counterfactual analysis (Rubin 1974) that the statistical literature reconnects with causality and establishes a statistical definition of causality. The work of Rubin gave momentum to the development and application of statistical models, or cause-effect analysis techniques, which in the last decades have expanded into various applications including the foundational statistical principles set in the early work of Wright (1921) in the field of genetics.

## 3 Approaches to causality research

### 3.1 Correlation studies and their limitations

Cause-effect analysis techniques presented in this review enable answering three types of causal questions: (1) identifying causes (i.e. *why* a singular event occurs), (2) assessing effects (i.e. the *what-if* type of question, referred to the change in effect of some change in the cause), and (3) describing

mechanisms [i.e. *how* some effects follow from a certain cause (Holland 2003)]. Before we begin the review of the mainstream approaches to causality research, here we provide a brief discussion on correlation studies. As pointed out by many in the literature (Pearl 2000b), correlation and causation should not be confused. Positive correlation may be defined probabilistically for two variables, $X$ and $Y$ , as follows:

$$P(Y|X) - P(Y)P(X) > 0, \quad (1)$$

meaning that the probability that $X$ and $Y$ occur jointly is larger than the product of probabilities for each occurring independently. Similarly, negative correlation, can be defined as:

$$P(Y|X) - P(Y)P(X) < 0, \quad (2)$$

and the two variables, $X$ and $Y$ are uncorrelated if:

$$P(Y|X) = P(Y)P(X). \quad (3)$$

Correlation typically indicates that whenever $X$ occurs, there is a higher chance of observing $Y$ . A well-known example is that homelessness and crime rate are correlated, however, mere correlations do not provide a scientific evidence of whether homelessness causes crime, or that crime causes homelessness (Sugihara et al. 2012). The underlying cause could be another variable (e.g. unemployment) that may influence both.

## 3.2 Randomized experiment

### 3.2.1 Statistical differences in the outcomes of experimental studies

Experimental randomized studies , in contrast to correlation studies, provide an ideal means to inferring causality (Angrist and Pischke 2008).

In randomized experiments, individuals (or units) taken from a sufficiently large population are divided into two subgroups: one in which individuals receive a treatment (treatment group), and one in which

individuals do not receive a treatment (control group). Let us consider the case, in which a large number of similar cities are randomly divided into two groups. One group enforces a road space rationing and the other does not. We can define $T_i = \{0,1\}$, for all $i = 1,...,N$, as a binary random variable describing the treatment (e.g., enforcing a road space rationing or not enforcing a road space rationing). Let us define $Y$ as the variable to be explained, or response variable, such as the urban air quality.

The observed outcome for an individual $i$, $Y_i$, can be written as:

$$Y_i = \begin{cases} Y_{1i} \; if \; T_i = 1 \\ Y_{0i} \; if \; T_i = 0 \end{cases} \\ = Y_{0i} + (Y_{1i} - Y_{0i}) T_i. \quad (4)$$

In order for equation (4) to hold, Rubin (1978, 1980) defines the so-called stable-unit-treatment-value-assumption (SUTVA). The assumption implies that "a causal effect of one treatment relative to another for a particular experimental unit is the difference between the result if the unit had been exposed to the first treatment and the result if, instead, the unit had been exposed to the second treatment" (Rubin 1978). SUTVA rests on the idea that the potential outcome of one participant is not affected by the treatment applied to another participant. For example, one city institutionalizing a road space rationing policy does not affect another city in the experiment. Furthermore, it assumes that for each unit there is a single version of each treatment level (i.e. only one type of road space rationing of equal efficacy is used by all cities under study). The assumption introduced by Rubin (1980) holds if the value of $Y$ for any individual $i$ exposed to a treatment $T$ will be the same no matter what mechanism is used for the assignment of $T$ to $i$ for all individual participants and treatments (Rubin 1986) so that:

$$Y_i(T_1, T_2, ..., T_n) = Y_i(T_i). \quad (5)$$

The assumption is violated if multiple versions of the treatments or interferences (e.g. communication) between individual participants exist (Rubin 1986). The plausibility of the assumption has been subject

matter of a number of publications [we refer the reader to e.g. Sobel (2006) for more information on the issue]. It is, however, notable that this assumption is hardly plausible in a policy context, where a policy instrument is often modified or customized to the local or regional circumstances and policy outcomes are often benchmarked or publicized widely, directly or indirectly affecting others in the experiment. We will come back on this issue later in this review.

In the notation introduced in equation (4), $Y_{0i}$ is the potential outcome for an individual $i$ (e.g., air quality index, AQI for city *i*) had the individual not been exposed to the treatment (e.g., road space rationing), regardless of whether the individual is actually exposed to the treatment or not; whereas $Y_{1i}$ is the potential outcome had the individual been exposed to the treatment. In general, $Y_{1i} - Y_{0i}$ represents the causal effect of $T_i$ on $Y_i$ at the individual level. However, it is not possible to observe both potential outcomes simultaneously from any given individual (e.g., a particular city), since an individual is either exposed to treatment or control, not to both at the same time. Therefore, the aggregate causal effects and, in particular, the average causal effect (i.e. the average effect in the general population) is observed instead in reality.

The observed difference in average outcome (e.g., AQI) between the treatment group (e.g., cities enforcing road space rationing) and control (e.g., those not) can be expressed as $E\left[Y_i \middle| T_i = 1\right] - E\left[Y_i \middle| T_i = 0\right]$. For example, if the average AQI of the cities that exercise road space rationing is 5 and that for those not is 2 using a 1-to-10 quality scale (least to most severe pollution), the observed difference in average outcome becomes 3, which can be interpreted as a worsening effect. However, road space rationing is likely to be introduced to the cities with heavy traffic and air pollution in the first place, and therefore the observed difference in the AQI between the two groups cannot be directly translated into the causal effect of a road space rationing. This problem, referred to as 'selection bias', is elaborated further in the next section.

### 3.2.2 Rubin's causal model

The expected outcome of a group of individuals who were not exposed to the treatment can be expressed as $E\left[Y_{0i}\middle|T_i=0\right]$. Using the same example, this term shows the AQI of the cities that did not use road space rationing. The expected outcome for group of individuals that were exposed to the treatment, had the group not exposed to the treatment can be expressed as $E\left[Y_{0i}\middle|T_i=1\right]$. For example, this term would show the average severity of air pollution measured in AQI of those cities that exercise road space rationing, if they had not taken such a measure. Suppose that a group of cities have been using road space rationing. Suppose, further, that one can reverse the time and let the same group avoid using road space rationing. If this were possible, $E\left[Y_{0i}\middle|T_i=1\right]$ would be the current average AQI of these cities after reversing the time. However, this term is obviously not measureable. If it were measurable and if the treatment is independent of potential outcomes (i.e. with $T_i$ randomly assigned), the causal effect of the treatment, $E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=1\right]$, can be written as:

(6) $$\underbrace{E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=1\right]}_{\text{average treatment effect on the treated}}=\underbrace{E\left[Y_i\middle|T_i=1\right]-E\left[Y_i\middle|T_i=0\right]}_{\text{observed difference in response}}-\underbrace{E\left[Y_{0i}\middle|T_i=1\right]+E\left[Y_{0i}\middle|T_i=0\right]}_{\text{selection bias}}.$$

The term, $E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=1\right]$ represents the average causal effect of treatment for those who were treated (e.g., the difference in AQI as a result of using road space rationing). The term $E\left[Y_{0i}\middle|T_i=1\right]+E\left[Y_{0i}\middle|T_i=0\right]$ represents the selection bias (Angrist and Pischke 2008) that represents the fact that those who need treatment are more likely to seek treatment. For example, suppose that the average AQI of the cities that actually used road space rationing if they had not introduced road space rationing is 8, and that of those that did not is 2. In this case, the selection bias becomes 8 – 2 = 6, and therefore the right-hand-side of the equation becomes 3 – 6= – 3, meaning that the average road space rationing AQI improved on average by 3.

However, as noted earlier the term, $E\left[Y_{0i}\middle|T_i=1\right]$ cannot be directly observed or calculated. Therefore, one would have to find a counterfactual for this term in order to estimate the causal effect of the treatment in eq. (6) (Angrist and Pischke 2008). This can be obtained by the random assignment of $i$. Under the Rubin's causal model, the problem of spurious correlations discussed in the previous section can only be eliminated by using randomization of observations to the categories of a hypothesized causal factor (e.g., treatment vs. control) or by using a method that somehow mimics randomization process [(Morgan 2013); see section 3.3.1]. Randomization reduces the chance of intentional or unintentional bias, and it allows for effects and errors due to 'unaccounted-for' variables to act randomly, rather than consistently, affecting the response across treatments (Shaffer and Johnson 2008).

For example, random assignment, or 'randomizing' can be achieved by choosing the treatment and control groups with statistically equivalent level of AQIs. Random assignment makes the treatment $T_i$ independent of potential outcomes. In particular, $T_i$ is independent of $Y_{0i}$, thus allowing us to swap the terms $E\left[Y_{0i}\middle|T_i=1\right]$ and $E\left[Y_{0i}\middle|T_i=0\right]$ in the following expression:

$$\begin{aligned} &E\left[Y_i\middle|T_i=1\right]-E\left[Y_i\middle|T_i=0\right] \\ &=E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=0\right] \\ &=E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=1\right]. \end{aligned} \tag{7}$$

Given random assignment, Eq. (7) can be further reduced to:

$$\begin{aligned} &E\left[Y_{1i}\middle|T_i=1\right]-E\left[Y_{0i}\middle|T_i=1\right] \\ &=E\left[Y_{1i}-Y_{0i}\middle|T_i=1\right] \\ &=E\left[Y_{1i}-Y_{0i}\right]. \end{aligned} \tag{8}$$

The relationship identified in eq.(8) contains no selection bias, thus signifying, for example, that whether each individual city in the population under study has instituted a road space rationing policy or not, it

does not affect the identification of the causal effect. The effect of a randomly-assigned road policy on the city that implemented it is, in fact, the same as the effect of the road policy on a randomly chosen city.

### 3.3 Observational studies

For a while, much of the causality literature, in particular in the epidemiological, psychological and educational sciences (Campbell and Erlebacher 1970), has implied that only properly randomized experiments could lead to useful and trustable estimates of causal effects. However, as Rubin states (1974), such contention would be untenable if taken as applicable to all fields of science, given that much of the scientific development has been obtained for a big part of the past century without using randomized experiments. The statement still holds today, since randomized experiments are only feasible under certain conditions, and would probably be counter-productive in those contexts in which observational data is not immediately available. [2]

Conceptually, there are two major criticisms to Rubin's model. First, as discussed earlier, it is impossible to detect the individual causal effect, $Y_{1i} - Y_{0i}$, thus making the true causal effect impossible to detect (Russo et al. 2011). Putting this into a practical context, the same person (or city) cannot simultaneously take and not take a painkiller (or institute a policy) to observe the effect. In some cases, experiments can be done for the same unit over time. Second, Rubin's model is confined to a *Platonic heaven* situation, in which one can observe only average representations, rather than direct causal effects (Dawid 2007, p. 510).

At a more practical level, Rubin (1974) also noted that randomized studies cannot be widely applied when: (a) the cost of performing the equivalent randomized experiment to test for all potential alternatives (or treatments) is prohibitive; (b) there is a presence of ethical reasons according to which the treatments

[2] In their satire, Smith and Pell (2003) point out that the effectiveness of parachutes has never been proven using a randomized control trial.

cannot randomly assigned; or (c) the estimates based on the results from experiment indicate that it would require several years to be completed (Rubin, 1974).

For these reasons, researchers rely on observational data, i.e. data that were not generated using an experimental design. Observational data are obtained from surveys, longitudinal and panel data, censuses, and administrative records, and can vary both temporally and spatially (Christman 2008). Observational data are typically inexpensive to collect and are in plentiful supply (Iacus et al. 2012). Investigators using observational data (i.e. from observational studies) share the common objective of devising causal relations and, thus, face similar problems to *experimenters* (Cochran 2009). Complex interactions are also present in observational studies and can greatly complicate the interpretation of effects, although they reflect the inherent complexity of natural systems (Shaffer and Johnson 2008).

### 3.3.1 Matching methods and quasi-experimental designs

In the absence of a randomized experiment and when only observational data is available, cluster analysis techniques such as matching (Stuart 2010) allow for harnessing the benefits of Rubin's model by equating (or "balancing") the distribution of covariates in the treatment and control groups. Well-matched organized samples of the treatment and control groups can achieve such goal. The methods aim to replicate as closely as possible a randomized experiment, by pruning the observational dataset and making sure that the empirical distributions of covariates are similar (Ho et al. 2006, Stuart 2010). Treatment and control units are paired based on a number of observable pre-treatment covariates (i.e. observable characteristics).

The individuals in a group are paired solely for the purpose of obtaining the best possible estimate of the effect of a causal variable $T_i$ on an observed outcome $Y_i$. Using matching, differences in outcomes for units with different treatment levels but the same values for pre-treatment variables can be interpreted causally (Yang et al. 2015). For example, matching could be based on the probability of $T_i$ for each individual $i$ in the population, calculated as a function of $Q_k$, with $k = 1,...,V$, which represent the set of

background variables of interest, that is assumed to predict both $T_i$ and $Y_i$ (Morgan and Harding 2006). The matching procedure will select only matched sets of treatment and control cases that contain equivalent values for these predicted probabilities (Morgan and Harding 2006). The matching algorithm allows selecting from the joint distribution of $Q_{ik}$ and $Y_i$ only the information that is related to the causal variable (or treatment variable) $T_i$, and the procedure is conducted until the distribution of $Q_{ik}$ is equivalent for both the treatment and control cases, thus until the data are balanced, or matched (Morgan and Harding 2006).

Matching methods do not directly allow for making causal inferences, since they are data-processing algorithms not statistical estimators, thus they require the use of some type of causal estimator to make such inferences [e.g. testing the difference in means between $Y$ in the treatment and control groups; see (Iacus et al. 2012)]. As Stuart (2010) points out, after the analyst has created treatment and control groups with adequate balance, and designed the observational study, the analysis moves to the outcome interpretation stage. At this stage, the analysis will typically be limited to techniques of regression adjustment using matched samples and use regression-based techniques in combination with the matched samples. Matching methods, in fact, are best used in combination with regression models (see section 3.3.2), instrumental variables models, or structural equation models [SEM (Ho et al. 2006)].

Matching techniques have been widely used in economics (Abadie and Imbens 2006), medicine (Christakis and Iwashyna 2003), and sociology (Morgan and Harding 2006), among other fields of science [see also (Sekhon 2011)]. Commonly used matching methods include difference-in-differences matching (Abadie 2005), multivariate matching based on the Mahalanobis distance metric (Cochran et al. 1973), nearest neighbor matching (Rubin 1973), propensity score matching (Caliendo and Kopeinig 2008), genetic matching (Diamond and Sekhon 2012), and coarsened exact matching (Iacus et al. 2012) [see (Stuart and Rubin 2008) for a review]. Quasi-experimental designs using the treatment and control

duality also include difference in differences techniques used with longitudinal data, for which we refer the reader to (Abadie 2005, Athey and Imbens 2006, Donald and Lang 2007, Puhani 2012).

Observational studies become relevant if performed on all causally-important variables and on several control groups that are each representative of a potentially different bias (Rubin 1974). Observational studies do require the analyst to carefully study the process of data generation and the treatment/assignment mechanism (Iacus et al. 2012). In observational studies without randomization the analyst uses the design phase to help with approximating hypothetical randomized experiments. The so-called identification strategy describes the manner in which a researcher uses observational data not generated by a randomized trial to approximate a real experiment (Angrist and Pischke 2008). The use of an observational study allows estimating the average effect on the treated (or ATT) and the average treatment effect (or ATE), based on data availability (Stuart 2010).

#### 3.3.2 Regression-based causality

SEM have become a core method for assessing causality in the social sciences, especially for research questions that cannot be tackled by experimental testing (Pearl 2009). The variables of interest for causal research are for this reason also called *latent variables*, because of their inaccessibility through direct measurement without a substantial measurement error (Bollen 2002). In many cases, it is impossible or too expensive to conduct controlled experiments, but SEM allows for discovery of likely causal relations from observational data (Shimizu et al. 2006).

SEM can also be combined with graphical constructs that allow laying out the causal relationships under analysis pictorially. A particular kind of graph used in causal analysis is the directed acyclic graph (DAG) or Bayesian network (Pearl 1995, Morgan 2013). DAGs are visual representations of qualitative causal assumptions and can be related to probability distributions linked to the data under study and to causal frameworks.

Causal models are usually characterized by the presence of a set of explanatory variables or covariates $X$ (i.e. the putative causes) and a response variable $Y$ (i.e. the putative effect) in the form, for instance, of a simple structural equation:

(9) $$Y = \beta X + \varepsilon,$$

where $\beta$ is the causal effect on $Y$ for a one unit difference in $X$, representing the coefficient determining the extent of the influence of $X$ on $Y$, and $\varepsilon$ represents the errors, unmeasured factors, or all other influences on $Y$.

The interpretation of $\varepsilon$ and $\beta$ is not trivial. Error terms may be interpreted deterministically or epistemically (Russo 2009). In the first case, we may assume that errors represent the lack of knowledge of the analyst. Thus if complete knowledge would be in hand, a precise relationship, between $X$ and $Y$, could be determined without error. The SEM reports deterministic causal relations. In the epistemic acceptation of the concept, the SEM represents causal relations that are thought to be genuinely indeterministic, thus errors are to be modelled probabilistically (Russo 2009). This second acceptation is the one we hold in this review.

The parameter $\beta$ has in the context of SEM a causal interpretation, thus it should quantify the extent of the causality. Thus, we can define (Russo 2009):

(10) $$\beta = r\frac{\sigma_X}{\sigma_Y}.$$

The correlation coefficient $r$ can be calculated as the ratio between the covariance $\sigma_{XY}$ and the variances $\sigma_X$ and $\sigma_Y$:

(11) $$r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Similarly, we may proceed and calculate all $\beta s$ and $\delta s$ in the SEM.

Let us now consider the example below representing a generic bivariate regression equation:

(12) $$Y = \alpha + \beta X + \varepsilon$$

where $\alpha$ is the intercept and $\varepsilon$ is the error term. In a causal interpretation of Eq. (12) $\beta$ represents the structural causal effect that applies to all members of the population of interest. Thus, in addition to being linear, this equation says that the functional relationship of interest is the same for all members of the population. Logarithmic transformations or other functional transformations of the variables of interest in the model can be typically considered (Baiocchi 2012). The ordinary least squares estimator of the bivariate regression coefficient $\beta$ is then (Morgan and Winship 2007):

(13) $$\beta_{OLS} = \frac{\sigma_{YX}}{\sigma_X}.$$

The above is just an example of the application of regression techniques for the estimation of the regressors of interest. Regression techniques provide a good estimation of the causal parameters, if the error terms in SEM are uncorrelated with the regressor (see assumptions in section 4.1). The coefficient of determination $r^2$ may be used to evaluate the goodness of fit of the model. Example of regression techniques include least squares and partial least squares techniques (Wold 1982, Angrist and Imbens 1995, Tenenhaus et al. 2005, Esposito Vinzi et al. 2010). In the next section we focus on the causal interpretation of regression techniques and on the instrumental variable approach. Further applications of regression-based techniques include regression-discontinuity designs, for which we refer the reader to (Hahn et al. 2001, Imbens and Lemieux 2008, Lee and Lemieux 2010).

#### 3.3.2.1 Causal interpretation of regressions

We focus on this section on the causal interpretation of regressions as estimators of causality. We refer the reader to (Berk 2004, Gelman and Hill 2006, Morgan and Winship 2007, Angrist and Pischke 2008,

Freedman 2009, Hansen 2015) for a complete presentation of regression techniques and for a complete analysis of the limitations of such approaches.

Regressions do not necessarily hold a causal interpretation, and they can be simply interpreted as a descriptive tool or as "a technique to estimate a best-fitting linear approximation to a conditional expectation function that may be nonlinear in the population" (Morgan and Winship 2007). However, regression, if well specified, can provide information about the causal relation between $X$ and $Y$ . It is the more ambitious question of when a regression has causal interpretation that concerns us in this review, due to its applicability for complex systems under study for QSA. To arrive at a causal model from a regression model, the analyst aims to study how one variable would respond, if one intervened and manipulated other variables (Freedman 2009). This implies that the causal results from a regression-based cause-effect analysis depend on the hypothesis framework of the analyst. It is within this framework that causality can be determined.

Let us assume that $X_i$ is a vector of covariates that are associated in some way with a response variable $Y$ . The conditional expectation function (CEF) of $Y$ is denoted as $E\left[Y\middle|X_i\right]$ and denoted as $E\left[Y\middle|X_i=x\right]$ for any realization $x$ of $X_i$ [see (Angrist and Pischke 2008) for a formal definition and proof of theorems]. Least squares regression allows the calculation of a regression surface that is a best-fitting linear-in-the-parameters model of $E\left[Y\middle|X_i\right]$, thus of the association between $Y$ and any realization $x$ of $X_i$, minimizing the average squared differences between the fitted values and the true values of $E\left[Y\middle|X_i=x\right]$ (Morgan and Winship 2007, Angrist and Pischke 2008).

A regression can be considered causal when the CEF it approximates is causal, or when the CEF describes differences in average potential outcomes for a fixed reference population (Angrist and Pischke 2008). As discussed in section 3.2.1, experiments with random assignments ensure that the causal

variable of interest is independent of potential outcomes, thus the groups under comparison are effectively comparable. A core assumption for the causal interpretation of regression, is the conditional independence assumption [or CIA; see (Rosenbaum 1984, Lechner 2001, Angrist and Pischke 2008)], which is at the basis of most empirical work in economics. The CIA is required for a regression to identify a treatment effect. The experimental design introduced in section 3.2 ensures that the causal variable of interest is independent of potential outcomes, which guarantees that the groups being compared are truly comparable (Angrist and Pischke 2008). This notion can be embodied regressions that are causally interpreted. CIA, also called as selection-on-observables, determines that the covariates to be held fixed are assumed to be known and observed. As a consequence, according to this assumption the residual in the causal model is uncorrelated with the regressors. Regression can be used as an empirical strategy to turn the CIA into causal effects. Under CIA the covariates $X_i$ are held fixed for the causal inference to be valid. These control variables (or covariates) are assumed to be known and observed (Angrist and Pischke 2008).

Let us consider a generic causal model:

(14) $$f_i(B) = \alpha + \rho B + \eta_i,$$

where $B$ is a variable that can take on more than two values. The equation is linear and assumes the functional relationship under consideration being the same for all individuals in the population under study. Unlike the factor $\eta_i$ that captures all unobserved factors determining the outcome for each specific individual $B$ is not indexed per individual. The causal model, therefore, tells us the extent of $B$ for any value of $B$ and not for a specific realization $B_i$ . We can further specify the causal model for the individual case, thus we consider that the causal relationship between putative causes and response is likely to be different for each individual, as in:

(15) $$Y_i = \alpha + \rho B_i + \eta_i.$$

A classic example is that $B_i$ could be the number of years of schooling for a certain individual and $Y_i$ could represent the current salary for that individual (Angrist and Krueger 1992). Eq. (15) is similar to a bivariate regression model. However, it is Eq. (14) that explicitly associates in the model constructed by the analyst the coefficients in Eq. (15)with a causal relationship, thus establishing the causal association. The causal model determines that $B_i$ may be correlated with $f_i(B)$ and the residual term $\eta_i$.

We can, then, consider the vector of covariates $X_i'$.The random residual part of Eq.(15) $\eta_i$ can be decomposed under CIA into a linear function of observable characteristics $X_i'$ and an error term $\upsilon_i$ :

(16) $$\eta_i = X_i'\gamma + \upsilon_i,$$

where $\gamma$ is a vector of population regression coefficients that satisfies the relationship $E\left[\eta_i \middle| X_i\right] = X_i'\gamma$. The vector $\gamma$ is defined by the regression of $\eta_i$ on $X_i$, thus the residual $\upsilon_i$ and $X_i'$ are uncorrelated by construction [see (Angrist and Pischke 2008) for further details and proof of concept]. By virtue of CIA, we can define (Angrist and Pischke 2008):

(17) $$E\left[f_i(B)\middle| X_i, B_i\right] = E\left[f_i(B)\middle| X_i\right] = \alpha + \rho B + E\left[\eta_i \middle| X_i\right] = \alpha + \rho B + X_i'\gamma.$$

We can re-write the causal model as:

(18) $$Y_i = \alpha + \rho B_i + X_i'\gamma + \upsilon_i.$$

The residual in the causal model is uncorrelated with the regressors $B_i$ and $X_i$, thus $\rho$ effectively represents the causal effect of interest, allowing for the attribution of causal meaning to the regression. The selection of the right set of control variables is the subject of an extensive literature. We refer the reader to Angrist and Krueger (2001) and Angrist and Pischke (2008) for a detailed analysis of the matter.

426 **3.3.2.2 Instrumental variables and causality**

427 We have just seen how regressions can be causally interpreted within the boundaries of a specific model.

428 A major complication is the possibility that regressors and errors [e.g., $B_i$, $X_i'$, and $\upsilon_i$ in the example in

429 Eq.(18)] are correlated, thus undermining the statistical validity of the model. Under such condition,

430 regression estimates would lose their causal interpretation. For the causal interpretation to hold, the

431 regressors have to be asymptotically uncorrelated with the errors or residuals. The potential inconsistency

432 is determined by the fact that changes in $B_i$ are not only associated with changes in $Y_i$ but also with

433 changes in $\upsilon_i$.

434 We consider that the potential outcomes can be written as (Angrist and Pischke 2008):

435 (19) $$Y_i = \alpha + \rho B_i + A_i'\gamma + \upsilon_i.$$

436 Here $A_i'$ is a vector of control variables, which unlike $X_i'$ in the example in Eq. (18)is unobserved.

437 Instrumental variable methods (Heckman and Vytlacil 2001, Newey and Powell 2003, Firebaugh 2008,

438 Bollen 2012) allow the analyst to introduce an instrumental variable, say $Z_i$, that is correlated with the

439 causal variable of interest $B_i$, and uncorrelated with both $A_i'$ and $\upsilon_i$, such that $E[Z_i\upsilon_i] = 0$. Such a

440 condition is a special case of CIA introduced in the previous section. In this case it is the instrumental

441 variable $Z_i$ that is independent of potential outcomes, rather than the variable of interest $B_i$. It follows

442 then that the causal effect $\rho$ can be expressed as (Angrist and Pischke 2008 chap. 4):

443 (20) $$\rho = \frac{\frac{\sigma_{Y_iZ_i}}{\sigma_{Z_i}}}{\frac{\sigma_{B_iZ_i}}{\sigma_{Z_i}}}.$$

444 The equality in Eq. (20) is verified if:

- $Z_i$ has a clear effect on $B_i$;
- $Z_i$ affect $Y_i$ only by means of the causal variable $B_i$;
- $Z_i$ is independent of potential outcomes, so it is as good as if randomly assigned.

The consideration of instrumental variables allows for the causal interpretation of $\rho$ . Instrumental variables are identified case by case from the processes determining the variable of interest. For the example of the relationship between schooling level and earnings, Angrist and Krueger (1992) used the school start age of pupils as an instrumental variable. Instrumental variables solve the problem of missing or unknown controls. In many cases, in fact, the necessary control variables are typically unmeasured or simply unknown. In the absence of suitable instrumental variables in the system the causal framework does not hold.

There are some recognized pitfalls of the instrumental variable approach (Morgan and Winship 2007). In some cases the assumption that the instrumental variable does not have a direct effect on the response variable may be too strong. Even when such condition is verified, an instrumental variables estimator is biased in a finite samplesample(Morgan and Winship 2007). These pitfalls may influence the possibility of drawing causal inference from the results of a study (see section 4.1). The limitations of regression-based methods should be carefully considered for the causal analysis to be valid. A causal regression may be invalidated by omitting variables that both affect the dependent variable and are correlated with the variables that are studied in the causal regression model, by the way missing data is handled, and by the presence of potential biases determined by measurement errors (Allison 1999).

### 3.3.3 Applications

We survey here the application of regression-based techniques and combined matching and regression techniques in the field of sustainability.

Empirical analyses using causal regression techniques have been widely applied to study the relationship between trade openness, economic development and environmental quality (Stern 2004, Copeland and Taylor 2013). In the Environment Kutznets Curve literature, a considerable amount of studies deal with this relationship, treating environmental degradation measures as the dependent variables and income as the independent variable, and providing mixed results (Soytas et al. 2007).

Antweiler et al. (1998) find that international trade, although altering the pollution intensity of countries, creates small changes in pollution concentrations, especially of $SO_2$. The authors find evidence that both environmental regulations and capital-labor endowments determine $SO_2$ concentrations and conclude that openness and freer trade appear to be good for the environment. The study concludes that if an increase in trade openness generates a 1% increase in income and output then, as a result of scale and technology pollution does fall by approximately 1%. Cole and Elliott (2003) confirm both environmental regulation effects and capital-labor effects for $SO_2$ and suggest that these results do not necessarily hold for other pollutants, such as $NO_x$, biochemical oxygen demand (BOD) and $CO_2$, for which an increase in emissions is likely to happen as a result of freer trade.

Frankel and Rose (2005) study the effect of trade on the environment and use exogenous geographic determinants (i.e., lagged income, population size, rate of investment, and human capital formation) as instrumental variables to account for the endogeneirty of trade. The authors conclude that trade appears to have a beneficial effect on some measures of environmental quality. In particular, they conclude that trade significantly tends to reduce the concentrations of $SO_2$ and $NO_2$. Managi et al. (2009) find that trade is beneficial for OECD countries, while it has detrimental effects on $SO_2$ and $CO_2$ concentrations in non-OECD countries. A lower BOD is found in non-OECD countries. The detrimental impact is found to be larger in the long term, rather than in the short term.

A bulk body of research regards the accumulation of greenhouse gases (GHGs) in the atmosphere leading to climate change. Regression techniques of econometric inspiration are commonly applied for the study of the influence of climate change on a number of endpoints. The matter of adaptation under climate

change is analyzed using nonlinear regression in Schlenker and Roberts (2009). The author controls for precipitation, technological change, soils, and location-specific unobserved factors, and the results show a nonlinear relationship between temperature and soil yields. The relationship between mortality and changes in daily temperatures is described using regression techniques in Barreca et al. (2013). The authors document a remarkable decline in the mortality effect of temperature extremes in the 20th century in the United States, and point to air conditioning as a central determinant in the reduction of mortality risks associated with extreme temperatures. The exposure to extreme temperatures determined by climate change is linked to deleterious effects on fetal health, the decrease in birth weight, and an increase in the probability of low birth weight in Deschenes et al. (2009 p. 216). The analysis rests on a number of strong assumptions about data, including that the climate change predictions used in the regression model are correct. In a similar fashion, climate policy has been linked to increase in mortality and migration (Deschenes and Moretti 2009), fluctuations in the labor markets (Deschenes 2010), and reduced profits from agriculture in the United States (Deschenes and Greenstone 2007) and in California (Deschenes and Kolstad 2011). Conflicts and social instability have also been associated with climate change (Homer-Dixon 1991). Earlier studies have shown that random weather events, such as drought and prolonged heat waves, might at times be correlated with armed conflict in Africa (Miguel et al. 2004, Smith et al. 2007, Burke et al. 2009). Hsiang et al. (2011) show that a causal link between temperature and conflict does exist at various scales for relatively richer countries as well. The issue of causal links between climate and conflict is contentious (Cane et al. 2014, Raleigh et al. 2014). Buhaug (2010 p. 16480) investigated the scientific base of the claims and concluded that "a robust correlational link between climate variability and civil war do not hold up to closer inspection" when alternative statistical models and alternative measures of conflict are used. Hsiang and Meng (2014) reproduced the analysis of Buhaug (2010) and corrected the correct the statistical procedure for model comparison. The study concludes that the claim of Buhaug (2010) is inconsistent with the evidence presented, thus climate change does affect conflicts in Africa (Hsiang and Meng 2014).

The potential sustainable impacts of fair trade, eco-certification and eco-labelling have been amply studied using matching techniques in combination with regression techniques. Ruben et al. (2009) use data from coffee and banana co-operatives in Peru and Costa Rica and find, using propensity score matching, that fair trade improves access of farmers to credit and investments, and also affects their attitude towards risk. The participation in a fair trade system improved employment, as well as their bargaining power and trading conditions. The difference-in-differences identification strategy is used by Hallstein and Villas-Boas (2013) to test the efficacy of eco-labels in promoting sustainable seafood consumption. The study finds evidence that in a sample of ten stores in the San Francisco Bay area the implementation of an eco-label led to a significant decline in sales in the range of 15%-40% of certain classes of products with limited environmental sustainability. Miller et al. (2011) use difference-in-differences to test the impact of a scheme of cash transfer on food security in Malawi. The study presents evidence that food security is improved by the transfer of cash by the government to rural households in Malawi. Eco-certification is also the subject of the study of Blackman and Naranjo (2012). The study uses propensity score matching to control for selection bias and tests the impact of eco-certification on a high-value agricultural commodity, organic coffee from Costa Rica. The study finds that organic certification improves the environmental performance of coffee growers by reducing the use of chemicals and improving the environmental performance of management practices.

Matching techniques have been used also to check progress on poverty reduction and on other goals in the Millennium Development Goals (Sachs and McArthur 2005). Maertens et al. (2011) use a variety of matching techniques to test the impact of globalization on poverty reduction in Senegal. The study finds a significant positive impact of globalization on poverty reduction through employment creation and labor market participation. Setboonsarng and Parpiev (2008) test the impact of microfinance on the MDGs using data from a microfinance institution in Pakistan. Using difference-in-differences, the study finds that the lending program of the institution contributed to income generation activities that have a beneficial impact on the MDGs. Arun et al. (2006) use propensity score matching to test whether

microfinance reduces poverty in India and show that microfinance institutions have a significantly positive effect on poverty reduction. Arnold and co-authors (2010) draw on the potential outcome model for causal inference and use a matched cohort to test the relationship between health and development. In a matched sample of 25 villages in rural India the study finds a positive influence on health from new toilet construction, while no impact was found from height-for-age.

In the field of sustainable fisheries, Costello et al. (2008) apply propensity score matching to evaluate the benefits of tradable harvest quotas (i.e. catch shares) on preventing the collapse of global fish resources. The study finds that the implementation of catch shares halts, and even reverses, the global trend toward widespread collapse of fish resources. The results are confirmed using propensity score matching by the same research group (Costello et al. 2010).

Quasi-experimental designs have been used to evaluate the biodiversity and social impacts of conversation and protection practices. Linkie et al. (2008) evaluate the impact of protected area on the conservation of species in a large protected area in Indonesia. The study uses propensity score matching to compare the deforestation rates in villages around the protected area and villages not around the area. The study finds no evidence of a positive effect of the protected area on the reduction of deforestation. Nelson and Chomitz (2011) test the impact of protected areas in reducing fires in tropical forests in various regions. The study finds that protected areas substantially reduced fire incidence in Latina America, Asia, and Africa. Matching criteria in this study included the distance to road network, distance to major cities, elevations and slope, and rainfall. Andam and co-authors (2008) apply matching methods to evaluate the impact on deforestation of Costa Rica's renowned protected-area system between 1960 and 1997. The institution of protected areas reduces deforestation and 10% of forests would have disappeared without being protected. Ferraro and Hanauer (2014) use a quasi-experimental design to study the mechanisms through which the policies of establishing protected areas affect social and environmental outcomes. The authors analyze the causal effects of ecosystem conservation programs on environmental and social outcomes, by focusing on the mechanisms determining variations that arise in a

certain area after land-use restrictions have been put in place. The study uses an asset-based poverty index developed by Andam et al. (2010) and investigates the effect of protected areas on this index. Therefore, the population is divided into treatment and treated group, respectively the causal effect of protected areas and the people living around protected areas. After controlling for potential confounding variables and biases, the authors conclude that two-thirds of the poverty reduction in Costa Rica can be causally attributed to opportunities afforded by tourism, while changes in infrastructure or land cover had a little causal influence on the outcome.

## 3.4 Time-series methods

### 3.4.1 Granger-causality

The vast availability of time-series data has given rise to a plethora of methods aimed at understanding complex systems through studying their evolution in time. Time-series refer to data observed over a number of discrete time-steps. In such cases, one may assume that causes both precede and help predict their effects. We credit Wiener (1956) with the intuition that the causality of a (time-series) variable in relation to another can be measured by how well one variable helps to predict the other. We can say that variable $Y$ 'causes' variable $X$ if the ability to predict $X$ is improved by incorporating information about $Y$ in the prediction of $X$. The concept was later formulated by Granger (1969), leading to the establishment of the Wiener-Granger framework of 'causality'. Geweke made several other important contributions to the concept (Geweke 1982, 1984).

Let us consider two variables $X$ and $Y$. We can say that $X$ does not cause $Y$ if the conditional distribution of the effect Y only depends on the past history of $Y$ itself and not on the past history of the putative cause $X$.

Let us consider the history of $X$ and $Y$, respectively $\{X\}$ and $\{Y\}$

589 (21)
$$\{X_t\} = (X_t, X_{t-1}, X_{t-2}, ..., X_L)$$
$$\{Y_t\} = (Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_L)$$

590 with $L$ representing the number of time lags.

591 According to GC we say that $X$ does not *Granger-cause* $Y$ if $Y_t$ only depends on its history but not on
592 the history of $X$, alternatively expressed as (Granger 1969):

593 (22)
$$P(Y_t | X_{t-1}, Y_{t-1}) = P(Y_t | Y_{t-1}).$$

594 No relevant information can be extracted from the history of $X$ to assess the effect $Y$, thus the
595 predictability of $Y$ does not increase if $X$ is part of the universe of possible causative variables.
596 Conversely, if $X$ Granger-causes $Y$, one can generally say that the past of $X$ contains information that
597 helps predict the future of $Y$.

598 GC models typically use vector auto-regression (VAR) models to analyze multivariate time-series. VAR
599 models are simple constructions in which the value of a variable at a particular time is modeled as a linear
600 weighted sum of its own past and of the past of a set of other variables. Each variable is a vector
601 stochastic process representing a time-series. The structure of the VAR model provides information about
602 the forecasting ability of a variable or of a group of variables. Therefore, GC does not directly imply true
603 causality, but rather only implies forecasting ability (Zivot and Wang 2006).

604 Operationally, GC analysis rests on estimating and comparing the VAR models, given a set of time-series
605 data. Let us expand our example consider a third variable $Z$ together with $X$ and $Y$. We are still
606 interested in measuring whether $X$ Granger-causes $Y$. The analysis starts with the joint estimation of a
607 full VAR model for all the variables. A prediction/estimation error is computed for all the variables in the
608 set (Seth et al. 2015). A second VAR model is then estimated, omitting $X$ from the universe of all
609 possible causative variables. For each remaining variable a new set of prediction/estimation errors is

calculated. If the prediction/estimation error for $Y$ is significantly smaller for the full regression including $X$, then we may confidently state that $X$ Granger-causes $Y$, conditioned on $Z$. As reported in (Seth et al. 2015), the magnitude of the GC is given by the ratio of the variance of the prediction-error terms for the reduced regression and the full regressions.

The standard test of GC developed by Granger (Granger 1969) is based on a linear regression model:

$$Y_t = a_0 + \sum_{k=1}^{L} b_{1k} Y_{t-k} + \sum_{k=1}^{L} b_{2k} X_{t-k} + \xi_t, \quad (23)$$

where $\xi_t$ are uncorrelated random variables with zero mean and variance $\sigma^2$, $L$ is again the specified number of time lags, and the time is $t = L+1, \ldots, S$. The null hypothesis that $X_t$ does not Granger cause $Y_t$ is supported when $b_{2k} = 0$ for $k = 1, \ldots, L$, thus allowing Eq. (23) to reduce to:

$$Y_t = a_0 + \sum_{k=1}^{L} b_{1k} Y_{t-k} + \xi_t. \quad (24)$$

Test statistics that can be applied to test the hypothesis are reported in Hlaváčková-Schindler et al. (2007).

GC has a number of useful properties [see e.g. (Geweke 1982, Seth et al. 2015)], including that VAR models may be estimated using relatively simple computation algorithms (e.g. ordinary least squares). Furthermore, the analyst needs to make only minimal assumptions about the underlying physical mechanisms linking the variables under study as long as they rest on data that is suitable for VAR modelling. GC is based on the comparison of model errors; therefore, the analysis is applicable only to the case of stochastic variables, i.e. variables that can be modelled as having random variations, and to data that have variance and mean that are stable over time (Geweke 1984).

A number of limitations, including those identified by Granger himself (Granger 1969), are applicable to GC. Strictly speaking, what GC establishes is the fact that one event happens before another, which may or may not provide an evidence of a cause-effect relationship between them (Hu et al. 2011). GC is

typically a bivariate procedure in 2-dimensional systems. In. In the presence of a third variable that commonly causes changes in the two variables with a different time-lag, the model may falsely recognize the relationship between the two variables as a GC. Limitations of the approach also include the inapplicability of GC in the presence of non-linear, contemporaneous causal links (Russo 2009). A number of approaches have been developed to test for the strength of GC [see e.g. [see e.g. (Dolado and Lütkepohl 1996, Zapata and Rambaldi 1997, Clarke and Mirza 2006)].

### 3.4.2 Convergent Cross Mapping in dynamic nonlinear systems

A number of methods extended the Granger's concept to non-linear cases [see e.g. Ancona et al. (2004) for nonlinear bivariate time series; Baek and Brock (1992), Hiemstra and Jones (1994) for non-parametric GC]. We refer to Hlavácková-Schindler et al. (2007) for a thorough analysis of these methods and to Hu et al. (2011) for further methodological expansions that address limitations of GC

GC may also give ambiguous results in deterministic settings, especially for dynamic systems with weak to moderate coupling (Sugihara et al. 2012). The assumption of separability (see section 4.1) is not satisfied in such systems. In this case, as noted by Sugihara et al. (2012), if $X$ is a cause for $Y$, information about $X$ will be redundantly present in $Y$ and, as a consequence of Takens' theorem (Takens 1981), removing $X$ from the universe of all possible causative variables would not remove the information carried by $X$.

Work from Sugihara and co-authors (Deyle and Sugihara 2011, Sugihara et al. 2012, Clark et al. 2015) addresses the limitations of GC for the case of (1) non-separable systems, (2) weakly coupled variables, and (3) a presence of interactions among variables from external driving variables (e.g. ecological variables such as species from temperature, precipitation, and upwelling). The applications demonstrated in Sugihara et al. (2012) are particularly interesting to understand causality in dynamic systems that are common in ecology (e.g. predator-prey systems).

For the case of dynamic systems, time-series variables can be considered causally linked if they belong to the same dynamic system (Takens 1981, Deyle and Sugihara 2011, Sugihara et al. 2012). Under such consideration, each variable can identify the state of the other, thus information about the past of one variable can be recovered from the time-series of the other, and vice versa (Sugihara et al. 2012).

Let us consider again two time-series of length L:

$$\{X_t\} = (X_t, X_{t-1}, X_{t-2}, ..., X_L)$$
$$\{Y_t\} = (Y_t, Y_{t-1}, Y_{t-2}, \ldots, Y_L). \quad (25)$$

$X$ and $Y$ are said to be causally linked if they share the same common attractor manifold $M$, thus if they are part of the same dynamic system (Sugihara et al. 2012). The manifold can be defined as the system of coordinates constructed from lagged coordinates of the time-series variables, of $X$ and $Y$ using the history of the variables as defined in Eq.(25). Following Takens' theorem (Takens 1981), we can reconstruct the value of the manifold $M$ from a single observation variable of the system. Thus, we can generate a system of coordinates in the attractor manifold built from $X$, $M_X$, and a system of coordinates built from $Y$, $M_Y$. The convergent cross mapping [CCM; see (Sugihara et al. 2012, Clark et al. 2015, Ye et al. 2015) for a thorough analysis] approach tests for causality by measuring the extent to which the historical record of $Y$ values can reliably estimate states of $X$. Using CCM, the analyst can test at a time $t$ whether points located closely (i.e. with similar coordinates in the manifold) in $M_Y$ can be used to identify closely-located points in $M_X$(Sugihara et al. 2012). Using a nearest-neighbor algorithm (Sugihara and Mayf 1990) CCM allows attributing weights to points in the manifold $M_X$ and estimating the value of $Y$ using these weights, thus estimating the quantity $Y|M_X$. Finally, the causal effect can be estimated by measuring the correlation between $Y$ and $Y|M_X$ (McCracken and Weigel 2014). If high correlation is measured, the analyst may confidently use $Y$ to estimate $X$ or vice versa. Time-series

variables that are mutually coupled, in fact, allow for cross-mapping estimations in both directions, thus each variable can be estimated from the other.

CCM can also accommodate for the case of time-series variables that do not interact with each other, but which are both driven by a common variable. Information about the common cause, e.g. $Z$, can still be recovered from $X$ and $Y$. Such application is of particular interest for the study of non-dynamic and chaotic variables. Increasingly data availability estimates improve in accuracy, thus implying convergence in CCM and declining estimation error when cross-mapped estimates are calculated (Sugihara et al. 2012).

#### 3.4.3 Causality as information flow through transfer entropy

As described in the previous sections, GC measures causal influence statistically based on prediction by means, e.g. of a VAR in stochastic systems with separability, and CCM deals with dynamical systems where causal variables have synergistic effects. More recently, transfer entropy [TE; (Schreiber 2000a, Runge et al. 2012, Gencaga et al. 2015, Gómez-Herrero et al. 2015)] has gained traction, finding application in a wide range of fields [see e.g. (Katura et al. 2006, Wibral et al. 2013, Lehnertz and Dickten 2015)]. TE is an information-theoretic measure of the time-directed information flows between jointly dependent processes (Barnett and Seth 2014). Of particular interest is the application of TE to quantify the information flow or information transfer within complex and dynamic systems (Runge 2015). The application of TE is especially suitable for systems for which only time-series of measurements are available and for which the underlying mechanisms that would be needed to directly infer causal relations are poorly understood.

Still bearing in mind the idea of Wiener earlier reported, one would expect the relationship between two variables $X$ and $Y$ to be asymmetric and that the information flows in a direction from the source $Y$ to the target $X$ (Razak and Jensen 2014). TE is a measure of such directed information transfer between joint processes. TE is asymmetric, so that $TE(X \rightarrow Y) \neq TE(Y \rightarrow X)$. The difference indicates a

699 direction of information flow, which can be considered as a measure of potential causation from $X$ to $Y$
700 (Boba et al. 2015). In contrast with GC, TE is not framed in terms of prediction but in terms of resolution
701 of uncertainty (Barnett et al. 2009). Therefore, $TE(Y \to X)$ is the degree to which $Y$ provides
702 information to disambiguate the future of $X$ (i.e. to reduce the level of uncertainty on the future of $X$)
703 beyond the degree to which $X$ already disambiguates its own future (Barnett et al. 2009).

704 At the basis of TE is the notorious concept of differential entropy for a continuous random variable
705 introduced by Shannon (Shannon 2001). Shannon's theorem is a very general way of characterizing the
706 statistical dependency or shared information between two variables. Following the theorem, differential
707 entropy is defined as:

708 (26) $$H(X_i = x) = -\int p(x) \log_2 p(x)\, dx.$$

709 The discrete version of Eq. (26) can be defined for a discrete random variable $X$ with a domain of
710 definition $D_X$, having possible values $x$ with probability $p = p(x)$ (Boba et al. 2015). The value of the
711 entropy $H(X)$ measures the average amount of information gained from a measurement that specifies
712 one particular value $x$:

713 (27) $$H(X_i = x) = -\sum_{x \in D_X} p(x) \log_2 \left( p(x) \right).$$

714 The entropy of $X$ can be seen as a measure of the uncertainty of $X$ (Hlaváčková-Schindler et al. 2007).
715 We use $log_2$ and $log$ interchangeably hereafter.

716 Let us consider again two random variables, $X$ and $Y$, whose probability distributions are $p_x$ and $p_y$,
717 and whose outcomes are from a set $D := D_X \cap D_Y$. The interdependence between $X$ and $Y$ can be
718 calculated using the Kullback-Leibler divergence (Kullback and Leibler 1951):

719 (28) $$D_{KL}\left(p_X|p_Y\right)=\sum_{x\in D}p_X\left(x\right)log\frac{p_X\left(x\right)}{p_Y\left(x\right)}.$$

720 The concept of entropy rate $H_r$ of a process was developed by Schreiber (Schreiber 2000b) to include a
721 direction of the information flow and a chronological ordering. As in Boba et al. (2015), we may assume
722 that time intervals τ are equidistant and use counters like $n := [t / \tau]$ to enumerate the time points. $H_r$ can
723 be, then, defined as:

724 (29) $$H_r = -\sum_{x_{n+1},x_n^m} p\left(x_{x+1},x_n^m\right)log\left[p\left(x_{x+1}|x_n^m\right)\right].$$

725 In the formulation, $x_n^m$ is an m-tuple of measurements at time steps $n, n-1, \ldots, n-m+1$ . Let us employ
726 the same history length $m$ for both data sets $x_n$ and $y_n$ , and, thus, define TE as (Schreiber 2000b):

727 (30) $$TE\left(Y\rightarrow X\right)=\sum p\left(x_{x+1},x_n^m,y_n^m\right)log\frac{p\left(x_{x+1}|x_n^m,y_n^m\right)}{p\left(x_{x+1}|x_n^m\right)}.$$

728 According to Lizier and Prokopenko (2010), TE can be viewed as 'a conditional mutual information, […]
729 the average information in the source about the next state of the destination that was not already contained
730 in the destination's past states". TE is a directional, dynamic measure of information transfer, but still
731 remains a measure of observed (conditional) correlation rather than direct effect (Lizier and Prokopenko
732 2010). TE can be considered as a nonlinear generalization of GC for the case of Gaussian variables (i.e.
733 normally-distributed), and a formal link between GC and TE can be found in Barnett et al. (2009).
734 Conversely, one can define GC as an approximation to TE (Schreiber 2000a), which is exact for Gaussian
735 variables. The equivalence between GC and TE means that the former can be interpreted in terms of
736 information-theoretic bits-per-unit-time (Barnett et al. 2009).

### 3.4.4 Applications

#### 3.4.4.1 Granger-causality

The study of the GC between energy consumption, income and climate in countries of the globe is a well-studied topic (Wagner et al. 2016). The application of the method allows for using a time-series approach to study the dynamic link between $CO_2$ emissions and income. The approach allows for including in the analysis the temporal component of the relationship between the emissions and economic growth, unlike the case previously introduced on causal regression-based techniques. Methodological issues and inconsistencies with the application of GC in this context are well-documented (Bruns et al. 2014).

Omri et al. (2014) test the causality between $CO_2$ emissions, economic growth, and foreign direct investment (FDI) using panel data for a global panel of 54 countries over the period 1992-2011. The authors prove a bidirectional GC between FDI and $CO_2$ for all countries, except for those in Europe and North Asia. The study also finds a unidirectional GC between $CO_2$ and economic growth, except for the analyzed countries in the Middle East, North Africa, and sub-Saharian Africa, for which a bidirectional causality cannot be rejected. Pao and Tsai (2011a) analyze the period between 1992 and 2007 for the causality between $CO_2$ emissions, energy consumption, FDI, and GDP in Brazil, Russian Federation, India and China. The authors suggest that emissions are highly responsive to change in both energy consumption and GDP, but not in FDI. Furthermore, the authors find a unidirectional causality from energy consumption to emissions and a bi-directional GC between emissions and FDI.

Soytas et al. (2007) use a time-series of data for the period 1960-2004 and show that income does not Granger-cause carbon emissions in the U.S. in the long run, but energy use does, thus revealing that a decreased energy use can reduce carbon emissions. Chiou-Wei et al. (2008) use linear and nonlinear GC to test the relationship between energy consumption and economic growth for a sample of Asian newly industrialized countries and the U.S.. The study supports the neutrality hypothesis for the U.S., Thailand and South Korea. A unidirectional causality from economic growth to energy consumption is found between the Philippines and Singapore, while energy consumption may have affected growth in Taiwan,

762 Hong Kong, Malaysia and Indonesia. Menyah and Wolde-Rufael (2010) examine the causal relationship
763 among $CO_2$ emissions, renewable and nuclear energy consumption, and real GDP for the U.S. during the
764 period 1960–2007. The test results indicate no causality from renewable energy to $CO_2$ emissions and a
765 unidirectional causality from nuclear energy consumption to $CO_2$ emissions.

766 For the case of Central America, Apergis and Payne (2009) show that energy consumption and economic
767 growth Granger-cause $CO_2$ emissions, and that there is a bivariate causality between energy consumption
768 and $CO_2$ emissions. Long et al. (2015) investigate the relationships between energy consumption, carbon
769 emissions and economic growth in China from 1952 to 2012. The study finds a bivariate causality
770 between GDP and $CO_2$ emissions and between GDP and coal, gas, and electricity consumption, thus
771 suggesting that a change in the energy consumption structure of China would be required to meet climate
772 goals. Acaravci and Ozturk (2010) show that there is evidence of a causal relationship between energy
773 consumption, income and carbon emissions in Denmark, Germany, Greece, Iceland, Italy, Portugal and
774 Switzerland. Halicioglu (2009) indicates that, for the case of the country of Turkey, GC runs in both
775 directions between $CO_2$ emissions and income both in the short term and long term. The authors, thus,
776 state that it is possible to forecast the future levels of these variables from the past levels of each other.
777 Ang (2008) confirms the existence of a causality running from economic growth to energy consumption
778 growth, both in the short-run and the long-run, for the country of Malaysia. GC is verified in this context
779 also for the case of ASEAN-5 countries [Malaysia, Indonesia, Singapore, the Philippines and Thailand
780 (Chandran and Tang 2013)], Tunisia (Fodha and Zaghdoud 2010), India (Ghosh 2009), and Brazil (Pao
781 and Tsai 2011b).

782 A number of studies used GC to test the relationship between urbanization and carbon emissions. Hossain
783 (2011) finds no evidence of causality in the short and long run for nine emerging industrialized countries
784 for the period 1971-2007. Al-mulali et al. (2012) study 19 countries and find a positive causal relationship
785 between urbanization and $CO_2$ emissions in 84% of the cases. The same groups confirm these results for
786 the middle eastern and north African countries in the MENA group for the years 1980-2009 (Al-mulali

and Fereidouni 2013). Long-run bidirectional causality is found between urbanization and electricity consumption in a study by Liddle and Lung (2014) on 105 countries for the period 1971-2009. Wang et al. (2016a) find that urbanization Granger-causes carbon emissions in the long- and short-run for the ASEAN countries and for the BRICS (i.e. Brazil, Russia, India, China and South Africa) countries (Wang et al. 2016b) within the period 1985–2014.

GC has found applications also in the field of climatology. GC is here used to test the connection between atmospheric $CO_2$ concentrations and global mean temperature. Young et al. (1991) show a causal relationship between atmospheric $CO_2$ concentrations and global sea surface temperatures. Sun and Wang (1996) suggest that global surface temperature Granger-causes global $CO_2$ emissions, but past temperatures do not significantly improve the predictability of current $CO_2$ concentrations. Tol and De Vos (1998) demonstrate that there is a robust statistical relationship between the records of the global mean surface air temperature and the atmospheric concentration of $CO_2$ over the period 1870–1991. Stern and Kaufmann (2013) find that find that both natural and anthropogenic forcing cause temperature change and also that temperature causes greenhouse gas concentration changes. The causal link between temperature and $CO_2$ concentration is tested in (2013) using ice core data from 800,000 to 6,000 years ago. The authors show strong evidence that $CO_2$ concentration Granger-causes temperature, as well as that temperature Granger-causes carbon dioxide concentration. Triacca (2005) shows the limits of the applications of GC in this context and proves that past observations of $CO_2$ concentration do not significantly improve the predictability of current temperature. Further discussions on the application of GC to these systems are available in the literature (Palus and Vejmelka 2007, Kampen 2010, Smirnov and Mokhov 2015).

#### 3.4.4.2 Convergent cross mapping

Sugihara et al. (2012) apply CCM and find causal evidence of the influence of climate forcing on populations of sardines and anchovies in the California Current. Clark et al. (2015) further expand the concept to include spatial variation in time-series data in order to accommodate for the case of limited

availability of data for a specific location. CCM has been applied also in the field of climate science. van Nes et al. (2015) applied CCM and found a marked positive feedback effect from temperature variability on greenhouse-gas concentrations. Tsonis et al. (2015) studied the association between cosmic rays and global temperature in the 20th-century observational record without finding measurable evidence of a causal effect linking cosmic rays to the overall 20th-century warming trend. The authors, however, do find a significant causal effect of cosmic rays on short term variability in global temperatures. The reproduction of the study by Luo et al. (2015) could not reproduce and confirm the findings.

#### 3.4.4.3 Transfer entropy

TE has found limited application in the sustainability sciences. Kumar and Ruddell (2010) use the technique to identify feedbacks between vegetation and climate components, using observations from a network of monitoring towers across North America. The study finds that the relationship between variability and information production leads to the emergence of ordered organization in the overall system.



## 4 Cause-effect analysis techniques in the context of QSA

The ability for causal models to answer these questions in the context of QSA largely depends on whether (1) the assumptions used by the technique are relevant in the context of QSA, and (2) whether it is viable to construct the model considering data availability. Different causal models may exhibit strengths and weaknesses for particular QSA applications depending on the relevance of the assumptions and the availability of data, requiring the modelers to carefully evaluate the most suitable ones considering the objectives and constraints of the study at hand.

### 4.1 Evaluation of major assumptions of cause-effect analysis techniques for QSA

**Error! Reference source not found.** lists the major assumptions used for cause-effect analysis techniques introduced in section 3 [see also (Russo 2009)].

[Add Table 1 about here]

836 #### 4.1.1 Linearity

837 The assumption of linearity needs to be carefully considered when applying the techniques in the context
838 of QSAs, and when interpreting the results of the analysis. The impossibility of building complex
839 nonlinear models does not fully impede to explain complex systems and provide interpretable causal
840 mechanisms. Linear models, in fact, can also support exploratory science and help identify trends and
841 dynamics in systems that are complex in reality [see also (Hofman et al. 2017) on predictive models
842 accuracy]. QSA typically deals with highly-nonlinear phenomena, thus the assumption of linearity would
843 question the strength of the causal-effect relationship measured assuming a linear relationship. The
844 application of QSA to the SDGs, for example, will likely require the analysis of complex non-linear
845 interconnected systems of policy interventions and effects. Understanding such actions and their
846 consequences under this condition requires incorporating all potential non-linearities into the causal
847 model.

848 In the context of SDGs, SDG number 14, which states "Conserve and sustainably use the oceans, seas and
849 marine resources for sustainable development", includes the specific target of "reducing the loss of
850 marine species" (target 14.2 in the SDGs). The loss of marine species is determined by numerous factors
851 and by inter-species relationships that can only be partially grasped using linear relationships. When time-
852 series data is available, GC assumes linearity, while CCM is shown to be applicable to dynamic, non-
853 linear systems, thus avoiding the necessity to linearize relationships that are not linear in the real world.
854 The applicability of this technique in the context of QSA needs to be further tested. Therefore, ES, linear
855 OS, and GC have limited applicability for the QSAs of a system that involves significant nonlinear
856 processes.

857 The assumption of separability and additivity regards the application of OS for causal inference. Under
858 this assumption, the changes in the whole system under study can be understood as a sum of the changes
859 by single constituents of the system. This assumption may not always be verified in the context of QSA,
860 since individual constituents of a system that is being studied under a QSA may present specific

properties that would result in a different behavior of the whole system if combined with other constituents. In the context of sustainability policy, for instance, government may decide to support a cleaner technology by means of tax rebates, which in turn would typically have a beneficial effect on the penetration of the technology in the market. Said rebates may be successful in promoting the diffusion of this more environmentally desirable technology, while they may totally fail in some cases. This can be due, for instance, to the strong and stable presence of a dominant alternative technology that has already reached maturity, or to the lower cost of another alternative technology because of the contingent low cost of raw materials, or to the irrational preference of consumers towards another preferred technology. The consideration in the analysis of elements, such as market dynamics and irrationality of decision-making, may bring upon the system unpredictable changes and unexpected behaviors that stem from the interaction between these components over time. These interactions and changes could not be anticipated by simply integrating the single elements into the analysis.

#### 4.1.2 Separability and additivity

The assumptions of separability and additivity also relate to the assumption of no omitted variables and no systematic measurement errors to which all techniques are bound. Due to the complexity and the type interactions considered in QSA that often involve socio-economic transformations, it is likely that the analysis may have one or more unaccounted factors. For example, SDG number 7 states “Ensure access to affordable, reliable, sustainable and modern energy for all”. A goal like this cuts across various socio-economic and technological variables of our society including population, income, energy technologies, natural resource availability, price, social welfare, pollution, public health, legal systems, and their changes and interactions with each other, to name just a few (see section 3.4.1, where energy consumption is connected to the emission of pollutants such as $SO_2$, $NO_x$). The existence and directionality of such causal relationships were measured in most cases by means of GC (see section 3.4.1). When using such studies for the purpose of predicting future emissions of pollutants, however, the existence of other intervening factors such as the existence of emission standards, regulations and their

level of stringency, or availability and affordability of emission mitigation technologies in those countries may come into play, depending on their significance in explaining the relationships between energy consumption and pollutant emissions. Similarly, ES, OS, CCM, and TE should possibly include these factors if used to predict future changes.

#### 4.1.3 Stable-unit-treatment-value

Earlier in this review, we have stressed the importance of extending the causal relationship between variables from a single case or unit in a system to a population of individuals/units, thus to the entire system under assessment. These criteria fall within SUTVA. The assumption is necessary to all techniques considered, but is especially relevant for the causal interpretation of regressions. The assumption of stability ensures that the causal factor is supposed to operate homogenously in the population of interest. The size of the sample directly relates to the data availability for QSA (see section 4.2). A sufficient sample size is, thus, fundamental to generalize the findings for any of the techniques that we presented. State-level policies on sustainability and their outcomes are limited by the fact that there are only about couple of hundred states in the world and that the circumstances under which those states operate, in terms of population, climate, endowment of wealth and resources, etc. are very different. This problem is exacerbated when considering the fact that much of the sustainability-related policies have been experimented mostly in high-income countries, leaving few empirical evidences associated with low-income countries. This problem can be mitigated by the virtue of 'natural experiments'(Hsiang et al. 2011, Gans et al. 2013, Hanna and Oliva 2015), while identifying a natural experiment that is directly applicable to the question at hand is often challenging.

The behavior of the entire population cannot be understood by the behavior of individual constituents in the presence of interactions among them (see section 3.2.1), which may display emergent properties at the system's level (Odum and Barrett 1971, Bar-Yam 1997). A failure to recognize that a certain system has properties that are emergent can lead to a fallacious identification of a cause-effect relationship. In this

case, it is recommended to explore other approaches such as quantitative systems modelling, socio-technical transition analysis, and initiative-based learning as elaborated in Turnheim et al. (2015).

#### 4.1.4 Parameters and error term

In regression-based causal models, the error term is uncorrelated with the regressors, allowing the analyst to interpret how much of the causal impact is escaping her control due to the factor $\varepsilon$. This assumption applies to ES and OS techniques. This assumption for QSA means that all unknowns may be referred to the error term, without the risk of missing, e.g. a third screening factor $Z$ when analyzing variable $X$ and the related outcome $Y$ . For the case of techniques that use time-series data, the core assumption is that a variable $X$ includes unique useful data to predict the state of a latent variable $Y$ , so that variable $X$ is causally linked to $Y$ . The assumption entails that no other variable should be relevant to describe such cause-effect chain, thus the potential bias from unobservable variables can only be identified if all potential other influential variables are considered.

Techniques that involve the use of time-series, namely GC, CCM and TE, can be used for causality assessments under the assumption that the cause contains unique information about the effect under study that could not be possibly obtained studying the time-series data of another potential cause. Such assumption would mean that it could be possible to univocally relate effects to a single cause. Such assumption may prove difficult to test in the general context of application of QSA, in which before identifying a single cause-effect relationship a number of time-series should be simultaneously tested to check for the potential information transfer and causal links between them. For this reason, the applicability of these cause-effect analysis techniques should be tested in practical case studies, e.g. related to the SDG. Practical applications would show the extent to which the assumption of causal inclusion could be always demonstrated.

### 4.2 Applicability

Applicability of a technique should give a due consideration to the relevance of the assumptions that they require. Even if the assumptions used in the presented cause-effect analysis techniques are relevant for the particular problems with which a QSA is concerned, a technique may show limited applicability due to other limitations, such as data availability. As seen in section 3.2.1, Rubin's model, for example, is an ideal approach to cause-effect analysis. The model shows that causality can be inferred from a well-designed randomized experiment (Holland 1986a). Even though such experimental design would be preferable, the applicability of Rubin's model in the context of QSA is limited due to the data requirement that any large-scale sustainability problem require to consider (see Table 2).

[Add Table 2 about here]

The data required to conduct and experiment in the fashion suggested by Rubin's approach is often prohibitive, due to the difficulty of easily assigning individuals to treatment and control groups, and also due to the difficulty of verifying alternative versions of the world that the counterfactual nature of Rubin's model assumes. Unlike typical treatments considered in empirical studies, in the context of sustainability policy, we have earlier noted that policy actions tend to evolve after they are launched, thus limiting the possibility of guarantying that the sample data under study is stable throughout the experimental design.

Also, in Rubin's model only the factors that are treated in an experiment can be considered as legitimate potential causes (Bhrolcháin and Dyson 2007). This implies that in all those contexts in which attributes need to be taken into account the potential outcome model cannot be applied. This interpretation would rule out personal attributes such as sex, or voluntary actions, such as deciding to start recycling waste, as potential causes. In other words, there would be "no causation without manipulation" (Holland 1986b p. 959), so only factors that could, in principle, be manipulated should be considered as potential causes. By ruling out factors that are not treated in an experiment from the pool of potential causes, the claim may confuse the factors that *can be* causes with those that *can be shown to be* causes experimentally (Bhrolcháin and Dyson 2007). Many phenomena that are thought to be potential causes in QSAs

sciences, in fact, cannot be manipulated. And even if an intervention of the analyst would be possible, it would still be difficult, if not impossible, to find a counterfactual that could be selected as a control. Such a specification would rule out experimentation from QSA, as well as from the social and natural sciences. For this reason, the concept of causality is not dependent upon conditions of manipulability (Russo et al. 2011), and the capacity of an analyst to manipulate a cause is irrelevant to whether a factor is or could be a cause in an experimental setting. Using a truly experimental design in QSA context comes with challenges as we have highlighted. The absence of an experimental design like Rubin's does not render the identification of causality impossible. We explored a set of techniques that can be used to assess causality and sustain causal claims also in cases in which Rubin's model is unsuitable.

#### 4.2.1 Observational studies and QSA

Methods based on observational data, such as matching techniques and causal regressions, under the analyzed set of assumptions can provide an alternative means of assessing causality in the context of QSA. We presented a number of examples that show that a well-defined causal regression allows for cause-effect analysis even under the absence of an experimental design. The "no causation without manipulation" criteria, again, would seem to rule out causal inference from observational studies (Angrist and Pischke 2008). Over the years, many have questioned this standpoint. Glymour (1986), for instance, finds the statement an unnecessary restriction, and objections can be found in the demography (Bhrolcháin and Dyson 2007) and economic literature (Angrist and Pischke 2008). Pearl (2009) argues that the essential ingredient for causation is the capacity of some variable to respond to variation in other variables, regardless of the manipulation restriction imposed on the analyst. There certainly is, according to Pearl, causation without manipulation, such as the gravitational force of the moon that causes tides [see also Goldthorpe, (2001); (Pearl 2009)]. In the systems typically under assessment in QSA, manipulation does not rule out, then, the use of observational studies. Moreover, observational data is increasingly available and global efforts to collect and classify data are undergoing. Applications of OS should be further tested.

Time-series techniques benefit from a vast availability of time-series data also in the QSA domain. GC has found wide application to study the GDP-energy-pollution nexus (Wagner et al. 2016). When interpreting the results of GC in the context of QSA, one should remember that the technique poorly performs in the presence of non-linear, contemporaneous causal systems, and its results are ambiguous when dynamic systems with weak to moderate coupling are under assessment (Sugihara et al. 2012). This element does pose some factual limitations in the use of GC for QSA, also given the number of interconnected systems that are typically under scrutiny.

The applicability of CCM should be further tested for QSA, since it would allow for the consideration of dynamic, strongly coupled systems also under conditions of limited data availability. In the context of sustainable energy of interest for SDG number 7, for instance, such approach could be beneficial to test the relationship between renewable energy deployment and climate-change mitigation using the relatively limited data available for renewable energy sources, which have been monitored only in the last two decades. Similarly, TE could support such assessments, since it is easily applicable in the context of complex dynamic systems; however, the limited application in real cases does not allow for total appreciation of the full potential of TE for QSA at the current stage of knowledge.

In order for cause-effect analysis techniques to be applicable for QSA, causal explanations should not be derived with statistical methods alone (Goldthorpe 2001). The role of the analyst in the form of background knowledge and theories, otherwise considered a subject-matter input, is key to design a causality study for QSA. To this end, all methods come with limitations and in all cases the *caveat* of the presence of strong assumptions is valid, as we have pointed out for each method in the previous section. In some cases, little evidence is available to the analyst when the claim is made that a causal model accurately estimates causal effects, and sometimes causal models can be wrong. We hold here the definition that of Berk et al. (2013) coined for SEM, stating that a causal model is “a quantitative theory of how the data were generated, in which a statistical formalization for random variables is combined with a causal account derived from subject-matter knowledge”. For QSA, the statistical properties of the

observed data model and the assumptions need to be carefully scrutinized before inferring causal relationships in order to avoid spurious results (Hlaváčková-Schindler et al. 2007). Global tests of model fit (e.g. the likelihood ratio chi-square test) should be combined with local tests [e.g. partial correlations and tetrad tests (Bollen and Ting 2000)], which allow for a more accurate model diagnosis and fine-tuning (Bollen and Pearl 2013). Also, more replications of causal models should be then combined as needed in different settings and among different populations in order to strengthen the causal claims (Brand and Thomas 2013).

#### 4.2.2 Dynamic and temporal inconsistencies

The causal models we described typically deal with historical data, and try to map causal mechanisms based on relationships between inputs and outcomes that have typically matured over a stretch of time in the past. However, the preferences of decision-makers measured at one point in time are not fixed, and they are not necessarily representative of future preferences. The so-called dynamic or temporal inconsistency (Thaler 1991, Loewenstein and Prelec 1992) refers to the apparent disagreement between decisions taken or preferences expressed at a certain time, and that change at a later stage. This incongruence of behavior determines a situation in which plans that are made (or policy that are developed) at one point in time may not attain the same level of outcome later (Pearce et al. 2003). This is true for the case of individuals, and even more for the case of societies that usual satisfy weaker rationality and congruence conditions than individuals (Pearce et al. 2003). In the context of QSA, the issue of dynamic inconsistencies is especially relevant, as the temporal dimension involved tends to be a longer term future, over which causal relationships may change. The existence of dynamic inconsistency indicates that policy recommendations based on a QSA that uses a causal relationship need to be re-evaluated over time to fully account for the changes in preference over time.

## 5 Outlook: causality for quantitative sustainability assessment

The ability to identify causality between policy instruments and sustainability performance indicators, including those listed under the SDGs, provides decision makers with a powerful tool. We surveyed a

number of cause-effect analysis techniques and their applications in the field of sustainability (in section 3, and also see Appendix A for a synthesis). Our review demonstrates that in the field of QSAs causality could provide a means to understand phenomena and their causes, as well as to investigate the effective policy measure to achieve desired outcomes. The use of cause-effect analysis techniques has important benefits also in analyzing the effect of a policy option (e.g. the level of air pollution abatement due to the introduction of road space rationing in a city). Such techniques can be used by decision makers to prioritize policy options, and understand the expected outcomes and potential unintended consequences.

We used SDGs to contextualize causal research and its applicability in QSA. In this context, we showed the challenges of understanding how various processes in the systems under study by QSA are connected and how specific policy or management actions can help address SDGs. Our review also discussed a number of caveats when applying cause-effect analysis techniques to QSA. The application of cause-effect analysis techniques is still a challenge in the field of sustainable development. Existing are confined to limited idealistic and isolated cases that are still far from the breadth of analysis that is required by the SDGs. Existing frameworks, as they stand, therefore, may offer limited insights for the questions as complex as SDGs. Overcoming such barriers would require a cross-disciplinary effort to develop an operational framework for quantifying the sustainability consequences of policy actions addressing the needs of QSA, discussed in this review. In the meantime, it is inevitable that QSA practitioners will have to make use of the imperfect information on causality from e.g., mechanistic models or from the use of the cause-effect analysis techniques presented in this review under the assumptions used and the limitations associated with them.

The application of causality for QSA also poses a number of challenges that relate to both the assumptions and the data requirement. Assumptions need to be tested, and the related uncertainties should be analyzed in order for a cause-effect analysis to be valid (Maxim and van der Sluijs 2011). The results should be carefully interpreted in light of the assumptions.

In relation to data availability, we discussed both experimental and observational techniques. Although experimental studies are always the preferred option, they are often impractical, and analysts may be reliant on observational studies alone. Formal experiments are of limited application for many of the most pressing questions of social relevance such as those directly related to climate change, large-scale agricultural intensification or habitat loss (Stephens et al. 2014). When it comes to irreversible, large-scale changes that SDGs and corresponding policy actions are interested in, obviously, experiments are neither practical nor desirable. Furthermore, deriving an unbiased estimate using indirect approaches, such as natural experiments, comes with its own challenges and limitations (Cucurachi and Suh 2015). Adequate non-experimental settings are, then, required to avoid these questions remaining elusive in the absence of experimental evidence. Sufficient survey data and time-series data are available for regression-based causality and time-series cause-effect analysis techniques, respectively, and new methodological developments, such as CCM and TE, allow for dealing with causality under those circumstances in which the analysist faces dearth of data.

Non-experimental cause-effect analysist techniques are powerful but do have limitations and need to accurately be interpreted before causal claims can be safely held to the test. Robust and unambiguous results are difficult to obtain, and many causal claims are yet to be tested with an experimental or quasi-experimental design. Simple correlation at times is taken as a direct measure of the strength of the causal mechanism, or a certain *driver* $X$ is identified as determinant for a certain response $Y$ (e.g. international trade as a driver of biodiversity loss) without testing the relationship using a formal cause-effect analysis technique. The latter does not exclude any complex model from providing valuable information to decision makers, but only suggests caution when causal claims are not supported by experimental models or carefully designed causal models fitted on observational data with all the related assumptions and limitations.

We acknowledge that the framework of hypotheses, assumptions, limitations and the boundaries of the causal analysis matter when claiming the existence of a causal relationship, or when an analysis aims at

verifying the existence of causal mechanisms. Still, new avenues of research can open opportunities for the development of improved cause-effect analysis techniques that build upon the body of knowledge that we have analyzed in this review.

We highlighted the difficulty of framing cause-effect research at the country level, due to the limited number of countries available as a sample, and due to the sheer differences that exist among them. A more promising level of analysis could have cities, instead, at the core. Socio-technical transitions that work at a single city level can be more easily compared with other cities with similar characteristics. To this end, the phenomenon of urban experimentation (Bulkeley and Broto 2013, Rijke et al. 2013, Voytenko et al. 2016) that has been gaining momentum in the last years does have the potential to support change at the global level in the context of SDGs (see also SDG number 9 on resilient cities). New datasets are being compiled based on the local experience of cities, and further efforts should be supported to collect and categorize this information. These data can be used to set up causal quasi-experimental and observational designs to support decision-making for sustainability and the SDGs.

The increased availability of data to which urban experimentation contributes joins the possibility to access and use large data sets. The challenges to causally analyze the so-called Big Data to reveal patterns, trends, and associations, especially related to human behavior and interactions, are enormous (Shiffrin 2016). The task is daunting and requires efforts across multiple disciplines, from the economics to all the sub-disciplines of the sustainability sciences (Athey 2017). Machine learning techniques, among others, have been suggested as a promising joint topic from which such effort could begin (Varian 2016).

QSA sciences, such as industrial ecology, should combine the benefits of an increased availability of data (Xu et al. 2015) with the benefits of understanding causality. Qualitative approaches such as causal loop diagrams (Asif et al. 2015, Laurenti et al. 2015, Efroymson et al. 2016) should be combined with quantitative techniques, and also QSA scientist should reach to other disciplines to contribute to perfect the existing cause-effect analysis techniques. Until new techniques and a new framework for causality in QSA would be available, scientists would still have to resort to alternative techniques (e.g. agent-based

modelling) that, although useful and at times sufficient for single assessments, cannot provide a substantive basis for causal inference at the level required by the SDGs.

In the sub-field of LCA (Hellweg and Mila i Canals 2014), causality is of particular interest in the context of the consequential approach to LCA (Zamagni et al. 2012). Following the consequential approach, the LCA analyst defines changes in a product system and models how the physical and social processes will be triggered by such changes in the system (Earles and Halog 2011). In short, consequential LCA assumes that changes to some parts of the life cycle inventory system lead to a series of consequences through chains of cause–effect relationships (Curran et al. 2005). Concepts borrowed from economic modelling, such as elasticities of supply and demand or general and partial equilibrium models, have been proposed in existing consequential LCA studies or have been recommended for inclusion to simulate responses to changes in the system (Ekvall and Andrae 2006). Practical applications that include these theories in LCA are still limited (Suh and Yang 2014), and only recently alternative solutions have been proposed, for instance using optimization techniques to anticipate changes in a system (Duchin and Levine 2011). We add to the research agenda of the consequential LCA community the necessity to study cause-effect relationships by means of causality theory and methods, thus contributing to increasing the value of the results.

The adoption of causality in all QSAs and the necessity to control causal dynamics and mechanisms requires involvement of experts from a variety of disciplines that can help disentangle those mechanisms and design appropriate causal studies. Collaborations among multiple disciplines will help development of new computational models, methods, and tools to tackle the various aspects of sustainability (Gomes 2009, Marvuglia et al. 2015). Such effort will progressively allow for understanding of the interconnections between SDGs and the results of QSAs beyond the sole use of indicators (Mirshojaeian and Kaneko 2012).

**Acknowledgements**

This article was developed under Assistance Agreement No. 83557901 awarded by the U.S. Environmental Protection Agency to University of California Santa Barbara. It has not been formally reviewed by EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the Agency. The authors thank Professor Olivier Deschenes (UCSB) and Jessica Perkins (UCSB) for helpful comments on an earlier version of this manuscript.

## References


Abadie, A. 2005. Semiparametric difference-in-differences estimators. Rev. Econ. Stud. **72**(1): 1–19.

Abadie, A., and Imbens, G.W. 2006. Large Sample Properties of Matching Estimators. Econometrica **74**(1): 235–267.

Acaravci, A., and Ozturk, I. 2010. On the relationship between energy consumption, CO 2 emissions and economic growth in Europe. Energy **35**(12): 5412–5420.

Aldrich, J. 1995. Correlations genuine and spurious in Pearson and Yule. Stat. Sci.: 364–376.

Allison, P. 1999. Multiple regression: A primer. *Edited by*P.F. Press.

Al-mulali, U., and Fereidouni, H. 2013. Exploring the relationship between urbanization, energy consumption, and CO 2 emission in MENA countries. Renew. Sustain. Energy Rev. **23**: 107–112.

Al-mulali, U., and Sab, C. 2012. The impact of energy consumption and CO 2 emission on the economic and financial development in 19 selected countries. Renew. Sustain. Energy Rev.

Ancona, N., Marinazzo, D., and Stramaglia, S. 2004. Radial basis function approach to nonlinear Granger causality of time series. Phys. Rev. E **70**(5): 56221. APS.

Andam, K., and Ferraro, P. 2008. Measuring the effectiveness of protected area networks in reducing deforestation. Proc. Natl. Acad. Sci. **105**(42): 16089–16094.

Andam, K.S., Ferraro, P.J., Sims, K.R.E., Healy, A., and Holland, M.B. 2010. Protected areas reduced poverty in Costa Rica and Thailand. Proc. Natl. Acad. Sci. U. S. A. **107**(22): 9996–10001. doi:10.1073/pnas.0914177107.

Ang, J. 2008. Economic development, pollutant emissions and energy consumption in Malaysia. J. Policy Model.

Angrist, J., and Krueger, A.B. 2001. Instrumental variables and the search for identification: From supply

and demand to natural experiments.

Angrist, J.D., and Imbens, G.W. 1995. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. J. Am. Stat. Assoc. **90**(430): 431–442. Taylor & Francis.

Angrist, J.D., and Krueger, A.B. 1992. The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples. J. Am. Stat. Assoc. **87**(418): 328–336. Taylor & Francis Group.

Angrist, J.D., and Pischke, J.-S. 2008. Mostly harmless econometrics: An empiricist's companion. Princeton university press.

Antonakis, J., Bendahan, S., Jacquart, P., and Lalive, R. 2014. Causality and endogeneity: Problems and solutions. Oxford Handb. Leadersh. Organ.: 93–117.

Antweiler, W., Copeland, B., and Taylor, M. 1998. Is free trade good for the environment?

Apergis, N., and Payne, J.E. 2009. Energy consumption and economic growth in Central America: Evidence from a panel cointegration and error correction model. Energy Econ. **31**(2): 211–216. doi:10.1016/j.eneco.2008.09.002.

Arnold, B., and Khush, R. 2010. Causal inference methods to study nonrandomized, preexisting development interventions. Proc. Natl. Acad. Sci. **107**(52): 22605–22610.

Arun, T., Imai, K., and Sinha, F. 2006. Does microfinance reduce poverty in India? Propensity score matching based on a national-level household data. *In* Economics Discussion Paper, The University of Manchester.

Asif, F., Rashid, A., Bianchi, C., and Nicolescu, C. 2015. System dynamics models for decision making in product multiple lifecycles. Resour. Conserv. Recycl. **101**: 20–33.

Athey, S. 2017. Beyond prediction: Using big data for policy problems. Science (80-. ). **355**(6324): 483–485.

1180 Athey, S., and Imbens, G. 2006. Identification and inference in nonlinear difference□in□differences
1181 models. Econometrica **74**(2): 431–497.

1182 Baek, E., and Brock, W. 1992. A general test for nonlinear Granger causality: Bivariate model. Iowa State
1183 Univ. Univ. Wisconsin Madison Work. Pap.

1184 Baiocchi, G. 2012. On dimensions of ecological economics. Ecol. Econ. **75**: 1–9.

1185 Barnett, L., Barrett, A.B., and Seth, A.K. 2009. Granger causality and transfer entropy are equivalent for
1186 Gaussian variables. : 1–10. doi:10.1103/PhysRevLett.103.238701.

1187 Barnett, L., and Seth, A.K. 2014. The MVGC multivariate Granger causality toolbox: A new approach to
1188 Granger-causal inference. J. Neurosci. Methods **223**: 50–68. Elsevier B.V.
1189 doi:10.1016/j.jneumeth.2013.10.018.

1190 Barreca, A., Clay, K., Deschenes, O., Greenstone, M., Shapiro, J.S., and Deschenes, O. 2013. Adapting to
1191 Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the
1192 20th Century. NBER Work. Pap. (January): 46. doi:10.3386/w18692.

1193 Barsky, R., and Kilian, L. 2004. Oil and the Macroeconomy since the 1970s.

1194 Bar-Yam, Y. 1997. Dynamics of complex systems. Vol. 213. *In* Addison-We. Reading, MA.

1195 Berk, R., Brown, L., George, E., and Pitkin, E. 2013. What you can learn from wrong causal models. *In*
1196 Handbook of causal analysis for social research. pp. 403–424.

1197 Berk, R.A. 2004. Regression analysis: A constructive critique. Sage.

1198 Bhrolcháin, M.N., and Dyson, T. 2007. On causation in demography: Issues and illustrations. Popul. Dev.
1199 Rev. **33**(1): 1–36. JSTOR.

1200 Blackman, A., and Naranjo, M. 2012. Does eco-certification have environmental benefits? Organic coffee
1201 in Costa Rica. Ecol. Econ. **83**: 58–66.

1202 Boba, P., Bollmann, D., Schoepe, D., Wester, N., Wiesel, J., and Hamacher, K. 2015. Efficient

computation and statistical assessment of transfer entropy. Front. Phys. **3**(March): 1–9. doi:10.3389/fphy.2015.00010.

Bollen, K., and Ting, K. 2000. A tetrad test for causal indicators. Psychol. Methods **5**(1): 3.

Bollen, K.A. 2002. Latent variables in psychology and the social sciences. Annu. Rev. Psychol. **53**(1): 605–634. Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA.

Bollen, K.A. 2012. Instrumental variables in sociology and the social sciences. Annu. Rev. Sociol. **38**: 37–72. Annual Reviews.

Bollen, K.A., and Pearl, J. 2013. Eight myths about causality and structural equation models. *In* Handbook of causal analysis for social research. Springer. pp. 301–328.

Brand, J., and Thomas, J. 2013. Causal effect heterogeneity. *In* Handbook of causal analysis for social research. pp. 189–213.

Brander, M., Tipper, R., Hutchison, C., and Davis, G. 2008. Consequential and attributional approaches to LCA: a Guide to policy makers with specific reference to greenhouse gas LCA of biofuels. Econom. Press (April): 1–14.

Brunner, P., and Rechberger, H. 2004. Practical handbook of material flow analysis. Int. J. Life Cycle **9**(5): 337–338.

Bruns, S., Gross, C., and Stern, D. 2014. Is there really Granger causality between energy use and output? Energy J. **35**(4).

Buhaug, H. 2010. Climate not to blame for African civil wars. Proc. Natl. Acad. Sci. **107**(38): 16477–16482. doi:10.1073/pnas.1005739107.

Bulkeley, H., and Broto, V.C. 2013. Government by experiment? Global cities and the governing of climate change. Trans. Inst.

Burke, M.B., Miguel, E., Satyanath, S., Dykema, J. a, and Lobell, D.B. 2009. Warming increases the risk

of civil war in Africa. Proc. Natl. Acad. Sci. U. S. A. **106**(49): 20670–20674. doi:10.1073/pnas.0907998106.

Caliendo, M., and Kopeinig, S. 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. **22**(1): 31–72. doi:10.1111/j.1467-6419.2007.00527.x.

Campbell, D.T., and Erlebacher, A. 1970. How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. Compens. Educ. A Natl. debate **3**: 185–210. Brunner/Mazel New York.

Cane, M. a., Miguel, E., Burke, M., Hsiang, S.M., Lobell, D.B., Meng, K.C., and Satyanath, S. 2014. Temperature and Violence. Nat. Clim. Chang. **4**(April): 234–235. doi:10.1038/nclimate2171.

Chandran, V., and Tang, C. 2013. The impacts of transport energy consumption, foreign direct investment and income on CO 2 emissions in ASEAN-5 economies. Renew. Sustain. Energy Rev. **24**: 445–453.

Chaumont, A., Nickmilder, M., and Dumont, X. 2012. Associations between proteins and heavy metals in urine at low environmental exposures: evidence of reverse causality. Toxicol. Lett. **210**(3): 345–352.

Chiou-Wei, S., Chen, C., and Zhu, Z. 2008. Economic growth and energy consumption revisited—evidence from linear and nonlinear Granger causality. Energy Econ. **30**(6): 3063–3076.

Chong, A., and Calderon, C. 2000. Causality and feedback between institutional measures and economic growth. Econ. Polit. **12**(1): 69–81.

Christakis, N.A., and Iwashyna, T.J. 2003. The health impact of health care on families: A matched cohort study of hospice use by decedents and mortality outcomes in surviving, widowed spouses. Soc. Sci. Med. **57**: 465–475. doi:10.1016/S0277-9536(02)00370-2.

Christman, M.C. 2008. Statistical modeling of observational data with spatial dependencies. J. Wildl. Manage. **72**(1): 23–33. doi:10.2193/2007-295.

Clark, A.T., Ye, H., Isbell, F., Deyle, E.R., Cowles, J., Tilman, G.D., and Sugihara, G. 2015. Spatial

convergent cross mapping to detect causal relationships from short time series. Ecology **96**(5): 1174–1181.

Clarke, J., and Mirza, S. 2006. A comparison of some common methods for detecting Granger noncausality. J. Stat. Comput. **76**(3): 207–231.

Cochran, W. 2009. Planning and analysis of observational studies, Vol.232. John Wiley & Sons.

Cochran, W.G., Rubin, D.B., Sankhyā, S., Indian, T., Series, A., and Cochean, B.W.G. 1973. Controlling Bias in Observational Studies : A Review All use subject to. SANKHYA INDIAN J. Stat. Ser. A **35**(4): 417–446.

Cole, M., and Elliott, R. 2003. Determining the trade–environment composition effect: the role of capital, labor and environmental regulations. J. Environ. Econ. … **46**(3): 363–383.

Conti, G., Heckman, J., and Urzua, S. 2010. The education-health gradient. Am. Econ. Rev. **100**(2): 234. NIH Public Access.

Copeland, B., and Taylor, M. 2013. Trade and the environment: Theory and evidence. Princeton University Press.

Costanza, R., Groot, R. de, and Sutton, P. 2014. Changes in the global value of ecosystem services. Glob. Environ. **26**: 152–158.

Costello, C., Gaines, S., and Lynham, J. 2008. Can catch shares prevent fisheries collapse? Science (80-. ). **321**(5896): 1678–1681.

Costello, C., Lynham, J., Lester, S., and Gaines, S. 2010. Economic incentives and global fisheries sustainability. Resource **2**.

Creutzig, F., Popp, A., Plevin, R., Luderer, G., Minx, J., and Edenhofer, O. 2012. Reconciling top-down and bottom-up modelling on future bioenergy deployment. Nat. Clim. Chang. **2**(5): 320–327. Nature Publishing Group. doi:10.1038/nclimate1416.

1272 Cucurachi, S., and Suh, S. 2015. A Moonshot for Sustainability Assessment. Environ. Sci. Technol.:
1273 150731104030000. doi:10.1021/acs.est.5b02960.

1274 Curran, M.A., Mann, M., and Norris, G. 2005. The international workshop on electricity data for life
1275 cycle inventories. J. Clean. Prod. **13**(8): 853–862. Elsevier.

1276 Cutler, D.M., and Lleras-Muney, A. 2010. Understanding differences in health behaviors by education. J.
1277 Health Econ. **29**(1): 1–28. Elsevier.

1278 Cutler, D.M., Lleras-Muney, A., and Vogl, T. 2008. Socioeconomic status and health: dimensions and
1279 mechanisms.

1280 Dawid, a P. 1979. Conditional Independence in Statistical Theory. J. R. Stat. Soc. Ser. B **41**(1): 1–31.

1281 Dawid, A. 2007. Counterfactuals, hypotheticals and potential responses: a philosophical examination of
1282 statistical causality.

1283 Deschenes, O. 2010. Climate Policy and Labor Markets. Bus. Econ. **45**(3): 152–157.

1284 Deschenes, O., and Greenstone, M. 2007. The Economic Impacts of Climate Change: Evidence from
1285 Agricultural Output and Random Fluctuations in Weather. Am. Econ. Rev. **102**(7): 3761–3773.
1286 doi:10.1257/aer.102.7.3761.

1287 Deschenes, O., Greenstone, M., and Guryan, J. 2009. Climate Change and Birth Weight. Am. Econ. Rev.
1288 Pap. Proc. **99**(2): 211–217. doi:10.2307/25592401.

1289 Deschenes, O., and Kolstad, C. 2011. Economic impacts of climate change on California agriculture.
1290 Clim. Change **109**: 365–386. doi:10.1007/s10584-011-0322-3.

1291 Deschenes, O., and Moretti, E. 2009. Extreme weather events, mortality, and migration. Rev. Lit. Arts
1292 Am. **91**(November): 659–681. doi:10.1162/rest.91.4.659.

1293 Deyle, E.R., and Sugihara, G. 2011. Generalized theorems for nonlinear state space reconstruction. PLoS
1294 One **6**(3). doi:10.1371/journal.pone.0018295.

1295 Diamond, A., and Sekhon, J.S. 2012. Genetic Matching for Estimating Causal Effects: A General
1296 Multivariate Matching Method for Achieving Balance in Observational Studies. Rev. Econ. Stat.
1297 **95**(July): 121010082253002. doi:10.1162/REST_a_00318.

1298 Dolado, J., and Lütkepohl, H. 1996. Making Wald tests work for cointegrated VAR systems. Econom.
1299 Rev. **15**(4): 369–386.

1300 Donald, S., and Lang, K. 2007. Inference with difference-in-differences and other panel data. Rev. Econ.
1301 Stat. **89**(2): 221–233.

1302 Duchin, F., and Levine, S. 2011. Sectors may use multiple technologies simultaneously: The rectangular
1303 choice-of-technology model with binding factor constraints. Econ. Syst. Res.

1304 Earles, J.M., and Halog, A. 2011. Consequential life cycle assessment: a review. Int. J. Life Cycle Assess.
1305 **16**(5): 445–453. doi:10.1007/s11367-011-0275-9.

1306 Efroymson, R., Kline, K., Angelsen, A., and Verburg, P. 2016. A causal analysis framework for land-use
1307 change and the potential role of bioenergy policy. Land use policy **59**: 516–527.

1308 Ekvall, T., and Andrae, A. 2006. Attributional and consequential environmental assessment of the shift to
1309 lead-free solders (10 pp). Int. J. Life Cycle Assess. **11**(5): 344–353.

1310 Esposito Vinzi, V., Chin, W.W., Henseler, J., and Wang, H. 2010. Handbook of partial least squares:
1311 Concepts, methods and applications. Heidelberg, Dordrecht, London, New York: Springer.

1312 Ferraro, P.J., and Hanauer, M.M. 2014. Quantifying causal mechanisms to determine how protected areas
1313 affect poverty through changes in ecosystem services and infrastructure. Proc. Natl. Acad. Sci. U. S.
1314 A. **111**(11): 4332–7. doi:10.1073/pnas.1307712111.

1315 Firebaugh, G. 2008. Seven Rules for Social Research. ERIC.

1316 Fischer-Kowalski, M., and Swilling, M. 2011. Decoupling: natural resource use and environmental
1317 impacts from economic growth.

1318 Fodha, M., and Zaghdoud, O. 2010. Economic growth and pollutant emissions in Tunisia: an empirical
1319 analysis of the environmental Kuznets curve. Energy Policy.

1320 Fowler, J.H., Christakis, N.A., and others. 2008. Dynamic spread of happiness in a large social network:
1321 longitudinal analysis over 20 years in the Framingham Heart Study. Bmj **337**: a2338. BMJ
1322 Publishing Group Ltd.

1323 Frankel, J. a., and Rose, A.K. 2005. Is Trade Good or Bad for the Environment? Sorting Out the
1324 Causality. Rev. Econ. Stat. **87**(1): 85–91. doi:10.1162/0034653053327577.

1325 Freedman, D.A. 2009. Statistical models: theory and practice. Cambridge University Press.

1326 Gans, W., Alberini, A., and Longo, A. 2013. Smart meter devices and the effect of feedback on residential
1327 electricity consumption: Evidence from a natural experiment in Northern Ireland. Energy Econ.

1328 Gelman, A., and Hill, J. 2006. Data analysis using regression and multilevel/hierarchical models.
1329 Cambridge University Press.

1330 Gencaga, D., Knuth, K., and Rossow, W. 2015. A Recipe for the Estimation of Information Flow in a
1331 Dynamical System. Entropy **17**(1): 438–470. Multidisciplinary Digital Publishing Institute.
1332 doi:10.3390/e17010438.

1333 Geweke, J. 1982. Measurement of Linear Dependence and Feedback between Multiple Time Series. J.
1334 Am. Stat. Assoc. **77**(378): 304–313. doi:10.1080/01621459.1982.10477803.

1335 Geweke, J.F. 1984. Measures of conditional linear dependence and feedback between time series. J. Am.
1336 Stat. Assoc. **79**(388): 907–915. doi:10.2307/2288723.

1337 Ghosh, S. 2009. Electricity supply, employment and real GDP in India: evidence from cointegration and
1338 Granger-causality tests. Energy Policy **37**(8): 2926–2929.

1339 Glymour, C. 1986. Comment: Statistics and metaphysics. J. Am. Stat. Assoc. **81**(396): 964–966. Taylor &
1340 Francis.

1341 Goldthorpe, J.H. 2001. Causation, statistics, and sociology. Eur. Sociol. Rev. **17**(1): 1–20. Oxford Univ
1342 Press.

1343 Gomes, C. 2009. Why Health Information Technology Doesn't Work. Bridg. **39**(4): 5.

1344 Gómez-Herrero, G., Wu, W., Rutanen, K., Soriano, M., Pipa, G., and Vicente, R. 2015. Assessing
1345 Coupling Dynamics from an Ensemble of Time Series. Entropy **17**(4): 1958–1970. Multidisciplinary
1346 Digital Publishing Institute. doi:10.3390/e17041958.

1347 Good, I.J. 1963. A Causal Calculus. Br. Soc. Philos. Sci. **197**(4869): 753–753. doi:10.1038/197753a0.

1348 Granger, C.W.J. 1969. Investigating causal relations by econometric models and cross-spectral methods.
1349 Econom. J. Econom. Soc.: 424–438. JSTOR.

1350 Groot, R. De, Fisher, B., Christie, M., and Aronson, J. 2010. Integrating the ecological and economic
1351 dimensions in biodiversity and ecosystem service valuation. *In* (TEEB): Ecological and Economic
1352 Foundations. Earthscan.

1353 Grossman, M. 2008. The relationship between health and schooling. East. Econ. J. **34**(3): 281–292.
1354 Nature Publishing Group.

1355 Guinée, J. 2002. Handbook on life cycle assessment operational guide to the ISO standards. Int. J. life
1356 cycle Assess.

1357 Haberl, H., Erb, K., and Krausmann, F. 2007. Quantifying and mapping the human appropriation of net
1358 primary production in earth's terrestrial ecosystems. Proc. **104**(31): 12942–12947.

1359 Hahn, J., Todd, P., and Klaauw, W. Van der. 2001. Identification and estimation of treatment effects with
1360 a regression-discontinuity design. Econometrica **69**(1): 201–209.

1361 Hák, T., Janoušková, S., and Moldan, B. 2016. Sustainable Development Goals: A need for relevant
1362 indicators. Ecol. Indic. **60**: 565–573.

1363 Halicioglu, F. 2009. An econometric study of CO 2 emissions, energy consumption, income and foreign

trade in Turkey. Energy Policy **37**(3): 1156–1164.

Hallstein, E., and Villas-Boas, S. 2013. Can household consumers save the wild fish? Lessons from a sustainable seafood advisory. J. Environ. Econ. ….

Hanna, R., and Oliva, P. 2015. The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. J. Public Econ. **122**: 68–79. doi:10.1016/j.jpubeco.2014.10.004.

Hansen, B.E. 2015. Econometrics. University of Wisconsin.

Heckman, J.J., Humphries, J.E., Veramendi, G., and Urzua, S.S. 2014. Education, health and wages.

Heckman, J.J., and Vytlacil, E.J. 2001. Instrumental variables, selection models, and tight bounds on the average treatment effect. Springer.

Hellweg, S., and Mila i Canals, L. 2014. Emerging approaches, challenges and opportunities in life cycle assessment. Science (80-. ). **344**(6188): 1109–1113. doi:10.1126/science.1248361.

Hertwich, E. 2010. Assessing the environmental impacts of consumption and production: priority products and materials. IRP UNEP.

Hertwich, E.G., Gibon, T., Bouman, E. a., Arvesen, A., Suh, S., Heath, G. a., Bergesen, J.D., Ramirez, A., Vega, M.I., and Shi, L. 2014. Integrated life-cycle assessment of electricity-supply scenarios confirms global environmental benefit of low-carbon technologies. Proc. Natl. Acad. Sci.: 201312753. doi:10.1073/pnas.1312753111.

Hiemstra, C., and Jones, J.D. 1994. Testing for linear and nonlinear Granger causality in the stock price-volume relation. J. Finance **49**(5): 1639–1664. Wiley Online Library.

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., and Bhattacharya, J. 2007. Causality detection based on information-theoretic approaches in time series analysis. Phys. Rep. **441**(1): 1–46. Elsevier.

Ho, D.E., Imai, K., King, G., and Stuart, E.A. 2006. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. Polit. Anal. **15**(3): 199–236.

doi:10.1093/pan/mpl013.

Hoekstra, a. Y., and Mekonnen, M.M. 2012. The water footprint of humanity. Proc. Natl. Acad. Sci. **109**(9): 3232–3237. doi:10.1073/pnas.1109936109.

Hofman, J., Sharma, A., and Watts, D. 2017. Prediction and explanation in social systems. Science (80-. ). **355**(6324): 486–488.

Holland, P. 2003. Causation and race. (January): 1–26.

Holland, P.W. 1986a. Statistics and Causal Inference. J. Am. Stat. Assoc. **81**(396): 945–960. doi:10.2307/2289064.

Holland, P.W. 1986b. Statistics and causal inference. J. Am. Stat. Assoc. **81**(396): 945–960. Taylor & Francis.

Homer-Dixon, T.F. 1991. On the threshold: environmental changes as causes of acute conflict. Int. Secur.: 76–116. JSTOR.

Hossain, M. 2011. Panel estimation for CO 2 emissions, energy consumption, economic growth, trade openness and urbanization of newly industrialized countries. Energy Policy **39**(11): 6991–6999.

Hsiang, S.M., and Meng, K.C. 2014. Reconciling disagreement over climate-conflict results in Africa. Proc. Natl. Acad. Sci. U. S. A. **111**(6): 2100–3. doi:10.1073/pnas.1316006111.

Hsiang, S.M., Meng, K.C., and Cane, M. a. 2011. Civil conflicts are associated with the global climate. Nature **476**(7361): 438–441. Nature Publishing Group. doi:10.1038/nature10311.

Hu, S., Dai, G., Worrell, G., and Dai, Q. 2011. Causality analysis of neural connectivity: critical examination of existing methods and advances of new methods. IEEE Trans. neural networks **22**(6): 829–844.

Hughes, T., Carpenter, S., and Rockström, J. 2013. Multiscale regime shifts and planetary boundaries. Trends Ecol.

1410 Hume, D. 1902. Enquiries Concerning the Human Understanding: And Concerning the Principles of
1411 Morals. *In* English ed. Clarendon Press.

1412 Huppes, G., Koning, A., Suh, S., and Heijungs, R. 2006. Environmental Impacts of Consumption in the
1413 European Union: High-Resolution Input-Output Tables with Detailed Environmental Extensions. J.
1414 Ind. **10**(3): 129–146.

1415 Iacus, S.M., King, G., and Porro, G. 2012. Causal inference without balance checking: Coarsened exact
1416 matching. Polit. Anal. **20**(1): 1–24. doi:10.1093/pan/mpr013.

1417 Imbens, G., and Lemieux, T. 2008. Regression discontinuity designs: A guide to practice. J. Econom.
1418 **142**(2): 615–635.

1419 International Council for Science, I.S.S.C. 2015. Review of Targets for the Sustainable Development
1420 Goals: The Science Perspective.

1421 ISO. 2006. ISO 14044: Environmental Management — Life Cycle Assessment — Requirements and
1422 Guidelines. Environ. Manage. **3**: 54.

1423 Kahneman, D., and Krueger, A.B. 2006. Developments in the measurement of subjective well-being. J.
1424 Econ. Perspect. **20**(1): 3–24. American Economic Association.

1425 Kampen, J.K. 2010. A methodological note on the making of causal statements in the debate on
1426 anthropogenic global warming. Theor. Appl. Climatol. **104**(3-4): 423–427. doi:10.1007/s00704-010-
1427 0355-y.

1428 Kang, J., and Larsson, R. 2013. What is the link between temperature and carbon dioxide levels? A
1429 Granger causality analysis based on ice core data. Theor. Appl. Climatol. **116**(3-4): 537–548.
1430 doi:10.1007/s00704-013-0960-7.

1431 Katura, T., Tanaka, N., Obata, A., Sato, H., and Maki, A. 2006. Quantitative evaluation of interrelations
1432 between spontaneous low-frequency oscillations in cerebral hemodynamics and systemic

1433 cardiovascular dynamics. Neuroimage **31**(4): 1592–1600. Elsevier.

1434 Kreuzer, M. 2016. Assessing causal inference problems with Bayesian process tracing: the economic
1435 effects of proportional representation and the problem of endogeneity. New Polit. Econ.: 1–11.

1436 Kullback, S., and Leibler, R.A. 1951. On information and sufficiency. Ann. Math. Stat.: 79–86. JSTOR.

1437 Kumar, P., and Ruddell, B. 2010. Information driven ecohydrologic self-organization. Entropy **12**(10):
1438 2085–2096.

1439 Laplace, P. 1902. A philosophical essay on probabilities. *In* English ed. J. Wiley, New York.

1440 Laurenti, R., Singh, J., Sinha, R., Potting, J., and Frostell, B. 2015. Unintended Environmental
1441 Consequences of Improvement Actions : A Qualitative Analysis of Systems ' Structure and
1442 Behavior. Syst. Res. Behav. Sci. (January 2014). doi:10.1002/sres.2330.

1443 Laurenti, R., and Sinha, R. 2015. Towards Addressing Unintended Environmental Consequences: A
1444 Planning Framework. Sustain. Dev. **24**: 1–17.

1445 Lechner, M. 2001. Identification and estimation of causal effects of multiple treatments under the
1446 conditional independence assumption. Springer.

1447 Lee, D., and Lemieux, T. 2010. Regression discontinuity designs in economics. J. Econ. Lit. **48**(2): 281–
1448 355.

1449 Lehnertz, K., and Dickten, H. 2015. Assessing directionality and strength of coupling through symbolic
1450 analysis: an application to epilepsy patients. Philos. Trans. R. Soc. London A Math. Phys. Eng. Sci.
1451 **373**(2034): 20140094. The Royal Society.

1452 Lenzen, M., Moran, D., Kanemoto, K., Foran, B., Lobefaro, L., and Geschke, a. 2012. International trade
1453 drives biodiversity threats in developing nations. Nature **486**(7401): 109–112.
1454 doi:10.1038/nature11145.

1455 Liddle, B., and Lung, S. 2014. Might electricity consumption cause urbanization instead? Evidence from

heterogeneous panel long-run causality tests. Glob. Environ. Chang.

Linkie, M., Smith, R., and Zhu, Y. 2008. Evaluating biodiversity conservation around a large Sumatran protected area. Conserv. Biol. **22**(3): 683–690.

Lizier, J.T., and Prokopenko, M. 2010. Differentiating information transfer and causal effect. Eur. Phys. J. B **73**(4): 605–615. Springer.

Loewenstein, G., and Prelec, D. 1992. Anomalies in intertemporal choice: Evidence and an interpretation. Q. J. Econ.

Long, X., Naminse, E.Y., Du, J., and Zhuang, J. 2015. Nonrenewable energy, renewable energy, carbon dioxide emissions and economic growth in China from 1952 to 2012. Renew. Sustain. Energy Rev. **52**: 680–688. doi:10.1016/j.rser.2015.07.176.

Luo, M., Kantz, H., Lau, N.-C., Huang, W., and Zhou, Y. 2015. Questionable dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. Proc. Natl. Acad. Sci. U. S. A. **112**(34): E4638–9. doi:10.1073/pnas.1510571112.

Maertens, M. 2011. Globalisation and poverty in Senegal: a worst case scenario? Eur. Rev. Agric. Econ.

Malik, O.A., Hsu, A., Johnson, L.A., and de Sherbinin, A. 2015. A global indicator of wastewater treatment to inform the Sustainable Development Goals (SDGs). Environ. Sci. Policy **48**: 172–185. Elsevier.

Managi, S., Hibiki, A., and Tsurumi, T. 2009. Does trade openness improve environmental quality? J. Environ. Econ. Manage. **58**(3): 346–363.

Mancini, L., Benini, L., and Sala, S. 2015. Resource footprint of Europe: Complementarity of material flow analysis and life cycle assessment for policy support. Environ. Sci. Policy **54**: 367–376. Elsevier.

Marvuglia, A., Benetto, E., and Murgante, B. 2015. Calling for an Integrated Computational Systems

1479 Modelling Framework for Life Cycle Sustainability Analysis. J. Environ. Account. Manag. **3**(3):
1480 213–216.

1481 Matthews, E., Amann, C., and Bringezu, S. 2000. The weight of nations. *In* Material Outflows from.
1482 Washington, DC.

1483 Maxim, L., and van der Sluijs, J.P. 2011. Quality in environmental science for policy: Assessing
1484 uncertainty as a component of policy analysis. Environ. Sci. Policy **14**(4): 482–492. Elsevier.

1485 McCracken, J., and Weigel, R. 2014. Convergent cross-mapping and pairwise asymmetric inference.
1486 Phys. Rev. E **90**(6).

1487 Mei-chu, W. 1987. Tests of causality and exogeneity between exports and economic growth: the case of
1488 Asian NICs. J. Econ. Dev. **12**(2).

1489 Menyah, K., and Wolde-Rufael, Y. 2010. CO 2 emissions, nuclear energy, renewable energy and
1490 economic growth in the US. Energy Policy.

1491 Michalsky, M., and Hooda, P.S. 2015. Greenhouse gas emissions of imported and locally produced fruit
1492 and vegetable commodities: A quantitative assessment. Environ. Sci. Policy **48**: 32–43. Elsevier.

1493 Miguel, E., Satyanath, S., and Sergenti, E. 2004. Economic Shocks and Civil Conflict: An Instrumental
1494 Variables Approach. J. Polit. Econ. **112**(4): 725–753. doi:10.1086/421174.

1495 Mill, J.S. 1856. A System of Logic, Ratiocinative and Inductive, Being a Connected View of the
1496 Principles, and the Methods of Scientific Investigation. JW Parker.

1497 Miller, C., Tsoka, M., and Reichert, K. 2011. The impact of the Social Cash Transfer Scheme on food
1498 security in Malawi. Food Policy.

1499 Mirshojaeian, H., and Kaneko, S. 2012. Causality between pillars of sustainable development: Global
1500 stylized facts or regional phenomena? Ecol. Indic. **14**(1): 197–201. Elsevier Ltd.
1501 doi:10.1016/j.ecolind.2011.07.005.

1502 Morgan, S.L. 2013. Handbook of causal analysis for social research. Springer.

1503 Morgan, S.L., and Harding, D.J. 2006. Matching Estimators of Causal Effects. Sociol. Methods Res. **35**:
1504 3–60. doi:10.1177/0049124106289164.

1505 Morgan, S.L., and Winship, C. 2007. Counterfactuals and Causal Inference. Cambridge university press.

1506 Nelson, A., and Chomitz, K. 2011. Effectiveness of strict vs. multiple use protected areas in reducing
1507 tropical forest fires: a global analysis using matching methods. PLoS One.

1508 van Nes, E.H., Scheffer, M., Brovkin, V., Lenton, T.M., Ye, H., Deyle, E., and Sugihara, G. 2015. Causal
1509 feedbacks in climate change. Nat. Clim. Chang. **5**(5): 445–448. Nature Publishing Group.
1510 doi:10.1038/nclimate2568.

1511 Ness, B., Urbel-Piirsalu, E., Anderberg, S., and Olsson, L. 2007. Categorising tools for sustainability
1512 assessment. Ecol. Econ. **60**(3): 498–508.

1513 Newey, W.K., and Powell, J.L. 2003. Instrumental variable estimation of nonparametric models.
1514 Econometrica: 1565–1578. JSTOR.

1515 Odum, E., and Barrett, G. 1971. Fundamentals of ecology. Vol. 3. Saunders, Philadelphia.

1516 Omri, A., Nguyen, D., and Rault, C. 2014. Causal interactions between CO 2 emissions, FDI, and
1517 economic growth: evidence from dynamic simultaneous-equation models. Econ. Model.

1518 Oreopoulos, P., and Petronijevic, U. 2013. Making college worth it: A review of research on the returns to
1519 higher education.

1520 Oreopoulos, P., and Salvanes, K.G. 2011. Priceless: The nonpecuniary benefits of schooling. J. Econ.
1521 Perspect.: 159–184. JSTOR.

1522 Palus, M., and Vejmelka, M. 2007. Directionality of coupling from bivariate time series: how to avoid
1523 false causalities and missed connections. Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. **75**(5 Pt 2):
1524 056211. American Physical Society. doi:10.1103/PhysRevE.75.056211.

1525 Pao, H., and Tsai, C. 2011a. Multivariate Granger causality between CO 2 emissions, energy
1526 consumption, FDI (foreign direct investment) and GDP (gross domestic product): evidence from a.
1527 Energy.

1528 Pao, H., and Tsai, C. 2011b. Modeling and forecasting the CO 2 emissions, energy consumption, and
1529 economic growth in Brazil. Energy **36**(5): 2450–2458.

1530 Pearce, D., Groom, B., Hepburn, C., and Koundouri, P. 2003. Valuing the future. World Econ. **4**(2): 121–
1531 141.

1532 Pearl, J. 1995. Causal diagrams for empirical research. Biometrika **82**(4): 669–688. Biometrika Trust.

1533 Pearl, J. 2000a. Causal Inference Without Counterfactuals. J. Am. Stat. Assoc. **95**(450): 428–431.

1534 Pearl, J. 2000b. Causality. : 1–386. doi:citeulike-article-id:3888442.

1535 Pearl, J. 2009. Causality. Cambridge University Press.

1536 Pearson, K. 1911. The Grammar of Science, 3rd. London, Engl. A. C. Black **153**.

1537 Peters, G. 2010. Carbon footprints and embodied carbon at multiple scales. Curr. Opin. Environ. Sustain.
1538 **2**(4): 245–250.

1539 Plevin, R., and Delucchi, M. 2014. Using attributional life cycle assessment to estimate climate-change
1540 mitigation benefits misleads policy makers. J. Ind.

1541 Puhani, P. 2012. The treatment effect, the cross difference, and the interaction term in nonlinear
1542 “difference-in-differences” models. Econ. Lett. **115**(1): 85–87.

1543 Raleigh, C., Linke, A., and O’Loughlin, J. 2014. Extreme temperatures and violence. Nat. Clim. Chang.
1544 **4**(2): 76–77. Nature Publishing Group. doi:10.1038/nclimate2101.

1545 Razak, F.A., and Jensen, H.J. 2014. Quantifying “causality”in complex systems: understanding transfer
1546 entropy.

1547 Rijke, J., Farrelly, M., Brown, R., and Zevenbergen, C. 2013. Configuring transformative governance to
1548 enhance resilient urban water systems. Environ. Sci. Policy **25**: 62–72.

1549 Rimer, B. 1998. Correlation is not causation. Am. J. Public Health **88**(5): 832–835.

1550 Rockström, J., Steffen, W., and Noone, K. 2009. Planetary boundaries: exploring the safe operating space
1551 for humanity.

1552 Rosenbaum, P.R. 1984. Testing the conditional independence and monotonicity assumptions of item
1553 response theory. Psychometrika **49**(3): 425–435. Springer.

1554 Ruben, R., Fort, R., and Zúñiga-Arias, G. 2009. Measuring the impact of fair trade on development.

1555 Rubin, D.B. 1973. Matching to remove bias in observational studies. Biometrics **29**(1): 159–183.
1556 doi:10.2307/2529684.

1557 Rubin, D.B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies.
1558 doi:10.1037/h0037350.

1559 Rubin, D.B. 1978. Bayesian inference for causal effects: The role of randomization. Ann. Stat.: 34–58.
1560 JSTOR.

1561 Rubin, D.B. 1980. Comment. J. Am. Stat. Assoc. **75**(371): 591–593. Taylor & Francis.

1562 Rubin, D.B. 1986. Which Ifs Have Causal Answers. Statistics (Ber). **81**(396): 961–962.

1563 Runge, J. 2015. Quantifying information transfer and mediation along causal pathways in complex
1564 systems.

1565 Runge, J., Heitzig, J., Petoukhov, V., and Kurths, J. 2012. Escaping the curse of dimensionality in
1566 estimating multivariate transfer entropy. Phys. Rev. Lett.

1567 Russell, B. 1912. On the notion of cause. *In* Proceedings of the Aristotelian society. pp. 1–26.

1568 Russo, F. 2009. Causality and Causal Modelling in the Social Sciences. *In* Causality and Causal Modlling

in the Social Sciences. doi:10.1007/978-1-4020-8817-9.

Russo, F., and Williamson, J. 2007. Interpreting causality in the health sciences. **21**(2): 157–170. doi:10.1080/02698590701498084.

Russo, F., Wunsch, G., and Mouchart, M. 2011. Inferring Causality through Counterfactuals in Observational Studies: Some Epistemological Issues. Bull. Methodol. socioloique **111**(1978): 43–64. doi:10.1177/0759106311408891.

Sachs, J., and McArthur, J. 2005. The millennium project: a plan for meeting the millennium development goals. Lancet.

Salmon, W. 1980. Probabilistic causality.

Schindler, D., and Hilborn, R. 2015. Prediction, precaution, and policy under global change. Science (80-. ). **347**(6225): 953–954.

Schlenker, W., and Roberts, M.J. 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. Proc. Natl. Acad. Sci. U. S. A. **106**(37): 15594–15598. doi:10.1073/pnas.0906865106.

Schreiber, T. 2000a. Measuring information transfer. Phys. Rev. Lett. **85**(2): 461. APS.

Schreiber, T. 2000b. Measuring Information Transfer. : 0–3. doi:10.1103/PhysRevLett.85.461.

Sekhon, J. 2011. Multivariate and Propensity Score Matching. J. Stat. Softw. **42**(7): 52. doi:10.1.1.335.7044.

Setboonsarng, S., and Parpiev, Z. 2008. Microfinance and the millennium development goals in Pakistan: impact assessment using propensity score matching.

Seth, a. K., Barrett, a. B., and Barnett, L. 2015. Granger Causality Analysis in Neuroscience and Neuroimaging. J. Neurosci. **35**(8): 3293–3297. doi:10.1523/JNEUROSCI.4399-14.2015.

Shaffer, T., and Johnson, D. 2008. Ways of learning: observational studies versus experiments. J. Wildl.

Manage.

Shannon, C.E. 2001. A mathematical theory of communication. ACM SIGMOBILE Mob. Comput. Commun. Rev. **5**(1): 3–55. ACM.

Shiffrin, R.M. 2016. Drawing causal inference from Big Data. Proc. Natl. Acad. Sci. **113**(27): 7308–7309. National Academy of Sciences. doi:10.1073/pnas.1608845113.

Shimizu, S., Hoyer, P.O., Hyvärinen, A., and Kerminen, A. 2006. A linear non-Gaussian acyclic model for causal discovery. J. Mach. Learn. Res. **7**: 2003–2030. JMLR. org.

Simon, H., and Iwasaki, Y. 1988. Causal ordering, comparative statics, and near decomposability. J. Econom. **39**(1): 149–173.

Smirnov, D.A., and Mokhov, I.I. 2015. Relating Granger causality to long-term causal effects. Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. **92**(4): 042138. American Physical Society. doi:10.1103/PhysRevE.92.042138.

Smith, D., Vivekananda, J., and Alert, I. 2007. A climate of conflict. (November): 1–48. doi:ISBN 978-91-586-8071-2.

Smith, G., and Pell, J. 2003. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. BMJ Br. Med. J. **327**(7429): 1459.

Sobel, M.E. 2006. What do randomized studies of housing mobility demonstrate? Causal inference in the face of interference. J. Am. Stat. Assoc. **101**(476): 1398–1407. Taylor & Francis.

Soytas, U., Sari, R., and Ewing, B. 2007. Energy consumption, income, and carbon emissions in the United States. Ecol. Econ. **62**(3): 482–489.

Steffen, W., and Richardson, K. 2015. Planetary boundaries: Guiding human development on a changing planet. Science (80-. ). **347**(6223): 1259855.

Stephens, P. a., Pettorelli, N., Barlow, J., Whittingham, M.J., and Cadotte, M.W. 2014. Management by

proxy? The use of indices in applied ecology. J. Appl. Ecol.: 1–6. doi:10.1111/1365-2664.12383.

Stern, D.I. 2004. The Rise and Fall of the Environmental Kuznets Curve. World Dev. **32**(8): 1419–1439. doi:10.1016/j.worlddev.2004.03.004.

Stern, D.I., and Kaufmann, R.K. 2013. Anthropogenic and natural causes of climate change. Clim. Change **122**(1-2): 257–269. doi:10.1007/s10584-013-1007-x.

Stuart, E., and Rubin, D. 2008. Best practices in quasi-experimental designs. *In* Best practices in quantitative methods. pp. 155–176.

Stuart, E.A. 2010. Matching Methods for Causal Inference: A Review and a Look Forward. Stat. Sci. **25**(1): 1–21. doi:10.1214/09-STS313.

Sugihara, G., May, R., Ye, H., Hsieh, C. -h., Deyle, E., Fogarty, M., and Munch, S. 2012. Detecting Causality in Complex Ecosystems. Science (80-. ). **338**(6106): 496–500. doi:10.1126/science.1227079.

Sugihara, G., and Mayf, R. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature **344**(6268): 734.

Suh, S. 2009. Handbook of input-output economics in industrial ecology. Springer Science & Business Media.

Suh, S., and Yang, Y. 2014. On the uncanny capabilities of consequential LCA. Int. J. Life Cycle Assess. **19**(6): 1179–1184. doi:10.1007/s11367-014-0739-9.

Sun, L., and Wang, M. 1996. Global warming and global dioxide emission: an empirical study. J. Environ. Manage.

Suppes, P. 1970. A probabilistic theory of causality. North-Holland Publishing Company.

Suweis, S., Carr, J. a., Maritan, A., Rinaldo, A., and D'Odorico, P. 2015. Resilience and reactivity of global food security. Proc. Natl. Acad. Sci. **112**(22): 201507366. doi:10.1073/pnas.1507366112.

1638 Suweis, S., Rinaldo, A., Maritan, A., and Odorico, P.D. 2013. Water-controlled wealth of nations. Proc.
1639 Natl. Acad. Sci. **110**(11): 4230–4233. doi:10.1073/pnas.1222452110/-
1640 /DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1222452110.

1641 Takens, F. 1981. Detecting strange attractors in turbulence. Springer.

1642 Tenenhaus, M., Vinzi, V.E., Chatelin, Y.-M., and Lauro, C. 2005. PLS path modeling. Comput. Stat. Data
1643 Anal. **48**(1): 159–205. Elsevier.

1644 Thaler, R. 1991. 'Some Empirical Evidence on Dynamic Inconsistency. Quasi Ration. Econ.

1645 Tol, R., and Vos, A. De. 1998. A Bayesian statistical analysis of the enhanced greenhouse effect. Clim.
1646 Change.

1647 Triacca, U. 2005. Is Granger causality analysis appropriate to investigate the relationship between
1648 atmospheric concentration of carbon dioxide and global surface air temperature? Theor. Appl.
1649 Climatol. **81**(3-4): 133–135. Springer.

1650 Tsonis, A.A., Deyle, E.R., May, R.M., Sugihara, G., Swanson, K., Verbeten, J.D., and Wang, G. 2015.
1651 Dynamical evidence for causality between galactic cosmic rays and interannual variation in global
1652 temperature. Proc. Natl. Acad. Sci. U. S. A. **112**(11): 3253–6. doi:10.1073/pnas.1420291112.

1653 Tukker, A., Huppes, G., Oers, L., and Heijungs, R. 2006. Environmentally extended input-output tables
1654 and models for Europe. Available from https://openaccess.leidenuniv.nl/handle/1887/11433
1655 [accessed 9 June 2016].

1656 Turnheim, B., Berkhout, F., and Geels, F. 2015. Evaluating sustainability transitions pathways: Bridging
1657 analytical approaches to address governance challenges. Glob. Environ. Chang. **35**: 239–253.

1658 United Nations. 2014. Sustainable development goals. United Nations sustainable development
1659 knowledge platform. Available from http://sustainabledevelopment.un.org/?menu=1300 [accessed
1660 17 April 2015].

1661 Varian, H.R. 2016. Causal inference in economics and marketing. Proc. Natl. Acad. Sci. **113**(27): 7310–
1662 7315. doi:10.1073/pnas.1510479113.

1663 Voytenko, Y., McCormick, K., and Evans, J. 2016. Urban living labs for sustainability and low carbon
1664 cities in Europe: Towards a research agenda. J. Clean. Produ **123**: 45–54.

1665 Wagner, L., Ross, I., Foster, J., and Hankamer, B. 2016. Trading Off Global Fuel Supply, CO 2
1666 Emissions and Sustainable Development. PLoS One.

1667 Wang, Y., Chen, L., and Kubota, J. 2016a. The relationship between urbanization, energy use and carbon
1668 emissions: evidence from a panel of Association of Southeast Asian Nations (ASEAN) countries. J.
1669 Clean. Prod. **112**: 1368–1374.

1670 Wang, Y., Li, L., Kubota, J., Han, R., Zhu, X., and Lu, G. 2016b. Does urbanization lead to more carbon
1671 emission? Evidence from a panel of BRICS countries. Appl. Energy **168**: 375–380.
1672 doi:10.1016/j.apenergy.2016.01.105.

1673 Weinzettel, J., Hertwich, E.G., Peters, G.P., Steen-Olsen, K., and Galli, A. 2013. Affluence drives the
1674 global displacement of land use. Glob. Environ. Chang. **23**(2): 433–438. Elsevier Ltd.
1675 doi:10.1016/j.gloenvcha.2012.12.010.

1676 Whiteman, G., Walker, B., and Perego, P. 2013. Planetary boundaries: Ecological foundations for
1677 corporate sustainability. J. Manage.

1678 Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., Lindner, M., Lizier, J.T., and
1679 Vicente, R. 2013. Measuring information-transfer delays. PLoS One **8**(2): e55809. Public Library of
1680 Science.

1681 Wiedmann, T., and Minx, J. 2007. A Definition of ' Carbon Footprint. Science (80-. ). **1**(01): 1–11.

1682 Wiener, N. 1956. The theory of prediction. Mod. Math. Eng. **1**: 125–139. McGraw-Hill New York.

1683 Wold, H. 1982. Soft modelling: the basic design and some extensions. Syst. under Indirect Obs. Part II:

36–37. North Holland.

Wright, S. 1921. Correlation and causation. J. Agric. Res. **20**(7): 557–585. Washington.

Xu, M., Cai, H., and Liang, S. 2015. Big Data and Industrial Ecology. J. Ind. Ecol. **19**(2): 205–210. doi:10.1111/jiec.12241.

Yang, S., Imbens, G., and Cui, Z. 2015. Propensity Score Matching and Subclassification in Observational Studies with Multi-level Treatments. arXiv Prepr.

Ye, H., Deyle, E.R., Gilarranz, L.J., and Sugihara, G. 2015. Distinguishing time-delayed causal interactions using convergent cross mapping. Sci. Rep. **5**: 14750. Nature Publishing Group. doi:10.1038/srep14750.

Young, P., Ng, C., Lane, K., and Parker, D. 1991. Recursive forecasting, smoothing and seasonal adjustment of non-stationary environmental data. J. Forecast. **10**(1-2): 57–89.

Zamagni, A., Guinée, J., Heijungs, R., Masoni, P., and Raggi, A. 2012. Lights and shadows in consequential LCA. Int. J. Life Cycle Assess. **17**(7): 904–918. doi:10.1007/s11367-012-0423-x.

Zapata, H., and Rambaldi, A. 1997. Monte Carlo evidence on cointegration and causation. Oxford Bull. Econ. **59**(2): 585–298.

Zivot, E., and Wang, J. 2006. Vector autoregressive models for multivariate time series. Model. Financ. Time Ser. with S-PLUS: 385–429. Springer.

1702

**Table 1. Set of assumptions of cause-effect analysis techniques**

| Assumption | Definition | Cause-effect analysis technique | | | | |
|---|---|---|---|---|---|---|
| | | ES* | OS | GC | CCM | TE |
| Linearity | Causal relationships are linear in the parameters and in the explanatory variable $X$ and the response variable $Y$ . | x | (x)† | x | | |
| Separability and additivity | The response variable $Y$ is the sum of explanatory variables $X$ and the errors $\varepsilon$. | | x | | | |
| No omitted variables and no systematic measurement errors | The analyst assumes that the factors considered in the analysis sufficiently represent the causal relationship at stake, and the measurement errors in the variables are random (i.e., not systematic) and are within a tolerable range. Under the hypothetical condition that expected $\varepsilon$is equal to zero, the causal effect $Y$ is explained by the explanatory variables $Xs$ . | x | x | x | x | x |
| Stable-unit-treatment-value-assumption | The potential outcome of one individual of a population is not affected by the treatment applied to another individual of the population. | x | x | x | x | x |
| No correlation within and between explanatory variables and | Explanatory variables and errors are not correlated with each other. | x | x | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| errors | | | | | | |
| Cause inclusiveness | A cause contains unique information about an effect not available elsewhere. | | | x | x | x |

**[*] ES = Experimental design based on Rubin's model; OS = Observational designs based matching and regression techniques; GC = Granger-causality; CCM = Convergent Cross Mapping; TE = Transfer entropy. [†] In the case of non-linear regression, linearity assumption does not apply to OS**

**Table 1. Synthesis of cause-effect analysis techniques for QSA research**

| Cause-effect analysis technique | Data requirements | Advantages | Disadvantages |
|---|---|---|---|
| ES* | High | • Most trusted approach<br>• Quantifiable strength of the causal relationship | • Cost of implementation<br>• Complexity of systems under analysis<br>• Limited possibility of integrating assumptions in QSA<br>• Limited possibility of replicating a study |
| OS | Medium | • Good alternative to experimental designs<br>• Ease of application<br>• Ease of access to data | • Risk of data manipulation<br>• Limited possibility of randomization for certain variables<br>• Limited possibility of generalizing findings from sample to the full population |
| GC | Medium | • Extensive literature<br>• Ease of application<br>• Ease of access to time-series data | • Inapplicability in the presence of non-linear, contemporaneous causal systems<br>• Ambiguous results for dynamic systems with weak to moderate coupling |

| | | | |
|---|---|---|---|
| CCM | Low | • Short time-series data<br>• Applicability to dynamic systems<br>• Applicability to non-linear, contemporaneous causal systems | • Limited application to contexts with multiple causal variables<br>• Untested application in the sustainability realm |
| TE | Medium | • Applicability to dynamic systems<br>• Easy to model | • It is only a measure of observed (conditional) correlation rather than direct effect<br>• Untested application in the sustainability realm |

* **The acronyms in the table refer to ES = Experimental design based on Rubin's model; OS = Observational designs based on matching and regression techniques; GC = Granger Causality; CCM = Convergent Cross Mapping; TE = Transfer entropy**