

# **Data Driven Methods for Granger Causality and Contemporaneous Causality with Non-Linear Corrections: Climate Teleconnection Mechanisms**

T. Chu<sup>1</sup>, D. Danks<sup>1,2</sup> and C. Glymour<sup>1,2</sup>

## **Abstract**

We describe a unification of old and recent ideas for formulating graphical models to explain time series data, including Granger causality, semi-automated search procedures for graphical causal models, modeling of contemporaneous influences in times series, and heuristic generalized additive model corrections to linear models. We illustrate the procedures by finding a structure of exogenous variables and mediating variables among time series of remote geospatial indices of ocean surface temperatures and pressures. The analysis agrees with known exogenous drivers of the indices, not assumed in the analysis. Automated search applied to the residuals after regressing each series on its lags and the lags of its Granger causes yields a graphical model of “comtemporaneous” causal relations identical with the qualitative graphical relations from the time series. A similar analysis produces reasonable results when applied to candidate climate indices obtained by a clustering method from sea surface temperature and sea level pressure data.

1. Institute for Human and Machine Cognition, University of West Florida
2. Department of Philosophy, Carnegie Mellon University

Research for this paper was supported by the National Aeronautics and Space Administration Intelligent Data Understanding program and the Office of Naval Research. We thank Drs. Rama Nemani, Alessio Monetta and Peter Spirtes for valuable assistance.

## 1. Introduction.

For stationary times series, Granger (1969) proposed a well-known and since often used definition of “ $X_t$  is a cause of  $Y_t$ , with respect to other series  $Z_t$ .”: the variance of  $Y_t$  conditional on values of all lags of  $X_t$ ,  $Y_t$  and  $Z_t$  is smaller than the variance of  $Y_t$  conditional on all values of all lags of  $Y_t$  and  $Z_t$ . In practice Granger’s criterion has been applied with some bound on the lags. Granger also proposed that  $X$  is an instantaneous cause of  $Y$  with respect to  $Z_t$  if and only if the variance of  $Y_t$  conditional on  $X_t$  and all lags of  $X_t$ ,  $Y_t$  and  $Z_t$  is smaller than the variance of  $Y_t$  conditional on all lags of  $X_t$ ,  $Y_t$ , and  $Z_t$ . For stationary systems with four or more time series components, Granger and Swanson (1998), proposed to determine “contemporaneous causation” in the form of a linear chain by regressing each time series variable on all lags of all variables considered and using the residuals to test for vanishing differences of products of correlations among foursomes of variables, so-called vanishing tetrad differences, e.g.,  $\rho_{ij}\rho_{kl} - \rho_{ik}\rho_{jl} = 0$ . Each distinct linear ordering of dependencies among 4 or more Gaussian variables implies a unique set of vanishing tetrad differences provided no correlations vanish (Glymour, et al., 1997). Using search procedures for directed acyclic graphical linear models (Spirtes, et al., 2000), Bessler (20??) and Hoover (2003) generalized Granger and Swanson’s procedure to allow search for contemporaneous linear systems among all partial ordering of the dependencies among the variables. Monetta (2003) derived the correction needed for the fact that the correlations are obtained as residuals from a regression. We will show that similar methods can be used to generate the Granger causal structure of a time series.

Various objections to all of these procedures include: (1) associations and conditional associations among the variables may be due to unrecorded variables, and so may not be causal in the specified sense ( Kaufman, et al, 2003); (2) relationships may be non-linear; (3) independencies and conditional independencies may arise because of canceling, or nearly canceling, influences rather than from the absence of any causal connection;

except when both a time ordering and the absence of unrecorded confounding variables is assumed, in the absence of much stronger assumptions relating correlations and causal connections (Zhang, 2003,) no confidence intervals are possible for the results of search over causal relations (Robins, et al., 2003). Using an example of climate teleconnections, this paper unifies these several methods while addressing aspects of issues (1) and (2).

Given an n-bound on the lags, using a sequence of conditional independence tests, Granger causal models for Normal variates can be specified semi-automatically as graphical models (Spirtes et al 2000, Bessler 2002) by producing n.lagged series for each variable, and constraining search by time-order. The procedure, or “constraint based,” search, recognizes many cases in which associations among endogenous variables are due to unobserved variables, and uses no prior information except the assumption of Gaussian distributions and linearity, and the time order. Alternative searches by Bayesian methods are possible and also converge, probability 1, under the assumption that there are no unobserved confounding series and no cancellations of influences. (Meek, Chickering) We then correct the graphical and parametric structure using additive model regression, and apply the Bessler/Hoover strategy to model contemporaneous causal relations with a similar correction procedure..

**2. Climate Teleconnection Data and Model Constraints.** Climate teleconnections are associations of geospatially remote climate phenomena produced by atmospheric and oceanic processes. The most famous, and first established teleconnection, is the association of El Nino/Southern Oscillation (ENSO) with the failure of monsoons in India. A variety of associations have been documented among sea surface temperatures (SST), atmospheric pressure at sea level (SLP), land surface temperatures (LST) and precipitation over land areas. REFERENCES. Since the 1970s data from a sequence of satellites have provided monthly (and now daily) measurements of such variables, at resolutions as small as 1 square kilometer. Measurements in particular spatial regions have been clustered into time indexed indices for the regions, usually by principal components analysis, but also by other methods. Climate research has established that some of these phenomena are exogenous drivers of others, and has sought physical

mechanisms for the teleconnections. REFERENCES We consider here whether constraints on such mechanisms can be obtained by data-driven model selection from time series of ocean indices.

We consider data sets from two sources. The first data set consists of the following 6 ocean climate indices, recorded monthly from 1958 to 1999, each forming a time series of 504 time steps:

- QBO (Quasi Biennial Oscillation): Regular variation of zonal stratospheric winds above the equator
- SOI (Southern Oscillation): Sea Level Pressure (SLP) anomalies between Darwin and Tahiti
- WP (Western Pacific): Low frequency temporal function of the ‘zonal dipole’ SLP spatial pattern over the North Pacific.
- PDO (Pacific Decadal Oscillation): Leading principal component of monthly Sea Surface Temperature (SST) anomalies in the North Pacific Ocean, poleward of 20° N
- AO (Arctic Oscillation): First principal component of SLP poleward of 20° N
- NAO (North Atlantic Oscillation) Normalized SLP differences between Ponta Delgada, Azores and Stykkisholmur, Iceland

Some connections among these variables are reasonably established, but are not assumed in the analysis that follows. In particular, SO and NAO are thought to be exogenous drivers, and NAO is thought to influence WP.(References). The second data set, obtained from different climate indices, will be briefly described later.

The conditions for a plausible model of climate mechanisms are remarkably constrained: The relations, if any, among the climate indices must be at least approximately Gaussian and linear else the search for a Granger causal model will fail—typically producing a spaghetti of associations or nearly complete graph. The indices themselves, which are functions of measurements over a spatial region, must be strongly correlated with

whatever physical variables actually produce the associations, else no conditional independence relations will be found among the indices, tending again to produce a complete graph. Poor indices that include irrelevant spatial regions or that aggregate inappropriately will not produce graphically sparse models. The graphical structure between lags must be reproduced as indices are advanced, that is if  $X_1 \rightarrow Y_2$  is found in the search, the search output should also contain  $X_2 \rightarrow Y_3$ , etc. The autocorrelation of a time series must not be too large, else other variables will be independent of a given variable conditional on the lags of the latter. Finally, if, as in monthly measurements of surface temperature and pressure, the sampling rate is longer than the expected time required for propagation of influence, the graphical structure obtained from the search for simultaneous causal relations should have the same qualitative structure as the graphical relations found for Granger causality.

### **3. A Granger Causal Model from PC Search and Additive Model Correction**

The PC algorithm (Spirtes, et al., 1993; 2000) is a search procedure for equivalence classes of graphs—with equivalence defined by equality of conditional independence relations implied by the Markov Assumption (each variable in a directed acyclic graph representation is independent of its non-descendants in the graph conditional on values of its parent in the graph) constraining relations between graphical topology and joint distributions on the random variables that are vertices of the graph (Lauritzen 1996). Assuming (1) the Markov property; (2) the converse of the Markov property, sometimes called Faithfulness (e.g., all conditional independence relations in the joint distribution follow from the Markov Assumption for the true causal graph); (3) causal sufficiency—all sources of covariation among recorded variables are recorded in the data; and (4) i.i.d. sampling, the algorithm is pointwise consistent. (Spirtes, et al., 200) The output of the algorithm is a mixed graph with directed edges and undirected edges, representing an equivalence class of directed acyclic graphs (Pearl 2000). The algorithm will nonetheless correctly identify (in the large sample limit) by doubly directed edges pairs of recorded variables associated by unmeasured covariates, but it is not complete for that property. Doubly directed edge, e.g.,  $X \rightarrow Y \leftrightarrow Z \leftarrow W$  will arise when  $X$ ,  $Z$  and

respectively,  $Y, W$  are not adjacent and  $X, Z$  are dependent conditional on  $Y$  and  $Y, W$  are dependent conditional on  $Z$ .

The implementation of PC algorithm we use ( the TETRAD IV program: <http://www.phil.cmu.edu/projects/tetrad>) assumes either a joint Gaussian or a joint multinomial distribution (we use the former), and allows the user to specify constraints of time order (directed edges will not be produced from later to earlier variables), required edges and forbidden edges. Our analysis pre-specifies only the time order among the lagged and unlagged series.

The assumption of a linear model significantly simplifies the task of testing for conditional independence, but this assumption is often, sometimes badly, violated in reality. Some (conditional) dependencies might be missed if we only use tests for zero (partial) correlation. For example, If  $Y = X^2 + \epsilon$ , where  $X, \epsilon \sim N(0, 1)$ , then  $\text{Corr}(X, Y) = 0$ . In order to recover the possibly missing (conditional) dependencies, we need a more general test method. Let the joint density  $\{X, Y, Z\}$  be  $f_{XYZ}(x,y,z)$ , the joint density of  $\{X,Z\}$  be  $f_{XZ}(x,z)$ , the joint density of  $\{Y,Z\}$  be  $f_{YZ}(y,z)$ , and the marginal density of  $Z$  be  $f_Z(z)$ . We could test if  $X$  and  $Y$  are independent given  $Z$  using the Hellinger distance between  $f_{XYZ}(x,y,z)f_Z(z)$  and  $f_{XZ}(x,z)f_{YZ}(y,z)$ . However, this approach requires the estimation of joint density, and hence is subject to dimensionality problems. Instead, in this study, we choose a heuristic approach based on an additive model, a semi-parametric regression method.

Recall that we could test the Granger causality by testing for zero regression coefficients. That is, estimate the following model with ordinary least regression for  $i = 0, \dots, p$  and  $j = 1, \dots, p$ ,

$$(2) Y_t = c + \sum_i \alpha_i X_{t-i} + \sum_j \beta_j Y_{t-j} + u_t$$

Then  $X$  does not Granger-cause  $Y$  iff  $\alpha_i = 0$  for  $i = 0, \dots, p$ . Assuming a linear model, this test is essentially a conditional independence test, because a predictor variable with 0

coefficient is independent of the response variable conditional on the other predictors. This equivalence, unfortunately, cannot be generalized to arbitrary nonlinear models: In general, if the  $X$  and  $Y$  are conditional independent given  $\{Z_1, \dots, Z_k\}$ , then the conditional expectation of  $X$  given  $Y$  and  $\{Z_1, \dots, Z_k\}$  is the same as the conditional expectation of  $X$  given  $\{Z_1, \dots, Z_k\}$ ; the converse, however, does not hold. For example, if  $Z = X^2 + Y^2 + \varepsilon_Z$ , where  $X$ ,  $Y$ , and  $\varepsilon_Z$  are independent, then  $X$  and  $Y$  should be dependent given  $Z$ . But regressing  $X$  on  $Y$  and  $Z$  using a nonparametric regression method, such as a surface smoother, we would find that neither  $Y$  nor  $Z$  is a significant term in predicting  $X$ . (Of course, the residuals of  $X$  would depend on  $Y$  and  $Z$ .) Given such a limitation, and the fact that a general regression method that assumes little about the functional relations among the variables in a causal system, such as surface smoother, is also subject to dimensionality problems, it seems to make little sense to use regression as a test for conditional independence for data not generated from a linear model.

The main advantage of the regression approach for independence testing, however, is that a general regression method can be approximated by efficient methods without dimensionality problems. One of such methods is the additive model. With the assumption that the response variable is the sum of smooth univariate functions of the predictors plus an independent error term, additive models transform the problem of estimating an  $n$ -dimensional surface to a problem of estimating  $n-1$ -dimensional curves. Although it is only an approximation of the general regression method, the method still could capture, in many cases, the nonlinear relation between the response variable and the predictors. Accordingly, we supplement a fully automated linear model search with additive model regression to search for the Markov blanket of each variable. (For each represented variable  $X$  in a set  $\mathbf{V}$ , its Markov blanket with respect to  $\mathbf{V}$  is the smallest subset  $M$  of  $\mathbf{V}$  such that  $X$  is independent of  $\mathbf{V} \setminus M$  conditional on  $M$ .) We do not perform an entirely new search for causal structure based solely on additive model regressions because in a causal system generated from additive models, the additive model regression approach could miss some conditional independencies. For example, consider a causal system with two exogenous variable  $X_1$  and  $X_2$ . Let  $W = X_1 + X_2 + \varepsilon_W$ ,  $Y = W^2 + \varepsilon_Y$ ,  $U = \log(X_1) + \varepsilon_U$ ,  $V = \log(X_2) + \varepsilon_V$ , and  $Z = U + V + \varepsilon_Z$ . Although  $Y$  and  $Z$

are independent conditional on  $X_1$  and  $X_2$ , regressing  $Y$  against  $X_1$ ,  $X_2$ , and  $Z$ , we would find that  $Z$  is a significant predictor for  $Y$ .

Instead, we use additive regression to correct the causal patterns generated by PC algorithm with linear model assumption. Each of the variables will be regressed against all the other variables and their lags. The set of variables that are significant predictors for the response variable will be considered as a subset of the Markov Blanket for the response variable in the causal graph. The general additive correction is carried out according to the following two heuristics:

- 1) Stricter test level: In order to reject the null hypothesis (that a variable is not a significant predictor for the response variable), the test statistic must have a smaller  $p$  value than that used in the linear model search.
- 2) Conflict resolution: If a set of tests  $\{T_1, \dots, T_k\}$  imply that  $X$  and  $Y$  are dependent given a set of conditioning variables  $\phi$ , and the conditional independence test  $T_0$  fails to reject the null that  $X$  and  $Y$  are independent given  $\phi$ ,  $T_0$  will be ignored as long as the dependency implied by the other tests is not extremely strong ????

The first heuristic is designed to allow for the fact that additive model regression might miss some conditional independencies, while the second heuristic allows for the possibility that the (WHICH LINEAR OR AM??) regression approach could miss some conditional dependencies.

Once a final model is obtained, we fit a piecewise linear model.

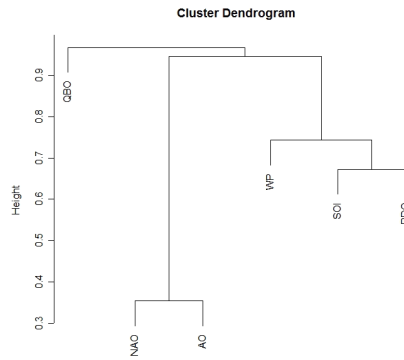
### 3.1 Clustering analysis

To contrast causal with non-causal methods, we perform a hierarchical clustering of the 6 original variables using a distance between  $x$  and  $y$  defined as:

$$1 - |corr(x, y)|$$

and using the average pairwise distance between points in two subclusters as the distance between the two subclusters, and get the following graph:

Cluster dendrogram of the 6 ocean indices



**Figure 1**

The clustering suggests that the variables could be put into three groups according to their association, but reveals nothing about which variables are exogenous and which endogenous variables are intermediate between exogenous variables and other endogenous variables.

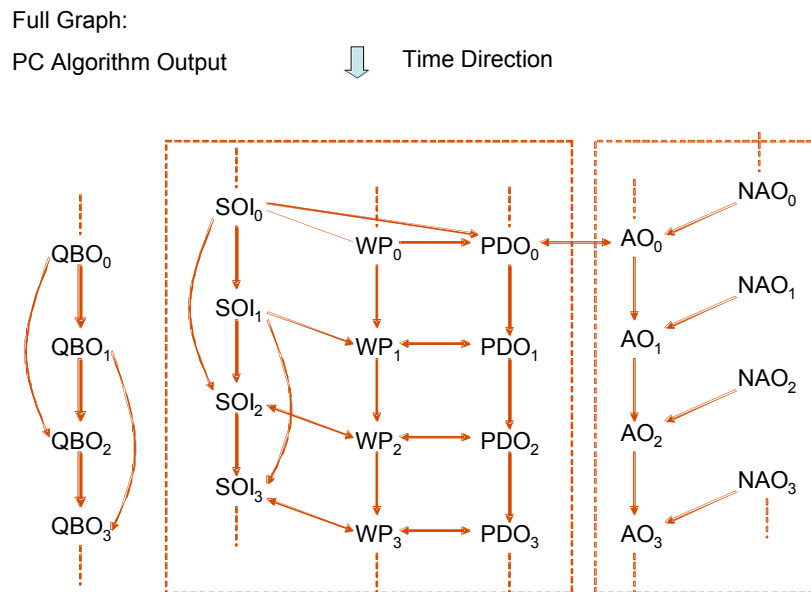
### 3.2 PC search

Three lagged series are produced from each time series:

- $QBO_0 = \{q_1, q_2, \dots, q_{501}\}$
- $QBO_1 = \{q_2, q_3, \dots, q_{502}\}$
- $QBO_2 = \{q_3, q_4, \dots, q_{503}\}$
- $QBO_3 = \{q_4, q_5, \dots, q_{504}\}$
- $SOI_0 = \{s_1, s_2, \dots, s_{501}\}$

- ... ..
- $NAO_3 = \{n_4, n_5, \dots, n_{504}\}$

With a multivariate normal assumption, we test the conditional independencies using the  $t$  test for zero partial correlation with significance level  $\alpha = 0.01$ . The result is shown in figure 2.



**Figure 2**

### 3.3 Additive model corrections

Some of the indices and their lags, such as  $NAO_2$  and  $WP_3$ , clearly are not linearly related (figure 3). Each of the  $QBO_3$ ,  $SOI_3$  and  $NAO_3$ , time series were therefore regressed against all other variables, assuming an additive model. The regression results then are interpreted as conditional independence tests.

In the general additive regressions, we choose an  $\alpha$  level of 0.001. The result removed only one edge (from  $SOI$  to  $PDO$ ), reoriented all the doubly directed edges (the

conditional dependencies noted above for doubly directed edges did not hold), and added a nonlinear relation between NAO and WP and a slightly nonlinear relation between SOI and QBO. The result is shown in figure 4.

Nonlinear relation between WP<sub>3</sub> and NAO<sub>2</sub>

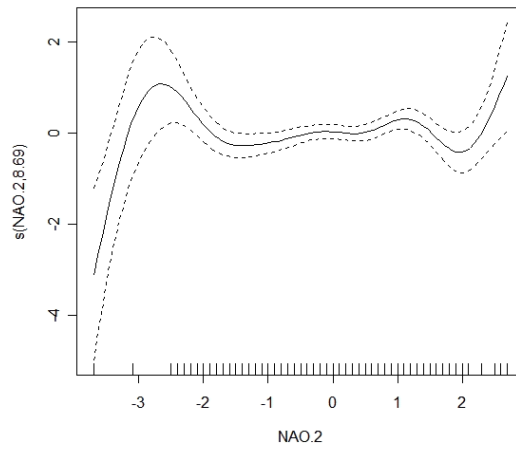
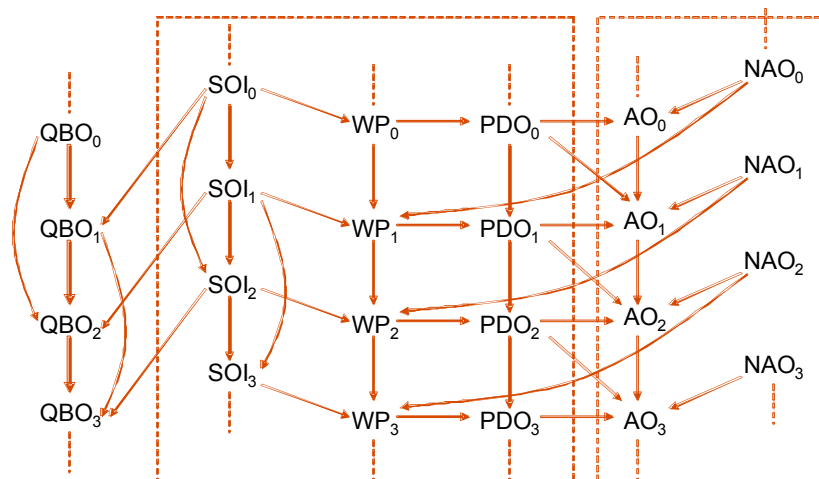


Figure 3

AM corrected pattern  
for 6 ocean indices

Time Direction



**Figure 4**

### 3.4. Simultaneous causation

Each time series variable was regressed on the lags of all the variables, and the residual correlation matrix was given to the PC algorithm. The contemporary structure obtained among the variables is shown in figure 5.



**Figure 5**

Ignoring the orientation of the edges, the resulting structure agrees with the structure shown in the corrected PC graph for the time series, but differs slightly from the uncorrected PC graph. After correcting by additive model regression, we get the structure shown in figure 5.



**Figure 6**

Correcting the statistical tests for the fact that the correlation matrix was produced by regression had no effect on the results.

## 4. Analysis with New Climate Indices

Steinbach et al 2003 propose to identify new climate indices using a Shared Nearest Neighbor (SNN) clustering algorithm. In total 107 clusters were found from the SST data, with another 25 clusters from SLP data. They suggest that each centroid of a SST cluster, or the difference between any pair of centroids of SLP clusters, could be treated as a candidate climate index. So in total they produce 407 candidate climate indices.

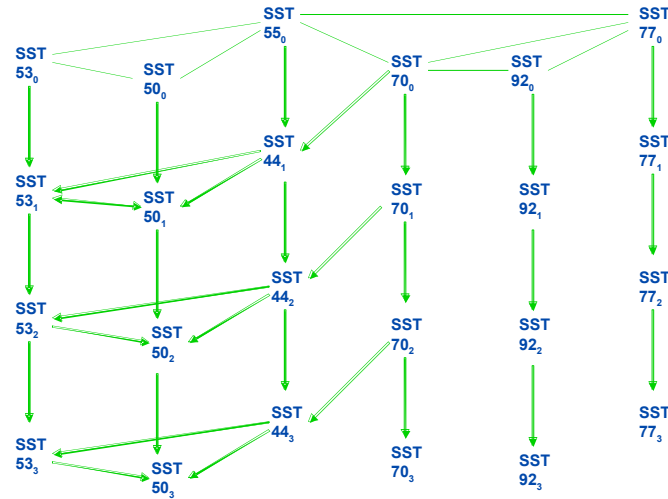
Unlike the established indices, such as SOI, which could be considered as good proxies to the factors of the climate causal system, we have no idea if there is an underlying causal system of which these new candidate indices are proxies. To make the causal inference easier, we pick indices with consistent piece wise correlations: Two columns  $X_1, \dots, X_n$  and  $Y_1, \dots, Y_n$  have consistent correlation if the correlation between  $X_a, \dots, X_{a+k}$  and  $Y_a, \dots, Y_{a+k}$  does not vary significantly for different  $a$  and  $k$ . The thought is that candidate indices with consistent piecewise correlation tend to have simple functional relations, and hence possibly simple causal relations, while moderate correlation among the candidate indices makes the tests for conditional independence more stable.

Values for coefficients are given in the appendix. .COMMENT ON WHY SOME OF THESE ARE NEGATIVE!!!

#### **4.1 SST based candidate indices.**

We first choose 6 SST based indices, named as SST44, SST50, SST53, SST70, SST77, and SST92. Performing the PC search over the indices and their lags with significance level set to 0.01, we get:

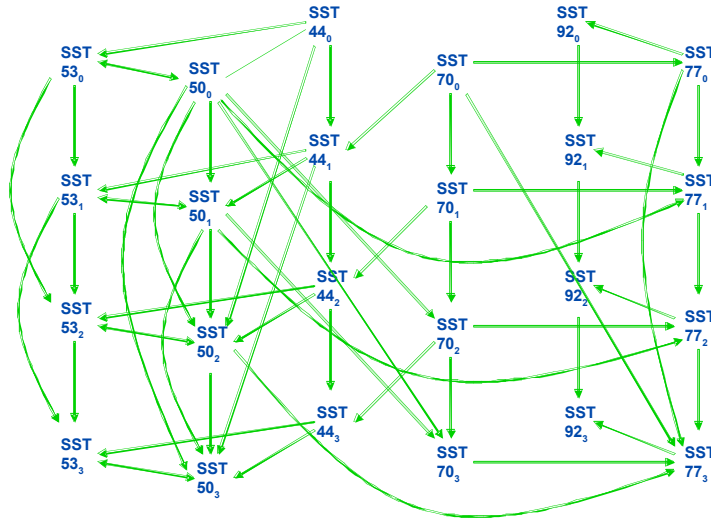
## PC pattern for 6 SST based candidate indices



**Figure 7**

This pattern looks fine except for the relations among SST70, SST92, SST77, where three edges appear at the uppermost row but not present at the low rows. This is caused by the high autocorrelations between these variables and their one month lags, and the relatively low correlation among different variables. Apply the additive model correction, we get a new pattern, which is much more complicated than the PC pattern. This suggests that, despite of our selection, the relations among these candidate indices are highly nonlinear.

### AM corrected pattern among 6 SST based candidate indices



**Figure 8**

Search for the simultaneous causality by regressing the variables against their lags, and then applying the PC algorithm, with significance level set to 0.01, to the residuals, we get the following result. Compared to the AM corrected graph, the edge between SST92 and SST77 is missing. However, if we increase the significance level to 0.011, that edge would be included in the graph.

### Contemporary causal relation among 6 SST based candidate indices



**Figure 9**

#### 4.2 Mixture of SLP and SST based candidate indices.

Another data set we selected including 2 SST based candidate indices and 5 SLP based candidate indices. Using the PC algorithm with significance level 0.01, we get:

PC pattern for the mixture of SLP and SST based candidate indices

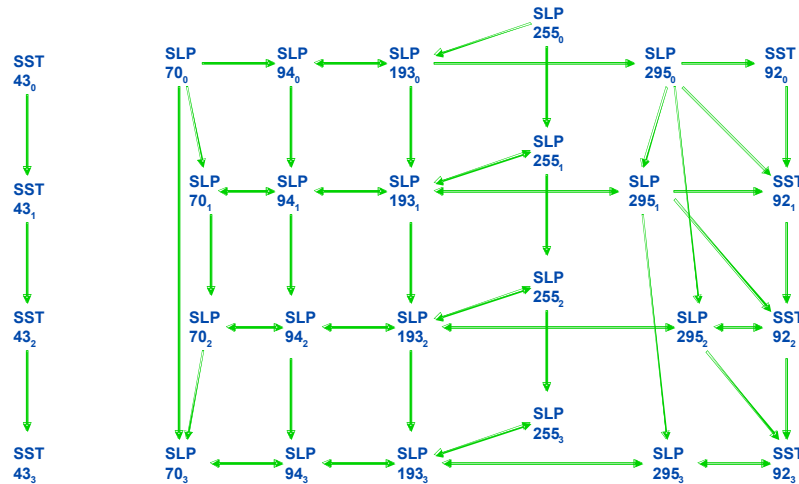
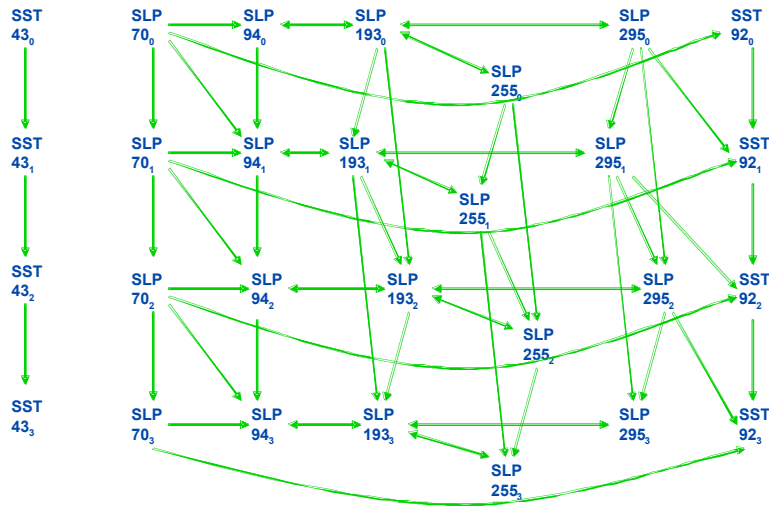


Figure 10

We notice that some causal relations implied by the PC pattern are not plausible: SLP70<sub>0</sub> is a direct cause of SLP70<sub>3</sub>, but not a direct cause of SLP70<sub>2</sub>; SLP295<sub>1</sub> is not a direct cause of SLP295<sub>2</sub>, and SLP295<sub>2</sub> not a direct cause of SLP295<sub>3</sub>. Applying the additive model correction, we get:

## AM corrected pattern for the mixture of SLP and SST based candidate indices

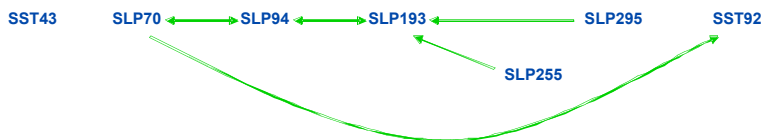


**Figure 11**

The problems about the PC pattern mentioned above are corrected in the new graph. However, we also observe that the edges among SLP94, SLP193, SLP255, and SLP295 cannot be oriented without contradicting some of the conditional independencies implied by the additive model regression results.

Search for the simultaneous causality with significance level for the PC algorithm set to 0.01, we get the following pattern which is a good match to the corrected PC pattern.

## Contemporary causal relation among the mixture of SLP and SST based candidate indices



## Figure 12

### 5. Discussion

This study argues that semi-automated model search for causal structure in time series can be combined with semi-parametric regression method to extract causal relations from time series data. The particular example is to ocean climate indices, but the component procedures have been individually applied to econometric data with some success, suggesting that the criteria for successful application of the joint procedures are statistical and causal rather than domain specific.

The analysis of the 6 established ocean indices derives causal relations part of whose specification (the exogeneity of SOI and NAO, for example) is in agreement with common expert opinion; other elements of the model are novel but not controverted. The model search and specification procedure makes no use of the expert information except in the formation of climate indices.. The fact that the additive model corrected PC patterns match well with the simultaneous relations obtained from the residuals of regression the variables against their lags provides further support for the approach.

The study also shows that the quality of indices is essential. The established ocean indices used here shown a wide range of autocorrelations. In contrast, the 107 candidate climate indices obtained by clustering based on SST data all show high autocorrelation, which could be problematic for automated search algorithms of any kind. As a consequence, the PC pattern for the ocean indices requires relative fewer corrections, while the PC patterns generated from the candidate indices need extensive additive model corrections, and are more likely to leave the directions of some connections undecided.

Our approach is modular, and its two main components, i.e., PC algorithm and additive model correction, could be replaced by other comparable methods. Thus, with appropriate data generated from appropriate mechanisms, related analyses could be

conducted under weaker assumptions. For example, the PC algorithm could be replaced by a currently available pointwise consistent algorithm that is more complete for unrecorded confounders (the FCI algorithm), although the latter algorithm often finds (appropriately) that the end marks of many edges cannot be determined. In the case at hand, no time series variables show a mutual influence. When mutual influences occur, the analysis of simultaneous causation should result in a cyclic graph, and a pointwise consistent algorithm for Normal variates for such cases is available (Richardson and Spirtes 1996) to replace the PC algorithm. Similarly, the additive model regression method could also be replaced by other regression methods that are good approximations to the general regression methods but not subject to the curse of dimensionality.

At the current stage, several key parameters of our analysis are chosen more or less arbitrarily. These include the significance level for the conditional independence test used in the PC algorithm, the significance level for determining if a predictor should be included in the Markov blanket of the response variable, as well the conflict resolution procedure in the additive model correction step. As a direction of further study, we are considering the development of an internal score for the final causal pattern derived from the time series data, and the use of this score to automate the selections of the above mentioned key parameters.

## Appendix

Response Variable	Predictor	Predictor Range	Coefficient
QBO	QBO.2	$-\text{inf} < \text{QBO.2} < \text{inf}$	1.668503
QBO	QBO.1	$-\text{inf} < \text{QBO.1} < \text{inf}$	-0.7513252
QBO	SOI.2	$\text{SOI.2} < 0$	-0.2179054
QBO	SOI.2	$\text{SOI.2} > 0$	0.1443253
SOI	SOI.2	$-\text{inf} < \text{SOI.2} < \text{inf}$	0.5225828
SOI	SOI.1	$\text{SOI.1} < -4.8$	-0.1979299
SOI	SOI.1	$\text{SOI.1} > -4.8$	0.2421996
WP	WP.2	$-\text{inf} < \text{WP.2} < \text{inf}$	0.1777548
WP	SOI.3	$-\text{inf} < \text{SOI.3} < \text{inf}$	-0.1286399
WP	NAO.2	$\text{NAO.2} < -2.5$	1.794230
WP	NAO.2	$-2.5 < \text{NAO.2} < -1.3$	-0.7410115

WP	NAO.2	$-1.3 < \text{NAO.2} < 1.1$	0.1816928
WP	NAO.2	$1.1 < \text{NAO.2} < 2$	-0.6978301
WP	NAO.2	$\text{NAO.2} > 2$	1.609102
PDO	PDO.2	$-\text{inf} < \text{PDO.2} < \text{inf}$	0.8172722
PDO	WP.3	$\text{WP.3} < 0$	0.1927047
PDO	WP.3	$\text{WP.3} > 0$	0.03967585
AO	AO.2	$\text{AO.2} < 0.546707$	0.1095149
AO	AO.2	$\text{AO.2} > 0.546707$	0.4394317
AO	PDO.3	$\text{PDO.3} < -0.32$	-0.1264911
AO	PDO.3	$-0.32 < \text{PDO.3} < 1.69$	-0.4570953
AO	PDO.3	$\text{PDO.3} > 1.69$	0.1594810
AO	PDO.2	$\text{PDO.2} < -1.47$	-0.2333332
AO	PDO.2	$\text{PDO.2} > -1.47$	0.2856299
AO	NAO.3	$-\text{inf} < \text{NAO.3} < \text{inf}$	0.605332

Response Variable	Predictor	Predictor Range	Coefficient
QBO.3	QBO.2	-inf – inf	1.668503
	QBO.1	-inf – inf	-0.7513252
	SOI.2	-inf – 0	-0.2179054
		0 – inf	0.1443253
SOI.3	SOI.2	-inf – inf	0.5225828
	SOI.1	-inf – 4.8	-0.1979299
		-4.8 – inf	0.2421996
WP.3	WP.2	-inf – inf	0.1777548
	SOI.3	-inf – inf	-0.1286399
	NAO.2	-inf – 2.5	1.794230
		-2.5 – 1.3	-0.7410115
		-1.3 – 1.1	0.1816928
		1.1 – 2	-0.6978301
		2 – inf	1.609102
PDO.3	PDO.2	-inf – inf	0.8172722
	WP.3	-inf – 0	0.1927047
		0 – inf	0.03967585
AO.3	AO.2	-inf – 0.546707	0.1095149
		0.546707 – inf	0.4394317
	PDO.3	-inf – 0.32	-0.1264911
		-0.32 – 1.69	-0.4570953
		1.69 – inf	0.1594810
	PDO.2	-inf – 1.47	-0.2333332
		-1.47 – inf	0.2856299
	NAO.3	-inf – inf	0.605332

## References

Bessler, D. (2002) *Directed Graphs*

[http://agecon.tamu.edu/faculty/bessler/WebPage/Quito\(2a\).ppt](http://agecon.tamu.edu/faculty/bessler/WebPage/Quito(2a).ppt)

- Demiralp, S. and Hoover, K. (2003) "Searching for the Causal Structure of a Vector Autoregression," *Oxford Bulletin of Economics and Statistics* 65(supplement), 2003
- Ertoz, L., Steinbach, M., and Kumar, V. (2001) "Finding Topics in Collections of Documents: A Shared Nearest Neighbor Approach," *Proceedings of Text Mine'01, First SIAM International Conference on Data Mining*, 2001
- Glymour, C., Scheines, R., Spirtes, P., and Kelly, K. (1987) *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science and Statistical Modeling*, Academic Press, 1987.
- Hoover, K. (2004) *Automatic Inference of the Contemporaneous Causal Order of a System of Equations*. Working Paper.
- Lauritzen, S. (1996) *Graphical Models*, Oxford University Press, Oxford UK.
- Monetta, A. (2003) *Graphic Models for Structural Vector Autoregressions*, Laboratory of Economics and Management Working Paper Series. Pisa, Italy.
- Pearl, J. (2000) *Causality*, Cambridge University Press, Cambridge UK.
- Richardson, T., and Spirtes, P. (1996) *Automated discovery of linear feedback models*, Technical Report CMU-75-Phil.
- Robins, J., Scheines, R., Spirtes, P., and Wasserman, L. (2003) "Uniform Consistency in Causal Inference", *Biometrika* in press.
- Spirtes P., Glymour, C. and Scheines, R. (2000). *Causation, Prediction, and Search*, 2nd ed. New York, N.Y.: MIT Press.
- Steinbach, M., Tan, P., Kumar, V., Klooster, S., and Potter, C. (2003), *Discovery of Climate Indices Using Clustering*, KDD 2003.
- Swanson, N.R. and C.W.J. Granger, 1997, *Impulse Response Functions Based on a Causal Approach to Residual Orthogonalization in Vector Autoregressions*, *Journal of the American Statistical Association*, 92, 357-367.
- Zhang, J. (2003) "Strong Faithfulness and Uniform Consistency in Causal Inference," *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2003