



ECOLOGICAL SOCIETY OF AMERICA

Ecology/Ecological Monographs/Ecological Applications

PREPRINT

This preprint is a PDF of a manuscript that has been accepted for publication in an ESA journal. It is the final version that was uploaded and approved by the author(s). While the paper has been through the usual rigorous peer review process of ESA journals, it has not been copy-edited, nor have the graphics and tables been modified for final publication. Also note that the paper may refer to online Appendices and/or Supplements that are not yet available. We have posted this preliminary version of the manuscript online in the interest of making the scientific findings available for distribution and citation as quickly as possible following acceptance. However, readers should be aware that the final, published version will look different from this version and may also have some differences in content.

The doi for this manuscript and the correct format for citing the paper are given at the top of the online (html) abstract.

Once the final published version of this paper is posted online, it will replace the preliminary version at the specified doi.



Spatial 'convergent cross mapping' to detect causal relationships from short time-series

For submission to Ecology as a Statistical Report

23 December 2014

Running head: Detecting causation in short time-series

Adam Thomas Clark^{1*}, Hao Ye², Forest Isbell^{1,3}, Ethan R Deyle², Jane Cowles¹,
David Tilman^{1,4}, and George Sugihara²

**Corresponding author adam.tclark@gmail.com, ¹University of Minnesota, Dept. of Ecology, Evolution, and Behavior, 1987 Upper Buford Circle, Saint Paul, MN 55108, USA, ²Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA, ³Department of Plant Biology, 2502 Miller Plant Sciences, University of Georgia, Athens, GA 30602, ⁴Bren School of Environmental Science and Management, University of California, Santa Barbara, 2400 Bren Hall, Santa Barbara, CA 93106, USA*

Abstract

Recent developments in complex systems analysis have led to new techniques for detecting causal relationships using relatively short time-series – on the order of thirty sequential observations. Although many ecological observation series are even shorter – perhaps fewer than ten sequential observations – these shorter time-series are often highly replicated in space (i.e. plot replication). Here, we combine the existing techniques of convergent cross-mapping (CCM) and dewdrop regression to build a novel test of causal relations that leverages spatial replication, which we call multispatial CCM. Using examples from simulated and real-world ecological data, we test the ability of multispatial CCM to detect causal relationships between processes. We find that multispatial CCM successfully detects causal relationships with as few as five sequential observations, even in the presence of process noise and observation error. Our results suggest that this technique may constitute a useful test for causality in systems where experiments are difficult to perform and long time-series are not available. This new technique is available in the `multispatialCCM` package for the R programming language.

Key words: Causality, convergent cross-mapping, dewdrop regression, multispatialCCM

Introduction

Detecting causal relationships in complex systems is one of the fundamental and most challenging goals of science. Convergent cross mapping (CCM) has recently been introduced as a practical numerical approach for identifying causal relationships in weakly coupled nonlinear systems (Sugihara et al. 2012). This approach is potentially of great utility in ecology, where many systems appear to be weakly coupled and complex, and can therefore be difficult to analyze using traditional techniques (e.g. repeated measures or mixed effects ANOVA, GAMs, neural networks, or auto-regressive models). Typically, CCM can be applied to time-series of

roughly 30 or more sequential observations (Sugihara et al. 2012). However, ecological time-series are often even shorter. In this report, we present a method that expands the application of CCM to very short time-series that are spatially replicated (e.g. data from multiple plots).

CCM is based on an algorithm that compares the ability of lagged components of one process to estimate the dynamics of another. In ecology, these processes might represent time-series observations of environmental data, such as temperature, or of species data, such as population abundance. There are three basic ways in which two processes can be causally-linked: (i) neither influences the others' temporal dynamics, and the variables are therefore *causally unrelated*, (ii) a *forcing process* influences the temporal dynamics of a *response process*, but the response process has no effect on the forcing process in return, meaning that there is uni-directional causality, or (iii) there is bi-directional causality where each variable influences the others' dynamics. In practice, the ability of any method to accurately distinguish between these cases depends not only on the strength of underlying relationships, but also on the amount of data available, the presence of process noise and observation error, and the sensitivity and assumptions of the method. Here we explore cases for which the underlying strength and direction of causal interactions, process noise, and observation error are known, and use these to assess the detection limits of CCM when applied to short, replicated observational records.

As described by Sugihara *et al.* (2012), the dynamics of a uni-directional forcing process can be more accurately estimated using information from a response process than is true for the reverse case. This counterintuitive phenomenon occurs because the response process necessarily contains information about the forcing process whereas the reverse may not be true. For example, consider the dynamics of a species responding to an exogenous environmental variable such as length of day. In a weakly-coupled system, many processes influence species abundance,

and consequently, day length alone will not be a good predictor of the organism's endogenous abundance dynamics. However, changes in day length (the forcing process) necessarily propagate to and influence the dynamics of the species (the response process), and day length therefore *can* be predicted from species abundance. Thus, observed states of a forcing process should (1) be significantly better described by observed states of a response process than would be expected by random chance, and (2) the accuracy of that description should improve with increasing time-series length, as this increases the amount of information available with which to estimate system dynamics (see example in Figure 1a). This joint criterion therefore provides a metric to test for significant causal relationships among processes. Algorithmic and visual descriptions of this approach are available in the supplement of Sugihara *et al.* (2012).

Although CCM may perform relatively well on short time-series of 30+ observations, longer time-series are generally preferable, particularly when causal relations are weak or process noise and observation error are large (Sugihara *et al.* 2012). However, ecological time-series are often much shorter, and experiments and observations generally focus on spatial rather than temporal replication. Thus, many potential applications of CCM may not be feasible using the currently published framework. Nevertheless, spatially replicated data do contain temporal information, and it seems plausible that sufficient replication might compensate for brevity in time-series. Here, we present a method (multispatial CCM) to recover this information for application in CCM through a bootstrapping technique that has previously been applied to single variable simplex projection (dewdrop regression; Hsieh *et al.* 2008). Dewdrop regression combines information from many short time series of systems that share similar dynamical forms to predict their future dynamics. We adapt this method for CCM in order to leverage the spatial breadth in ecological datasets. We also present the first comprehensive test of detection power

and error rates for CCM across a broad range of forcing strengths, observation error, and process noise magnitudes.

Methods

Multispatial CCM algorithm:

To test the predictive ability of variables, we use simplex projection, an application of Takens' Theorem (Sugihara 1990). Simplex projection predicts the dynamics of a process that is part of a larger system by using information from multiple lagged observations of that single process. Consider as an example a predator-prey system, where both predator and prey dynamics influence one another. Because interacting processes contain information about one another, accurate predictions of future predator population dynamics can be made based on knowledge of previous predator dynamics, even if no information on prey populations is available (Schaffer 1985, Sugihara and May 1990, Sugihara 1994). Three consecutive years of predator population dynamics might be compared to three-year periods from historical data. To predict next year's population size, simplex projection identifies a subset of observed three-year dynamics that are most similar to the current three-year trend, and averages predictions across these observations weighted by their similarity. Recent work has shown these techniques to be applicable under a wide range of circumstances (Sugihara and May 1990, Deyle and Sugihara 2011), and perform well both on simulated and observed data (Perretti et al. 2013, Deyle et al. 2013).

Classical simplex projection uses a single long time-series, which means that most observations have many corresponding historical observations, and many lagged time steps that can be used to make predictions and estimates. In the multispatial approach that we propose here, we instead leverage information from many short time-series, by drawing samples from the pool of all spatially-replicated observations. We then look across these samples to find observations

with similar historical dynamics, and use a weighted average of those observations to estimate expected dynamics for subsequent time steps. Based on these estimates, we can characterize how well the dynamics of a putative forcing process are described by the dynamics of a suspected response process, and thus assess their causal relatedness.

To test for significant causal relationships between two candidate processes, the multispatial CCM algorithm proceeds in five steps, which are described briefly below. An annotated implementation of these steps for the R programming environment (R 2014) is available in the `multispatialCCM` package available through CRAN, and worked examples of applying these functions to real and simulated data are available in Appendix A.

(1) Determine the best embedding dimension for the analysis:

In simplex projection, the accuracy of predictions varies as a function of *embedding dimension* (E). E describes the size of the time-windows (i.e. number of time steps) that are used for prediction ($E = 3$ in the predator prey example above). One can estimate an optimal embedding dimension by using simplex projection to test the ability of a process to predict its own dynamics through "leave-one-out cross-validation" (i.e. removing one observation from the time series and using the rest of the time series to predict its dynamics).

In multispatial simplex projection, we sample across many short time series from spatial replicates. Because of this, it is important that time lags should not cross "gaps" between plots in the composite time-series. Even though multispatial CCM assumes that data from multiple plots come from the same dynamical system, observations from the end of the time-series of one plot will have no bearing on the beginning of the time-series of another plot. As a result, lagged dimensions are only considered when they come from the same plot (i.e. do not cross any "gaps"), which therefore limits the potential embedding dimension E for an analysis. For

example, if only five observations are recorded in each plot, then E cannot exceed four, as the analysis requires four observations to describe historical dynamics, followed by one observation against which to test estimated values.

(2) Test for nonlinearity and stochastic noise:

Because CCM applies in coupled nonlinear systems, it is important that the system under consideration is not purely random, in which case other tests for causality (e.g. Granger's test, see Detto et al. 2012), rather than CCM, might instead be considered (Sugihara et al. 2012).

Furthermore, it is important to test that dynamics are not so dissimilar among plots, or so strongly influenced by stochastic noise, that causal links cannot be recovered. To accomplish this, the same simplex projection algorithm described in (1) can be applied. Using the E identified with techniques described above, predictions should be made based on historical dynamics for observations that are increasingly far into the future. If the system is nonlinear, then predictive power should significantly decrease with increasing prediction time (Sugihara 1994). Predictive power for short time intervals should also be better than expected by random chance. Note that this is a "necessary but insufficient" criterion for determining nonlinearity. For a completely random process, predictive ability should be equally low regardless of prediction interval. However, an autocorrelated system dominated by stochastic noise could show the same decreasing pattern that we use to detect nonlinearity. In these cases, CCM should (correctly) return that there is no causal forcing in either direction, as increasing information about the system should not increase predictive ability.

(3) Calculate two processes' abilities to describe each others dynamics using CCM:

In addition to varying with E , predictive skill also varies as a function of *library length* (L). L describes the number of historical observations that are used to generate predictions. In

multispatial CCM, L increases both with increased observation time in each spatial replicate, and with increased number of spatial replicates included in the composite time series. As L increases, it is more likely that previous trends will have been observed that are similar to current trends, thus improving the predictive ability of simplex projection. Because processes that are responding to the effects of a forcing process should be good indicators of the dynamics of the forcing process (just as species abundance should be a good indicator of day length in the example above), the strength of that description should consequently increase as L increases, approaching some finite limit for very large L (Sugihara et al. 2012).

(4) Use bootstrapping to leverage spatial information:

To prevent the order in which spatial replicates are sampled from influencing the outcome of the test, we use a simple, nonparametric bootstrapping routine. Given n spatial replicates, we draw n samples (with replacement) from among all spatial replicates, and repeat tests for the best E , the relationship between predictive power and interval, and for CCM for this re-sampled assemblage. We then repeat this procedure for many iterations, in order to average across many potential combinations and orderings of spatial data. This also provides estimates for uncertainty around the predictive power estimated in each test. In the annotated examples we present, we use 100 iterations per test, and for each iteration calculate the Pearson correlation coefficient ρ , comparing predicted estimates from the CCM algorithm to observed values at each library length. For analysis of real data, the number of iterations required will almost always be greater than 100 (e.g. we use 1000 replicates for our real world examples). We discuss how to determine whether sufficient iterations have been conducted in Appendix A.

(5) Test whether these predictions indicate a significant causal relationship:

Finally, it is necessary to determine whether the calculated ρ is significantly greater than

zero, and whether it increases significantly with L . Testing for this increase can be complicated by two factors. First, because the relationship between ρ and L is often not well-described by simple parametric curves, statistical tests such as regression based on particular functions may not be a good indicator of significance in all cases. Second, a very rapid rise of ρ with L may be an indication of synchrony (see Sugihara et al. 2012 and citations therein). Synchrony arises when forcing of one process by another is so strong that the response process dynamics become subordinate to the forcing process. Consequently, both processes become good predictors of one another's dynamics, even if forcing is solely uni-directional, which can confound CCM analysis and should not necessarily be taken as an indication of a causal link.

In the analyses presented here, we interpret CCM results for which ρ is greater at the longest L available than at the shortest L available, and where ρ at the longest L is greater than zero as indicating causal forcing. Note that the longest L is determined by data availability, whereas the shortest is determined by E . To test for statistical significance of this signal, we use the nonparametric bootstrapping conducted in (4) to determine whether the ρ vs. L relationship passes both of these criteria for at least 95% of bootstrapped iterations.

Simulating system dynamics:

We applied the multispatial CCM method to a simulated dynamical system as considered in Sugihara *et al.* (2012). The model describes competition between two species, where X and Y signify the abundance of the two species, t and $t+1$ are times, r_X and r_Y are species' intrinsic growth rates, α describes the effect of species X on the dynamics of species Y , and β describes the effects of species Y on species X :

$$X(t+1) = X(t) (r_X - r_X X(t) - \beta Y(t))$$

$$Y(t+1) = Y(t) (r_Y - r_Y Y(t) - \alpha X(t))$$

We used the parameters $\beta = 0$ and $\alpha \geq 0$ for all simulations presented here. This corresponds to a system where process Y is influenced by process X , but X is not influenced by Y , except for the case where $\alpha = 0$, in which case neither process influences the other.

To introduce process noise into our simulation, we drew r_X from a random normal distribution centered on 3.8 with standard deviation σ_P , and r_Y from a random normal distribution centered on 3.5 also with standard deviation σ_P for each simulation of each plot. Consequently, species dynamics varied among plots based on σ_P , but did not vary through time within a single plot. To introduce observation error into the system that was proportional to species abundance, we multiplied each observed value of X and Y by a random, log-normally distributed variable with mean 1 (0 on the log scale) and standard deviation σ_O after simulating the entire time-series. Consequently, the observation error did not change the actual trajectory of the processes, but did alter the values that were used for subsequent analyses.

To test the ability of multispatial CCM to detect different causal signals across a range of forcing strengths and magnitudes of process noise and observation error, we simulated the system for five different values of α spaced evenly between 0 and 2.5, for five values of σ_P spaced evenly between 0 and 0.2, and for five values of σ_O spaced evenly between 0 and 0.2. In these simulations, the average values for $X(t)$ and $Y(t)$ are both around 0.7. Thus, across the range of α values that we tested (0.000, 0.625, 1.250, 1.875, and 2.500), roughly 0%, 7.1%, 14.1%, 21.2%, and 28.2% of the dynamics of $Y(t)$ were determined by the dynamics of $X(t)$ respectively. Similarly, the range of σ_P and σ_O values that we tested (0.00, 0.05, 0.10, 0.15, and 0.20) corresponded to average differences of roughly 0%, 1.6%, 3.1%, 4.6%, and 6.0% respectively in individual species' growth rates among plots (for σ_P), and 0.0%, 5.9%, 11.4%, 17.0%, and 23.4% respectively between observed and actual values of $X(t)$ and $Y(t)$ (for σ_O).

We compared these parameters in a fully factorial design ($5 \times 5 \times 5 = 125$ combinations), simulated each parameter set 1200 times, and bootstrapped each simulated library 100 times. This resulted in 15 million individual multispatial CCM tests for causality in each direction. In each simulation, we chose the first value for X and Y in each plot from a random, uniform distribution between 0 and 1. We considered three scenarios, representing different kinds of plot data. In scenario 1, we simulated 20 observations from each of 100 separate systems, representing 20 sequential observations in 100 "plots". Second, we simulated 10 observations from each of 100 plots, and third, 5 observations in each of 100 plots to assess the effect of shrinking time-series length on multispatial CCM performance. These time periods roughly correspond to the number of annual observations that might be expected from an established Long Term Ecological Research site, from a single long-term research project, and from a PhD dissertation respectively. In the Appendix A, we consider a further set of cases in the same model with bi-directional causation (i.e. where $\beta > 0$).

Real world examples:

We also apply CCM to two examples of real world data. First, we consider a short time-series (4-7 sequential samples in each of 72 spatial replicates) describing the dynamics of invading plant species and soil nitrate in the "Big Biodiversity" experiment at Cedar Creek. Experimental results suggest that soil nitrate dynamics force invasion dynamics, while the reverse should not be true (Fargione and Tilman 2005). Second, we consider a longer time series (62 replicates of about 20 observations each) recording annual dynamics for summer precipitation and aboveground biomass for the cool-season, drought-intolerant grass *Elymus* (*Agropyron*) *repens*. In this system, *E. repens* dynamics should be forced by precipitation, but the reverse should not be true. Eight further examples, and detailed methods, diagnostic tests,

231 and descriptions of the data, are presented in the Appendix A.

232 **Results**

233 Our analyses of simulated data using multispatial CCM correctly recovered causal forcing
234 of process Y by process X under most cases considered for all three plot lengths that we tested,
235 particularly when observation error and process noise were small (Figure 2a). It also performed
236 well for both real world examples, correctly detecting the presumed direction of causal
237 association for both systems (Figure 1b-c).

238 When no observation error or process noise were present, the test correctly indicated
239 causal relations in almost every instance. In most cases where σ_O (strength of observation error)
240 was less than or equal to 0.05, failures to detect causal signals were below 5% (and were
241 generally around 0%), though for very high observation error ($\sigma_O = 0.2$), causal signals were
242 almost never detected. Detection strength also declined in most cases with process noise, though
243 the effect was not as large as for observation error. In general, detection rates remained higher
244 even with process noise when time-series were long (maximum library length = 20), and for
245 moderately strong forcing ($\alpha = 1.875$). Performance for bi-directional simulations was similar
246 (Figure A16 in Appendix A).

247 The multispatial algorithm correctly indicated no forcing of process X by process Y (i.e.,
248 in the "wrong direction") in most cases as well (Figure 2b). For weak forcing ($\alpha = 0.625$), error
249 rates were around 0% for all levels of observation error and process noise. For systems with
250 stronger forcing, false detection rates rose when observation error was not present, particularly
251 when process noise was high, reaching 60-100% for the case with the strongest forcing ($\alpha = 2.5$).
252 When no forcing was present ($\alpha = 0$), the test correctly indicated no causal forcing in most cases.
253 In all cases with any observation error ($\sigma_O > 0.2$), failure rates were below 5%. Without

observation error or process noise, failure rates were somewhat higher, but remained below 20%. The test performed worst when there was no observation error, but process noise was strong, and showed failure rates as high as 80% in both directions, particularly for long libraries.

Discussion

Multispatial CCM appears to perform well for moderately short and very short time-series (20, 10, and 5 observations per plot) when there is no observation error or process noise, and when forcing is weak. Without error or noise, false detection rates (type I error) were less than 5% for all forcing strengths considered, except for the strongest forcing strength tested. Failures to detect causal links (type II error) were also around 0% when process noise and observation error were absent. These results indicate that multispatial CCM performs best under the same sort of circumstances that classical CCM does: low observation error and process noise relative to forcing strength (Sugihara et al. 2012). Encouragingly, errors only increased moderately as observational periods decreased, suggesting that short time-series can be analyzed without a large decrease in test performance. Note that shorter observational periods in our simulations also lead to shorter L (i.e. total number of observations), as we retained the same number of spatial replicates among all tests. Some of the decrease in algorithm performance could therefore be alleviated by increasing spatial replication to compensate for decreasing observational period.

The test is conservative when observation error and process noise are present, and thus failures to detect causal links in systems where large observation error is suspected should not necessarily be taken as evidence for a lack of causal relationship. When observation error and process noise are present, there appears to be a trade-off between type I and type II errors. Increased observation error decreases correct detections of causality, but also decreases spurious detections. Process noise slightly decreased the rate of detection of causal signals, but only when

forcing was strong, and had mixed effects on false detection rates. Our results suggest a steep decline in algorithm performance around $\sigma_O = 0.1$, which represents roughly a 11.4% difference between true and observed values in our system.

The test is anti-conservative when there is strong forcing, particularly in the absence of error and noise, and thus detection of apparently bi-directional causality in systems where strong forcing is suspected might instead simply be uni-directional causality. These detection failures are demonstrative of a failure of our convergence detection method, not necessarily of the CCM algorithm itself, and are likely a result of synchrony: As processes become tightly coupled, each tends to become a good estimator of the others' dynamics even though causation is unidirectional (Sugihara et al. 2012). Signals that are confounded by synchrony are characterized by a sharp rise in ρ over the first few time steps, followed by a long flat plateau (see Figure 3e in Sugihara et al. 2012), which is therefore a useful diagnostic that can warn of confounding. When synchrony is the product of an exogenous variable (e.g. joint response to temperature), it can often be minimized by taking “first differences” of the cross-correlated variables, where the preceding observation is subtracted from each observation prior to analysis (Granger and Newbold, 1974). Where it cannot be eliminated (i.e. very strong forcing), synchrony may preclude data from CCM analysis.

The performance of multispatial CCM, based on a simple simulation of competing species, does indicate that algorithm performance is sensitive to properties of the dynamical systems under investigation. We chose this functional form because of its computational simplicity, and because it makes the direction and strength of interactions, process noise, and observation error easy to interpret and quantify. However, a variety of other dynamical forms were considered by Sugihara *et al.* (2012), with very little variation in algorithm performance for

classical CCM. Encouragingly, the algorithm performs well for a number of real-world cases, suggesting that at least some real systems meet the assumptions required for CCM analysis. In particular, synchrony is rarer than in the simple simulated system that we consider. Nevertheless, where the assumptions of multispatial CCM are not met or where manifolds are sparsely sampled (e.g. a system with an exogenous time trend or “non-stationary” processes), the method is unlikely to work. In the Appendix A, we include five examples that our diagnostic tests show do not meet all the required assumptions, and show how these problems lead to method failures (in addition to five examples where the method performs well).

For a wide range of forcing strengths, and for moderate levels of observation error and process noise, multispatial CCM performs well in detecting causality, even for time-series as short as 5 observations per plot. Our findings suggest that multispatial CCM is a useful analytical tool for empirical studies, and can help inform questions about causal links in systems where experimental studies are not possible or available, or where the results from experiments are difficult to interpret. Based on our experience with the technique and the results that we present here, we suggest that multispatial CCM can be applied as a useful analytical method, but caution that results based on the convergence detection algorithm that we present here are often very conservative. Further study with other dynamical systems will be important for determining the general applicability of multispatial CCM, and for identifying convergence detection algorithms that are less conservative for a wide variety of applications. We look forward to future examples of multispatial CCM applications, and to corresponding with parties that are interested in applying it to novel systems.

Acknowledgements: We thank R. Barnes, S. Binder, M. Burgess, M. Kosmala, C. Lehman, C. Perretti, K. Thompson, and P. Wragg for insightful discussions about this project. We also thank

B. Inouye for helping us restructure our original manuscript, and T.J. Hefley and an anonymous reviewer for helpful comments, particularly their suggestions to add real world analysis examples. Computing time for ATC was supported through NSF/XSEDE and the University of Minnesota Supercomputing Institute. Data collection was supported by the NSF LTER program including DEB-0620652 and DEB-1234162, and by the Cedar Creek Ecosystem Science Reserve and the University of Minnesota. ATC, ED, and HY were supported by the NSF GRFP. ED was supported by the EPA STAR Fellowship. GS and HY were supported by NSF DEB-1020372. ED, GS, and HY were supported by NSF-NOAA CAMEO program Grant NA08OAR4320894. Additional support to GS was provided by the Sugihara Family Trust, the Deutsche Bank-Jameson Complexity Studies Fund, and the McQuown Chair in Natural Sciences.

References

- Detto, M., A. Molini, G. Katul, P. Stoy, S. Palmroth, and D. Baldocchi. 2012. Causality and Persistence in Ecological Systems: A Nonparametric Spectral Granger Causality Approach. *American Naturalist* 179:524–535.
- Deyle, E. R., M. Fogarty, C. Hsieh, L. Kaufman, A. D. MacCall, S. B. Munch, C. T. Perretti, H. Ye, and G. Sugihara. 2013. Predicting climate effects on Pacific sardine. *PNAS* 110:6430–35.
- Deyle, E. R., and G. Sugihara. 2011. Generalized Theorems for Nonlinear State Space Reconstruction. *PLOS ONE* 6.
- Fargione, E.J., and D. Tilman. 2005. Diversity decreases invasion via both sampling and complementarity effects. *Ecology Letters* 8:604-11.
- Granger, C.W.J., and P. Newbold. 1974. Spurious regressions in econometrics. *Journal of Econometrics* 2:111-120.

346 Hsieh, C., C. Anderson, and G. Sugihara. 2008. Extending Nonlinear analysis to short ecological
347 time series. *American Naturalist* 171:71–80.

348 Perretti, C. T., S. B. Munch, and G. Sugihara. 2013. Model-free forecasting outperforms the
349 correct mechanistic model for simulated and experimental data. *PNAS* 110:5253–5257.

350 R Core Development Team. 2014. R: a language and environment for statistical computing. R
351 Foundation for Statistical Computing.

352 Schaffer, W. M. 1985. Order and Chaos in Ecological Systems. *Ecology*. 66:93-106.

353 Sugihara, G. 1994. Nonlinear forecasting for the classification of natural time-series.

354 *Philosophical Transactions of the Royal Society A* 348:477–495.

355 Sugihara, G., and R. M. May. 1990. Nonlinear forecasting as a way of distinguishing chaos from
356 measurement error in time-series. *Nature* 344:734-741.

357 Sugihara, G., R. May, H. Ye, C. Hsieh, E. Deyle, M. Fogarty, and S. Munch. 2012. Detecting
358 Causality in Complex Ecosystems. *Science* 338:496-500.

359 **Ecological Archives Material:**

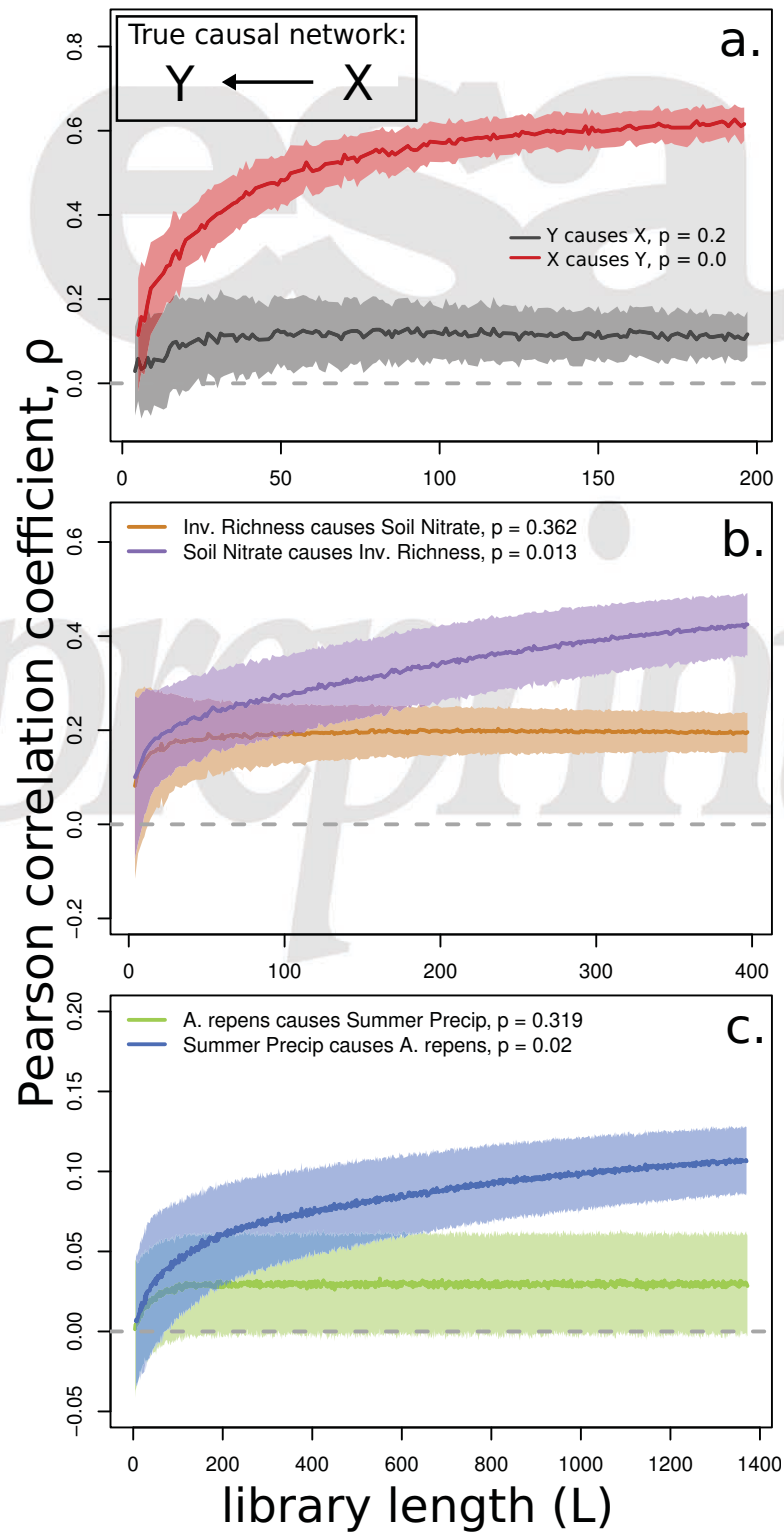
360 **Appendix A:** Supplemental examples of multispatial CCM analyses

361 **Supplement:** Data for real-world examples, and source code for the `multispatialCCM` R
362 package

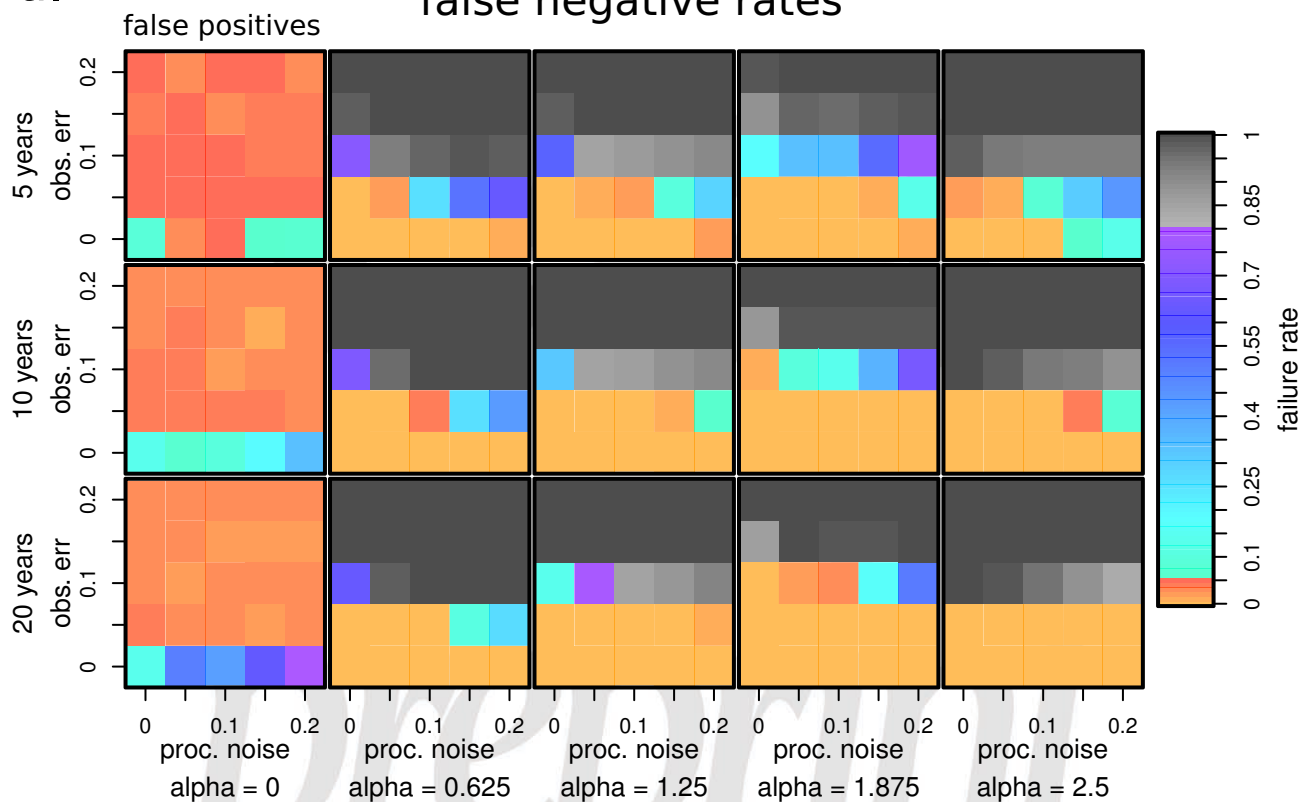
Figure legends

Figure 1: Example applications of multispatial CCM. Lines and shaded regions show mean \pm one standard deviation, from bootstrapped iterations. Causal forcing is indicated when (1) the Pearson correlation coefficient ρ is significantly greater than zero for large library length L , and (2) that ρ increases significantly with increasing L . (a) A simulated system of interacting processes where X causes Y , but Y does not cause X , as correctly indicated by the test (causal network shown in inset). Based on 20 observations in 10 plots, as in the `multispatialCCM` R package example files. (b) A real world example (Figure A14b in Appendix A) comparing soil nitrate and invading plant species richness in a plant biodiversity experiment at Cedar Creek, indicating that nitrate dynamics force invader species richness, but the reverse is not true. (c) A second example (Figure A11a in Appendix A) comparing annual summer precipitation (June–August) and *Elymus (Agropyron) repens* aboveground biomass dynamics, indicating that biomass is forced by precipitation, but the reverse is not true. See Appendix A for details.

Figure 2: Failure rates of multispatial CCM. Panel (a) shows the failure to detect an existing causal effect, while panel (b) shows the false detection of a non-existent effect. Colors indicate proportion of simulations for which the algorithm gave the incorrect result, with orange indicating 0–5% failures, blue indicating 5–80%, and grey/black indicating 80–100%. Horizontal axis in all panels shows increasing rates of process noise (σ_P), at five evenly-spaced levels between 0 and 0.2. Vertical axis in all panels shows increasing rates of observation error (σ_O), at five evenly-spaced levels between 0 and 0.2. From left to right, panels vary in forcing strength of Y by X (α), at five evenly-spaced levels between 0 and 2.5. From top to bottom, panels vary in number of time steps included in the simulation (100 plots with 5 observations each, 10 observations each, and 20 observations each).



a. false negative rates



b. false positive rates

