Vector Autoregression, Cross-Correlation, and Cross-Recurrence Quantification Analysis: A Case Study in Social Cohesion and Collective Action

Megan Chiovaro (megan.chiovaro@uconn.edu)
Department of Psychological Sciences, Center for the Ecological Study of Perception and Action,
University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269 USA

Leah C. Windsor (lcwells@memphis.edu)
Institute for Intelligent Systems, University of Memphis,
365 Innovation Drive, Suite 303, Memphis, TN 38152

Alexandra Paxton (alexandra.paxton@uconn.edu)
Department of Psychological Sciences, Center for the Ecological Study of Perception and Action,
University of Connecticut, 406 Babbidge Road, Unit 1020, Storrs, CT 06269 USA

Abstract

As time series analysis continues to capture the interest of cognitive and behavioral researchers, it is increasingly important to evaluate these methods and compare their respective insights. Here, we evaluate three popular analyses: vector autoregression, cross-correlation, and cross-recurrence quantification analysis. Using social cohesion data derived from Twitter and daily counts of real-world events during the Arab Spring, we present a case study using these methods and evaluate their benefits, limitations, and differences in results. We propose that researchers interested in time series analysis consider these differences and use multiple methods to assure reliability of their results.

Keywords: vector autoregression; cross-correlation; cross-recurrence quantification analysis; time series analysis; social dynamics

Introduction

Time series analysis has grown increasingly popular over the past several decades across the behavioral and psychological sciences. While traditional analyses of human behavior and cognition tended to focus on single means or standard deviations, recent multidisciplinary efforts have pushed researchers to see the variability and fluctuations of their time series—whether in measurements of motor behavior or brain areas or political movements—as not only useful but *vital* information about their phenomena of interest (Carello & Moreno, 2005). These time series methods are especially important when studying noisy, real-world data to understand the evolution and *co*-evolution of systems over time.

As researchers move toward naturalistic time series analysis, many tend to focus on using one single type of method. Prioritizing a particular analysis across various experiments can be beneficial for testing reliability of results, but using multiple methods can help answer new theoretical questions and extend the knowledge of how systems interact. However, adoption of other kinds of time series analyses is often limited by knowledge of different methods and their relations to one another. In this paper, we evaluate three common time series methods: vector autoregression, cross-correlation, and cross-recurrence quantification analysis.

Most prominently used in the field of econometrics, vector autoregression (VAR; Sims, 1980) allows for the investigation of the relationship between multiple signals. While its forecasting benefits have made it attractive to economists, the social sciences have begun to harness it as well. Some of the primary users of VAR are political scientists, who employ it to answer questions of causal inference (Freeman et al., 1989). It has been used to evaluate the coevolution of presidential approval ratings, economic factors, and leader discourse (Love & Windsor, 2018).

A fundamental signal processing method, cross-correlation analysis (CC; Shumway & Stoffer, 2006) quantifies the correlation between two signals in and across time. CC has been used to investigate psychological, financial, mechanical, and computational phenomena, amongst others. Given its relative simplicity, it is often a first choice for those getting started with time series analysis.

Originating from dynamical systems theory, cross-recurrence quantification analysis (CRQA; Zbilut et al., 1998) quantifies the repeating (or *recurrent*) patterns of two co-evolving time series, and it has become one of the most popular nonlinear analyses in the social sciences. In psychological science, it has been used to investigate phenomena such as emotion dynamics (Main et al., 2016) and gaze coupling (Richardson & Dale, 2005). CRQA has the *ability* to capture unique nonlinearities of systems (e.g., Brick et al., 2018) that are often overlooked by linear analyses, although differences in metrics do not necessarily imply nonlinearity.

The current paper aims to act as a practical guide for selecting between and applying these methods. We build on a growing literature of time series comparisons, such as time-delay stability comparisons of cross-correlation and joint recurrence quantification analysis for the estimation of coupling strength (Tolston et al., 2020), comparisons of linear time series analysis methods in the investigation of nonverbal synchrony (Schoenherr et al., 2018), and conceptual comparisons of different time series analyses (Gates & Liu, 2016). Here, we compare two linear analyses and a nonlinear analysis—all of which have the ability to identify leading-following behavior of two time series—to investigate a coupled system of real-world political action and online social cohesion during the Arab Spring. We discuss the convergence and divergence of results and the implications of the differences. Through this work, we hope to guide researchers in selecting appropriate methods to best answer their theoretical questions.

Time Series Analyses

Vector Autoregression (VAR)

Vector autoregression is a linear model that captures the relationship between multiple variables across time (Sims, 1980). As an extension of univariate autoregression, VAR can account for multivariate data through the use of vectors. Each endogenous variable is given its own linear equation that includes lagged values of itself, lagged values of the other variables in the model, and an error term. The equations are all of a particular *order*, which characterizes the number of time lags used in the model. A VAR equation with one variable and n time lags can be expressed as:

$$y_t = c + \Phi_1 y_{t-1} + \cdots + \Phi_n y_{t-n} + e_t$$

where y is the variable, c is a constant, Φ are the weights of the terms, and e is an error term. A first-order VAR with two variables can be expressed as:

$$\begin{aligned} y_{1,t} &= c + \Phi_{1,1} y_{1,t-1} + \Phi_{1,2} y_{2,t-1} + e_{1,t} \\ y_{2.t} &= c + \Phi_{2,1} y_{1,t-1} + \Phi_{2,2} y_{2,t-1} + e_{2,t} \end{aligned}$$

The number of terms in the equations will always be identical and will equal two (the error and constant terms) plus the number of time lags times the number of terms. These equations allow modeling of each variable based on its past values and the past values of other variables in the model.

Assumptions The primary assumption of VAR is that the time series are stationary. Stationary processes are those which underlying joint probability distribution, mean, and variance remain the same over time. While many time series are non-stationary, there are methods for assuring stationarity (see Process subsection, below). Two other assumptions are that the variables have been sampled consistently at a regular interval and that the time series are of the same scale.

Process When conducting VAR, it is essential to first assess the stationarity of the time series. This can be done using the Augmented Dickey-Fuller (ADF) Test. If an ADF test indicates that a time series is non-stationary, differencing can be used to remove the trend. Typically, taking the first difference removes linear trends, while taking the second difference removes quadratic trends.

Once all of the time series intended for VAR are stationary, an optimal lag length must be determined, allowing for maximal information given a preference for model parsimony. There are four common information criteria used to determine optimal lag: Akaike (AIC), Schwarz-Bayes (SC or BIC), Akaike's Final Prediction Error (FPE), and Hannan-Quinn (HQ). Of these four, AIC appears to be most commonly used in the social sciences, while SC is the most conservative. The VAR model can then be constructed using the selected optimal lag(s) with various statistical programs, including the vars package (Pfaff, 2007) in R (R Core Team, 2020). Once the model has been built, the modulus of each root is checked to assure that it is less than one, implying that the model is stable.

One of the most common structural analyses used on VAR models is Granger causality, indicating which variable "Granger causes" changes in the other. The underlying assumption of Granger causality is that variable Y can be better predicted when accounting for the history of both Y and some other variable X, rather than just accounting for Y alone (Freeman, 1983). In this instance, it would be said that X Granger causes Y. Alternatives to Granger causality include impulse-response functions (IRFs) and forecast error variance decomposition (FEVD; Freeman, 1983).

Limitations Although researchers can handle nonstationarity in time series through differencing before running VAR, the resulting time series is inherently different from the original. As opposed to being an evolution of a *value* over time, the new time series is an evolution of the *change in value* over time. In the instance where there are both non-stationary and stationary time series going into the model, some time series would then be differenced, while others would not. One could overcome this by taking the difference of both time series, to assure that all are of the same scale, but this may change the theoretical questions that can be answered with the analyses.

In addition to the limitations posed by the assumption of stationarity, VAR is a linear model and cannot capture nonlinearities that may be essential to the system. Researchers should check for nonlinearities in the system to assure that VAR is not overlooking critical dynamics. Additionally, scaling from bivariate to multivariate VAR requires corrections for multiple comparisons.

Benefits A significant advantage of VAR over other methods considered here is its ability to analyze as many time series as desired. There are no limitations to how big the vector in the model is. However, it is important to recognize that the use of many time series may cause an increased amount of error to be incorrectly accounted for. VAR also allows for various kinds of structural analyses to analyze the patterns of causal behavior in the system.

Cross-Correlation (CC)

As a measure of similarity across time at multiple time lags, cross-correlation can capture the leading-following behavior of two time series (Gates & Liu, 2016). At each lag, a correlation coefficient is calculated and can be tested for significance. A significant coefficient indicates that one time series leads the behavior of the other by s time steps.

Assumptions CC assumes that time series are stationary. This can again be overcome by taking the difference of the time series to remove the trend. Stationarity should be checked again afterwards to assure the trend was removed.

Process First, a correlation coefficient is calculated determining how well the time series predict one another. The equations for the function are:

$$c_{ij}(t) = \frac{1}{n} \sum_{s=\max(1,-t)}^{\min(n-t,n)} \left[X_i(s+t) - \overline{X_i}\right]\left[X_j(s) - \overline{X_j}\right]$$

$$r_{ij}(t) = \frac{c_{ij}(t)}{|c_{ij}(0)|}$$

where c is the covariance coefficient, t is time, s is the lag, X is the current observation of the specified time series, $\overline{X}$ is the mean value of the specified time series, and r is the cross-correlation coefficient (Gates & Liu, 2016; Shumway & Stoffer, 2006). After a coefficient is calculated, one time series is shifted, and the coefficient is calculated again. The result is a series of correlation coefficients and their probability of occurring by chance, measuring which time series leads the other and by how many time steps.

Limitations CC is also a linear model and may overlook inherent nonlinearities fundamental to the system. There is also limited opportunity for interpretation of output, given that the only result is a series of coefficients and their significance, unlike VAR and CRQA.

Benefits Of the three methods considered here, CC is perhaps the simplest analysis to understand and implement. There are no input parameters to select aside from the number of lags (also required with VAR and CRQA), leaving less room for experimenter error, and its output is simple to interpret.

Cross-Recurrence Quantification Analysis (CRQA)

As an extension of recurrence quantification analysis (which projects a system onto itself; Trulla et al., 1996), CRQA allows for the dynamical analysis of two co-evolving time series. Whenever the two time series are in the same state, either in synchronous time or across time, CRQA documents it and analyzes the patterning of the shared states.

Assumptions For categorical CRQA, the two time series must be in the same unit. If the two time series are not in the same unit, they can be transformed (e.g., through creating quantiles). It is also assumed that the two time series going into the analysis have been sampled at a rate that corresponds to their native timescale. For example, a system that changes every hour must be sampled more frequently than a system that changes once every day. Under- or oversampling can skew the results of the analysis.

Process Categorical CRQA captures the shared evolution of two categorical time series. By projecting two time series onto one another, we are able to quantify the interaction of the systems. CRQA identifies shared states, both in synchronous and asynchronous time, as *recurrent points*. These points may be sequential in time, resulting in *trajectories*. These are extended periods where the two time series get "stuck" together in a particular state. While the continuous case of CRQA requires the estimation of multiple parameters (Riley & Van Orden, 2005), no parameter estimation is required for the categorical case. When using the crqa package (Coco & Dale, 2014) in R, this means that the *delay* can always be set to 0; the *embedding dimension* to 1; the *radius* to any value higher than 0 and lower than 1; and the *Theiler window* to 0. Further discussion of these parameters' uses in the continuous case can be found in Riley and Van Orden (2005).

Using these parameters, the next step is constructing a *cross-recurrence plot* (CRP). In this visual representation, the two time series are plotted against one another on the *x*- and *y*- axes, and for every *x*,*y* coordinate where the two have the same observation value, a point is plotted to indicate recurrence. The plot can be both visually inspected for the texture and pattern of the system by looking at the trajectories and box-like areas and quantified with various metrics. The leading-following behavior of the time series is analyzed through another visual representation, known as a *diagonal recurrence profile* (DRP; Main et al., 2016). DRPs are conceptually and visually similar to cross-correlation plots. In a DRP, recurrence rate is on the *y*-axis, and time lag is on the *x*-axis; a synchronous relationship would display a peak at 0, while a leading-following relationship would show a peak at a negative or positive lag value.

Metrics Nine quantitative measurements can be derived from the CRP to conduct inferential statistics. The first metric is *recurrence rate* (RR), the proportion of recurrent points to all possible places where there could have been a point. RR is a general measure of how much the two systems are sharing the same state, but it does not incorporate information about the structure or location of the points. *Determinism* (DET) is a measure of the structure of recurrence, formalized as the proportion of points that occur in diagonal trajectories. It indicates how deterministic or how random the coupled systems' co-evolution is. Total *number of lines* (NRLINE) also indicates stability and scales with structure, as an increased number of lines suggests stronger coupling. Max line (*maxL*) is the longest shared diagonal trajectory and is a measure of attractor strength (Pellecchia et al., 2005). Additional metrics are *entropy* (ENTR), *normalized entropy* (nENTR), *laminarity* (LAM), and *trapping time* (TT). For more detail, see Coco and Dale (2014).

Limitations First, compared to VAR and CC, CRQA has fewer educational resources available, perhaps because its use is somewhat limited in the cognitive science community. Second, CRQA does not inherently produce significance metrics. Because there is no null comparison, one must be generated in order to perform inferential statistics. This can be done by conducting approximate permutation tests (e.g., Chiovaro et al., *under review*), where a large set of pseudo-time series is generated by randomly sampling from the original time series without replacement. This set is then subjected to CRQA, and the resulting metrics are used as the null distribution, although RR must be excluded from plot-wise analyses (like those here), as it remains constant.

Benefits Unlike CC and VAR, CRQA is a nonlinear analysis. Complex interactions of coupled systems can produce nonlinear dynamics. While VAR and CC would overlook these patterns, CRQA can provide multiple windows into these nonlinearities through various metrics. Because the linear case is a subset of the nonlinear case (Carello & Moreno, 2005), CRQA is capable of capturing both linear *and* nonlinear patterns in co-evolution. Additionally, there are also many extensions of the analysis, such as windowed and joint recurrence, all of which can be used for categorical or continuous data.

Methods

Corpus

The goal of these analyses was to analyze the co-evolution of online social cohesion (via Twitter data) and real-world action (via event data). We use a corpus originally presented in Chiovaro and colleagues (*under review*), where we investigate if the frequency of real-world events of varying intensities (e.g., positive events such as treaties and negative events such as bombings) coincides with the real-time linguistic cohesion on Twitter. Although a full accounting of the corpus is outside of the scope of the current work, we present a brief review of these components below.

Social Cohesion The first set of data used was a corpus of 3,443,742 English-language Syrian tweets dated between March 31–June 15, 2012, that used the keywords “syria,” “syrian,” “damascus,” “homs,” “al-assad,” and “sunni.” The corpus contained 63.2% original tweets, 33.5% retweets, and 3.4% replies. The corpus included an average of 44,723.92 tweets per day and a total of 574,104 unique users. The scripts for collecting and compressing data can be found at https://github.com/chbrown/twilight.

To quantify the social cohesion within the tweet corpus, we sorted the tweets in chronological order, grouped them into successive groups of five, and then used a process conceptually similar to term frequency-inverse term document frequency (Chiovaro et al., *under review*). Essentially, the social cohesion metric quantified two tweets as more similar if the two tweets shared more words that occurred fewer times in the corpus overall. This method allowed us to measure how similar tweets were in their language over time. We then averaged the social cohesion values within the groups of five to get a group average, and then averaged across the groups for each day, yielding a daily social cohesion metric. To make the social cohesion data comparable to the second dataset, the social cohesion time series was converted into deciles.

Events The second set of data were event data obtained from the 2012 Integrated Crisis Early Warning System (ICEWS; Boschee et al., 2015). The dataset was filtered to remove incomplete and incorrectly formatted data. A filter was then applied to include only events that were directed at Syria (i.e., Syria-target events; n = 6,300). We then counted all Syria-target events for each day, yielding a daily event count. To make the event data comparable to the social cohesion data, the event time series was also converted into deciles.

Analyses

VAR, CC, and categorical CRQA were all conducted using the deciled social cohesion and event time series. For VAR and CC, ADF tests identified social cohesion and event time series as nonstationary; we took the first difference of each time series, resulting in stationary time series. We used the original nonstationary time series for CRQA, as it is capable of handling such data. All code for the analyses can be found at https://www.github.com/mchiovaro/time-series-analyses.

Results and Discussion

Vector Autoregression (VAR)

SC suggested an optimal lag of 1, while AIC suggested an optimal lag of 6. We conducted both models (Table 1). The moduli of the roots indicated that both models were stable. Using SC criteria, neither social cohesion nor event time series Granger caused one another. Under AIC, social cohesion did not Granger cause the frequency of events, but the daily event counts did Granger cause the level of social cohesion on Twitter. (As noted previously, alternatives to Granger causality include IRFs and FEVD, but these are outside the scope of the current work.)

	F	p-value
Cohesion (SC)	0.0398	0.8422
Events (SC)	2.51	0.1153
Cohesion (AIC)	1.222	0.3
Events (AIC)	2.549	0.02352

Table 1. Granger causality results for social cohesion and count of events using two information criteria, SC and AIC.

Cross-Correlation (CC)

Cross correlation analysis of social cohesion and frequency of real-world events revealed significant correlations for -2 and -4 lags (Fig. 1). In other words, the changes in social cohesion levels on Twitter led changes in the number of events happening on the ground by 2 and 4 days. This may suggest that individuals used Twitter to rally and facilitate plans for action; alternatively, the Syrian government may have responded to increased Twitter activity with additional actions. Based on the current data, these possibilities cannot be disentangled. Further analyses involving the specific parties who initiated these events and more detailed content analyses are needed to answer these questions.

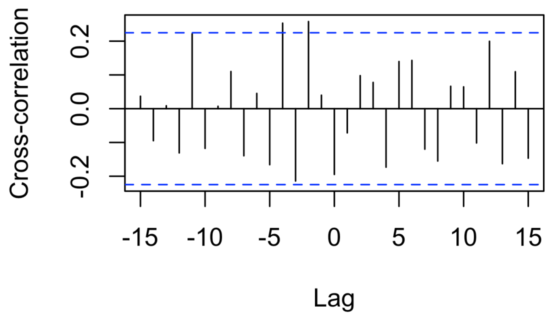

Figure 1. Cross-correlation plot. Horizontal dashed blue lines indicate confidence intervals (otherwise known as *conventional cross-correlation limits*, ccsls), but for limitations of using ccsls, see Dean & Dunsmuir (2016).

Cross-Recurrence Quantification Analysis (CRQA)

The CRP (Fig. 2) shows many shared trajectories between the time series and a shift in texture about the 32nd time point, where there appears to be an increase in the number of lines.

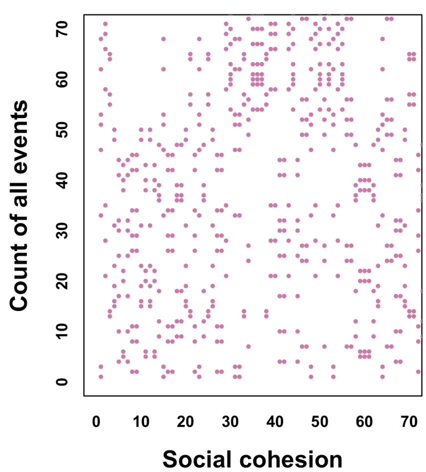

Figure 2. Recurrence plot for Twitter social cohesion and daily count of events. Points represent shared states of the two time series (in deciles).

CRQA metrics were tested for significance against a distribution of 1,000 permuted time series, with a customary alpha of .05. In other words, significant metrics were higher than would be expected by chance—in this case, observed in our permuted time series in no more than 50 simulations. We focused here only on DET, NRLINE, and maxL (Table 2).

The social cohesion and event data often got "stuck" together in the same relative states across time (DET). In addition, they entered these shared states more frequently than chance (NRLINE). Together, these results suggest that the two systems are coupled in their behavior. The trend toward significance for maxL suggests indicates that it was approaching a stable relationship, though it did not reach statistical significance.

metric	value	p	sig.
DET	25.714	0.000	***
NRLINE	71.000	0.000	***
maxL	4.000	0.051	.
L	2.155	0.157	
ENTR	0.452	0.115	
rENTR	0.412	0.410	
LAM	39.496	0.000	***
TT	2.136	0.276	

Table 2. CRQA results for social cohesion and event count. RR is not included due to the nature of approximate permutation tests (Chiovaro et al., *under review*).

The DRP (Fig. 3) showed no reliable leader between the two time series, as shown by the jagged line through the majority of lags investigated. The peaks seen here indicate an increased amount of recurrence at particular time lags, similar to how CC indicated stronger correlation strength at different time lags. For example, social cohesion led at 7, 11, and 15 days, with a large spike at the 15-day lag, and the count of all events led at 5 and 9 days. Comparing CRQA and CC results, we saw converging evidence of spikes in social cohesion leading real-world events by 2 and 4 days, although the patterns for CRQA were more complex than for CC.

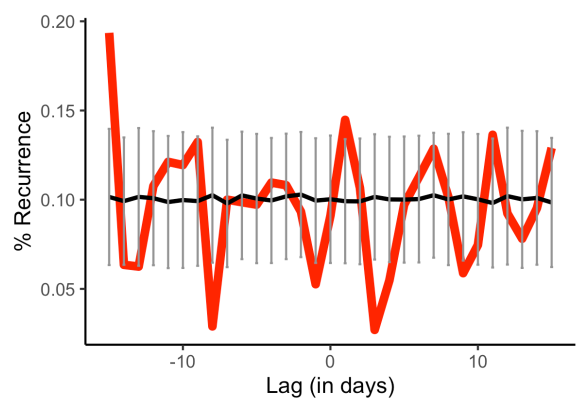

Figure 3. Diagonal recurrence profile for Twitter social cohesion and daily count of events. The red line is the DRP for the actual data. The black line is the DRP for the mean of the 1,000 permuted time series (bars: ±1 SD).

Comparing Time Series Analyses

Within the context of a growing interest in time series analyses in psychological and cognitive science, we here presented a direct comparison of three different methods on the same time series. We did so to highlight the importance of the underlying theoretical questions and structures of the available data when deciding among these methods.

When to Choose Vector Autoregression VAR can answer how two or more variables covary over time, allowing for investigations of multivariate time series. Uniquely among the analyses considered here, VAR can uncover the causal relationships among time series. However, it assumes stationarity (which can be addressed by altering the original time series) and requires estimation of important parameters. It also is a symmetric model, requiring the same number of lags to be used across all equations. Given the dynamic nature of coupled systems subjected to time series analysis, it may be more beneficial to use the more flexible method of asymmetric vector autoregression (AVAR; Keating, 2000). The parameter estimates in AVAR also typically have smaller standard errors than those resulting from VAR. VAR may be most appealing when investigating causality between two (or more) time series that have linear and stationary dynamics.

When to Choose Cross-Correlation CC can answer how only two variables covary over time, but its simplicity—requiring no parameter estimation—is its strength. Unlike VAR, CC cannot determine causal structure, but it can identify leader-follower patterns within specified levels of temporal proximity. Like VAR, CC is also a stationary, linear analysis and cannot capture nonlinearities. CC may be most useful when examining similarities and leader/follower patterns between two linear, stationary time series.

When to Choose Categorical CRQA Categorical CRQA's strength lies in its ability to uncover even *nonlinear* dynamics of two time series. Without requiring parameter estimation, it can help researchers answer questions about the structure, stability, and patterning of the interaction of two variables over time. Its nine metrics provide many angles for analyzing the relationship between the two variables, and the DRP can capture leading-following behavior. Unfortunately, CRQA cannot answer questions about causality. Categorical CRQA may be most beneficial when quantifying linear and nonlinear patterns in the co-evolution of two time series.

Examining Converging and Diverging Results Our analyses of the same data using these three methods uncovered some similar and some dissimilar patterns. VAR suggested that there was either no relationship or only slight Granger causality of social cohesion by the frequency of real-world events across the entire time series. CC suggested the opposite (though non-causal) relationship, with the social cohesion time series appearing to lead the changes in the number of daily events at a shorter timescale (±15 days). Finally, CRQA identified the two systems as coupled but without a clear leader-follower relationship, instead showing a back-and-forth trade-off within the same time window.

One potential contributor to the differences in these patterns of results may be the difference between leading/following and causality. In other words, although both CC and CRQA identified leading/following patterns, these patterns may not have been truly (Granger) causal in nature. The pairing of causal (VAR) and non-causal (CRQA and CC) methods may provide meaningful information about two levels of the system, capturing both fluctuations in the time series and deeper causal relations.

Another potential contributor to the differences in leading/following patterns may be due to the fact that VAR assumes that causal patterns in each time series occur at the same number of lags. However, it would not be surprising if the two time series—one being a measure of language on social media and the other being a count of real-world events—required different lags. Thus, causality may be more appropriately captured with AVAR.

Finally, differences between the linear (VAR and CC) and nonlinear (CRQA) analyses may have been affected by the data in question, as the data were transformed from analyses of *states* to analyses of *relative changes in state* to satisfy the stationarity assumption of the linear analyses. This attests to the importance of considering necessary data transformations when choosing analyses. Given the nonstationarity of some of the time series, we took the first difference and used these transformed time series for VAR and CC, but we used the untransformed (nonstationary) time series for CRQA. Differencing could have overshadowed dynamic changes in the observations that were captured by CRQA, which does not require stationarity. Thus we suggest that researchers using nonstationary time series consider the impacts of differencing on their data, and potentially utilize only analyses that assume stationarity or only those that do not require it.

Conclusion and Future Directions

Our case study suggests that the ability of a specific time series analysis to identify time-lagged relations among time series is highly dependent upon the scale of the lags and, perhaps, nonlinearities in the system. However, further investigation is necessary to understand the diverging results identified across these three types of analyses. Future comparisons should focus on analyses of simulated data, in order to analyze each method's ability to capture different patterns of psychological and behavioral data. Such simulation studies may inform a framework for selecting among these and other time series methods.

References

Brick, T. R., Gray, A. L., & Staples, A. D. (2018). Recurrence quantification for the analysis of coupled processes in aging. *The Journals of Gerontology: Series B*, *73*(1), 134-147.

Boschee, E., Lautenschlager, J., O'Brien, S., Shellman, S., Starz, J., & Ward, M. (2015). *ICEWS Coded Event Data* (Version V27) [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/28075

Carello, C., & Moreno, M. (2005). Why nonlinear methods. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 1-25). National Science Foundation.

Chiovaro, M., Windsor, L. C., Windsor, A., & Paxton, A. (*under review*). Online social cohesion reflects real-world group action in Syria during the Arab Spring.

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An r package. *Frontiers in Psychology*, *5*, 510.

Dean, R. T., & Dunsmuir, W. T. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior research methods*, *48*(2), 783-802.

Freeman, J. R. (1983). Granger causality and the times series analysis of political relationships. *American Journal of Political Science*, 327-358.

Freeman, J. R., Williams, J. T., & Lin, T. M. (1989). Vector autoregression and the study of politics. *American Journal of Political Science*, 842-877.

Gates, K. M., & Liu, S. (2016). Methods for quantifying patterns of dynamic interactions in dyads. *Assessment*, *23*(4), 459-471.

Keating, J. W. (2000). Macroeconomic modeling with asymmetric vector autoregressions. *Journal of macroeconomics*, *22*(1), 1-28.

Love, G. J., & Windsor, L. C. (2018). Populism and popular support: vertical accountability, exogenous events, and leader discourse in Venezuela. *Political Research Quarterly*, *71*(3), 532-545.

Main, A., Paxton, A., & Dale, R. (2016). An Exploratory Analysis of Emotion Dynamics Between Mothers and Adolescents During Conflict Discussions. Emotion. Advance online publication. http://dx.doi.org/10.1037/emo0000180

Pellecchia, G. L., Shockley, K., & Turvey, M. T. (2005). Concurrent cognitive task modulates coordination dynamics. *Cognitive science*, *29*(4), 531-557.

Pfaff, B. (2007). Package vars. http://cran.r-project.org/web/packages/vars/index.html.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*(6), 1045-1060.

Riley, M. A., & Van Orden, G. C. (Eds.). (2005). *Tutorials in contemporary nonlinear methods for the behavioral sciences*. National Science Foundation.

Schoenherr, D., Paulick, J., Worrack, S., Strauss, B. M., Rubel, J. A., Schwartz, B., ... & Altmann, U. (2019). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior research methods*, *51*(1), 361-383.

Shumway, R. H., & Stoffer, D. S. (2006). Time series analysis and its applications (2nd ed.). New York, NY: Springer.

Sims, C. (1980). Macroeconomics and reality. Econometrica, 48, 1–48. doi: 10.2307/ 1912017

Tolston, M. T., Funke, G. J., & Shockley, K. (2020). Comparison of Cross-Correlation and Joint-Recurrence Quantification Analysis Based Methods for Estimating Coupling Strength in Non-linear Systems. *Frontiers in Applied Mathematics and Statistics*, *6*, 1.

Trulla, L. L., Giuliani, A., Zbilut, J. P., & Webber Jr, C. L. (1996). Recurrence quantification analysis of the logistic equation with transients. *Physics Letters A*, *223*(4), 255-260.

Zbilut, J. P., Giuliani, A., & Webber Jr, C. L. (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, *246*(1-2), 122–128