Dizzying Dyadics

Differences in Assessing Dyadic Relationships with Automatic and Manual Event Classification

Tyler Ashoff
April 15, 2020

Abstract

This study seeks to understand the level of agreement between two human and machine coded event databases, the COPDAB dataset and Cline Center Historical Phoenix dataset respectively. By investigating the dyadic relationips of six countries between 1955 and 1978 and varying the level of aggregation, we show that there are conditions where this agreement can be strong. However, this is not always the case. In many instances especially at lower levels of aggregation, versee this agreement weaken drastically.

Introduction

Dyadic relationships are important to understand in international relations. Tracking how countries respond to each other and exogenous shocks is critical to understanding what drives these interactions. To make this analysis possible there needs to be a consistent and reliable source for measuring these relationships. Event databases seek to fill this role by classifying events and scoring them based on the conflictual or cooperative nature of each event. Historically these events were human coded and more recently they are automated by performing text analysis on news stories.

How accurately each of these methods classify events is an ongoing topic of research. As is defining the scales on which the intensity of the events is measured. This study though, does not seek to compare these databases to ground truth. There we examine how the agreement between human coded and automatically coded datasets in measuring dyadic relationships between 6 countries between 1955 and 1978. Conducting our analysis at varying time scales we can determine the effect of granularity on this agreement. Seesess this agreement we calculate the Pearson and Spearman correlations and the dynamic time warping distance to capture similarity between signals despite possible misalignment.

Ideally, this research can be used to make more informed decisions about event database usage and understand the conditions where agreement between human and automatically coded datasets largely agree.

Data

Two conflict datasets were used in this study, one manually coded and another automatically coded. These records contain more detailed information about the events than we will use in this study. We are concerned with only four fields: Actor, Target, Score and Date. The Actor and Target variables for assessing membership in dyadic relationships, Score for assessing the intensity of event, and Date to track the relationship over time.

The manually coded dataset is the COPDAB dataset developed by Edward Azar covering the years 1948 to 1978 (Azar 1948). Coding of these events and how coders were trained is discussed in greater detail in the original paper. Each event is scored on a 1 to 15 scale, with 1 representing a major governmental program and 15 representing the highest level of structural violence. These values were then weighted using the international weighting scale (Azar and Havener 1976). These weightings with two modifications are as shown in Table 1. These modifications from the original scale are 9-15 to negative values to indicate conflict and 8 equals 0 instead of 1 to better represent neutrality.

Table 1:	Augmented	International	Event	Weighting
Table 1.	riusiiiou	TITUCI II CUI CII CUI	L V CIII	* * C151101115

Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Weight	92	47	31	27	14	10	6	0	-6	-16	-29	-44	-50	-65	-102

The automatically coded dataset is the Cline Center Historical Phoenix Event Data NYT 1945-2005 (Althous et al. 2019). These data are collected using the PETRARCH-2 software to extract event information from news stories. The scoring for these events is based on the Golstein scale. A -10-8.3 scale with -10 representing a military attack, 0 representing a neutral event, and 8.3 representing military assistance (Golstein 1992).

The weighting of each of each scale was decided by a panel of judges asked to weight the relative intensity of events. While not discussed at length here, more information may be found in the original papers by Golstein and Azar (Golstein 1992) (Azar 1948).

Both of these datasets were chosen because they are well established and have a long overlapping period. While the datasets overlap from 1948-1978, we restrict this time frame slightly to 1955-1978. This period is chosen because post-war reconstruction is generally considered complete in 1955 and it is the first full year the German Democratic Republic was a sovereign country as discussed in Pollins' "Does Trade Still Follow the Flag?" (Pollins 1989).

Manipulations of the data

The score data are scaled and missing data is mean imputed. While the techniques discussed in subsection Statistical Techniques are invariant to scale, this manipulation made visualization of the data much cleaner. It is true that the magnitude of the dynamic time warping distance is affected by scale, however the rank order of these distances is not. Since the rank order is

of primary importance in this inquiry we are not concerned with the change in magnitude associated with this transformation.

While complete data was ensured in the yearly assessment of dyadic relationships, this was not the case for other time scales. For the Monthly and Weekly analyses, and their associated simple moving averages, there were instances of incomplete data. This problem was especially pronounced as the level of aggregation decreased. The techniques used require the same number of observations in each time series, so mean imputation was used to fill missing values. While there are significant potential issues with mean imputation, more complex methods such as MICE are left for future analysis.

Methods

Dyad Selection

This investigation focuses on six states: USA, USSR, China, Canada, East-Germany, and West-Germany. The dyadic relationships between these countries are chosen based on data availability. If both datasets have at least one observation per year for a given dyad (before imputation), this dyad is included in the analysis. It is important to clarify how an observation was defined. An observation was considered part of the dyad, say USA-USSR, only when the actor was the USA and the target was the USSR or vice versa. Domestic events when the target and actor are the same were not included. The selected dyads are as seen in Table 2.

Table 2: Dyads included in analysis

Dyads	
USA - USSR USA - China USA - West-Germany USA - Canada USSR - China	
USSR - West-Germany USSR - Canada East-Germany - West-Germany	7

These dyadic relationships are analyzed at varying levels of granularity. Each dyad in Table 2 is assessed at the levels described in Table 3.

The simple moving averages of described in Table 3 are an average of the current time step and the n-1 time steps preceding for window size n. For example the 2 week average is an average of the current week and the previous week, the 6 month average is an average of the current month and the 5 months preceding. The initial observations where the window could not be calculated fully use only the data available. For example while calculating the 6 month averages, the window for March 1955 only includes January, February, and March.

Table 3: Time frames included in dyadic analysis

Time.frame
Year
6 Month Simple Moving Average
3 Month Simple Moving Average
2 Month Simple Moving Average
Month
3 Week Simple Moving Average 2 Week Simple Moving Average Week

Statistical Techniques

To assess the similarity of the time series, three metrics were employed: Pearson correlation, Spearman correlation, and Dynamic Time Warping (DTW). The traditional Pearson correlation was of primary interest, however Spearman correlation and DTW are included to investigate non-linearity and data shifts.

Pearson correlation is defined as follows:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y}$$

where cov() is the covariance and σ is the standard deviation of the respective datasets.

Spearman correlation is similar to the Pearson correlation, but instead of using the data directly, it calculates the Pearson correlation of the ranked datasets.

DTW is especially useful for comparing signals that may vary in time and speed (Anon n.d.). Which works well for our purposes because it is not guaranteed that the events are aligned, meaning the same event might be reported later in different datasets. In contrast with the Pearson correlation, this property gives us another tool for assessing similarity in the face of non-linearity. While the algorithm is not detailed here, the distance is essentially an cumulative squared distance between points with optimized pairings. Importantly, this cumulative distance is dependent on the number of points in the time series. To make comparison across varying granularities, the DTW distance presented is a normalized value.

The Pearson and Spearman correlations fall in an easily interpretable range [-1,1]. With 1 indicating total linear correlation, 0 no linear correlation, and -1 total negative correlation. DTW distances do not have an analogous scale, instead it has the range, $[0,\infty)$ with no universal thresholds. Instead we must compare the distances to each other as we look for similarity with smaller values indicating increasing agreement.

Results

Presented in this section are a subset of the full analysis. While analysis was performed on every time frame described in Table 3, only the analysis of yearly, monthly, and 6 month average data accompanied by selected plots will be discussed. These time frames are illustrative of the over all findings and will suffice in the place of an exhaustive discussion. A full accounting of the correlations and the associated time series plots can be found in Appendix A and B respectively.

Yearly Dyads

The yearly dyadic relationships are the primary interest of this study. In Table 7 we see the Pearson and Spearman correlations with their associated p-values and the normalized dynamic time warp distance. While the range of values is fairly large, there is good correlation between the datasets for a few dyads. The USA-USSR dyad shows high correlation, and the USSR-China, USSR-Canada, and USA-China dyads have good correlations well.

Table 4: Yearly Correlations

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.813	1.41e-06	0.866	4.55e-08	0.438
USA-China	0.438	3.21e-02	0.342	1.02e-01	0.642
USA-West-Germany	0.142	5.09e-01	0.129	5.49e-01	0.647
USA-Canada	0.190	3.74e-01	0.338	1.06e-01	0.790
USSR-China	0.681	2.49e-04	0.705	1.22e-04	0.492
USSR-West-Germany	0.406	4.91e-02	0.410	4.69e-02	0.585
USSR-Canada	0.591	2.38e-03	0.503	1.21e-02	0.603
East-Germany-West-Germany	0.069	7.50e-01	0.118	5.82e-01	0.684

Figures 1 and 2 are the time series of the dyads with the highest and lowest correlations respectively. In Figure 1, the time series of the USA-USSR dyad, we see the high correlation clearly. The COPDAB and Phoenix datasets are largely in agreement for these two superpowers in the heat of the Cold War. Being able to track this highly visible and contentious relationship is a good sign that under the right conditions, these datasets capture similar dynamics.

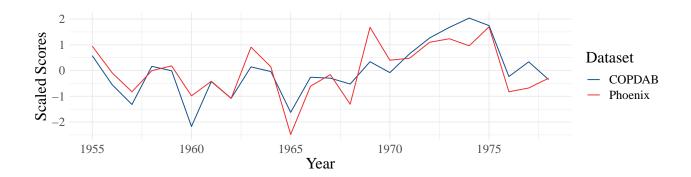


Figure 1: Yearly values for USA and USSR

Figure 2 shows the time series of the East and West Germany dyad which has a correlation near 0. However, when examining the graph it appears that the datasets largely agree. Interestingly we can see that the relationship follows a similar trend as the USA-USSR relationship. This should be expected since the two states were highly aligned with the two superpowers.

Figure 2 highlights a short coming of the correlation coefficients. Some points are shifted between the two datasets, yet the measured relationship exhibited by the time series appear similar. In combination with Figure 3 we see motivation for the use of dynamic time warping as a useful metric

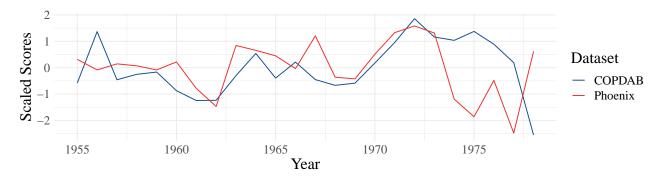


Figure 2: Yearly values for East and West Germany

Figure 3 shows the time series of the USA-Canada dyad. The datasets appear to disagree more for this dyad than the German dyad, yet we see a much higher correlation. The DTW distance reflects this intuition, the value is smaller for the German dyad than the USA-Canada dyad, indicating greater similarity. While this certainly does not prove that DTW is a better measure in all situations, it does highlight potential short comings of analysis of correlations alone.

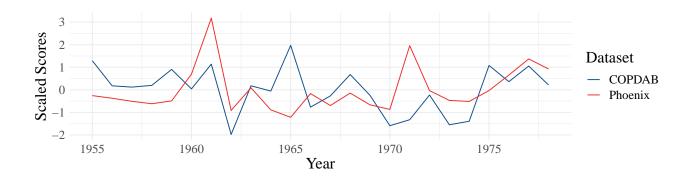


Figure 3: Yearly value for USA and Canada

Monthly Dyads

While levels of aggregation below one month were examined, the results become increasingly less reliable. At the Monthly level of aggregation we begin to see the adverse effects of missing data for dyads with lower data density.

Table 8 shows a similar ordering of correlation coefficients with the larger dyads recording higher correlations as before. Figure 4 shows that while there is significantly more noise in the signal than at yearly aggregation, the signal is still visible for the two superpowers. It is in Figures 5 and 6 that we see the effects of missing values. Due to the mean imputation approach, missing values will be filled with 0, the mean for scaled data. In these figures we can see a flat lining effect as missing values become more prevalent. Due to this, more advanced imputation techniques are required for reliable results across all dyads.

These examples also serve to highlight the short comings of dynamic time warping. At this monthly aggregation level, it appears that the USSR-Canada similarity far out paces the yearly values. This again is a relic of missing values. Because of the flat lining effect the curves this value does not reflect an accurately measured relationship. This stresses again how over reliance on a single metric may lead to erroneous conclusions.

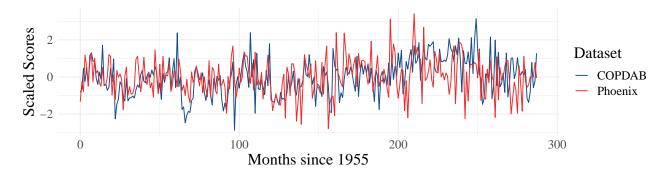


Figure 4: Monthly values for USA and USSR

Table 5: Monthly Correlations

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.589
USA-China	0.182	1.90e-03	0.210	3.22e-04	0.587
USA-West-Germany	0.064	2.80e-01	0.023	6.94e-01	0.572
USA-Canada	0.001	9.91e-01	0.016	7.88e-01	0.561
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.494
USSR-West-Germany	0.139	1.80e-02	0.129	2.84e-02	0.493
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.261
East-Germany-West-Germany	-0.011	8.54 e-01	0.033	5.76e-01	0.430

Dataset

— COPDAB
— Phoenix

Months since 1955

Figure 5: Monthly values for East and West Germany

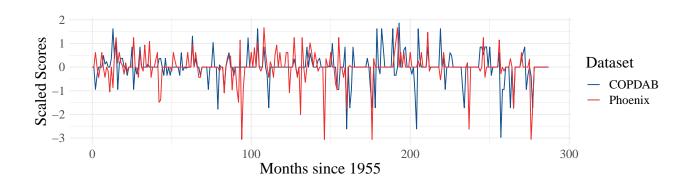


Figure 6: Monthly value for USSR and Canada

Six Month Simple Moving Average Dyads

While monthly aggregation is limited by missing values, it appears that a six month simple moving average effectively mitigates the flat line effect, while increasing granularity compared

to yearly aggregation. In Figure 7, a similar relationship as the one seen in yearly aggregation is clear. However the signal becomes much noisier with this increased granularity. And while the flat line effect is less pronounced, it is still present as we can see from Figure 9

The correlation has increased compared to the monthly aggregation, but more interestingly, we see a huge improvement in DTW distance across the board. From inspection of the plots it appears that the two datasets are capturing similar dyadic relationships. The discrepancy between correlation coefficients and dynamic time warping may indicate misalignment. While the datasets capture the relationships, they may be shifted or delayed compared to one another. Another explanation for this increased similarity may be a damped signal. The values from the six month aggregation are confined primarily to a single standard deviation from the mean. Whereas the monthly and yearly values exceeded two or three standard deviations. The improved DTW distance may simply be a reflection of this tighter score distribution. Further analysis is needed to determine which effect is dominant, which is not pursued here.

Table 6: Monthly Correlations - 6 Month Average

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.221
USA-China	0.361	2.56e-10	0.339	3.69e-09	0.291
USA-West-Germany	0.064	2.80e-01	0.023	6.94e-01	0.197
USA-Canada	0.126	3.23e-02	0.164	5.35e-03	0.171
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.202
USSR-West-Germany	0.228	9.29e-05	0.236	5.07e-05	0.210
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.123
East-Germany-West-Germany	0.192	1.08e-03	0.118	4.49e-02	0.152

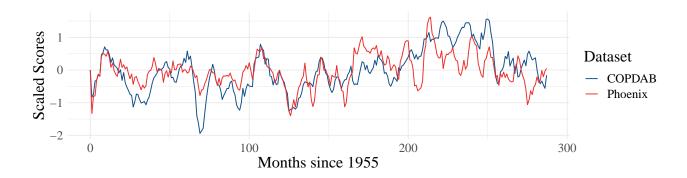


Figure 7: Six month average value for USA and USSR

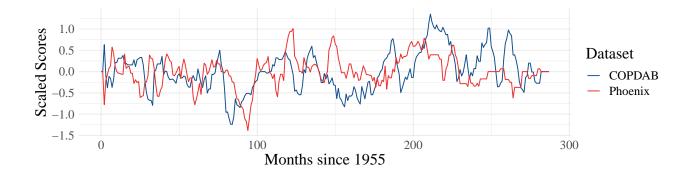


Figure 8: Six month average value for East and West Germany

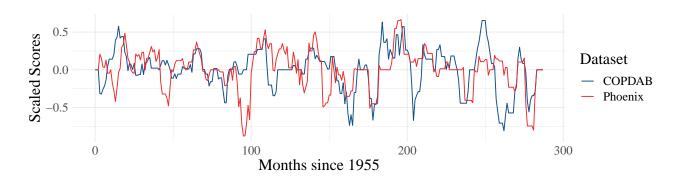


Figure 9: Six month average value for USSR and Canada

Conclusion

The yearly dyadic relationships showed that there is significant agreement between the COPDAB and Phoenix datasets. While efforts to investigate smaller time frames were inconclusive, the moving average approach shows promise for achieving increased signal clarity and mitigating the effect of missing data. In all cases though, it is apparent that the datasets have higher agreement when more data is available. Recall from the Dyad Selection Section that our dyads were selected with the criteria that they must have at least one observation per year. This is a low bar to clear but still some dyads were not included in analysis. It is not a surprise that all but one dyad includes either the USA or USSR during the cold war period. We have seen that with highly observed dyads like the USA-USSR dyad, high levels of agreement are possible. However, given the variety of results from the dyads clearing our low bar, that these datasets agree in all circumstances is not guaranteed.

te to Prof. Greico

I am sorry about blowing the 20 page limit out of the water with my appendices. I kept the body of the paper to 10 pages as you suggested for potential publication. But I wanted to

include the other figures in case you feel they might be good additions or substitutions for the figures included.

References

- Althaus, Scott, Joseph Bajjalieh, John F. Carter, Buddy Peyton, and Dan A. Shalmon. 2019. "Cline Center Historical Phoenix Event Data." *Cline Center for Advanced Social Research* v.1.2.0(University of Illinois Urbana-Champaign).
- Anon. n.d. "Dynamic Time Warping with Time Series Shachia Kyaagba Medium."
- Azar, Edward E. 1948. "Conflict and Peace Data Bank (COPDAB)."
- Azar, Edward E., and Thomas N. Havener. 1976. "Discontinuities in the Symbolic Environment: A Problem in Scaling." *INTERNATIONAL INTERACTIONS* Vol. 2:231–46.
- Golstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *The Journal of Conflict Resolution* 36(2):369–85.
- Pollins, Brian M. 1989. "Does Trade Still Follow the Flag?" The American Political Science Review 83(2):465–80.

Appendix

A: Correlations

Table 7: Yearly Correlations

	Pearson		Spe	arman	DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.813	1.41e-06	0.866	4.55e-08	0.438
USA-China	0.438	3.21e-02	0.342	1.02e-01	0.642
USA-West-Germany	0.142	5.09e-01	0.129	5.49e-01	0.647
USA-Canada	0.190	3.74e-01	0.338	1.06e-01	0.790
USSR-China	0.681	2.49e-04	0.705	1.22e-04	0.492
USSR-West-Germany	0.406	4.91e-02	0.410	4.69e-02	0.585
USSR-Canada	0.591	2.38e-03	0.503	1.21e-02	0.603
East-Germany-West-Germany	0.069	7.50e-01	0.118	5.82e-01	0.684

Table 8: Monthly Correlations

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.589
USA-China	0.182	1.90e-03	0.210	3.22e-04	0.587
USA-West-Germany	0.064	2.80e-01	0.023	6.94 e-01	0.572
USA-Canada	0.001	9.91e-01	0.016	7.88e-01	0.561
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.494
USSR-West-Germany	0.139	1.80e-02	0.129	2.84e-02	0.493
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.261
East-Germany-West-Germany	-0.011	8.54 e-01	0.033	5.76 e-01	0.430

Table 9: Monthly Correlations - 2 Month Average

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.391
USA-China	0.226	1.07e-04	0.242	3.36e-05	0.447
USA-West-Germany	0.064	2.80e-01	0.023	6.94e-01	0.392
USA-Canada	-0.004	9.47e-01	0.022	7.11e-01	0.375
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.385
USSR-West-Germany	0.209	3.47e-04	0.237	4.84e-05	0.379
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.204
East-Germany-West-Germany	0.045	4.44e-01	0.057	3.39e-01	0.326

Table 10: Monthly Correlations - 3 Month Average

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.316
USA-China	0.280	1.33e-06	0.277	1.74e-06	0.397
USA-West-Germany	0.064	2.80e-01	0.023	6.94 e-01	0.289
USA-Canada	0.042	4.83e-01	0.045	4.45e-01	0.283
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.306
USSR-West-Germany	0.232	6.96e-05	0.254	1.24e-05	0.316
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.175
East-Germany-West-Germany	0.068	2.52e-01	0.070	2.33e-01	0.263

Table 11: Monthly Correlations - 6 Month Average

	Pe	Pearson		arman	DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.249	1.87e-05	0.272	2.72e-06	0.221
USA-China	0.361	2.56e-10	0.339	3.69e-09	0.291
USA-West-Germany	0.064	2.80e-01	0.023	6.94 e-01	0.197
USA-Canada	0.126	3.23e-02	0.164	5.35e-03	0.171
USSR-China	0.237	4.84e-05	0.269	3.51e-06	0.202
USSR-West-Germany	0.228	9.29e-05	0.236	5.07e-05	0.210
USSR-Canada	0.223	1.33e-04	0.185	1.60e-03	0.123
East-Germany-West-Germany	0.192	1.08e-03	0.118	4.49e-02	0.152

Table 12: Weekly Correlations

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.072	1.13e-02	0.106	1.73e-04	0.5492
USA-China	0.123	1.17e-05	0.146	2.28e-07	0.3453
USA-West-Germany	0.027	3.38e-01	0.017	5.42e-01	0.3122
USA-Canada	0.016	5.68e-01	-0.010	7.19e-01	0.3552
USSR-China	0.144	3.24e-07	0.146	1.99e-07	0.3009
USSR-West-Germany	0.048	9.07e-02	0.019	5.01e-01	0.2570
USSR-Canada	0.090	1.49e-03	0.069	1.46e-02	0.0744
East-Germany-West-Germany	0.033	2.42e-01	0.028	3.14e-01	0.1875

Table 13: Weekly Correlations - 2 Week Average

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.072	1.13e-02	0.106	1.73e-04	0.3757
USA-China	0.130	3.96e-06	0.160	1.21e-08	0.2854
USA-West-Germany	0.027	3.38e-01	0.017	5.42e-01	0.2642
USA-Canada	0.012	6.64 e-01	-0.022	4.31e-01	0.2799
USSR-China	0.144	3.24e-07	0.146	1.99e-07	0.2493
USSR-West-Germany	0.105	1.94e-04	0.079	5.39e-03	0.2147
USSR-Canada	0.090	1.49e-03	0.069	1.46e-02	0.0715
East-Germany-West-Germany	0.056	4.91e-02	0.050	7.66e-02	0.1655

Table 14: Weekly Correlations - 3 Week Average

	Pearson		Spearman		DTW
Dyad	rho	p-value	rho	p-value	normalized
USA-USSR	0.072	1.13e-02	0.106	1.73e-04	0.2932
USA-China	0.145	2.54e-07	0.158	2.08e-08	0.2425
USA-West-Germany	0.027	3.38e-01	0.017	5.42e-01	0.2150
USA-Canada	-0.008	7.90e-01	-0.026	3.52e-01	0.2381
USSR-China	0.144	3.24e-07	0.146	1.99e-07	0.2142
USSR-West-Germany	0.134	2.08e-06	0.113	5.77e-05	0.1789
USSR-Canada	0.090	1.49e-03	0.069	1.46e-02	0.0696
East-Germany-West-Germany	0.079	5.30e-03	0.057	4.29e-02	0.1467

B: Plots

Yearly values for dyad: ['USA' 'USSR']



