

Predicting the Onset of Intra-State Conflict in Eurasia

BY

Kenton Derek Van, Jr.

Submitted to the graduate degree program in Global and International Studies and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

Chairperson Dr. Mariya Omelicheva

Dr. Eric Hanley

Mr. Richard Ground

Date Defended: 8/14/2012

The Thesis Committee for Kenton Derek Van, Jr.
certifies that this is the approved version of the following thesis:

Predicting the Onset of Intra-State Conflict in Eurasia

Chairperson Dr. Mariya Omelicheva

Date approved: 8/14/2012

ABSTRACT

While much of the contemporary quantitative research in intra-state conflicts focuses on explaining the causes of these conflicts using large-N studies, these models have limited forecasting power, i.e., they cannot be used in making predictions about whether or not an intra-state conflict will occur with a high degree of certainty. To that end this paper utilizes a forecasting model developed by D'Orazio, Yonamine and Schrodte and applies it for predicting the onset of intra-state conflict in Eurasia, which encompasses the former Soviet republics of Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan. The region was chosen due to a high incidence of intra-state conflict and proximity to the states analyzed by D'Orazio. Sample size was kept small due to manpower limitations. Event data was collected from news sources and a bivariate logit model was used to analyze patterns in the event data sequences in the 12 weeks preceding onsets of conflict and peace from 1990-2011. Although the results indicated that the model was not able to predict onsets of conflict with a high level of accuracy, this level of predicting power is attributed to the limitations of the original data rather than the model itself. The Eurasian model showed patterns similar to that of the original model, providing support for the observation that conflicts do not have a similar escalation pattern. Rather each conflict tends to be preceded by an event sequence, which is dissimilar both to peace sequences and other conflict onsets. In other words, the events leading up to intra-state conflicts tend to be unique to each conflict; there is no one pattern of events which leads to conflict.

ACKNOWLEDGEMENTS

I would like to extend a special thank you to Dr. Philip Schrodtt and Vito D'Orazio, upon whose work this paper is built. They provided vital insights, data and code and without their assistance this paper would not have been possible. I would also like to thank the members of my committee, Dr. Omelicheva, Dr. Hanley and Richard Ground for their guidance and support throughout this project.

ACRONYMS AND ABBREVIATIONS

CONIS- Conflict Information System

DARPA- Defense Advanced Research Projects Agency

EAV- Euclidean Absolute Value

Eurasian dataset- Dataset generated for this thesis

GTDS- Ground Truth Data Set

ICEWS- Integrated Crisis Early Warning System

KEDS- Kansas Event Data System

TABARI- Textual Analysis By Augmented Replacement Instructions

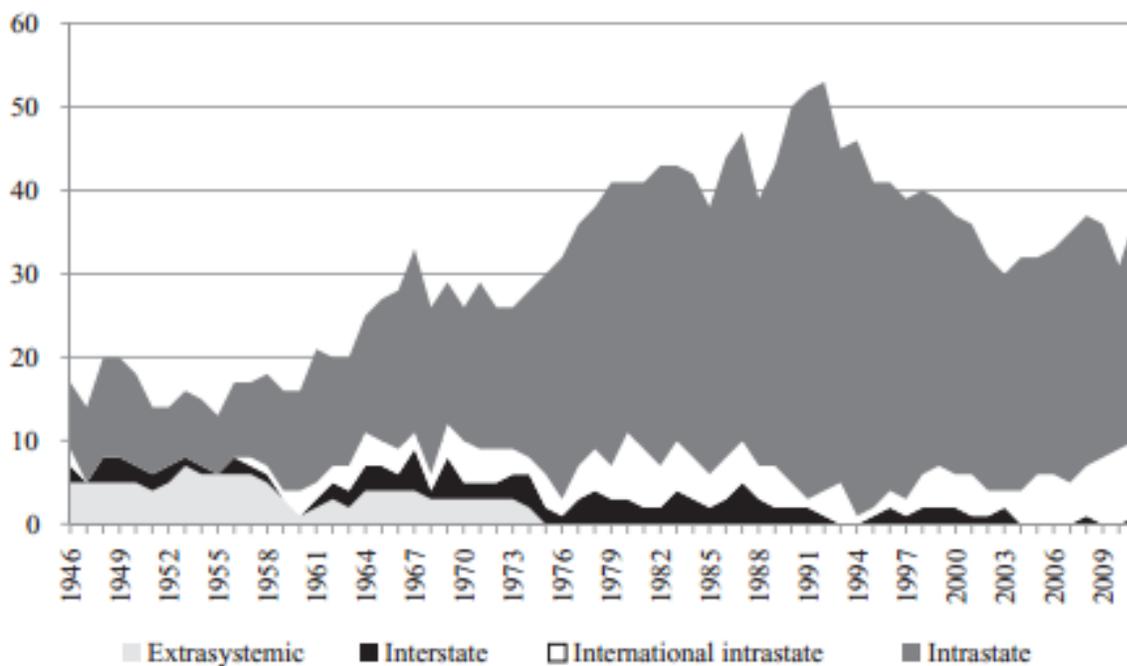
UCDP- Uppsala Conflict Data Program

VRA- Virtual Research Associates

INTRODUCTION

There are two broad categories of international conflict: inter-state conflicts and intra-state conflicts. Inter-state conflicts occur when the main actors are governments and can be fought across multiple countries; intra-state conflicts are conflicts between a government and a non-governmental party, with or without aid from outside parties.¹ Although these conflicts may draw in other regional or world state actors the conflict is primarily internal to a certain country. While the level and incidence of inter-state wars has decreased in the second part of the 20th century, intra-state conflicts rose appreciably after the late 1980s, until recently, attracting mounting interest and concern from both politicians and scholars.²

Figure 1³: Armed Conflicts, 1946-2011



Source: Themnér, Lotta and Wallensteen, Peter. "Armed Conflicts, 1946-2011." *Journal of Peace Research*. Vol. 49, No. 4. 568.

¹ "Definitions." *Uppsala Conflict Data Program*. Uppsala University. <http://www.pcr.uu.se/research/ucdp/definitions/>.

² Englehart, Neil and Kurzman, Charles. "Welcome to World Peace." *Social Forces*. Vol. 84, No. 4. Pg 1958-1959

³ Themnér, Lotta and Wallensteen, Peter. "Armed Conflicts, 1946-2011." *Journal of Peace Research*. Vol. 49, No. 4. 568.

This thesis focuses on intra-state conflict. Its goal is to contribute to the study of intra-state conflict through the development of a new regional database of event data on the onset of intrastate conflict in Eurasia and testing a new method of predicting conflict onset in the region using this event dataset. I present and test a model, to predict onsets of intra-state conflict in Eurasia 1 to 2 months into the future using the research methodology developed by Schrodtt, D’Orazio and Yonamine in “Predicting Intra-state Conflict Onset”⁴.

I begin with a justification of a need for new methodology that will enable scholars to forecast this type of conflict. Next, I provide an overview of the model and methods developed for predicting conflicts by D’Orazio, Yonamine and Schrodtt, including a discussion of event data and how it is well-suited for this line of research. I then introduce my research agenda and discuss the data collection, data set creation, and analysis. Finally I give the results of my analysis and compare those results to the existing literature.

Explanation vs. Prediction of Intra-State Conflict

Although there has been extensive research in the field of intra-state conflict, only a relatively small proportion of that has been conducted with the explicit goal of predicting conflict onsets. Qualitative research tends to employ small N studies which provide in depth understanding of specific conflicts but lack predictive power. Quantitative research, on the other hand, has tended to focus on large N studies, analyzing annual data using a variety of statistical techniques. A wide variety of explanatory variables are thrown together and analyzed; some variables turn out to be statistically significant, but most of the time they tend to be weak with

⁴ D’Orazio, Vito, Yonamine, James and Schrodtt, Philip. *Predicting Intra-state Conflict Onset- An Event Data Approach Using Euclidian and Levenshtein Distance Measures*. Pennsylvania State University. 2011. <http://eventdata.psu.edu/papers.dir/MPSA11.ICEWS.sequence2.1.pdf>.

wide confidence intervals. As a result, "...those with deep qualitative knowledge of the subject are rarely persuaded by conclusions from quantitative works."⁹

What these two methods have in common, at least in the current study of intra-state conflicts, is that both focus on *explanation* of the causes of intra-state conflict instead of *prediction*. In fact, Paul Collier, who recently conducted a large-scale statistical analysis of the factors affecting civil conflict, took pains to point out that his model should not be used for prediction but can show typical structural features of countries engaged in such conflicts.¹⁰

Indeed, Ward, Greenhill and Bakke show that Collier's model performs poorly at predicting civil war onsets and suggest that searching for statistically significant relationships is not the way for models to demonstrate their applicability to the real world.¹¹ They argue that scholars should improve quantitative models by employing out of sample testing, or in other words, prediction. Further, the authors note that even though researchers may caution that their results are not to be used for prediction, policy prescriptions are based on such academic work. Such work should be tested through using predictive heuristics in order to avoid erroneous conclusions by policy makers.¹²

In reality though, there should be no distinction between explanation and prediction. Schrodtt notes that as the social sciences are a *scientific* discipline, prediction and explanation are treated as the same and that the absence of prediction is "prescientific."¹³ According to Hempel and Oppenheim, "... an explanation of a particular event is not fully adequate unless its

⁹ King, Gary, Beck, Nathaniel, and Zeng, Lengche. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review*. Vol. 94, No. 1. March 2000. Pg. 21.

¹⁰ Collier, Paul. *The Bottom Billion*. Oxford: Oxford University Press. 2007, Pg. 19.

¹¹ Ward, Michael D., Greenhill, Brian D. and Bakke, Kristin M. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research*. Vol. 47 No. 4. Pg. 365.

¹² *Ibid.*, Pg. 364.

¹³ Schrodtt, Philip. "Seven Deadly Sins of Contemporary Quantitative Political Analysis." Presented at annual meeting of the American Political Science Association, 2010.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1661045###. Accessed 4/16/12. Pgs. 4-7.

explanans, if taken account of in time, could have served as a basis for predicting the event in question.”¹⁴ That is, an explanation is incomplete if it can’t be used for prediction. Many of the existing approaches cannot be used for prediction due to unsophisticated methodologies and data collection techniques. One type of data however, event data, has been used for prediction and is the focus of much of the contemporary intra-state conflict prediction research.

Overview of Event Data

This thesis uses event data to create a model predicting intra-state conflict. Event data are nominal or ordinal codes which record the descriptions of interactions between two actors in various news sources. They are an attempt to break down complex relationships reported in the open sources into simple parts which can then be analyzed.¹⁵ Data is gathered from stories produced by major news outlets like Reuters, The New York Times and Agence France Presse and generally follows the format:

<DATE> <ACTOR1><INTERACTION TYPE><ACTOR2>

Though studies may record much more information about each event, each will contain at minimum this core information.¹⁶

For example, a researcher might come across a news story that reads “*British and American war planes bombed Dresden last night, killing 25,000.*” The resulting codes would be:

<02131945><GBR><BOMB><GER>

<02131945><USA><BOMB><GER>

¹⁴ Hempel, Carl G. and Oppenheim, Paul. "Studies in the Logic of Explanation." *Philosophy of Science* 15,2. Pgs. 138-139.

¹⁵ Schrodtt, Philip. *Twenty Years of the Kansas Event Data System Project*. University of Kansas. Written 11/9/2006. Accessed 11/24/11. <http://eventdata.psu.edu/utilities.dir/KEDS.History.0611.pdf>. Pg. 1.

¹⁶ At least this has been my experience, though I should note that I have primarily looked at newer, machine coded events created by KEDS, TABARI and the VRA reader. Older data sets may have less information, though I doubt that their data sets don’t include these four basic fields. The World Event/Interaction Survey (WEIS), created in the 1960s-70s records five fields for each event: actor, target, date, action category and arena. <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/05211/detail#scope-of-study>

This would indicate that two events were reported. First, on February 13, 1945 Great Britain bombed Germany, and on the same date, the US also bombed Germany. Each event datum represents exactly one reported interaction where one actor does something to another actor. Note that “interaction” doesn’t necessarily mean that the two actors physically interact. Interactions can also be verbal, as in one actor denouncing the actions of another. Also, since one interaction can involve more than two actors, multiple events can be generated from a single interaction, as noted in the example above.

An event data set consists of a large number of such events, arranged chronologically. Researchers attempt to analyze the data in order to reveal relevant patterns in the actions or relationships between the actors. Ultimately the goal is to be able predict how actions by one actor will affect the decisions of another. Event data is well-suited for the purposes of predicting behavior for two main reasons. First, data can be aggregated in smaller intervals than traditional models. Interaction patterns can be analyzed on a daily basis, as opposed to variables like GDP which is calculated quarterly or annually. Larger aggregations can hide important details; smaller units of measurement allow for more precise analyses. The same holds true for the actor aggregation level. Event data can specify individual actors while traditional statistical analyses aggregate actors at the state level. Second, it makes intuitive sense that past interactions will influence future interactions. Event data is an excellent way to record these interactions and it lends itself easily for analysis using existing statistical techniques.

Major research using event data in international relations began in the 1960s and was “[h]istorically... coded by legions of bored undergraduates flipping through copies of *The New York Times*.”¹⁷ With the advent of personal computing, however, coding has switched from manual, human coding of events to machine coding. Machine coding has two major advantages

¹⁷ *Twenty Years*, 1.

over human coding. First, it is consistent. Since the computer program runs exactly the same analysis on each story it eliminates the errors that arise from differing interpretations, fatigue and bias. Second, machine coding is much faster than human coding.¹⁸

Since machine coding takes specialized software there have historically been very few choices for researchers. One of the earliest was the Kansas Event Data System (KEDS),¹⁹ developed by Schrodtt, and succeeded by Textual Analysis By Augmented Replacement Instructions²⁰ (TABARI), also written by Schrodtt. The other major coding software was developed at Harvard University and later spun off into its own company, Virtual Research Associates (VRA). Since the VRA reader is proprietary and there aren't any other coding programs available²¹ KEDS and TABARI are the only viable choice for most researchers looking to create their own data sets.²²

Machine coding has led to several changes in the nature of event data sets. First, since machine coding is much faster than human coding, larger data sets can be made. WEIS, for example, coded 98,043 events using five fields (actor, target, date, action category and arena) from 1966-1978.²³ The 10 million International Dyadic Events dataset created by VRA, on the

¹⁸ Ibid.

¹⁹ "KEDS." *Penn State Event Data Project*. <http://eventdata.psu.edu/software.dir/keds.html>. Accessed 10/5/2011.

²⁰ TABARI." *Penn State Event Data Project*. <http://eventdata.psu.edu/software.dir/tabari.html>. Accessed 10/5/2011.

²¹ Not to my knowledge, at least. Rice University has developed their own coding software but it isn't available for public use. <http://www.cs.rice.edu/~devika/conflict/projects.html>

²² I myself am using TABARI to generate the Eurasia data sets. The dataset used in *Predicting Intra-State Conflict* was generated using JABARI, an improved Java version of TABARI. Unfortunately that program was developed by Lockheed Martin and is proprietary.

²³ McClelland, Charles. *World Event/Interaction Survey (WEIS) Project, 1966-1978*. Inter-university Consortium for Political and Social Research [distributor]. Accessed 11/24/11. <http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/05211/detail#scope-of-study>

other hand, contains over a hundred times that and includes two more levels of specification for each actor.²⁴

Second, there has been a shift in actor specificity. Earlier data sets tended to focus on state-level dyads, with the inclusion of some major international organizations. More recent data sets have become more specific in their actor coding. For example, in the VRA dataset mentioned above there are two more levels of specificity for each actor. This means that instead of assigning an interaction to the state as a whole, researchers are able to attribute that action to the specific group actor (such as a political party, or a rebel group). Though it is likely that this level of specificity was too difficult to code and analyze manually there was also a theoretical drive behind the switch. Since the 1990s international conflict has shifted from inter-state conflict to mainly intra-state conflict. Therefore, in order to remain relevant, researchers needed a way to look at intra-state processes; while interesting, inter-state dyads aren't as relevant today.

Third, increased computer power has allowed for more complex analyses to be conducted on event data. While traditional statistical analysis is still very much used, political scientists have borrowed a number of pattern recognition techniques from the field of computer science in order to find relationships that would otherwise have been extremely difficult to discover. The techniques used range from temporal dependencies²⁵ to rule-based patterns modeled after Stephen Wolfram's "New Kind of Science"²⁶ to hidden Markov models.²⁷

²⁴ King, Gary and Lowe, Will. *10 Million International Dyadic Events*. Accessed 11/24/11. <http://gking.harvard.edu/data>.

²⁵ Kovar, Klaus, Furnkranz, Johannes, Petrak Johann and Pfahringer, Bernhard. *Searching for Patterns in Political Event Sequences: Experiments with the KEDS Database*. *Cybernetics and Systems*. 31:649-668.

²⁶ Hudson, Valerie, Schrodt, Philip and Whitmer, Ray. *A New Kind of Social Science: The Path Beyond Current (IR) Methodologies May Lie Beneath Them*. *A New Kind of Social Science*. Accessed 11/25/11. <http://nkss.byu.edu/papers.html>.

²⁷ Schrodt, Philip. *Pattern Recognition of International Crises using Hidden Markov Models*. In "Political Complexity: Nonlinear Models of Politics." Ed. Diana Richards. Pg. 296-328. Ann Arbor: University of Michigan Press. <http://eventdata.psu.edu/papers.dir/schro97.pdf>.

The use of computer science techniques on political science event data is in its infancy and there isn't yet an understanding of which temporal aggregation of data (daily, weekly, monthly, etc.) should be used, much less an established method for recognizing and interpreting patterns in actor interactions. At this stage the research is much like tossing spaghetti against the wall to see what sticks- a researcher will choose an algorithm that seems promising and apply it to a data set then move on to the next algorithm.²⁸ Given more time it is likely that this line of research will yield a standard set of analytical tools and methods but the field is not yet at that point.

Review of "Predicting Intra-State Conflict"

The prediction model used comes from "Predicting Intra-State Conflict Onset: An Event Data Approach Using Euclidean and Levenshtein Distance Measures,"²⁹ by Schrodt, D'Orazio and Yonamine³⁰. The goal of "Predicting Intra-State Conflict" is to apply sequence analysis techniques borrowed from computer science in order to analyze event data and develop a model that can predict the onset of intra-state conflict one to two months into the future. The authors gathered news stories from 29 Asian states for 1998-2010 and used DARPA's³¹ Ground Truth Dataset³² (GTDS) to determine the onset of intrastate conflicts in those states during that period.

²⁸ I am not criticizing this in any way. Research at this stage is a matter of trial and error, not refinement.

²⁹ Note: There are two versions of *Predicting Intra-State Conflict*, using two different titles. *Predicting Intra-State Conflict* was the original title of the paper and was the version used as the basis of this thesis. After the analysis for this thesis was completed however, the authors came out with a new version of their paper called *Kickoff to Conflict: A Sequence Analysis of Intra-State Conflict-Preceding Event Structures*. This new version has several changes, most noticeably the omission of the Baseline Peace Archetype (BPA) and the use of only the Euclidean distance measure. This thesis will refer to the works as *Predicting Intra-State Conflict* and will keep the use of the three different distance measures but will omit the work done on the BPA since D'Orazio indicated that there were methodological issues with it.

³⁰ Predicting Intra-State Conflict

Note: The above is the publically available version of the paper, but is not the most current. The current version of *Predicting Intra-State Conflict* can be obtained from the authors.

³¹ Defense Advanced Research Projects Agency. This agency is responsible for developing new technologies for the military. www.darpa.mil

³² Available upon request from the author.

This research was conducted as part of DARPA's Integrated Crisis Early Warning System (ICEWS) and the dataset created for it is called the "ICEWS" dataset in this thesis.

Schrodt, D'Orazio and Yonamine's work contributes to the field of international conflict in three major ways. First, their data set, at two million events, is the second largest event data set created so far³³ and uses the most news sources.³⁴ While the 10 Million International Dyadic Events database only used Reuters news stories to generate the event data³⁵, the ICEWS data set used sources from 75 international and regional news services.³⁶

Second, the authors applied sequence analysis techniques, specifically Euclidian and Levenshtein distance measures to determine the similarity of periods preceding the onset of intra-state conflict. In order to do so, all events in a given state were aggregated into one of four categories: verbal cooperation, material cooperation, verbal conflict and material conflict. Furthermore, for theoretical reasons actors within a state were classified as either government (GOV) rebels (REB) or other (OTH). GOV-REB as well as REB-OTH interactions were examined.³⁷

The aggregated data was then compared in 12 week sequence pairs. Pairs were chosen at random consisted of one of 3 types: pairs in which both sequences immediately preceded the onset of conflict, pairs in which one preceded conflict and the other preceded peace and pairs in which both sequences preceded peace.³⁸ Analysis was conducted using Euclidian, Euclidean Absolute Value and Levenshtein sequence analysis.

³³ As far as I am aware only the 10 Million International Dyadic Events database is larger.

³⁴ This dataset will now be referred to as the ICEWS dataset. ICEWS stands for Integrated Crisis Early Warning System and is the DARPA project that funded this research.

³⁵ King, Gary.

³⁶ Predicting Intra-State Conflict, 5.

³⁷ *Ibid.*, 6-8. This is in contrast to most other event data studies which aggregate actors at the state level and analyze state:state interactions.

³⁸ *Ibid.*, 11.

Euclidean distance is used to measure the distance between two vectors and is calculated by.³⁹

$$\sqrt{\sum_{i=1}^{12} (A_{ij} - B_{ij})^2}$$

where A_{ij} is the event count of type j ⁴⁰ at week i for actor A. So the value of the sequence:

A: 1 5 8 5 3 8 9 6 5 4 8 6 8

B: 5 8 5 3 8 9 6 5 4 8 6 8 9

is ten. A second version of Euclidean distance is called the Euclidan Absolute Value (EAV).

EAV is similar to Euclidean distance, but the instead of squaring, absolute values of the differences in the sequences are added together. Levenshtein distance, on the other hand, is calculated similarly but determines three different Euclidian distances and assigns the lowest value as the Levenshtein distance. First, the normal Euclidian distance is calculated. Then one sequence is shifted to the left by one place and the resulting distance calculated. Finally the original data is shifted to the right once place. This will result in a smaller distance if one of the sequences is similar, but temporally shifted. For example, in the above example actor B's counts are identical to A's, but are shifted to the left by one place. The Levenshtein distance, therefore, is significantly lower than the Euclidian distance.⁴¹

Finally, results of their analysis indicated that periods of peace have similar event type counts and that sequences preceding peace had significantly different event counts from periods

³⁹ Ibid., 12.

⁴⁰ Verbal cooperation, material cooperation, verbal conflict or material conflict.

⁴¹ Ibid., 12-13.

which preceded conflict. Interestingly, periods preceding conflict weren't similar to each other, suggesting that there is not one single event count pattern to intra-state conflict onset.

THESIS RESEARCH DESIGN

The goal of this research is to develop a model, which can accurately predict onsets of intra-state conflict in Eurasia 1 to 2 months into the future. For the purposes of this thesis Eurasia will be treated as consisting of three South Caucus states: Armenia, Azerbaijan and Georgia, and five Central Asian republics: Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan. I apply the analytical methods developed by Schrod, D'Orazio and Yonamine in "Predicting Intra-state Conflict Onset" to the Eurasian states for the period of January 1991 through October 2011.

Most intra-state conflict research using event data focuses on more volatile regions, such as the Middle East, Sub-Saharan Africa and Central and South Asia. By looking at intra-state conflicts in Eurasia this research will contribute to the current event-based intra-state conflict literature in several ways including: creation of an Eurasian event data database, testing the robustness of the prediction methodology used in "Predicting Intra-State Conflict" and examining whether periods leading up to civil conflict in different parts of the world exhibit the same event data characteristics.

There are three main stages to this research: raw data collection, event data production and event data analysis. Since this research involved the creation of a brand-new dataset certain assumptions and decisions were made which may affect the accuracy of the data and are discussed below.

Raw Data Collection

This stage of the research was dominated by two main questions. First, how many years of data should be collected? ICEWS data was collected for 1998-2010 but that doesn't necessarily mean that the same timeframe is appropriate for this project.

Although the ICEWS dataset covers the period from 1998-2010 the decision was made to examine 1991-Oct 2011.⁴² The dissolution of the Soviet Union makes a natural beginning point for the study of the Eurasian states and many of their intra-state conflicts occur before 1998. If these early years weren't included then there wouldn't have been enough onsets of conflict for the analysis. It is likely that the 1998 beginning point of the ICEWS dataset is more due to the limits of certain news sources or merely an arbitrary starting point rather than due to some theoretical reason.

Second, how many news sources should be used? The *10 Million International Dyadic Events* dataset on one hand only used Reuters⁴³ while the ICEWS dataset used 75 international and regional sources. This is important because it makes sense that the number and type of sources used would influence the "coarseness" of the data. International news agencies most likely only report news stories that are deemed sufficiently important and/or interesting enough for a global audience while smaller regional news agencies cover smaller, seemingly less important events. It may be the case that these smaller events significantly affect the pre-conflict event type count so that an analysis conducted with "fine" data would be able to predict a conflict whereas a similar analysis conducted with "coarse" data would not.

In the end it was decided to use sources from four major international newswire services: Reuters, Agence France Presse, Associated Press and United Press. The first two were chosen

⁴² Oct 2011 doesn't have any significance of its own, it is merely the date that the data was collected.

⁴³ They did use multiple Reuters sources, but only one at a time.

because they are used as sources in a majority of the machine-generated event data literature. The last two were chosen in order to make sure that no major news stories had been missed; between the four sources every single noteworthy story in the Eurasian region should be included.

There were three main reasons for not collecting stories from any smaller regional news sources. First, not every country had a regional news source. Second, the news sources that did exist didn't cover the entire 1991-2011 timeframe. If regional sources had been used for some periods and states but not others it would have distorted the data and might have led to misleading results. Finally, the amount of time needed to download and process the additional sources was a limiting factor, given that unlike the VRA and ICEWS projects, there was only one person working on creating this event data set.

Event Data Production

A total of 74,000 stories were collected from the four newswire services, including 35,000 from Agence France Presse alone. In order to get the most relevant stories possible a number of filters were applied before downloading the stories. For example, stories had to be set in one of the Eurasian states and sports and entertainment stories were omitted. After processing the stories with TABARI and removing duplicate events there were 54,301 events. Removing all stories in which both actors were not Eurasian brought the total down to 15,182 events. Finally, both actors were from the same state in 9,531 events.

TABARI, written by Dr. Schrodtt, was used to process the raw news stories into machine-readable event data because it and its predecessor, the Kansas Event Data System (KEDS) have been extensively used in event data research. "Predicting Intra-State Conflict" used a slightly modified version of TABARI to process their data called JABARI⁴⁴ but this software is

⁴⁴ So called because the program is written in Java.

proprietary to Lockheed-Martin and not available for public use. VRA Reader, the other major event data processing software, is also proprietary.

Once the event data had been generated it was necessary to determine whether or not there was enough data to conduct an analysis. Since relatively few event data datasets have been created using computer software there are no standards for the “density” of a dataset in events per state year.

Ideally every event would be recorded in a dataset but this is impossible for several reasons. Recall that an event datum is a record of an interaction between two actors and that these events are recorded by news sources. At every stage of the event data collection process information is lost and leads to the following formula:

$$D = X * R * C * T$$

Where:

D = Number of events in data set

X = Number of relevant events

R = % of events recorded by news sources

C = % of relevant news stories collected

T = % of events in news stories correctly coded

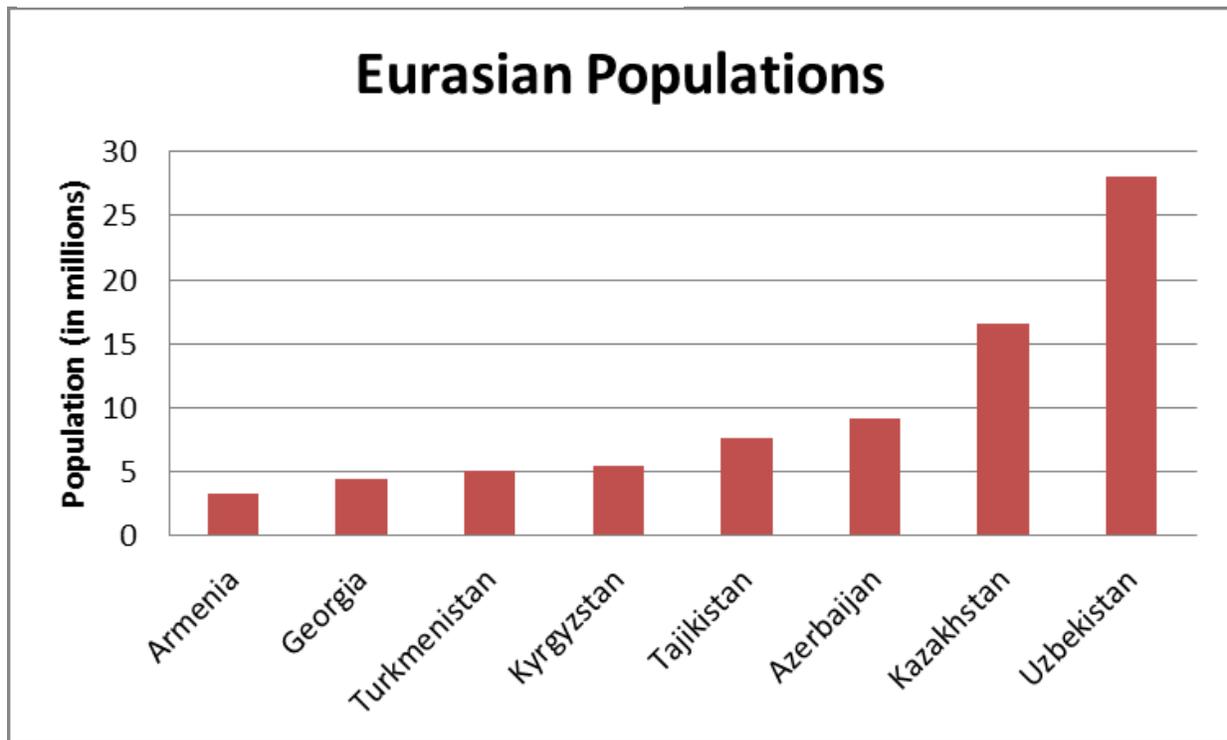
Only the final value, D, is known. There is no way to actually solve the equation for any of the other variables, but the formula does explain why event data sets almost certainly won't equal the actual number of relevant events. Of the four variables a researcher only has control over two of them- C (% of relevant news stories collected), and to some extent, T(% of events in news stories correctly coded). The percentage of relevant news stories collected is the easiest to control because the actor and verb dictionaries used can be improved.

Since it's not possible to directly determine how complete the Eurasian dataset is, an indirect method comparing it to two other datasets was used. The two chosen for comparison were VRA's 10 Million International Dyadic Events and the ICEWS dataset. VRA's dataset was chosen because it covers the entire world from 1990-2004 and provides a way to compare the number of Eurasian events generated for this project to the number generated by the VRA Reader.

VRA's dataset averages 55.5 events per state year where both of the actors were a Eurasian state. The Eurasian dataset generated for this project had an average of 90.37 events per state year, but unfortunately a direct comparison of the two numbers is misleading. VRA's dataset records almost no intra-state events; in virtually all of the events each actor represents another state. The Eurasian dataset, on the other hand, has a large percentage of intra-state events, at approximately 60 per state year. This leaves only 30 events per state year that were inter-state meaning that half of the events coded by the VRA Reader didn't show up in the Eurasian data set.

A more relevant comparison is the ICEWS data set since it was generated using similar software and data collection methods. It would be misleading though to only look at the average intra-state events per state year of all the ICEWS countries. The ICEWS dataset, after all, consists of 29 states including China, Russia and Indonesia. It makes sense that larger countries will have more actual events (X) and therefore more news coverage (R), so comparing those countries to the much smaller Eurasian ones would be inappropriate.

Figure 2: Population of Eurasian states



A more appropriate comparison is to only look at the ICEWS states that fall within the same population range as the Eurasian states. Eurasia ranges from a minimum of 3.3 million persons in Armenia to 28 million in Uzbekistan. Since Uzbekistan's population is much larger than that of the rest of the region, the population range was capped at Kazakhstan's 16.6 million.

Five ICEWS states met the population limits: Cambodia, Laos, New Zealand, Papua New Guinea and Singapore. As shown in figure 3, the comparable ICEWS states range from a low of 19 events in a year (Papua New Guinea, 2009) to 2138 (Cambodia, 1998). Eurasian states range from 2 events in a year (Tajikistan, 1997) to 559 (Georgia, 1993). Although on average the Eurasian states have a lower event count than the ICEWS states, there is some overlap. For example, Georgia's average intra-state event count per state year is 177, putting it ahead of 3 of the 5 comparable ICEWS states.

The Eurasian states have roughly similar event counts to those of similar size in the ICEWS dataset. The difference might be due to several factors including fewer sources and less available manpower. It also may be the case that the Eurasian states are simply covered less frequently in the news than some of the ICEWS states.

Figure 3: Event Count Comparison between selected ICEWS states and the Eurasian states.

<i>Intra-State Events</i>										
	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
Cambodia	-	-	-	-	-	-	2138	1202	431	302
Laos	-	-	-	-	-	-	65	103	147	118
New Zealand	-	-	-	-	-	-	94	139	148	143
Papua New Guinea	-	-	-	-	-	-	84	116	86	49
Singapore	-	-	-	-	-	-	229	300	324	277
Armenia	53	30	18	9	39	19	38	60	29	36
Azerbaijan	206	213	87	48	43	34	126	66	69	63
Georgia	355	559	132	103	88	84	139	101	71	138
Kazakhstan	13	20	20	35	17	25	54	47	38	46
Kyrgyzstan	7	9	10	11	9	10	28	55	50	24
Tajikistan	55	79	51	48	106	160	64	51	35	71
Turkmenistan	12	9	15	7	7	19	13	15	36	34
Uzbekistan	19	14	7	17	16	11	17	45	47	66
<i>Intra-State Events</i>										
	2002	2003	2004	2005	2006	2007	2008	2009	2010	Average
Cambodia	258	555	473	351	475	301	252	337	-	1018
Laos	65	89	80	52	111	72	60	47	-	108
New Zealand	166	127	155	138	165	121	141	86	-	131
Papua New Guinea	49	67	72	42	31	31	42	19	-	84
Singapore	283	327	357	304	332	80	63	51	-	283
Armenia	14	28	22	4	6	28	55	18	12	33
Azerbaijan	51	70	34	75	41	42	26	31	33	96
Georgia	79	232	319	82	116	161	712	187	60	177
Kazakhstan	49	52	46	45	42	45	18	29	29	32
Kyrgyzstan	59	11	12	149	87	81	17	53	253	21
Tajikistan	53	45	30	23	35	2	10	18	17	72
Turkmenistan	49	28	25	15	42	28	30	17	10	17
Uzbekistan	39	20	34	122	52	33	20	18	17	26

*Only full years are included

Event Data Analysis

This stage of the research entailed finding a source of data for the Eurasian conflict onset dates and actually running the model. Finding conflict onset data for Eurasia was necessary because the Ground Truth Dataset (GTDS) used by “Predicting Intra-State Conflict” was created especially for the ICEWS project and as such only covered the 29 Asian states from 1998-2010. Unfortunately DARPA did not release the criteria that they used to determine whether any particular state month experienced intra-state conflict.

After looking through the available datasets, the Uppsala Conflict Data Program (UCDP)⁴⁵ was chosen because it is a well-known and reputable data source, covered all of the Eurasian states for the chosen timeframe and had a clear definition of conflict onset. The UCDP’s definition of conflict onset is, “An armed conflict is a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year.”⁴⁶

Since the GTDS doesn’t include their criteria for determining intra-state conflict onsets it was necessary to compare the actual conflict onsets given by the GTDS to the conflict onsets given by the UCDP to figure out how compatible the two are. This is important because this research builds upon the work of “Predicting Intra-State Conflict” and it is necessary to know whether the conflict onsets predicted by the ICEWS model are the same as the conflicts predicted by the Eurasian data.

⁴⁵ Themnér, Lotta and Wallensteen, Peter, "Armed Conflict, 1946-2010." *Journal of Peace Research* 48(4) 2011. http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/

⁴⁶ “Definitions: Conflict, armed.” *Department of Peace and Conflict Research, Uppsala University*. <http://www.pcr.uu.se/research/ucdp/definitions/>. Accessed 11/5/11.

The conflict onsets from the GTDS and the UCDP for the ICEWS states from 1998-2010 were compared and found to have significant differences. First, while the GTDS listed 53 conflicts, the UCDP only recorded 21. There weren't any exact conflict matches either. The closest onset dates between the two were off by three months. Fourteen states were listed as having conflict in GTDS while the UCDP only had 8, though all UCDP states were included in the GTDS. Also, the proportion of events was dissimilar. For example, the GTDS counted only three conflicts in India while the UCDP recorded seven in the period in question.

The fact that the GTDS has more than double the recorded incidences of conflict onset indicate that it uses a much more sensitive definition of conflict onset than the UCDP does. Also, since the GTDS and the UCDP didn't have any conflict onsets in common, it is highly likely that the criteria used by the GTDS for determining conflict onset is not based on a battle deaths count like the UCDP is. This is further supported by the fact that the GTDS breaks down conflict onsets into four categories,⁴⁷ rebellion, insurgency, violence and domestic crisis.

Although an attempt was made to find a different source of conflict onset data that was more similar to the GTDS, none was found. The Correlates of War dataset, for example, only had data through 2007⁴⁸ while CONIS wasn't released at the time. Since the UCDP dataset was the most appropriate, the decision was made to continue research using its conflict onset data for the Eurasian states. It is important to note though that the use of the UCDP dataset means that the results of the model used in "Predicting Intra-State Conflict" cannot be directly compared to the results from the Eurasian model. The two are predicting two separate events, though both call these events "conflict onset."

⁴⁷ Actually five, but only four are relevant to this research. The fifth is international crisis.

⁴⁸ "Intra-State War Data (v4.0)." *Correlates of War*. <http://www.correlatesofwar.org/>. Accessed 12/1/11.

In addition to using a different data source for the conflict onset dates, one other major change was the elimination of the government-rebel and rebel-other actor filters for the event data. As mentioned above, the original “Predicting Intra-State Conflict” model employed a filter which only kept events in which one agent was classified as a member of the government and the other a member of a rebel group (GOV-REB) or one agent was a member of a rebel group (REB-OTH).

The assumption behind this filter was that interactions between the state and rebel groups and between rebel groups and non-governmental organizations and individuals are the driving force behind conflict escalation.⁴⁹ Though this assumption makes sense, it was unable to be used with the Eurasian dataset. Employing the filter on the Eurasian dataset resulted in the loss of almost all of the events. Out of a total of 9,531 Eurasian intra-state events, only 227 remained after using the filter. The model did not have sufficient data to run with the filter in place. A look at the Eurasian event data indicate that the reason the filter omits the vast majority of the events is due to a paucity of events with rebel actors. That is not to say that rebel actors weren’t included in the dictionaries used to classify the event data, the actors just weren’t classified as rebels. For example, in the case of Georgia, actors from Ossetia weren’t given the “REB” label, but instead an “OSS” label indicating that they were from Ossetia.

Given that it wasn’t possible to continue with the filter the decision was made to conduct the analysis without the filter. It was recognized that the addition of the extra events would likely add noise to the data, making it harder for the model to find a pattern and thus predict the onset of conflicts, though the addition of that extra data made sure that all relevant events were included in the analysis.

⁴⁹ Predicting Intra-State Conflict, 6.

Since the Eurasian model couldn't be directly compared to the results from the ICEWS model because the actor filter wasn't used and the conflict onset dates were obtained from the UCDP instead of the GTDS, two additional analyses were run, for a total of three. The first was an analysis of the original ICEWS data without the GOV-REB filter and using conflict onset data from the UCDP. The second was the original ICEWS data with the GOV-REB filter and using conflict data from the UCDP. It was hoped that by showing the effects of those two modifications on the original data that it might be possible to make an educated guess as to what the results of the Eurasian model would look like had it been possible to run it without those two modifications.

The analysis of the Eurasian event data set was run using the same techniques as the model created in "Predicting Intra-State Conflict," in order to test its robustness and applicability as well as to determine if there are any patterns to the intra-state conflict escalation process. A set of replication steps and scripts⁵⁰ were obtained from Schrodtt and D'Orazio and their methods were followed as closely as possible, excepting of course the elimination of the GOV-REB filter.

As mentioned above, the model used three different distance measurement techniques, Euclidean, Euclidean Absolute Value (EAV) and Levenshtein. Multiple distance measurement techniques were used to determine which was the most accurate.

The Eurasian event data, as produced by TABARI, is a sequence of individual event datum, each listing the date of the event, the two actor codes and a verb code representing the interaction between the two agents. The Conflict and Mediation Event Observations (CAMEO) code set, developed by Philip Schrodtt, was used to classify the actors and interactions.⁵¹

⁵⁰ Available upon request from Vito D'Orazio.

⁵¹ Schrodtt, Philip. "CAMEO: Conflict and Mediation Event Observations Event and Actor Codebook." *Event Data Project, Department of Political Science, Pennsylvania State University*. Version 1.1b3, March 2012. <http://eventdata.psu.edu/cameo.dir/CAMEO.Manual.1.1b3.pdf>.

Events were grouped into four categories: verbal cooperation, material cooperation, verbal conflict and material conflict. An event count structure was used instead of assigning a numerical value to each action in order to address the problem of peaceful and aggressive actions canceling each other out. For example, in the Goldstein scale peaceful actions such as negotiations, receive a positive score while aggressive ones receive a negative score, which could lead to events canceling each other out. The

Figure 4: List of Eurasian conflict onsets.

Conflict Onset Weeks		
State	Year	Week
Azerbaijan	1991	35
Azerbaijan	1994	48
Azerbaijan	2005	27
Georgia	1991	35
Georgia	2004	18
Georgia	2008	18
Tajikistan	1992	5
Tajikistan	1997	48
Tajikistan	2010	22
Uzbekistan	1998	48
Uzbekistan	2003	48

four chosen categories on the other hand are comprehensive and mutually exclusive and have been used as far back as 1980 for event counts.⁵²

Data was aggregated at the weekly level in order to produce the event counts. The original study aggregated data at the weekly and monthly levels; however the difference in accuracy between them was small enough that only the weekly aggregation was tested on the Eurasian data set. Once an event count had been established for each week of the study period the conflict onset data was merged with the event data. Weeks during which conflict either began or was ongoing were coded with a “1” while peace weeks were coded with a “0”. A random sample of 28 peace sequences, that is, peace weeks preceded by at least 12 weeks of

⁵² For exact definitions of the four categories see Predicting Intra-State Conflict, 7.

peace, were selected, along with all 11 of the conflict onsets.⁵³ It should be noted that while the UCDP has more than 11 conflicts listed in the Eurasian region during the 1990-2011 timeframe, several of those conflicts began during while another conflict was occurring in the same state. Since “Predicting Intra-State Conflict” only counted conflict weeks as onsets if they were preceded by at least 3 months of peace, these conflicts were not included in this analysis.

Once the data were merged and the conflict onset and peace sequences selected, the analysis was conducted. The first part of the analysis was to calculate the vector distances between the 12 week event data count sequences using each of the three distance measures. Dyads were classified into three types, 0, 1 and 2. Type 0 dyads were sets in which both sequences were peace sequences. In type 1 dyads, one sequence was a conflict onset and the other a peace. Finally, both sequences in type 2 dyads were conflict onsets.

The data was then randomly separated into two groups, a training set and a set for testing the model’s predictions. Since this model is trying to find patterns which can be used to predict conflict onset, if conflicts escalate in a similar manner the distances in type 2 dyads should be small. Similarly, if conflict onsets are preceded by event structures that are dissimilar to those found during peace times the distances of type 1 dyads should be large.

A bivariate logistic regression model was used with the event sequence distances as the regressors. The model yields four predicted possibilities since the sequence in each state can either be a peace sequence or a conflict sequence. Assigning a 0 for peace and a 1 for a conflict onset sequence the four probabilities are (0,0), (0,1), (1,0) and (1,1) where the first number represents the first state in the dyad and the second number the second state. The probability (0,0) therefore is a type 0 dyad and (1,1) is a type 2 dyad. The (0,1) and (1,0) probabilities are added together to produce the probability of a type 1 dyad.

⁵³ This approximately 2-1 ratio of peace:conflict sequences was adapted from the original model.

Once the model was trained the second half of the sequences were used to verify the accuracy of the model’s predictions. Each sequence was given to the model which classified the sequence as either a type 0, 1 or 2 dyad. That prediction was then compared to the actual classification of the sequence.⁵⁴

RESULTS

Figure 5: Overall model accuracy.

Overall		
Distance Measure	Source	Correctly Classified
Euclidean	Eurasia	55.43%
	Non-Filtered ICEWS w/ UCDP	47.67%
	Filtered ICEWS w/ UCDP	57.51%
Euclidean Absolute Value	Eurasia	57.61%
	Non-Filtered ICEWS w/ UCDP	45.08%
	Filtered ICEWS w/ UCDP	54.92%
Levenshtein	Eurasia	54.35%
	Non-Filtered ICEWS w/ UCDP	48.19%
	Filtered ICEWS w/ UCDP	46.63%

Figure 5 shows the overall accuracy of the model, utilizing the three different distance measures, Euclidean, Euclidean Absolute Value (EAV) and Levenshtein and three different sets of data.

The first data set, Eurasia, is the event data collected from the eight Eurasian countries, using the Uppsala Conflict Data Program’s dataset to identify conflict onset dates. The Eurasian data set did not include the GOV-REB filter. The second, “Non-Filtered ICEWS w/ UCDP” is the original ICEWS data without the GOV-REB filter and using the UCDP conflicts instead of

⁵⁴ For a more in-depth explanation of the model see the most recent version of the paper by D’Orazio and Yonamine, titled “Kickoff to Conflict: A Sequence Analysis of Intra-State Conflict-Preceding Event Structures,” pgs. 15-17.

the GTDS conflict onsets. Finally, the third data set, “Filtered ICEWS w/ UCDP” is the original ICEWS data with the GOV-REB filter but using UCDP conflict onsets. Note that the discussion of the results will begin with a look at the Eurasian model data, so even though data for all three of the model runs will be included, unless specifically stated the results referred to are for the Eurasian model and data.

The results shown in figure 5 indicate that all of the models had roughly the same level of overall accuracy, around 50%. In addition, all of the data gathered from the model indicates that the three distance measures perform roughly the same, though in some cases the two Euclidean measures outperform the Levenshtein distance measure. While there isn’t any conclusive evidence to show that the Levenshtein measure performs worse than the two Euclidean measures, the fact that it would underperform the other two make theoretical sense.

The Levenshtein distance measure shifts the sequence up and down one spot in either direction in order to minimize the total sequence distance. If the Levenshtein measure produces smaller sequence distances then there is a greater chance that conflict onset sequences will be misclassified as peace sequences, which, as will be discussed further on, tend to have smaller event counts.⁵⁵ Since all three distance measures are essentially the same, the discussion will mostly focus on the Euclidean distance measure, though the results from each will be included in the tables.

Four measures of model accuracy were examined for the three dyad types to determine the strengths and weaknesses of the model. Figure 6 shows the sensitivity results from each of the models and dyads. Sensitivity is the number of correctly classified positives divided by the total number of positives. For example, looking at the type 0 dyad, the sensitivity is equal to the

⁵⁵ It is notable that D’Orazio and Yonamine in the updated version of the paper, “Kickoff to Conflict,” omit the EAV and Levenstein distance measures. It is not necessary to run all three since they produce roughly similar results.

number of times that the model correctly predicted a type 0 dyad, divided by the total number of type 0 dyads in the data.

Figure 6: Model sensitivity

Sensitivity				
Distance Measure	Source	Type "0"	Type "1"	Type "2"
Euclidean	Eurasia	86.17%	27.10%	0.00%
	Non-Filtered ICEWS w/ UCDP	32.89%	75.28%	0.00%
	Filtered ICEWS w/ UCDP	61.84%	71.91%	7.14%
Euclidean Absolute Value	Eurasia	88.30%	29.68%	0.00%
	Non-Filtered ICEWS w/ UCDP	23.68%	77.53%	0.00%
	Filtered ICEWS w/ UCDP	53.95%	73.03%	10.71%
Levenshtein	Eurasia	81.38%	30.32%	0.00%
	Non-Filtered ICEWS w/ UCDP	36.84%	73.03%	0.00%
	Filtered ICEWS w/ UCDP	48.68%	59.55%	3.57%

The results show that while the model performed rather well and classified approximately 86% of the total type 0 dyads correctly, only 27% of the type 1 dyads and none of the type 2 dyads were correctly labeled. The low accuracy for dyads including a conflict onset sequence is a major problem for the applicability of this model to real-world situations. After all, the model is supposed to be used to predict the onset of conflict and as it stands though it can predict peace rather well it performs poorly at predicting conflict onsets.

Specificity, on the other hand, is the number of false classifications divided by the total false instances in the data. Here, for example, a false type 0 dyad is a true type 1 or 2 dyad, so the specificity is the number of correctly classified type 1 and 2 dyads divided by the total number of type 1 and 2 dyads. As expected, type 1 dyads show low specificity due to the fact that the model has difficulty correctly classifying type 1 and 2 dyads. At the same time the specificity for type 1 and 2 dyads rise due to the inclusion of the high-accuracy type 0 dyad.

Figure 7: Model specificity

Specificity				
Distance Measure	Source	Type "0"	Type "1"	Type "2"
Euclidean	Eurasia	23.33%	76.06%	59.48%
	Non-Filtered ICEWS w/ UCDP	57.26%	24.04%	55.76%
	Filtered ICEWS w/ UCDP	56.41%	47.12%	67.27%
Euclidean Absolute Value	Eurasia	25.56%	77.93%	61.81%
	Non-Filtered ICEWS w/ UCDP	58.97%	17.31%	52.73%
	Filtered ICEWS w/ UCDP	58.12%	42.31%	64.24%
Levenshtein	Eurasia	26.11%	71.83%	58.31%
	Non-Filtered ICEWS w/ UCDP	55.56%	26.92%	56.36%
	Filtered ICEWS w/ UCDP	46.15%	36.54%	54.55%

The third measure of model accuracy, positive predictive value, is defined as the number of correct positive classifications divided by the total number of positive classifications. In this case, the positive predictive value of a type 0 dyad is the number of type 0 dyads that were classified correctly, divided by the total number of dyads that were actually classified as type 0. Approximately half of the dyads that were classified as types 0 and 1 were actually types 0 and 1. The asterisk in the type 2 column however indicates that the model never predicted a type 2 dyad. Since it's not possible to divide by zero an asterisk is used instead of 0%.

Figure 8: Model positive predictive value

Positive Predictive Value				
Distance Measure	Source	Type "0"	Type "1"	Type "2"
Euclidean	Eurasia	55.67%	54.55%	*
	Non-Filtered ICEWS w/ UCDP	45.45%	48.55%	*
	Filtered ICEWS w/ UCDP	70.15%	53.78%	28.57%
Euclidean Absolute Value	Eurasia	57.04%	59.74%	*
	Non-Filtered ICEWS w/ UCDP	42.86%	45.70%	*
	Filtered ICEWS w/ UCDP	69.49%	52.42%	30.00%
Levenshtein	Eurasia	55.84%	50.00%	*
	Non-Filtered ICEWS w/ UCDP	48.28%	48.15%	*
	Filtered ICEWS w/ UCDP	53.62%	45.69%	12.50%

Figure 9: Model negative predictive value

Negative Predictive Value				
Distance Measure	Source	Type "0"	Type "1"	Type "2"
Euclidean	Eurasia	66.23%	61.17%	93.21%
	Non-Filtered ICEWS w/ UCDP	63.04%	60.00%	85.49%
	Filtered ICEWS w/ UCDP	76.98%	66.22%	86.02%
Euclidean Absolute Value	Eurasia	71.43%	62.54%	93.21%
	Non-Filtered ICEWS w/ UCDP	61.59%	52.38%	85.49%
	Filtered ICEWS w/ UCDP	73.88%	65.22%	86.34%
Levenshtein	Eurasia	62.77%	60.58%	93.21%
	Non-Filtered ICEWS w/ UCDP	64.44%	58.62%	85.49%
	Filtered ICEWS w/ UCDP	68.55%	53.25%	85.41%

The final measure used is negative predictive value. Negative predictive value is the number of correctly classified negatives divided by the correctly classified negatives plus the incorrectly classified negatives. Taking the type 0 dyad as an example, the negative predictive value is the number of type 1 and 2 predictions that were correct, divided by the number of correct type 1 and 2 predictions plus the number of incorrectly classified type 0 sequences. Types 0 and 1 have a negative predictive score of about 60% while the value of type 2 is at 93%. The high value for type 2 indicates that 93% of the time when the model says that a pair is not a type 2 dyad, it is actually not a type 2 dyad. Since no sequences were actually classified as type 2, this number also shows the percent of the onset sequences that were actually type 2.

The next three figures show the distribution of event distances based on dyad type.

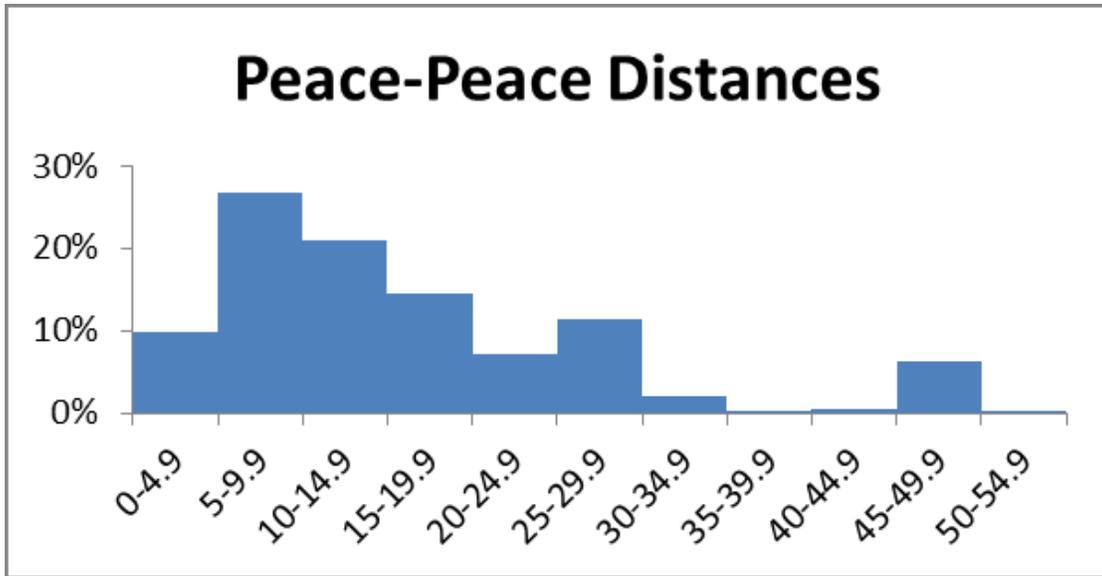


Figure 11: Distribution of type 1 total sequence distances

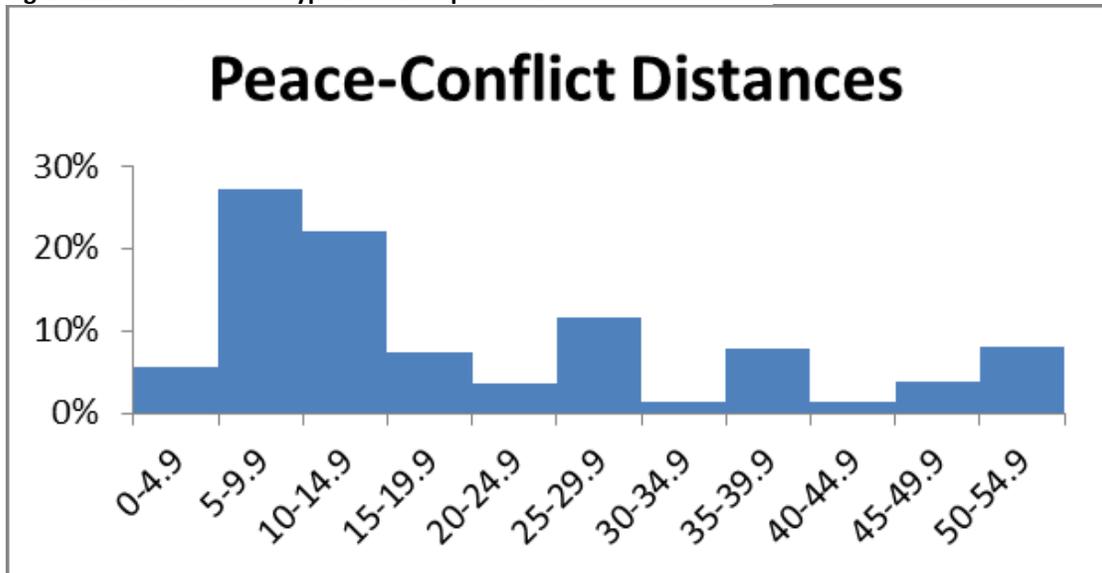
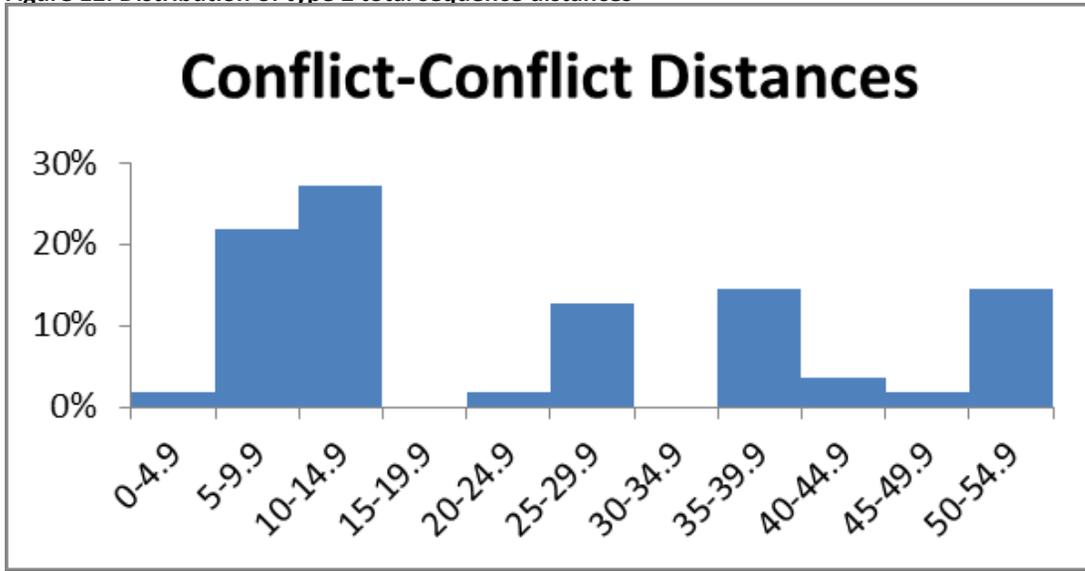


Figure 12: Distribution of type 2 total sequence distances



As can be seen in the distance charts, the type 0 dyads tend to have similar event count distances, concentrated toward the lower end of the scale. Type 1 dyads, while still mostly concentrated toward the lower end of the scale is skewed right with some higher sequence distances. Type 2 dyads show a similar pattern to type 1 dyads though with a slightly higher percentage of sequences with greater distances.

Note that these charts do not represent the total event counts for each type, but the distance between the sequences. The figures therefore do not necessarily show that there are fewer events on average in peace sequences, but that the sequences leading up to peace weeks are similar. Likewise, the fact that the type 2 dyads are spread out across the x axis indicates that while some conflicts escalate in a similar manner, many do not. In addition, the type 1 dyad chart shows that many conflict sequences display similar event count patterns as do peace sequences, though a smaller number do not.

Figure 13: Eurasian model average sequence distance

Average Distance						
Distance Measure	Dyad Type	V Coop	M Coop	V Confl	M Confl	Total
Euclidean	Type 0	6.23	2.11	3.70	4.36	16.39
	Type 1	8.25	2.63	4.70	4.63	20.22
	Type 2	10.33	3.13	5.64	4.98	24.08
Euclidean Absolute Value	Type 0	12.85	3.39	7.36	8.84	32.44
	Type 1	18.94	4.94	10.18	10.15	44.20
	Type 2	24.95	6.36	12.98	11.24	55.53
Levenshtein	Type 0	7.06	2.62	4.56	5.50	19.74
	Type 1	8.74	3.74	6.04	6.48	25.01
	Type 2	10.33	4.75	7.33	7.40	29.80

The figure above shows that in each of the distance measures, type 0 dyads have the smallest distances in all four categories, followed by type 1 and then type 2 dyads.

Interestingly though, while the absolute event sequence distance increases from type 0 to 2 in each of the four event categories, not all of the categories increase at the same rate. As figure 14 shows, while the percentage of events that are classified as either material cooperation or verbal conflict remains steady in all three dyad types, the percentage of events that are

Figure 14: Eurasian model average sequence distance percentages

Distribution of Total Event Count by Type					
Distance Measure	Dyad Type	V Coop	M Coop	V Confl	M Confl
Euclidean	Type 0	38.02%	12.85%	22.55%	26.57%
	Type 1	40.81%	13.02%	23.25%	22.91%
	Type 2	42.91%	13.01%	23.41%	20.66%
Euclidean Absolute Value	Type 0	39.62%	10.46%	22.69%	27.24%
	Type 1	42.85%	11.17%	23.02%	22.96%
	Type 2	44.92%	11.46%	23.38%	20.24%
Levenshtein	Type 0	35.77%	13.27%	23.11%	27.85%
	Type 1	34.96%	14.97%	24.15%	25.92%
	Type 2	34.66%	15.92%	24.59%	24.83%

classified as material conflict actually *decrease*. The difference is made up by an increase in the percentage of events classified as verbal cooperation, except by the Levenshtein measure, in which an increase in material cooperation makes up most of the difference.

So in absolute terms peace-peace sequences are most similar to each other while peace-conflict sequences are a bit less similar and conflict-conflict sequences are the least similar to each other in all measures. At the same time, the fact that the percentage of the total sequence distance contributed by verbal cooperation increases from type 0 to type 2 dyads indicates that much of the variability in conflict escalation sequences comes from verbal cooperation. In other words, the lead up to some conflicts is full of dialog while others are lacking. The decrease from types 0 to 2 in material conflict on the other hand indicate that its average distance remains relatively stable; all three types of dyads exhibit similar material conflict event distances.

Comparison and Interpretation of Results

Despite the fact that the Eurasian model was only able to correctly classify approximately half of the out of sample test cases the model actually supports the findings of the original ICEWS model. This is because though the accuracy is lower, the model demonstrates the same general patterns as the original and the reasons for the lowered accuracy are mostly understood and are not a function of the model itself.

Figure 5, which gives the overall accuracy of the three models, shows that the Euclidean and EAV distance measures show a 10% reduction in model accuracy between the ICEWS filtered model to the non-filtered model.⁵⁶ Also note that the filtered model was the only model run that actually predicted any type 2 dyads. The increased noise in the data that results from eliminating the GOV-REB filter seems to have a negative effect on the overall accuracy of the

⁵⁶ Interestingly, the Levenshtein model shows a slight *increase* in model accuracy. The reason for this is unknown and doesn't make sense theoretically. Most of the interpretation will focus on the two Euclidean measures since they are the measures used for interpretation of results in the original research.

model, which vindicates D’Orazio and Yonamine’s original assumption. It is reasonable to assume therefore, that had the Eurasian data been detailed enough to run the filter that the accuracy of the Eurasian model would be several percentage points higher.

Another possible reason for the reduction in model accuracy is the use of the UCDP instead of GTDS for conflict onset dates, as discussed above. The accuracy of the filtered ICEWS model using UCDP tends to be lower than the same model using GTDS onsets. For example, the sensitivity of the type 0 dyads is at 100%⁵⁷ for the GTDS model while only 61% in the UCDP model. This holds true for most measures, though there are a couple cases where the accuracy actually increased slightly. Overall though, what this shows is that the definition of “intra-state conflict” used matters greatly to the accuracy of this type of model. The battle-deaths method for determining conflict onset as used by the UCDP seems to be difficult to predict. Unfortunately the definition used by the GTDS, as mentioned above, isn’t known so it’s difficult to say what phenomenon the model is actually good at identifying.

One other possible source of error was briefly mentioned earlier; that is, the fact that some of the conflict onsets listed in the UCDP database were hidden by previous conflicts in the same state. Take, for example, the case of Georgia. According to the UCDP⁵⁸ Georgia experienced a civil conflict in South Ossetia which began on December 22, 1991 and lasted through July 14, 1992. On August 14, 1992 the UCDP records Georgia entering into a civil conflict with Abkhazia but because the previous conflict had ended less than three months previous this could not be counted as a conflict onset for the purposes.

This highlights the fact that the structure of this model implicitly assumes that each state only has one possible source of intra-state conflict. Events are aggregated at the state level, so

⁵⁷ Kickoff to Conflict, 40.

⁵⁸ Themnér.

even if a state such as Georgia has multiple rebel or separatist groups there is no way to separate the event patterns leading up to each conflict. Instead the data from both groups is merged together which may make the processing of finding patterns much more difficult. It is possible that a state with multiple rebel groups may be in the process of signing a peace agreement with one but getting ready to enter into armed conflict with another at the same time. Most likely those two processes would have distinct event count sequences but because the data is aggregated at the state level the sequences would be merged and the pattern rendered incomprehensible. A disaggregated analysis would not have been possible with the Eurasian data set given that even with the aggregated GOV-REB filter on there wasn't enough data to run an analysis, but something similar might work with larger data sets, such as ICEWS, though it might require substantial editing of the actors dictionary because instead of classifying a rebel as REB, each rebel would have to be classified according to his or her specific rebel group.

One final reason that helps to explain the diminished performance of the model is the fact that the model was run on an extremely small population. In total there were only 11 conflict onsets and data from only four countries was included.⁵⁹ In contrast the original ICEWS model had data from 15 states with 53 conflict onsets⁶⁰ and a much larger pool of event data sources. Fewer conflicts mean that there was less data to train the model with since the number of peace sequence selected was based on the amount of conflict sequences and fewer sources means that relevant events may have been missed. Both of those combined most likely played a role in decreasing the accuracy of the Eurasian model as compared to the ICEWS version.

Despite the decreased accuracy, the Eurasian model displays the same trends as the original, lending support to D'Orazio and Yonamine's observations about the nature of conflict.

⁵⁹ Although data from all Eurasian states was collected, only data from states which experienced at least one case of intra-state conflict was included, as per the original model design.

⁶⁰ Kickoff to Conflict, 37.

First, the distribution of sequence distances based on dyad type followed similar patterns, though the scale is different. The largest sequence distance was 180 in the ICEWS model⁶¹ while the max of the Eurasian model was only 55. In both the ICEWS and Eurasian sets however, the majority of the type 0 sequences were clustered at the low end of the scale, supporting the claim that peace sequences tend to have similar event counts to each other. Likewise, types 1 and 2 showed an increase in total sequence distance in both models. This supported the finding that while peace and conflict sequences tend to differ from each other, conflicts don't have a single escalation pattern. Each conflict seems to be unique.

Second, D'Orazio finds that the ICEWS model over predicts type 0 dyads and has trouble distinguishing between dyad types 1 and 2⁶². The Eurasian model shows the same pattern. As is shown above, the model correctly classified 86% of the type 0 dyads but only 27% of the type 1 dyads and none of the type 2 dyads. At the same time the positive predictive value for types 1 and 2 were around 55%, indicating that only half of the predictions of types 1 and 2 were correct. There is no positive predictive value for type 2 dyads since none were predicted.

These findings provide additional support to the claim that weeks of peace are preceded by similar event structures while the weeks preceding conflict onsets tend to vary. In addition, some conflicts seem to be preceded by similar event structures as evidenced by the cluster of small sequence distances in figure 11, the type 1 dyads sequence distribution. D'Orazio labels these "hidden conflict"- those conflicts which either aren't preceded by relevant events or for which events aren't recorded.⁶³ Such conflicts are frequently mislabeled as peace sequences as evidenced by the over prediction of type 0 dyads.

⁶¹ Ibid, 29.

⁶² Ibid, 21.

⁶³ Ibid.

CONCLUSION

The goal of this research was to accurately predict onsets on intra-state conflict in the Eurasian region. Overall the model was able to make predictions with about 55% accuracy, but there was a large variation between the accuracy of peace predictions and conflict predictions. This means that practically, this model, using this data set, is not able to make predictions which can be used to inform policy decisions.

This doesn't mean, however, that the model design is flawed. It is more likely that the data used for the event sequences and conflict onsets merely needs to be improved. Accuracy would also likely improve if the number of countries studied increased. In addition, despite decreased accuracy the Eurasian model still showed the same conflict patterns as the original ICEWS model and lends support to the notion that conflict onsets tend to have unique event sequence patterns. Peace-peace sequences showed decreased sequence distances, indicating that peace times display similar event patterns. The fact that this same pattern was found in both the original ICEWS study and the Eurasian model lends support to the notion that this may be a pattern inherent to global intra-state conflict, not just a regional characteristic, though it should be noted that the Eurasian region does border some of the states studied in the original research. Aside from confirming the findings of the original ICEWS project, this research contributed to the literature through the development of a new event data database and tested this model on a new region of the globe.

There are several areas where this research can be continued. First, the definition of "intra-state conflict" needs to be clearly defined. Determining the differences between the UCDP and GTDS definitions of that term may help researchers to better understand why event

data was more accurate in predicting instances of the latter and not the former. At the very least it is essential to know exactly what is being predicted.

Second, standards for event data research need to be developed. Specifically, how “dense” does a dataset need to be before one can reasonably assume that the majority of relevant events have been included? Third, a way needs to be found to differentiate the “hidden conflicts” from peace sequences. D’Orazio and Yonamine suggest using sentiment analysis on social media sites to determine shifting popular sentiment preceding conflict onset that may occur even when no relevant events are recorded.⁶⁴

⁶⁴ Ibid.

WORKS CITED

- Collier, Paul. *The Bottom Billion*. Oxford: Oxford University Press. 2007.
- “Definitions.” *Uppsala Conflict Data Program*. Uppsala University.
<http://www.pcr.uu.se/research/ucdp/definitions/>.
- “Definitions: Conflict, armed.” *Department of Peace and Conflict Research, Uppsala University*.
<http://www.pcr.uu.se/research/ucdp/definitions/>. Accessed 11/5/11.
- D’Orazio, Vito and Yonamine, James. *Kickoff to Conflict: A Sequence Analysis of Intra-State Conflict-Preceding Event Structures*. Pennsylvania State University. Not yet published.
- D’Orazio, Vito, Yonamine, James and Schrodt, Philip. *Predicting Intra-state Conflict Onset- An Event Data Approach Using Euclidian and Levenshtein Distance Measures*. Pennsylvania State University. 2011.
<http://eventdata.psu.edu/papers.dir/MPSA11.ICEWS.sequence2.1.pdf>.
- Englehart, Neil and Kurzman, Charles. “Welcome to World Peace.” *Social Forces*. Vol. 84, No. 4.
- Hempel, Carl G. and Oppenheim, Paul. "Studies in the Logic of Explanation." *Philosophy of Science* 15,2.
- Hudson, Valerie, Schrodt, Philip and Whitmer, Ray. *A New Kind of Social Science: The Path Beyond Current (IR) Methodologies May Lie Beneath Them*. A New Kind of Social Science. Accessed 11/25/11. <http://nkss.byu.edu/papers.html>.
- “Human Security Report, Part I: The Causes of Peace.” *Human Security Report Project*. Accessed 4/16/12. <http://www.hsrgroup.org/human-security-reports/20092010/text.aspx>.
- “Intra-State War Data (v4.0).” *Correlates of War*. <http://www.correlatesofwar.org/>. Accessed 12/1/11.
- “KEDS.” *Penn State Event Data Project*. <http://eventdata.psu.edu/software.dir/keds.html>. Accessed 10/5/2011.

- King, Gary and Lowe, Will. *10 Million International Dyadic Events*. Accessed 11/24/11.
<http://gking.harvard.edu/data>.
- King, Gary, Beck, Nathaniel, and Zeng, Lengche. "Improving Quantitative Studies of International Conflict: A Conjecture." *American Political Science Review*. Vol. 94, No. 1. March 2000.
- Kovar, Klaus, FurnKranz, Johannes, Petrak Johann and Pfahringer, Bernhard. *Searching for Patterns in Political Event Sequences: Experiments with the KEDS Database*. *Cybernetics and Systems*. 31:649-668.
- McClelland, Charles. *World Event/Interaction Survey (WEIS) Project, 1966-1978*. Inter-university Consortium for Political and Social Research [distributor]. Accessed 11/24/11.
<http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/05211/detail#scope-of-study>
- Schrodt, Philip. "CAMEO: Conflict and Mediation Event Observations Event and Actor Codebook." *Event Data Project, Department of Political Science, Pennsylvania State University*. Version 1.1b3, March 2012.
<http://eventdata.psu.edu/cameo.dir/CAMEO.Manual.1.1b3.pdf>.
- Schrodt, Philip. "Seven Deadly Sins of Contemporary Quantitative Political Analysis." Presented at annual meeting of the American Political Science Association, 2010.
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1661045##. Accessed 4/16/12.
- Schrodt, Philip. *Pattern Recognition of International Crises using Hidden Markov Models*. In "Political Complexity: Nonlinear Models of Politics." Ed. Diana Richards. Ann Arbor: University of Michigan Press. <http://eventdata.psu.edu/papers.dir/schro97.pdf>.
- Schrodt, Philip. *Twenty Years of the Kansas Event Data System Project*. University of Kansas. Written 11/9/2006. Accessed 11/24/11.
<http://eventdata.psu.edu/utilities.dir/KEDS.History.0611.pdf>.
- TABARI." *Penn State Event Data Project*. <http://eventdata.psu.edu/software.dir/tabari.html>. Accessed 10/5/2011.
- Themnér, Lotta and Wallensteen, Peter, "Armed Conflict, 1946-2010." *Journal of Peace Research* 48(4) 2011.
http://www.pcr.uu.se/research/ucdp/datasets/ucdp_prio_armed_conflict_dataset/
- Themnér, Lotta and Wallensteen, Peter. "Armed Conflicts, 1946-2011." *Journal of Peace Research*. Vol. 49, No. 4. 565-575.
- Ward, Michael D., Greenhill, Brian D. and Bakke, Kristin M. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research*. Vol. 47 No. 4.