Data-based Computational Approaches to Forecasting Political Violence

Philip A. Schrodt, James Yonamine, and Benjamin E. Bagozzi

1 Introduction and Overview

The challenge of terrorism dates back centuries if not millennia. Until recently, the basic approaches to analyzing terrorism—historical analogy and monitoring the contemporary words and deeds of potential perpetrators—have changed little: the Roman authorities warily observing the Zealots in first-century Jerusalem could have easily traded places with the Roman authorities combatting the Red Brigades in twentieth century Italy.

This has changed with the exponential expansion of information processing capability made possible first by the development of the digital computer, followed by the phenomenal growth in the quantity and availability of machine-readable information made possible by the World Wide Web. Information that once circulated furtively on hand-copied sheets of paper (or papyrus) is now instantly available—for good or ill—on web pages which can be accessed at essentially no cost from anywhere in the world. This expansion of the scope and availability of information in all likelihood will change the dynamics of the contest between organizations seeking to engage in terrorism and those seeking to prevent it. It is almost certainly too early to tell which group will benefit more—many of the new media are less than a decade old—but the techniques of processing and evaluating information will most certainly change.

This chapter provides an extensive overview of inductive statistical and computational methodologies used in the analysis and forecasting of political violence, and some of the challenges specific to the issue of analyzing terrorism. It is intended for the non-specialist, but assumes a general familiarity with data and computational methods. Our purpose is not to exhaustively explore any of these methods—each

P.A. Schrodt (✉) • J. Yonamine • B.E. Bagozzi
Political Science, Pennsylvania State University, University Park, PA 16801, USA
e-mail: schrodt@psu.edu; jxy190@psu.edu; beb196@psu.edu

V.S. Subrahmanian (ed.), *Handbook of Computational Approaches to Counterterrorism*,
DOI 10.1007/978-1-4614-5311-6_7,

technique would typically require tens or hundreds of pages—but instead to provide a sufficient introduction to the basic concepts and vocabulary, as well as a very extensive sets of references, so that the reader can explore further on his or her own.

The psychologist and philosopher William James, in his Lowell Institute lectures in 1906, subsequently published under the title *Pragmatism: A New Name for Some Old Ways of Thinking* notes that the fundamental split in philosophy, dating to the very origins of the field, is between "rationalists" who seek to find an intellectual structure that will reveal a proper order in the world, and "empiricists," who take the disorder of the observed world as a given and simply try to make as much sense of it as they can. More than a century later, we find exactly the same split in formal approaches in the social sciences: The rationalist position is expressed in deductive approaches such as game theory, expected utility models, systems dynamics and agent-based models, which seek to explain behavior from a set of a priori first-principles and their consequent emergent properties. The empiricist approach is found in inductive statistical and computational data-mining approaches which extract structured information from large sets of observed data. Both approaches are well-represented in this volume [38, 48, 103, 105, 117, 123, 151, 152] but this review will focus only on the empirical approaches.

Throughout this chapter, we will generally be looking at models which focus on forecasting and understanding political violence in general, not just approaches to terrorism per se. This is done for two reasons. First, due to a combination of data limitations and a paucity of interest in the research community prior to the 2000s, the number of quantitative studies of terrorism was quite limited and focused on a relatively small number of approaches [149]. However, the large-scale research efforts did consider various forms of political violence more generally. Second, most of those more general methods are at least potentially applicable to the study of terrorism—in fact, many are generally applicable to almost any behavior for which large amounts of data are available—albeit we will frequently caveat these possibilities with concerns about some of the atypical aspects of terrorism, in particular the distinctions between methods appropriate to rare-events analysis and those dealing with high-frequency behaviors. Finally, in many instances, there is a close relationship between situations of general political violence such as civil war and state failure, and conditions which encourage the formation and maintenance of terrorist groups, so political instability is of considerable interest on its own.

We will first provide a general overview the types of data commonly used in technical forecasting models, then consider the two broad categories of computational models: statistical and algorithmic. Within statistical modeling, we assess the strengths and weaknesses of the widely-used ordinal least squares and logit methods, conventional time series approaches, vector-autoregression models, event-history models, and a variety of techniques involved in rare-events analysis. Within computational modeling, we consider supervised classification methods and unsupervised clustering methods, and then look more specifically at the issue of using event sequences for forecasting. Finally, we will briefly consider some of the recent developments in computational network analysis.

1.1 The Development of Technical Political Forecasting

There is both a push and a pull to the current interest in technical political forecasting. As with many aspects of the Information Revolution of the past several decades, some of this is driven purely by technology and the availability of data, computational power and readily accessible software. We can do things that we couldn't do before, and clearly the impact of intellectual curiosity and technical itches to scratch is a factor.

But if the technological factors were the only driver, we could satisfy that by clicking on virtual cows and broadcasting our prowess on Facebook (http://www.wired.com/magazine/2011/12/ff_cowclicker/all/1). Instead, at the level of serious policy analysis—mess up counter-terrorism, and real people die, with no option to reset—the motivation is different: humans are appallingly bad at predicting human behavior. The pathbreaking studies of Philip Tetlock [161], who assessed the performance of 284 expert political forecasters in the assessments of more than 80,000 predictions over a 20 year period, found that their accuracy was barely better than chance, and the record for some of the most publicly-visible forecasters, who play to a crowd anxious for dramatic predictions, is actually worse than what could be achieved by a dart-throwing chimpanzee. Tetlock's results confirmed for the political realm—albeit in much greater detail—similar findings in economics and psychological assessment [94, 111] but were nonetheless a stunning indictment of the "subject matter expert" as a prognosticator.

Tetlock's work was particularly timely given the intellectual climate for the development of statistical and computational models of political violence. The earliest systematic studies of political violence using modern statistical methods and standards for data collection date from the middle of the twentieth century, in the work of British meteorologist Lewis Richardson [131]. This approach gained considerable momentum in the 1960s with the "behavioral revolution" in political science, which used the emerging technologies of social science statistics, digital computing, and machine-readable data to begin the systematic assessment of various theoretical propositions about political violence.

These found, for the most part, that the patterns in the qualitative "wisdom literature" were far less generalizable than their advocates believed—perhaps unsurprising in a world that had experienced two civilization-shattering wars in only three decades—with the results on the "democratic peace" [12, 15, 81] being the notable exception. This period saw the initial development of many large-scale databases relevant to the study of political violence, including the Correlates of War (http://www.correlatesofwar.org/) on international and civil war, Polity (http://www.systemicpeace.org/polity/polity4.htm) on state political characteristics, and early event data sets [8, 36, 99, 110] on political interactions in general. The data sets collected specifically for the study of political behavior were supplemented by readily-available global data on demographic, development, economic and trade characteristics of states provided by organizations such as the United Nations (UN), International Monetary Fund (IMF) and World Bank.

The U.S. Department of Defense Advanced Research Projects Agency (DARPA) funded a series of projects for the development of statistical forecasting in the 1970s [4, 44, 85], and parallel efforts, using both statistical and computational ("artificial intelligence") methods continued in the political science community under National Science Foundation funding [39, 76, 87, 160]. However, these early efforts were not particularly successful in the production of robust, timely, and accurate forecasts. In retrospect, neither the data nor methods available at the time were sufficient for the task. By the 1980s U.S. government efforts had largely ended, though the basic research in the academic community continued, as did some efforts by international non-governmental organizations [45, 75, 79, 132].

The U.S. government re-engaged with the development of technical political forecasting in the mid-1990s, motivated in part by the considerable progress in the development of methods and data since the earlier efforts, and by dramatic intelligence failures by human analysts on such events as the end of the Cold War [64, 65], state failure in Somalia and the former Yugoslavia, and the Rwandan genocide. The most conspicuous effort was the U.S. multi-agency State Failures Project [60, 61], later renamed the Political Instability Task Force (PITF) [74], a long-term collaboration between government analysts, contractors, and academic researchers. After some initial missteps involving efforts which were overly complex, PITF developed models that predicted various forms of political instability with 70–80% accuracy[1] with a 2-year time horizon and global coverage. In the mid-2000s, DARPA returned to the issue with the Integrated Conflict Early Warning System (ICEWS) [117, 118], which achieved similar levels of accuracy.

As a consequence of these various efforts, technical forecasting of political violence is now quite well developed and the subjects of numerous efforts in both North America and Europe [18, 27, 33, 82–84, 118, 165, 167]. While most of these studies look at violence and irregular political activity such as coups, a quantitative literature has also emerged on the topic of terrorism; extended reviews of this literature can be found in [66, 126, 127, 150, 169].

2 Data Sources

As noted in the previous section, one of the major factors fueling the development of computation methods for the analysis and forecasting of political violence has been the availability of machine-readable data sets. Collection of data sets has now been underway for close to five decades—the Correlates of War Project, for example, began in 1963—and are now the product of thousands of hours of careful research, refinement and coding, and have been used in hundreds of studies. In other cases—automated coding of atomic event data and analysis of the new social media—the

[1]Defined here in terms of the percentage of conflict-onset cases correctly classified.

collections are relatively new and unexplored, but potentially provide very large amounts of data in near-real-time. These data are generally available on the web, either through academic archives such as the Inter-University Consortium for Social and Political Research (http://www.icpsr.umich.edu/icpsrweb/ICPSR/) and Harvard's Dataverse Network (http://thedata.org/), through government and IGO sources such as USA.gov (http://www.usa.gov/Topics/Reference-Shelf/Data.shtml) and the United Nations Statistics Division (http://unstats.un.org/unsd/databases.htm) or—increasingly—through individual web sites established by the projects collecting the data.

In this section, we will briefly describe some general types of data that have been used in models of political violence, with comments on a few of the strengths and weaknesses of each type. As with any typology, not all of the data sets fit clearly into a single category, but most will.

2.1 Structural Data

Structural data refer to characteristics of a single political unit: typically this is a nation-state but it could also be a province or city, or a militarized group. At the nation-state level, classical structural measures include socio-economic variables such as population, literacy and infant mortality rates, gross domestic product (GDP), GDP per capita, percent urban population, ethnic composition, and the Gini index measure of income inequality. Specialized political science data sets such as Polity [104] and Freedom House [62] provide an assessment of a large number of dimensions of governance, often at a very high level of detail: for example the Institutions and Elections Project [130] has coded 127 characteristics of government and electoral institutions for all countries with populations over 500,000 from 1972–2005.

At the (militarized) group level, structural data is available for a wide variety of terrorism-related actors [6, 112, 153]. For instance, the Minorities at Risk (MAR) and Minorities at Risk Organizational Behavior (MAROB) datasets each provide extensive structural information on the social characteristics (e.g., ideology, religion, population) of politically active communal groups throughout the world [6, 112]. Finally, note that structural data can be used either directly or indirectly. GDP and population are good examples of direct measures: they are primarily used to adjust for the economic capacity and number of people in a country. Probably the best known indirect measure is infant mortality rate (IMR), which has consistently emerged as a one of the most important predictors of political instability. This is not due to any tendency of individuals whose children have died to immediately go out and revolt, but rather because IMR has proven, in multiple independent studies, to be a good indirect measure of a combination of poverty and the failures of a government to deliver basic services.

2.2 Dyadic Data

Dyadic data deal with *relations* between two entities. As above, these are usually states but could also be individuals or organizations. Trade data—notably the IMF's *Direction of Trade* [89]—are some of the most commonly used. Again, depending on the application, these can be used as either direct or indirect measures: trade is important in its own right, but measures like bilateral trade can also reflect the extent to which two states have common interests. Data on shared alliances, joint international organization (IGO) membership and military involvement (as allies or antagonists) are also readily available and are among the commonly used concepts used to test traditional international relations theories.

In recent years, the most important development in dyadic measures has been geospatial, which we discuss in Sect. 5.2. While the concept of political geography is quite old, only recently have political-geographic measures reached a level where they can be used systematically. However, political violence has a very strong geospatial component—shared borders are consistently the single most important predictor of whether states will engage in conflict, and overrides issues such as commonalities or differences in religion, language and culture [15], and high levels of terrorism tend to strongly cluster in both time and space—and these methods are gaining increasing attention [34, 35, 37, 129, 162].

2.3 Atomic Event Data

"Atomic" events data—usually called simply "event data" in the literature are basic units of political interaction coded from news sources that provide the date, source, target, event-type of an interaction. These have a long history in the quantitative analysis of conflict, with the early large data sets, the World Event Interaction Survey (WEIS) [109, 110] and the Conflict and Peace Data Bank (COPDAB) [8, 9] dating to the 1960s. While the original data were laboriously coded by (bored) students working with paper and microfilm, starting in the 1990s collection shifted to automated coding from machine-readable news wires [20, 67, 123, 138, 143], which resulted in data sets coded in near-real-time containing millions of events [72, 118, 143]. Along these lines, a number of more specialized event data sets have recently been developed to specifically code terrorism events, and employ a mixture of human and automated coding techniques to do so. Most notably, the National Counterterrorism Center's Worldwide Incidents Tracking System (WITS), the Global Terrorism Database (GTD), and the Terrorism in Western Europe Events Data set (TWEED) all provide wide coverage and detailed codings of terrorist attacks and related events [114, 120, 163].

Most event data sets use systematic typologies to reduce the nearly infinite variety of event descriptions possible in natural language reports to a manageable number of categories that a computational model can be process. All of these typologies have about 20 "macro" categories and around 100 more detailed subcategories. While the original event data collections focused almost exclusively on nation-states, contemporary systems provide higher levels of substrate aggregation, sometimes down to the level of individuals [68, 147]. While most studies in the past aggregated the events into a single cooperation-conflict dimension using a scale [73], more recent approaches have created composite events out of patterns of the atomic events [88, 140] or looked at sequences of events [50].

2.4 Composite Event Data

"Composite" data are those which code the characteristics of an extended set of events such as a war, crisis, or terrorist incident: the Correlates of War project (COW; http://www.correlatesofwar.org/) is the archetype; International Crisis Behavior (http://www.cidcm.umd.edu/icb/) is a more recent example as is the Global Terrorism Database [reference to **The Global Terrorism Database, 1970–2010** Gary LaFree, Laura Dugan, Chap. 1 in this volume]. And multiple characteristics of the incident are coded. Composite event data are typically coded from a number of reports of the event—as distinct from atomic event data, which generally looks at single sentences—and typically code a large number of characteristics of the event. At present, composite event data are usually coded by humans, though machine-assisted coding approaches are becoming increasingly prominent due primarily to automated "data field extraction" methods, able to rapidly locate information from text, such as the number of individuals killed or the amount of aid promised.

In some instances, it may also be possible to define composite events such as "civil war" by using patterns of the atomic events [88]. This would also make the differences between definitions used by various project unambiguous (or at least comparable) and allow the composite events to be easily constructed out of existing data sets rather than starting every new project from the beginning, and dramatically reduce the cost of this type of data.

2.5 Social Media and Other Unstructured Data Sources

The new social media—web sites, blogs, chat rooms, Twitter, Facebook, and other information easily available on the web—represent an emerging frontier in data collection. The advantages of these sources are clear: they can provide direct access to information about groups through the materials that they voluntarily make available, they are easily accessible at nearly zero cost, and with automated natural language processing, they can be analyzed in near-real-time. Social communication media such as Twitter and Facebook also provide extremely finely-grained, minute-by-minute data. In addition, a number of analysts have suggested that the social media themselves are changing the character of political mobilization, as demonstrated by the events of the "Arab Spring," particularly in Tunisia and Egypt.

Skeptics, however, have pointed out that social media also have some disadvantages. While access to the Web is increasing rapidly, it is still far from universal, and social media in particular tend to be disproportionately used by individuals who are young, economically secure, and well-educated. In areas with strong authoritarian regimes, notably China, there are substantial (though not uniformly successful) efforts to control the use of these media, and a government agency with even modest resources can easily create a flood of false posts, sites and messages. While there is *some* political content in these media, the vast bulk of the postings are devoid of political content—`OMG, Bieber fever!!!`—and what relevant content does exist may be deliberately or inadvertently encoded in a rapidly-mutating morass of abbreviations and slang almost indecipherably to conventional NLP software. (This contrasts to the news stories used to encode atomic event data, which generally are in syntactically correct English.) Finally, a number of analysts have argued that the critical communications development for political mobilization is the cell phone, both for voice and texting, rather than the Web-based media.

2.6 The Challenges of Data Aggregation

In most instances, conflict data comes pre-aggregated: no assembly required. However, researchers interested in analyzing specific events extracted form the text (which is common in event data studies as well as social network analysis), must make critical aggregation decisions in three areas to convert the raw text or 'event triplets' into a format suitable to their models of choice: (1) actors (2) actions, (3) temporal. (In contrast, techniques such as sentiment analysis and unsupervised clustering algorithms (see Sect. 4.2.2), are equipped to analyze text in its raw format.

2.6.1 Actors

Most major event data datasets—including WEIS, CAMEO, ICEWS, and VRA—code the source and target actors for each event in the data set. However, many of these actors may be irrelevant to the specific outcome-of-interest. For example, a study focusing on Israeli-Palestinian conflicts would not want to include events between Aceh rebels and the Indonesian army, as these are not relevant to the conflict of interest. Although excluding Indonesian rebel activity is obvious in this case, more difficult decisions exist, such as whether or not to include events between members of the Lebanese and Syrian armies, or the governments of the United States and Iran in a study of conflict between Israel-Palestine. Yonamine [170] provides a more detailed description of event data aggregation.

2.6.2 Actions

Event datasets use a numerical code to reflect the specific type of event that is occurring between the two actors. Since the numerical codes do not carry intrinsic value, researchers manipulate the code to reflect meaningful information. The majority of extant event data literature either scales all events, assigning them a score on a conflict-cooperation continuum or generates event counts reflecting the number of events that occur within conceptually unique categories. The Goldstein Scale [73], which is the most commonly used scaling technique within the event data literature [73, 77, 125, 136, 138], assigns a value from a -10 to $+10$ conflict/cooperation scale, with -10 reflecting the most conflictual events and 10 indicating the most cooperative.

Despite the preponderance of the Goldstein scale, a number of other studies [140, 141, 144] utilize count measures. Duvall and Thompson [51] put forth the first event data count model by placing all events into one of four conceptually unique, mutually exclusive categories: verbal cooperation, verbal conflict, material cooperation, material conflict. Although this count approach is more simplistic than scaling methods, [135, 144] find strong empirical results using this count method of action aggregation.

2.6.3 Temporal

Finally, scholars must temporally aggregate data in order to perform empirical analyses at levels appropriate for their theory or empirical models of choice. All of the previously mentioned event data sets code the exact day on which events occur. As the specific time-of-day that events occurred is not reported, events must at the very minimum be aggregated to the daily level [125, 135, 144], though weekly [26, 148], monthly [136, 165], quarterly [93], and annual level aggregations are common within the literature. By aggregated, we mean that the events occurring within the selected temporal length must be jointly interpreted. Common approaches are to calculate the sum or the mean of events that occur within the chosen temporal domain.

3 Statistical Approaches

Most of the work on forecasting political conflict has used statistical modeling, since this has a much longer history in political science than algorithmic and machine learning approaches. While the bulk of these studies have focused on simply interpreting coefficient estimates within the "frequentist" mode of significance testing, a method which has proven to have little utility for predictive modeling [165], more recent work has taken prediction seriously, both using classical time series models and more recently a substantial amount of work using vector autoregression (VAR) models. In addition, recent work has focused on one of the most challenging aspects of forecasting, particularly when applied to counterterrorism: the fact that these events occur very rarely. While this presents a serious challenge to the method, a number of sophisticated methods have been developed to deal with it.

3.1 Cross-Sectional Regression and Logit

By far the most common statistical model in contemporary research—probably accounting for at least 80% of the studies—are variants on ordinary least squares regression and the closely-related logit model. Ordinary least squares regression uses equations of the form

$$Y_i = \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \varepsilon_i = x_i'\beta + \varepsilon_i, \qquad i = 1, \ldots, n, \tag{1}$$

Logistic regression, in contrast, is used to predict a values between 0 and 1—typically interpreted as a probability—and does this by using an equation of the form

$$Y_i = \frac{e^z}{e^z + 1} = \frac{1}{1 + e^{-z}} \tag{2}$$

where the variable "z" is usually defined as

$$z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots + \beta_k x_k, \tag{3}$$

While these approaches have been used extensively in the academic literature, their practical utility is, unfortunately, quite limited due to a reliance on the "frequentist" significance testing approach. Extended discussions of this issue can be found in [3, 70, 96, 137] but briefly, significance testing was originally developed in the early twentieth century to study problems where the "null hypothesis" of a variable having no effect was meaningful (the archetypical example is whether a new medicine has a different effect than a placebo). In some political science applications, this is still valid—for example, determining whether a forthcoming election is a "statistical tie" based on opinion polling—but in most models a variable is generally not included unless there are theoretical (or common sense) reasons to assume that it will have at least *some* effect. Because this is all that the significance test is assessing—a non-zero effect, not an effect that has some meaningful impact—the information it provides is limited and all but negligible for predictive problems [165].

Although the frequentist approach can be useful in weeding out variables that might seem to be important but in fact are not, in contemporary models that tend to be complex, even this should be interpreted with caution. Linear models are particularly susceptible to problems of "colinearity" when the independent variables x_i are correlated—as is often the case, particularly in models where a researcher is experimenting with several measures of a diffuse concept such as “economic development” or “authoritarianism”—and in the extreme (but not uncommon) case, colinearity can cause the estimated direction of a variable to be the opposite of what it would be in the absence of the colinearity. Specification error—leaving out a variable that in fact has a causal link to the dependent variable Y_i—is another issue, and has the unfortunate impact of inflating the estimated coefficients of the variables which are in the model.

Finally, the existing literature tends to use complex models, often with a dozen or more variables, despite the fact that a large set of literature, going back as early as the 1950s [111], with the lesson repeated by [3] about a decade ago, indicates that simple models are preferable, for at least two reasons. First, complex models tend to “fit the error,” providing overly-optimistic assessments of the accuracy of the model based on the existing data, with those estimates *decreasing* the accuracy of the model once new data are available. Second, the nearly inevitable presence of colinearity in non-experimental social science variables tends to increase the variance of the estimated coefficients as the number of independent variables increase.

On the positive side, the strengths and weaknesses of linear models have been studied for decades and are well-understood. The problems noted above are widespread but not inevitable, and models which have confirmed the same result in a wide variety of formulations—for example on the impact of geographical contiguity, joint-democracy and trade on interstate conflict, and the impact of economic development on political instability—are probably robust. There has also been a gradual shift away from the classical frequentist approaches to the “Bayesian” approach [71, 92], which simply uses the data to adjust estimate of the distribution of the coefficient, rather than employing either/or significance tests, as well as methods such as matching cases within a sample [121] which can improve the accuracy of the coefficient estimates.

3.2 Classical Time Series

Classical time series models predict the future value of a continuous variable based on some combination of the past values, usually with a focus on reducing the systematic error in the model to “white noise”, that is, errors with a mean of zero and the correlation between any ε_t and ε_s is equal to zero. The classical reference is [22], an open-source introductory is available at http://statistik.mathematik.uni-wuerzburg.de/timeseries/, and an extended treatment can be found in [78]. Examples in the literature on forecasting political conflict include [79, 125, 145, 146]; [124] provides a general reference with respect to applications in political science.

The notation AR(“p”) refers to the autoregressive model of order “p”. The AR(“p”) model is written

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \tag{4}$$

where $\varphi_1, \ldots, \varphi_p$ are the parameters of the model, c is a constant and ε_t is white noise. In most applications, the error terms ε_t are assumed to be independent, identically-distributed, and sampled from a normal distribution.

The notation $MA(q)$ refers to the moving average model of order q:

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{5}$$

where the $\theta_1, \ldots, \theta_q$ are the parameters of the model, μ is the expectation of X_t (often assumed to equal 0), and the $\varepsilon_t, \varepsilon_{t-1}, \ldots$ are white noise error terms.

The notation $ARMA(p, q)$ refers to the model with p autoregressive terms and q moving-average terms, where a moving average is defined by . This model contains the $AR(p)$ and $MA(q)$ models,

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{6}$$

In practical terms, classical time series models are useful for phenomena where the current value of the variable is usually highly dependent on past values. This "incremental behavior" is characteristic of a great deal political behaviors, notably public opinion, budgeting, and most conflict behavior: once a civil war or insurgency gets going, it is likely to continue for a while, and once peace is firmly established, that is also likely to the maintained. In such situations, the best predictor of the variable at time t is its value at time $t - 1$ Consequently, these models are very widely used.

Incremental models, however, have a critical weakness: by definition, they cannot explain sudden change, whether isolated incidents or the onset and cessation of protracted behaviors. An autoregressive model can have a very good overall predictive record but still miss the beginning and end of a conflict, and these may be of greatest interest to decision-makers.

The classical time series literature places a great emphasis on the characteristics of the error terms. While these are treated in the literature as random variables, in practical terms, much of the "error" is simply the effects of variables that were not included in the model. Because most social and demographic indicators and many economic indicators are also very strongly autoregressive—for example indicators such as infant mortality rate and literacy rarely change more than a percentage point year to year—these "errors" will strongly correlate with their lagged values, hence the interest in the MA and $ARMA$ models. Unfortunately, disentangling the effects of autoregressive variables and autoregressive errors is extraordinarily difficult in many circumstances, which in turn has led to the development of vector-autoregressive models, discussed in Sect. 3.3.

Finally, one will occasionally encounter questions about "co-integration" and "non-stationarity" in time series models [113]. This issue generally arises in situations where exponential growth is occurring in the variable, as can happen in situations of hyper-inflation, individuals who are really lucky in their choice of stocks or really unlucky in their choice of real estate markets. While conflict data will switch between *modes*—peace, crisis and war—and this is a form of non-stationarity (which refers to a change in the statistical characteristics of the errors), this is not an issue of co-integration. Unfortunately, the tests for co-integrated time series have a very high probability of incorrectly identifying it in situations where the autoregressive coefficient is close to 1.0—that is, the variables change very little from time period to time period—and hence conflict time series can easily be incorrectly identified as co-integrated.

A second common problem encountered in time series with conflict data are "time series cross sections" (TSCS), where data are collected across a number of cases across a period of time (for example, incidents of terrorism by country-year). Because these cases are not independent—in particular, contemporary trans-national terrorist groups operate in multiple countries—this will affect the estimates of the coefficients and methods that do not adjust for this can give misleading results [10, 11].

3.3 Vector Autoregression Models

A relatively recent modification of the classical time series approach has been vector autoregression (VAR) models. Originally developed by Sims [154] and used primarily for financial forecasting, VAR models were also used in some of the earliest applications of sophisticated econometric modeling to assessment of terrorist strategies [52–54, 58] and have been used in a number of policy-oriented studies in the post-9/11 period [28, 29, 55–57, 59], as well as more general applications to conflict early warning [26, 30, 63].

A p-th order VAR, denoted $VAR(p)$, is

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t \tag{7}$$

where c is a $k \times 1$ vector of constants, A_i is a $k \times k$ matrix (for every $i = 1, \ldots, p$) and e_t is a $k \times 1$ vector of error terms.

The distinctive characteristic of VAR approaches is their emphasis on response functions rather than on the interpretation of individual coefficients. VAR takes as a given the fact that if the specified model is correct, the relations between the variables—or technically speaking, their colinearity—will make the uncertainty in the estimated coefficients very large even when the predictive accuracy of the system as a whole is quite high. Consequently, rather than trying to interpret those coefficients, an analyst looks at the effects—including error bands—of an exogenous change propagating over time through a series of inter-dependent variables. The downside of VAR is that it is very data-intensive and requires data measured consistently across a large number of time points.

A more recent variation on VAR models has been the addition of "switching" components, which allow for the impact of distinct regimes on the behavior of the system [155]. While originally developed for the study of foreign exchange markets, this is clearly applicable to political conflict situations where the system is likely to behave differently in times of peace than during times of war. Bayesian versions have also been developed [26] and applied to cases such as the conflict in the Middle East [25, 27].

3.4 Event-History and Survival Models

A different approach to conventional time series and VAR, but somewhat similar to the probabilistic interpretation of logit models, is the event-history or survival model approach [24, 86]. Rather than trying to predict the value of a behavior at a specific time, these models predict *probabilities* of events, and in particular look at the change in probability over time of the next event occurring. The models were thoroughly developed in the bio-medical literature to model disease prognosis—hence the term "survival"—prior to their application in political science. See [23, 41, 42] for applications to political conflict.

The "survival function" is defined as $S(t) = \Pr(T > t)$ where t is time, T is a random variable indicating when the event—for example the next terrorist attack, or the end of a terrorist organization—occurs, and Pr is probability. In many models, the focus of the analysis is the "hazard function"

$$\lambda(t)\,dt = \Pr(t \le T < t + dt \mid T \ge t) = \frac{f(t)\,dt}{S(t)} = -\frac{S'(t)\,dt}{S(t)} \tag{8}$$

which the event rate at time t conditional on survival until time T, $T \ge t$. From these functions, information such as the average number of events in a given period or average lifetime can be derived.

Of particular interest in survival analysis is the overall shape of the hazard function. For example, a survival function which levels off indicates that once an interval has occurred without an event occurring, it is less likely to occur again, whereas a survival function that declines linearly means that the occurrence or nonoccurrence provides no information. Other forms are also possible—for example many biological organisms exhibit a U-shaped hazard function, meaning that mortality is high in the early stages of life, then becomes low for a period of time, then rises again as the organism reaches its natural age limit.

A key downside of survival models is the difficulty that humans have in working with probabilities [94, 161]. In addition, the models are relatively new to the study of political conflict, due to earlier constraints of computational feasibility and the availability of appropriate estimation software. However, with these constraints removed, there are likely to be more such applications in the future.

3.5 *Rare-Events Models*

In most studies of political violence, the occurrence of violence is rare relative to the instances wherein violence did not occur. This poses a number of challenges to the empirical analysis of terrorism. When the absolute number of observations (or actual terrorism-events) in a sample is low, the rarity of terrorism can bias coefficient estimates and underestimate the probability of terrorist events. Even When one has a large sample of terrorist incidents, the relative rarity of these events and the preponderance of zeroes that this rarity entails can pollute samples with "non-event" observations that effectively have zero likelihood of *ever* experiencing terrorist events. This "zero inflation" contaminates terrorism-event data with multiple (latent) data generating processes, which when unaccounted for, can similarly bias estimates and predicted probabilities—often again by understating the true occurrence of terrorism.

Within logit models of binary terrorism data (see Sect. 3.1 above), scholars have shown that information content deficiencies of these sorts are primarily a function of two dynamics: the total number of observations within one's sample ("N"), and the mixture of 1's and 0's therein [97]. In such instances, estimates will be biased downward due to an underestimation of the variance around $Pr(Z|Y = 1)$ and common methods for computing predicted probabilities will exacerbate this bias through their disregard for estimation uncertainty [97]. While such concerns may not at first appear to be detrimental to 'big data' analyses, they are nevertheless highly relevant to large-N conflict studies given the growing importance of in-sample and out-of-sample forecasting techniques in the discipline [12, 166]. Indeed, even for large-N event data sets, the practices of subdividing data sets into smaller training, test and validation-data sets will frequently necessitate that scholars address these concerns in order to achieve stable and comparable estimates of conflict forecasts across sub-samples.

SOMA behavior models and CAPE models—which we discuss in more detail in Sect. 4.4 below—provide one potential solution to this problem [107, 157]. These models account for behavioral changes, and have been found to be especially useful in modeling the behaviors of terrorist groups under the specific rare-event contexts outlined above. Within the political science literature, an alternative approach is the 'rare-events logit' model, which addresses these challenges by treating rare-events as a sampling design problem. Provided that the dataset has a representative (or complete) sample of actual terrorism events, this approach calls for researchers to then augment their samples of observed terrorism events with a comparable, random sub-sample of all corresponding non-events. When this strategy is consistent with the assumptions of 'case-control' or 'case-cohort' sampling approaches, unbiased coefficient estimates and predictions can be derived by adjusting logit estimates and predicted probabilities for the true ratio of events to non-events within the population (or an estimate thereof) during model estimation [97]. This makes such an approach particularly advantageous for empirical studies of terrorism that (i) employ in-sample and out-of-sample forecasting techniques (ii) focus on a specific region or country over a short period of time, or (iii) require additional variable coding or data collection that, due to time or money constraints, cannot be completed for an entire sample of interest.

A second challenge to the analysis of rare-event data arises from the corresponding preponderance of zero (non-event) observations within such data sets. As recent studies note [126], many zero-observations in terrorism-event data sets correspond to country or region cases that have no probability of experiencing a terrorist-event in any period of interest. In these instances, empirical analyses of conflict-event data risk conflating two distinct types of zero-observations; one for which the probability of terrorism is non-zero but terrorism-events nevertheless didn't occur, and one wherein the probability of terrorist-events is consistently zero. Given that the latter set of zero-cases often arise as the result of covariates that correlate or overlap with one's primary independent variables of interest, ignoring zero-inflation processes of these sorts not only leads to an underestimation of terrorist events generally, but can also bias specific coefficient estimates in indeterminate directions. To correct for these biases, one must conditionally model *both* the zero-inflation process *and* the main outcome of interest.

Zero-inflated mixture-models specifically address these very problems, most notably within the contexts of event-count models such as Poisson or negative binomial estimators. In essence, these models employ a system of two equations to estimate the combined probability of an observation (i) being inflated and (ii) experiencing an event-outcome of interest—usually by including separate but overlapping covariates as predictors for each respective equation. For example, one could expand the logit model presented in Eq. 13 above to the zero-inflated logit approach by incorporating a second, "inflation-stage" logit equation as so,

$$f(z,w) = \left[\frac{1}{1+e^{-w}} + \left(1 - \frac{1}{1+e^{-w}}\right)\left(1 - \frac{1}{1+e^{-z}}\right)\right]^{1-Y_i} * \left[\left(1 - \frac{1}{1+e^{-w}}\right)\left(\frac{1}{1+e^{-z}}\right)\right]^{Y_i} \tag{9}$$

where the variable "w" represents the additional set of inflation-stage covariates;

$$w = \beta_0 + \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \cdots + \beta_k v_k, \tag{10}$$

which may or may not overlap with z [13]. Zero-inflated models thereby add an additional layer of nuance to the empirical modeling of conflict-events by estimating *both* the propensity of ever experiencing an event of interest *and* the likelihood of experiencing an event of interest conditional on being able to do so. This allows one to use ex-ante observable and theoretically informed covariates to account for the probability that a given zero observation is 'inflated', and to then probabilistically discount these zeroes' leverage within one's primary analysis without dropping these observations entirely. While such zero-inflated modeling approaches have been most extensively applied to political violence count-data [14, 40, 126], zero-inflated models have also recently been developed and applied by conflict-researchers to a variety of other limited dependent variables [83, 159, 168].

4 Algorithmic Approaches

Although more traditional "statistical" models still dominate quantitative studies of political conflict, "algorithmic" approaches have proven effective, thus gaining momentum not just within political science [21, 139] but also in other disciplines. For example, computer scientists have developed the CONVEX [157], CAPE [106], and SOMA [107] tools to forecast terrorist group behavior. While we will follow common practice in using the "statistical" vs. "algorithmic" distinction to differentiate between methodologies, there is overlap between the two definitions. For example, linear regression is consider a canonical "statistical" approach, but as we describe in Sect. 4.1.1, it is also a straightforward example of a supervised linear "algorithm".

In general, by algorithmic approaches, we refer to specific models (such as neural networks or random forests) or techniques (like bagging and boosting) that attempt to leverage computational power to specify and train models through iterative resampling techniques and to build and assess out-of-sample predictions, rather than obtaining a single estimate of the model coefficients. Algorithmic approaches can provide a number of benefits over statistical models and are particularly relevant to forecasting terrorism for at least the following four reasons.

First, machine learning algorithms are often better suited than many traditional statistical models at handling 'big data' data sets with large numbers of independent variables that potentially exceed the number of observations. Second, these algorithms are also less dependent on rigid assumptions about the data generating process and underlying distributions. Third, as opposed to some statistical models, many machine learning algorithms were specifically designed to generate accurate predictions, and do this exceedingly well. Finally, a number of the algorithmic approaches approximate the widely used qualitative method "case-based reasoning" [95, 108, 116] which match patterns of events from past cases to the events observed in a current situation, and then use the best historical fit to predict the likely outcome of the current situation; [134, Chap. 6] gives a much more extended discussion of this approach. This similarity to the methods of human analysts accounted for these methods originally being labeled "artificial intelligence" in some of the early studies.

Indeed, major trends in the empirical study of political violence, such as the 'big data' revolution and an increasing interest in predictive models, mean that algorithmic approaches will likely become increasingly popular in the coming years. In the following sections, we address some of the most relevant machine learning algorithms for forecasting political violence. Following standard practices, we divide algorithmic approaches into two general, though not mutually exclusive categories, supervised and unsupervised algorithms, with an additional discussion of sequence analysis techniques.

4.1 Supervised Cross-Sectional Classification Methods

Using the language of political science, supervised algorithms require an observed dependent variable value—be it binomial, continuous, or categorical—and at least one independent variable for each observation. The vast majority of quantitative studies of political violence employ data that meet the requirements of supervised algorithms. In the following section, we discuss the two supervised algorithmic approaches that have received the most attention in the political violence literature, linear models and neural networks, and also discuss tree-based approaches that have yet to take hold in the study of political violence but have proven successful at forecasting in other disciplines.

4.1.1 Linear Approaches

Though generally thought of as a statistical approach, liner regression is also one of the most common and straightforward examples of a supervised learning algorithm. The dominance of linear regression among the broader group of existing supervised linear approaches is not fully unwarranted, since linear regression contains nice attributes, such as being "BLUE"—the Best Linear Unbiased Estimator. Additionally, and perhaps more importantly, linear regression approaches are widely taught, straightforward to interpret, and easy to implement with one line of code (or a few clicks of a drop down menu) in virtually every statistical package.

Consistent with the widespread use of the linear method, other supervised approaches similarly utilize a line (in <3 dimensions) or a hyper plane (in >2 dimensions), such as linear discriminant analysis (LDA) and support vector machines (SVM). LDA is a supervised linear classification algorithm that produces nearly identical results to linear regression when Y takes on two classes, but can outperform multinomial logit and probit models as the number of classes exceeds 2 [80, pp. 106–114]. SVMs are an additional linear classification supervised algorithm that is most commonly used in problems such as text classification when the number of covariates (i.e. words in the text example) exceeds the number of observations. [49, 142] demonstrate the strength of SVMs in classifying articles for the Militarized Interstate Disputes 4.0 (MID4) project, while [90] uses SVM to extract events of terrorism from text.

4.1.2 Neural Networks

Aside from traditional linear regression approaches, neural networks have arguably received the most attention in quantitative political violence literature among supervised learning algorithms due almost entirely to Beck, King and Zeng's important article on forecasting interstate conflict [12]. (Neural networks can also be applied in an unsupervised setting: see [98].) Neural networks were originally created to model complex data characterized by a large number of nodes that receive and process input functions and then transfer information to other nodes, such as neurons in the case of the human brain.

Although most work with neural networks seems far removed from terrorism, [12] articulately explain how supervised neural networks are not only an appropriate algorithmic approach to predicting violence but can also be applied as a straightforward extension to logistic regression. Like a logistic regression, neural networks can be applied to a traditional TSCS dataset with a binary dependent variable to generate predicted probabilities that are interoperated identically to the π parameter of logistic regression. However, the primary advantage of a neural network approach is that they are able to account for the potential of 'massive nonlinear interaction effects' that may causally link the independent variables to the outcome of interest without having to directly specify interactive or non-linear terms to the model as required in a logistic. [12] demonstrate that the neural network approach consistently outperforms logistic regression in out-of-sample accuracy measures. Though we are unaware of neural networks being applied in studies of terrorism, it is likely that doing so could yield similar improvements in predictive accuracy.

4.1.3 Tree-Based Algorithms

Though yet to gain traction in the political violence literature, tree-based approaches [80, pp. 305–317, 587–604] are commonly used to generate accurate predictions in a host of other disciplines, including finance, medicine, computer science, and sociology. In brief, tree-based algorithms operate by iteratively partitioning the data into smaller sub-sets based on a break-point in a particular independent variable. This process is repeated with the goal of creating bins of observations with similar Y_i values (for continuous data) or class (for categorical data).

We highlight three important factors that contribute to the strong predictive accuracy of tree-based approaches. First, trees can be used to forecast both continuous (i.e. "regression trees") and binomial (i.e. "classification trees") dependent variables. Second, tree-based approaches, such as random forests, are not sensitive to degrees of freedom and can handle more independent variables than observations. Third, leading tree-based approaches incorporate iterative re-sampling, weighting, and model averaging strategies like bagging and boosting techniques [156], which tend to enhance accuracy and stability vis-à-vis other supervised learning classification and regression algorithms. In social science applications of tree-based algorithms, [16] and [17] demonstrates that random forests can help generate accurate forecasts of violent crime rates in the United States. Despite the scarcity of tree-based approaches in political science, scholars in other disciplines ranging from ecology (see [43]) to transportation studies (see [158]) have generated accurate predictions using this approach. As the quantitative political violence literature continues to progress, it will behoove scholars to continue experimenting with supervised forecasting algorithms that have demonstrated their value in other disciplines. We believe that tree-based approaches, especially boosted regression trees (BRTs) and boosted random forests, are particularly useful for scholars attempting to forecasting terrorism.

4.2 Unsupervised Methods

Unlike supervised learning algorithms that train a model based on relationships between a matrix of covariates and a corresponding vector of observed dependent variables for each observation, unsupervised approaches are applied to datasets for which dependent variables are 'latent' and therefore not directly provided. Though the concept of unsupervised algorithms may seem abstract, useful applications exist in the study of political violence. We describe two types of particularly relevant unsupervised learning algorithms in the following sections dimensionality reduction and clustering approaches. (Hidden Markov Models (HMMs) can also be classified as an unsupervised learning algorithm, but we present it in Sect. 4.3 as a sequence-based approach.)

4.2.1 Dimension Reduction

Factor analysis is a general algorithmic approach to reduce the dimensionality of matrix of potential covariates by identifying latent attributes. Principal components analysis (PCA) is a more restricted form of factor analysis and [91] is among the oldest (circa 1901) and most commonly employed dimension reduction tools in the social science [69]. In brief, the PCA algorithm works by analyzing relationships among the covariates and using these to create a new set of orthogonal variables called "principal components" that reflect the latent attributes and can be used in place of the original set of covariates in a quantitative model. Certain models, like linear regression, are sensitive to degrees of freedom and cannot operate when a dataset contains more covariates than observations. In these instances, a researcher can implement PCA or a similar approach to sufficiently reduce the dimensionality of covariates to enable the linear regression to converge [2].

4.2.2 Clustering

Clustering approaches such as k-means and Latent Dirichlet Allocation are similar to dimension reduction in that they attempt to identify latent classes amongst a set of observations, but differ in that they identify discrete, rather than continuous solutions like PCA and FA [69]. Within the machine learning literature, k-means approaches are among the most commonly used clustering algorithms, though their application to the study of political violence has been scarce [80, pp. 509–520].

However, [139] demonstrate that k-means can successfully identify latent clusters within event data that identify phases of violence in the Middle East.

Latent Dirichlet Allocation [19] is another clustering algorithm that is primarily applied to raw text. In the typical Latent Dirichlet Allocation application to document classification, each document is assumed to be a mixture of multiple, overlapping *latent topics*, each with a characteristic set of words. Classification is done by associating words in a document with a pre-defined number of topics most likely to have generated the observed distribution of words in the documents.

The purpose of LDA is to determine those latent topics from patterns in the data, which are useful for two purposes. First, to the extent that the words associated with a topic suggest a plausible category, they are intrinsically interesting in determining the issues found in the set of documents. Second, the topics can be used with other classification algorithms such as logistic regression, support vector machines or discriminant analysis to classify new documents.

Despite the surface differences between the domains, the application of Latent Dirichlet Analysis to the problem of political forecasting is straightforward: it is reasonable to assume that the stream of events observed between a set of actors is a mixture of a variety political strategies and standard operating procedures (for example escalation of repressive measures against a minority group while simultaneously making efforts to co-opt the elites of that group). This is essentially identical to the process by which a collection of words in a document is a composite of the various themes and topics, the problem Latent Dirichlet Analysis is designed to solve. As before, the objective of Latent Dirichlet Analysis will be to find those latent strategies that are mixed to produce the observed event stream. These latent factors can then be used to convert full event stream to a much simpler set of measures.

The Latent Dirichlet Analysis approach is similar in many ways to the hidden Markov approach (Sect. 4.3). In both models, the observed event stream is produced by a set of events randomly drawn from a mixture of distributions. In an HMM, however, these distributions are determined by the state of a Markov chain, whose transition probabilities must be estimated but which consequently also explicitly provides a formal sequence. An Latent Dirichlet Analysis, in contrast, allows any combination of mixtures, without explicit sequencing except to the extent that sequencing information is provided by the events in the model.

4.3 Sequence Development: Hidden Markov Models

Hidden Markov models (HMM) are a type of stochastic signaling model that has become increasingly popular in computing the probability that a noisy sequence was generated by a known model, most notably in speech recognition and protein sequence comparison. A detailed discussion and application of the method to the problem of forecasting terrorism can be found in [123] in this volume; other applications to political forecasting include [21, 133, 135, 144].

An HMM is a variation on the well-known Markov chain model, one of the most widely studied stochastic models of discrete events. Like a conventional Markov chain, a HMM consists of a set of discrete states and a matrix $A = a_{ij}$ of transition probabilities for going between those states. In addition, however, every state has a vector of observed symbol probabilities, $B = b_j(k)$ that corresponds to the probability that the system will produce a symbol of type k when it is in state j. The states of the HMM cannot be directly observed and can only be inferred from the observed symbols, hence the adjective "hidden."

In empirical applications, the transition matrix and symbol probabilities of an HMM are estimated using an iterative maximum likelihood technique called the Baum-Welch algorithm which finds values for the matrices A and B that locally maximize the probability of observing a set of training sequences. Once a set of models has been estimated, they can be used to classify an unknown sequence by computing the probability that each of the models generated the observed sequence. The model with the highest such probability is chosen as the one which best represents the sequence.

The application of the HMM to the problem of classifying international event sequences is straightforward. The symbol set consists of the event codes taken from an event data set such as IDEA [123] or CAMEO [143]. The states of the model are unobserved, but have a close theoretical analog in the concept of crisis "phase" [101]. Different political phases are distinguished by different distributions of observed events from the event ontology. Using CAMEO coding as an example, a "stable peace" would have a preponderance of cooperative events in the 01–10 range which codes cooperative events; a crisis escalation phase would be characterized by events in the 11–15 range (accusations, protests, denials, and threats), and a phase of active hostilities would show events in the 18–22 range, which codes violent events.

An important advantage of the HMM is that it can be trained by example rather than by the deductive specification of rule. Furthermore, HMMs require no temporal aggregation. This is particularly important for early warning problems, where critical periods in the development of a crisis may occur over a week or even a day. Finally, indeterminate time means that the HMM is relatively insensitive to the delineation of the start of a sequence: It is simple to prefix an HMM with a "background" state that simply gives the distribution of events generated by a particular source (e.g. Reuters/IDEA) when no crisis is occurring and simply cycle in this state until something important happens.

4.4 Sequence Analysis: Sequence Matching

Sequence analysis is a method that assesses the degree of similarity between two or more $1xN$ vectors, commonly populated by numerical or categorical inputs. Scientists first developed the approach as a tool to locate matches between different genomic D.N.A. sequences [31, 46] and it has since been applied to various disciplines including sociology [1] and demographics [102]. Although existing

Table 1 Archetype sequence the precedes an attack against the base

Variable	Week$_{t-4}$	Week$_{t-3}$	Week$_{t-3}$	Week$_{t-1}$	Week$_t$
Level of civilian casualties	1	2	3	4	N/A
Number of ambush attempts against patrol units	3	3	3	0	N/A
Attack on base?	No	no	No	No	Yes

studies using sequence analysis primarily focus on matching rather than forecasting, Martinez et al. [106], Silva et al. [157], and D'orazio et al. [50], do utilize sequence analysis-based algorithms for the explicit goal of prediction.

4.4.1 Archetypal Sequence Matching

D'orazio et al. suggest an archetype-driven approach to sequence analysis with the explicit goal of predicting political violence that uses sequence analysis to build features on which a statistical or algorithmic model produces out-of-sample forecasts. For this archetype approach be succeed, it requires the existence of a distinct pattern of events (i.e. an archetype) that tends to precede a given outcome-of-interest. If such an archetype exists, then a model may be able to forecast the outcome of interest based on the extent to which a sequence whose outcome is unknown (for example, events being observed in real time) is similar to the archetype sequence. The process of using this archetype-driven sequence analysis approach for prediction follows four main steps:

1. Define the archetype(s)
2. Calculate distances
3. Train a model
4. Build out-of-sample forecasts

To more clearly explain this approach, we use a hypothetical example to demonstrate each of the four steps above. Imagine our goal is to predict a binary variable indicating whether or not an American military base will be attacked in the upcoming week. Looking back, officers notice patterns in the level of civilian casualties (scale of 1–4, 4 being most severe) and the number of attempted ambushes on patrol missions that have occurred in peaceful weeks that precede an attack on the base. Table 1 completes <Step 1> as it reflects the pattern of events that tends to precede an attack against the base in terms of archetypical sequences. If officers observe 4 week progressions in real time that are sufficiently similar to the archetypical sequence of the level of civilian casualties and the number of ambush attempts on patrols, then they can confidently predict that an attack against the base will occur in the next week.

Just how similar must the observed sequences be to the archetypical sequences in order to predict that the following week will experience an attack against the base? To address this, a researcher must first complete <Step 2>, calculating the mathematical degree of similarity (i.e. distance) between the archetypes and

$$Distance_{civilian_casualty} = \sqrt{\sum_{i=1}^{4} (Archetype_{civiliancasulaty(i)} - Sequence_{civiliancasualty(i)})^2}$$

$$Distance_{ambush_attempts} = \sqrt{\sum_{i=1}^{4} (Archetype_{ambushattempts(i)} - Sequence_{ambusattemps(i)})^2}$$

Fig. 1 Building co-variates with Euclidean distance

set of training sequence of equal length (i.e. 4). Euclidean distance is one of the most common and robust measures, though many other distance measures exist. Figure 1 demonstrates how Euclidean is applied to calculate distances–which serve as the covariates in `<Step 3>`–between training sequences and the archetypical sequences.

The two distances, $Distance_{civiliancasualty}$ and $Distance_{ambushattempts}$ are calculated for every observation (i.e. 4 week sequence) in the training set. To complete `<Step 3>`, we choose a logistic regression, which is suitable to a parsimoniously specified model with a binary dependent variable. To train the logistic model, estimate the β values that maximize the likelihood function below.

$$L(\beta|y) = \prod_{i=1}^{N} \frac{n_i!}{y_i!(n_i - y_i)!} \pi_i^{y_i} (1 - \pi_i)^{n_i - y_i} \tag{11}$$

To complete `<Step 4>` and build actual forecasts based on the archetype-driven sequence analysis approach, we apply the two β estimates that result from (Eq. 6) to the logistic regression formula in order to calculate $f(z)$, which reflects the likelihood that the week following the 4-week period used to generate the $Distance_{civiliancasualty}$ and $Distance_{ambushattempts}$ distances from the archetypes will experience an attack on the base.

$$z = \beta_0 + \beta_1 Distance_{civiliancasualty} + \beta_2 Distance_{ambushattempts} \tag{12}$$

$$f(z) = \frac{1}{1 + e^{-z}} \tag{13}$$

To the extent that archetypical sequences that tend to precede an outcome-of-interest exist, sequence analysis may be an effective tool of prediction that can provide leverage over other approaches.

4.4.2 Convex Algorithms

In addition to the sequential archetype approach, Martinez et al. [106] put forth two alternative, sequence-based algorithms (Convex$_N^k N$ and ConvexMerge) that have proven successful at predicting terrorist activities. The key difference between these algorithms and the previously discussed archetypical approach is that the former do not assume the presence of an archetype. Thus, rather than calculating the distances between out-of-sample sequences with unknown outcomes and an archetypical sequence, the Convex algorithm calculates the various distance measure between each out-of-sample sequence and *all* in-sample sequences. In order to form out-of-sample predictions, the Convex algorithm determines the K-number of the most similar sequences based on various distance measures. If all of the k-nearest neighbor sequences have the same dependent variable, then the predicted value for that out-of-sample sequence is assigned as the dependent variable value of the single nearest neighbor. When this occurs, the $\mathrm{Convex}_N^k N$ and ConvexMerge approaches generate identical predictions. However, when the dependent variable values of the nearest neighbor sequences are not identical, $\mathrm{Convex}_N^k N$ rounds to the nearest integer while ConvexMerge assigns more value to more similar sequences. In addition to Martinez et al., Sliva et al. [157] apply both variants of the Convex algorithm to terrorist group behavior.

5 Network Models

We label our final category "network models:" these are approaches that specifically consider the relationships between entities, either according to their interactions (SNA models) or location (geospatial models). As with many of the algorithmic applications, network models have only become feasible in the last decade of so, as they require very substantial amounts of data and computational capacity. Consequently the number of applications at present is relatively small, though these are very active research areas.

5.1 Social Network Analysis Models

Social network analysis (SNA) in a type of graph theory that models systems in terms of nodes, which are the unit of analysis and can be anything from states in the international system to members of a terrorist cell, and edges that connect the nodes. The edges can reflect any type of interactions that occur between the nodes, such as levels of international trade if states are the nodes or amount of communication if the nodes reflect members of a terror cell. Though still an emerging method, SNA is becoming particularly attractive approach for studying terrorism [122]. For example, [7] use SNA to model the organizational structure of the terrorist groups responsible for the 2008 Mumbai attacks and [47] develop an edge-based network to forecast the source and target locations of transnational terrorist attacks. We speculate three potential explanations for its rapid growth.

First, terrorist scholars are often interested in information regarding within-group dynamics of terrorist organizations—such as chains of command, lines of communication, and policy preferences of leadership—which SNA may reveal more effectively than statistical or algorithmic approaches. Second, the rise of social networking platforms like Facebook, Twitter, less publicized online forums, and any other modes of communication open source or otherwise, has greatly increased the amount of information that can be used to populate a social network. Third, SNA has been almost simultaneously embraced by the intelligence, policy, and academic communities—a rare feat for a new methodology. These factors, as well as its ability to deliver unique insights, is likely to continue driving its growth among quantitative studies of terrorism.

5.2 Geo-spatial Models

Similar to our discussion of temporal dynamics above (Sects. 3.2 and 3.3), the incorporation of spatial dynamics into statistical models of terrorism can often enhance modeling and forecasting capabilities. Political violence frequently clusters or diffuses in geographic space [35, 164, 166], and failure to account for these dependencies can bias inferences [5, 164]. Spatial statistics allow one to model the existence of dependencies or influences across spatial units with a dependency matrix that measures geographic distances between units, which can then be used to specify a spatial autoregressive model with spatially lagged dependent variable or a spatially correlated error structure model [5, 164]. A number of recent terrorism studies employ such approaches [115, 119]. Scholars have also tailored these spatial statistics to the study of binary dependent variables—akin to the binary terrorism events mentioned above—through the use of (e.g.) autologistic models that incorporate distance-based connectivity matrices [164, 166]. Finally, spatial forecast methods employing feature reduction techniques have recently been applied to terrorism events in efforts to incorporate the explanatory power of high dimensional geographic spaces into probabilistic forecasting models, and show great promise in this regard [32, 128].

6 Conclusion

In 1954, the psychologist Paul Meehl [111] published a path-breaking analysis of the relative accuracy of human clinical assessments versus simple statistical models for a variety of prediction problems, such as future school performance and criminal recidivism. Despite using substantially less information, the statistical models either outperformed, or performed as well as, the human assessments in most situations. Meehl's work has been replicated and extended to numerous other domains in the subsequent six decades and the results are always the same: the statistical models win. Meehl, quoted in Kahneman [94, Chap. 21], reflecting on 30 years of studies, said, "There is no controversy in the social sciences which shows such a large body of qualitatively diverse studies coming out in the same direction as this one."

This has not, of course, removed the human analysts from trying—and most certainly, *claiming*—to provide superior performance. These are justified in the face of overwhelming empirical evidence that they are inaccurate using a wide variety of excuses—dare we call them "pathologies"?—that Kahneman [94, Part 3] discusses in great detail in the section of his book titled, appropriately, "Overconfidence." It is not only that simple statistical models are superior to human punditry, but as Tetlock [161] established, the most confident and a well-known human analysts actually tend to make the worst predictions. Furthermore, in most fields, the past performance of analysts—notoriously, the well-compensated stock-pickers of managed mutual funds—provides essentially no guide to their future performance.

At present, there is very little likelihood that human punditry, particularly the opinionated self-assurance so valued in the popular media, will be completely replaced by the unblinking assessments of computer programs, whether on 24-hour news channels or in brainstorming sessions in windowless conference rooms. Humans are social animals with exquisite skills at persuasion and manipulation; computer programs simply are far more likely to provide the correct answer in an inexpensive, consistent and transparent manner.

Yet with the vast increase in the availability of data, computational power, and the resulting refinement of methodological techniques, there is some change in the works. While the covers of investment magazines are adorned with the latest well-coffed guru who by blind luck has managed to have an unusually good year, in fact algorithmic trading now accounts for all but a small fraction of the activity on financial exchanges. Weather forecasts are presented on television by jocular and attractive individuals with the apparent intelligence of tadpoles, but the forecasts themselves are the result of numerical models processing massive quantities of data. In the political realm, sophisticated public opinion polling replaced the intuitive hunches of experts several decades ago, and polls are so rarely incorrect that it is a major news event when they fail. Even that last bastion of intuitive manliness, the assessment of athletes, can be trumped by statistical models, as documented in the surprisingly popular book and movie *Moneyball* [100].

Similar progress in the adoption of models forecasting political violence, particularly terrorism, is likely to be much slower. As we have stressed repeatedly, one of the major challenges of this research is that political violence is a rare event, and acts of terrorism in otherwise peaceful situations—the "bolt out of the blue" of Oklahoma City, 9/11/2001, Madrid 2004 and Utøya Island, Norway in 2011—are among the rarest. Consequently even if the statistical models are more accurate—and there is every reason to believe that they will be—establishing this will take far longer than is required in a field where new predictions can be assessed by the day, or even by the minute. In addition, a long series of psychological studies have shown that human risk assessment is particularly inaccurate—and hence its validity overestimated—in low probability, high risk situations, precisely the domain of counter-terrorism. Getting the technical assessments in the door, to say nothing of getting them used properly, will not be an easy task, and the initial applications will almost certainly be in domains where the events of interest are more frequent, as we are already seeing with the success of the Political Instability Task Force. But as this chapter has illustrated, the challenges are well understood, a plethora of sophisticated techniques await experimentation, and the appropriate data is readily available, all that is needed is the will and skill to apply these.

Acknowledgements This research was supported in part by a grant from the U.S. National Science Foundation, SES-1004414, and by a Fulbright-Hays Research Fellowship for work by Schrodt at the Peace Research Institute, Oslo (http://www.prio.no).

References

1. Abbott A (1995) Sequence analysis: new methods for old ideas. Ann Rev Sociol 21:93–113
2. Abdi H, Williams LJ (2010) Principal components analysis. Wiley Interdiscip Rev 2(4): 433–459
3. Achen C (2002) Toward a new political methodology: microfoundations and ART. Ann Rev Pol Sci 5:423–450
4. Andriole SJ, Hopple GW (1984) The rise and fall of events data: from basic research to applied use in the U.S. Department of Defense. Int Interact 10(3/4):293–309
5. Anselin L (1988) Spatial econometrics: methods and models. Kluwer, Dordrecht
6. Asal V, Pate A, Wilkenfeld J (2008) Minorities at risk organizational behavior data and codebook version 9–2008. http://www.cidcm.umd.edu/mar/data.asp
7. Azad S, Gupta A (2011) A quantitative assessment on 26/11 mumbai attack using social network analysis. J Terror Res 2(2):4–14
8. Azar EE (1980) The conflict and peace data bank (COPDAB) project. J Confl Resolut 24: 143–152
9. Azar EE, Sloan T (1975) Dimensions of interaction. University Center for International Studies, University of Pittsburgh, Pittsburgh
10. Beck N, Katz. JN (1995) What to do (and not to do) with time- series cross-section data. Am Pol Sci Rev 89:634–634
11. Beck N, Katz JN, Tucker R (1998) Taking time seriously: time-series-cross-section analysis with a binary dependent variable. Am J Pol Sci 42(4):1260–1288
12. Beck N, King G, Zeng L (2000) Improving quantitative studies of international conflict: a conjecture. Am Pol Sci Rev 94(1):21–35
13. Beger A, DeMeritt JH, Hwang W, Moore WH (2011) The split population logit (spoplogit): modeling measurement bias in binary data. http://ssrn.com/abstract=1773594. Working paper
14. Benini AA, Moulton LH (2004) Civilian victims in an asymmetrical conflict: operation enduring freedom, afghanistan. J Peace Res 41(4):403–422
15. Bennett DS, Stam AC (2004) The behavioral origins of war. University of Michigan Press, Ann Arbor
16. Berk R (2009) The role of race in forecasts of violent crime. Race Soc Probl 1(4):231–242
17. Berk R, Sherman L, barnes G, Kurtz E, Ahlman L (2007) Forecasting murder within a population of probationers and parolees: A high stakes application of statistical learning. J R Stat Soc 172(1):191–211
18. Besley T, Persson T (2009) Repression or civil war? Am Econ Rev 99(2):292–297
19. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 2:993–1022
20. Bond D, Jenkins JC, Taylor CLT, Schock K (1997) Mapping mass political conflict and civil society: Issues and prospects for the automated development of event data. J Confl Resolut 41(4):553–579
21. Bond J, Petroff V, O'Brien S, Bond D (2004) Forecasting turmoil in Indonesia: an application of hidden Markov models. Presented at the International Studies Association Meetings, Montréal

22. Box GE, Jenkins GM, Reinsel GC (1994) Time series analysis: forecasting and control, 3rd edn. Prentice Hall, Englewood Cliffs
23. Box-Steffensmeier J, Reiter D, Zorn C (2003) Nonproportional hazards and event history analysis in international relations. J Confl Resolut 47(1):33–53
24. Box-Steffensmeier JM, Jones BS (2004) Event history modeling: a guide for social scientists. Analytical methods for social research. Cambridge University Press, Cambridge
25. Brandt PT, Colaresi MP, Freeman JR (2008) The dynamics of reciprocity, accountability and credibility. J Confl Resolut 52(3):343–374
26. Brandt PT, Freeman JR (2006) Advances in Bayesian time series modeling and the study of politics: Theory testing, forecasting, and policy analysis. Pol Anal 14(1):1–36
27. Brandt PT, Freeman JR, Schrodt PA (2011) Real time, time series forecasting of inter- and intra-state political conflict. Confl Manage Peace Sci 28(1):41–64
28. Brandt PT, Sandler T (2009) Hostage taking: understanding terrorism event dynamics. J Policy Model 31(5):758–778
29. Brandt PT, Sandler T (2010) What do transnational terrorists target? Has it changed? Are we safer? J Confl Resolut 54(2):214–236
30. Brandt PT, Williams JT (2007) Multiple time series models. Sage, Thousand Oaks
31. Brochet J, Marie-Paule-Lefrance, Guidicelli V (2008) Imgt/v-quest: the highly customized and integrated system for ig and tr standardized v-j and v-d-j sequence analysis. Nucl Acids Res 36:387–395
32. Brown D, Dalton J, Hoyle H (2004) Spatial forecast methods for terrorist events in urban environments. Intell Secur Inform 3073:426–435
33. Buhaug H (2006) Relative capability and rebel objective in civil war. J Peace Res 43(6): 691–708
34. Buhaug H, Lujala P (2005) Accounting for scale: measuring geography in quantitative studies of civil war. Pol Geogr 24(4):399–418
35. Buhaug H, Rod JK (2006) Local determinants of African civil wars, 1970–2001. Pol Geogr 25(6):315–335
36. Burgess PM, Lawton RW (1972) Indicators of international behavior: an assessment of events data research. Sage, Beverly Hills
37. Cederman LE, Gleditsch KS (2009) Introduction to special issue of "disaggregating civil war". J Confl Res 24(4):590–617
38. Choi K, Asal V, Wilkenfeld J, Pattipati KR (2012) Forecasting the use of violence by ethno-political organizations: middle eastern minorities and the choice of violence. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
39. Choucri N, Robinson TW (eds) (1979) Forecasting in international relations: theory, methods, problems, prospects. Freeman, San Francisco
40. Clark DH, Regan PM (2003) Opportunities to fight: a statistical technique for modeling unobservable phenomena. J Confl Resolut 47(1):94–115
41. Cunningham DE (2011) Barriers to Peace in Civil War. Cambridge University Press, Cambridge
42. Cunningham DE, Gleditsch KS, Salehyan I (2009) It takes two: a dyadic analysis of civil war duration and outcome. J Confl Resolut 53(4):570–597
43. Cutler DR, Edwards TC Jr, Beard KH, Cutler A, Hess KT, Gibson J, Lawler J (2007) Random forests for classification in ecology. Ecology 88(11):2783–2792
44. Daly JA, Andriole SJ (1980) The use of events/interaction research by the intelligence community. Policy Sci 12:215–236
45. Davies JL, Gurr TR (eds) (1998) Preventive measures: building risk assessment and crisis early warning. Rowman and Littlefield, Lanham
46. Dereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the vax. Nucl Acids Res 12(1):387–395

47. Desmarais BA, Cranmar SJ (2011) Forecasting the locational dynamics of transnational terrorism: a network analysis approach. In: Paper presented at the 2011 European Intelligence and Security Conference, Athens. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6061174&tag=1
48. Dickerson JP, Simari GI, Subrahmanian VS (2012) Using temporal probabilistic rules to learn group behavior. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
49. D'Orazio V, Landis ST, Palmer G, Schrodt PA (2011) Separating the wheat from the chaff: application of support vector machines to mid4 text classification. Paper presented at the annual meeting of the midwest political science association, Chicago
50. D'Orazio V, Yonamine J, Schrodt PA (2011) Predicting intra-state conflict onset: an events data approach using euclidean and levenshtein distance measures. Presented at the midwest political science association, Chicago. Available at http://eventdata.psu.edu
51. Duval RD, Thompson WR (1980) Reconsidering the aggregate relationship between size, economic development, and some types of foreign policy behavior. Am J Pol Sci 24(3): 511–525
52. Enders W, Parise GF, Sandler T (1992) A time-series analyssis of transnational terrorism. Def Peace Econ 3(4):305–320
53. Enders W, Sandler T (1993) The effectiveness of anti-terrorism policies: a vector-autoregression-intervention analysis. Am Pol Sci Rev 87(4):829–844
54. Enders W, Sandler T (2000) Is transnational terrorism becoming more threatening? A time-series investigation. J Confl Resolut 44(3):307–332
55. Enders W, Sandler T (2005) After 9/11: is it all different now? J Confl Resolut 49(2):259–277
56. Enders W, Sandler T (2006) Distribution of transnational terrorism among countries by income classes and geography after 9/11. Int Stud Q 50(2):367–393
57. Enders W, Sandler T (2006) The political economy of terrorism. Cambridge University Press, Cambridge
58. Enders W, Sandler T, Cauley J (1990) Assessing the impact of terrorist-thwarting policies: an intervention time series approach. Def Peace Econ 2(1):1–18
59. Enders W, Sandler T, Gaibulloev K (2011) Domestic versus transnational terrorism: Data, decomposition, and dynamics. J Peace Res 48(3):319–337
60. Esty DC, Goldstone JA, Gurr TR, Harff B, Levy M, Dabelko GD, Surko P, Unger AN (1998) State failure task force report: phase II findings. Science Applications International Corporation, McLean
61. Esty DC, Goldstone JA, Gurr TR, Surko P, Unger AN (1995) State failure task force report. Science Applications International Corporation, McLean
62. Freedom House (2009) Freedom in the world. http://www.freedomhouse.org/reports
63. Freeman JR (1989) Systematic sampling, temporal aggregation, and the study of political relationships. Pol Anal 1:61–98
64. Gaddis JL (1992) International relations theory and the end of the cold war. Int Secur 17:5–58
65. Gaddis JL (1992) The United States and the end of the cold war. Oxford University Press, New York
66. Gassebner M, Luechinger S (2011) Lock, stock, and barrel: a comprehensive assessment of the determinants of terror. Public Choice 149:235–261
67. Gerner DJ, Schrodt PA, Francisco RA, Weddle JL (1994) The machine coding of events from regional and international sources. Int Stud Q 38:91–119
68. Gerner DJ, Schrodt PA, Yilmaz Ö (2009) Conflict and mediation event observations (CAMEO) Codebook. http://eventdata.psu.edu/data.dir/cameo.html
69. Ghahramani Z (2004) Unsupervised learning. In: Advanced lectures on machine learning. Springer, Berlin/New York, pp 72–112
70. Gill J (1999) The insignificance of null hypothesis significance testing. Pol Res Q 52(3): 647–674
71. Gill J (2003) Bayesian methods: a social and behavioral sciences approach. Chapman and Hall, Boca Raton

72. Gleditsch NP (2012) Special issue: event data. Int Interact 38(4)
73. Goldstein JS (1992) A conflict-cooperation scale for WEIS events data. J Confl Resolut 36:369–385
74. Goldstone JA, Bates R, Epstein DL, Gurr TR, Lustik M, Marshall, MG, Ulfelder J, Woodward M (2010) A global model for forecasting political instability. Am J Pol Sci 54(1):190–208
75. Gurr TR, Harff B (1994) Conceptual, research and policy issues in early warning research: an overview. J Ethno-Dev 4(1):3–15
76. Gurr TR, Lichbach MI (1986) Forecasting internal conflict: A competitive evaluation of empirical theories. Comp Pol Stud 19:3–38
77. Hämerli A, Gattiker R, Weyermann R (2006) Conflict and cooperation in an actor's network of Chechnya based on event data. J Confl Resolut 50(159):159–175
78. Hamilton J (1994) Time series analysis. Princeton University Press, Princeton
79. Harff B, Gurr TR (2001) Systematic early warning of humanitarian emergencies. J Peace Res 35(5):359–371
80. Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer, New York
81. Hegre H, Ellingson T, Gates S, Gleditsch NP (2001) Toward a democratic civil peace? democracy, political change, and civil war, 1816–1992. J Policy Model 95(1):34–48
82. Hegre H, Ostby G, Raleigh C (2009) Property and civil war: a disaggregated study of liberia. J Confl Resolut 53(4):598–623
83. Hill DW, Bagozzi BE, Moore WH, Mukherjee B (2011) Strategic incentives and modeling bias in ordinal data:the zero-inflated ordered probit (ziop) model in political science. In: Paper presented at the new faces in political methodology meeting, Penn State, 30 April 2011. http://qssi.psu.edu/files/NF4Hill.pdf
84. Holmes JS, Pineres SAGD, Curtina KM (2007) A subnational study of insurgency: farc violence in the 1990s. Stud Confl Terror 30(3):249–265
85. Hopple GW, Andriole SJ, Freedy A (eds) (1984) National security crisis forecasting and management. Westview, Boulder
86. Hosmer DW, Lemeshow S, May S (2008) Applied survival analysis: regression modeling of time to event data. Series in probability and statistics. Wiley, New York
87. Hudson V (ed) (1991) Artificial intelligence and international politics. Westview, Boulder
88. Hudson VM, Schrodt PA, Whitmer RD (2008) Discrete sequence rule models as a social science methodology: an exploratory analysis of foreign policy rule enactment within Palestinian-Israeli event data. Foreign Policy Anal 4(2):105–126
89. International Monetary Fund. Statistics Dept. (2009) Direction of trade statistics yearbook, 2008. http://www.imf.org/external/pubs/cat/longres.cfm?sk=22053.0
90. Inyaem U, Meesad P, Haruechaiyasak C (2009) Named-entity techniques for terrorism event extraction and classification. In: Paper presented at the eighth international symposium on natural language processing. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5340924&tag=1
91. Jackman RW, Miller RA (1996) A renaissance of political culture? Am J Pol Sci 40(3): 632–659
92. Jackman S (2009) Bayesian analysis for the social sciences. Wiley, Chichester
93. Jenkins CJ, Bond D (2001) Conflict carrying capacity, political crisis, and reconstruction. J Confl Resolut 45(1):3–31
94. Kahneman D (2011) Thinking fast and slow. Farrar, Straus and Giroux, New York
95. Khong YF (1992) Analogies at war. Princeton University Press, Princeton
96. King G (1986) How not to lie with statistics: avoiding common mistakes in quantitative political science. Am J Pol Sci 30(3):666–687
97. King G, Zeng L (2001) Logistic regression in rare events data. Pol Anal 9(2):12–54
98. Lee K, Booth D, Alam P (2005) A comparison of supervised and unsupervised neural networks in predicting bankruptcy of korean firms. Expert Syst Appl 29:1–16
99. Leng RJ (1987) Behavioral correlates of war, 1816–1975. (ICPSR 8606). Inter-University Consortium for Political and Social Research, Ann Arbor

100. Lewis M (2011) Moneyball. Norton, New York
101. Lund MS (1996) Preventing violent conflicts: a strategy for preventive diplomacy. United States Institute for Peace, Washington, DC
102. Malo MA, Munoz-Bullon F (2003) Employment status mobility from a life-cycle perspective: a sequence analysis of work-histories in the bhps. Demogr Res 9:119–162
103. Mannes A (2012) Qualitative analysis & computational techniques for the counter-terror analyst. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
104. Marshall MG, Jaggers K, Gurr TR (2010) Polity iv project: political regime characteristics and transitions, 1800–2010. Center for Systemic Peace, Vienna, VA
105. Martinez MV, Silva A, Simari GI, Subrahmanian VS (2012) Forecasting changes in terror group behavior. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
106. Martinez V, Simari G, Sliva A, Subrahmanian V (2008) Convex: similarity-based algorithms for forecasting group behavior. IEEE Intell Syst 23:51–57
107. Martinez V, Simari G, Sliva A, Subrahmanian V (2008) The soma terror organization portal (stop): Social network and analytic tools for the real-time analysis of terror groups. In: Liu H, Salerno JJ, Young M (eds) Social computing, behavioral modeling, and prediction. Springer/Wien, Norderstedt, pp 9–18
108. May ER (1973) "Lessons" of the past: the use and misuse of history in American foreign policy. Oxford University Press, New York
109. McClelland CA (1967) World-event-interaction-survey: a research project on the theory and measurement of international interaction and transaction. University of Southern California, Los Angeles, CA
110. McClelland CA (1976) World event/interaction survey codebook (ICPSR 5211). Inter-University consortium for political and social research, Ann Arbor
111. Meehl P (1954) Clinical and statistical prediction: a theoretical analysis and a review of the evidence. University of Minnesota Press, Minneapolis
112. Minorities at Risk Project (2009) Minorities at risk dataset. Center for International Development and Conflict Management, College Park. http://www.cidcm.umd.edu/mar/
113. Murray MP (1994) A drunk and her dog: an illustration of cointegration and error correction. http://www-stat.wharton.upenn.edu/~steele/Courses/434/434Context/Co-int%egration/Murray93DrunkAndDog.pdf
114. National Consortium for the Study of Terrorism and Responses to Terrorism (START) (2011) Global terrorism database [data file]. http://www.start.umd.edu/gtd
115. Neumayer E, Plümper T (2010) Galton's problem and contagion in international terrorism along civilizational lines. Confl Manage Peace Sci 27(4):308–325
116. Neustadt RE, May ER (1986) Thinking in time: the uses of history for decision makers. Free Press, New York
117. O'Brien S (2012) A multi-method approach for near real time conflict and crisis early warning. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
118. O'Brien SP (2010) Crisis early warning and decision support: Contemporary approaches and thoughts on future research. Int Stud Rev 12(1):87–104
119. Öcal N, Yildirim J (2010) Regional effects of terrorism on economic growth in turkey: A geographically weighted regression. J Peace Res 47(4):477–489
120. Oskar Engene J (2007) Five decades of terrorism in europe: the tweed dataset. J Peace Res 44(1):109–121
121. Pearl J (2009) Understanding propensity scores. In: Causality: models, reasoning, and inference, 2nd edn. Cambridge University Press, Cambridge/New York
122. Perliger A, Pedahzur A (2011) Social network analysis in the study of terrorism and political violence. Pol Sci Pol 44(1):45–50

123. Petroff V, Bond J, Bond D (2012) Using hidden markov models to predict terror before it hits (again). In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
124. Pevehouse JC, Brozek J (2010) Time series analysis in political science. In: Box-Steffensmeier J, Brady H, Collier D (eds) Oxford handbook of political Methodology. Oxford University Press, New York
125. Pevehouse JC, Goldstein JS (1999) Serbian compliance or defiance in Kosovo? Statistical analysis and real-time predictions. J Confl Resolut 43(4):538–546
126. Piazza J (2011) Poverty, minority economic discrimination and domestic terrorism. J Peace Res 48(3):339–353
127. Piazza J, Walsh JI (2009) Transnational terrorism and human rights. Int Stud Q 53:125–148
128. Porter MD, Brown DE (2007) Detecting local regions of change in high-dimensional criminal or terrorist point processes. Comput Stat Data Anal 51(5):2753–2768
129. Raleigh C, Linke A, Hegre H, Karlsen J (2010) Introducing ACLED: an armed conflict location and event dataset. J Peace Res 47(5):651–660
130. Regan P, Clark D (2010) The institutions and elections project data collection. http://www2.binghamton.edu/political-science/institutions-and-elections%-project.html
131. Richardson LF (1960) Statistics of deadly quarrels. Quadrangle, Chicago
132. Rupesinghe K, Kuroda M (eds) (1992) Early warning and conflict resolution. St. Martin's, New York
133. Schrodt PA (2000) Pattern recognition of international crises using hidden Markov models. In: Richards D (ed) Political complexity: nonlinear models of politics. University of Michigan Press, Ann Arbor, pp 296–328
134. Schrodt PA (2004) Detecting united states mediation styles in the middle east, 1979–1998. In: Maoz Z, Mintz A, Morgan TC, Palmer G, Stoll RJ (eds) Multiple paths to knowledge in international relations. Lexington Books, Lexington, pp 99–124
135. Schrodt PA (2006) Forecasting conflict in the Balkans using hidden Markov models. In: Trappl R (ed) Programming for peace: computer-aided methods for international conflict resolution and prevention. Kluwer, Dordrecht, pp 161–184
136. Schrodt PA (2007) Inductive event data scaling using item response theory. Presented at the summer meeting of the society of political methodology. Available at http://eventdata.psu.edu
137. Schrodt PA (2010) Seven deadly sins of contemporary quantitative analysis. Presented at the american political science association Meetings, Washington, DC. http://eventdata.psu.edu/papers.dir/Schrodt.7Sins.APSA10.pdf
138. Schrodt PA, Gerner DJ (1994) Validity assessment of a machine-coded event data set for the Middle East, 1982–1992. Am J Pol Sci 38:825–854
139. Schrodt PA, Gerner DJ (1997) Empirical indicators of crisis phase in the Middle East, 1979–1995. J Confl Resolut 25(4):803–817
140. Schrodt PA, Gerner DJ (2004) An event data analysis of third-party mediation. J Confl Resolut 48(3):310–330
141. Schrodt PA, Gerner DJ, Abu-Jabr R, Yilmaz, Ö, Simpson EM (2001) Analyzing the dynamics of international mediation processes in the Middle East and Balkans. Presented at the American Political Science Association meetings, San Francisco
142. Schrodt PA, Palmer G, Hatipoglu ME (2008) Automated detection of reports of militarized interstate disputes using the svm document classification algorithm. Paper presented at American Political Science Association, San Francisco
143. Schrodt PA, Van Brackle D (2012) Automated coding of political event data. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
144. Shearer R (2006) Forecasting Israeli-Palestinian conflict with hidden Markov models. Available at http://eventdata.psu.edu/papers.dir/Shearer.IP.pdf
145. Shellman S (2000) Process matters: Conflict and cooperation in sequential government-dissident interactions. Journal of Conflict Resolution **15**(4), 563–599
146. Shellman S (2004) Time series intervals and statistical inference: the effects of temporal aggregation on event data analysis. Secur Stud 12(1):97–104

147. Shellman S, Hatfield C, Mills M (2010) Disaggregating actors in intrastate conflict. J Peace Res 47(1):83–90
148. Shellman S, Stewart B (2007) Predicting risk factors associated with forced migration: An early warning model of Haitian flight. Civil Wars 9(2):174–199
149. Silke A (ed) (2004) Research on terrorism: trends, achievements and failures. Frank Cass, London
150. Silke A (2009) Contemporary terrorism studies: issues in research. In: Jackson R, Smyth MB, Gunning J (eds) Critical terrorism studies: a new research agenda. Routledge, London
151. Silva A, Simari G, Martinez V, Subrahmanian VS (2012) SOMA: Stochastic opponent modeling agents for forecasting violent behavior. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
152. Simari GI, Earp D, Martinez MV, Silva A, Subrahmanian VS (2012) Forecasting group-level actions using similarity measures. In: Subrahmanian V (ed) Handbook on computational approaches to counterterrorism. Springer, New York
153. Simi P (2010) Operation and structure of right-wing extremist groups in the united states, 1980–2007. http://www.icpsr.umich.edu/icpsrweb/NACJD/studies/25722/detail
154. Sims CA (1980) Macroeconomics and reality. Econometrica 1, 634–4
155. Sims CA, Waggoner DF, Zha TA (2008) Methods for inference in large multiple-equation Markov-switching models. J Econom 146(2):255–274
156. Siroky DS (2009) Navigating random forests and related advances in algorithmic modeling. Stat Surv 3:147–163
157. Sliva A, Subrahmanian V, Martinez V, Simari G (2009) Cape: Automatically predicting changes in group behavior. In: Memon N, Farley JD, Hicks DL, Rosenorn T (eds) Mathematical methods in counterterrorism. Springer/Wien, Norderstedt, pp 253–269
158. Sun S, Zhang C (2007) The selective random subspace predictor for traffic flow forecasting. IEEE Trans Intell Transp Syst 8(2):367–373
159. Svolik MW (2008) Authoritarian reversals and democratic consolidation. Am Pol Sci Rev 102(2):153–168
160. Sylvan DA, Chan S (1984) Foreign policy decision making: perception, cognition and artificial intelligence. Praeger, New York
161. Tetlock PE (2005) Expert political judgment: how good is it? How can we know? Princeton University Press, Princeton
162. Urdal H (2008) Population, resources and violent conflict: A sub-national study of india 1956–2002. J Confl Resolut 52(4):590–617
163. U.S. National Counterterrorism Center (2009) Worldwide incidents tracking system (wits). http://wits.nctc.gov/
164. Ward MD, Gleditsch KS (2002) Location, location, location: an mcmc approach to modeling the spatial context of war and peace. Pol Anal 10(3):244–260
165. Ward MD, Greenhill BD, Bakke KM (2010) The perils of policy by p-value: predicting civil conflicts. J Peace Res 47(5):363–375
166. Weidman NB, Ward MD (2010) Predicting conflict in space and time. J Confl Resolut 54(6):883–901
167. Weidmann NB, Toft MD (2010) Promises and pitfalls in the spatial prediction of ethnic violence: a comment. Confl Manage Peace Sci 27(2):159–176
168. Xiang J (2010) Relevance as a latent variable in dyadic analysis of conflict. J Pol 72(2): 484–498
169. Young JK, Findley MG (2011) Promise and pitfalls of terrorism research. Int Stud Rev 13:411–431
170. Yonamine JE (2012) Working with even data: a guide to aggregation. Available at http://jayyonamine.com/wp-content/uploads/2012/06/Working-with-Event-Data-A-Guide-to-Aggregation-Choices.pdf