Can We Predict Politics? Toward What End?

Michael D. Ward

Duke University

Shipwrecks and Security Studies

In 1901, off the Greek island of Antikythera, a ship pulled into a bay to wait out a storm. After the storm was over, its divers discovered an ancient shipwreck containing many valuable antiquities, including jewelry, coins, statues, and pottery. One item was a lump of corroded bronze and wood. Everything was carted off to the National Museum of Archeology in Athens. In 1902, an archeologist noticed that the corroded lump had what appeared to be gears in it. He assumed that it was some sort of astrological clock, but it appeared to be too far advanced given the dating of the other items it was found with, which were initially dated to about 150 BCE, and it was ignored for five decades. Several years later, it was X and γ rayed, resulting in images of eighty-two different fragments of the device.

By the end of 2014, it had been established that the mechanism dated back further, to about 250 BCE. Many consider it to be the first computing machine (Freeth et al. 2006). This device was able to calculate and display celestial cycles, including phases of the moon, as well as a solar calendar. Perhaps more importantly, it was able to predict eclipses—seen at the time as omens. In fact, many think of the Antikythera Mechanism as an *omen prediction device*, with all the other predictions it makes simply by-products of its true purpose.

The main purpose of this mechanism was to generate accurate predictions, and whoever used it could do so without a detailed theoretical knowledge of Hipparchosian astronomy as applied to irregular phases of ellipsoid orbits. Thus, it was pure prediction and did not "explain" anything. At the same time, it embodied detailed engineering that was based on a theoretical mechanism that provided exquisite details of planetary orbits in general as well as

specifically. For some, that might suffice as explanation. It seems clear that this accurate prediction device was possible only because it was based on a deep understanding of celestial orbits. What is particularly surprising is that this level of astronomical prediction was apparently lost in the shipwreck and did not reappear in Europe for over a millennium. In retrospect, it seems unlikely that this device would simply disappear, given how powerful it must have been in 250 BCE. Yet, as far as we know, this is exactly what happened.

Jumping ahead a thousand years, we see that prediction is deeply embedded in the philosophy of science. Although in hiding in the aftermath of the French Revolution, Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet, wrote in *Historical View of the Progress of the Human Mind* (1795, translated):

> If man can, with almost complete assurance, predict phenomena when he knows their laws, and if, even when he does not, he can still, with great expectation of success, forecast the future on the basis of his experience of the past, why, then, should it be regarded as a fantastic undertaking to sketch, with some pretense to the truth, the future destiny of man on the basis of his history. The sole foundation for belief in the natural sciences is this idea, and the general laws directing the phenomena of the universe, known or unknown, are necessary and constant.

This theme continues to the current day through Auguste Comte (1846), John Stuart Mill (1843), A. J. Ayer (1936), Carl Hempel (1935), Karl Popper (1935), Thomas Kuhn (1962), and even Jon Elster (1989, 2007), who notes:

> To predict that less of a good will be bought when its price goes up, there is no need to form a hypothesis about human behavior. Whatever the springs of individual action—

ration, traditional, or simply random—we can predict that people will buy less of a good simply because they can afford less of it. Here there are several mechanisms that are constrained to lead to the same outcome, so that for predictive purposes there is no need to decide among them. Yet for explanatory purposes the mechanism is what matters. It provides understanding whereas prediction at most offers control. (1989, 9)

The argument is that if you can develop models that provide an understanding—without a teleology of why things happen—you should be able to generate predictions that will not only be accurate, but may also be useful in a larger societal context. The basic argument is essentially that it should be possible to develop a predictive science of human behavior. It is debatable whether the so-called regularities are constant or changing.[1]

During the height of the war in Vietnam, in the mid-1960s, a graduate student at Stanford began to systematically study the details of the war, down to the level of troop numbers on all sides, number of bombing sorties, casualties, troop attrition, kill ratios (then a big indicator in McNamara's Defense Department), and political support in the United States and South Vietnam for the incumbent administration and the war itself, along with the number of Viet Cong and North Vietnamese defectors. Statistical models of the relationships among these variables were constructed and validated toward the goal of generating predictions (from computer simulations) about the escalation of the war. Jeffrey S. Milstein published his efforts in *Dynamics of the Vietnam War: A Quantitative Analysis and Predictive Computer Simulation* in 1974, and his working papers on this topic were widely read in the academic and policy communities. This was one of the first uses of prediction in the study of international relations (IR) and security studies. By 1974, articles were starting to promote forecasting in the realm of world politics (Choucri 1974), and a few years later, there was a volume devoted to the topic (Choucri and Robinson 1978).

There have been a number of scholarly efforts at prediction in the realm of international and domestic conflict. (Ironically, the Vietnam War was considered international by many comparativists, but a civil conflict by many IR scholars). One of the first calls for greater attention was by Herman Weil (1974), then working in the defense-consulting sector. After a few efforts in the later 1970s, an article by Gurr and Lichbach (1986)

took Gurr's model of *Why Men Rebel* (1970) and used it to make out-of-sample forecasts of where there would be future conflicts. It did so explicitly to test the theoretical model of conflict that had been initially developed by Ted Robert Gurr. It forecasts the number of days of protests as well as the number of deaths in protests in ten countries. But this study generated forecasts that used data measured in five-year periods, and the period from 1971 to 1975 was forecast based on data for about 1970. Nonetheless, it was one of the first empirically oriented studies to focus explicitly on forecasts. Until very recently, most of this thread of work in security studies had been lost, or if not lost, at least abandoned.

Cioffi-Revilla (1996) provided a published prediction of what was likely to happen in the first Iraq War. It was the only published prediction that I can find, or remember, though it took a while to get published. By now, there is a bevy of recent efforts that include a forecasting component (Gleditsch and Ward 2000, 2010, 2013; King and Zeng 2001; Ward and Gleditsch 2002; Hegre 2008; Weidmann and Ward 2010; de Mesquita 2011; Metternich et al. 2013; Ward et al. 2013; Koubi and Böhmelt 2014; Pilster and Böhmelt 2014; Schutte 2014). Why has this thread been so sparse in security studies? Two words: Kenneth Waltz.

Many social scientists see a sharp distinction between *explanation* on the one hand and *prediction* on the other. Indeed, this distinction is often sharp enough that it is argued that doing one of these things cuts you out of doing the other. Kenneth Waltz (1997) is a good example of this belief. A long-standing, but incorrect, example has been the weather, about which it has been famously argued that while you can understand the various components of the weather system—the evaporation of water, its collection in clouds, the changing temperatures that result in lightning, and the like—having this explanation does not enable you to make actual predictions about the weather. It is argued that because the contexts are sufficiently varied and numerous, they would defeat our ability to predict the weather. Mill argued against this idea, the then-prevailing opinion that tidology would never be a precise endeavor. He proposed that weather prediction could, in principle, be successful and become an exact science. Guess what? Against all the naysayers, he was right.

Nonetheless, the idea that prediction and understanding are different has persisted and in fact is widespread in the realm of security studies. In a simple way this is true, because you can develop a predictive system without necessarily understanding all the details that are in play. A classic example often offered is that

---

1   Others—Hume and Heidegger, for example, as well as those working on recent hermeneutics—had a different view that I do not elaborate herein.

an individual can be an excellent pool player without being able to explain the basic equations of motion. But this example seems disingenuous, and solutions to the equations may take different forms. Perhaps, a kind of "intuitive" understanding may supplant a more "analytic" understanding. We know that analytic knowledge is harder to communicate to others than is intuitive information, but it may be that both types of knowledge are useful and even accurate. Moreover, who knows exactly what is going on in the nervous system of an excellent pool player at a level below cognition?

It is clear that the distinction between these two ideas about knowledge is prominent, but it is very hard to determine where it originates. Note the following training guide on telling the two apart: http://bit.ly/1DErHSD. This comes from the US National Institutes of Health Office of Behavioral and Social Science Research and is part of their online training curriculum that helps practitioners determine what is useful for prediction and explanation. This visual guide distinguishes between predictive and explanatory concepts related to the risk of teen pregnancy, allowing practitioners to "learn" which category best describes different variables such as race. The software not only classifies factors into the mutually exclusive categories of explanation and prediction, but it also grades you on how well you can distinguish both.

It turns out that most of the examples of what we cannot predict have been overturned by scientific inquiry, accompanied by clever and concerted research programs. In part, that was Popper's (1935) point. Lewis Fry Richardson dreamed in 1922 of thousands of individuals making thousands of calculations in order to do weather prediction. Richardson had just spent over two years making the necessary calculations for a single day, May 20, 1910. He described his dream:

> After so much hard reasoning, may one play with fantasy? Imagine a large hall like a theatre, except that the circles and galleries go right round through the space usually occupied by the stage. The walls of this chamber are painted to form a map of the globe. The ceiling represents the north polar regions, England is in the gallery, the tropics in the upper circle, Australia on the dress circle and the Antarctic in the pit. A myriad of computers are at work upon the weather of the part of the map where each sits, but each computer attends only to one equation or part of an equation.

A picture of that dream was remarkably prescient, *Richardson's Dream*, which had 60,000 "computers" (individuals with slide rules) working simultaneously to produce a weather forecast. Detailed data on a wide variety of variables, coupled with rapid calculations of some fairly "simple" equations, have produced weather predictions that are sufficiently accurate that there is now a derivatives market in weather.[2]

On the other hand, many scholars (but few others) will tell you that we need more theory. Doubtless they are right. Few of them really mean "theory" in the sense that I reserve for the term. Few of them mean "theory" in the sense of analytical narratives. Many of them mean "detailed, plausible stories" about how stuff occurs. Long ago in a galaxy far away from here, James Caporaso (echoing Arthur Stinchcombe) taught me that any good social scientist should be able to come up with a variety of plausible stories about how something comes about, about what "causes" things, not just one. Unfortunately, much of security studies appears to stop at this junction. At the time, Caporaso intended it as a challenge to guide research. I have come to appreciate that it is also a curse. In short, most scholars seem to want more theories that answer the question of "why" something occurs. A wide swath of published articles in security studies broadly includes a section on "theory," and even more empirical studies are pilloried for weak or absent theory.[3]

What is typically meant by "more theory" falls into a way to create more nuanced understandings of the world. Often, what is desired is an attempt to find ways to encapsulate parts of the theory in a broader picture aimed at reconciling empirical findings with what is prescribed by the theory. Vasquez (1997) details how this typically works in security studies and elsewhere in his widely cited examination of Waltz's balancing proposition. He details a longish list of recommendations, which he claims is degenerate in that it avoids attempts at refutation whether they are logical, definitional, or historical. Waltz's response (1997) misses the point but does offer up another defense: theories are meant for explanations, and he claims that "tests are always problematic."

In his most recent article in *Security Studies*, Daniel Altman (2015) proposes "a new theory of false optimism as a cause of war." Along with formal models, Altman points to Pearl Harbor as illustrative evidence that Japan's decision-making in 1941 was "an instance of false optimism that lead to war, in this case war against a country with nearly ten times its military potential." Altman argues that when alternative strategies are proposed, those that are most optimistic are most

2 Among many derivative market products for weather, see http://www.climetrix.com.

3 By wide swath, I mean virtually all.

likely to be chosen and those that are most pessimistic are likely to be avoided. But this "theory" is not really examined empirically, nor are there counter examples.

Theory is also used as a goal of process tracing. Consider Mahoney's claim that process tracing can be used for both theory testing and theory development, developed in many places but available in Mahoney (2015), in which a scholar is interested in the putative, generic question "What *X* caused *Y* in case *Z*?" or for example "What are the possible causes or alternative explanations of World War I? … [T]his [is] the *theory construction task* and process tracing is one tool for pursuing it." The more prosaic task is to determine whether this generated theory was validated in a specific case, although herein the case is already pretty specific. There are many guidelines, including reasoning logically (surely a good thing), having knowledge of specific cases (also beneficial, but potentially misleading), and having "good knowledge of relevant preexisting theories and generalizations." The latter is doubtless useful, perhaps to know which ideas to throw away and which to build on. For Mahoney, and many others, a theory appears to be a list of possible causes or explanations.

For many others, "theory" is what is missing in extant research on security studies or simply what they want to focus on. Beckley (2015, 9) notes that there is a paucity of studies that focus on theory building in the realm of alliance studies. He constructs two competing "perspectives": entanglement theory that alliances drag states into war and "freedom of action theory," suggesting that loopholes avoid entanglements. Although anointed as theories, Beckley often refers to each of these simple, long-standing ideas in world affairs as "perspectives." Another approach to "theory" is to refine extant explanations. A terrific example of this can be found in Liff and Ikenberry (2014), who focus on the theory of "security dilemmas" as a guide to understanding military competition in the Asian Pacific. They take the idea of military competition back to Herz (1950) and Jervis (1978), though I would date it back to Richardson (1960), which was written much earlier (first published in 1948, but written in the 1930s) and published posthumously. Liff and Ikenberry suggest that "on both sides" states prefer to avoid consumption of security-producing goods (militaries, wars, treaties) if they can secure credible commitments without them. Mistrust and uncertainty generate an action–reaction dynamic that is contextualized in a number of ways, (notably those introduced by Jervis [1978]), and point to the importance of distinguishing between putatively offensive versus defensive postures, actions, and goods. This article expands the set of things to examine in order to determine this dynamic balance, adding considerable nuance along the way (Liff and Ikenberry 2014, 61):

> To begin, borrowing from recent scholarship critiquing the balancing literature, we expand the scope of the metrics typically employed in scholarly debates on the security dilemma to include a broader selection of internal and external policy measures aimed at enhancing a state's military capabilities. Indeed, various military force development and force employment measures aimed at enhancing military capabilities often overlooked in the existing literature can significantly influence a state's perceptions of its would-be adversary's intentions. Such measures include qualitatively improving military capabilities through modernization, innovation, or rationalization; transforming force structure or posture to confront changing threats; tightening military ties with other states short of new, formal mutual defense pacts through joint exercises and training, hosting or rotating foreign forces, co-locating military facilities, expanding interoperability and joint contingency planning, and sharing intelligence and military technology. In addition, we argue that independent of shifts in material power, leadership rhetoric and political statements can generate insecurity in others. For example, statements seen by one side as supporting the status quo may be interpreted by the other party as offensive and threatening.

Each of these may be useful exercises, but whether they merit the term "theory" or validate the repudiation of other strategies is open to interpretation. They do add nuance.

In a soon-to-be-classic piece, Healy (2015) notes that this sort of theoretical nuance has (also) gained a foothold in sociology. He finds this dysfunctional, not because theory or nuance is bad, but like Vasquez, because it hides ideas from critical examination. This nuance is fractal in that it appears in calls for higher as well as lower levels of abstraction. On the one hand, we see calls for the nuance of ever more detailed empirical analyses that illustrate which theory needs to be contextualized and amended to be useful. An opposite approach is to demand a theory of such a level of generality that it does not apply to anything precisely observable in the world. Healy's third nuance is the epitome of contextualization, "the insinuation that your sensitivity to nuance is a manifestation of one's distinctive … ability to grasp and express the richness, texture, and flow of social reality itself. This is the nuance of the *connoisseur*" (3). Healy concludes that sociology is "gutted" with nuance, and nuance needs to be avoided, at least for now. I have a similar view of the nuance of theory in security studies.

I am here to suggest that less is more.[4] Thus, let me be the first to call for less theory in security studies. We should winnow the many, many such "theories" that occupy the world of security studies.

Instead, we need more predictions. We need these predictions for four reasons. First, we need these predictions to help us make relevant statements about the world around us. We also need these predictions to help us throw out the bad "theories" that continue to flourish. These predictions will help drive our research into new areas, away from moribund approaches that have been followed for many decades. Finally, and perhaps most important, predictions will force us to keep on track. It is hard to imagine that in ten years hence, the theories I have highlighted above will have been evaluated in terms of their accuracy. It is easy to go back in the literature ten years and find similar pronouncements that were left hanging out there in the realm of JSTOR alone. At the end of the last century, John Vasquez and Kenneth Waltz debated the value of predictions in terms of their implications for theory (Vasquez 1997; Waltz 1997). As it turns out, this debate, while interesting, is not relevant for our discussion because neither Vasquez nor Waltz mean predictions in the sense of a forecast, but rather in the sense of an empirical regularity or fact that can be induced or deduced by the theoretical framework. Indeed, most of the "predictions" in this debate relate to events in the early twentieth century.

Predictions in the Larger World

In the 1980s, Philip Tetlock was concerned—like many—about the possibility of a nuclear conflict between the United States and Soviet Union. Indeed, this was the main focus of most security studies during the Cold War. Popular movies and scholarly journals highlighted what time it was on the so-called "doomsday" clock. Tetlock wanted to understand why the pundits made statements about the future that were all over the place. Frustrated by the fact that pundits justified their inaccuracies with revisions of history, timing, and other excuses, he collected forecasts—an amazing number of them (∼25,000)—and started to keep track and grade their accuracy. Tetlock's book, *Expert Political Judgment* (2005, 2010), showed that pundits as well as

4 I agree that we need more real theories. More generally, a theory is an explanation of some aspect of the natural world that is logically consistent and generates implications that are empirically falsifiable through observation and/or experimentation. A theory is not a simple conjecture or hypothesis. Many social science conjectures and hypotheses are called theories by their authors.

social scientists making predictions were equally and totally bad at it. An oft-cited refrain from the book is that "dart throwing chimps" were equally accurate to experts. Indeed, the most well qualified of experts often turned out to be the worst forecasters. Unlike Bill Ascher (1979), who made similar points earlier, Philip Tetlock (2005) undertook to improve forecasting methodology.

With his team of statisticians and social scientists—including Barbara Mellers and Don Moore—Tetlock was able to propose an adventurous and ambitious program of research for the then newly formed IARPA (Intelligence Research Advanced Projects Agency), a research arm of the US intelligence community focused on addressing hard problems with open-source efforts between scientists and those in the intelligence community. The Good Judgment Project (goodjudgmentproject.org) was the result. Essentially, the effort is to provide specific forecasting questions that (a) have a discernible outcome (e.g., Will Kim Jung Un preside at the May 1 parade in Pyongyang in 2014?) and (b) have a precise time of resolution. These two characteristics of the forecast can be judged independently from the forecast itself (a problem when forecasts are self-evaluated). About 4,000 people have signed up for this tournament. They are also trained to use the computing platform. Each individual is made part of a group, and these groups as well as the individuals compete in terms of their accuracy (measured largely by Brier scores).

Two major lessons are gleaned from the experiments that are embedded in these tournaments. One is that understanding the base rates at which some phenomena occur is very important. Without an understanding of what the base rate is, it becomes difficult to tell whether an event is "normal" or exceptional. Is a protest of 1,000 people in Tahrir Square in Cairo something that is unusual, or does it happen on a monthly basis, for example? The second lesson follows from the first: How do you identify an exception? Just these two skills, which can be easily taught, cause a change in thinking about what is likely to happen. And it turns out, they also help individuals make more accurate forecasts. Two other things are important. One of them is the simple act of keeping track of your successes and failures, something that is unlikely to receive widespread assent by media stars and pundits. The second is more prosaic: Aggregations help reduce bias and uncertainty. The classic example of this is the 1906 livestock fair in which 800 individuals guessed the weight of an ox on display. No single individual got the right answer, but according to Galton, both the mean and the median of these 800 estimates were within one percent of the correct weight (Galton 1907).

This leads to four points:

1. Base rates are important contexts for predictions.
2. Exceptions to base rates are important to identify; however, what an exception is remains complicated.
3. Keeping track of successes and failures is important.
4. Having a lot of answers may be the only way to get the right answer, but doing so may require some aggregation from a variety of perspectives.

How are Tetlock's forecasters doing in this competition with prediction markets, subject matter experts, and other unspecified "work products" of the intelligence community? So far, there are five independent research teams, of which Tetlock's group is one. This experiment involves over one million judgments from over 10,000 participants. This is based on about one hundred real events each year, for example "Will Iran sign an IAEA Structured Approach document before June 1, 2014?" A systematic comparison of the various approaches is underway. In short:

- Aggregations of trained forecasters beat subject matter experts by about fifty percent, and they also beat prediction markets.
- Prediction markets significantly out-perform subject matter experts. Domain expertise is less important than problem solving and belief updating.
- The top forecasters—the so-called super forecasters—tend to be the same individuals each year.

Another IARPA project is also interesting to examine. It is known as the OSI, Open Source Indicators, project, in which the basic idea is that classified information is possibly inferior to open-source materials for making important predictions. Thus, an attempt has been made to undertake detailed, precise forecasts of the type often engaged in within the intelligence community, while only using materials found in the unclassified world. The goal is to develop and test methods for the automated analysis of publicly available data to anticipate or detect significant social events. Such events include disease outbreaks, political instability, and elections. The goal is kind of a version of Google FluTrends on steroids: beat the news by fusing early indicators of events from diverse data (Ramakrishnan et al. 2014). They are less interested in theory, but do base their predictions on exactly the kind of data that are widely used in the social sciences.

Some recent predictive successes include:

1. Riots after impeachment of Paraguay's president (2012)

2. The so-called "Brazilian Spring" (June 2013)
3. Hantavirus outbreaks in Argentina and Chile (2013)
4. Venezuelan student uprising (Feb 2014)

What is remarkable here is that these predictions are graded by independent subject matter experts reading the local, non-English text. This suggests that not just trained individuals, but also trained statistical models might be able to generate accurate and useful predictions.

Recent Work

Over the past eight years, there has been an effort to create a forecasting platform for decision-making within the defense and intelligence communities, again based on open sources. This is known as W-ICEWS, the Worldwide Integrated Crisis Early Warning System, often abbreviated ICEWS. It has a lot of components, but the most pertinent is the suite of models, developed by social scientists, that forecast major instability events around the world with high accuracy. The basic idea is to use data that are detailed but aggregated to the month. These data are both structural and behavioral. The behavior data are event data produced using a tested ontology of categories and an automated procedure for constructing word graphs of stories order to glean context. The actor dictionary and verb dictionary are updated monthly, and the data are made available on a monthly basis. As of March 2015, these data are publicly available (Boschee et al. 2015; Lautenschlager et al. 2015; Lustick et al. 2015).

Each of three teams creates several models of political instability events. These events are rebellions, insurgencies, ethnic violence, and domestic and international crises. Several theme models are developed using contemporary social science literature. The themes each address different substantive arguments. For example, in developing models for domestic international crises, we generated themes that captured (a) the performance of the macro economy, (b) the demographics of each country, including youth bulges and other demographic factors, (c) the structure of the political system, (d) the institutional infrastructure of each country both politically and economically, (e) the scope and extent of rebellious activity, (f) the interactions between the standing government and various factions of opposition, and (g) the extent of economic vitality (or stagnation) in neighboring countries. Each of these themes produced predictions based on a limited but useful set of forces thought variously to be important. Without going into a lot of detail, we use these themes to produce predictive distributions, which we then combine using a special version

Table 1. Current metrics for out-of-sample results for W-ICEWS ensemble models

	Insurgency (%)	Rebellion (%)	Domestic political crisis (%)	Ethnic religious violence (%)	Dyadic international crisis (%)
Precision	89	84	68	98	96
Recall	80	88	49	71	95

of Ensemble Bayesian Modeling Averaging, which weights each theme in proportion to its ability to make accurate predictions in a test set of data set aside for this calibration test (Raftery et al. 2005; Vrugt et al. 2006; Vrugt, Diks, and Clark 2008; Fraley et al. 2010; Fraley, Raftery, and Gneiting 2010; Sloughter, Gneiting, and Raftery 2010; Montgomery, Hollenbach, and Ward 2013, 2014, 2015). Our predictions are then combined in the same way, by an independent group with the predictions made in different ways by different modelers (all, as it turns out, political scientists).

The so-called ICEWS component models developed at Duke (Ward et al. 2013; Ward 2015) produce accurate forecasts, and as expected, the ensemble forecasts are generally more accurate, in terms of having both fewer false negatives and positives than any individual model. They are combined with other models within the ICEWS project, again using Bayesian ensemble techniques. The result is a model that preserves transparency, yet maximizes prediction. At the same time, it is based on explanations for political stability that are found in the literature. The result of this approach has been successful and more accurate than other efforts that tried to find the best *single* model. Table 1 illustrates these results, which pertain to six-month predictions for 167 countries each month.

What we have learned in the project so far is that aggregation of different diverse models produces better predictions than models that are singular in their explanation of what is occurring in the world. Further, this principle seems to scale up as well as down, and different models can benefit from aggregation as well as diverse approaches. It may be that there is no one best model from a forecasting perspective. Moreover, we find that like the weather, other forms of complicated phenomena that exhibit violent variation—conflict and crisis—can also be explained and predicted.

There are many more political forecasting projects that I will not discuss in detail at present, including the attempt to forecast conflict decades into the future. These include Choucri (1974), Hildebrand, Laing, and Rosenthal (1976), Choucri and Robinson (1978), Freeman and Job (1979), and Singer and Wallace (1979) in the 1970s, and Vincent (1980), Gurr and Lichbach (1986), Cioffi-Revilla

(1996), Davies and Gurr (1998), and Pevehouse and Goldstein (1999) in the 1980s and 1990s.

More recently, in this century, there have been many ongoing efforts at forecasting.[5] In April 2015, new predictions of weather and conflict in Africa at the 1 deg grid (Witmer et al. 2015) were presented at a conference at PRIO (Peace Research Institute Oslo) along with a variety of other new studies (Beger and Ward 2015; Brandt 2015; Colaresi 2015; Gleditsch 2015; Hegre, Høyland, and Nygård 2015; Hunziker 2015; Schneider 2015; Schutte 2015) all focusing explicitly on prediction of conflict (See also the Joshi et al. scenarios about improved governance [Joshi, Hughes, and Sisk 2015]).

There will be many successes (and failures) in the years ahead as the policy and scholarly world turns to a more rigorous examination of their understandings through predictions, not just locally in elections, but globally as well in a wider range of human endeavors.

The bottom line with theory is that for the most part, it looks back at the world and describes how things occurred. Often, it attempts to develop a teleology of why things occurred. Rarely is there a track record kept of whether the theory was borne out in cases not initially considered and where there is some ex ante evaluation, the theoretical response is often to add greater nuance to encompass the situations or cases that were not well covered. Theory building does not generally lead to new ideas. This is a well-known facet of inquiry that has been presented and debated for decades, if not longer. Predictions, failed as well as successful, have greater potential for moving the needle on what we know. The cure for bad theory is more theory; the cure for bad predictions is better and often different understandings of what is implied by the underlying (and changing) social dynamics.

5   See Gleditsch and Ward (2000), de Mesquita (2002, 2009, 2011), Feder (2002), O'Brien (2002), Ward and Gleditsch (2002), de Marchi, Gelpi, and Grynaviski (2004), Brandt, Colaresi, and Freeman (2008), Hegre (2008), Schneider, Gleditsch, and Carey (2010), Weidmann and Ward (2010), Ulfelder (2012), Gleditsch and Ward (2013), Hegre et al. (2013), Blair, Blattman, and Hartman (2014), and Brandt, Freeman, and Schrodt (2014).

Data Constraints

One of the difficulties of doing predictions is that data are often hard to obtain for the future. More practically, all quantitative and quality investigations are limited to information about the past. The typical response to this kind of approach is to employ all the available information, because more information is always better. In the quantitative world, this results in an undesirable dependence between the data collected and the model posited. This dependence typically results in overfitting in which the model has been sculpted to the available data but might not characterize additional data (such as the future or different cases). The same can happen qualitatively wherein new cases are sought to fit different aspects of the theory being examined or constructed.

Even if we do not have time to wait for data on the future to appear as a way of examining the external validity of our explanations, we still have the past. Ideally, we should divide our cases and samples into different groups, one of which we use to estimate the model and evaluate its characteristics as we refine or estimate it statistically. These cases are known as the training cases. A second set of observations, known as the test data, can be then examined to see if they also conform to the theory/model. This procedure, by now a gold standard in most of science, is known as cross-validation and serves as a way to anneal one's investigations against being too tied to the data at hand.

What Are the Pros and Cons of a Predictive Approach to Social Science?

The main pro is that the predictive enterprise helps us evaluate how well we are doing so that we can improve our understanding of the world. It is the gold standard of a scientific approach. We do not yet have an experimental framework for many important subjects. As a result, it is important to make sure we can get the same kind of results with new information that we got with the data we began investigating. That means we have to either save some data back (a great idea), use the future to see how well our modeled understandings perform, or preferably both. There are only nascent traditions of this in the social sciences at present. Keeping track of your success is not collecting significant coefficients. *Keeping track matters.* One consequence is that we cannot just keep using the same data over and over. And over. One reason that many hate predictions is that talking heads make many predictions in the media, but few of them ever keep track of how well they are doing. Their goal is somewhat akin to a venture capitalist's make enough bets that eventually one of them is correct enough that you get to make a lot more bets. Ascher (1979) long ago showed that the talking heads were most often wrong. This is still true. You would think that they would get better over time, but there is little evidence that this is the case.

We also will be driven into making more precise investigations once we start to predict. We will not be satisfied with annual data for most things. Nor will we necessarily be satisfied with national-level information because it becomes even more apparent in the predictive domain that the world is neither flat nor homogeneous. As a result, we should get more precise understandings of how things play out in our social world. At the same time, we have to recognize that our predictions are probabilistic and contain a large amount of uncertainty, more so than in other endeavors.

As a result of these two aspects, better and more precise understandings of our social world, it is possible to be more relevant to decision makers at all levels. This does not mean just inside the beltway. It also means decision makers at CDC (Centers for Disease Control) as well as those in non-governmental organizations around the world.

What are the cons? Several arguments are usually brought to the fore.

1. *The world is inherently unpredictable.* But we are studying it anyway. Go figure. The refrain to this litany is often "but I know what is going to happen in this instance." Maybe, but let us keep track and find out if you are right. This is the talking heads premise, and it is demonstrably false. Making cause and effect statements about politics does imply that politics is in part, at least, predictable.

2. *This will empower the establishment and impoverish those without power.* Actually, it might. But at the same time, it provides ways in which those outside the capitals can also affect the future. Why will prediction be more valuable to the establishment than it is to the rest of society? Is the same thing true of explanation and substantive knowledge? Western society is based in part on the idea that knowledge is a valuable thing for all. It is true that some take more advantage of it than others. But having open and available knowledge can be important for many diverse groups. As an example, we might think that clandestine organizations can benefit from open knowledge, even precise actionable knowledge, but as we think these thoughts, most of us might not be thinking about transnational activist networks but rather large governmental organizations. However,

it is clear that non-governmental actors are also consumers of knowledge.

3. *It is possible to predict things without true understanding or knowledge.* Sure, but rarely is this true for any but the simplest of systems. I am reminded of the wonderful essay by Calvin Trillin describing the chicken on Mott Street in New York's Chinatown that played and always won Tic-Tac-Toe. The chicken did not understand the game. But the chicken always won.[6] It is ridiculous to suggest that we have models that predict as accurately as the (now gone) Mott Street chicken but have the same understanding of the "chicken" game. Even if it were the case (and I repeat it is not), could the opposite really be true? If you have deep understanding of the world, should you not be able to generate accurate predictions of how it will work in situations you have not seen before? Will proximate effects lead us toward distant causes much like proximate causes can lead us to distant effects?

4. *We will disrupt the space-time continuum.* If we can predict conflict, for example, we will be able to prevent it. Or start it where we want. And then we will no longer be able to predict conflict. One of my models attempts to predict where there will be coups de état and other types of irregular regime changes on a monthly level. Maybe if Muhammadu Buhari, who assumed the presidency of Nigeria at the end of May 2015, sees our manuscript, he might be able to prevent any irregular leadership change from occurring in the next six months. And maybe that would be bad. Or good. But in any case, I am pretty sure that the Nigerian president is already aware of the fragility of the Nigerian political landscape. This is a frequent type of criticism of forecasting. I think we can wait for this to become a real problem before we stop trying to develop better understandings of the world.

5. *Real social effects occur glacially. Predictions will be focused on epiphenomenal changes that won't matter in the long run.* How do you know?

In summary, we need less theory because most theory is an attempt to rescue or adapt extant theory. We need more predictions in order to keep track of how well we understand the world around us. They will tell us how good our theories are and where we need better explanations. Predictions are like cell phones. First, they seem arcane and bizarre. Then, in a few short years, there is no one around

---

6 And if it did not, the winner was ridiculed for being marginally smarter than a chicken.

who remembers life without cell phones and your kids use them in ways you don't understand. The 2012 US presidential election was the first that was famously and accurately predicted. But it will be the first of many. All future voters will vote in an era in which accurately predicting the election will be the norm, not the exception, though as in the UK in 2015, there will be exceptions. This will have consequences for democracy. In the same way that having a product recommended to us on the web is now normal, this will become the new normal. Data science (and more data) will guide us to a better understanding of our future than we have now. Whether you are involved with commercial organizations, local government, non-governmental organizations (NGOs), the federal government, or international organizations, prediction will be part of the daily ebb and flow of information, and we shall become used to seeing accurate predictions about a wide variety of political phenomena.

But, as we get more accurate, will we be able to begin manipulating outcomes? Engineer results? It may not seem like it to everyone, but political beliefs are malleable. In fact, survey researchers are feigning shock at discovering that political surveys tend to politicize respondents. Republicans can turn into Democrats, and vice versa. Will we get equally adept at predicting what kinds of information, interactions, and initiatives will turn the tide in a particular election? Facebook and Google—and many other less famous firms—think so, and are gearing up for the 2016 election with tools that go way beyond surveys that can be used for that purpose.

What do we expect to see in the global security system? First, we know that a wider variety of actors will be consuming and generating data on their activities. Indeed, we know that a wider variety of national and non-state actors are using predictive models of behavior that might broadly be considered in the realm of political violence, ranging from strikes and protests to attacks and casualties. China, Russia, the European Union, and the UK, as well as the United States and the United Nations and the North Atlantic Treaty Organization, all have substantial forecasting capabilities in the global realm. Some of these forecasts are made on a weekly basis, and others are more long term, looking out a couple of decades. But the fact that predictive heuristics are now part and parcel of normal statecraft is important and recent. Moreover, non-state actors are not to be left out and are beginning to evolve toward greater foresight. Consider that an International Red Cross that is able to predict domestic conflicts will be better able to pre-position supplies and expertise to deal with the human toll of such conflicts. A monitoring of violent

government actions that can predict the safest time to remove NGO personnel in conflict zones is another example of a predictive tool that can affect the global security system in new and beneficial ways.

The global security system is complicated, multilayered, unknown, and changing. In some ways, it is exactly like the solar system. However, it may change more quickly but maybe less dramatically.[7] By developing explanations and subjecting them to critical evaluations, we learned more. We can learn more again. We can build Antikythera Mechanisms. Though they may or may not tell us about the next sociopolitical eclipse in the international security system, they are likely to help us get rid of bad ideas.

Acknowledgments

References

Altman, Daniel. 2015. "The Strategist Curse: A Theory of False Optimism as a Cause of War." *Security Studies* 24(2): 284–315.

Ascher, William. 1979. *Forecasting, an Appraisal for Policy-Makers and Planners*. Baltimore, MD: Johns Hopkins University Press.

Ayer, Alfred Jules. 1936. *Language, Truth, and Logic*. London: Gollancz.

Beckley, Michael. 2015. "The Myth of Entangling Alliances: Reassessing the Security Risks of U.S. Defense Pacts." *International Security* 39(4): 7–48.

Beger, Andreas, and Michael D. Ward. 2015. "Irregular Leadership Changes in 2015." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Blair, Robert, Christopher Blattman, and Alexandra Hartman. 2014. *Predicting Local-Level Violence*. SSRN. http://ssrn.com/abstract=2497153.

Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz, and Michael D. Ward. 2015. *ICEWS Coded Event Data* (Harvard Dataverse Network [Distributor] V1 [Version]). http://dx.doi.org/10.7910/DVN/28075.

Brandt, Patrick. 2015. "Forecasting Conflicts: Long and Short Term Predictions Based on Different Training Set Considerations." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Brandt, Patrick T., Michael Colaresi, and John R. Freeman. 2008. "The Dynamics of Reciprocity, Accountability, and Credibility." *Journal of Conflict Resolution* 52(3): 343–74.

Brandt, Patrick T., John R. Freeman, and Philip A. Schrodt. 2014. "Evaluating Forecasts of Political Conflict Dynamics." *International Journal of Forecasting* 30(4): 944–62.

Choucri, Nazli. 1974. "Forecasting in International Relations: Problems and Prospects." *International Interactions* 1(2): 63–68.

Choucri, Nazli, and Thomas W. Robinson. 1978. *Forecasting in International Relations: Theory, Methods, Problems, Prospects*. San Francisco: W. H. Freeman.

Cioffi-Revilla, Claudio. 1996. "On the Likely Magnitude, Extent, and Duration of an Iraq-UN War." *Journal of Conflict Resolution* 35(3): 387–411.

Colaresi, Michael. 2015. "Using Explicit Model Criticism to Build and Improve an Open Source, Policy-Relevant Forecast of Civil Conflict." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Comte, Auguste. 1846. *General View of Positivism*. London: Trübner.

Davies, John L., and Ted Robert Gurr. 1998. *Preventive Measures: Building Risk Assessment and Crisis Early Warning Systems*. Lanham, MD: Rowman & Littlefield Publishers.

de Caritat (Marquis de Condorcet), Marie Jean Antoine Nicolas. 1795. *Création de Esquisse d'un tableau historique des progrès de l'esprit humain*. Translated and edited by de Caritat (Marquis de Condorcet), Marie Jean Antoine Nicolas. Philadelphia: Lang and Uttick.

de Marchi, Scott, Christopher Gelpi, and Jeffrey D. Grynaviski. 2004. "Untangling Neural Nets." *American Political Science Review* 98(2): 371–78.

de Mesquita, Bruce Bueno. 2002. *Predicting Politics*. Columbus: Ohio State University Press.

——. 2009. *The Predictioneer's Game*. New York: Random House.

——. 2011. "A New Model for Predicting Policy Choices: Preliminary Tests." *Conflict Management and Peace Science* 28(1): 65–87.

Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

——. 2007. *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Feder, Stanley A. 2002. "Forecasting For Policy Making in the Post–Cold War Period." *Annual Review of Political Science* 5: 111–25.

Fraley, Chris, Adrian E. Raftery, and Tilmann Gneiting. 2010. "Calibrating Multimodel Forecast Ensembles with Exchangeable and Missing Members Using Bayesian Model Averaging." *Monthly Weather Review* 138(1): 190–202.

Fraley, Chris, Adrian E. Raftery, J. McLean Sloughter, and Tilmann Gneiting. 2010. *ensembleBMA: Probabilistic Forecasting using Ensembles and Bayesian Model Averaging* (R package version 4.5). http://CRAN.R-project.org/package=ensembleBMA.

---

7 Though there is a genre of science fiction that suggests otherwise. Stephenson (2015).

Freeman, John R., and Brian L. Job. 1979. "Scientific Forecasts in International Relations: Problems of Definition and Epistemology." *International Studies Quarterly* 23(1): 113–43.

Freeth, Tony, Yanis Bitsakis, Xenophon Moussas, John H. Seiradakis, A. Tselikas, H. Mangou, M. Zafeiropouou, R. Hadland, D. Bate, A. Ramsey, M. Allen, A. Crawley, P. Hockley, T. Malzbender, D. Gelb, W. Ambrisco, and M. G. Edmunds. 2006. "Decoding the Ancient Greek Astronomical Calculator Known as the Antikythera Mechanism." *Nature* 444(Supplement 7119): 587–91.

Galton, Francis. 1907. "Vox populi." *Nature* 75(1949): 450–51.

Gleditsch, Kristian Skrede. 2015. "Predictoing UCDP/PRIO Civil Wars: Expanding the Inequality and Grievance Model with Event Data." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Gleditsch, Kristian Skrede, and Michael D. Ward. 2000. "War and Peace in Space and Time: The Role of Democratization." *International Studies Quarterly* 44(1): 1–29.

——. 2010. "Contentious Issues and Forecasting Interstate Disputes." Presented at the 2010 Annual Meeting of the International Studies Association. New Orleans.

——. 2013. "Forecasting Is Difficult, Especially the Future: Using Contentious Issues to Forecast Interstate Disputes." *Journal of Peace Research* 50(1): 17–31.

Gurr, Ted Robert. 1970. *Why Men Rebel*. Princeton, NJ: Princeton University Press.

Gurr, Ted Robert, and Mark Irving Lichbach. 1986. "Forecasting Internal Conflict: A Competitive Evaluation of Empirical Theories." *Comparative Political Studies* 19(3): 3–38.

Healy, Kieran. 2015. "Fuck Nuance." Working Paper. Accessed December 2, 2015. http://kieranhealy.org/files/papers/fuck-nuance.pdf.

Hegre, Håvard. 2008. "Gravitating toward War: Preponderance May Pacify, but Power Kills." *Journal of Conflict Resolution* 52(4): 566–89.

Hegre, Håvard, Bjørn Høyland, and Håvard Nygård. 2015. "Dynamic Prediction of Armed Conflict, 2015–2065." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Hegre, Håvard, Joakim Karlsen, Håvard Mokleiv Nygård, Håvard Strand, and Henrik Urdal. 2013. "Predicting Armed Conflict, 2011–2050." *International Studies Quarterly* 57(2): 250–70.

Hempel, Carl G. 1935. "Über den Gehalt von Wahrscheinlichkeitsaussagen." *Erkenntnis* 5: 228–60.

Herz, John H. 1950. "Idealist Internationalism and the Security Dilemma." *World Politics* 2(January): 157–80.

Hildebrand, David K., James D. Laing, and Howard Rosenthal. 1976. "Prediction Analysis in Political Research." *American Political Science Review* 70(2): 509–35.

Hunziker, Philip. 2015. "Online Change Detection for Conflict Event Data: Towards an Early Alarm System." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Jervis, Robert. 1978. "Cooperation under the Security Dilemma." *World Politics* 30(2): 167–214.

Joshi, Devin, Barry B. Hughes, and Timothy Sisk. 2015. "Improving Governance for the Post-2015 Sustainable Development Goals: Scenario Forecasting the Next 50 Years." *World Development* 70: 286–302.

King, Gary, and Langche Zeng. 2001. "Improving Forecasts of State Failure." *World Politics* 53(4): 623–58.

Koubi, Vally, and Tobias Böhmelt. 2014. "Grievances, Economic Wealth, and Civil Conflict." *Journal of Peace Research* 51(1): 19–33.

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Lautenschlager, Jennifer, Steve Shellman, and Michael D. Ward. 2015. *ICEWS Coded Event Aggregations* [Harvard Dataverse Network (Distributor) V1 (Version)]. http://dx.doi.org/10.7910/DVN/28117.

Liff, Adam P., and G. John Ikenberry. 2014. "Racing toward Tragedy? China's Rise, Military Competition in the Asia Pacific, and the Security Dilemma." *International Security* 2(Fall): 52–91.

Lustick, Ian, Sean O'Brien, Steve Shellman, Timothy Siedlecki, and Michael D. Ward. 2015. *ICEWS Events of Interest Ground Truth Data Set* [Harvard Dataverse Network (Distributor) V1 (Version)]. http://dx.doi.org/10.7910/DVN/28119.

Mahoney, James. 2015. "Process Tracing and Historical Explanation." *Security Studies* 24(2): 200–218.

Metternich, Nils W., Cassy Dorff, Max Gallop, Simon Weschle, and Michael D. Ward. 2013. "Anti-Government Networks in Civil Conflicts: How Network Structures Affect Conflictual Behavior." *American Journal of Political Science* 57(3): 892–911.

Mill, John Stuart. 1843. *A System of Logic*. London: John W. Parker, West Strand.

Milstein, Jeffery S. 1974. *Dynamics of the Vietnam War: A Quantitative Analysis and Predictive Computer Simulation*. Columbus: Ohio State University Press.

Montgomery, Jacob M., Florian M. Hollenbach, and Michael D. Ward 2012. "Improving Predictions Using Ensemble Bayesian Model Averaging." *Political Analysis* 20(3): 271–91.

——. 2013. *Ensemble BMA Forecasting*. http://cran.r-project.org/web/packages/EBMAforecast/EBMAforecast.pdf.

——. 2014. "Calibrating Ensemble Forecasting Models with Sparse Data in the Social Sciences." *International Journal of Forecasting*, in press.

——. 2015. "Calibrating Ensemble Forecasting Models with Sparse Data in the Social Sciences." *International Journal of Forecasting* 31(3): 930–42.

O'Brien, Sean P. 2002. "Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis." *Journal of Conflict Resolution* 46(6): 791–811.

Pevehouse, Jon C., and Joshua S. Goldstein. 1999. "Serbian Compliance or Defiance in Kosovo? Statistical Analysis and Real-Time Predictions." *Journal of Conflict Resolution* 43(4): 538–46.

Pilster, Ulrich, and Tobias Böhmelt. 2014. "Predicting the Duration of the Syrian Insurgency." *Research & Politics* 1(2). doi:10.1177/2053168014544586.

Popper, Karl. 1935. *Logik der Forschung: Zur Erkenntnistheorie der Modernen Naturwissenschaft*. Number Band 9. Berlin: Springer Verlag.

Raftery, Adrian E., Tilmann Gneiting, Fadoua Balabdaoui, and Michael Polakowski. 2005. "Using Bayesian Model Averaging to Calibrate Forecast Ensembles." *Monthly Weather Review* 133(5): 1155–74.

Ramakrishnan, Naren, Patrick Butler, Sathappan Muthiah, Nathan Self, Rupinder Paul Khandpur, Parang Saraf, Wei Wang, Jose Cadena, Anil Vullikanti, Gizem Korkmaz, Chris J. Kuhlman, Achla Marathe, Liang Zhao, Ting Hua, Feng Chen, Chang-Tien Lu, Bert Huang, Aravind Srinivasan, Khoa Trinh, Lise Getoor, Graham Katz, Andy Doyle, Chris Ackermann, Ilya Zavorin, Jim Ford, Kristen Maria Summers, Youssef Fayed, Jaime Arredondo, Dipak Gupta, and David Mares. 2014. "'Beating the News' with EMBERS: Forecasting Civil Unrest using Open Source Indicators." *CoRR* abs/1402.7035. http://arxiv.org/abs/1402.7035.

Richardson, Lewis F. 1960. *Arms and Insecurity*. Chicago and Pittsburgh: Quadrangle/Boxwood.

Schneider, Gerald. 2015. "The Oracle or the Crowd? Experts versus Markets in Forecasting Conflict and Cooperation." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Schneider, Gerald, Nils Petter Gleditsch, and Sabine C. Carey. 2010. "Exploring the Past, Anticipating the Future: A Symposium." *International Studies Review* 12(1): 1–7. http://search.ebscohost.com/login.aspx?direct=true&db=poh&AN=48465758&site ehost-live&scope=site.

Schutte, Sebastian. 2014. "Geography, Outcome, and Casualties: A Unified Model of Insurgency." *Journal of Conflict Resolution* 59(6): 1101–28.

———. 2015. "Regions at Risk: Predicting Conflict Zones in African Insurgencies." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.

Singer, J. David, and Michael D. Wallace. 1979. *To Augur Well: Early Warning Indicators in World Politics*. Beverly Hills, CA: Sage Publications.

Sloughter, J. McLean, Tilmann Gneiting, and Adrian E. Raftery. 2010. "Probabilistic Wind Speed Forecasting Using Ensembles and Bayesian Model Averaging." *Journal of the American Statistical Association* 105(489): 25–35.

Stephenson, Neal. 2015. *Seveneves*. New York: Harper Collins.

Tetlock, Philip. 2005. *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.

———. 2010. "Second Thoughts about *Expert Political Judgment*: Reply to the Symposium." *Critical Review* 22(4): 467–88.

Ulfelder, Jay. 2012. "Forecasting Onset of Mass Killings." Paper presented at the Annual Northeast Political Methodology Meeting at New York University.

Vasquez, John A. 1997. "The Realist Paradigm and Degenerative versus Progressive Research Programs: An Appraisal of Neotraditional Research on Waltz's Balancing Proposition." *American Political Science Review* 91(4): 899–912.

Vincent, Jack E. 1980. "Scientific Prediction versus Crystal Ball Gazing: Can the Unknown Be Known?" *International Studies Quarterly* 24(3): 450–54.

Vrugt, Jasper A., Martyn P. Clark, Cees G. H. Diks, Qinyun Duan, and Bruce A. Robinson. 2006. "Multi-Objective Calibration of Forecast Ensembles Using Bayesian Model Averaging." *Geophysical Research Letters* 33: L19817.

Vrugt, Jasper A., Cees G. H. Diks, and Martyn P. Clark. 2008. "Ensemble Bayesian Model Averaging Using Markov Chain Monte Carlo Sampling." *Environmental Fluid Mechanics* 8(5): 579–95.

Waltz, Kenneth N. 1997. "Evaluating Theories." *American Political Science Review* 91(4): 913–17.

Ward, Michael D. 2015. Monthly Predictions of Conflict in 167 Countries, March 2015. Monthly Report 4/2015 mdwardlab.com. http://mdwardlab.com.

Ward, Michael D., and Kristian Skrede Gleditsch. 2002. "Location, Location, Location: An MCMC Approach to Modeling the Spatial Context of War and Peace." *Political Analysis* 10(3): 244–60.

Ward, Michael D., Nils W. Metternich, Cassy L. Dorff, Max Gallop, Florian M. Hollenbach, Anna Schultz, and Simon Weschle. 2013. "Learning from the Past and Stepping into the Future: Toward a New Generation of Conflict Prediction." *International Studies Review* 16(4): 473–644.

Weidmann, Nils B., and Michael D. Ward. 2010. "Predicting Conflict in Space and Time." *Journal of Conflict Resolution* 54(6): 883–901.

Weil, Herman M. 1974. "Domestic and International Violence: A forecasting approach." *Futures* 6(6): 477–85.

Witmer, Frank D. W., John V. O'Loughlin, Andrew M. Linke, Arlene Laing, and Andrew Gettelman. 2015. "Climate-Sensitive Violent Conflict Forecasts for Sub-Saharan Africa, 2015–2050." In *Conference on Forecasting and Early Warning of Conflict and Political Instability*. Oslo: Peace Research Institute.