

Introduction: Forecasting in peace research

Håvard Hegre

Uppsala University & Peace Research Institute Oslo (PRIO)

Nils W Metternich

Department of Political Science, University College London

Håvard Mogleiv Nygård

Peace Research Institute Oslo (PRIO)

Julian Wucherpfennig

Hertie School of Governance

Journal of Peace Research
2017, Vol. 54(2) 113–124
© The Author(s) 2017
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0022343317691330
journals.sagepub.com/home/jpr



Abstract

Prediction and forecasting have now fully reached peace and conflict research. We define forecasting as predictions about unrealized outcomes given model estimates from realized data, and predictions more generally as the assignment of probability distributions to realized or unrealized outcomes. Increasingly, scholars present within- and out-of-sample prediction results in their publications and sometimes even forecasts for unrealized, future outcomes. The articles in this special issue demonstrate the ability of current approaches to forecast events of interest and contributes to the formulation of best practices for forecasting within peace research. We highlight the role of forecasting for theory evaluation and as a bridge between academics and policymakers, summarize the contributions in the special issue, and provide some thoughts on how research on forecasting in peace research should proceed. We suggest some best practices, noting the importance of theory development, interpretability of models, replicability of results, and data collection.

Keywords

forecasting, out of sample evaluation, peace research, prediction, theory testing

No matter how I turn it over in my mind, the number one task of peace research always turns out to be that of prediction [...] (J David Singer, 1973)

Although rarely articulated explicitly, ultimately being able to forecast peace and conflict is a fundamental motivation for peace research (Singer, 1973). Indeed, reliable forecasting or early-warning systems that could indicate risks before conflict erupts or escalates would make it possible to prepare for, intervene in, or build resilience against deadly conflicts (Harff, 2003). Moreover, by designing forecasts as contingent on policy interventions, they could be powerful tools to guide policy: by comparing forecasts of the expected risk or intensity of armed conflict given different conflict-preventing policies, the risks

associated with particular interventions can be more fully understood, and the case for (non-)intervention made evidence based.

Forecasting peace and conflict was long viewed with considerable skepticism and often considered unfeasible (e.g. Stephens, 2012). However, new data projects, new theories, and innovative methods – as demonstrated in this special issue – are taking us closer to generating conflict forecasts that are sufficiently precise to be policy relevant. We focus on forecasts of phenomena that are sufficiently regular and frequent to support the

Corresponding author:
havard.hegre@pcr.uu.se

estimation of statistical models that typically requires large-N datasets. 'Black swans' (Taleb, 2007) such as the onset of a world war or the collapse of a superpower are typically too infrequent to qualify as such, and rather tend to fall into the realm of 'judgemental forecasters' (Tetlock, 2005). We do briefly discuss how large-N work can be made relevant for rare, high-impact events, provided that such events can be credibly construed as an agglomeration of smaller, more regular events.

What does the term 'forecasting' mean in peace and conflict research? The usage in the literature varies somewhat. We here define forecasts as *predictions about unrealized outcomes given model estimates from realized data*. 'Early-warning systems' we define as systematic procedures set up to provide regular forecasts for conflict-related events along the lines of, for instance, daily weather forecasts. 'Prediction' is a more general concept, and refers to the assignment of a probability distribution to an outcome based on such model estimates, but may be applied to realized as well as unrealized outcomes. More colloquially, forecasts are predictions about tomorrow given information we have about what has happened up to today. This means two inputs are required to make forecasts: realized data and estimators; and one output is produced: predictions.

Understood this way, forecasting and prediction have now fully arrived in the field of peace and conflict research (Schneider, Gleditsch & Carey, 2011; Metternich & Gleditsch, forthcoming). Increasingly, scholars present prediction results and forecasts (O'Brien, 2010; Brandt, Freeman, & Schrodt, 2011; Schrodt, Yonamine & Bagozzi, 2013), while specialized conferences and workshops are addressing forecasting frameworks. This focus is mirrored in the policymaking world that has benefited from scholarly work on forecasting (King & Zeng, 2001; Harff, 2003; Goldsmith et al., 2013; Bell et al., 2013). Beyond structural (e.g. Beger, Dorff & Ward, forthcoming; Goldstone et al., 2010) and time-series forecasting designs (e.g. Metternich et al., 2013; Brandt, Colaresi & Freeman, 2008), rational choice (e.g. Bueno de Mesquita, 2011) and judgmental forecasts (e.g. Tetlock, 2005) have made their way from the academic to the policy world. In fact, many large international organizations and governments rely on regional or global forecasts of conflict in order to address humanitarian, military, and political crises. However, the quality of these efforts outside of academia is hard to assess, as they are often not transparent or replicable, if not outright secret, and the methodologies employed have rarely been subjected to the scrutiny of academic peer review. Despite the recent surge of large prediction and forecasting efforts

(Boschee et al., 2015; Doyle et al., 2014; Goldstone et al., 2010; De Groeve, Hachemer & Vernaccini, 2014), the discipline has not overcome the challenge with a lack of shared standards as well as tools for assessing and comparing predictive performance (Ward, Greenhill & Bakke, 2010; Brandt, Freeman, & Schrodt, 2014; Carment, 2003). In addition to demonstrating the ability of current approaches to forecast events of interest, and their implications for prediction-based public policies, this special issue contributes to filling this gap by laying out best practices in conflict forecasting.

A brief history of forecasting in peace research

Systematic conflict forecasting is not new and is deeply rooted in the systematic study of peace and conflict (Choucri, 1974; Bueno de Mesquita, Newman & Rabushka, 1985; Gurr & Lichbach, 1986; Bremer, 1987). We find it useful to think of the history of forecasting in the peace research literature as three generations of studies.

The first generation of conflict prediction was inspired by the work of Sorokin ([1957] 1962), Richardson (1960a), and Wright ([1942] 1965). It was heavily influenced by the foundation of the Correlates of War Project in 1963 aiming to systematically accumulate scientific knowledge about war (Small & Singer, 1982). Early-warning purposes were explicitly among the aims of this effort (Singer & Wallace, 1979). Early events-data projects (that collect data on individual events of the size typically reported by an individual news report) also highlighted forecasting (e.g. Azar et al., 1977). These efforts, pioneered by Azar (1980) and McClelland & Hoggard (1968), provided templates for collecting fine-grained data sufficiently effective to approximate real-time conflict early warning.

This first enthusiasm for conflict prediction faded, however, and throughout the 1970s and early 1980s explicit efforts to use statistical models to predict or warn against armed conflict were relatively rare.¹

The second generation of conflict prediction contributed especially two critical innovations. Bueno de Mesquita (1980, 1983, 1984) made explicit the link between theory and conflict prediction by using game-theoretical models to predict armed conflict as well as other foreign and domestic policy events. In addition, from the late 1980s Philip Schrodt has been building statistical models based on extensive news source data to predict armed conflict. Schrodt (1988, 1991) used methods from

¹ Zinnes & Muncaster (1984) and Ward (1984) are among the few cases of forecasting efforts during this period.

artificial intelligence and machine learning, including neural networks, to predict state-based conflict. Such methods are now increasingly being used in the discipline. Schrodt was also a pioneer in moving away from the widely used country-year datasets constructed from the Correlates of War collection of data and similar sources.

Schrodt, Davis & Weddle (1994) introduced algorithms to automatically classify and code political events based on large numbers of news articles. These techniques have since been further refined and now allow the discipline to use increasingly more fine-grained data to code both dependent and independent variables. While the country-year format pushed the discipline forward (Gurr & Lichbach, 1986; Harff & Gurr, 1998; Gurr & Moore, 1997; Beck, King & Zeng, 2000), empirical analysis and forecasts alike are increasingly cast on a daily, weekly, or monthly level (e.g. Schrodt & Gerner, 2000; Brandt, Freeman & Schrodt, 2011; Doyle et al., 2014). This is reflected in the increasing demand for spatio-temporally disaggregated event data (Cederman & Gleditsch, 2009; Weidmann & Ward, 2010).

The focus on early warning garnered substantial interest in the policy community. The third generation of conflict prediction thus started with the development of the US government-financed State Failure Task Force (SFTF, later re-named the Political Instability Task Force PITF). The goal of the PITF was to predict a long range of political instabilities ranging from coups and revolutions to armed conflict two years before they occurred. Goldstone et al. (2010) conclude that the PITF studies 'have substantially achieved that objective'. Beginning in the mid to late 2000s, conflict prediction became a very active subdiscipline of conflict research and is now increasingly seen as a 'mainstream' effort by the wider scientific community (Schneider, Gleditsch & Carey, 2011). This has been aided by the realization, most succinctly communicated by Ward, Greenhill & Bakke (2010), that prediction often is a better way of evaluating research than more traditional significance and *p*-value based approaches, a discussion we return to below.

Prediction is now used throughout the discipline of peace and conflict research. Greatly helped by the advances in computationally intensive methods to collect and analyze data, researchers increasingly follow Phil Schrodt in using automated event coded data from news wires to study, for instance, how public opinion affects the Israeli–Palestinian conflict (Brandt, Colaresi & Freeman, 2008), or whether news data can be used to predict the outbreak of the First World War (Chadefaux, 2014).

The focus is not confined to armed conflict, but extends to predicting irregular leadership transfers (e.g. Beger, Dorff & Ward, forthcoming), one-sided violence (e.g. Scharpf et al., 2014), nonviolent movements (e.g. Chenoweth & Ulfelder, 2017), and many other forms of political violence (Ward et al. 2013) and its consequences. These studies have in common that they use data at a granular level (sometimes days or months instead of years) to predict conflict in the short term. Other studies rely on country-year data to produce long-range predictions. Hegre et al. (2013, 2016) forecast civil conflict many decades into the future, as do Witmer et al. (2017) in this issue. Forecasting is a thriving sub-discipline in peace and conflict studies and will, we forecast, continue to grow in the coming years. The most important questions regarding the future of forecasting in peace research pertain to its shape, not its importance. We believe that the interplay between theory and forecasting will increasingly take place alongside data and methods development as crucial elements of forecasting approaches in our discipline.

The shape of forecasting to come

The bulk of quantitative peace and conflict research has traditionally been interested in *explaining* the relationship between explanatory factors and outcomes of interest. Yet, the majority of applied statistical studies in our discipline focus on estimating marginal effects (along with their standard error) while almost completely disregarding the evaluation of model predictions. Forecasting puts the ability of researchers to generate predictions or predicted probability distributions for outcomes such as war, civil conflict, or one-sided violence at the forefront of the research agenda and seeks to limit the traditional exclusive reliance on statistical significance to assess scientific progress. Prediction can take many different shapes and with this special issue we want to take the opportunity to highlight areas where it increases our ability to explain and areas where researchers face a trade-off between prediction and explanation.

When evaluating the relationship between prediction and explanation it is important to recognize the different purposes of forecasting. Forecasting can help researchers to test, improve, and build their theories. However, forecasting not only fulfills scientific objectives; it also enables policymakers to formulate evidence-based policies regarding peace and security issues. Forecasts can help designing policies or act merely as an early-warning tool. Below we discuss both purposes.

Prediction to evaluate theory

Several of the articles in this issue show the utility of prediction for evaluating or testing theories or hypotheses. Ward, Greenhill & Bakke (2010) forcefully argue that the almost exclusive emphasis on classical hypothesis testing and analysis of p -values has undermined efforts to improve predicting the outcomes we are actually interested in (e.g. peace, armed conflict, or war). Although this problem has been discussed for decades,² there is an increasing awareness in the social sciences that the focus on statistical significance sometimes promotes findings that capture very small effects with limited ability to predict the outcomes of interest. Prediction provides one answer to this debate because researchers can evaluate the extent to which explanatory factors deemed theoretically important improve the prediction of the outcome. Hence, theoretically derived factors that are consistently associated with better predictions should increase the researcher's confidence about their substantive meaningfulness.³

Prediction is also a guard against overfitting, when combined with out-of-sample and cross-validation approaches. In this issue, for example, Blair, Blattman & Hartman (2017) use survey data to predict local-level violence. They compare a wide range of models and show that a more parsimonious model outperforms more extensive models. The problem is that p -values alone generally say little about the real-world impact of a variable or the concept it operationalizes. The danger is that an exclusive reliance on p -values in combination with larger datasets drives the discipline to identify an ever-growing list of increasingly marginal variables, since p -values are directly related to sample size. This is a problem that will only be compounded by the increasing availability of 'big data' sources. Contributors to this special issue show that their estimated effects are not just statistically significant, but also matter substantively, as their models help improve out-of-sample predictive performance.

Hegre, Nygård & Ræder (2017) extend this advantage through using forecasting techniques to study the size and intensity of the conflict trap. They argue that previous studies that focus on single parameters have underestimated the effect of the conflict trap. This also points to another benefit of prediction. Peace researchers, and indeed social scientists more generally, usually build

'models' to explain particular phenomena. In model testing, however, focus is often restricted to a single, or a few, parameters of interest. Prediction, in contrast, lends itself more easily to evaluating the power of an entire model, but also to the assessment of the predictive power of particular variables.

From a philosophy of science standpoint, we argue that the role of theory is central to the explanation and prediction of social behavior and that forecasting may help us to rigorously test theories. Based on Hempel, Schrodtt (2014: 290) takes this line of thought very far when arguing that 'explanation in the absence of prediction is not scientifically superior to predictive analysis, it isn't scientific at all! It is, instead, "pre-scientific"'. In the social sciences, where complete isolation of causal factors and their precise measurement are virtually impossible, this statement is likely too strong. Explanation in the absence of prediction is certainly possible, as is prediction without explanation (cf. Elster, 1989: 8–10; Tetlock, 2005: 14–15). Still, prediction can be a powerful tool to help us develop and improve theoretical explanations of conflict and peace as a supplement to hypothesis testing. Conversely, theoretical reasoning is essential to improve the predictive power of models without limiting their interpretability (for an example, see Gleditsch & Ward, 2013). By analyzing the characteristics of forecasts that do particularly well or particularly poorly in the out-of-sample evaluation, we can learn about the features of our models and theories that improve our understanding of the empirical data. Colaresi & Mahmood (2017), in this special issue, propose a modeling framework adapted from machine learning – build, compute, critique, and think – for doing just that.

This learning process promises more sophisticated models of how measurable explanatory factors are related to outcomes. Factors or their combinations are typically meaningless in themselves; it is theory that attributes meaning. Indeed, many of the 'usual suspect' variables are proxies that on their own cannot exert causal effects. For example, a high 'infant mortality rate' is a robust predictor of political instability, but arguably only as a proxy for the theoretical concept of 'weak state capacity'.⁴ Since factors also rarely predict with high precision in isolation, theoretical and empirical models that succeed in capturing the contingent and interactive nature of individual factors are likely to do better in this type of evaluation. Likewise, the failure of a single factor to

² Galtung (1967: 372–373), for instance, notes it explicitly and argues for rather calling p -value tests 'tests of generalizability' rather than 'tests of significance'.

³ In the case of forecasting as defined above, readers will have to wait until the forecasting period is over to make such evaluations.

⁴ Thanks to Phil Schrodtt and Nils Weidmann for input to this discussion.

improve prediction does not necessarily mean that it has no place in a social-science model. This, we think, would be a welcome although challenging aspect of the evaluation of theory through prediction.

In addition, many predictions are not directly causal, but instead reflect ‘signals’. Gohdes & Carey (2017) in this issue, for instance, show that killings of journalists are regularly precursors to increased repression. Canaries in a coal-mine can be used as early-warning signals, but the causal relationship between signal and outcome goes through an unobserved third variable: toxic gas leakages in the coal-mine case, and changes in government’s willingness to use extreme measures in the repression case.

Another caveat pertains to the role of reverse causation, when there are theoretical and empirical reasons to believe that there is also a causal effect of Y on X. In this case, X could be a good predictor, but this will be difficult to discern. In our reading, our current understanding of such problems is partial at best, although reduced form models that solve for such endogeneity are a promising way forward.

Forecasting and outcome prediction

A trade-off between *explanation* and *prediction* arises when researchers are simply interested in increasing the predictive power of their models, and not primarily concerned with the understanding of the data-generating mechanisms that is driving, for instance, peace duration, conflict escalation, or war onset. This trade-off applies particularly to the area of machine learning where the combination of computational power and the availability of big data have produced highly flexible and non-parametric methods. While being extremely flexible in the sense that machine-learning algorithms can adapt to non-linear and higher-order relationships, this can come at the price of reduced interpretability. In machine learning approaches, tracing back the most important predictors can be difficult. For instance, methods relying on ensemble techniques are powerful because they average over multiple models (Montgomery, Hollenbach & Ward, 2012). Unless great care is taken in how to specify models and report the results (see Ward & Beger, 2017 for helpful suggestions), the contributions of individual components in an ensemble are difficult to discern. If such methods mis-classify important instances of peace or conflict, it may be difficult to identify parts of the model that improve its predictions.⁵

⁵ Colaresi & Mahmood (2017), Colaresi, Hegre & Nordkvelle (2016), and Ward & Beger (2017) provide some guidance for how

Prediction and forecasting efforts are most useful when they help us understand when our theoretical model hits the mark, when it is (widely) off, and the extent of uncertainty associated with attained insights. Informative predictions ask for explanation, and any explanation worth its weight should predict. Neural networks, random forests, and other non-parametric approaches often exhibit severe limitations when it comes to their ability to generate meaningful policy advice, simply because these tend to obscure what and how to manipulate the real world to avoid undesirable outcomes, including conflict.⁶

With proper attendance to such interpretability issues, machine-learning techniques will obviously play an increasingly important role in the future, especially when it comes to early-warning systems that do not necessitate a full understanding of why an outcome of interest is about to unfold. Reflecting the current main trends in the conflict-forecasting literature, this special issue highlights the role of simpler statistical estimators as recognizable representations of the theoretically deduced data-generating mechanisms. If we have a good albeit partial grasp of the ‘true’ model for how peaceful relations transform into conflictive relations (or vice versa), we may represent this in an estimator and a model specification and make very precise forecasts. If the forecasts are not accurate, we may first consider revising the modeling approach so that it conforms more closely to the theoretical model. Since all parts in the models have a deductive basis, possible improvements to the model are relatively easy to identify (provided the theory is a correct representation of the real world). From this perspective, forecasting is seen simply as a part of the scientific process of building and improving theories.

Forecasting to bridge the gap between basic and applied research

This special issue stresses that forecasting enables conflict researchers to bridge the gap between basic and applied research. Forecasting political instabilities (Goldsmith et al., 2013), regime change, mass killings (Harff & Gurr, 1998), and war (Hegre et al., 2013, 2016) are important preconditions for implementing adequate policy responses, building resilience, and preparing early

to help identify the contribution of individual explanatory factors in ensembles.

⁶ Machine-learning procedures can still be useful for theory development where they identify where predictions from a particular model succeed or fail. See Colaresi & Mahmood (2017) for a discussion.

action. Given resource constraints, when reacting to conflict around the world it is important that policymakers can assess risks, calculate costs and benefits, and condition their responses accordingly. Translating basic political science research into forecasting tools is therefore an important avenue of bridging public policymaking with the academic community.

Ultimately, the goal for the international community should be to prevent armed conflict. In 2015, the UN member states conducted a large-scale review of the tools and approaches used to respond to violence (UN, 2015). The overarching conclusion from that review was that the UN system paid lip service to prevention, but had not really invested anything near close to the necessary knowledge or resources to it. Anticipation is at the heart of efficient prevention of armed conflict. By continuously improving forecasting tools, peace research will be delivering an important public good to the international community.

But for prediction to be useful, it must be embedded in a theory of how the processes that are modeled operate, and how they affect the outcome of interest. For example, Hegre, Nygård & Hultman (2016) forecast how the global incidence of war changes with various UN peacekeeping policies, while Cederman, Gleditsch & Wucherpfennig (2017) examine how various forms of accommodative policies toward ethnic groups contributed to the decline of ethnic civil war after the end of the Cold War.

However, forecasts are clearly not certain statements about the world. Just like weather forecasts, conflict predictions provide some informed guidance about possible scenarios. Forecasts do not tell decisionmakers what they should do, but rather what is likely to happen if they do nothing. Current research is also exploring how best to assess the consequences of possible future public policy interventions (Hegre et al., 2016; Weidmann & Salehyan, 2013; Clayton & Gleditsch, 2014). These approaches are still in their infancy and their ethical implications need to be further considered in a broader debate. They will also become more useful when moving away from simple point predictions. Point predictions are often accompanied by uncertainty estimates, but an alternative is to produce full density forecasts such as Bayesian posterior probabilities. For density forecasting, the goal is rather to forecast the full underlying probability density function of the data-generating mechanism over the outcomes of interest.⁷

A particular concern here is the role of prediction as self-fulfilling or self-containing prophecies. Chadeaux (2017) in this issue discusses how, for instance, forecasts indicating an increased risk of war might prompt countries to attack now, perhaps before a power shift, so that the initial predictions are invalidated. However, he also notes that an improved ability to anticipate conflict is more likely to have the opposite effect. States that underestimate the risk of war may behave more recklessly or demand larger concessions in negotiations than those that have more appropriate estimates, that will take steps to reduce the risk.⁸ As such, war will to some extent always be 'in the error term' (Gartzke, 1999). At the individual level, social competence implies an ability to anticipate the reactions and behaviors of others, including the ability of foreseeing hostile interactions among other members of the individual's social group. Well-functioning social groups continuously use such forecasts to adapt behavior and reduce friction. The set of armed conflict forecasting efforts in this issue is the systematic, large-scale, data-driven analogy to such social skills. If we are able to anticipate violent behavior, we obviously are in a better place to react to it. Clearly, high-quality forecasts of conflict can be misused just as psychopaths misuse their social skills, but that does not in any way invalidate the importance of prediction by itself. Moreover, transparency about methods and techniques as well as public availability of this helps safeguard against misuse.

Policymakers may in particular want to have reliable forecasts of unexpected, high-impact events – the 'black swans' (Taleb, 2007). Forecasting extremely rare events such as world wars using statistical methods is unfeasible.⁹ Judgmental forecasts (Tetlock, 2005) are likely to be more useful in such cases. However, large, unusual wars can be seen as a large cluster of smaller events of more normal types. Statistical approaches that identify typical temporal and spatial escalation patterns may produce warnings about situations that have the potential to become very deadly. When combined with low-level events data, the development of dynamic simulations in Hegre et al. (2013) and Hegre, Nygård & Ræder (2017) is one suggested approach to achieve this.¹⁰ Another promising avenue is provided by the demonstration that the severity of wars and other forms of political violence follows power-law distributions – that is, that the

⁸ See Richardson (1960b: 12) for a similar reasoning.

⁹ See Ward & Beger (2017) for a discussion of the rare-events problem.

¹⁰ The ViEWS project (<http://www.pcr.uu.se/research/views/>) is developing such a model.

⁷ Thanks to Patrick Brandt for input on this.

probability that a war escalates from 1,000 to 10,000 deaths is the same as that of moving from 100,000 to 1 million (Richardson, 1960a; Cederman, 2003; Clauset, Young & Gleditsch, 2007). A better understanding of why wars display this regularity would help anticipate infrequent but extremely deadly quarrels.¹¹ Moreover, a move toward density forecasting where one can focus on the extreme tails of the forecasting distribution will help focus on the most extreme outcomes.

The contributions

The articles in this special issue discuss several of these topics in more detail. Below the contributions are discussed in alphabetical order.

Blair, Blattman & Hartman (2017) show that individual-level survey data from selected locations in Liberia can be harnessed to provide useful forecasts of post-survey risk of violence in the towns the respondents reside in. Moreover, when comparing a number of different model specifications employing a variety of analytical techniques, they find the most parsimonious model to outperform the others. This result is of interest to builders of real-time early-warning systems, since it indicates that they may be able to do well even when monitoring a limited number of indicators.

Cederman, Gleditsch & Wucherpfennig (2017) revisit an explicit forecast articulated by Gurr after the end of the Cold War, namely that thanks to a rise in governments' accommodative politics towards ethnic groups, ethnic civil war would be declining in the years to come. With the benefit of more than a decade of new data and analyses that mimic the postulated causal mechanisms, Cederman et al. find support for Gurr's conjecture about the 'waning of ethnic warfare'. Moreover, they find that this decline appears to have been driven by politics of accommodation and compromise.

Chadefaux (2017) complements the other articles in the issue by analyzing the extent to which financial markets have been able to anticipate the onset of interstate wars. They do not seem to have succeeded historically: yields on government bonds systematically increase after wars start. Moreover, the article shows that a model that uses government bond yields as predictors of interstate wars is poorly calibrated in the sense that it systematically underestimates the risk of conflict onset. A model that uses news sources as predictor for the same set of outcomes is better calibrated.

Chiba & Gleditsch (2017) explore whether dynamic information about mobilization and the behavior of actors from event data can help improve an existing forecast model of civil war that relies on relatively static measures of horizontal inequality (Buhaug, Cederman & Gleditsch, 2014). While their findings suggest some support, the contribution of events data to improving predictive power is somewhat limited.

Colaresi & Mahmood (2017) draw on lessons from machine learning and propose an adapted Box's loop (Blei, 2014) to guide conflict researchers in building better models to explain conflict. The loop consists of four iterative steps: build, compute, critique, and think. The central idea is that researchers should explicitly incorporate and use model criticism to improve their models instead of relying on robustness tests to, essentially, shield their models from criticism. Colaresi & Mahmood (2017) illustrate the utility of this framework by illustrating how it can improve out-of-sample forecasts for armed conflict.

Daxecker & Prins (2017) use forecasting to study the relationship between the availability of lootable resources and armed conflict. They focus on maritime piracy and argue that such piracy is one potent way in which rebel groups can finance rebellion. They use data from Africa and East Asia and show that including dynamic factors measuring piracy improves predictive performance of their models, compared to a baseline excluding these, both in and out of sample.

Gohdes & Carey (2017) examine whether incidents of journalist killings – interpreted as a sign of deteriorating respect for human rights – can help predict subsequent increases in repression. Analyzing a new dataset, their results show that especially in countries with limited repression initially, journalist killings are frequently followed by human rights deterioration.

Hegre, Nygård & Ræder (2017) construct forecasts for the incidence of armed conflict into the future in order to study the conflict trap. The focus is on comparing forecasts under different scenarios rather than on the forecasts themselves. The study shows that an onset of a new armed conflict in a country substantively increases the long-term expected incidence of conflict. Correspondingly successful de-escalation of conflict has large positive, long-term effects. As such, this forecasting exercise can inform decisions regarding how much effort to invest in conflict prevention.

Schneider, Hadar & Bosler (2017) explore the conflict-forecasting potential of economic indicators. They show that tourism sector stocks on the Tel Aviv Stock Exchange perform better as predictors of whether ceasefire agreements in the Levant hold or not than a

¹¹ See Pinker (2011, ch. 5) for a discussion of possible explanations.

careful codification of the assessments of experts in leading newspapers following these conflicts.

Ward & Beger (2017) present a near real-time six-month forecast model of irregular leadership changes for most countries in the world. Their approach relies on ensemble Bayesian model averaging that combines seven different thematic models, each based on a split-population model that disentangles incidence and timing of leadership changes. Overall, the approach yields high out-of-sample accuracy. In addition, Ward & Beger (2017) also reflect upon issues of prediction and forecasting in peace science more generally.

Weidmann & Schutte (2017) demonstrate that night-lights data can be exploited at much finer resolutions than before. While these data enable very good predictions of economic performance at the country level, they improve the prediction of household-level wealth even more. This is an important contribution because it shows how remote-sensing data can be used to predict outcomes of interest in areas where we otherwise have limited data sources. By combining them with spatial data on violence, state reach, health care, and many other issues, it is possible to examine the economic preconditions and consequences of politics at the local level.

Witmer et al. (2017) provide geographically disaggregated forecasts of violent conflict patterns in Africa, projecting for the period 2015–65. Their forecasts integrate climatic and sociopolitical factors, and include various forecasts under alternative scenarios for climate change, political rights, and population growth. Among others, they find that if political rights will continue to improve, then this can neutralize effects of population growth and rising temperatures that would otherwise drive conflict.

Forecasting the future direction of forecasting

The articles in this special issue suggest that forecasting will play an increasingly important role in peace research. They also indicate some particularly fruitful avenues for research. We consider the following important:

Methods/best practices

A benefit of the null-hypothesis significance testing (NHST) research framework is that it has given the scientific community a problematic but shared standard for how to evaluate evidence. For prediction and forecasting, a similar evaluation framework presently does not exist. This special issue takes steps in the direction of agreeing on a set of common standards to evaluate predictions. We do not want, however, to arrive at a new '*p*-value' system for prediction. Instead, the discipline needs to learn to live with

a more flexible system that communicates multiple aspects of a model's performance and reflects ambiguities in its evaluation. Indeed, a recurrent argument in this special issue is that any single-statistic evaluation of results is likely to lead to a suboptimal accumulation of knowledge.

The articles in this special issue point to some clear best practices that should start to lay the foundation for a more stable shared framework for prediction, and, eventually, for establishing empirical evidence for theoretical frameworks. In particular:

1. Researchers need to look at predictions both in-sample and out-of-sample even if their main focus is a traditional empirical analysis and they have no ambition to provide forecasts. Evaluating (out-of-sample) predictions is necessary to harness their power to guard against overfitting, and to ensure that we focus on substantive impact. True out-of-sample forecasts, such as those provided by Ward & Beger (2017) in this special issue, should be the ideal.
2. The various measures used to evaluate predictive performance have different strengths and weaknesses (see Brandt, Freeman & Schrodt, 2014). In this special issue, most contributors report variations of the Brier score, the area under the ROC curve (AUC/AUROC) and precision-recall (P-R) curves. These metrics are flexible and have well-understood strengths and weaknesses. Researchers should not rely on a single metric and further research needs to more thoroughly investigate and propose ways of weighing or combining these under various conditions.¹² Monte Carlo experiments exploring the properties of such metrics would be very welcome.
3. Prediction and forecasts should be used to further the efforts to ensure replicability of scientific results. Transparency at all steps of the chain that produces the prediction is crucial to guarantee that results can be replicated.¹³
4. Predictions need to be presented visually in ways that are meaningful to readers. Colaresi &

¹² For the more restricted case of developing better forecasts we agree with Wang et al. (2016) on the need to develop open and shared test beds against which different forecasts can be tested in a comparable manner. The construction of such test beds through ImageNet has been integral for the tremendous success of the field of machine processing and recognition of images.

¹³ This point is even more relevant given how important peace research and the *Journal of Peace Research* have been in promoting replicability in research.

Mahmood (2017), for instance, develop a novel way of visualizing which cases fit the model, and which are left unexplained. Similarly, contributors such as Chiba & Gleditsch (2017) and Witmer et al. (2017) use maps to efficiently visualize the geographic variation of the predictions as well as predictive power plots. Weidmann & Schutte (2017) use customized figures to illustrate how their predictions differ across estimators.

Theory development

As discussed above, forecasting and out-of-sample evaluation put theories to a different type of test than classical hypothesis testing. Correspondingly, an increased focus on forecasting calls for reformulation of existing theories and raises new questions that, in turn, call for new theory-building efforts. In particular, prediction shifts focus from statistical to substantive significance, and it more easily allows us to gauge the real-world impact of theories. It also allows us to get back to focusing on the performance of a model or a theory instead of focusing on single parameters. This, arguably, can also serve important bridge-building efforts between quantitative and qualitative researchers.

Interpretability

We have highlighted the trade-off between interpretability and forecasting performance and argued for emphasizing the former. It is possible to improve on one front without necessarily sacrificing the other, however, and future forecasting efforts should strive to strengthen the ex-post interpretability of flexible and non-parametric forecasting approaches (e.g. machine-learning based forecasts).

Data

In order to produce timely and truly useful forecasts, we need high-quality data. Particularly promising for the development of ambitious armed conflict early-warning systems are the ACLED (Raleigh et al., 2010) and UCDP-GED (Sundberg & Melander, 2013) data projects that are both very detailed and have frequent and regular update schedules. One ambitious effort to employ these data sources for early warning is the Violence Early Warning System (ViEWS) project in which several of the authors of this special issue are involved.¹⁴ Moreover, to provide consistent forecasts beyond the

immediate future, data collected must adhere to definitions that are constant over both time and space. This requires further efforts into standardization. Past research at the country-year level benefited greatly from the shared standard for what constitutes the unit of analysis developed by Gleditsch & Ward (1999), based on the pioneering work by the Correlates of War Project. Currently, the lack of similar standards for structuring conflict data involving different actors at various spatiotemporal scales impedes scientific progress (Tollefsen, Strand & Buhaug, 2012).

Prediction also brings to the fore the well-known problem of missing data. It is more important than ever to get a good handle on missing-data problems, since it is hard to obtain a forecast for a unit for which we lack crucial data. An innovation to these issues proposed in this special issue is to use forecasting techniques to improve data with estimated values, as Weidmann & Schutte (2017) do when they use light emission to predict economic wealth.

Acknowledgements

Thanks to Henrik Urdal, Patrick Brandt, Philip Schrodt, Nils Petter Gleditsch, and the authors of the special issue for helpful comments. We also want to thank everyone who participated in a forecasting workshop that was held at PRIO. All participants helped to shape and inform this special issue. We also want to thank the reviewers of this special issue for their insightful feedback.

Funding

The work of Hegre and Nygård was funded by the Research Council of Norway, project 217995/V10, and Metternich acknowledges support from the Economic and Social Research Council (ES/L011506/1).

References

- Azar, Edward E (1980) The Conflict and Peace Databank (COPDAB) project. *Journal of Conflict Resolution* 24(1): 143–152.
- Azar, Edward E; RD McLaurin, Thomas Havener, Craig Murphy, Thomas Sloan & Charles H Wagner (1977) A system for forecasting strategic crises: Findings and speculations about conflict in the Middle East. *International Interactions* 3(3): 193–222.
- Beck, Nathaniel; Gary King & Langche Zeng (2000) Improving quantitative studies of international conflict: A conjecture. *American Political Science Review* 94(1): 21–35.

¹⁴ See: <http://www.pcr.uu.se/research/views/>. A pilot for the project is presented in Colaresi, Hegre & Nordkvelle (2016).

- Beger, Andreas; Cassy L Dorff & Michael D Ward (forthcoming) Ensemble forecasting of irregular leadership change. *Research & Politics* 1(3). DOI: 10.1177/2053168014557511.
- Bell, Sam R; David Cingranelli, Amanda Murdie & Aalper Caglayan (2013) Coercion, capacity, and coordination: Predictors of political violence. *Conflict Management and Peace Science* 30(3): 240–262.
- Blair, Robert A; Christopher Blattman & Alexandra Hartman (2017) Predicting local violence: Evidence from a panel survey in Liberia. *Journal of Peace Research* 54(2): 298–312.
- Blei, David M (2014) Build, compute, critique, repeat: Data analysis with latent variable models. *Annual Review of Statistics and Its Applications* 1(1): 203–232.
- Boschee, Elizabeth; Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, James Starz & Michael D Ward (2015) *ICEWS Coded Event Data*. Harvard Dataverse, V10 (<http://dx.doi.org/10.7910/DVN/28075>).
- Brandt, Patrick; Michael Colaresi & John Freeman (2008) The dynamics of reciprocity, accountability and credibility. *Journal of Conflict Resolution* 52(2): 343–374.
- Brandt, Patrick; John Freeman & Philip A Schrodt (2011) Real time, time series forecasting of inter- and intra-state political conflict. *Conflict Management and Peace Science* 28(1): 41–64.
- Brandt, Patrick; John Freeman & Philip A Schrodt (2014) Evaluating forecasts of political conflict dynamics. *International Journal of Forecasting* 30(4): 944–962.
- Bremer, Stuart A (1987) *The Globus Model: Computer Simulation of Worldwide Political and Economic Developments*. Boulder, CO: Westview.
- Bueno de Mesquita, Bruce (1980) An expected utility theory of international conflict. *American Political Science Review* 74(4): 917–931.
- Bueno de Mesquita, Bruce (1983) *The War Trap*. New Haven, CT: Yale University Press.
- Bueno de Mesquita, Bruce (1984) Forecasting policy decisions: An expected utility approach to post-Khomeini Iran. *PS: Political Science & Politics* 17(2): 226–236.
- Bueno de Mesquita, Bruce (2011) A new model for predicting policy choices and preliminary tests. *Conflict Management and Peace Science* 28(1): 65–87.
- Bueno de Mesquita, Bruce; David Newman & Alvin Rabushka (1985) *Forecasting Political Events: The Future of Hong Kong*. New Haven, CT: Yale University Press.
- Buhaug, Halvard; Lars-Erik Cederman & Kristian Skrede Gleditsch (2014) Square pegs in round holes: Inequalities, grievances, and civil war. *International Studies Quarterly* 58(2): 418–431.
- Carment, David (2003) Assessing state failure: Implications for theory and policy. *Third World Quarterly* 24(3): 407–427.
- Cederman, Lars-Erik (2003) Modeling the size of wars: From billiard balls to sandpiles. *American Political Science Review* 97(1): 135–150.
- Cederman, Lars-Erik & Kristian Skrede Gleditsch (2009) Introduction to special issue on 'Disaggregating Civil War'. *Journal of Conflict Resolution* 53(4): 487–495.
- Cederman, Lars-Erik; Kristian Skrede Gleditsch & Julian Wucherpfennig (2017) Predicting the decline of ethnic civil war: Was Gurr right and for the right reasons? *Journal of Peace Research* 54(2): 262–274.
- Chadefaux, Thomas (2014) Early warning signals for war in the news. *Journal of Peace Research* 51(1): 5–18.
- Chadefaux, Thomas (2017) Market anticipations of conflict onsets. *Journal of Peace Research* 54(2): 313–327.
- Chenoweth, Erica & Jay Ulfelder (2017) Can structural conditions explain the onset of nonviolent uprisings? *Journal of Conflict Resolution* 61(2): 298–324.
- Chiba, Daina & Kristian Skrede Gleditsch (2017) The shape of things to come? Expanding the inequality and grievance model for civil war forecasts with event data. *Journal of Peace Research* 54(2): 275–297.
- Choucri, Nazli (1974) Forecasting in International Relations: Problems and prospects. *International Interactions* 1(2): 63–86.
- Clauset, Aaron; Maxwell Young & Kristian Skrede Gleditsch (2007) On the frequency of severe terrorist events. *Journal of Conflict Resolution* 51(1): 58–87.
- Clayton, Govinda & Kristian Skrede Gleditsch (2014) Will we see helping hands? Predicting civil war mediation and likely success. *Conflict Management and Peace Science* 31(3): 265–284.
- Colaresi, Michael & Zuhaib Mahmood (2017) Do the robot: Lessons from machine learning to improve conflict forecasting. *Journal of Peace Research* 54(2): 193–214.
- Colaresi, Michael; Håvard Hegre & Jonas Nordkvelle (2016) Early ViEWS: A prototype disaggregated, open-source Violence Early-Warning System. Presented to the American Political Science Association annual convention, Philadelphia, 1 September (<http://www.pcr.uu.se/research/views/papers-publications/>).
- Daxecker, Ursula & Brandon C Prins (2017) Financing rebellion: Using piracy to explain and predict conflict intensity in Africa and Southeast Asia. *Journal of Peace Research* 54(2): 215–230.
- De Groeve, Tom; Peter Hachemer & Luca Vernaccini (2014) The global conflict risk index (GCRI). A quantitative model. JRC Scientific and Policy Reports. Joint Research Center of the European Commission, Ispra, Italy (<http://publications.jrc.ec.europa.eu/repository/handle/JRC92293>).
- Doyle, Andy; Graham Katz, Kristen Summers, Chris Ackermann, Ilya Zavorin, Zunsik Lim, Sathappan Muthiah, Patrick Butler, Nathan Self, Liang Zhao, Chang-Tien Lu, Paul Khandpur Rupinder, Youssef Fayed & Naren Ramakrishnan (2014) Forecasting significant societal events using the Embers streaming predictive analytics system. *Big Data* 2(4): 185–195.

- Elster, Jon (1989) *Nuts and Bolts for the Social Sciences*. New York: Cambridge University Press.
- Galtung, Johan (1967) *Theory and Methods of Social Research*. Oslo: Universitetsforlaget/London: Allen & Unwin.
- Gartzke, Erik (1999) War is in the error term. *International Organization* 53(3): 597–587.
- Gleditsch, Kristian Skrede & Michael D Ward (1999) A revised list of independent states since the Congress of Vienna. *International Interactions* 25(4): 393–413.
- Gleditsch, Kristian Skrede & Michael D Ward (2013) Forecasting is difficult, especially the future: Using contentious issues to forecast interstate disputes. *Journal of Peace Research* 50(1): 17–31.
- Gohdes, Anita R & Sabine C Carey (2017) Canaries in a coalmine? What the killings of journalists tell us about future repression. *Journal of Peace Research* 54(2): 157–174.
- Goldsmith, Benjamin E; Charles R Butcher, Dimitri Semenovich & Arcot Sowmya (2013) Forecasting the onset of genocide and politicide: Annual out-of-sample forecasts on a global dataset, 1988–2003. *Journal of Peace Research* 50(4): 437–452.
- Goldstone, Jack A; Robert H Bates, David L Epstein, Ted Robert Gurr, Michael B Lustik, Monty G Marshall, Jay Ulfelder & Mark Woodward (2010) A global model for forecasting political instability. *American Journal of Political Science* 54(1): 190–208.
- Gurr, Ted Robert & Mark Irving Lichbach (1986) Forecasting internal conflict: A competitive evaluation of empirical theories. *Comparative Political Studies* 19(1): 3–38.
- Gurr, Ted Robert & Will H Moore (1997) Ethnopolitical rebellion: A cross-sectional analysis of the 1980s with risk assessments for the 1990s. *American Journal of Political Science* 41(4): 1079–1103.
- Harff, Barbara (2003) No lessons learned from the Holocaust? Assessing risks of genocide and political mass murder since 1955. *American Political Science Review* 97(1): 57–74.
- Harff, Barbara & Ted Robert Gurr (1998) Systematic early warning of humanitarian emergencies. *Journal of Peace Research* 35(5): 551–579.
- Hegre, Håvard; Halvard Buhaug, Katherine V Calvin, Jonas Nordkvelle, Stephanie T Waldhoff & Elisabeth Gilmore (2016) Forecasting civil conflict along the shared socioeconomic pathways. *Environmental Research Letters* 11(5): 054002.
- Hegre, Håvard; Joakim Karlsen, Håvard Møkleiv Nygård, Håvard Strand & Henrik Urdal (2013) Predicting armed conflict 2010–2050. *International Studies Quarterly* 55(2): 250–270.
- Hegre, Håvard; Håvard Møkleiv Nygård & Lisa Hultman (2016) Evaluating the effectiveness of UN peacekeeping operations. Typescript, Uppsala University (https://havard.hegre.files.wordpress.com/2013/09/pko_prediction_march2016.pdf).
- Hegre, Håvard; Håvard Møkleiv Nygård & Ranveig Flaten Ræder (2017) Evaluating the scope and intensity of the conflict trap: A dynamic simulation approach. *Journal of Peace Research* 54(2): 243–261.
- King, Gary & Langche Zeng (2001) Improving forecasts of state failure. *World Politics* 53(4): 623–658.
- McClelland, Charles A & Gary D Hoggard (1968) *Conflict Patterns in the Interactions Among Nations*. Los Angeles, CA: University of Southern California.
- Metternich, Nils W & Kristian Skrede Gleditsch (forthcoming) Forecasting in International Relations. In: *Oxford Bibliographies*. Oxford: Oxford University Press. DOI: 10.1093/OBO/9780199743292-0179 (<http://www.oxfordbibliographies.com/view/document/obo-9780199743292/obo-9780199743292-0179.xml>).
- Metternich, Nils W; Cassy Dorff, Max Gallop, Simon Weschle & Michael D Ward (2013) Antigovernment networks in civil conflicts: How network structures affect conflictual behavior. *American Journal of Political Science* 57(4): 892–911.
- Montgomery, Jacob M; Florian M Hollenbach & Michael D Ward (2012) Improving predictions using ensemble Bayesian model averaging. *Political Analysis* 20(3): 271–291.
- O’Brien, Sean P (2010) Crisis early warning and decision support: Contemporary approaches and thoughts on future research. *International Studies Review* 12(1): 87–105.
- Pinker, Steven (2011) *The Better Angels of our Nature: The Decline of Violence in History and Its Causes*. New York: Penguin.
- Raleigh, Clionadh; Håvard Hegre, Joakim Karlsen & Andrew Linke (2010) Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research* 47(5): 651–660.
- Richardson, Lewis Fry (1960a) *Statistics of Deadly Quarrels*. Pittsburgh, PA: Boxwood.
- Richardson, Lewis Fry (1960b) *Arms and Insecurity: A Mathematical Study of the Causes and Origins of War*. Pittsburgh, PA: Boxwood.
- Scharpf, Adam; Gerald Schneider, Anna Nöh & Aaron Clauset (2014) Forecasting the risk of extreme massacres in Syria. *European Review of International Studies* 1(2): 50–68.
- Schneider, Gerald; Nils Petter Gleditsch & Sabine Carey (2011) Forecasting in International Relations: One quest, three approaches. *Conflict Management and Peace Science* 28(1): 5–14.
- Schneider, Gerald; Maya Hadar & Naomi Bosler (2017) The oracle or the crowd? Experts versus the stock market in forecasting ceasefire success in the Levant. *Journal of Peace Research* 54(2): 231–242.
- Schrodt, Philip A (1988) Artificial intelligence and the study of International Politics. *American Sociologist* 19(1): 71–85.
- Schrodt, Philip A (1991) Prediction of interstate conflict outcomes using a neural network. *Social Science Computer Review* 9(3): 359–380.
- Schrodt, Philip A (2014) Seven deadly sins of contemporary quantitative political analysis. *Journal of Peace Research* 51(2): 287–300.
- Schrodt, Philip A & Deborah J Gerner (2000) Cluster-based early warning indicators for political change in the contemporary Levant. *American Political Science Review* 94(4): 803–817.

- Schrodt, Philip A; Shannon G Davis & Judith L Weddle (1994) Political science: KEDS – A program for the machine coding of event data. *Social Science Computer Review* 12(4): 561–587.
- Schrodt, Philip A; James Yonamine & Benjamin E Bagozzi (2013) Data-based computational approaches to forecasting political violence. In: VS Subrahmanian (ed.) *Handbook of Computational Approaches to Counterterrorism*. New York: Springer, 129–162.
- Singer, J David (1973) The peace researcher and foreign policy prediction. *Papers of the Peace Science Society (International)* 21: 1–13.
- Singer, J David & Michael D Wallace (1979) *To Augur Well: Early Warning Indicators in World Politics*. Beverly Hills, CA: Sage.
- Small, Melvin & J David Singer (1982) *Resort to Arms: International and Civil Wars, 1816–1980*. Beverly Hills, CA: Sage.
- Sorokin, Pitirim Aleksandrovič ([1957] 1962) *Social and Cultural Dynamics*. Piscataway, NJ: Transaction.
- Stephens, Jacqueline (2012) Political scientists are lousy forecasters. *New York Times* 23 December (<http://www.nytimes.com/2012/06/24/opinion/sunday/political-scientists-are-lousy-forecasters.html>, accessed 21 December 2016).
- Sundberg, Ralph & Erik Melander (2013) Introducing the UCDP Georeferenced Event Dataset. *Journal of Peace Research* 50(4): 523–532.
- Taleb, Nassim Nicholas (2007) *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.
- Tetlock, Philip (2005) *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton, NJ: Princeton University Press.
- Tollefsen, Andreas Forø; Håvard Strand & Halvard Buhaug (2012) PRIO-GRID: A unified spatial data structure. *Journal of Peace Research* 49(2): 363–374.
- UN (2015) *The Challenge of Sustaining Peace. Report of the Advisory Group of Experts* (<https://www.un.org/en/peacebuilding/pdf/150630%20Report%20of%20the%20AGE%20on%20the%202015%20Peacebuilding%20Review%20FINAL.pdf>).
- Wang, Wei; Ryan Kennedy, David Lazer & Naren Ramakrishnan (2016) Growing pains for global monitoring of societal events. *Science* 353(6307): 1502–1503.
- Ward, Michael D (1984) Differential paths to parity: A study of the contemporary arms race. *American Political Science Review* 78(2): 297–317.
- Ward, Michael D & Andreas Beger (2017) Lessons from near real-time forecasting of irregular leadership changes. *Journal of Peace Research* 54(2): 141–156.
- Ward, Michael D; Brian D Greenhill & Kristin M Bakke (2010) The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research* 47(4): 363–375.
- Ward, Michael D; Nils W Metternich, Cassy L Dorff, Max Gallop, Florian M Hollenbach, Anna Schultz & Simon Weschle (2013) Learning from the past and stepping into the future: Toward a new generation of conflict prediction. *International Studies Review* 15(4): 473–490.
- Weidmann, Nils B & Idean Salehyan (2013) Violence and ethnic segregation: A computational model applied to Baghdad. *International Studies Quarterly* 57(1): 52–64.
- Weidmann, Nils B & Sebastian Schutte (2017) Using night light emissions for the prediction of local wealth. *Journal of Peace Research* 54(2): 125–140.
- Weidmann, Nils B & Michael D Ward (2010) Predicting conflict in space and time. *Journal of Conflict Resolution* 54(6): 883–901.
- Witmer, Frank DW; Andrew M Linke, John O'Loughlin, Andrew Gettelman & Arlene Laing (2017) Subnational violent conflict forecasts for sub-Saharan Africa, 2015–65, using climate-sensitive models. *Journal of Peace Research* 54(2): 175–192.
- Wright, Quincy ([1942] 1965) *A Study of War*, 2nd edition. With a commentary on war since 1942. Chicago, IL: University of Chicago Press.
- Zinnes, Dina A & Robert G Muncaster (1984) The dynamics of hostile activity and the prediction of war. *Journal of Conflict Resolution* 28(2): 187–229.
- HÅVARD HEGRE, b. 1964, Dr Philos in Political Science (University of Oslo, 2004); Dag Hammarskjöld Professor of Peace and Conflict Research, Uppsala University (2013–); Research Professor, Peace Research Institute Oslo (PRIO) (2005–); recent articles in *Environmental Research Letters*, *Journal of Conflict Resolution*, and *Research and Politics*; current research interests: democracy, development, and armed conflict; climate change and conflict; forecasting of armed conflict.
- HÅVARD MOKLEIV NYGÅRD, b. 1983, PhD in Political Science (University of Oslo, 2014); Senior Researcher, Peace Research Institute Oslo (PRIO) (2014–); recent articles in the *American Journal of Political Science*, *World Politics*, and *Journal of Conflict Resolution*; current research interests: the political economy of violence and social order, forecasting of armed conflict, and instability.
- NILS W METTERNICH, b. 1979, PhD in Government (University of Essex, 2011); Senior Lecturer (Associate Professor), Department of Political Science, University College London (2013–); current research interests: forecasting, spatial econometrics, and network analysis; recent publications have appeared in the *American Journal of Political Science*, *Journal of Conflict Resolution*, and *Political Science Research and Methods*.
- JULIAN WUCHERPFENNIG, b. 1980, PhD in Political Science (ETH Zurich, 2011); Assistant Professor, Hertie School of Governance (2016–); current research interests: power-sharing and civil war, conflict processes and political violence; recent publications in the *American Political Science Review* and *American Journal of Political Science*.