Journal of Computational Science 72 (2023) 102074

Contents lists available at ScienceDirect

Journal of Computational Science

journal homepage: www.elsevier.com/locate/jocs

An automated pattern recognition system for conflict

Thomas Chadefaux

Department of Political Science, Trinity College Dublin, College Green, Dublin 2, Ireland

ARTICLE INFO

Keywords:
Conflict
Pattern recognition
Classification
War
Forecasting
Computational diplomacy

ABSTRACT

This article introduces an automated pattern recognition system for conflict. The monitoring system aims to uncover, cluster, and classify temporal patterns of escalation to improve future forecasts and better understand the causes of escalation toward war. It identifies important temporal patterns in conflict data using novel pattern detection methods and new data. These patterns are used to forecast conflict, with live predictions released in real time. Finally, the discovery of recurring motifs—prototypes—can inform new or existing theoretical frameworks. In this article, I discuss the methodological innovations required to achieve these goals and the path to creating an autonomous conflict monitoring system. I also report on promising results obtained using these methods, which show that they perform well on true out-of- sample forecasts of the count of the number of fatalities per month from state-based conflict. The monitoring system has important implications for computational diplomacy, as it can alert diplomats of geopolitical risks.

1. Introduction

There have been more than 350 wars since the start of the 20th century, leading to at least 35 million battle deaths and countless more civiliancasualties [1].[1] Large-scale political violence still kills hundreds every day across the world. International conflicts and civil wars also lead to forced migration, disastrous economic consequences, weakened political systems, and poverty. The recurrence of wars despite their tremendous economic, social, and institutional costs, may suggest that we are doomed to repeat the errors of the past. Does history indeed repeat itself? In other words, are there dangerous temporal patterns of escalation and conflict onset that we should understand to avoid conflict?

This article outlines a monitoring system for patterns of conflict escalation. Its aims are to uncover, cluster, and classify these patterns in meaningful ways to improve future forecasts. In particular, are there recurring pre-conflict patterns? Certain indicators may follow a typical path – a motif – prior to conflict events (inter- or intra-state). Or are the variables associated with conflict largely chaotic and hence inherently unpredictable? Temporal patterns can be searched in the observable actions that international leaders and actors take prior to conflict events, as well as in their perceptions. This can be done at multiple levels of resolution – the minute, the month, the year – and using original data on financial assets, news articles, and diplomatic cables.

Existing work and methods used in social sciences have made substantial progress toward understanding the causes of conflict and even toward forecasting its onset [2–7]. Of particular interest have been international conflicts [8,9], civil wars [10], coups [11,12], and mass killings [13,14]. This growing interest in forecasting in the academic community has been matched by increasing expectations from international organizations, the military, and the intelligence communities, who are working closely with academics to avoid repeating past intelligence failures and misestimations of the costs and risks of war. The availability of increasingly fine-grained spatio-temporal data, in particular, has allowed more refined predictions [15,16]. Data sources range from stock market prices [17,18] to news reports [19], urban violence [20], climate data [21,22], or night-light emissions [23].

However, recent improvements in the availability of fine-grained time series related to conflict now allow us to use techniques which, so far, have been limited to data-rich fields such as finance or speech recognition, where identifying patterns is key. Now is therefore the ideal time to extend existing methodological approaches in political science and related fields (e.g. international relations) to include not only observations treated as independent, but also sequences of observations (a series of data points that occur in successive order over some period of time). This approach has long been championed by qualitative researchers, who often trace process and uncover sequences of events over time. However, qualitative work is often limited in its scope and its ability to uncover patterns in large data. Quantitative approaches, on the other hand, can deal with large quantities of data, but have so far have largely relied on measures of correlation. In short, these

E-mail address: chadefat@tcd.ie.

[1] This number includes all inter-, intra-, extra- and non-state wars with at least 1000 battle deaths each but excludes civilian casualties and battle deaths from smaller conflicts.

https://doi.org/10.1016/j.jocs.2023.102074
Received 5 December 2022; Received in revised form 12 May 2023; Accepted 15 May 2023
Available online 5 June 2023

approaches attempt to match observations on a one-to-one basis – often using complex combinations of variables – but fail to detect patterns which may be stretched, distorted, or generally highly non-linear. Yet the world of conflict is in fact non-linear. Escalation or brinkmanship may follow distinct patterns of warming and cooling in the protagonists' relationship, and this possibly long-term and nonlinear dynamic will not be captured by standard methods. Sequences of events matter perhaps just as much as the value of given covariates.

Second, can we exploit these patterns for prediction purposes? A relevant approach to address this question would be to cluster and classify sequences to understand where tensions are headed—escalation, diffusion, or decline.

Finally, the uncovered patterns can be used for the generation of new theories about conflict processes by identifying the key sequences and combinations thereof that are particularly dangerous. Most work involving complex dynamics remains largely theoretical, in large part because complex, non-linear dynamics involving escalation or diffusion are difficult to study empirically with existing methods. The suggested approach aims not only to inform existing models and theory by learning about the ebbs and flow of international relations, but is also a first step toward generating new theories of conflict processes by identifying the key sequences and their combinations – a grammar of patterns – that are particularly dangerous.

In this article, we introduce a pattern detection system, PaCE (Patterns of Conflict Escalation), which aims to move away from the current tendency in the social sciences to rely on covariance structures of the raw signals (e.g., regression), and to supplement these typical approaches with clustering and prototyping methods to extract shapes and better understand the patterns of escalation into war. This provides a valuable alternative to existing approaches, which are typically unable to treat time series as a whole and therefore often fail to match them with similar patterns on longer-term horizons.

The pattern detection system is also novel in the data it can use. Patterns of geopolitical risk escalation (interstate and intrastate) can be inferred from largely unexplored sources such as: (a) several decades of news articles (using Lexis-Nexis and Factiva); (b) centuries of financial market data (government bond yields data since 1800 from Global Financial Data, and decades of minute-level stock prices from Tick Data Market); or (c) detailed diplomatic records for many European states (e.g. British Documents on the Origin of the First World War; Documents Diplomatiques Français). These long-term and temporally fine-grained data would make it possible to evaluate the pattern of escalation over different time-scales—the century, the year, and the minute. Fine-grained data on the actual conflict events can be obtained from disaggregated conflict data such as ACLED (Armed Conflict Location and Event Dataset—[24]); event data from the Integrated Crisis Early Warning System (ICEWS) Dataverse[2] and Phoenix near-real-time data; and data on the timing of Palestinian rocket launches from Israel's Home Front Command. Understanding these patterns and generating prototypes of escalation structures can improve our ability to forecast future events and help us better understand the pattern of misperceptions surrounding the onset of conflicts.

To our knowledge, no current system directly addresses these questions. Real-time crisis monitoring systems have emerged both in academia and in the policy area, but they rely on methods which (i) do not attempt to measure how fundamentally chaotic (and therefore how predictable) their data is, and (ii) do not treat entire sequences as units of analysis and hence may fail to identify important geometric shape and redundancies in the data.

Below, we detail the state-of-the art of current research and outline the efforts that are required for each of these objectives. We discuss preliminary results, some of the methodological challenges faced, and the approach to follow.

[2] https://dataverse.harvard.edu/dataverse/icews

2. Significance and contribution of the proposed project to the field of conflict forecasting

Advances in computational methods have made it possible to analyze larger and broader sources of information in real time, thereby moving from structural to short-term measures of tensions and other markers of conflict. TABARI (Textual Analysis by Augmented Replacement Instructions), for example, uses the lead sentence of wire service reports (e.g. Reuters, Agence France press, etc.) to generate such event data [25]. The World-Wide Integrated Crisis Early Warning System (ICEWS) is currently the most prominent of these event datasets. Sponsored by the Defense Advanced Research Projects Agency in the United States, it provides a detailed database of political events at the sub-daily and sub-national level [26]. The Violence Early-Warning System (ViEWS) project [27] also directly aims to build a real-time early warning system for war. Here, we propose a warning system, which we call Patterns of Conflict Escalation (PaCE), that goes to go beyond existing work by collecting and analyzing even finer-grained data—up to the minute level using financial data; and including private information using diplomatic cables. This goes beyond existing approaches, which have usually been limited to monthly observations and public data. This added precision allows us to uncover more patterns and at different levels of resolution.

Methodologically, existing approaches in the social sciences typically rely on analyzing the relationship between two sequences by matching each observation in one sequence to another in the other series. They would thus match the current situation with cases that share similar sets of covariate values. However, this approach is likely to miss important relations between the sequences because it does not make use of patterns and geometric regularities. The reliance on one-to-one analysis of each observation (e.g., country-year) is unable to incorporate the more complex dynamics of escalation that typically arise from geopolitical tensions and may take place over different time spans. Most real-world interactions cannot simply be understood as a function of the current state of a variable. Modeling the type of complex back and forth, up and down, aggression–appeasement–further aggression that are common in international relations is difficult. Existing models such as Autoregressive Integrated Moving Average models (ARIMA) may include lags, but are unable to account for this complexity. This may explain why, in existing approaches, the lagged dependent variable is almost always the best predictor, even with long lags and numerous covariates and complex interactions thereof—the immediate past best explains the current state of the system. PaCE addresses this by treating entire sequences of events as units of analysis. This makes it possible to extract motifs that take place at different time resolutions and that are possibly distorted.

To understand the importance of considering sequences as a whole rather than as a set of more or less independent observations, consider a simple illustrative example. Suppose that we observe a pre-conflict process, such as the evolution of the price of a financial asset; a quantified representation of speeches; or a number of terrorist attacks (top plot in Fig. 1; vertical dotted lines denote the timing of the conflict event). We are then interested in finding similar patterns in the future, with the expectation that the similarities help anticipate the likely outcome of that sequence of events. Unfortunately, a logistic regression or a random forest applied to the data points is unable to detect any pattern here. This is regardless of the number of lags, first differences, or splines included [28]. The problem is not the logit itself, but rather that we do not treat sequences as the unit of observation. Instead, decomposing the series into sequences, we realize that the same pattern repeats three times, albeit over different durations (second row). Isolating each one, it then becomes clear that they are all noisy versions of a general motif of a form that could be summarized as 'up 1, down 1/2, up 1'. Correlation would also fail here, because it requires the same number of observations. This is unfortunate, as escalation, power shifts, or terrorism may take place over different time horizon

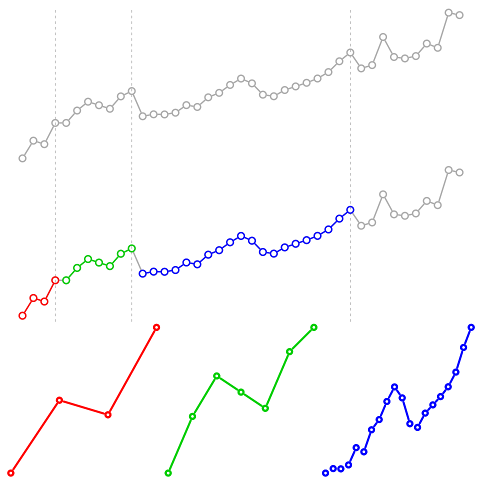


Fig. 1. The red, green and blue sequences are all instances of the same prototype. Yet, standard methods (e.g., logit) do not detect the onset of war (vertical lines).

yet share a general underlying pattern. That pattern would be missed using standard approaches.

The importance of sequences has long been recognized in fields such as bioinformatics and molecular biology, where 'motifs' refer to repeated sequences in a DNA sequence or in amino acids (International Human Genome Sequencing Consortium, 2001). These sequences may also extend at different scales [29], just as we argue that pre-conflict patterns may emerge on different time scales. The idea that sequences might exhibit recurring patterns is not entirely new in social sciences either. Marx's historical materialism is a theory of repeating motifs throughout history [30], and Kondratieff waves describe long-term economic cycles [31]. In economics, the '*J* curve' describes a country's trade balance following a devaluation of its currency. These approaches, however, are ad hoc. The researcher hypothesizes a pattern, then proceeds to look for it in empirical data, but it is likely that many patterns will remain undetected.

Others have, on the contrary, sought to disprove the existence of these patterns. The efficient market hypothesis, for example, shows that any recurrent motif in markets would quickly be exploited by traders and therefore disappear. In conflict studies, Richardson also looked for patterns in the timing of the onset and the duration of wars, yet found that they are governed by a Poisson process—i.e., that they were random (Richardson, 1945). Yet this finding relies on coarse data on the year of onset of interstate conflict. A potentially useful approach to uncover patterns—or the absence thereof, is to (a) gather finer-grained time series related to conflict events, from financial data to news and diplomatic documents; and (b) apply recent methodological developments in machine learning on the clustering of time series to look for hidden patterns. The combination of data orders of magnitude larger than the ones available to, say, Richardson, and recent methods, allows to better address the question whether war does indeed repeat itself or rather is largely stochastic.

This approach is fundamentally different from existing work in political science. It uses recent advances in machine-learning methods and in particular geometry-based pattern recognition methods and Bayesian clustering methods. These methods have the potential to significantly improve conflict forecasts, and therefore to help decision-makers – international organization, non-governmental organizations, and governments – allocate resources where they are needed. Moreover, this approach would radically change the way conflict is studied and supplement existing methods – which largely rely on one-to-one comparison of observations – with ones better able to incorporate the dynamics of conflict. This would allow scholars to better test complex theories, such as formal models, which emphasize the possibly long-term and complex dynamics of interactions rather than the value of certain covariates [32]. This approach would also form a bridge between quantitative and qualitative methods, since the repeating patterns uncovered could match some of the more complex structures typically only uncovered in qualitative approaches. Perhaps more importantly, the findings would also address a long-standing question in philosophy and history: does History repeat itself?

As an illustration of the promise of this approach, I report below the results of a live forecasting competition over the true (unknown) future for a period of six months. I find that the proposed approach significantly improves upon the existing benchmark.

3. Initial results

To illustrate the potential of this approach, I show below the results of a conflict forecasting exercise on the 'true' future—i.e., on data never seen before. A competition was organized with the goal to predict the emergence and evolution of civil wars in Africa [33,34]. The variable of interest was the number of monthly fatalities caused by state-based violence in each region in Africa. State-based violence, or armed conflict, refers to "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year". [35].

The competition took place in October 2020. Learning data was provided for the period 1989–October 2020, and participants were asked to generate forecasts for the following six months, i.e. November 2020 to April 2021. The unit of analysis was the PRIO grid [36]—a unit of space which divides the globe into square cells with a resolution of 0.5 *x* 0.5 decimal degrees (about 50 km at the equator).

The output to predict was the change in the log of the number of conflict-related fatalities in a given cell-month. More precisely, the outcome to be predicted was:

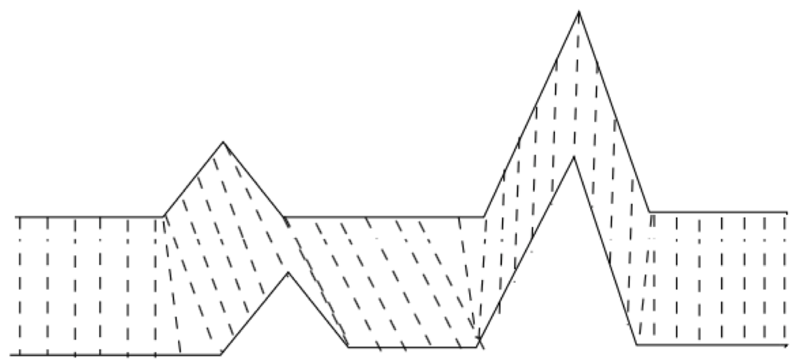
$$\Delta_s ln(Y_{i,t}+1) = ln(Y_{i,t}+1) - ln(Y_{i,t-s}+1),$$

where Y is the number of fatalities recorded by the UCDP and aggregated monthly to the grid-cell level.

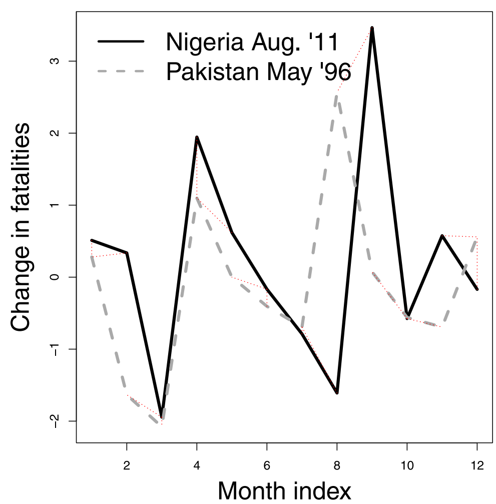
Entries in the competition were evaluated against a benchmark trained on 40 features, including geographic information, population, time since prior episodes of violence, distance to the capital, socio-economic indicators (e.g., literacy rate), oil production, vegetation, agriculture, historical variables (e.g. time since independence), and so on—these features were also available to participants, although the results below did not make use of them. Adding them to the model would improve the results even further.

The models were trained for six steps, i.e., for $Y_{t+2}, \ldots, Y_{t+7}$. The models were trained using features that are respectively $\{2, \ldots, 7\}$ months old. True forecasts for months Oct. 2020–April 2021 were then generated using data up to August 2020 (for more on the benchmark, see [34]). We had available monthly from 1989 to 2020 for each PRIO-grid cell, for a total of 378 months of observations each. Our first step was to split these 378 observations into sequences of one year. The next step was to compare these 12-month sequences to each other. The idea is that sequences that behave similarly may lead to similar outcomes. A distance metric was calculated between each pair of sequences (excluding future observations). For example, the 12-month sequence for Nigeria's casualties is compared to the 12-month sequence for Pakistan from June 1995 to May 1996, as well as to the 12-month sequence for Tunisia from Jan. 1995 to December 1995, and so on.

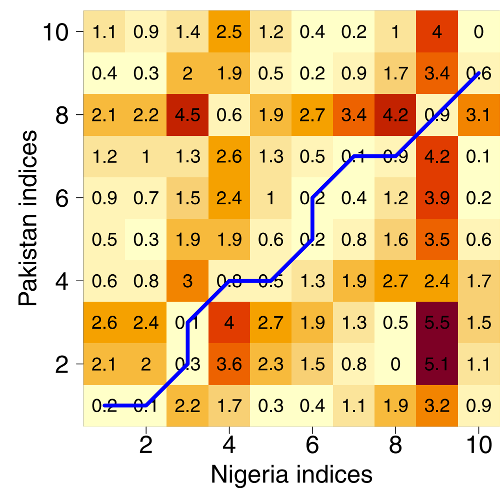
The distance metric we used was based on dynamic time warping (see Fig. 2). For each pair of sequences, a cost matrix is calculated, which represents the distance between each point. Dynamic Time Warping then identifies the path with the smallest total cost (Fig. 2(c)). This total cost is used as a measure of distance between each pair of sequences.



(a)



(b)



(c)

Fig. 2. Dynamic time warping. (a) Illustration of the DTW method. Similar shapes are not aligned in the time axis. As a result, Euclidean distance leads to a large distance, because it assumes a one-to-one alignment in the two sequences. Instead, a nonlinear dynamic time warped alignment (i.e., a Euclidean distance based on the warped series) reveals the proximity between the two sequences; (b) Point-wise comparison of sequences $s_{\text{Nigeria},380}$ and $s_{\text{Pakistan},197}$. Alignments are displayed in red; (c) Cost matrix and warping curve (blue). Each cell shows the Euclidean distance between observation X_i and X_j. The warping path is the one with the smallest cumulative distance.

Our forecasts were made by a weighted average of the closest matches to sequences of 12 observations (a year) in past data. For example, the best match for Sudan's East Darfur state in the 12 months of 2012 is Kenya's Rift Valley Province in 1994, which is therefore given a large weight.[3] The predicted change in fatalities for sequence $s_{i,t}$, denoted by $\hat{f}_{i,t}$, is simply calculated as the average of the futures of past sequences, weighted by their similarity (i.e., the inverse of their distance):

$$\hat{f}_{i,t} = \frac{1}{N} \sum_{j \neq i, t_j < t_i} \left(\frac{1}{d_{ij}} f_{j,t_j} \right), \tag{1}$$

where $d_{i,j}$ is the distance between series s_i and s_j

Intuitively, $\hat{f}_{i,t}$—the estimated future of sequence i—is obtained by comparing sequence $s_{i,t}$ to past sequences $s_{j,t_j<t_i}$. The future of $s_{i,t}$ is then forecasted as a weighted average of the future of those (past) comparison sequences. Past sequences $s_{j,\cdot}$ that are similar (i.e., those with a small d^w_{ij}) are expected to lead to similar outcomes and hence are assigned a large weight, whereas dissimilar sequences (large d^w_{ij}) receive a low weight (i.e., $1/d^w_{ij}$ is small).

We repeat this procedure for every prio-grid and predict two to seven months ahead for each. We find that our results significantly improve upon existing approaches. Using a standard regression (OLS) or a random forest (RF) with suitable covariates yields good results. However, the addition of information about patterns using dynamic time warping significantly reduced the Mean Absolute Error (MAE = $\sum_i (\hat{y}_{i,t+s} - \hat{y}_{i,t+s})$) not only for the past data, but also for the prediction period, which at the time were entirely in the future. The results (Fig. 3) show that our analysis ('Patterns') significantly improves upon both the benchmark model and a baseline model that predicts the future on the basis of the past ('Lag' model defined as $Y_{t+s} = Y_t$).[4]

These results are encouraging. They show the importance of accounting not only for the raw value of covariates, but also for their patterns over time. Existing approaches, from regression to random forests and neural networks, have a hard time dealing with sequences that may vary in speed and shape. Accounting for the geometry of a sequence or its shape allows us to capture additional information that is not available to other methods. Furthermore, the performance obtained here is based on the shape of the sequences only. Incorporating

[3] Chadefaux [37] describes the methodology in greater detail. We focus here on the out-of-sample results that were not available in [37].

[4] The difference between the lag model and our model is not significant for predictions two and three months ahead. This is mainly due to the small number of these predictions.

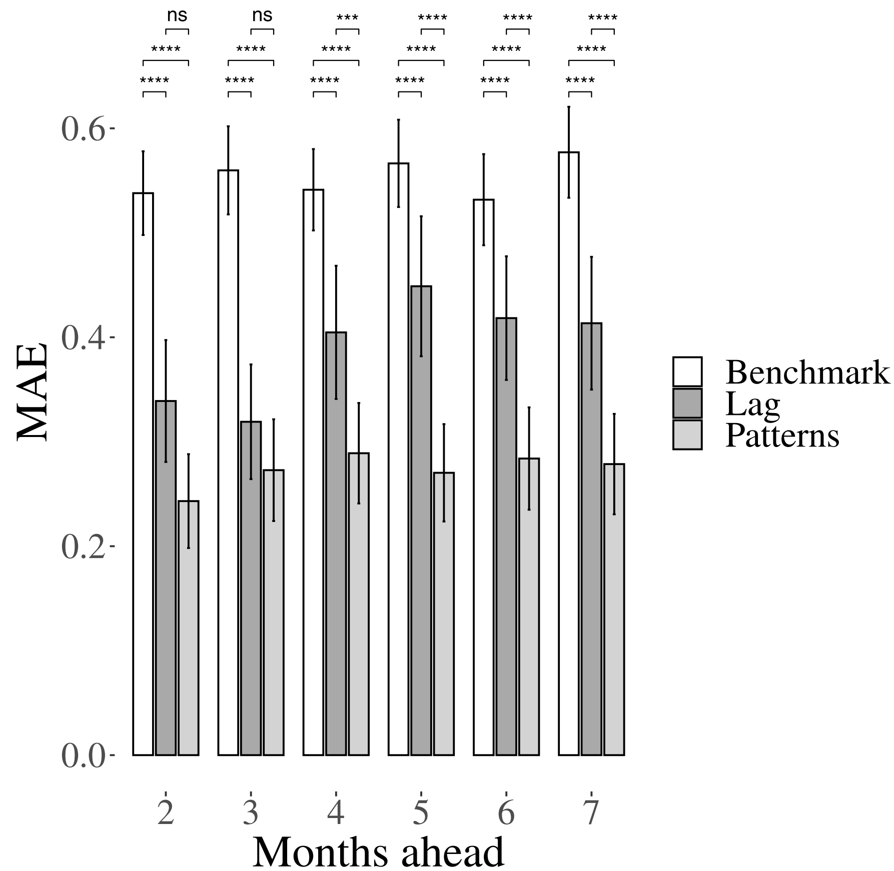


Fig. 3. Mean Absolute Errors for state-based violence forecasts, October 2020–April 2021.

these predictions into an ensemble model of other predictions using additional covariates may improve results further [34].

4. Developing an automated conflict warning system

These initial results pave the way for developing a broader and more ambitious warning system (see Fig. 4). We aim to infer patterns of geopolitical risk escalation (interstate and intrastate) from three largely unexplored sources over two centuries: (a) several decades of news articles (using Lexis-Nexis and Factiva); (b) centuries of financial market data (government bond yields data since 1800 from Global Financial Data, and decades of minute-level stock prices from Tick Data Market); and (c) detailed diplomatic records for many European states (e.g. British Documents on the Origin of the First World War; Documents Diplomatiques Français). These long-term and temporally fine-grained data allow us to evaluate the pattern of escalation over different time-scales—the century, the year, and the minute. Fine-grained data on the actual conflict events can be obtained from disaggregated conflict data from ACLED (Armed Conflict Location and Event Dataset—[24]); event data from the Integrated Crisis Early Warning System (ICEWS) Dataverse[5] and Phoenix near-real-time data; and data on the timing of Palestinian rocket launches from Israel's Home Front Command.

4.1. Data

For this project, we need time series long enough to extract patterns. This means either fine-grained data (i.e., many data points over a possibly short period) or data of long duration (years or decades) when the data is coarse. Of particular interest is detailed information about conflict events and their unfolding over time, as well as time series that may serve as early indicators for these events. All of these series are complemented by the relevant covariates identified in the literature (e.g., neighboring conflicts, GDP, ethnic polarization, etc.).

While it is not possible to gather all possible conflict related data or indicators of conflict, our data collection strategy is to combine data that varies along two main dimensions: the availability of the data to the public (publicly or privately observed indicators); and its objectivity (factual/perceived indicators). We also note the importance of geographic variables to understand and predict conflict events. Forecasts should also incorporate relevant geographical covariates (e.g., diffusion effects) into the ensemble model.

	Factual	*Perceived*
Public info.	Conflict events (CoW, ACLED, UCDP, ICEWS) Palestinian rocket launches GDP, trade, foreign aid, etc	News reports Financial assets Exchange rates
Private info.	Diplomatic cables (event analysis) Military spending	Diplomatic cables (sentiment analysis)

Conflict data. Data on conflict events include state-based armed conflict, non-state conflict, and one-sided violence (with variation depending on the availability of covariates and the subject). The focus is on armed conflict involving consciously conducted and planned political campaigns rather than spontaneous violence. The main unit of analysis are be events where armed force was used by an organized actor against another organized actor, or against civilians, resulting in at least one direct death. This includes the prediction of individual incidents, as

[5] https://dataverse.harvard.edu/dataverse/icews

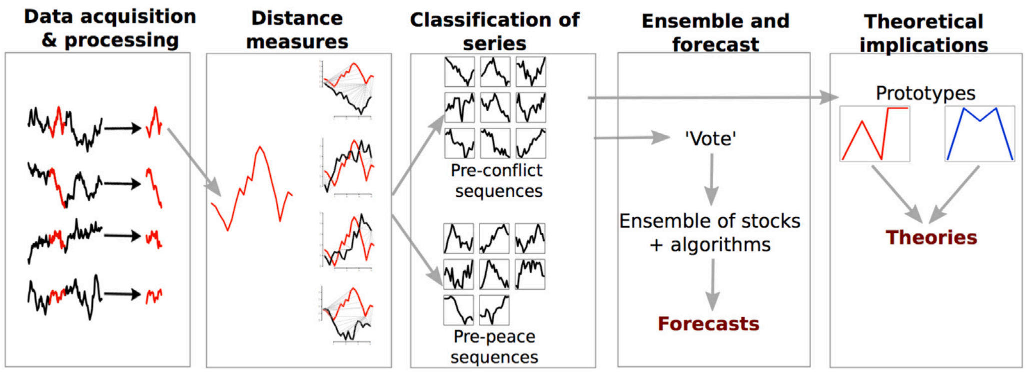


Fig. 4. Outline of PaCE's workflow.

well as of the general onset of conflicts, both at the monadic and the dyadic levels. Data with worldwide coverage (typically with a time resolution of a day) include: the Correlates of War (CoW) for historical data on interstate conflict; the Armed Conflict Location & Event Data Project (ACLED), which collects real-time and historical data on political violence and protest events; and mainly the Uppsala Conflict Data Program (UCDP/PRIO Armed Conflict Dataset and GED), the main provider of data on organized violence. For civil wars, we use UCDP's data when possible, and ACLED otherwise. Two other very detailed records of political and conflict events can also be used for a larger coverage of events related to conflict: DARPA's Integrated Crisis Early Warning System (ICEWS) and the Phoenix dataset from the University of Texas, Dallas. We also make use of data with more local coverage but greater temporal precision. These data are particularly useful to examine the patterns of conflict at the intra-day level (e.g., minute- or even second-level resolution using tick market data). For example, Israel issues rocket alerts to citizens who live within the range of these attacks. The alerts have a precision of one minute and are available back to 2010.

Early indicators of war. Conflict events are typically the end product of a possibly long gestation process. Prior to acting, actors often ponder their options, deliberate, or engage in public debate. These processes are not reflected in conflict data, but we would still like to observe the evolution of these processes—i.e., not only the discrete points at which actions are taken, but also the actors' continuous assessment of risk and how this assessment is evolving over time. In other words, we also want to observe participants' perceptions of conflict, as they provide finer-grained and possibly more truthful representations of the evolution of risk. In particular, we attempt to classify three main types of time series into pre-war or peaceful: (i) financial market data; (ii) news articles; and (iii) diplomatic documents.

Financial Assets. Financial markets are the ideal time series for a monitoring system because they combine the forecasts of actors who have a financial stake in making accurate predictions [38–40]. Securities are traded in a way that reflects the investors' beliefs about the probability of a certain event occurring. Large events such as wars are economically and financially costly, and market participants will therefore strive to anticipate them as early as possible and to react accordingly. For example, bonds are likely to be sold in anticipation of a war, as will be the stocks of industries most likely to be affected by it [18]. As a result, financial assets often respond strongly to the expected occurrence of violence. We use three main types of financial data: (i) Government bond yields are good proxies because they depend on the perceived sovereign risk. Wars generate two main kinds of sovereign risks for investors: full or partial default, and inflation. These risks imply that a bondholder who expects war should demand a higher yield today [18]. (ii) Stock prices of industries affected by war are also likely to react in advance of conflict events. They also have the advantage of being available at the minute- or even the tick-level. (iii) Finally, exchange rates can also be used as they are also sensitive to geopolitical tensions and risks and are more readily available for countries without a developed financial market.

News Reports. We also rely on the analysis of newspaper articles. News has the advantage of directly reporting observers' perceptions of events, and hence can be more specific than financial data. One challenge, of course, is that news reports are less easily quantifiable than asset prices. The text of each article is processed according to a standard procedure in text mining approaches: first, remove common words such as 'and' or 'that'; second, lemmatize and stem the words (i.e., attacking, attacked, attacker all become 'attack'). We then apply standard methods such as the Latent Dirichlet Allocation (LDA) to model topics [41,42] and more advanced dynamic topic models to estimate changes in the salience of themes over time [43,44]. The distribution of these topics over time forms the basis of our time series.

Diplomatic Documents. Our final set of time series comes from historical diplomatic documents. A large and comprehensive corpus of telegrams, letters and reports is available for many European countries since the 1870s. France and Great Britain, in particular, but also Germany, Austria, Italy, and others have made a conscious effort to publish as comprehensively as possible their diplomatic archives. The diplomatic cables covered include various messages from diplomats and politicians to their Capital. This includes information on both events (facts) as well as on the diplomat's opinions on them (perceptions). The telegrams are often classified as confidential and typically show the candid opinion of the sender—for example often commenting on their counterparts' trustworthiness; the impression they had from their meetings; and direct thoughts about strategy. These cables span many years—a hundred for France and Great Britain (1870s–1970s), and include a vast number of documents—typically hundreds each month—which would allow us to draw robust inferences.

Most volumes are available online in machine-readable form. These unstructured text files are then be passed through a series of pre-processing steps prior to data analysis. First, we exploit formatting patterns within the text to automatically parse the content of the diplomatic cable entries from the volumes. Next, we extract the date of each document found within the content, using regular expressions when possible, and crowd-coding by online workers via CrowdFlower otherwise.

Finally, we process the text to reduce the dimensionality of the data while preserving meaningful and predictive terms. This gives us time series – one for each extracted topic (as was the case for news reports) – that can then be passed on to our classifier for shape extraction.

Early indicators of war. Finally, we collect standard panel data on indicators such as GDP and its growth, military spending, democracy scores, military spending, trade, foreign aid, primary commodity exports as a percentage of GDP, or income inequality. The underlying idea is to uncover potential patterns in the evolution of these variables, and to use them as predictors of conflict.

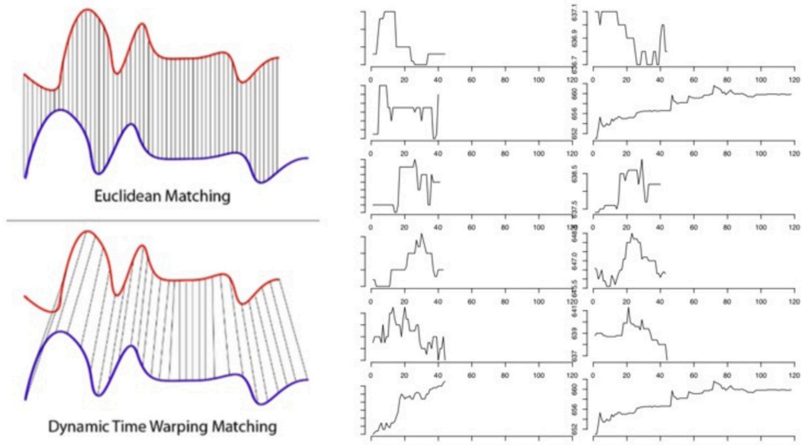


Fig. 5. Dynamic time warping (left) and illustrative matches (right).

4.2. Pattern classification and clustering

We aim to (i) identify the patterns that are associated with periods preceding conflict; and (ii) cluster and classify them. We apply these procedures to two main types of series: events, and perceptions.

- Patterns of Events. Are there patterns in the observable actions that actors take prior to conflict? Here we use vast amounts of data on conflict timing—mostly at the daily level (e.g., ACLED, UCDP, ICEWS), but in some cases as precise as the minute-level (see section on data above). With these data, we ask (a) whether, irrespective of model or perception, there actually exist patterns. This is done using model-free measures of predictability such as entropy measures (e.g., permutation entropy). And (b) if patterns exist, can they be clustered and classified into meaningful sequences and categories that help understand where tensions are headed—escalation, diffusion, or decline?
- Patterns of Perception. Are there identifiable patterns in observers' perceptions of geopolitical risk? In particular, are there systematic biases and specific shapes to the evolution of these perceptions? History is replete with examples of miscalculations, from World War I to the 2003 US invasion of Iraq. But are these stories representative of a larger pattern? Unfortunately, little is known about how well wars are anticipated. Do observers make recurring mistakes and are there patterns in these misperceptions? We examine these questions by looking for patterns in data on perceptions derived from: financial markets (the crowd's perception); news articles (the "expert's" perception); and diplomatic documents (the policy-maker's perception). In Fig. 3, for example, we consider the evolution of Israeli stocks and try to cluster them into pre-attack/not-pre-attack according to their shape signature (original on the left; best match on the right).

4.2.1. Distance measures

The first step before clustering and classifying is to calculate the distance between each sequence. This involves splitting each time series into small sequences, and searching for similar sequences in past series (both in the same class of data—e.g., finance—or in other types as well—e.g., news). To measure similarity, PaCE relies on two main types of methods:

First, we rely on model-free approaches. These can use the longest common subsequence (applied to symbolic representations of the time series) or landmark similarities (e.g., local minima/maxima, inflection points). One of the main approaches here is Dynamic Time Warping (Berndt & Clifford, 1994), an algorithm which measures the similarity between two temporal sequences that may vary in speed—which is particularly appropriate for the problem of escalation and the ups and downs of conflict, as these may take place over shorter or longer time periods but still present the same pattern (Fig. 5a). This type of problem cannot be easily addressed by standard, moment-based approaches to time series. As an example, Fig. 5b shows how past financial prices (here., those preceding violent attacks on the left) can be matched with current financial price patterns (on the right). Based on the quality of these matches, today's series can be classified as pre-attack or not. We note that given the large amount of data involved, particularly with financial prices, computational challenges require novel approaches to the problem. Here we rely in particular on [45]'s approach, which among others uses various techniques of early abandoning whereby unsuitable matches are abandoned early in the search. Moreover, weighted DTW is be used to avoid giving undue weight to outliers, which can occur in the field of conflict research because of imperfect data [46].

Second, we aim to reduce the dimensionality of the series. The idea here is that time series can be decomposed into their components—for example, a long-term trend plus seasonal variation plus a cyclical component and finally an irregular component (the residual). These transformations of the time series can help to search for patterns. For example, the Discrete Fourier Transform [47,48], the Discrete Wavelet Transform [49,50], Adaptive Piecewise Constant Approximation [51,52], Singular Value Decomposition [53,54], Piecewise Aggregate Approximation [51], or piecewise linear approximation [55]. Finally, discretization allows us to use the large number of existing algorithms for the efficient manipulation of symbolic representations. In particular, Symbolic Aggregate Approximation [56] allows the discretization of original time series into symbolic strings.

4.2.2. Clustering and feature extraction

The second step is to extract the defining features that characterize pre-conflict situations. Our approach is to cluster the time series according to both their shape and complexity. In particular, we can reduce the time series to a meaningful description in a low-dimensional feature space by means of geometric approximations, or by taking an expansion over a basis of functions such as splines or wavelets. This is in direct contrast to existing approaches in the social sciences which, directly or indirectly (typically through some regression model) use clustering based on the Euclidean space spanned by the raw signals of the series. These existing approaches are not sufficient for the kind of nonlinear signals that are common in conflict data. All pairwise distances obtained enter a matrix, which is then fed to a clustering algorithm. This clustering procedure results in an assignment of each pattern to a group of similar patterns (i.e., those with a small distance). These clusters can then be used both for forecasting and theory-building purposes. Computational power is a challenge. Even for small datasets, calculating all pairwise distances between subsequences does not scale well.

For that reason, we use a number of speed-up techniques, including indexing [57], lower-bounding [58], pruning [59], early abandoning and the matrix profile [60], in addition to making use of Trinity College's high performance computing cluster.

4.3. Predicting conflict

It is now well known that statistically significant variables may fail to improve a model's prediction out of the estimation sample [10], typically because they overfit the in-sample data. Out-of-sample forecasting avoids this trap by measuring the quality of forecasts, rather than simple in-sample significance. Forecasting can also be an important tool for theory building [61]. Our main tool for external validation is therefore be out-of-sample forecasting. This involves both backwards (historical) testing and true out-of-sample testing using future data over the lifetime of the project.

Backward Forecasting. Backward Forecasting is performed on the onset of inter- and intra-state conflict events for the past. We first define a 'learning' window (e.g., 1800–1850) within which we extract the relevant shapes and measures of distance. Using this learning set, we infer which patterns are particularly dangerous, and estimate the risk of war in the following time period (e.g., 1851). This is done in two steps: first, the distance between the sequences in the testing set and all sequences in the learning set are calculated. For example, this may mean that the weekly price pattern of French bonds in 1851 is compared to the price pattern of German bonds in 1850, 1849, 1848. This is done using each of the distance measures described above.

The result is a distance metric for each possible pair of sequences and each algorithm—an enormous amount of information. Different data sources and different algorithms yield different estimates. Each of these estimates contains new information, and our second step is to combine each distance measure into an ensemble forecast of all distances. For example, we might have learned from the learning set that the combination of the distance between a certain stock price and another price yields particularly good predictions. We therefore rely heavily on that estimate, and less on other. The ensemble allows us to aggregate these various measures and make the best use of each of their contributions.

Live Forecasting. 'Live' forecasting differs from backward forecasting in that it predicts a future that is truly unknown, and not "as if" unknown. For this project, for example, assuming a start year of 2020, we would make forecasts for $2021, 2022, \dots, 2026$. Such true forecasts have the benefit of being free from biases originating from data selection, statistical biases, or selective releases which affect ex-post tests. Second, it would provide policy-makers with strong evidence that our results are useful. This is particularly relevant today, as various governments and international organizations (e.g., the UN, the World Bank, ECOWAS, the European Commission) have implemented their own early warning systems (using existing methods).

5. Discussion: Theoretical implications of the extracted features

Extracting and clustering patterns allows us to make forecasts and improve on existing work. However, we also strive to interpret the patterns we extract into categories that are meaningful from a theoretical point of view. In particular, the uncovered patterns – e.g. 'three ups-one down-two ups' – may be fruitfully related to (i) existing theoretical approaches to escalation and the causes of war and (ii) to new theoretical understandings thereof. Scholars of international relations have long understood that sequences of events matter and are not simply the sum of their individual components. Formal models of conflict, in particular, clearly specify an order in which events take place and what the outcome of these steps is (e.g., [62]). They also understand that the time between each of these steps is flexible. t and t+1 may refer to seconds, months or decades, but the underlying structure is the same. However, empirical analysis has thus far been unable to directly address these two facets of strategy. Instead, empirical models treat events more or less in isolation (with at best a few lags), thereby ignoring sequences; and they impose strict units of time on actions—typically the year. PaCE would address both issues. By directly modeling the sequence as the unit of analysis, rather than the individual observation, PaCE is able to match entire sequences of time; and by using flexible methods such as dynamic time warping, we can match sequences by stretching and scaling them over different time scales.

A risk of this approach is to extract shapes that are in fact noise. By sheer chance, we do expect recurring shapes to emerge from time series which, in fact, contain no information. An important step is therefore to compare the rate of occurrence of a particular sequence to the expected rate in the absence of signal. To that end, one can first generate synthetic time series with parameters comparable to the ones of our empirical series (e.g., ones with the same autocorrelation coefficients) and obtain a distribution of the expected number of sequences in a particular time series. This in turn allows us to obtain a p-value, as it were, of each temporal sequence, by comparing it to that theoretically expected distribution.

These shapes can then be used for theory building. Instead of relying on ad hoc theories about the sequence of events and testing them individually, PaCE aims to build a repository of shapes – a grammar – which can be used as building blocks for new theories. We expect some of these patterns to match existing models, such as bargaining games of conflict, models of arms competition, deterrence, diplomacy and signaling, power change and war. We also expect unknown patterns to emerge and can use them to build new theories.

Overall, the extraction of temporal can help us bridge the gap between qualitative and quantitative research. Qualitative work can uncover complex historical processes and is able to make connections between analogous situations that would be nearly impossible for algorithms to find in the absence of immense amounts of data—data which is not available to most social sciences. However, qualitative research is limited in scope by its human cost and the practical impossibility of scaling up the analyses. Quantitative research, on the other hand, easily scales up, but is typically unable to make flexible connections between cases. The approach proposed here – a quantitative approach but with a flexible understanding of time and clustering of similar pattern – can contribute to bridge that gap and to uncover more meaningful patterns at lower cost.

The findings are important for computational diplomacy. Identifying dangerous sequences can be critical to the ability of diplomats to recognize risky situations for what they are, and to engage in negotiations with a better understanding of the underlying risks. Moreover, the ability to identify dangerous sequences would mean that diplomats no longer solely rely on experts and their selection of relevant past comparison—'a Munich moment', or 'a risk of entrapment similar to 1914'—but rather on a more exhaustive list of all relevant comparison, and their degree of similarity. Rather than relying solely on a possibly spotty identification of relevant past event, the system would be able to be queried for specific patterns and their occurrences in the past, including in contexts that diplomats may not necessarily think about. As such, diplomacy can be substantially enhanced by this computational tool.

While the methods and tools we have discussed hold significant potential for enhancing our understanding of conflict and improving diplomacy, they are not without potential risks. These methods, if placed in the wrong hands, could be used for detrimental purposes. For instance, actors with malicious intent could manipulate this information to instigate conflicts, exacerbating tensions rather than diffusing them. They might identify patterns and exploit them to their advantage, creating unrest or destabilizing regions. Moreover, the predictive power of these tools could be used to anticipate and circumvent diplomatic efforts, or even to craft strategies that take advantage of predicted international responses. This could involve the orchestration of actions that follow or deliberately break identified patterns, thereby misleading observers or controlling the narrative.

It is therefore crucial that these tools are used responsibly, with a clear understanding of their potential implications. This involves not only the responsible use and interpretation of the generated data but also the ethical considerations surrounding its collection and dissemination. The development and application of appropriate safeguards, regulations, and standards for usage are vital. This could include protocols for data access, guidelines for interpretation, and measures to prevent misuse. In addition, ongoing dialogue and cooperation among researchers, policy makers, and practitioners are necessary to ensure the benefits of these tools are maximized, while their risks are mitigated. The potential for misuse highlights the need for transparency, ethical considerations, and robust debate in the development and application of these innovative methods in the field of conflict analysis and diplomacy.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101002240).

References

[1] M.R. Sarkees, F. Wayman, Resort to War: 1816-2007, Cq Press, 2010.
[2] L.-E. Cederman, N. Weidmann, Predicting armed conflict: Time to adjust our expectations? Science 355 (6324) (2017) 474–476.
[3] T. Chadefaux, Conflict forecasting and its limits, Data Sci. 1 (1) (2017) 1–11.
[4] H. Hegre, N. Metternich, H. Nygürd, J. Wucherpfennig, Introduction: Forecasting in peace research, J. Peace Res. 113–124 (2017).
[5] B. Mesquita, Predicting Politics, Ohio State University Press, 2002.
[6] B. Mesquita, The Predictioneer's Game: Using the Logic of Brazen Self-Interest to See and Shape the Future, Random House, New York, 2009.
[7] G. Schneider, N. Gleditsch, S. Carey, Forecasting in international relations: One quest, three approaches, Confl. Manag. Peace Sci. 28 (1) (2011) 5–14.
[8] N. Beck, G. King, L. Zeng, Theory and evidence in international conflict: A response to de Marchi, Gelpi, and Grynaviski, Am. Polit. Sci. Rev. 98 (2) (2004) 379–389.
[9] K. Gleditsch, M. Ward, Forecasting is difficult, especially about the future: Using contentious issues to forecast interstate disputes, J. Peace Res. 50 (1) (2013) 17–31.
[10] M. Ward, B. Greenhill, K. Bakke, The perils of policy by P-value: Predicting civil conflicts, J. Peace Res. 47 (4) (2010) 363–375.
[11] J.A. Goldstone, R.H. Bates, D.L. Epstein, T.R. Gurr, M.B. Lustik, M.G. Marshall, J. Ulfelder, M. Woodward, A global model for forecasting political instability, Am. J. Polit. Sci. 54 (1) (2010) 190–208.
[12] M. Ward, A. Beger, Lessons from near real-time forecasting of irregular leadership changes, J. Peace Res. 54 (2) (2017) 141–156.
[13] N. Rost, Will it happen again? On the possibility of forecasting the risk of genocide, J. Genocide Res. 15 (1) (2013) 41–67.
[14] J. Ulfelder, Forecasting onsets of mass killing, 2012.
[15] P. Brandt, J. Freeman, P. Schrodt, Evaluating forecasts of political conflict dynamics, Int. J. Forecast. 30 (4) (2014) 944–962.
[16] N. Weidmann, M. Ward, Predicting conflict in space and time, J. Confl. Resolut. 54 (6) (2010) 883–901.
[17] G. Schneider, M. Hadar, N. Bosler, The oracle or the crowd? Experts versus the stock market in forecasting ceasefire success in the Levant, J. Peace Res. 54 (2) (2017) 231–242.
[18] T. Chadefaux, Market anticipations of conflict onsets, J. Peace Res. 54 (2) (2017) 313–327.
[19] T. Chadefaux, Early warning signals for war in the news, J. Peace Res. 51 (1) (2014) 5–18.
[20] R. Bhavnani, K. Donnay, D. Miodownik, M. Mor, D. Helbing, Group segregation and urban violence, Am. J. Political Sci. 58 (1) (2014) 226–245.
[21] H. Hegre, H. Buhaug, K. Calvin, J. Nordkvelle, S. Waldhoff, E. Gilmore, Forecasting civil conflict along the shared socioeconomic pathways, Environ. Res. Lett. 11 (5) (2016).
[22] F. Witmer, A. Linke, J. O'Loughlin, A. Gettelman, A. Laing, Sub-national violent conflict forecasts for sub-Saharan Africa, 2015–2065, using climate-sensitive models, J. Peace Res. 54 (2) (2017) 175–192.
[23] N. Weidmann, S. Schutte, Using night light emissions for the prediction of local wealth, J. Peace Res. 125–140 (2016).
[24] C. Raleigh, A. Linke, H. Hegre, J. Karlsen, Introducing ACLED: an armed conflict location and event dataset: special data feature, J. Peace Res. 47 (5) (2010) 651–660.
[25] P. Schrodt, TABARI: Textual analysis by augmented replacement instructions, 2009.
[26] S. O'Brien, Crisis early warning and decision support: Contemporary approaches and thoughts on future research, Int. Stud. Rev. 12 (1) (2010) 87–104.
[27] H. Hegre, M. Allansson, M. Basedau, M. Colaresi, M. Croicu, H. Fjelde, F. Hoyles, L. Hultman, S. Högbladh, R. Jansen, et al., ViEWS: A political violence early-warning system, J. Peace Res. 56 (2) (2019) 155–174.
[28] N. Beck, J.N. Katz, R. Tucker, Taking time seriously: Time-series-cross-section analysis with a binary dependent variable, Am. J. Political Sci. 42 (4) (1998) 1260–1288.
[29] G. Bernardi, B. Olofsson, J. Filipski, M. Zerial, J. Salinas, G. Cuny, M. Meunier-Rotival, F. Rodier, The mosaic genome of warm-blooded vertebrates, Science 228 (4702) (1985) 953–958.
[30] K. Marx, A contribution to the critique of political economy, in: Marx Today, Springer, 2010, pp. 91–94.
[31] N.D. Kondratieff, The long waves in economic life, Review (Fernand Braudel Center) (1979) 519–562.
[32] T. Chadefaux, What the enemy knows: Common knowledge and the rationality of war, Br. J. Polit. Sci. 49 (2018) 1–15.
[33] P. Vesco, H. Hegre, M. Colaresi, R.B. Jansen, A. Lo, G. Reisch, N.B. Weidmann, United they stand: Findings from an escalation prediction competition, Int. Interact. (2022) 1–37.
[34] H. Hegre, P. Vesco, M. Colaresi, Lessons from an escalation prediction competition, Int. Interact. 48 (4) (2022) 521–554.
[35] N.P. Gleditsch, P. Wallensteen, M. Eriksson, M. Sollenberg, H. Strand, Armed conflict 1946–2001: A new dataset, J. Peace Res. 39 (5) (2002) 615–637.
[36] A.F. Tollefsen, H. Strand, H. Buhaug, PRIO-GRID: A unified spatial data structure, J. Peace Res. 49 (2) (2012) 363–374.
[37] T. Chadefaux, A shape-based approach to conflict forecasting, Int. Interact. 48 (4) (2022) 633–648.
[38] K.J. Arrow, R. Forsythe, M. Gorham, R. Hahn, R. Hanson, J.O. Ledyard, S. Levmore, R. Litan, P. Milgrom, F.D. Nelson, et al., The promise of prediction markets, Science 320 (5878) (2008) 877.
[39] J. Berg, F. Nelson, T. Rietz, Prediction market accuracy in the long run, Int. J. Forecast. 24 (2) (2008) 285–300.
[40] J. Wolfers, E. Zitzewitz, Prediction markets in theory and practice, 2006.
[41] D. Blei, A. Ng, M. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (Jan) (2003) 993–1022.
[42] H. Mueller, C. Rauh, Reading between the lines: Prediction of political violence using newspaper text, Am. Polit. Sci. Rev. 112 (2) (2018) 358–375.
[43] D. Blei, J. Lafferty, Dynamic topic models, in: Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 113–120.
[44] A. Bhadury, J. Chen, J. Zhu, S. Liu, Scaling up dynamic topic models, in: Proceedings of the 25th International Conference on World Wide, 2016, pp. 381–390.
[45] T. Rakthanmanon, B. Campana, A. Mueen, G. Batista, B. Westover, Q. Zhu, E. Keogh, Searching and mining trillions of time series subsequences under dynamic time warping, in: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data, 2012, pp. 262–270.
[46] Y. Jeong, M. Jeong, O. Omitaomu, Weighted dynamic time warping for time series, 2011, pp. 2231–2240.
[47] R. Agrawal, C. Faloutsos, A. Swami, Efficient similarity search in sequence databases, in: International Conference on Foundations of Data Organization and Algorithms, 1993, pp. 69–84.
[48] J. Paparrizos, L. Gravano, Fast and accurate time-series clustering, ACM Trans. Database Syst. 42 (2) (2017) 1–49.
[49] K.-P. Chan, A.W.-C. Fu, Efficient time series matching by wavelets, in: Proceedings 15th International Conference on Data Engineering, Cat. No. 99CB36337, IEEE, 1999, pp. 126–133.
[50] B. Su, X. Ding, H. Wang, Y. Wu, Discriminative dimensionality reduction for multi-dimensional sequences, IEEE Trans. Pattern Anal. Mach. Intell. 40 (1) (2017) 77–91.
[51] E. Keogh, K. Chakrabarti, M. Pazzani, S. Mehrotra, Dimensionality reduction for fast similarity search in large time series databases, Knowl. Inf. Syst. 3 (3) (2001) 263–286.
[52] R.J. Hyndman, E. Wang, N. Laptev, Large-scale unusual time series detection, in: 2015 IEEE International Conference on Data Mining Workshop, ICDMW, IEEE, 2015, pp. 1616–1619.

[53] F. Korn, H. Jagadish, C. Faloutsos, Efficiently supporting ad hoc queries in large datasets of time sequences, in: ACM Sigmod Record. Vol. 26, 1997, pp. 289–300.
[54] J. Chen, K. Li, H. Rong, K. Bilal, K. Li, S.Y. Philip, A periodicity-based parallel time series prediction algorithm in cloud computing environments, Inform. Sci. 496 (2019) 506–537.
[55] Y. Morinaka, M. Yoshikawa, T. Amagasa, S. Uemura, The L-index: An indexing structure for efficient subsequence matching in time sequence databases, in: Proc. 5th Asia-Pacific Conf. on Knowledge Discovery and Data Mining, 2001, pp. 51–60.
[56] S.J. Wilson, Data representation for time series data mining: time domain approaches, Wiley Interdiscip. Rev. Comput. Stat. 9 (1) (2017) e1392.
[57] I. Assent, P. Kranen, C. Baldauf, T. Seidl, Anyout: Anytime outlier detection on streaming data, in: International Conference on Database Systems for Advanced Applications, Springer, 2012, pp. 228–242.
[58] W. Luo, H. Tan, H. Mao, L. Ni, Efficient similarity joins on massive high-dimensional datasets using MapReduce, in: 2012 IEEE 13th International Conference on Mobile Data Management, MDM, 2012, pp. 1–10.
[59] Y. Li, M. Yiu, Z. Gong, Quick-motif: an efficient and scalable framework for exact motif discovery, in: 2015 IEEE 31st International Conference on Data Engineering, ICDE, 2015, pp. 579–590.
[60] C.-C.M. Yeh, Y. Zhu, L. Ulanova, N. Begum, Y. Ding, H.A. Dau, D.F. Silva, A. Mueen, E. Keogh, Matrix profile I: all pairs similarity joins for time series: a unifying view that includes motifs, discords and shapelets, in: 2016 IEEE 16th International Conference on Data Mining, ICDM, IEEE, 2016, pp. 1317–1322.
[61] G. Shmueli, To explain or to predict? Statist. Sci. 25 (3) (2010) 289–310.
[62] J.D. Fearon, Rationalist explanations for war, Int. Org. 49 (3) (1995) 379–414.

Thomas Chadefaux, professor in Political Science, Trinity College Dublin. Thomas received a Ph.D. in Political Science from the University of Michigan in 2009. He previously worked as an assistant professor at the University of Rochester (2009–10) and as a postdoctoral research at ETH Zurich. His research interests include the causes of war, conflict forecasting, and political methodology. He is currently the principal investigator of ERC grant PaCE: Patterns of Conflict Escalation (2022–26).