



ACLED CAST Methodology

The ACLED Conflict Alert System (CAST) is a global prediction tool that forecasts the number of political violence events that will occur each month, for the next six months, in every country around the world. Specifically, the tool predicts the ACLED *battles, explosions/remote violence, and violence against civilians* event types at the country and first administrative division level (i.e. province), in addition to producing an overall global forecast. New predictions are published each month for the following six months, alongside accuracy metrics for previous forecasts. All results are available for download as an [Excel file](#) or via an [API](#).

ACLED CAST is unique in its use of comprehensive ACLED data, its global, national, and sub-national scope, its disaggregated prediction of a variety of violent events, and its emphasis on the occurrence of these events and their escalation rather than only on conflict onset (i.e. violence in previously peaceful areas). It presents results as a discrete number of forecasted events rather than a probability of any violence taking place. In this way, the user is able to judge the specific scope and change in conflict, at both an aggregate level and at the event type level.

ACLED CAST allows users to view forecasted events in a specific area and, importantly, disentangle the factors driving each forecast and see how they compare to past trends. It is intended to be used in tandem with ACLED's [Early Warning Dashboard](#), which includes a range of additional resources that allows users to contextualize recent events in a particular country or subnational region within the larger conflict environment.

Model Construction

Predictors

ACLED CAST relies on a variety of indicators to generate its predictions. Most of the indicators are derived from ACLED data, as it provides the most comprehensive and up-to-date information about political violence and demonstration events, actors, locations, and fatalities globally. Events coded by ACLED do not have to meet a minimum fatality level for inclusion. Indicators include event counts from the previous month in neighboring administrative divi-



sions, interactions between different types of actors, fatality rates, and *strategic development* events, such as whether a peace agreement was signed within the last six months.

In addition to the indicators derived from ACLED data, ACLED CAST uses population estimates from [WorldPop](#), as well as infant mortality rates from the [Center for International Earth Science Information Network](#) as a proxy for economic development.

For a full list of indicators, their definitions, and which type of models they are included in, see Appendix A.

Subnational Models

ACLED CAST structures the data at the first administrative division (ADMIN1)-month level, where ADMIN1s correspond to first administrative divisions like districts, provinces, and states, depending on the country. The data enter a machine learning pipeline that emphasizes an accurate – yet interpretable – output and contextually informed predictions.

Data since 2018 are first split by country and a single country is selected.¹ The data for the selected country are then split into two sets: a training set and a hold-out set. The training data are used to tune the machine learning algorithms, while the hold-out data are used to assess forecasting performance over a six-month period. The training data are further partitioned into four time-series cross-validation splits, which sequentially hold out six months of data. These partitions simulate six-month forecasts using the historical data and guard against overemphasizing recent anomalies in the data that are not predictive of future violence.

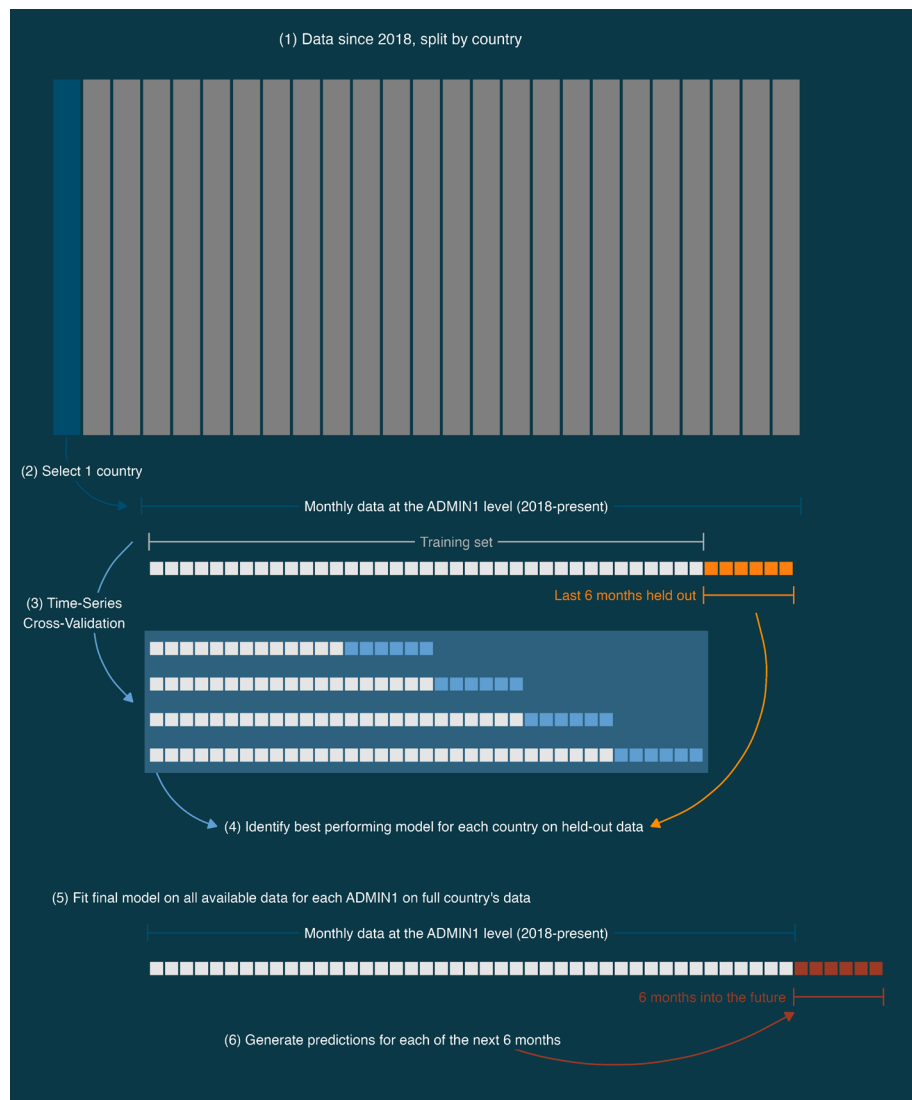
ACLED CAST then tunes a series of machine learning algorithms to the time-series cross-validation folds for the selected country. Four types of models are fit to the data for each outcome type (*battles*, *explosions/remote violence*, and *violence against civilians events*): a random forest model and an xgboost model, each using a “simple” and “complex” list of predictors.² In this way, ACLED CAST tests hundreds of candidate models that are designed to identify complex interactions among predictors over time, and selects the best-performing model for each country based upon its predictive performance in the held-out data.

After identifying the best-performing models for each country, ACLED CAST fits a final model

¹ Over half of ACLED countries have data available dating back to 2018, making it a useful threshold. It also allows for ample data on which to train the models. For countries that do not have data dating back to 2018, all available years of data are used.

² For details on which predictors are included in each list, see Appendix A.

for each country using the full data since 2018, with no held-out data. The finalized models are then used to generate forecasts for the coming six months. This is repeated for every country in the ACLED dataset.



In cases where an ADMIN1 averages fewer than one event a month, the model is still trained on the data but the results are omitted from the dashboard. In the event that a low-event ADMIN1 is predicted to have more than five events in the next month (i.e. a case of 'conflict onset'), the results are reported in the dashboard and marked as 'Onset' in the 'Predicted change' column.



Aggregation

All subnational results per country are aggregated to produce the country-level prediction. This provides more accurate, granular forecasts for a country that take into account the unique features of each administrative division. It also allows users to consider both subnational predictions as well as a country-level total forecast that intuitively reflects a sum of those predictions: the country-level forecasts provided in the ACLED CAST dashboard and results file are the summed values of all of the component ADMIN1 divisions; any discrepancies are due to rounding. Similarly, the global forecast provided by ACLED CAST is an aggregation of the various country forecasts.

Results

Monthly Reporting

ACLED CAST results are released on the first or second Thursday of each month after complete data for the previous month have been published. For example, the April 2023 update will be released on 6 April 2023, because all events from March 2023 will be published on 3 and 4 April 2023.³ This schedule allows the ACLED CAST models to run on complete data from the past month, producing more accurate and consistent results.

Results are reported for the upcoming six-month period and updated monthly. For example, the March 2023 update will include forecasts for March, April, May, June, July, and August 2023. The April 2023 results will update the April-August forecasts with fresh data, and also produce a September 2023 forecast.

Results are made available through the ACLED [CAST dashboard](#) as well as in a downloadable file on the main ACLED [CAST webpage](#) and via [API](#). The results file includes the forecasted events for all months (beginning at the time of launch, March 2023), as well as the number of observed events for each month once the month is complete. Note that the observed values reflect the number of events in the ACLED dataset at the time stated in the *timestamp* column. ACLED is a living dataset, meaning that the numbers in the *observed events* columns may not match the number of events for that month in the main ACLED dataset, depend-

³ ACLED data are updated at the start of every week, through the most recent Friday. Data for Africa, the Middle East, and Latin America & the Caribbean are updated on Mondays, and data for the Asia-Pacific, Europe & Central Asia, and the United States & Canada are updated on Tuesdays. Please note that years of historical coverage vary across countries and regions. A full list of country and time period coverage is available [here](#).



ing on future updates and date of access. The observed values are intentionally kept static in order to provide the most accurate error rates for the CAST models – allowing users to compare forecasts to the observed values that existed in the data *at the time the forecast was made*.

Change From Designated Moving Average

In the ACLED CAST dashboard, all forecasts are presented alongside the moving average for that country or administrative division. The time horizon of the moving average is customizable by the user. This comparison point is provided in order to contextualize the forecasts within the larger scope of a country or administrative division's conflict environment. For example, if the 12-month moving average is selected, the dashboard will show the selected month's forecast alongside the 12-month moving average for that country or administrative division. This shows, on average, how many *battles, explosions/remote violence, or violence against civilians* events occurred in that area in the last 12 months. The time range for the moving average is customizable to allow for dynamic conflict environments; if, for example, a country has a recent spike in conflict events, the user may wish to set the moving average to just the last one to three months.

The ACLED [Early Warning Dashboard](#) can provide further context for a country or administrative division. This dashboard provides various types of information about a country or administrative division's conflict situation and history. By using this tool in conjunction with ACLED CAST, users can put the forecasts for an area in context with the recent developments or trends in that space. The Early Warning Dashboard shows the present and past context of a conflict, while ACLED CAST forecasts the future of that conflict.

What's Driving the Forecasts?

A key feature of ACLED CAST is its demonstration of how various indicators shape the final forecast for an administrative division. The 'What's Driving the Forecasts?' section of the ACLED CAST dashboard shows how a forecast begins at the 'baseline' number of events (i.e. the number of events that would be forecasted based solely on the history of that administrative division, excluding the effect of any other indicator). It then displays the indicators with the largest contribution to the forecast, and how many events were added or subtracted from the final forecast because of them.⁴

⁴ The explanations of each forecast are generated using the model-agnostic approach described [here](#).



In this way, users can see the evolution of a forecast, from starting assumption ('baseline') to the final forecasted value. It makes it clear which indicators play a large part in predicting events in certain administrative divisions over others, which is a level of transparency that is useful for understanding a conflict context in more detail.

At the country level, the 'What's Driving the Forecasts?' section shows the aggregation of these important indicators. For each indicator, it shows what percentage of the country's ADMIN1-level predictions relied heavily on it.

Accuracy

Transparency is an important element of the ACLED CAST methodology. Conflict forecast models are never 100% accurate, and users need to be aware of their limitations when using their results in real-world applications. Accuracy rates are also not consistent across all countries or administrative divisions in the data – a stable conflict with an established pattern of events will have more accurate forecasts compared to a new, highly volatile conflict with no established pattern of events. For this reason, the ACLED CAST dashboard contains an 'Accuracy' section that reports on the accuracy of the forecasts for previous months. These figures are updated monthly as the observed event counts for that month are complete. In other words, in April 2023, the 'Accuracy' section will display what was forecasted for March 2023 and how many events actually occurred in March 2023.

The 'Accuracy' section contains several metrics that can be used to gauge the accuracy of forecasts for a given country or administrative division. The mean error and mean absolute error show the difference between the forecasted value and the observed number of events. The mean absolute error shows the absolute difference, while the mean error shows whether events are under or over-counted.

Similarly, the mean percent error and the mean absolute percent error represent the *percentage* difference between the forecasted event count and the observed event count. Again, the mean absolute percent error represents the absolute difference between the two figures, while the mean percent error shows the directionality of the difference.

ACLED CAST provides both types of metrics – the raw number and the percentage value – because both can be valuable in different situations. In cases with a very low number of observed events, a prediction that is off by just one event can register an extremely high percentage error. Conversely, in situations where there are regularly a large number of events, it can be more useful to use the percentage error rate as it puts a potential 100+ event error into context.



Access

Access to the ACLED CAST dashboard and a downloadable file of monthly predictions is available to users who have registered in the ACLED Access Portal. Access to ACLED CAST is free to the public for extensive, *though not unlimited*, use. The extent of free access for users is determined based on the category of the user (as documented during the registration process). Once the category of user is determined, for those who are not Media, an existing Sponsor or Partner, or an ACLED Network member, free access will include 10 logins to the ACLED CAST dashboard. Those who wish to gain further access to ACLED CAST are invited to contact ACLED to discuss their needs. Please send any questions to access@acleddata.com.

All use of ACLED CAST must abide by the ACLED [Terms of Use & Attribution Policy](#). If you wish to reproduce or republish a visual, graphic, or map from ACLED CAST (rather than creating an original image using raw data) for non-commercial purposes, please cite ACLED CAST using the following format:

ACLED. (DD Month YYYY). "ACLED Conflict Alert System." *Armed Conflict Location & Event Data Project (ACLED)*. <https://acleddata.com/early-warning-research-hub/conflict-alert-system/>
© 2023 ACLED All rights reserved. Used with permission from ACLED. Accessed (DD Month YYYY).



Appendix A: Glossary

NAME	DESCRIPTION	MODEL(S)
Battles	Violent interactions between two politically organized armed groups at a particular time and location. Battles can occur between armed and organized state, non-state, and external groups, and in any combination therein. There is no fatality minimum necessary for inclusion. Included as <i>Battles (t-1)</i> when used as a predictor.	Simple & Complex
Violence Against Civilians	Violent events where an organized armed group deliberately inflicts violence upon unarmed non-combatants (civilians). Included as <i>Violence Against Civilians (t-1)</i> when used as a predictor.	Simple & Complex
Explosions/ Remote Violence	One-sided violent events in which the tool for engaging in conflict creates asymmetry by taking away the ability of the target to respond. Included as <i>Explosions/Remote Violence (t-1)</i> when used as a predictor.	Simple & Complex
(t-1)	Represents the value of the accompanying predictor from the previous month. Example: <i>Battles (t-1)</i> represents the number of Battles events from the previous month.	
Protests	A public demonstration in which the participants do not engage in violence, though violence may be used against them. Events include individuals and groups who peacefully demonstrate against a political entity, government institution, policy, group, tradition, businesses or other private institutions. Included as <i>Protests (t-1)</i> when used as a predictor.	Simple & Complex
Riots	Violent events where demonstrators or mobs engage in disruptive acts, including but not limited to rock throwing, property destruction, etc. They may target other individuals, property, businesses, other rioting groups or armed actors. Included as <i>Riots (t-1)</i> when used as a predictor.	Simple & Complex
Excessive Force Against Protesters	Events where individuals are engaged in a peaceful protest and are targeted with violence by an actor leading to (or if it could lead to) serious/lethal injuries. Included as <i>Excessive Force Against Protesters (t-1)</i> when used as a predictor.	Simple & Complex

Total Violence	Sum total of all organized violence event times in a given month (total of battles, violence against civilians, and explosions/remote violence). Included as the 6 month moving average and standard deviation when used to generate predictors.	Complex only
Fatalities	The number of reported fatalities which occurred during an event. Included as <i>Fatalities (t-1)</i> when used as a predictor.	Complex only
Actor Concentration	A Herfindahl-Hirschman Index tracking the number of active conflict actors and for how many violent events each is responsible. Included as <i>Actor Concentration (t-1)</i> when used as a predictor.	Complex only
Actor Interactions	Interactions among the actors in an event, grouped by the types of actors engaged in the event. Group types include State Forces, Rebel Groups, Political Militias, Identity Militias, Rioters, Protesters, Civilians, and External/Other Forces. An event involving a state military and an armed rebel group, for example, would be a State Forces-Rebel interaction. Events with only one actor are coded as sole actions (e.g., State Forces sole action). Included as the <i>Interaction-Type (t-1)</i> when used as predictors, where the variable represents the sum total of events within that actor type in the month prior. For example, <i>State Forces-Rebel Interactions (t-1)</i> is the count of state military and rebel group interactions in an ADMIN1 the month before.	Complex only
Violence in Neighbors	Battles, Explosions/Remote violence, and Violence against civilians events in neighboring ADMIN1 locations. Included as <i>Violence in Neighbors (t-1)</i> when used as a predictor.	Simple & Complex
Strategic Developments	Contextually important information regarding the activities of violent groups that is not itself recorded as political violence, yet may trigger future events or contribute to political dynamics within and across states. Included as <i>Strategic Developments (t-1)</i> when used as a predictor.	Simple & Complex
Agreements	Agreements between different actors (such as governments and rebel groups) within the previous six months. Examples include peace agreements/talks, ceasefires, evacuation deals, prisoner exchanges, negotiated territorial transfers, prisoner releases, surrenders, repatriations, etc. Included as <i>Agreements (t-1)</i> when used as a predictor.	Complex only
Time Trend	A set of variables capturing many potential temporal trends, including those specific to years, months, and quarters, as well as a linear monthly trend since 2018.	Simple & Complex

Infant Mortality	Estimated ADMIN1-level infant mortality rates, derived from the Center for International Earth Science Information Network's (CIESIN) Global Subnational Infant Mortality Rates rasters.	Complex only
Population	Estimated ADMIN1-level population, derived from WorldPop population rasters	Complex only
Temporal Weights	Importance weights for each ADMIN1-level observations, where observations more recent in time are more heavily weighted.	Simple & Complex
Relative Time Comparison	Time period in the past to compare forecasted violence with.	
Predicted Change Category	Level of predicted change from selected time period.	
Forecast Date	Month within the next 6 months to view forecast.	
State Forces	Collective actors that are recognised to perform government functions, including military and police, over a given territory. Referred to as "Military" in interactions.	
Rebel Groups	Political organizations whose goal is to counter an established national governing regime by violent acts.	
Political Militias	Armed, organized groups with political goals that use violence to advance those goals. Unlike rebel groups, political militias generally do not actively seek to topple or replace the national government using violence, though some are organized in opposition to government authority.	
Identity Militias	Armed and violent groups organized around a collective, common feature including community, ethnicity, region, religion, or, in exceptional cases, livelihood. Therefore, identity militias captured in the ACLED dataset include those reported as tribal, clan, communal, ethnic, local, community, religious, and livelihood militias.	
Rioters	Individuals who engage in violence during demonstrations and mob violence events. Violence can be directed against people, property, or both.	

Protesters	Peaceful, unarmed demonstrators. Although protesters are nonviolent, they may be the targets of violence by other groups (e.g. security institutions, private security firms, or other armed actors).	
Civilians	Civilians, in whatever number or association, are victims of violent acts within ACLED as they are, by definition, unarmed and, hence, vulnerable. Some normally armed actors may be coded as civilians if they are targeted with violence in situations where they are caught unarmed.	
External/Other Forces	International organizations, state forces active outside of their main country of operation, private security firms and their armed employees, and hired mercenaries acting independently.	