



Version 2 (12.2019)

ACLED Methodology and Coding Decisions around the Conflict in Afghanistan

The current main conflict in Afghanistan began in 2001, following the post-9/11 US-led North Atlantic Treaty Organization (NATO) invasion and the subsequent overthrow of the Taliban government. In 2003, the Taliban announced it had officially regrouped, and a year later announced it had begun an insurgency under Mohammed Omar against both foreign forces and the US-installed Transitional Islamic State of Afghanistan, and later the Islamic Republic of Afghanistan. The Taliban's goal is to regain political power over the country and to push out foreign forces. As of 2019, the conflict between the Taliban and joint domestic and international forces is at a relative stalemate, with all parties controlling a number of districts while many others continue to be contested ([Long War Journal, 2018](#)). However, the war is extremely active, with a consistent counter-insurgency campaign across the country. The Taliban's numbers are often bolstered by fighters coming from countries such as Pakistan, Uzbekistan, and Tajikistan, though the majority of them still identify as Pashtun ([Foreign Policy, 15 June 2016](#)). The group's allies include both local and international groups such as Tehrik-i-Taliban Pakistan (TTP), the Haqqani Network, and Al-Qaeda, in addition to a number of other smaller militant groups. The group allegedly receives financial support from a number of countries, and also heavily relies on profits from the opium trade; airstrikes by both government and international forces often target drug facilities with this in mind.

In addition to the ongoing contest with the Taliban, the Islamic State announced the formation of an Afghanistan/South Asia affiliate which refers to itself as Islamic State of Iraq and the Levant – Khorasan Province (ISIL-KP) in 2015. Its membership is composed of a number of foreign fighters in addition to those formerly affiliated with the TTP, Taliban, and other militant groups. The group pledged allegiance to the ISIS leader al-Baghdadi, and its goal is to incorporate "Greater Khorasan" (an historical region encompassing parts of Central Asia, Afghanistan, and Pakistan) as a province of the imagined Islamic Caliphate declared by the Islamic State. For this reason, ACLED refers to the group as Islamic State (IS), following the standard name in all other countries we cover with the group's activity. The presence of IS has added a new element to the conflict as a common enemy for both the Taliban and Afghan security forces (and NATO) to combat. While the Taliban's goal requires them to gain the trust of the people, IS will often carry out suicide attacks in crowded areas resulting in high civilian casualties (for more on that, see this [ACLED](#)



[analysis piece](#)). As of December 2019, Taliban and Afghan/NATO operations against IS have pushed the group back to its original stronghold of Nangarhar province – the group formally having a significant presence in both Kunar and Jowzjan.

Afghanistan presents some unique methodological challenges for the recording of political violence. These primarily concern media coverage, which often suffers from reporting biases and little access to Afghanistan’s most remote regions, or regions controlled by insurgent groups. We conclude that these biases lead to both inflation and deflation of fatalities, depending on who is reporting the events. Additionally, reports of conflict events are often vague, using general terms such as “security forces” and non-descript “militants” or “terrorists”, in addition to aggregating what may be several distinct events into single provincial or district overviews. This report aims to outline ACLED’s efforts to address these challenges and accurately capture manifestations of political violence in Afghanistan.

What does ACLED cover in Afghanistan?

ACLED’s coverage of political violence and protests in Afghanistan spans from January 2017 to the present, with continued weekly real-time data releases. Efforts are currently underway to expand and improve the current coverage through additional sourcing.

Which armed actors are recorded?

- **Government Forces** constitute a number of state-sponsored security organizations, including the military and police forces, in addition to a third group known as the **Afghan Local Police (ALP)**, which are a government-supplied paramilitary, local police force which acts as support to the Afghan National Police. For general “security forces” where the source neither specifies police nor military forces, **Military Forces of Afghanistan (XXXX)** is selected as the default. Military and police forces have been further broken down into popular sub-actors for more advanced actor analysis. For police, these include wings such as:

- **Police Forces of Afghanistan (XXXX) National Directorate of Security:** the primary intel agency in Afghanistan. Often the target of bombings. NDS is a common acronym.
- **Police Forces of Afghanistan (XXXX) Sangorians:** special forces designed to infiltrate Taliban cells with support from NDS.
- **Police Forces of Afghanistan (XXXX) Special Forces:** police commandos/SWAT.



For military forces, the only sub-actor coded thus far in the data is:

- **Military Forces of Afghanistan (XXXX) Special Forces:** often described as “commandos”. When engaging in airstrikes, they are accompanied by Military Forces of Afghanistan (XXXX) as an associate actor since it is the Airforce providing the planes. There is no sub-actor for the Airforce since analysis on this group can be done by simply sorting on airstrikes made by Military Forces of Afghanistan (XXXX). NATO are the only other group with airplanes in the current Afghan conflict, and they will be mentioned specifically when known to be involved in an event.

The actor **Militia (Pro-Government)** – often described as “arbakis” – which is a word that loosely translates to “militia” – are used to describe militia groups unofficially sanctioned by the Afghan government / NATO. While some sources place ALP under the “arbaki” term, others (including ACLED) have chosen to represent them as a distinct unofficial version of the ALP (untrained, non-officially sanctioned). These are separate from **Communal Militias**, which do not necessarily work with state forces and instead are armed communities/groups defending their larger communities and localities.

• **Militant/Insurgent Groups:** By far the most active insurgency in the country is led by the **Taliban**. Due to the size of the group, ACLED also records a sub-actor for this group called **Taliban - Red Unit**, which is a special forces/commando unit. Despite the group’s prevalence, ACLED does not make assumptions that unidentified militant groups mentioned in sources are the Taliban, unless the events occur in areas where known Taliban operations are occurring at the time (for example, during the siege of Ghazni city). This is because other insurgent groups are active in Afghanistan during the current period of coverage. Among these is the Khorasan branch of IS – referred to as **Islamic State (Afghanistan)** – which is also often the target of both Afghan security and NATO operations. While the Taliban is active throughout the country, the Islamic State’s presence is less widespread, with the majority of events involving them occurring in the provinces of Nangarhar, Kunar, and Jowzjan. Because both groups have similar goals, they often battle for supremacy in areas where their control overlaps.

For Ministry of Defense (MOD) press releases, the actor **Taliban and/or Islamic State (Afghanistan)** is often used for general “military operations” where the specific group is not mentioned. This is because we know that these large-scale operations are targeting one of these groups. **Unidentified Armed Group (Afghanistan)** is used for all other cases



where the armed group is not specifically named, since there are other groups operating within the country on a smaller scale.

Smaller groups operating throughout the state include:

- **Haqqani Network:** an offshoot of the Taliban which operates both independently and in support of other Taliban groups. Active in the eastern provinces of the country.
- **Al-Qaeda:** Global network of jihadists which formed as a response to the Soviet war in Afghanistan. Were targeted by the US-led invasion of Afghanistan on the pretense of being harbored by the Taliban government. Currently, they are active in provinces that border Pakistan, although their presence appears to be significantly lessened. They are allied to the Taliban.

To a lesser degree, small groups from Pakistan occasionally make an appearance within Afghanistan – possibly as a result of using the country as a base of operation. The most prominent of these groups is **Lashkar-e-Islam (LeI)**, who often clash with IS. Lastly, several political party militias are also present to a very small degree, including **Jamiat-e Islami** and **Hezbi Islami**.

• **External Forces:** There are a number of other state military forces which operate within Afghanistan, all of which can be sorted in the data by their interaction code (8) – a cross-ACLED code used for ‘other’ forces, including state forces operating outside their own countries. By far, the most prominent foreign force in Afghanistan is **NATO: North Atlantic Treaty Organization**, which is a joint force made up of several individual international militaries. They are often described by name, or simply as “foreign” forces. Occasionally, sources will refer to a specific NATO member (example: “supported by US airstrikes”), in which case the specific country’s military forces is represented as an Associate Actor to NATO; for example, **Military Forces of United States**. If NATO is supporting Afghan military forces, either NATO, or NATO *and* the specific country, are represented in the appropriate Associate Actor column. This format is also used for other coalition forces operating in other countries, such as the Global Coalition Against Daesh or Operation Restoring Hope. For more information, refer to this [ACLED piece](#) on NATO in Afghanistan.

Pakistani state forces occasionally fire shells across the border (Durand Line) into Afghanistan, especially into the provinces of Kunar and Nangarhar, which border the former Federally Administered Tribal Areas (FATA). Shells are often fired by both **Military Forces of Pakistan (XXXX)** and **Police Forces of Pakistan (XXXX) Frontier Corps**.



Although the sources rarely mention targets, and casualties are also rare in these events, it is assumed that the shelling acts as a deterrent for cross-border militant activity. In a few events, Pakistani and Afghani border forces have clashed.

How are events sourced?

Each week, ACLED researchers review approximately 60 sources in English and Dari/Farsi, in order to provide the most comprehensive database on political violence in Afghanistan.

A primary challenge for sourcing events in Afghanistan is that the country is not only the sight of an ongoing and deadly conflict, but also that a great deal of this conflict occurs in areas that remain inaccessible to most. The largely rural nature of Afghanistan, coupled with the ongoing conflict, restricts access to much of the country by journalists. In many cases, the primary conflict parties — the Afghan government and the Taliban — may be the only source of information for a particular event, reporting on this information in reports from the Afghan Ministry of Defense (MOD) and the Taliban’s Voice of Jihad (VOJ) website, respectively.

Indeed such groups, as they are parties to the conflict themselves, may have incentive to share biased information. As such, including them may introduce bias into the data, such as inflated or deflated fatality counts. However, not including them introduces another type of bias into the data — such as a biased view of where violence is happening, with events only appearing in areas where reputable sources have immediate access. Small-scale skirmishes or assaults in remote areas often occur in areas where more reputable sources lack access, so limiting information gathering to only information reported on by ‘reputable sources’ would mean such events would be missed, resulting in these spaces appearing ‘peaceful’ when they may not be.¹ As such, choosing to include or exclude information from these sources comes at a cost regardless of the final decision made.

ACLED finds that it is important to consider these events, and has determined that not necessarily all aspects of the information these sources report is biased (more on that below).

¹ Certain types of events may also be missed. Traditional English media, for example, may report targeting of civilians at the hands of rebels at a higher rate than that occurring at the hands of governments, despite the fact that the latter may be prevalent. For example, a [UN report](#) notes that Afghan and US forces killed more civilians than the Taliban in Afghanistan in the first half of 2019: something that has not been reported often in media. The Taliban, however, reports on a number of such events, several of which are corroborated by independent or international sources like the New York Times.



This is why national media journalists in the region too rely heavily on information from these sources as they are not able to access all parts of the country. Approximately three-fifths of the information used in ACLED's coding of disorder in Afghanistan comes from the MOD and VOJ. These sources are coded as source scale 'Other' in the data.

ACLED has found that these sources tend to be reliable in their coverage of things such as whether an event occurred or not. Even in cases where there may be a discrepancy between the two conflict parties in their reporting — e.g. the MOD denying the Taliban's capturing of a district — other sources, such as local media, may corroborate the occurrence of the event. This is why, unless otherwise determined, events reported by these armed groups in the Afghan context are considered to have happened, although specific details are assumed to be less reliable.

There is more discrepancy around who the instigator of the violence is and fatality estimates. ACLED does not specify who the instigator of a clash is, and our fatality specifications for Afghanistan are specific to the reports. Across all contexts, fatality numbers are the most biased aspect of reporting and ought to be considered carefully by users. In the Afghan context, where reports from the conflict parties themselves are heavily relied upon, this bias is even further exacerbated. To account for this, ACLED prioritizes fatality estimates reported by sources other than the MOD or VOJ, whenever available, in addition to taking other precautionary measures (see the [How does ACLED record fatalities in Afghanistan](#) section below).

While national media might use information from these conflict parties in their own reporting at times, the biases of traditional media are still evident — namely, the appetite of traditional media consumers for larger-scale events, with smaller skirmishes still tending to go unreported by such media outlets. This means that relying on national media alone will still result in missing a certain subset of events, even if national media relies on reporting from the MOD and VOJ too.

Per ACLED methodology, the ACLED source column will not note every single source/outlet that may have reported on a specific event; in the case of large-scale sensational events, the number of sources reporting on such an event could be in the hundreds, with information across outlets largely just repeated. Rather, ACLED only codes the name of every source which contributed unique information to the coding of the event. This means that if Source X reports on an event, and the researcher uses information from this source to code the event, the Source X is coded as the source for the event. If the researcher comes across the same



event reported by Source Y and Source Z, Source Y and Source Z will only be included in the source column if they contribute additional novel information that was not already captured by Source X. In the Afghan context, this means that if an event is reported on by VOJ, and the researcher uses the VOJ to code the event, VOJ is listed as the source. If national and subnational media sources are later consulted as they report on the same event, yet they include no new information (e.g. if they too originally got the information from the VOJ), then these other sources will not be added in the source column in the data as they did not provide any new information. (For more on ACLED methodology, see this [sourcing primer](#) or the [ACLED Codebook](#).)

In addition to information from the MOD and VOJ, ACLED also relies on local media sources. Each week, ACLED researchers review over 10 subnational sources in English and/or Dari. While only about two percent of information used in ACLED's coding of disorder in Afghanistan comes from subnational media sources (outlets such as local radio or TV stations), these sources typically cover reporting by local (district or provincial) authorities. The details around things such as fatality numbers, victims of violent events, or territorial transfers may differ from MOD reporting. As such, when available, such sources are a helpful tool in corroborating information from other sources. In a similar fashion, the independent website iCasualties is used to verify the deaths of NATO forces in Afghanistan when reported by other sources, such as VOJ. iCasualties gets its information from news reports and press releases from the US Department of Defense, CENTCOM, the MNF, and the British Ministry of Defence.

In addition to information from the MOD and VOJ, as well as subnational media sources, ACLED also relies on information from traditional media at the national, regional, and international scales. Just under one-third of the information used in ACLED's coding of disorder in Afghanistan comes from such traditional media sources. Each week, ACLED researchers review over 40 such media sources in English and/or Dari. International media sources tend to cover larger scale events. Exceptions to that include NYC Reporting's weekly Afghan War Casualty Report, which contains information that the news source's local correspondents confirm (and in this way is much more akin to more local reporting); this source has been helpful in corroborating a number of VOJ reports and in this way has also helped to lend legitimacy to ACLED's inclusion of VOJ as a source. Additionally, the Bureau of Investigative Journalism's reporting on US drone strikes provides both new events in addition to further information on strikes otherwise reported as simply 'unknown airstrikes'. A number of the primary national news outlets are covered as well, including: Bahktar, Khaama Press, Tolo News, and Pajhwok, of which the latter yields the most unique new



events.

Lastly, while ‘new media’ can be a powerful supplemental source in some contexts, it varies widely in its helpfulness in capturing trends across contexts. In Afghanistan, “social media is still in its infancy ... with a small and mostly homogenous user base of educated, relatively wealthy, predominantly male users” ([InterNews, 25 January 2018](#)). Similar patterns have been seen across other countries as well, even in cases where social media might be more popular — such as in Kenya ([Dowd et al., 2019](#)). Given the less representative sample and the fact that such tools are still not incredibly popular in the Afghan context, use of new media has been limited in sourcing information around disorder in Afghanistan by ACLED.

There has been some use of new media in this context, however. This includes information from LiveUAMap. LiveUAMap is an online news aggregator whose analysts and editors fact-check before displaying relevant conflict and disaster coverage through an interactive online map tool. ACLED has a partnership with LiveUAMap, information from which is used to supplement coverage of disorder across a number of countries of ACLED coverage. While the media sources used by LiveUAMap in the Afghan context are near identical to ACLED’s source list, their coverage of multiple verified Twitter accounts has been useful for covering events not captured by traditional or conflict party media. Mostly notably, these include smaller-scale attacks by the Islamic State (IS), and fighting between IS and the Taliban. Events coded based on information from LiveUAMap are noted as such in the source column.

Including LiveUAMap, fewer than one percent of events in the Afghan dataset come from ACLED’s local partners. In addition to LiveUAMap, this includes information from some of ACLED’s partners with global coverage. These include [Front Line Defenders](#), who reports on attacks against human rights defenders, as well as the [Aid Worker Security Database](#), who reports on attacks against aid workers.

Where does violence take place, and how are locations recorded in Afghanistan?

Political violence in Afghanistan is country-wide; however, the geography of the Hindu Kush Mountains and their extending ranges means that the majority of events occur in the crescent surrounding the central mountainous area. Conflict events are also heavily clustered in provinces through which the highway linking Jalalabad-Kabul-Kandahar-Lashkargah passes, in the east of the country.

ACLED has georeferenced events over 1,000 distinct locations across Afghanistan. These



include cities, towns, villages and other populated places, as well as natural locations like mountains and valleys. Depending on the precision of the sources and the size of the recorded location (town, district, or province), researchers will select the appropriate GeoPrecision code to reflect the precision of the geocoding. Due to the remoteness of many parts of Afghanistan, in addition to varying levels of government presence in the rural areas of the country, the majority of sources will only identify the district or province in which an event has occurred (which are coded at GeoPrecision 2 or 3, respectively).

A GeoPrecision code of 2 is also used when sources give the name of a village or general area, yet our researchers are unable to find coordinates for this specific location; in these cases, we code at the district capital and mention the village/area name in the event notes, unless further identifying information is provided to use in coding at a more natural location. Various tools, including GeoNames, OpenStreetMap, and OCHA atlases, are used in an attempt to find these specific locations; however, a wide range of transliterations and the presence of colloquial names for locations makes this a difficult process. With this in mind, appropriately using GeoPrecision codes can help to control for any urban bias in the data, especially during any analysis on conflict in rural versus urban settings.

Additionally, multiple unknown locations within the same district or province are combined into a single event if all other details are the same. For example, “On August 7, 2018, 13 Taliban militants were killed by Afghan security forces in the Balakhel and Paen areas of Kot district, Nangarhar” would be coded as two events in Balakhel and in Paen. However, if for any reason we are not able to specifically locate the Balakhel and Paen areas (due to different spelling, local slang, name changes, etc.), a single event would be coded in Kot district (GeoPrecision 2), as opposed to coding two identical events, each aggregated to Kot district to denote these events.

Along those lines, in an effort to code conservatively, similar events involving the same actors in the same location (with the same GeoPrecision) on the same day are considered to be the same event, even if casualties differ (due to the wide discrepancy in reported fatalities; for further info, see ***How does ACLED record fatalities in Afghanistan?*** below). Events coded with GeoPrecision 1 or 2 take precedence over events with GeoPrecision 3 (where geographic information is fuzziest and only the province is known). For example, if one report states that “21 Taliban militants were killed in military operations in Ghazni province”, and another report says that “18 Taliban militants were killed by soldiers in Ab Band district of Ghazni province” on the same day, the latter will be assumed to be the same event and chosen due to its more specific location given that the former “21 militants killed” event *could in fact be* the “18 Taliban militants killed” event occurring in Ab Band



specifically, with fatality numbers not matching. This is not a stretch given the nature of fatality estimates.

How does ACLED record fatalities in Afghanistan?

Fatality counting in regards to the Afghan insurgencies faces several obstacles, which make any effort to provide reliable estimates even more difficult. Scarce or biased reporting, as well as limited media access to the sites of violence, may indeed result in substantially different fatality estimates arising from the same event, uncertain figures, or one-sided coverage of conflict events in certain areas. Both the Taliban and Ministry of Defense have reason to avoid reporting their own losses while also inflating the losses of the other side; the war they are fighting is protracted and requires the continual support of the population and ongoing recruitment ([New York Times, 21 September 2018](#)). As such, fatality counts reported by independent or local sources take precedence when determining how many fatalities to code.

To avoid artificially increasing the number of reported fatalities, ACLED has taken steps to ensure that fatality estimates related to the conflict in Afghanistan are the most accurate possible and least subject to media biases. As approximately three-fifths of the information used in the Afghan dataset comes from the MOD and VOJ, and the fact that these sources are often the only ones reporting on events, recording fatalities as zero in all cases when the MOD and/or VOJ are the only source for an event would significantly undercount fatalities in Afghanistan. To ensure that fatality estimates for the war overall are closer to reality, yet still erring on the side of being more conservative and mitigating bias, ACLED seeks to avoid including fatality estimates from both the MOD and VOJ. Doing so would result in significantly over-counting fatalities, as both of these sources are assumed to be inflating fatality numbers. That being said, when no other source of information for an event exists, ACLED defers to coding estimates from the MOD rather than VOJ. ACLED assumes that, of the two sources, the Afghan government has more oversight than the Taliban when it comes to claims around fatalities, especially in larger attacks. Fatality estimates from VOJ are assumed too biased to include. This means that if VOJ is the only source of information for an event, fatalities for that event are recorded as zero. This protocol results in the total fatality estimates from these data to be closer to reality, yet still errs on the side of being more conservative. While the MOD fatality estimates included may be inflated, this is offset by the undercounting of fatalities stemming from events in which VOJ is the only source of information. *To reiterate, VOJ reports are coded with zero fatalities.*

Meanwhile, across all sources, researchers triangulate reported fatality counts to always



select the most conservative available, unless more recent or verified information is released. For example, Khaama Press reports that “On October 13, 2018, 12 civilians were killed by a NATO airstrike in Ghazni city”, while Agence France Presse reports that 10 were killed in the same incident. Here, we would choose 10 as the more conservative estimate between the two sources. However, if the UN sends a team to investigate and reports that 11 civilians were killed, the event will be updated to include this seemingly more reliable report;

If fatality estimates are unknown – and this happens often as many reports tend to be vague and only mention the occurrence of “deaths and injuries” or “losses” – ACLED uses a standard estimate of 10 fatalities for Afghanistan, or 3 when the event is known or likely to have caused fewer than 10 fatalities (e.g. the bombing of a motorcycle resulting in some fatalities, or a small skirmish at a remote checkpoint). Additional intermediate estimates are used to capture other inaccurate reported figures, such as dozens, scores, etc.

Fatalities by Actor: ACLED **does not** code fatality figures according to which group suffered fatalities because many source reports do not offer this level of detail, and when they do, this information may be biased; instead, ACLED reports on the total number of deaths reported from a conflict event. The only exception to this is in incidents involving civilians: because ACLED treats civilians as unarmed, non-combatants, the number of fatalities reported for each event involving only civilians as Actor 2 – typically coded under “Violence against civilians”, or under “Explosions/Remote violence” – is taken to be the number of civilians killed (with the exception of events like suicide bombings where the perpetrator’s death will also be included in the total fatality count). As such, **aggregate estimates of “civilian fatalities”² do not include civilians that may have died as ‘collateral damage’ during fighting between armed groups or as a result of the remote targeting of armed groups** (e.g. an airstrike hitting militant positions but that also kills civilians). These collateral damage fatalities are still recorded in both the Fatalities and Notes columns – meaning that any analysis of the total impact on civilians, including battles, will require additional modification on the part of the user, along with a number of assumptions as to how to attempt to disaggregate these total event fatality numbers.

Splitting Fatalities: Many sources will release a single fatality total referring to events occurring across a number of locations; for example, “On Dec 2, 33 Islamic State militants were killed during NATO airstrikes in the Rodat, Kot, and Khogyani districts of Nangarhar”. In

² ACLED has developed a data box that allows users to download all ‘Violence against civilians’ and ‘Explosions/Remote violence’ events targeting unarmed protesters and civilians into a single file. It is available for download at the bottom of the page [here](#).



this case, the fatalities are split equally over three events (in Rodat, Kot, and Khogyani), 11 fatalities for each location. Odd number totals will be split as evenly as possible; for example, 34 fatalities over three events would be split into 11, 11, and 12. The notes will be the same for each event and will state that the fatalities have been split. ***It is important to remember that though the Notes column entries may match across these events, these are not 'duplicates' but rather are unique events with unique locations coded.***