UCDP Candidate Events Dataset Codebook
Version 1.0
Uppsala Conflict Data Program

Department of Peace and Conflict Research
Uppsala University

This version authored by Stina Högbladh (2020)

Data extracted from UCDP systems on the 19th of every month

When using this data, please always cite:

Forthcoming Hegre, Håvard, Mihai Croicu, Kristine Eck, and Stina Högbladh, 2020. *Introducing the UCDP Candidate Events Dataset*", Research and Politics

When appropriate, also cite this codebook: Högbladh Stina, 2019, "UCDP Candidate Events Dataset Codebook, v.1.0", Department of Peace and Conflict Research, Uppsala University.

Contents

Introduction

The UCDP Candidate data includes the same variables as the UCDP GED dataset, the main difference being that UCDP Candidate has a monthly release cycle and UCDP GED is released annually. Not all events in UCDP Candidate will make it into the final UCDP GED and new events will be added. The UCDP Candidate is a dataset containing events that have a high probability of being included in the final annual version of UCDP GED.

The basic unit of analysis for the UCDP-Candidate dataset is the "event", i.e. an individual incident (phenomenon) of lethal violence occurring at a given time and place.

An event is defined the as: "*An incident where armed force was used by an organised actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date*".

UCDP Candidate can only be retrieved from the UCDP API for GED events; see http://ucdp.uu.se/apidocs/,
http://ucdpapi.pcr.uu.se/api/gedevents/<version>?<pagesize=x>&<page=x>

The current downloadable versions are:
18.0.1 to 18.0.12, 19.0.1 to 19.0.12, 20.0.1, 20.0.2, 20.0.3, the three last with data for January-March 2020, new releases will follow the same naming convention (year.0.month)

For example: http://ucdpapi.pcr.uu.se/api/gedevents/20.0.1?pagesize=1

Currently, only events for Africa are released in UCDP Candidate.

Definitions

The UCDP Candidate include data on the three core UCDP types of organized violence.

For definitions related to the three types State-based violence, Non-state violence and One-sided violence, please consult Appendix X

An event is defined as:

*An incident where **armed force** was used by an **organised actor** against **another organized actor, or against civilians, resulting in at least 1 direct death** at a **specific location** and a **specific date**.*

These are the specific elements of the definition:

1. **Armed force**: use of <u>arms</u> in order to promote the parties' general position in the conflict, resulting in deaths.
   - <u>arms</u>: any material means e.g. manufactured weapons but also sticks, stones, fire, water etc.

2. **Organized actor**: a government of an independent state, a formally organized group or an informally organized group according to UCDP criteria:

   a. **Government of an independent state**: The party controlling the capital of a state.

   b. **Formally organized group**: Any non-governmental group of people having announced a name for their group and using armed force against a government (state-based), another similarly formalized group (non-state conflict) or unorganized civilians (one-sided violence). The focus is on armed conflict involving consciously conducted and planned political campaigns rather than spontaneous violence.

   c. **Informally organized groups**: Any group without an announced name, but which uses armed force against another similarly organized group (non-state conflict), where the violent activity indicates a clear pattern of violent incidents that are connected and in which both groups use armed force against the other

3. **direct death** : a death relating to either combat between warring parties or violence against civilians.

UCDP Candidate provides three estimates for deaths for each event, thus creating an uncertainty interval:

   - a low estimate, containing the most conservative estimate of deaths that is identified in the source material;

3

- a best estimate, containing the most reliable estimate of deaths identified in the source material;

- a high estimate, containing the highest reliable estimate of deaths identified in the source material. Note that UCDP attempts to distinguish and not include unreasonable claims in the high estimate of fatalities, and tends to be highly conservative when counting fatalities[1].

In order for an event to exist, at least one dead needs to be registered in the high, best or low estimate.

4. *Specific location*: a name and one pair of latitude and longitude coordinates that relate to the geographical information specified in the source material.

The maximum (best) spatial resolution of the dataset is the individual village or town. The dataset is fully geocoded.

5. *Specific date*: a specified time period during which armed interactions cause at least 1 fatality. The normal temporal unit to which an event can be related is a 24-hour day starting at midnight.

The maximum (best) temporal resolution of the dataset is the day.

- In some cases it is impossible, based on the source material, to reduce the specific date to a single day as reporting only refers to wider time spans (multiple days) or information on the exact day is not clear. For these events, a wider time span is provided through the use of the *date_start, date_end and date_prec* variables.

Variables in the UCDP Candidate dataset

Variable name	Content	Type
id	A unique numeric id identifying each event.	integer
relid	Used as a unique sting in the dataset, the relid will change between the monthly and annual release if the event changes dyad or year.	string(255)
year	The year of the event.	integer
active_year	**1**: if the event belongs to an active conflict/dyad/actor-year where events with codestatus-clear have passed the 25 battle-related deaths threshold. **0**: otherwise	Integer
code_status	Clear Check Type of violence Check Dyad	string(50)

---

[1] For a more elaborate discussion on aspects concerning point 1-4, please refer to UCDP Codebooks for State-Based Armed Conflicts, Non-state Conflicts and One-Sided Violence.

	Check Deaths Check Incompatibility Check Vague/biased source Check Geography Check Other	
type_of_violence	Type of UCDP organized violence: **1**: state-based conflict **2**: non-state conflict **3**: one-sided violence	Integer
conflict_dset_id	This id column is only in the dataset to ensure the same structure between old and new UCDP GED outputs **please use conflict_new_id**. UCDP/PRIO conflict id for state based conflicts as per the UCDP/PRIO Armed Conflict Dataset. UCDP dyad id code for non-state dyad as per the UCDP Non-State Dataset. UCDP actor id code for the one-sided violence actor as per the UCDP One-Sided Dataset. For violence, not part of the datasets the same rule applies, for events involving actors where UCDP lacks credible information on the identity type of violence -XXX followed by the Gleditsch and Ward numeric identifier of the country where the event takes place is used. The conflict_dset_id is not unique and can clash between different types of violence.	string(255)
conflict_new_id	A unique conflict id code for each individual conflict in the dataset. Please use this variable to aggregate, subset or filter by conflict within GED.	integer
conflict_name	Name of the UCDP conflict to which the event belongs. For statebased conflicts all dyads and episodes fought over the same incompatibility has the same conflict name. For non-state conflicts and one-sided violence, this is the same as the dyad name.	string(9999)
dyad_dset_id	This id column is only in the dataset to ensure the same structure between old and new UCDP GED outputs **please use dyad_new_id**. UCDP dyad id code for state based dyad as per the UCDP Dyadic Dataset. UCDP dyad id code for non-state dyad as per the UCDP Non-State Dataset. UCDP actor id code for the one-sided violence actor as per the UCDP One-Sided Dataset. For violence, not part of the datasets the same rule applies, for events involving actors where UCDP lacks credible information on the identity type of violence -XXX followed by The Gleditsch and Ward numeric identifier of the country where the event takes place is used.	string(255)

	Note that the same dyad_id in GED can represent two separate dyads, as the dyad_id is unique only within each type of violence (i.e. a non state dyad can an identical id with a state-based dyad).	
dyad_new_id	A unique dyad id code for each individual dyad in the dataset. Please use this variable to aggregate, subset or filter by dyad within GED.	integer
dyad_name	Name of the conflict dyad creating the event. A dyad is the pair of two actors engaged in violence (in the case of one-sided violence, the perpetrator of violence and civilians). The two sides are separated by an ASCII dash (e.g. Government of Russia - Caucasus Emirate, Taleban – civilians).	string(9999)
side_a_dset_id	The unique ID of side A, for actors where UCDP lacks credible information on the identity XXX followed by the Gleditsch and Ward numeric identifier of the country where the event takes place is used.	string(255)
side_a_new_id	A unique id of side A.	integer
side_a	The name of Side A in the dyad. In state-based conflicts always a government. In one-sided violence always the perpetrating party.	string(9999)
side_b_dset_id	The unique id of side B, for actors where UCDP lacks credible information on the identity XXX followed by the Gleditsch and Ward numeric identifier of the country where the event takes place is used.	string(255)
side_b_new_id	A unique ID of side B.	integer
side_b	The name of Side B in the dyad. In intrastate armed conflicts (state-based) always the rebel group, in interstate armed conflicts (state-based) the government last in alphabetical order. In one-sided violence always "civilians". In non-state conflicts the name of the organized group last in alphabetical order (or that were last when the dyad was first entered in the data).	string(9999)
number_of_sources	Number of total sources entered, containing information for an event.	integer
source_article	References to the names, dates and titles of the source material from which information on the event is gathered.	text
source_office	The name of the organizations publishing the source materials.	text
source_date	The dates the source materials were published on.	text
source_headline	The titles of the source materials.	text
source_original	The name or type of person or organization from which the information about the event originates in the original report. This can also be the name of a news outlet. e.g. "police", "Lt. Col. Johnson", "eyewitnesses", "rebel spokesman".	string(9999)
where_prec	The precision with which the coordinates and location assigned to the event reflects the location of the actual event.	integer

	1: exact location of the event known and coded. 2: event occurred within at maximum a ca. 25 km radius around a known point. 3: only the second order administrative division where an event happened is known. That administrative division is coded with a point representing it (typically the centroid). 4: only the first order administrative division where an event happened is known. That administrative division is coded with a point representing it (typically the centroid). 5: Linear features (e.g. a long river, a border, a longer road or the line connecting two locations further afield than 25 km) or a fuzzy polygon without defined borders (informal regions, large radiuses etc.). Or an administrative region that is not an admin 1 or 2 and larger so that the point might be further than 25 km from the centroid. A representation point is chosen for the feature and employed. 6: only the country where the event took place in is known. 7: event in international waters or airspace.	
where_coordinates	Name of the location to which the event is assigned. Fully standardized and normalized.	string(9999)
where_description	An extracted snippet of text from the source material describing the location. Is often left empty if where_coordinates contain all relevant information.	text
adm_1	Name of the first order (largest) administrative division where the event took place	string(9999)
adm_2	Name of the second order administrative division where the event took place.	string(9999)
latitude	Latitude (in decimal degrees).	numeric(9,6)
longitude	Longitude (in decimal degrees).	numeric(9,6)
geom_wkt	An Open Geospatial Consortium textual representation of the location of each individual point. Formatted as OGC WKT (*well known text*) without SRID.	string(9999)
priogrid_gid	The PRIO-grid cell id (**gid**) in which the event took place. Compatibility with PRIO-grid (Tollefsen, 2012) is guaranteed for both PRIO-grid **1** and **2**. **Warning:** We associate every point to the PRIO-grid that contains it, even if the point is in another country than the one officially. assigned to the respective PRIO-grid cell through their majority area rule. It is your responsibility to make sure the covariates for the PRIO-grid cell are correct for each event. Further, for the same reason, DO NOT, under any circumstances, first clip out (subset) PRIO-grid by country before merging with UCDP GED as data loss will certainly occur. Refer to your copy of the PRIO-grid for further details on PRIO-grid's majority assignment rule (p.3).	integer
country	Name of the country in which the event takes place.	string(999)
region	Region where the event took place. One of following: Africa, Americas, Asia, Europe, Middle East	string(999)

event_clarity	1 (high) for events where the reporting allows the coder to identify the event in full. That is, events where the individual happening is described by the original source in a sufficiently detailed way as to identify individual incidents, i.e. separate activities of fighting in a single location: Example of such reporting: *"2 people where killed in Banda Aceh town on the 9th of December in fighting between the government and GAM when a car exploded in a main market."* 2 (lower) for events where an aggregation of information was already made by the source material that is impossible to undo in the coding process. Such events are described by the original source only as aggregates (totals) of multiple separate activities of fighting spanning over a longer period than a single, clearly defined day. Examples of such reporting: "The Ukrainian government informs that 29 people have died in the past six days in a number of clashes with the separatists along the line of conflict".	integer
date_prec	How precise the information is about the date of an event. **1**: exact date of event is known; **2**: the date of the event is known only within a 2-6 day range. **3**: only the week of the event is known **4**: the date of the event is known only within an 8-30 day range or only the month when the event has taken place is known **5**: the date of the event is known only within a range longer than one month but not more than one calendar year.	integer
date_start	The earliest probable date when the event has taken place.	Date YYYY-MM-DD
date_end	The last probable date when the event has taken place.	Date YYYY-MM-DD
deaths_a	The best estimate of deaths sustained by side a. Always 0 for one-sided violence events.	integer
deaths_b	The best estimate of deaths sustained by side b. Always 0 for one-sided violence events.	integer
deaths_civilians	The best estimate of dead civilians in the event. For non-state or state-based events, this is the number of collateral damage resulting in fighting between side a and side b. For one-sided violence, it is the number of civilians killed by side a.	integer
deaths_unknown	The best estimate of deaths of persons of unknown status.	integer
best	The best (most likely) estimate of total fatalities resulting from an event. It is always the sum of deaths_a, deaths_b, deaths_civilians and deaths_unknown.	integer
high	The highest reliable estimate of total fatalities.	integer
low	The lowest reliable estimate of total fatalities.	integer

8

gwnoa	The Gleditsch and Ward numeric identifier of the country for the actor on side A if a government.	integer
gwnob	The Gleditsch and Ward alphabetic identifier for the actor on side B if a government.	string(3)
	For a more detailed overview of the variables please consult: Croicu, Mihai and Ralph Sundberg, 2017, "UCDP GED Codebook version 18.1", Department of Peace and Conflict Research, Uppsala University	

Main difference between UCDP Candidate and UCDP GED

The main difference between UCDP Candidate and UCDP GED is that UCDP Candidate relaxes some of UCDPs criteria for inclusion in order to release the data early without having extra time to research an event. A code status is used to flag the events included with a higher uncertainty (see a description of the different code statuses below). This means that events are included which with current information do not fully meet the definition of an event, but which have a high likelihood of doing so with additional information.

Earlier UCDP datasets have always relied on the 25 deaths in a calendar year threshold for inclusion in its datasets. In UCDP-GED this threshold is relaxed to include years with fewer deaths, but only for dyads having reached the level of inclusion in at least one year. UCDP Candidate relaxes this criterion further and include all events that meet the definition of an event regardless of the annual threshold.

Code status

Clear is used when we know what parties are involved, that the actors are organized enough (see definition of organized actor), that there is a reported number of deaths, that we trust the source, that we have geo-coded the location, and for state-based violence that we have identified a stated incompatibility. The clear events are included in UCDP GED if the dyad has ever reached the 25-fatality threshold for inclusion.

Check Type of Violence: In many conflict areas all types of UCDP organized violence are present at the same time. The code status *check type of viole*nce is used when we know the number of deaths, but the deaths may be the result of more than one type of violence and we cannot yet disaggregate the deaths by type due to insufficient information.

Check Dyad: In some cases more than one dyad fights in an area, in cases where we know the *type of violence* the dyad can still be unclear. These events are flagged with check dyad, the event can still be attributed to the most probable dyad.

Check Deaths is used:

1. if information on deaths is likely to be revised;
2. if information on large attacks is still unclear in terms of fatalities;
3. if information is of a "summary" nature, and it is likely that the coder will have to make new calculations as more information with higher clarity becomes available.

Check Incompatibility is used when we have a new state-based dyad where we have not yet looked for an incompatibility, or the incompatibility is unclear.

Check Vague Or Biased Source is used when we have reports of violent deaths between actors or by an unclear actor but the source is known to be biased in some significant way.

Check Geography is used when an event is geocoded but the coder thinks that the geocoding can become more precise if more geographic sources are consulted.

Check is used when it is unclear if the same event already has been coded but with slightly different information.

The code status will signal the highest degree of uncertainty in cases were more than one aspect is unclear. The order set for the degree of uncertainty is:

- Check type of violence
- Check dyad
- Check incompatibility
- Check vague or biased source
- Check deaths
- Check geography
- Check

For further UCDP definitions please consult the "Definitions" section of UCDPs webpage available at http://pcr.uu.se/research/ucdp/definitions/

Data Collection Methods

The UCDP Candidate is manually curated and compiled, with automatic assistance in data retrieval, filtering, data storage and manipulation, as well as data validation. The following sources are consulted for the monthly update of UCDP Candidate:

Algeria, Tunisia, Morocco, Western Sahara

Factiva: Agence France Presse, All Africa, Associated Press Newswires, BBC Monitoring Middle East, BBC Monitoring Africa, BBC Monitoring Media, Reuters

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Terrorism Monitor, UN Mission report to Western Sahara

Angola, Republic of Congo, Zimbabwe, Uganda, Rwanda

Factiva: Agence France Presse, All Africa, BBC Monitoring Africa, Reuters

Factiva Extra search: Portugese factiva search for FLEC

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: LRA crisis tracker

Benin, Cape Verde, Equatorial Guinea, Gabon, Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Sierra Leone, Togo, Cote d Ivoire, Senegal

Factiva: Agence France Presse, All Africa, BBC Monitoring Africa, Reuters, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Botswana, Comoros, Lesotho, Madagascar, Malawi, Mauritius, Mozambique, Namibia, Swaziland, Tanzania, Zambia, South Africa

Factiva: Agence France Presse, All Africa, BBC Monitoring Africa, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Burkina Faso, Mali, Mauritania:

Factiva: Agence France Presse, All Africa, Associated Press Newswires, BBC Monitoring Africa, BBC Monitoring Middle East, Reuters, Xinhua,

Factiva Extra search: French Burkina 24

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Sahel Memo, CTC Sentinel articles, Terrorism Monitor

Burundi:

Factiva: All Sources

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: ITEKA: https://www.ligue-iteka.bi/, UN Independent Panel

Cameroon, Chad, Niger, Nigeria

Factiva: Agence France Presse, All Africa, BBC Monitoring Newsfile, BBC Monitoring Africa, Reuters, Vanguard (Nigeria)

Factiva Extra search: French Journal du Cameroun.com

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Bareta News, CTC Sentinel articles, Nigeria watch, Terrorism Monitor, Vox Peccavi

Central African Republic

Factiva: All Sources

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: UN Panel of Experts and MINUSCA reports, LRA crisis tracker, RJDH - reports in French http://rjdh.org/ (tag VIOLENCES), OCHA weekly https://reliefweb.int/updates?source=1503&country=54&format=10#content, Humanitarian Response - documents and data

Democratic Republic of the Congo:

Factiva: All Sources

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Radio Okapi, UN Group of Experts Reports, Congo Research Group, LRA crisis tracker

Egypt:

Factiva: Agence France Presse, Associated Press Newswires, BBC, Reuters, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Terrorism Monitor,

Ethiopia, Eritrea, Djibouti:

Factiva: Agence France Presse, All Africa, BBC Monitoring Africa, BBC Middle East, Reuters, Xinhua

Factiva Extra search: Extra search with "OLF or ONLF" for all sources.

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Kenya, Somalia:

Factiva: Agence France Presse, All Africa, BBC Monitoring Africa, Reuters, The Star Kenya, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Terrorism Monitor, CTC Sentinel, Secretary General's Reports (Somalia)

Libya:

Factiva: Agence France Presse, Associated Press Newswires, BBC Monitoring European, BBC Monitoring Middle East, BBC Monitoring Central Asia, BBC Monitoring Newsfile, Reuters, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: UN Secretary General's Reports, Terrorism Monitor, CTC Sentinel, UNSMIL Human Rights Report on Civilian Casualties, Eye on ISIS in Libya (http://eyeonisisinlibya.com/)

South Sudan, Sudan:

Factiva: Agence France Presse, All Africa, BBC monitoring Africa, BBC Monitoring Middle East, Reuters, Sudan Tribune, Xinhua

Extra Sources: Africa Confidential, Africa Research Bulletin, Amnesty International, Crisis Watch, Human Rights Watch, International Crisis Group

Extra Sources area specific: Africa Centre for Justice and Peace Studies (ACJPS) including South Sudan Human Rights Monitor, CTSAMM reports, LRA crisis tracker, OHCHR (UN Human Rights Council), reports from the UN-missions in the countries, SUDO (UK), Nuba Reports, Small Arms Survey

Format availability

UCDP Candidate data is only available as a web-service through a RESTful API under the DaaS paradigm. Descriptions on how to use the UCDP API are available at http://ucdp.uu.se/apidocs/

References

Hegre, Hegre, Mihai Croicu, Kristine Eck, and Stina Högbladh, *2018 Introducing the UCDP Candidate Events Dataset and the VIEWS Outcomes dataset*", Department of Peace and Conflict Research, Uppsala University. Available at: http://www.pcr.uu.se/digitalAssets/653/c_653796-l_1-k_ucdp_candidate_views_outcomes.pdfhttp://pcr.uu.se/digitalAssets/717/c_717328-l_1-k_ucdp_candidate_views_outcomes.pdf

Sundberg, Ralph, and Erik Melander, 2013, "Introducing the UCDP Georeferenced Event Dataset", *Journal of Peace Research*, vol.50, no.4, 523-532

UCDP GED Codebook version 19.1:
 http://ucdp.uu.se/downloads/ged/ged191.pdf

UCDP/PRIO Conflict Dataset Codebook :
http://ucdp.uu.se/downloads/ucdpprio/ucdp-prio-acd-191.pdf

UCDP Non-State Conflict Dataset Codebook:
http://ucdp.uu.se/downloads/nsos/ucdp-nonstate-191.pdf

UCDP One-Sided Dataset Codebook:
http://ucdp.uu.se/downloads/nsos/ucdp-onesided-191.pdf

UCDP Actor Dataset Codebook:
http://ucdp.uu.se/downloads/old/actor/ucdp-actor-dataset-22-2014.pdf

Appendix 1 UCDP definitions

Definition of State-based Armed Conflict

UCDP defines state-based armed conflict as: "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 a calendar year."

The separate elements of the definition are operationalized as follows:

(1) *Use of armed force:* use of arms, resulting in deaths.

    (1.1) *Arms:* any material means, e.g. manufactured weapons but also sticks, stones, fire, water etc.

(2) *25 deaths:* A minimum of 25 battle-related deaths per year and per dyad (see Item 3.3 in this definition) in an incompatibility.[2]

(3) *Party:* A government of a state or any opposition organization or alliance of organizations. UCDP distinguishes between primary and secondary parties. Primary parties are those that form an incompatibility by stating incompatible positions (see Item 5 in this definition). At least one of the primary parties is the government of a state.
Secondary parties are states that enter a conflict with troops to actively support one of the primary parties. The secondary party must share the position of the primary party it is supporting in the incompatibility.

    (3.1) *Government:* The party controlling the capital of a state.

    (3.2) *Opposition organization:* Any non-governmental group of people having announced a name for their group and using armed force to influence the outcome of the stated incompatibility (see Item 5 in this definition). The UCDP only deals with formally organized opposition. The focus is on armed conflict involving consciously conducted and planned political campaigns rather than spontaneous violence.

    (3.3) *Dyad:* A dyad consists of two conflicting primary parties. At least one of the primary parties must be the government of a state. In interstate conflicts, both primary parties are state governments.[3] In intrastate and extrasystemic conflicts, the non-governmental primary party includes one or more opposition organization(s). A conflict can include more than one dyad. If e.g. a government is opposed by three rebel groups over the same incompatibility, the conflict is made up of three dyads. Note that secondary parties (i.e. intervening states supplying troops to one of the primary parties) do not lead to the formation of additional dyads.

(4) *State:* A state is an internationally recognised sovereign government controlling a specific territory or an internationally unrecognised government controlling a

---

[2] Note that an incompatibility involving two opposition groups, each involved in clashes with the government resulting in 20 deaths, would not be recorded as a conflict (neither dyad reached the minimum casualty threshold), whereas an incompatibility involving a single opposition group that caused 25 battle-deaths would be included in the dataset.
[3] With three exceptions, the primary parties in interstate conflicts consist of only one actor on each side (and thus only constitute one dyad). See description of Location (Section 4.1) below for further information on these cases.

15

specified territory whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory.

(5) *Incompatibility concerning government or territory:* The incompatibility, as stated by the parties, must concern government and/or territory.

(5.1) *Incompatibility:* The stated general incompatible positions.

(5.2) *Incompatibility concerning government:* Incompatibility concerning type of political system, the replacement of the central government, or the change of its composition.

(5.3) *Incompatibility concerning territory:* Incompatibility concerning the status of a territory, e.g. the change of the state in control of a certain territory (interstate conflict), secession or autonomy (internal conflict).

As a country can experience several simultaneous conflicts, it is essential to differentiate between them, UCDP collects information on conflicts where the incompatibility, i.e. the general incompatible positions, concerns either government or territory or both. Note that the incompatibility expressed in terms of government or a specific territory is crude in the sense that possible underlying incompatibilities are not considered. In other words, the stated incompatibility is *what* the parties are (or claim to be) fighting over, but it says nothing about *why* the parties are fighting. However, classifying incompatibilities as stated manifestations of possible underlying incompatibilities or goals allows for the intersubjective compilation of the list of armed conflicts.

While a state can only experience one intrastate conflict over government in a given year, that same state can simultaneously be a primary party to one or more interstate conflicts over government and/or territory. In the case of intrastate territorial conflicts, multiple conflicts can be recorded over different territories in a state in a given year.

Definition of Non-State conflict

A non-state conflict is defined by the Uppsala Conflict Data Program (UCDP) as *"the use of armed force between two organized armed groups, neither of which is the government of a state, which results in at least 25 battle-related deaths in a year."*

The separate elements of the definition are operationalized as follows:

(1) *Use of armed force*: the use of arms, resulting in deaths.

(1.1) *Arms*: any material means, e.g. manufactured weapons but also sticks, stones, fire, water, etc.

(2) *25 deaths*: a minimum of 25 battle-related deaths per year

(2.1) *battle-related deaths:* deaths directly related to the use of armed force between the warring groups

(3) *Organized groups*: consists of either

(3.1) formally organized groups: any non-governmental group of people having announced a name for their group and using armed force against another

16

similarly formally organized group, *or*

(3.2) informally organized groups: any group without an announced name, but who uses armed force against another similarly organized group, where the violent activity meets the following requirement:

(3.2.a.) there is a clear pattern of violent incidents that are connected and in which both groups use armed force against the other

(4) *State*: a state is

(4.1) an internationally recognized sovereign government controlling a specified territory, *or*

(4.2) an internationally unrecognized government controlling a specified territory whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory.

(5) *Government*: the party controlling the capital of the state

Definition of One-sided violence

One-sided violence is the use of armed force by the government of a state or by a formally organized group against civilians which results in at least 25 deaths. Extrajudicial killings in custody are excluded.

The separate elements of the definition are operationalized as follows:

(1) *Use of armed force*: use of arms in order to exert violent force, resulting in death
(1.1) Arms: any material means, e.g. manufactured weapons but also sticks, stones,
fire, water, etc

(2) *25 deaths*: a minimum of 25 civilian deaths per year and per actor

(3) *Government*: the party controlling the capital of the state

(4) *Formally organized group*: any non-governmental group of people having announced a name for their group and using armed force

(5) *State*: a state is
(5.1) an internationally recognized sovereign government controlling a specified territory, *or*
(5.2) an internationally unrecognized government controlling a specified territory whose sovereignty is not disputed by another internationally recognized sovereign government previously controlling the same territory.

(6) *Civilians:* unarmed people who are not active members of the security forces of the state, or members of an organized armed militia or opposition group. Government

officials, such as members of parliament, governors, and councilors, are also excluded and are instead seen as representatives of the government of a state.

Note that the UCDP definition of civilian does not refer to the definition of civilian under international law.

(7) *Extrajudicial killings in custody*: when the government of a state kills a person in its custody
> (7.1) *custody*: when the person is located in a prison or another type of governmental facility

Appendix 2 Temporal Precision and Date Estimation Rules

This document specifies the qualifications for all temporal precision variable values according to the rules constructed by the UCDP for the GED. It also sets rules for interpretation of time-related expressions and estimation of events' start and end dates. The appendix presents concrete examples that guide temporal precision coding and date estimation procedures.

Estimation of Start and End Dates

1. Start and end dates of the events are set according to information in the original sources.

2. Ambiguous time-related expressions (e.g. past few days) are interpreted on the basis of the rules presented below. This ensures uniform estimation of the events' start and end dates throughout the entire GED.

3. If the source does not provide any information about the time period during which the event took place, dates are estimated for three days, counting backwards from the day of reporting and excluding the day of reporting:

   a. "24 rebel soldiers were killed";

   b. "Security forces stepped up operations against the largest insurgent group in Assam state, where a new government was set to take charge on Friday. A police spokesman said four members of the outlawed ULFA were killed in the battles";

   c. "10 bodies found buried in a mass grave in territory controlled by the ULFA rebels".

Temporal Precision 1 – Daily Precision of Time

18

1. If the exact date of an event is known the temporal precision code of 1 is applied. Such events have the same start and end dates that are precisely specified in the news sources either by dates, day names, hours or other specific temporal concepts:

    a. "14th January", "today", "yesterday", "last Tuesday" - date for specified day;

    b. "Monday night" - date for Monday;

    c. "Last night" - date for preceding day of reporting;

    d. "The other day"- date for the preceding day of reporting.

Temporal Precision 2 – Imprecise Time (2-6 days)

1. Temporal precision value of 2 should be used in those cases when start and end dates for events are of unspecified character, spanning more than one calendar day though no longer than six days, i.e. shorter than a week:

    a. "Recently", "recent attacks" - dates for 3 days preceding and not including the day of reporting;

    b. "Past/last few days" - dates for 3 days preceding and not including the day of reporting;

    c. "Around 2 July" - dates for three days, 1-3 July, with the stated date +/- one calendar day;

    d. "Over the weekend" - dates for Saturday and Sunday, if source does not include Friday in the concept of weekend and unless specific dates/days for the weekend are provided in the source;

    e. "Since the beginning of the week", "this week" - dates from Monday to the day of reporting;

    f. "Night between Sunday and Monday" - dates for 2 days;

    g. "Past 24 hours" - dates for the day of reporting and the preceding day;

    h. "Past 48 hours" - dates for the day of reporting and 2 preceding days;

    i. "Past 72 hours" - dates for the day of reporting and 2 preceding days;

    j. "Past 2 days" - dates for 2 days preceding and not including the day of reporting;

    k. "Since Thursday" – dates from Thursday until the day of reporting;

    l. "Five-day offensive" - dates for 5 days of fighting including the day of reporting;

19

m. "Continuous fighting between 13-16 February" - specified dates;

n. "Night-long battle" - dates for 2 days covering the whole night;

o. "Night of clashes" - dates for 2 days covering the whole night;

p. "Last 6 days of January" - dates for 25-30 January, including final date of month;

q. "Late last week" - dates for Friday to Sunday of the preceding week.

Temporal Precision 3 – Weekly Precision of Time

1. Temporal precision value of 3 should be used in those cases when start and end dates for events are specified to a certain week, but specific dates are not provided:

   a. "Last week" - dates for Monday-Sunday of the preceding week. Exceptions can be made if there are reasons to believe that the event took place during the week of the reporting (e.g. sometimes "a raid last week" reported on Sunday might refer to the period Monday-Saturday of the same week, then dates for Monday-Saturday of that week should be used);

   b. "Past week" - dates for 7 days including the day of the reporting, unless text indicates that past week refers to an ongoing week (starting Monday);

   c. "First week of August" - dates for August 1-7.

   d. "Week-old offensive" - dates for a week of fighting, 7 days, including the day of reporting;

Temporal Precision 4 – Monthly Precision of Time

1. Temporal precision value of 4 should be used in those cases when start and end dates for events are specified to a certain month, but specific dates are not provided:

   a. "Beginning of/early March" – March 1 to March 10/day of reporting;

   b. "Middle of March" – March 15 +/- 5 calendar days, i.e. March 10-20;

   c. "End of/late March" – March 15 to the last day of March/day of reporting;

   d. "A number of weeks", "recent weeks" - dates for 3 weeks counting backwards from the day of reporting;

   e. "Several weeks" – dates for 3 weeks;

f. "Earlier this month" – starting the 1st day of the month and ending on the day preceding the day of reporting;

g. "Last month" - dates for the month preceding the one on which the event was reported;

h. "A fortnight ago" - dates for preceding 14 days including the day of reporting.

Temporal Precision 5 – Annual Precision of Time

1. Temporal precision value of 5 should be used in those cases when start and end dates for events are specified to a certain year, but specific dates are not provided:

    a. "1995" - 1995-01-01 to 1995-12-31;

    b. "Last year" - dates covering the year, YYYY-01-01 to YYYY-12-31;

    c. "Past year" – All dates from the date of reporting back to YYYY-01-01

    d. "Early 1999" – 1999-01-01 to 1999-04-30;

    e. "Mid 1999" – 1999-05-01 to 1999-08-31;

    f. "Late 1999" – 1999-09-01 to 1999-12-31;

    g. "Past 3 months" - dates for 3 months counting backwards from the day of reporting (may not cross over into another calendar year);

    h. "Past few months" – dates for 3 months counting backwards from the date of reporting (may not cross over into another calendar year).

Appendix 3 Geo-precision Coding Rules

This document gives an overview of the coding rules for geo-precision codes coupled with examples and comments.

General rules

1. All geographical locations are coded with moderation with preference given to more certain locations even if they represent a higher level of aggregation over those locations which are less certain but represent a lower level of aggregation.

2. Unclear geographical references with several possible levels of aggregation are coded as the highest possible one. For instance, if there is a town, a district (ADM2) and a province (ADM1) of the same name and the source does not specify to which type of location it refers, then the location will be coded as ADM1.

3. If event location (camp, bridge, road etc.) has the same name as a certain suburb, town or village (e.g. Uppsala IDP camp and Uppsala town), the coordinates for that town or village should be used only if it is known that the event location is within or close to (within 25 km) that town or village. If information about the locations' proximity to that town or village is not available, the location is aggregated to the lowest available administrative division. For instance, if it is not known that Uppsala IDP camp is within 25 km from Uppsala town, coordinates for Uppsala municipality (ADM2) should be used.

4. If the source refers to a certain location (e.g. river, forest, lake, park, mountains etc.) that is not similar in size with a locality, or that is not a point, a representation point is created with precision 5. If that location lies within an ADM2 or ADM1, the ADM2 or ADM1 is attached to the representation point. Do not aggregate e.g. rivers or national parks to administrative divisions if representation points can be made.

5. When coding historical observations the GED uses the names of the administrative divisions in force at the time of the reporting. If the boundaries of ADM1 have changed over time in a country, the dataset uses estimated coordinates for older provinces based on the relevant seat of the ADM1 at the time of the event.

   A history of administrative changes is tracked internally by the UCDP system in a data structure referred to as a geotree. If you require access to such files, contact us.

22

Geo-precision 1

Geo-precision value of 1 is used if the location information corresponds exactly to the geographical coordinates available. Each pair of coordinates is also coupled with names for ADM1 and ADM2 when available.

1. "City", "town", "village" , "location", "locality"- centroid point coordinates;

2. "District", "quarter", "neighbourhood", "locality" (of town) - coordinates for town centroid point are applied here, and not the specific section of it, though the name and details are kept in text in parenthesis in "Where";

3. Air battles if location is clear, i.e. "a plane was shot down over Kitgum".

Geo-precision 2

If the location information refers to a limited area around a specified location, coordinates for that location together with the geo-precision value of 2 are used.

1. "Near/in the vicinity of/adjacent to/just outside/around Kitgum town" – coordinates for Kitgum town;

2. "Pietermaritzburg area" – coordinates for Pietermaritzburg town;

3. "Outskirts/suburbs of Bujumbura city" – since outskirts and suburbs are understood as relatively independent and distant entities coordinates for Bujumbura city should be used;

4. "17 km from Uppsala town" – if the event takes place within a distance of 25 km from a specified location, coordinates for that specified location are used;

5. "North of Luanda city", "southeast of Y mountain" - unspecified distances from a specified location are understood to be near the stated location;

6. "Bujumbura city towards Gishingano village" – if coordinates for Gishingano village can not be retrieved, then coordinates for Bujumbura city will be used;

7. "Niuland village near Dimapur town" - if coordinates for Niuland village are not available, but coordinates for Dimapur town exist, the latter are used;

8. "Dungu territory in DRC" – third level administrative divisions (ADM3), if small enough to have an approximate radius of 25 km or less, receive a precision code of 2.

Geo-precision 3

If the source refers to or can be specified to a larger location at the level of second order administrative divisions (ADM2), such as district or municipality, the GED uses centroid point coordinates for that ADM2. If these are not available, representation coordinates

for a town within that ADM2 are used. The name of the ADM2 in force at the time of reporting is recorded in the variable ADM2.

1. "Arusha district, Arusha province" - coordinates for Arusha district (ADM2);

2. "Burambi commune, Burundi" – coordinates for Burambi commune (ADM2);

3. Air battles if unclear location - if the battle takes place "over" a certain ADM2, coordinates for that ADM2 will be used;

Geo-precision 4

If the location information refers to a first order administrative division, such as a province (ADM1), the GED uses the coordinates for the centroid point of ADM1.

1. "Cibitoke province, Burundi" – coordinates for Cibitoke province (ADM1);

2. Air battles if unclear location - if the battle takes place "over" a certain ADM1, coordinates for that ADM1 are used;

3. If the ADM2 in which the event took place in unclear (e.g. different sources refer to different ADM2s in which the same event took place), the location is aggregated to the ADM1 level;

Geo-precision 5

Geo-precision value of 5 is used in these cases:

1. If the location information refers to parts of a country which are larger than ADM1, but smaller than the entire country such as "Southern Lebanon", "Northern Uganda". In these cases, a representation point is created for that part of the country and used as a representation of that area together with geo-precision value of 5. Note that these points are stored and reused consistently by the UCDP (thus, all events assigned to "Northern DR Congo" will have the same coordinates recorded).

2. If a pair of coordinates is estimated as a representation point for a linear, non-administrative polygon or fuzzy geographic feature (river, informal area, large lake etc.). For example, if the location is on the border between two countries and the location of such point is not precisely known, a pair of estimated coordinates will be used together with geo-precision value of 5. For example, "on the border between Uganda and Sudan" will be coded as "Uganda/Sudan border" with the coordinates for a selected point on the border between Uganda and Sudan; Note that these points are stored and reused consistently by the UCDP (thus, all events assigned to "Uganda/Sudan border" will have the same coordinates recorded).

3. If the location information refers to islands which are not an ADM1 or 2 of their own. For example, "Zanzibar island" will be understood as eastern part of Tanzania and receive geo-precision value of 5. If a pair of coordinates for that island is not available in the gazetteers, it can be represented by an ADM1 in that island.

24

4. If the location is not specific and need to be estimated (for example, "road between Pader and Kitgum", "along Aswa river" etc.), or the location is more than 25 km away from another location (for example, 75 km south of Kitgum town), then a representation point is created for that point. This is done even if the two points are located in the same ADM2 As such, if an event is described as taking place on "the road between Yei and Rasul in Yei district of Equatoria State", then a point is estimated on that road, with precision 5, with both the ADM1 (Equatoria state) and the ADM2 (Yei district) coded.

Geo-precision 6

If the location information refers to an entire country, centroid point coordinates of that country are used. Also, if the location is not provided/is unclear/refers to several locations which can not be split and covers the whole country and a particular activity area of the actor is not clear, centroid point coordinates of that country are used.

1. "Germany" - centroid point coordinates;

Geo-precision 7

If the event takes place over water or in international airspace, the geographical coordinates in the dataset either represent the centroid point of a certain water area or estimated coordinates according to similar techniques as presented above for geo-precision code 5.

For air events, precision code 7 is used only if the death is not the effect of or did not result in the airplane crashing (in such a case, 1-5 precision codes are used with the location of the crash).

1. "Southern ocean" – centroid point coordinates;

2. "Bay of Bengal" – centroid point coordinates;

3. "37 km off the coast from Stockholm city" – estimated coordinates for a point 37 km and 90 degrees off the coast of Stockholm.

4. "the minister was stabbed on an airplane en route to Delhi after departing Islamabad" – coordinates for Islamabad airport, precision code 7.