Social Conflict in Africa Database
Version 3.0

Codebook and coding procedures

Idean Salehyan
University of North Texas

&

Cullen Hendrix
College of William and Mary

Updated March 29, 2012

*(Before using SCAD 3.0, please read below about "geographic information.")

The Social Conflict in Africa Database (SCAD) contains information on protests, riots, strikes, and other social disturbances in Africa. Whereas conflict data is generally available for large-scale events such as civil and international war, the purpose of this dataset is to compile information on other types of social and political disorder. See www.scaddata.org for further information about the project.

Every country in Africa (with a population greater than 1 million) is covered, beginning in 1990. SCAD therefore includes all North African states in addition to Sub-Saharan Africa. The primary source of information for this dataset comes from searches of Associated Press and *Agence France Presse* newswires, as compiled bythe Lexis-Nexis news service (http://www.lexisnexis.com/).

Search methodology

This section describes our search procedure in Lexis-Nexis. The research staff conducted news searches and coded variables data using the following steps. Users wanting more information about particular events in the database can replicate the search methodology to find the original story.

Search Procedure

1. Using the power search interface, use the HLEAD option to search for the country name in the headline and lead paragraphs.
2. Use Boolean options to select additional terms. After selecting the country, search for AND "protest" OR "strike" OR "riot" OR "violence" OR "attack" as additional terms anywhere in the document.
3. In the "sources" field, choose, "Wire Service Stories".
4. Within the list of wire services, choose "Agence France Presse– English", "The Associated Press", "Associated Press Online", and "Associated Press Worldstream".
5. Select appropriate date ranges.

Sorting procedure

1. Begin with the oldest listed story and proceed chronologically. For each story, determine if this information is relevant to the data project. If not, discard the story.

2. During periods of civil conflict, defined by the start and end dates in the *Uppsala Armed Conflict Database* (see below for further details), do *not* code violent events associated with the civil conflicts. If Actor1, Actor2, or Actor3 is the government, a quasi-governmental organization (i.e., paramilitary organization), or the rebel organization, and these acts are directly linked to the civil conflict, the event should not be coded.
3. Sometimes, many articles will cover a single news story or event. Group these articles together as a single event. Avoid double or triple counting a single event if there are multiple articles chronicling a story. However, use the most recently cited and/or most common estimates for disputed information, such as the number of deaths.
4. Sometimes, a single article will cover multiple events. Determine if the main actor(s), target(s), and issue(s) are different, and if so, code these as distinct events.
5. Code a new event if the actor(s), target(s), and issue(s) are different or if there is substantial time that elapses between one event and another. If the actors, targets, and issues are the same, but there is a brief lull in activity, code it as the same event, so long as there is significant "momentum" linking the time periods.

Coding procedure

This section describes the particular variables in the database.

Identification scheme
The "country code" column provides the Correlates of War (http://www.correlatesofwar.org/) numeric designation of the country. The "id" column lists the event number. These fields are combined to generate a single, unique "event id" code.

Start Date
Lists the day, month, and year in DD/MM/YYY format for when the event begins. If the exact day cannot be identified, this is the best approximation of the start date. The start year (*styr*), start month (*stmo*), start day (*stday*) are listed in separate columns.

End Date
Lists the day, month, and year in DD/MM/YYYY format for when the event ends. If the exact day cannot be identified, this is the best approximation of the end date. The end year (*eyr*), end month (*emo*), end day (*eday*) are listed in separate columns.

Duration
The number of days the event lasted.

Etype(categorical)
Indicates the type of event according to the following coding scheme:

1 = **Organized Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct "other" group or government authorities. In this event, clear leadership or organization(s) <u>can</u> be identified.

2 = **Spontaneous Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct "other" group or government authorities. In this event, clear leadership or organization <u>cannot</u> be identified.

3 = **Organized Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct "other" group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) <u>can</u> be identified.

4 = **Spontaneous Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct "other" group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) <u>cannot</u> be identified.

5 = **General Strike.** Members of an organization or union engage in a total abandonment of workplaces and public facilities.

6 = **Limited Strike.** Members of an organization or union engage in the abandonment of workplaces in limited sectors or industries.

7 = **Pro-Government Violence (Repression)**: Distinct violent event waged primarily by government authorities, or by groups acting in explicit support of government authority, targeting individual, or "collective individual," members of an alleged opposition group or movement. Note that this event is initiated by the government or pro-government actors. See code for repression, below.

8 = **Anti-Government Violence**: Distinct violent event waged primarily by a non-state group against government authorities or symbols of government authorities (e.g., transportation or other infrastructures). As distinguished from riots, the anti-government actor must have a semi-permanent or permanent militant wing or organization.

9 = **Extra-government Violence**: Distinct violent event waged primarily by a non-state group targeting individual, or "collective individual," members of an alleged oppositional group or movement. As distinguished from riots, at least one actor must have a semi-permanent or permanent militant wing or organization. Government authorities are not listed as actors or targets.

10 = **Intra-government Violence:** Distinct violent event between two armed factions associated with different elements within the government. These include violence between two legally constituted armed units (e.g. clashes between police and military) or between unofficial militias associated with particular governmental leaders. This code includes events such as military coups.

Escalation (categorical)
Did the nature of the event change during its duration?

0 = **No Escalation.**

1 = **Organized Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct "other" group or government authorities. In this event, clear leadership or organization(s) <u>can</u> be identified.

2 = **Spontaneous Demonstration.** Distinct, continuous, and largely peaceful action directed toward members of a distinct "other" group or government authorities. In this event, clear leadership or organization <u>cannot</u> be identified.

3 = **Organized Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct "other" group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) <u>can</u> be identified.

4 = **Spontaneous Violent Riot.** Distinct, continuous and violent action directed toward members of a distinct "other" group or government authorities. The participants intend to cause physical injury and/or property damage. In this event, clear leadership or organization(s) <u>cannot</u> be identified.

5 = **General Strike.** Members of an organization or union engage in a total abandonment of workplaces and public facilities.

6 = **Limited Strike.** Members of an organization or union engage in the abandonment of workplaces in limited sectors or industries.

7 = **Pro-Government Violence (Repression)**: Distinct violent event waged primarily by government authorities, or by groups acting in explicit support of government authority, targeting individual, or "collective individual," members of an alleged opposition group or movement. Note that this event is initiated by the government or pro-government actors. See code for repression, below.

8 = **Anti-Government Violence**: Distinct violent event waged primarily by a non-state group against government authorities or symbols of government authorities (e.g., transportation or other

infrastructures). As distinguished from riots, the anti-government actor must have a semi-permanent or permanent militant wing or organization.

        9 = **Extra-government Violence**: Distinct violent event waged primarily by a non-state group targeting individual, or "collective individual," members of an alleged oppositional group or movement. As distinguished from riots, at least one actor must have a semi-permanent or permanent militant wing or organization. Government authorities are not listed as actors or targets.

        10 = **Intra-government Violence:** Distinct violent event between two armed factions associated with different elements within the government. These include violence between two legally constituted armed units (e.g. clashes between police and military) or between unofficial militias associated with particular governmental leaders. This code includes events such as military coups.

Actor 1 (text)
Records the social, political, or identity group (i.e. actor) directly involved in the event.

Actor 2 (text)
Records the social, political, or identity group (i.e. actor) directly involved in the event.

Actor 3 (text)
Records the social, political, or identity group (i.e. actor) directly involved in the event.

Target 1 (text)
Records the social, political, or identity group directly targeted by the event.

Target 2 (text)
Records the social, political, or identity group directly targeted by the event.

Cgovtarget(dichotomous)
Was the central government the target of the event?
Yes = 1, No=0

Rgovtarget(dichotomous)
Was a regional, provincial or local government the target of the event?
Yes = 1, No = 0

Npart(categorical)
Total number of participants in the event.
        1 = less than 10
        2 = 10 to 100
        3 = 101 to 1,000
        4 = 1,001 to 10,000
        5 = 10,001 to 100,000
        6 = 100,001 to 1,000,000
        7 = over 1,000,000
        -99 = unknown

Ndeath(cardinal numbers)
Records the best estimate of the number of persons killed in the event. If multiple estimates are given, use the mean number of reported deaths or the most recent figures. If the exact number is not given, use the following codes:
        -99 = unknown
        -88 = unknown but probably small (less than 10)

-77 = unknown but probably large (10 or more)

Repress (categorical)
Did the government use repression or violence against participants in the event?
       0 = no repression used
       1 = non-lethal repression used (e.g. tear gas, arrests, etc)
       2 = lethal repression used (there must be deaths reported)

Elocal(text)
Identifies the name of the locality where the event occurred.
[See below for more information on geocoding SCAD]

Locnum(categorical)
Coding of the event locality
       1 = Capital city
       2 = Other major urban area (population greater than 100,000)
       3 = Rural (including small towns, villages with population less than 100,000)
       4 = Multiple urban areas
       5 = Multiple rural areas
       6 = Province/region listed, exact location unknown
       7 = Nationwide. Effects several cities and rural areas
    -99 = location unknown

Issue 1 (categorical)
What was the first issue that was mentioned as the source of the tension/disorder?
       1 = elections
       2 = economy, jobs
       3 = food, water, subsistence
       4 = environmental degradation
       5 = ethnic discrimination, ethnic issues
       6 = religious discrimination, religious issues
       7 = education
       8 = foreign affairs/relations
       9 = domestic war, violence, terrorism
    10 = human rights, democracy
    11 = pro-government
    12 = economic resources/assets
    13 = other
    14 = unknown, not-specified

Issue 2 (categorical)
What was the second issue, if any, that was mentioned as the source of the tension/disorder
       1 = elections
       2 = economy, jobs
       3 = food, water, subsistence
       4 = environmental degradation
       5 = ethnic discrimination, ethnic issues
       6 = religious discrimination, religious issues
       7 = education
       8 = foreign affairs/relations
       9 = domestic war, violence, terrorism

10 = human rights, democracy
11 = pro-government
12 = economic resources/assets
13 = other
14 = unknown, not-specified

Issue 3 (categorical)
What was the third issue, if any, that was mentioned as the source of the tension/disorder
    1 = elections
    2 = economy, jobs
    3 = food, water, subsistence
    4 = environmental degradation
    5 = ethnic discrimination, ethnic issues
    6 = religious discrimination, religious issues
    7 = education
    8 = foreign affairs/relations
    9 = domestic war, violence, terrorism
10 = human rights, democracy
11 = pro-government
12 = economic resources/assets
13 = other
14 = unknown, not-specified

Issuenote(text)
Include a very brief description of the event.

Nsource(categorical)
Did more than one news article give information on the event?
1 = yes, 0 = no

Notes (text)
This provides additional information about the event. It lists any irregularities in the case, difficulties in coding, discrepancies between sources, etc.

Relationship with the Uppsala University Armed Conflict Database

Some events listed in the database serve as "placeholder rows" for armed conflicts, as listed by the Uppsala University Armed Conflicts Database (ACD). For these events, a negative event id code is given to distinguish them from SCAD events, indicating that our researchers did not compile this information. The start and end dates of ACD conflicts are given, along with brief summary information, using the coding classifications given above. For these events, the etype is given as -9, again to distinguish them from SCAD event types. For most purposes, users will find it acceptable to treat ACD armed conflicts as its own category of events within the SCAD framework. However, in the "notes" field, the ACD conflict id is provided as an aid to researchers who would like to merge the datasets or find additional information.

SCAD researchers excluded events that are clearly part of the civil conflict dynamic, such as when specific rebel actors are named. In some cases, however, it was difficult to determine whether or not the

event pertained to the ACD conflict. The column for "ACD questionable" assigns these ambiguous cases a code of 1.

Geographic Information – Important!

For the SCAD 3.0 update, we provide geographic information on event locations in the SCAD dataset. This version of the data is organized differently from previous versions of SCAD. Previously, each <u>event</u> was a unique case or "row" in the dataset. In the 3.0 version, each row corresponds to a particular event <u>location</u>. Therefore, an event which takes place in multiple locations is given the same event ID for each row, but individual locations are listed separately.

For example, event 6150020 is a riot in Algeria which takes place in three cities: Algiers, Constantine and Batna. The *elocal* column (event location) indicates the three cities in aggregate form, "Algiers, Constantine, and Batna." However, note that in SCAD 3.0 event 6150020 is assigned to additional rows, one for each city in which the riot takes place. All three locations have the same event id number, indicating that they are part of a single SCAD event. The *ilocal* column (individual location) separately lists each location. The *sublocal* column provides a numeric count of the individual locality, 1st location, 2nd location, 3rd location, etc.

Users not interested in longitude and latitude coordinates or additional geographic information are <u>advised to drop or filter out</u> all rows where *sublocal* does not equal 1 (i.e. only take the first entry), and ignore information in the *ilocal*, *longitude* and *latitude* columns. Failure to do so will result in over-counting the same event!

> It is important to note that all other corresponding information pertains to the event <u>as a whole</u> and not for that individual location. All other informational variables (e.g. deaths, repression, magnitude) are identically duplicated for each location and may or may not represent what occurred in that particular town or village. For instance, in the Algeria example above (6150020), the event as a whole killed two people; it is not the case that two people died in each of the three cities. Therefore, in calculating statistics such as the number of deaths or the number of participants, users should take care to use unique event ids rather than individual event locations. Doing otherwise would artificially create multiple counts of the same information.

Longitude and Latitude

GeoSCAD estimates, with as much precision possible, the *latitude* and *longitude* coordinates for each event location. We consulted Yahoo! Maps, www.geonames.org, www.fallingrain.com, and www.itouchmap.com/latlong to pinpoint event locations.

For major cities and towns, identifying the X/Y coordinates was relatively straightforward. However, smaller towns, villages, refugee camps, etc, sometimes did not appear in gazetteers

and required additional location cues—e.g. "10km south of Bujumbura"—and the given coordinates are rough approximations. At other times, the news report did not mention a specific location, but rather, the district or region was given (these correspond to *locnum* = 6 in the SCAD file). In such cases, we placed the point at the approximate center of the region. Point locations that are estimated rather than precise are flagged in the *location_precision* column. Care should be taken when performing spatial analyses of such cases.

For several events, the location is listed as being "nationwide" or "unknown" (locnum 7, -99). These events are assigned at or near the centroid, or geographic center point, of the country. For the sake of clarity, for location "unknown" events, the point was offset from "nationwide" by subtracting 0.5 from the longitude and latitude, with the exception of very small countries where smaller increments were used.

When further clarifications are needed, additional information about the location and geocoding procure are noted in the *geo_comments* column.

Finally, a very small share of events could not be georeferenced because the location could not be found or the description was too vague. Users may choose to drop such cases or assign them to points as they see fit.

Appendix 1. Corrections to the SCAD 2.0 data.

In addition to including geographic information, and updating SCAD through 2011, several corrections were made to the previous, 2.0 version of SCAD. These changes are documented below.

(1) One event was miscoded as event type 3. It has been changed to event type 9.
4040016

(2) The following events were erroneously coded as escalating to pro-government violence. Rather, they should have been coded as being repressed by government forces.
4390001
4520021
5530087
4350016
5600173
5600275

(3) This event involved Algerians in a protest which took place in France rather than Algeria itself. As events must take place within the geographic territory of the country in question, the event was deleted.
6150011

(4) The following events involved individual battles during the Liberian civil war. According to the SCAD criteria, such events are subsumed within the larger conflict and do not normally enter the database. Therefore, the following cases were deleted.
  4500092
  4500093
  4500094
  4500095
  4500096
  4500097
  4500098
  4500099
  4500100
  4500101
  4500102
  4500103
  4500104
  4500105

(5) This event was erroneously coded as occurring in Bundibugyo district, which is in Uganda, not the DRC. This event actually occurs in the Upper Nile Region of the DRC.
4900347

(6) This event was erroneously coded as occurring in "Azikiwe". It occurs in Onitsha.
4750693

(7) This event was erroneously coded as occurring in Jema'a. It occurs in Kanfanchan.
4750212

(8) In this event, Kinyandoni refugee camp is not in Rwanda. The event is reassigned to the DRC.
5170082
The event has been given a new event id, reflecting its location in the DRC.
4900411

(9) This event does not occur in Rwanda but in the Democratic Republic of the Congo. The event is reassigned to the DRC.
5170077
The event has been given a new event id, reflecting its location in the DRC.
4900412

(10) The following cases reflect a misspelling of the location name and have been changed.

Ethiopia: Harare →Harar
DRC: Kinangani→ Kisangani.
DRC: Kigalama→Kigarama.
Central African Republic: Gimari→Grimari
Central African Republic: Obob→ Obo.
Chad: Adeche→Abeche.
Ethiopia: Addis Ababs→ Addis Ababa
Ethiopia: Amob→ Ambo
Ethiopia: Eastern Ethiopis→ Eastern Ethiopia
Ghana: Katsumi → Kumasi
Kenya: Vihiga "distric→ district
Kenya: Saburu district, Sambuu village → both are "Samburu"
Nigeria: Abai→Abia
Nigeria: Abasa→Asaba
Nigeria: Damoa→Damoba
Nigeria: Nassarawa→Nasarawa
Sierra Leone: Kanema→Kenema
Somalia: Heraleh→Herale
Somalia: Qoryloley and Qoryloley District→ "Qoryoley"
South Africa: Kathlelong→Kathlehong
South Africa: Mabalni→Mabaleni
South Africa: Primose gold mine → Primrose gold mine
South Africa: Sharpsville → Sharpeville
Sudan: Malalka→Malakal
Togo :Gerin, Kouka→Gerin-Kouka
DRC: Kilanga→Kalanga

(11) Beginning in 2009, the Uppsala Armed Conflict Program lists a new civil war involving BokoHaram militants in Nigeria. Accordingly, we have removed earlier events involving BokoHaram but include a new row indicating an ongoing civil war.