

Codebook for the Militarized Interstate Dispute Location (MIDLOC-I) Dataset, v2.0

Paul Bezerra

National Security Affairs Department

U.S. Naval War College

paul.bezerra@usnwc.edu

Alex Braithwaite

School of Government & Public Policy

University of Arizona

abraith@email.arizona.edu

July 19, 2018

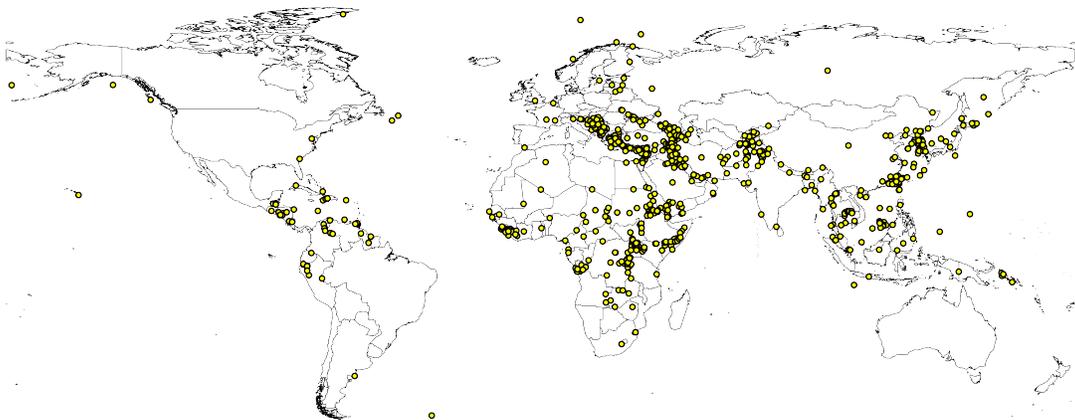


Figure 1: MID **Incidents**, 1993 - 2010

Overview

The Militarized Interstate Dispute Location (MIDLOC) dataset details the precise geographic location of Militarized Interstate Disputes (MIDs) in the post-Napoleonic era (from 1816 - 2010), complementing the Correlates of War (COW) Project's considerable empirical data collection. The MIDLOC-I v2.0 dataset offers point locations to represent each MID **incident** between 1993 and 2010. Following common convention, latitude and longitude point coordinates are recorded for each observation. Latitude is measured as the angular distance, in decimalized degrees (DD), of a point north or south of the equator. Similarly, longitude is measured as the angular distance, in decimalized degrees (DD), of a point east or west of the Prime (Greenwich) Meridian.

There are a total of 3,317 individual records in MIDLOC-I v2.0. This is comprised of recorded locations for 3,141 individual MID incidents from a combination of the MIDI v4.01 and the MIDA v4.2 (Palmer et al. 2015) datasets - approximately 95% data coverage.¹ 2,093 of these 3,317 incidents took place between 1993 and 2001 and 1,224 between 2002 and 2010. A total of just 176 MID incidents remain without a recorded location. **In addition to the MID-Incident data detailed, a separate and complementary dataset (MIDLOC-A v2.0) includes onset location data for the period 1816 - 2010.**

Updates since Version 1.1

MIDLOC v2.0 represents a significant update from v1.1 (Braithwaite 2010). In summary, the significant changes include: coding original geographic coordinates for MID v4.2 disputes between 2002 - 2010; revisiting and updating geographic coordinates for all disputes included in MIDLOC v1.1; and applying a new precision coding scheme with greater subnational nuance based upon the UCDP and AidData codebook on georeferencing aid (Strandow et al. 2011) across all disputes. We do not discuss each of these individual changes here in this document. However, we do report original codings as additional variables / columns in the dataset with the MIDLOC11_preamble in their names. Users can therefore identify changes to observations by comparing variables with MIDLOC2 and MIDLOC11 preambles.

As an overview of these changes, we offer a comparison of MIDLOC v1.1 and MIDLOC-I v2.0 precision codes for all incidents. This comparison offers the number of each precision code during their shared time frame, 1993 - 2001, as well as the 2002 - 2010 time frame for MIDLOC-I v2.0 only.²

Table 1: Precision Code Comparison

Code	MIDLOC v1.1	MIDLOC-I v2.0
1	Known point location	The coordinates correspond to an exact location, such as a populated place or a hill.
	1993 - 2001: 475	1993 - 2001: 312 2002 - 2010: 31

¹ In May 2018, COW released updated dispute-level data. We have used this data (MIDA v4.2) to adapt our population of MID incidents from MIDI v4.01. As a result, MIDLOC-I v2.0 includes only incidents for disputes, which are themselves included in MIDA v4.2. MIDLOC will continue to be updated as the larger MID “universe” is itself updated.

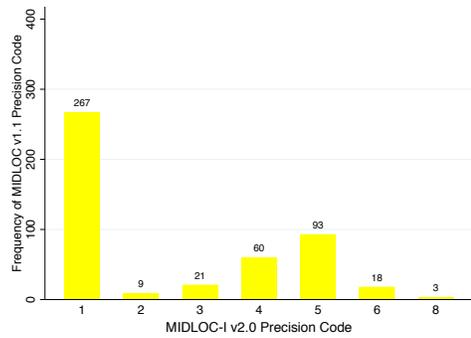
² MID v4.2 introduces a noticeable update to the population of disputes relative to MID v3.1 (Ghosn, Palmer, & Bremer 2004). Since MID v3.1 forms the basis of MIDLOC v1.1 and MID v4.2 forms the basis of MIDLOC-I v2.0, the number of disputes reflected in the two MIDLOCs between 1816 - 2001 are not equal.

2	Point location of proximity	The location is mentioned in the source as being “near” or “in the area” of an exact location. The coordinates refer to that adjacent, exact, location.
	1993 - 2001: 147	1993 - 2001: 237 2002 - 2010: 33
3	Point representation of sub-national unit/polygon	The location is, or is analogous to, a second-order administrative division (ADM2), such as a district, municipality, or commune.
	1993 - 2001: 619	1993 - 2001: 157 2002 - 2010: 34
4	Point along a line, such as a border or river or road	The location is, or is analogous to, a first-order administrative division (ADM1), such as a province, state, or governorate.
	1993 - 2001: 494	1993 - 2001: 186 2002 - 2010: 30
5	Point representation of a large national polygon	The location can only be related to estimated coordinates, such as when a location lies between populated places; along rivers, roads, and borders; or when sources refer to parts of a country greater than an ADM1, such as a national park, which spans across several provinces.
	1993 - 2001: 206	1993 - 2001: 897 2002 - 2010: 1,022
6	(None)	The location can only be related to an independent political entity. In this case, the coordinates represent the geographic centroid of the entity.
		1993 - 2001: 106 2002 - 2010: 70

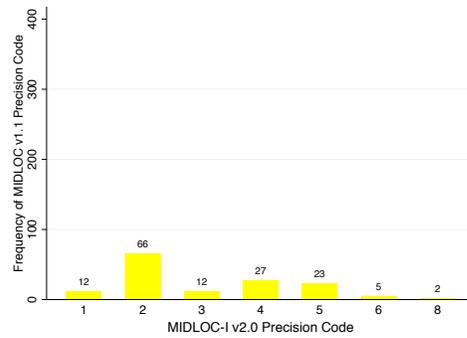
7	Unknown location 1993 - 2001: 200	(None)
8	(None)	The location is estimated to be the seat of an administrative division (local capital) or the national capital. 1993 - 2001: 22 2002 - 2010: 4
-99	(None)	No known location/location missing. 1993 - 2001: 176 2002 - 2010: 0

From this table, we would highlight three points of emphasis. First, in recoding the location text and measuring points from MIDLOC v1.1, we attempted to follow the more conservative approach suggested by the UCDP and AidData georeferencing codebook. This does result in 204 formerly precision 1 incidents being recoded to precision codes less precise than precision 1. But, second, the new precision scheme offers greater subnational nuance as it features three separate precision codes for varying subnational levels: ADM2 (3); ADM1 (4); and areas > ADM1 (5). Third, the process of recoding location text and measuring points in MIDLOC-I v2.0 greatly benefited from a variety of more recent sources, including cShapes (Weidmann, Kuse, & Gleditsch 2010), and the Global Administrative Areas (GADM; Hijmans et al. 2012) dataset, among others.

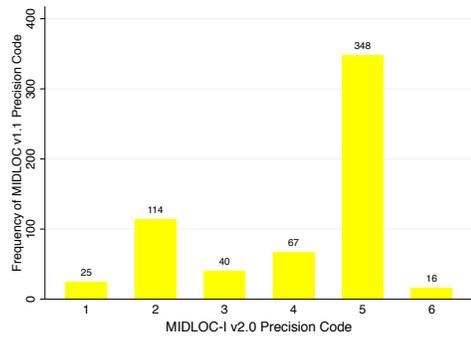
As suggested by the comparison of MIDLOC v1.1 and MIDLOC-I v2.0 precision codes, some movement between precision codes has occurred across the dataset. Figure 2 provides visual documentation of this movement by MIDLOC v1.1 precision code for the period 1993 - 2001. In general, many precision codes remain the same or similar, and the largest volume of changes are attributable to the increased subnational precision available in MIDLOC-I v2.0. Thus, we see a number of former precision 3s become precision 4s and 5s, and we, likewise, see a number of precision 4s become 5s and 3s.



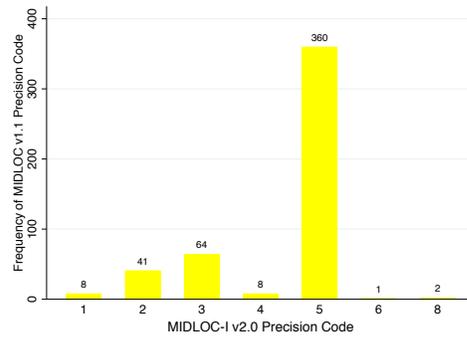
(a) Precision 1



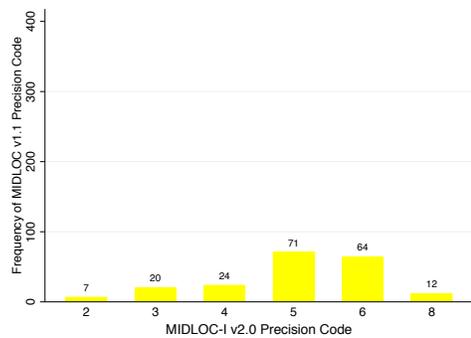
(b) Precision 2



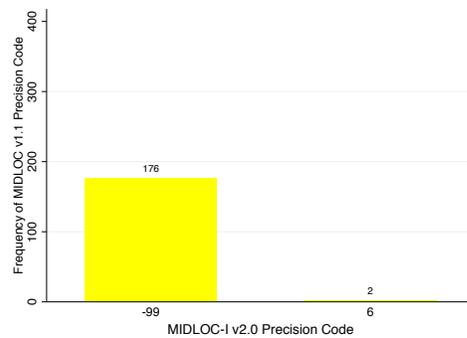
(c) Precision 3



(d) Precision 4



(e) Precision 5



(f) Precision 7

Figure 2: Redistribution of MIDLOC v1.1 Precision Codes to MIDLOC-I v2.0 Precision Codes

Citation

In any papers or publications that utilize this dataset, users are asked to give the version number and cite the articles of record for the dataset, as follows:

Braithwaite, A. 2010. MIDLOC: Introducing the Militarized Interstate Dispute Location dataset. *Journal of Peace Research*, 47(1), 91-98.

Bezerra, P., Braithwaite, A. 2018. Codebook for the Militarized Interstate Dispute Location (MIDLOC-A/I) Dataset, v2.0.

Data

The data are distributed in a flat text, comma-separated-variable (.csv) file. By default under Microsoft Windows, .csv files will open in Microsoft Excel. But the file may be opened in any program that reads text. In the statistical analysis program Stata, the file may be easily loaded using the command “insheet using MIDLOCI_2.0.csv”.

Variables

Table 2: Variables in MIDLOC-I v2.0.csv

Variable Name	Variable Description
Year	Start year of dispute
DispNum	MID Dispute Number
IncidNum	MID Incident Number
MIDLOC2_Location	Text comment on location of dispute or incident provided by coders of MID 2.1/MID 3.0/MID 4.0 projects
MIDLOC2_measuringpoint	Most precise known location from which MIDLOC observation is coded
MIDLOC2_XLongitude	Longitude in decimalized degrees
MIDLOC2_YLatitude	Latitude in decimalized degrees
MIDLOC2_precision	Indicator of geographic precision/resolution of coded location

For the basis of precision code scheme, see: Strandow et al. (2011) 1 = The coordinates correspond to an exact location, such as a populated place or a hill.

2 = The location is mentioned in the source as being “near” or “in the area” of an exact location. The coordinates refer to that adjacent, exact, location.

3 = The location is, or is analogous to, a second-order administrative division (ADM2), such as a district, municipality, or commune.

4 = The location is, or is analogous to, a first-order administrative division (ADM1), such as a province, state, or governorate.

5 = The location can only be related to estimated coordinates, such as when a location lies between populated places; along rivers, roads, and borders; or when sources refer to parts of a country greater than an ADM1, such as a national park, which spans across several provinces.

6 = The location can only be related to an independent political entity. In this case, the coordinates represent the geographic centroid of the entity.

8 = The location is estimated to be the seat of an administrative division (local capital) or the national capital.

-99 = No known location/location missing.

Onset	0 = Incident occurred subsequent to onset of MID
	1 = First known (i.e., onset) incident of MID
MIDLOC2_HowObtained	Text comment providing the source(s) of the event’s coordinates.
MIDLOC2_Precision_Comment	Text comment providing an explanation of the precision code.

MIDLOC2_General_Comment	Text comment providing any additional information relevant to the coordinates or the coordinates' sources.
PRIOGRID_Cell	Numeric variable providing the identifier for the PRIO-GRID cell whose centroid is nearest the noted coordinates. Please note, we include this variable as a matter of convenience for integrating MIDLOC with PRIO-GRID. However, we follow the UCDP-GED dataset in identifying the cell where the coordinates fall irrespective of countries involved, and irrespective of PRIO-GRID's majority area rule. For more information, please see both the PRIO-GRID Codebook (Tollefsen, Strand, & Buhaug 2012) and the UCDP-GED Codebook (Croicu & Sundberg 2017; Sundberg & Melander 2013).
MIDLOC11_location	Text comment on the location of the event from MIDLOC1.1 (available from 1993 - 2001 only).
MIDLOC11_midlocmeasuringpoint	Most precise known location from which MIDLOC1.1 observation is coded (available from 1993 - 2001 only).
MIDLOC11_latitude	Latitude in decimalized degrees from MIDLOC1.1 (available from 1993 - 2001 only).
MIDLOC11_longitude	Longitude in decimalized degrees from MIDLOC1.1 (available from 1993 - 2001 only).
MIDLOC11_precision	An indicator of geographic precision/resolution of coded location from MIDLOC1.1 (available from 1993 - 2001 only). 1 = point representation of a known point location (e.g., conflict located in town, city, base X) 2 = point representation of a location of proximity (e.g., conflict located close to town, city, base X)

3 = point representation of a sub-national unit/polygon (e.g., conflict located within region, district, area X)

4 = point representation of a location along a line (e.g., conflict located on border, river, road X)

5 = point representation of a large/national polygon (e.g., conflict located in country, sea, ocean X)

7 = no known location/ location missing

Other Notes

Consistent with MIDI v4.01, we have included incidents from disputes that began in 1992 and continued into 1993. Consistent with MIDA v4.2, we have elected to fold dispute 4087 into dispute 4022 as well as to fold dispute 4157 into dispute 4156. We have also chosen to add incident 4186001 into MIDLOC-I v2.0 even though it is not included in MIDI v4.01. We suspect its absence in MIDI v4.01 to be an oversight.

Additional Thanks

We are appreciative of Joshua Caldon, Christopher J. Fariss, Alessandro Guarino, and Nadav G. Shelef for kindly identifying discrepancies in earlier versions of MIDLOC. Any remaining errors are our own and may be sent to the authors at either paul.bezerra@usnwc.edu or abraith@email.arizona.edu.

References

- Braithwaite, A. 2010. MIDLOC: Introducing the Militarized Interstate Dispute Location dataset. *Journal of Peace Research*, 47(1), 91-98.
- Croicu, M. & Sundberg, R. 2017. UCDP-GED Codebook version 17.1, Department of Peace and Conflict Research, Uppsala University.
- Ghosn, F., Palmer, G., & Bremer, S. A. 2004. The MID3 dataset, 1993-2001: Procedures, coding rules, and description. *Conflict Management and Peace Science*, 21(2), 133-154.
- Hijmans, R., Kapoor, J., Wieczorek, J., Garcia, N., Maunahan, A., Rala, A., & Alex, M. 2012. GADM database of global administrative areas (version 2.0).
- Palmer, G., d'Orazio, V., Kenwick, M., & Lane, M. 2015. The MID4 dataset, 2002-2010: Procedures, coding rules, and description. *Conflict Management and Peace Science*, 32(2), 222-242.
- Strandow, D., Findley, M., Nielson, D., & Powell, J. (2011). The UCDP and AidData codebook on georeferencing aid: Version 1.1. Department of Peace and Conflict Research, Uppsala University.
- Sundberg, R., & Melander, E. 2013. Introducing the UCDP georeferenced event dataset. *Journal of Peace Research*, 50(4), 523-532.
- Tollefsen, A. F., Strand, H., & Buhaug, H. 2012. PRIO-GRID: A unified spatial data structure. *Journal of Peace Research*, 49(2), 363-374.
- Weidmann, N. B., Kuse, D., & Gleditsch, K. S. 2010. The geography of the international system: The CShapes dataset. *International Interactions*, 36(1), 86-106.