

Supplemental Appendix for *A Wiki-based Dataset of Military Operations with Novel Strategic Technologies (MONSTr)*

Authorship Anonymized

This appendix accompanies the paper “A Wiki-based Dataset of Military Operations with Novel Strategic Technologies (MONSTr)”. It provides supplemental information concerning the dataset criteria and covariates, its relationship to other datasets, validation, coding, and descriptive statistics for data used in the demo model.

Contents

1	Comparison of Dataset Definitions and Scope	2
2	Validation checks	3
3	Naval Case Coding Protocols and Comparisons	9
4	Descriptive Statistics	12

1 Comparison of Dataset Definitions and Scope

Table A1: Definitional Boundaries Across Datasets Capturing Military Interventions

Dataset →	<i>MONSTr</i>	<i>MIPS</i>	<i>IMI</i>	<i>MIP</i>	<i>RAND</i>
Criterion ↓					
<i>Political</i>	✓	✓	✓	×	×
<i>Size Threshold</i>	none	500	none	none	1000
<i>Target</i>	any foreign	any foreign	any foreign	state actor	any foreign
<i>Threats</i>	×	×	×	✓	✓
<i>Training Exercises</i>	×	×	×	✓	✓
<i>Troop Movement / Mobilization</i>	×	×	×	✓	✓
<i>Diplomatic Cover / Evacuation</i>	×	×	✓	✓	✓
<i>Humanitarian Force</i>	×	×	✓	✓	✓
<i>Unintentional Contingencies</i>	×	×	×	✓	✓
<i>Airspace Violations</i>	×	×	×	✓	✓
<i>Sanctions Enforcement</i>	×	×	×	✓	✓
<i>Irregular Troops</i>	*	×	×	✓	✓

Note: MONSTr features operations nested within interventions. This table demonstrates our scoping conditions for how we identify military interventions within which to collect component cases at the level of our unit of analysis.

2 Validation checks

To verify Wikipedia’s codings of the variables, we performed manual validation checks on a random 6% sample of the observations in the data prior to our work identifying operations. Since accuracy about Wikipedia’s coverage of operations and higher aggregation interventions is of interest, we found value in checking both. Consequently, the sample was drawn from all Wikipedia pages on US military activities, not just the observations later culled and used in the statistical model.

The authors consulted with personnel from all four of the major military services to identify reputable government sources that could contain information on military operations. For each observation, the following list of government sources was searched for a match based on the name of the military operation:¹

- Air Force Historical Research Agency
- Air Force Public Affairs Office
- Army Asymmetric Working Group
- Army Center for Military History
- Army Contemporary Operations Study Team
- Army Combined Arms Center

¹The date was also included in searches where necessary to produce accurate search results, especially given that Wikipedia naming conventions sometimes differ from the military’s. For example, the observation that Wikipedia identifies as the “Second Battle of Fallujah” is not known as the “Second” in official records, so dates are used to ensure the correct match.

- Center for Army Lessons Learned
- Marines Battle Study Packages

Once a match was located in a government source, variables were verified for the start and end date, location (either geo-coordinates or city/town name, government sources primarily providing the latter), belligerents (both allied partners and adversaries) and the military means used. In cases where none of the selected government or military sources contained information about an operation, reputable think tank sources or news outlets were used. To avoid circularity in validation, we were mindful to select news outlets that were not themselves cited in the Wikipedia article, although that was not possible in a few instances. Two Freedom of Information Act (FOIA) requests were also filed for After Action Reports (AAR), Battle Damage Assessments (BDA), and ship deck logs for dates and ships where known missile attacks occurred but little open source evidence existed. CENTCOM FOIA #18-0027 was returned without any information, as it was all deemed classified. AFRICOM FOIA #2017-78 was returned with partially redacted information on cruise missile targets from select naval vessels that validated some of the data on Wikipedia.

The random sample chosen for validation, verification sources used for each case, and data evaluations are provided in full below.

Observation: Deir ez-Zor Governorate clashes (April 2018)

- *Verification source:* [REACH NGO](#), [Reuters](#)
- *Evaluation:* exact matches

Observation: Insurgency in Khyber Pakhtunkhwa

- *Verification source:* [Congressional Research Service \(CRS\)](#), [Reuters](#)
- *Evaluation:* exact matches

Observation: Second Battle of Fallujah

- *Verification source:* [Marines Battle Study Packages](#)
- *Evaluation:* exact matches

Observation: Banja Luka incident

- *Verification source:* [Air Force Historical Research Agency](#)
- *Evaluation:* exact matches

Observation: Battle of Samarra (2004)

- *Verification source:* [Army Combined Arms Center](#)
- *Evaluation:* exact matches

Observation: Battle of Hit (2016)²

- *Verification source:* [Combined Joint Task Force - Operation Inherent Resolve \(CJTF-OIR\)](#), [Department Of Defense News](#)
- *Evaluation:* Exact matches from DOD News. Only dates and means verifiable from CFTF-OIR source.

²Also known as Operation Desert Lynx

Observation: Battle of Khasham

- *Verification source:* [CENTCOM Senate Committee Hearing](#)
- *Evaluation:* exact matches

Observation: Battle of Basra (2008)

- *Verification source:* [Army Center for Military History, Institute for the Study of Warfare](#)
- *Evaluation:* exact matches

Observation: Mosul offensive (2015)

- *Verification source:* [CJTF-OIR](#)
- *Evaluation:* exact matches

Observation: Battle of Suq al Ghazi

- *Verification source:* [New York Times](#)
- *Evaluation:* exact matches, no government or military record found

Observation: United States bombing of the Chinese embassy in Belgrade

- *Verification source:* [US State Department Public Statement](#)
- *Evaluation:* Belligerent unclear since event was an accident / GPS targeting mistake (Wiki describes the target as “disputed”). Sources confirm Wiki’s target coordinates.

Observation: Night raid on Narang

- *Verification source:* [Center for Civilians in Conflict](#), *New York Times*
- *Evaluation:* exact matches, no government or military record found

Observation: Operation Iron Saber

- *Verification source:* [Army War College \(AWC\) Monograph](#), *Washington Times*
- *Evaluation:* Wiki end date earlier than end date provided by media sources

Observation: Battle of Fallujah (2016)³

- *Verification source:* [CRS](#), [US Institute for Peace](#)
- *Evaluation:* exact matches

Observation: Battle of Samawah (2003)

- *Verification source:* [AWC monograph](#)
- *Evaluation:* exact matches

Observation: Battle of Abu Ghraib

- *Verification source:* [Air Force Public Affairs Office](#)
- *Evaluation:* exact matches

Observation: Operation Hammer Down

- **Verification source:** [Army Combined Arms Center](#)
- **Evaluation:** exact matches

³Code-named Operation Breaking Terrorism

Observation: Operation Freedom's Sentinel

- *Verification source:* [Defense Manpower Data Center's Defense Casualty Analysis System](#)
- *Evaluation:* exact matches

Observation: Fall of Mazar-i-Sharif

- *Verification source:* [Army Contemporary Operations Study Team](#)
- *Evaluation:* belligerent described by gov source as Taliban and al-Qaeda, described by Wiki as just al-Qaeda

Observation: Operation Dragon Strike

- *Verification source:* [Marine Corps University Press, Institute for the Study of Warfare](#)
- **Evaluation:** end date could not be verified - not described in external sources

Observation: Battle of Mosul (2016-2017)⁴

- *Verification source:* [Center for Army Lessons Learned](#)
- *Evaluation:* exact matches

Observation: Operation Odyssey Dawn (March 19 - April 27, 2011)

- *Verification source:* AFRICOM FOIA case number 2017-78
- *Evaluation:* exact matches for means used and date, BDA column with target information redacted

⁴Code-named Operation Eagle Strike

Observation: Operation Iraqi Freedom (March 21, 2003)

- *Verification source:* USS O’Kane (DDG 77) deck log
- *Evaluation:* exact matches for means used and date, target information not provided

Observation: Operation Inherent Resolve (April 7, 2017)

- *Verification source:* AFRICOM FOIA case number 2017-78, USS Porter (DDG 78) deck log
- *Evaluation:* exact matches for means used and date, BDA column with target information redacted

3 Naval Case Coding Protocols and Comparisons

The International Military Intervention dataset (IMI) ([Pearson and Baumann 1993](#); [Kisangani and Pickering 2008](#)) and Military Intervention Project (MIP) ([Kushi and Toft 2022](#)) feature interventions coded as naval means. We diverge from coding this domain for cases comprising MONSTr for two reasons. First, some interventions coded as naval are outside our definitional scope as humanitarian or disaster relief events. Second, we code events by the platform deployed to the target, not its origin or launchpoint. Thus, when a cruise missile is launched from a naval vessel, IMI and MIP code it as naval while MONSTr codes it as a “cruise missile” means. Similarly, when aircraft take off from a naval vessel to perform an aerial bombing mission we code this as an “aerial bombing” means. We are agnostic about

the launch points or service branch of these platforms, instead capturing the administration of means. For transparency, we detail these cases for both datasets.

IMI naval interventions

IMI contains 19 observations coded as naval means. Of those, ten do not meet our criteria for inclusion in MONSTr as humanitarian or disaster relief operations. They include (as listed in IMI):

US troops give humanitarian relief to Somalia

Operation Assured Response by US to evacuate 2444 people from Liberia

US evacuates American citizens from Albania

US evacuates civilians from Sierra Leone

US send peacekeepers to Liberia

US provides tsunami relief to Thailand

US provides tsunami relief to Indonesia

US provides tsunami relief to Sri Lanka

US troops build schools and provide medical aid in Haiti

US evacuates American citizens from Liberia during civil conflict

Of the remaining nine naval cases, three are coded as naval “shelling,” four as naval “intimidate,” and two as naval “transportation.” All “shelling” cases are coded in MONSTr as cruise missile means, and additional checks confirm that the nature of naval involvement was solely as the origin of the missile launch. Of the four “intimidate” cases, three concern

Operations Desert Storm and Desert Shield, where the naval intimidation involved repositioning vessels that were later used to launch cruise missiles or aircraft. Those three are coded in MONSTr as cruise missile and aerial bombing cases. The remaining “intimidate” case is

US restores democratically elected government in Haiti

which MONSTr codes as ground troops and paramilitary since their conveyance to the theater was the purpose of the vessel’s presence. The remaining two “transport” cases are

US buildup of troops in Kuwait after Iraq’s provocation

US aids in restoring order in Haiti

both of which MONSTr codes as ground troops for the same reason as the last.

MIP naval interventions

MIP has 33 observations coded as naval means. Twenty-four were determined outside the definitional criteria for inclusion in MONSTr because they were humanitarian or disaster relief operations or militarized events that, in our minds, did not constitute a US military operation (e.g., observations “Fishing Boats,” “Evacuating U.S. citizens,” “Venezuelan Refugees Humanitarian mission,” and “Detention of U.S. Service members”).

The remaining nine populate on Wikipedia and are thus included in MONSTr. MIP does not code the type of naval intervention like IMI does, instead featuring a binary variable indicating presence and a “MaxNavy” variable corresponding to the number of vessels. Eight

of these cases are coded in MONSTr as cruise missiles or bomber aircraft, launched from naval vessels. The sole case in our dataset that does not correspond to an aerial platform launch point from a naval vessel is Operation Uphold Democracy, which is the same 1991 Haitian coup attempt that IMI codes as a naval “intimidate” and that we code as ground troops.

4 Descriptive Statistics

Table A2 shows descriptive statistics for the primary new variables created in MONSTr.

This table exludes variables drawn from other datasets.

Table A2: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Drones	336	0.09	0.28	0	1
Air to air	336	0.02	0.15	0	1
Cruise missiles	336	0.04	0.20	0	1
Aerial bombing	336	0.43	0.50	0	1
Close air support	336	0.32	0.47	0	1
Ground troops	336	0.49	0.50	0	1
Paramilitary	336	0.26	0.44	0	1
Duration (days)	333	50.67	181.76	0	2,151
Days into parent (days)	336	1,625.26	1,475.81	−4	9,596
Ally count	336	1.57	2.35	0	32
Adversary count (state)	336	0.32	0.63	0	5
Adversary count (non-state)	336	1.49	2.52	0	39

All military means in the original dataset shown.

References

- Kisangani, Emizet F. and Jeffrey Pickering. 2008. “International Military Intervention, 1989-2005: Version 1.”.
- Kushi, Sidita and Monica Duffy Toft. 2022. “Introducing the Military Intervention Project: A New Dataset on US Military Interventions, 1776–2019.” *Journal of Conflict Resolution* p. 00220027221117546.
- Pearson, Frederic S. and Robert A. Baumann. 1993. “International Military Intervention, 1946-1988: Version 1.”.