Conflict Site 1989–2008:

A Geo-Referenced Dataset on Armed Conflict

Special data feature manuscript prepared for submission to

Conflict Management and Peace Science

7 April 2011

Word count 3,767

Introduction

A recent stream of research has argued that existing quantitative studies on intrastate conflicts too often suffer from a discrepancy between theory and empirical analysis. Empirical evidence shows that civil war typically is a local phenomenon and the theorized causal mechanisms tend to operate at sub-national levels. Nevertheless, most empirical analyses have been carried out on the state level, which might lead to biased results (Buhaug and Rød, 2006).

Naturally, some characteristics of a state, such as features of the central government, are constant across space and should therefore be measured on the state level, but other factors, for instance economic development and ethnic composition, exhibit substantive cross-regional variation. The appropriate level of aggregation is dependent on the purported causal mechanism and the research question.[1] One way of overcoming the aggregation problem is to shift focus from the state-level and carry out disaggregated statistical studies. Increased availability of data has made this possible and as a result there has been a spate of articles based on disaggregated data published in recent years. [2]

The Conflict Site project aims to contribute to the development of this field of research by providing a global dataset on the geographic location and scope of armed conflicts in the post-Cold War period. This is not the first attempt to quantify the location of conflicts within states; however, earlier efforts either suffer from lack

[1] See Buhaug and Lujala (2005) for a more thorough discussion.
[2] See for instance the 2009 *Journal of Conflict Resolution* special issue on "Disaggregating Civil War".

of temporal variation (Buhaug and Gates, 2002; Gleditsch et al., 2002), lack global coverage (Buhaug and Rød, 2006; Raleigh et al., 2010), or are limited to the point location of conflict outbreak (Braithwaite, 2010). By making this data available, the Conflict Site dataset will make it possible for researchers to carry out more fine grained analyses, in which the conflict itself or the conflict-affected area is the unit of analysis. The Conflict Site data can be combined with other disaggregated datasets in order to test existing theories and hypotheses on a level of aggregation that corresponds closely to the relevant causal mechanisms. This is a promising research agenda that offers great potential for theory development in the field of conflict research, and that is likely to contribute new and valuable knowledge about civil wars and related phenomena.

The paper proceeds with a brief description of the new dataset, including definitions and illustrations. Next, it presents some general trends of the global dispersion of conflict before providing a few examples of how the data may be applied to reveal new features about contemporary civil wars. A concluding section closes the paper.

The New Dataset

The original version of the Conflict Site data was compiled by Halvard Buhaug for use in Buhaug and Gates (2002) and subsequently released as part of the official UCDP/PRIO Armed Conflict Dataset (Gleditsch et al., 2002).[3] Later, the geocoding

---

[3] The UCDP/PRIO Armed Conflict Dataset defines a 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" (Harbom et al.,

component was separated from the main UCDP/PRIO project and released under its own name, the Conflict Site version 2.0 (Raleigh et al., 2006). These data have since been used in numerous studies (e.g., Buhaug and Lujala, 2005; Levy et al., 2005; Raleigh and Urdal, 2007; United Nations Environment Programme, 2002; Wood, 2010).

The allocation of geographic coordinates to each conflict facilitated geographically disaggregated empirical analysis where countries could be substituted by provinces, grid cells or conflict zones as units of analysis. At the same time, the earlier versions of the Conflict Site dataset suffered from a major limitation: a lack of temporal dimension. With very few exceptions, the size and location of the coded conflict zones did not vary over time; hence, the spatial data in essence reflected the maximum spatial extent of the conflicts during their course. This limitation was of little concern if one sought to understand geographic determinants of conflict prevalence and duration but it hindered systematic analyses of conflict onset and dynamics. Besides, the geocoding rested on a much narrower selection of source material (mostly *Keesing's World News Archive*) than the parent conflict dataset, which at least in principle means that some conflict zones may have been inaccurately coded.

The new version (3.0) of the Conflict Site dataset is a totally revised and recoded update that corresponds to the latest available version of the UCDP/PRIO Armed Conflict Dataset (Version 4-2010). One of the major improvements is that the data

2009). The variable *location* in the UCDP/PRIO Armed Conflict Dataset merely identifies "the government side of a conflict, and should not be interpreted as the geographical location of the conflict" (Harbom et al., 2009).

are structured in a conflict-year format, thereby adding a temporal – dynamic – aspect lacking in previous version. The time period covered is 1989-2008. In addition, a wider range of sources than in earlier versions have been used, including news reports from *Factiva* and a generous selection of case-specific literature. Moreover, the use of geographic information systems (GIS) software and databases has increased the precision in the geo-referenced coding.

Every conflict-year in the dataset is assigned a circular conflict zone, which is defined by a center point (*location*), given as latitude and longitude coordinates in decimal degrees, and a radius (*scope*) indicator that measures the distance from the center point to the most distant point in the conflict zone, rounded upwards to the nearest 50 kilometers (Figure 1). The conflict zone covers the area directly affected by a conflict. It includes all

- locations of reported armed encounters between the parties to the conflict
- territories occupied by the parties of the conflict (in intrastate conflicts, this applies only to territories occupied by the rebel side)
- locations of rebel bases (intrastate conflicts only).

In intrastate conflicts, only armed encounters, occupied territories and rebel bases *in the country whose government is challenged by a rebel group* are included in the conflict zone.[4] This is done to ensure consistency with the *location* indicator in the

---

[4] The Conflict Site codebook provides more detailed information on coding criteria. All data and documentation will be made available from the host institution's web page upon publication.

UCDP/PRIO Armed Conflict Dataset and to avoid selectively accounting for battle events in countries not hosting an armed conflict according to UCDP's criteria.

[FIGURE 1 ABOUT HERE]

To demonstrate the need for a geo-referenced conflict dataset, consider the civil war in Chechnya. In all major conflict datasets, Russia is coded as hosting a civil war in the mid-1990s. Does this mean that all of Russia was at war and that the fighting directly affected the whole population? No. According to the new Conflict Site dataset (Figure 2), the conflict in Chechnya in 1995 covered an area of some 67,000 km^2, less than 0.5 percent of Russia's total landmass (about 17,000,000 km^2). Accordingly, it makes little sense to "explain" this conflict by relying on conventional country-level regressors – e.g., ethno-linguistic fractionalization, GDP per capita, percent mountainous terrain – as proxies for phenomena that are inherently spatial (local dominance of ethnic minority, local poverty, local topography).

[FIGURE 2 ABOUT HERE]

As illustrated in Figure 1, a drawback with circular conflict zones is that they cover more territory than is actually affected by the conflict, including territories of

neighboring countries.[5] Due to the nature of armed conflict it may be impossible to gain information on the exact location of all armed encounters, occupied territories, and rebel bases during the course of the conflict. Moreover, as the quality of reporting is prone to varying across cases as well as over time, a very detailed coding procedure (e.g., allowing for uniquely shaped, dynamic conflict polygons) would very likely introduce systematic bias and inconsistent precision in the data. The choice of circular conflict zones is essentially a trade-off between precision and simplicity, and in most cases a circular conflict zone will give a good indication of the *core area of the conflict*. More information on coding decisions and problematic cases can be found in the codebook.

The Dispersion of Armed Conflicts across Time and Space

Data from the UCDP/PRIO Armed Conflict Dataset show that the world has become more peaceful since the end of the Cold War. The number of armed conflicts today is considerably lower than in the beginning of the 1990s, although there has been a slight increase the last few years (Harbom and Wallensteen, 2010). In terms of conflict severity, a downward trend is evident for the entire post-World War II period (Lacina and Gleditsch, 2005).

One possible explanation for the declining risk of death in battle could be that the average conflict has decreased in extent over time as a consequence of regimes becoming more successful in containing the fighting to small parts of the their

---

[5] In most GIS software packages the conflict zones can easily be clipped to follow state boundaries and coast lines, like in Figure 2.

territories. Figure 3 suggests that this is not the case. Although the number of conflicts is on the wane, the total conflict-affected area is not decreasing in a similar fashion.[6] Rather, the Conflict Site data reveal several peaks and dips, including a hump in 1997-2002, which is not found in the parent dataset. One reason for the discrepancy in trends is the fact that the high number of conflicts in the early 1990s was mainly caused by a sudden rise in short conflicts in Eastern Europe and the former Soviet Union, but since these wars were relatively small they exerted a limited influence on the total conflict-affected area. Another reason for the irregular conflict area pattern is that conflicts that cover large swathes of territory exert a disproportionate influence on the aggregate conflict area, whereas all conflict-years are equally counted in the UCDP/PRIO Armed Conflict Dataset. Influential cases in the shown time period include the Democratic Republic of the Congo, Angola, and Turkey. The increase in total conflict-affected area from 2005 onwards is consistent with the UCDP/PRIO Armed Conflict Dataset conflict count.

The overall lack of decline in conflict area is an interesting finding that adds nuance to the notion of the post-Cold War decline of armed conflict. Whether this implies that the average conflict instead has relocated to peripheral and less populated areas remains to be uncovered though it is also consistent with better medical care and increased internationalism over time that have contributed to shrinking costs of war (Human Security Report Project, 2010).

---

[6] The figure shows data where the circular conflict polygons have been clipped along national boundaries and coast lines. Map data from the CShapes dataset (Weidmann et al., 2010) were used throughout this paper. The area calculations were made in the PostGIS software, which uses a spherical model of the globe. Since some conflict zones overlap, e.g., intrastate wars in Ethiopia and India, parts of those countries' territory are counted more than once. However, this is unproblematic for this application, where the purpose is to point out a general trend.

[FIGURE 3 ABOUT HERE]

Next, we divide the conflict data by type and investigate their relative influence on the global conflict area. As Figure 4 shows interstate wars make up a tiny part of the data and are hardly discernable, which is mainly a function of the low number of interstate conflicts since the end of the Cold War. Intrastate and internationalized intrastate conflicts are the predominant form of war in the examined time period in terms of spatial coverage as well as absolute number of observations.

 [FIGURE 4 ABOUT HERE]

As a final visualization of trends in conflict sites, Figure 5 shows the dispersion of conflict across regions. The bulk of the temporal variation is driven by African conflicts, with the wars in Angola, the Democratic Republic of the Congo and Sudan as the most influential cases. The Middle East is also relatively volatile, where the conflicts in Iran, Iraq, and Turkey drive most of the variance. As can be seen, the European conflicts are extremely marginal in the spatial context and can hardly be discerned.

Has the world become more peaceful since the end of the Cold War? In terms of battle-related deaths and the number of unique armed conflicts: yes. In terms of the spatial extent of conflict zones, and, potentially, the number of people and economic

activities affected, the answer is not so straightforward. The Conflict Site data show that there is no direct, linear relationship between the number and severity of conflicts and the total size of the conflict zones, so the widely acknowledged downward trend since the early 1990s is not unequivocal. Interestingly, there is no correlation (0.08) between the size of a conflict zone and the number of battle-related deaths in a conflict-year. Cases like the Naxalite rebellion in India are illustrating for this lack of relationship: in 2005 only 199 battle-related deaths were recorded but since the battle events were not spatially concentrated, the conflict zone covered an area of more than 2,000,000 square kilometer. At the other end of the spectrum is the civil war in Rwanda, with 1125 battle-related deaths in a conflict zone of about 3,000 square kilometers in 1990. [7]

[FIGURE 5 ABOUT HERE]

Application: Spatial Characteristics of Separatist and Governmental Conflicts

The conflict-year format of the current version of the Conflict Site dataset captures the dynamic character of armed conflict, which was not possible in earlier versions. This allows for more nuanced studies in which the temporal aspect is taken into account. This section presents a simple analysis of the dynamics of internal conflicts

---

[7] Data on battle-related deaths are best estimates from the UCDP Battle Related Deaths Dataset (Uppsala Conflict Data Program, 2010)

that serves to demonstrate the applicability of the dataset and point to new directions for future research.

In the first systematic study of the spatial characteristics of civil wars, Buhaug and Gates (2002) argued that separatist conflicts are mainly confined to the areas the rebel groups seek to liberate whereas conflicts over the central power are more dynamic and cover larger parts of the country.[8] Due to the static nature of their conflict data, they were unable to offer empirical evidence to substantiate the claim. With the new Conflict Site dataset, we may for the first time assess and compare the spatial mobility of armed conflicts of different kinds.

A preliminary test of Buhaug and Gates' assumption can be carried out using GIS software to visualize the conflict zones across time. The UCDP/PRIO Armed Conflict Dataset incompatibility variable indicates if a conflict is fought over *territory*, i.e. a separatist conflict, or *government*, in which the objective of the rebels is to oust the central government (Harbom et al., 2009). A visual inspection of the data gives support to Buhaug and Gates: whereas conflicts over territory are more or less confined to the same geographic area throughout their duration, the data show that conflict over the state power are more dynamic, both in terms of location and scope, and cover larger swathes of territory. As an illustration, Figure 6 depicts the separatist conflict in Bangladesh 1989–1992 and the 1999–2002 period of the conflict over state power in Chad. Evidently, the war between the Chittagong Hill

---

[8] Coups are an exception since they usually take place in the capital only.

Tracts People's Coordination Association/Peace Force and the Bangladeshi government was only active in the Chittagong Hill Tracts, while the Movement for Democracy and Justice in Chad was more mobile and fought the government in large areas of the northern part of the country.[9]

 [FIGURE 6 ABOUT HERE]

The descriptive statistics in Table 1 offers a more general demonstration of the observed pattern: the absolute size of conflicts over governmental power is on average about three times larger than separatist wars. The relative size – the conflict zone as a proportion of the total land area – is about twice as large for government conflicts.

[TABLE 1 ABOUT HERE]

As a more comprehensive and satisfactory demonstration of the different dynamics of territorial and governmental conflicts, the paper next considers multivariate regression models. For the exploratory purposes of this paper, a rigorous theoretical framework is not necessary. Instead, the expectations are formalized directly by means of the following two hypotheses:

---

[9] The data on government conflicts in the UCDP/PRIO Armed Conflict Dataset are not dyadic and thus there can be several rebel groups active in a conflict-year. This partly explains why these conflicts are more dynamic and affect larger areas, but a closer investigation reveals that this is also true for conflicts over government with only one active rebel group, as in the example from Chad above. However, due to the unstable nature of rebel groups, an analysis on the dyad-level would not be optimal either. Splits, mergers, name changes, and alliances between groups are common and would cause delimitation problems.

H1: Conflicts over territory are smaller than conflicts over state power.

H2: Conflicts over territory are more static than conflicts over state power.

To test these expectations empirically, the conflict-years in the dataset are aggregated to create one observation for each conflict[10], which constitutes the unit of analysis. The observations are clustered by country to account for multiple conflicts within one country during the examined time period (1989-2008).

In order to test Hypothesis 1, two alternative dependent variables are constructed: the size of the conflict zone, measured in thousand square kilometers (*absolute size*), and the size of the conflict zone divided by that of the country (*relative size*). Since the conflict-years have been aggregated to one observation, the mean values are used. The principle independent variable is a dummy variable (*territory*) indicating if the incompatibility concerns territory (1) or government (0). Country size in thousand square kilometers (logged) (*ln country size*) and duration (*duration*), operationalized as a count of active conflict-years since 1989, are included as control variables. By nature, larger countries can harbor larger conflicts, and the country size sets an absolute limit to the size of the conflict zone. Duration should matter since protracted conflicts give insurgents time to spread the rebellion to larger areas. In addition, two of the "usual suspects" in the civil war literature are

---

[10] Each conflict is identified by the unique conflict ID in the dataset.

controlled for: population size in millions (logged) (*ln population*) and GDP per capita, measured in thousands of U.S. dollars current prices purchasing power parity terms (*GDP per capita*).[11]

The results are presented in Models 1–2 (Table 2) and are in line with the hypothesis. The incompatibility variable is highly significant and has the expected sign, which implies that separatist wars are smaller than conflicts over state power both in relative and absolute terms. As expected, country size and duration have a significant positive effect in Model 1; large countries and protracted wars increase the size of the conflict-affected area. The average conflict over government is about 164,000 square kilometers larger than separatist conflicts, *ceteris paribus.* Interestingly, country size has a reversed effect when regressed on the relative size indicator, suggesting that conflicts tend to cover a larger share of the territory of small countries. A high GDP per capita has a mitigating effect, probably at least partly because strong states are better at curbing internal strife.

Two different operationalizations of conflict dynamics are used to test Hypothesis 2. The dependent variable in Model 3 is the difference in size (in thousand square kilometers) between the area of the smallest and the largest conflict zone (*range*) during the course of the conflict. In Model 4, the dependent variable is the distance

---

[11] Data on country size come from Weidmann et al. (2010). Population and GDP data have been retrieved from the *World Development Indicators* (World Bank, 2010). Missing GDP data have been imputed from the *Penn World Table* (Heston et al., 2009) and the *World Economic Outlook* (International Monetary Fund, 2010). Mean value of population is used, but since GDP per capita is endogenous to the model and therefore likely to change during the course of a conflict the first value in the time series is used.

in kilometers between the two center points lying farthest apart (*mobility*). Since the unit of analysis is conflict, *range* and *mobility* are calculated within each conflict time series. The main explanatory variable (*territory*) and controls remain the same as in Models 1–2.

[TABLE 2 ABOUT HERE]

Hypothesis 2 also gets strong support from the regression analysis: separatist wars are significantly more static in both models, and the difference is quite large. Country size and duration remain highly significant, whereas the effects of GDP per capita and population size are not robust and fluctuate across the models.

Conclusion

A central tenet in all research is that the choice of unit of observation should reflect the phenomenon being studied. While recent efforts to disaggregate conflicts by location or events have laid the foundation for new and innovative research programs (see for instance Cederman and Gleditsch, 2009), these datasets suffer from one of two limitations: lack of a proper temporal dimension or lack of global coverage. The Conflict Site project alleviates this shortcoming by providing new and better disaggregated data on the location and scope of all armed conflicts in the post-Cold War era.

The simple analysis carried out above demonstrated that valuable new insights can be gained by using this data. A quick glance at the data revealed that the total

conflict-affected area has not decreased in a similar manner as the frequency or severity of war, and the data also showed substantial inter-continental differences in conflict size and trends. Furthermore, the analysis found that conflicts over state power are more widespread and dynamic in spatial terms than separatist conflicts. As the civil war scholarship increasingly demand high-resolution conflict data, the Conflict Site dataset should find much interest and, hopefully, contribute to further advancement of the field.

References

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

Buhaug, Halvard, and Scott Gates. 2002. The Geography of Civil War. *Journal of Peace Research* 39(4): 417-433.

Buhaug, Halvard, and Päivi Lujala. 2005. Accounting for scale: Measuring geography in quantitative studies of civil wars. *Political Geography* 24(4): 399-418.

Buhaug, Halvard, and Jan Ketil Rød. 2006. Local determinants of African civil wars. *Political Geography* 25(3): 315-335.

Cederman, Lars-Erik, and Kristian S. Gleditsch. 2009. Introduction to special issue on disaggregating civil war. *Journal of Conflict Resolution* 53(4): 487-495.

Collier, Paul, and Anke Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56(4): 563-595.

Fearon, James D., and David D. Laitin. 2003. Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97(1): 75-90.

Gleditsch, Nils Petter, Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Håvard Strand. 2002. Armed Conflict 1946-2001: A New Dataset. *Journal of Peace Research* 39(5): 615-637.

Harbom, Lotta, Håvard Strand, and Håvard M. Nygård. 2009. *UCDP/PRIO Armed Conflict Dataset Codebook Version 4-2009*. (www.ucdp.uu.se/database)

Heston, Alan, Robert Summers, and Bettina Aten. 2009. *Penn World Table Version 6.3*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. (http://pwt.econ.upenn.edu/php_site/pwt_index.php)

Human Security Report Project. 2010. *Human Security Report 2009/2010: The Causes of Peace and the Shrinking Costs of War*. Simon Fraser University, Canada: Human Security Report Project.

International Monetary Fund. 2010. *World Economic Outlook*. (http://www.imf.org/external/pubs/ft/weo/2010/02/weodata/index.aspx)

Lacina, Bethany, and Nils Petter Gleditsch. 2005. Monitoring trends in global combat: A new dataset of battle deaths. *European Journal of Population* 21(2-3): 145-166.

Levy, Marc A., Catherine Thorkelson, Charles Vörösmarty, Ellen Douglas, and Macartan Humphreys. 2005. Freshwater Availability Anomalies and Outbreak of Internal War: Results from a Global Spatial Time Series Analysis. Paper presented at the Human Security and Climate Change Workshop in Asker, Norway, 21-23 June 2005.

Raleigh, Clionadh, David Cunningham, Lars Wilhelmsen, and Nils Petter Gleditsch. 2006. Conflict Sites 1946–2005 Codebook. Peace Research Institute Oslo (PRIO).

Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joakim Karlsen. 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research* 47(5): 651-660.

Raleigh, Clionadh, and Henrik Urdal. 2007. Climate change, environmental degradation and armed conflict. *Political Geography* 26(6): 674-694.

Tomz, Michael, Jason Wittenberg, and Gary King. 2003. *CLARIFY: Software for Interpreting and Presenting Statistical Results.* Version 2.1. Stanford University, University of Wisconsin, and Harvard University. (http://gking.harvard.edu/)

United Nations Environment Programme. 2002. *Mountain Watch*. Cambridge, UK: UNEP World Conservation Monitoring Centre.

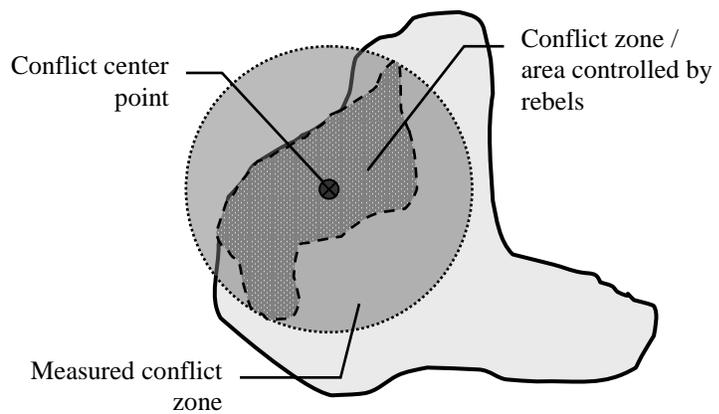Uppsala Conflict Data Program. 2010. *UCDP Battle-Related Deaths Dataset v.5-2010*. Uppsala University. (http://www.ucdp.uu.se)

Weidmann, Nils B., Doreen Kuse and Kristian S. Gleditsch. 2010. The Geography of the International System: The CShapes Dataset. *International Interactions* 36(1): 86-106.

Wood, Reed M. 2010. Rebel capability and strategic violence against civilians. *Journal of Peace Research* 47(5): 601-614.

World Bank, 2010. *World Development Indicators*. (http://data.worldbank.org/data-catalog/world-development-indicators)

Conflict center point

Conflict zone / area controlled by rebels

Measured conflict zone

Note: This figure was first used in Buhaug and Gates (2002).

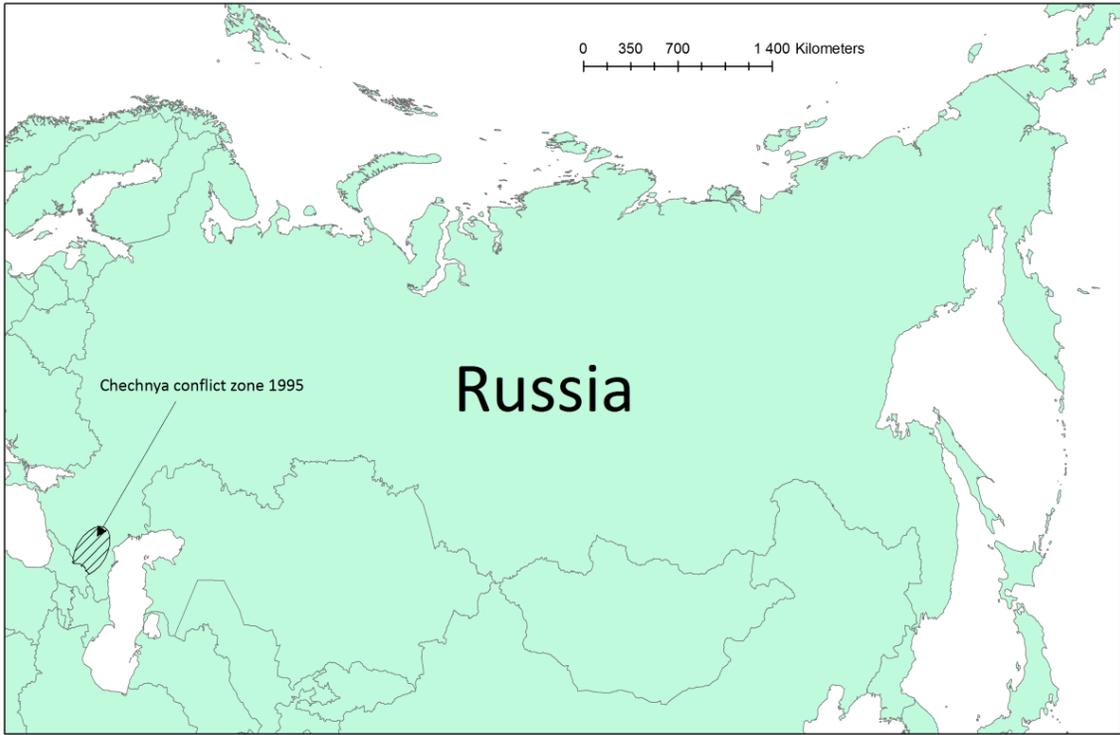Figure 1. Identifying the conflict zone

Figure 2. Locating the Chechnya civil war in Russia, 1995
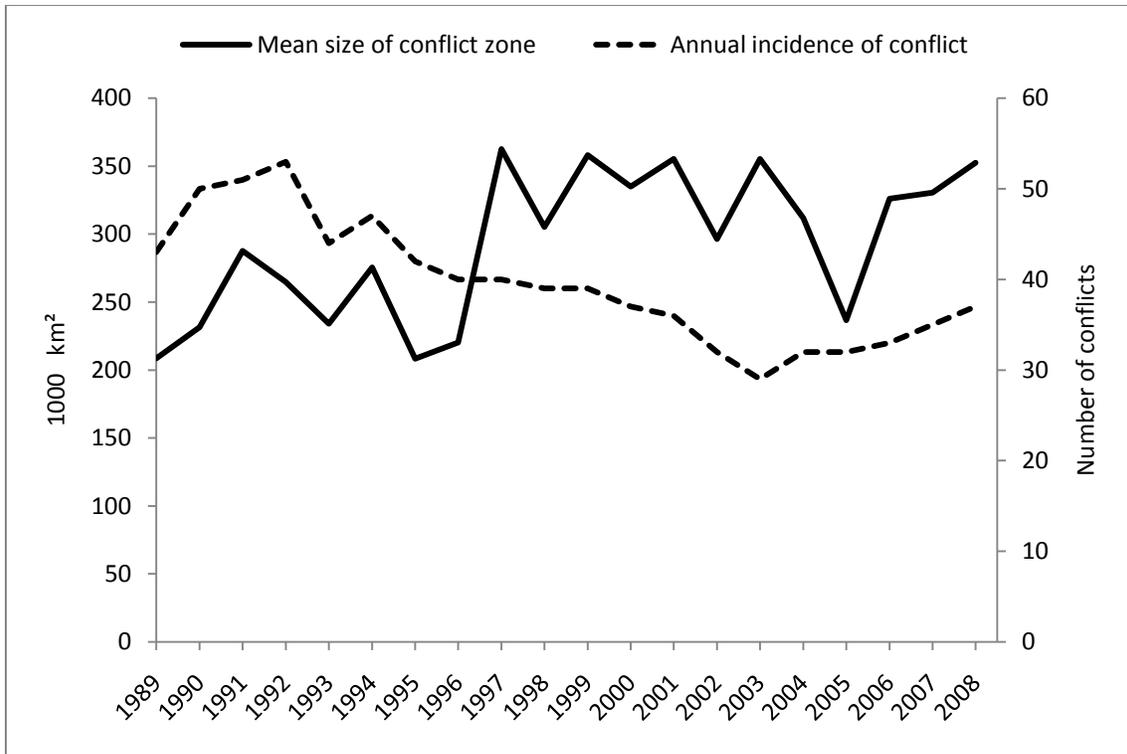
Figure 3. Trends in frequency and size of armed conflicts, 1989–2008
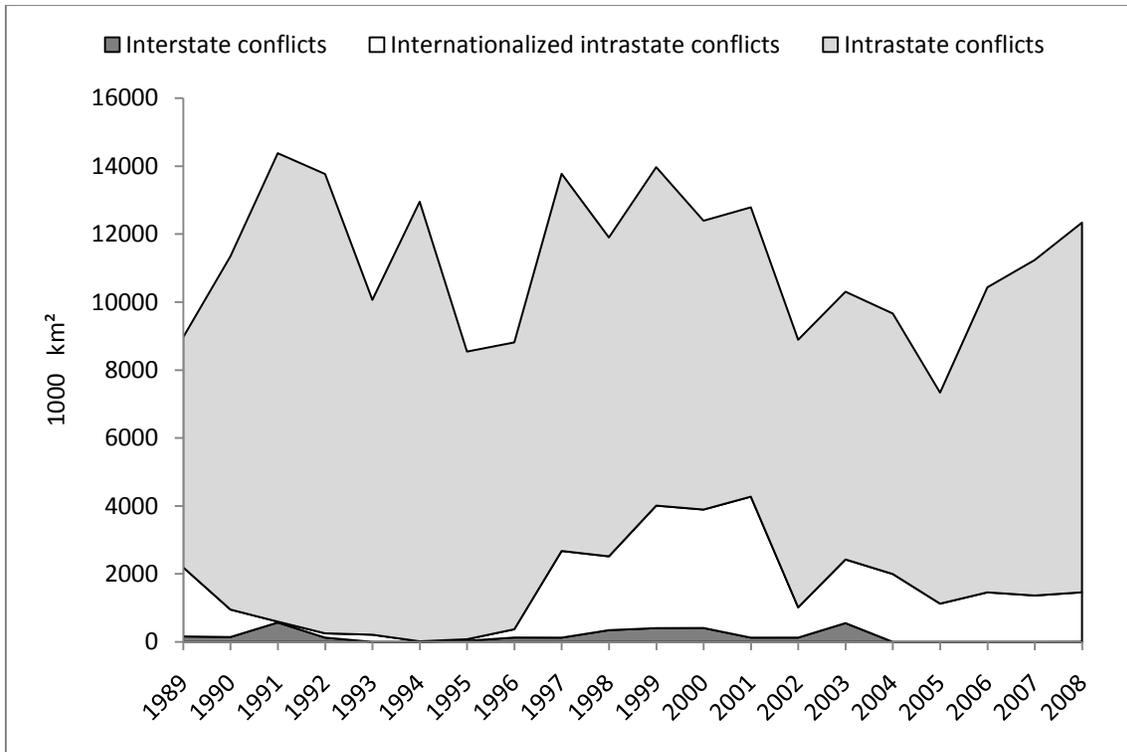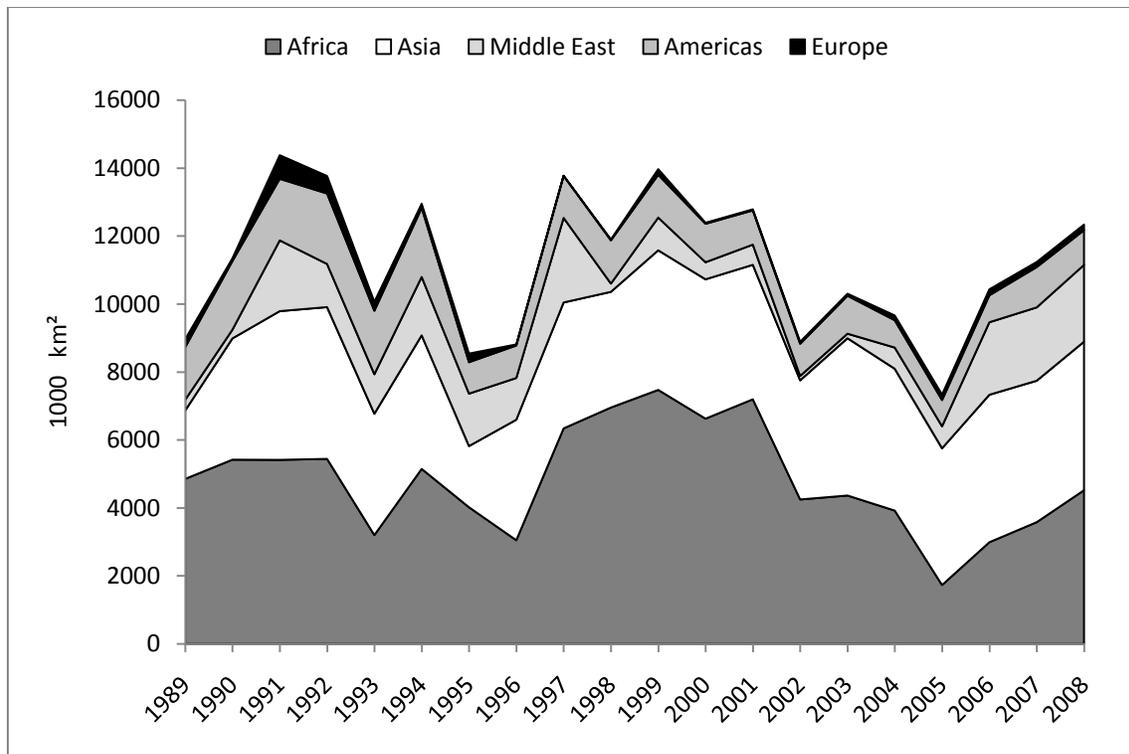
Figure 4. Conflict area by type

Figure 5. Conflict area by region

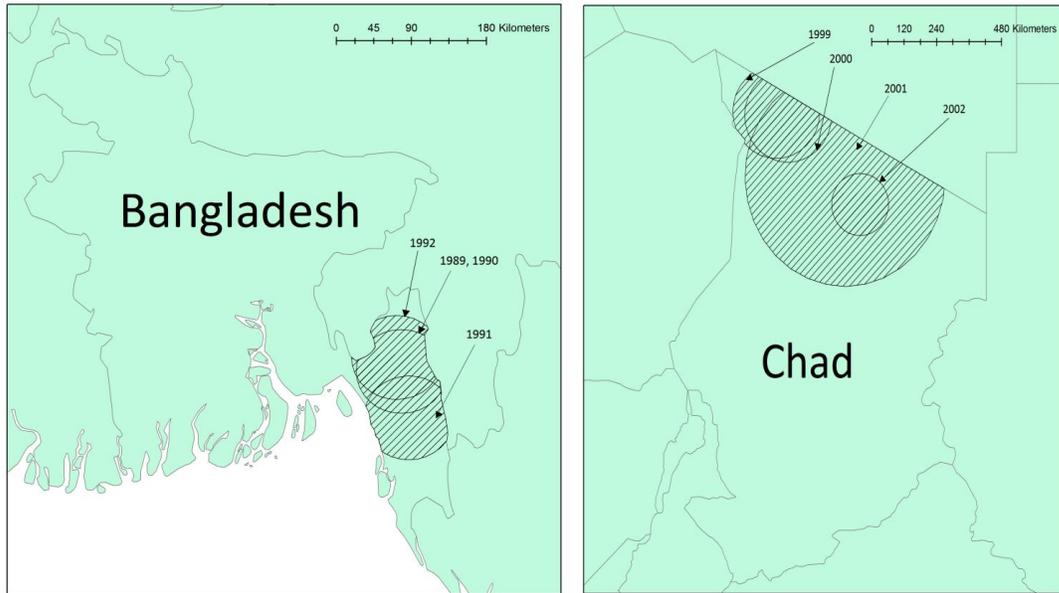Figure 6. Conflict zones in Bangladesh (territorial) and Chad (governmental)

Table 1. Descriptive statistics

	Obs.	Mean	Std. Dev.	Minimum	Maximum
Incompatibility: government					
Absolute size of conflict zone	380	439.975	546.126	1.053	2315.291
Relative size of conflict zone	380	0.529	0.330	0.0005	0.996
Incompatibility: territory					
Absolute size of conflict zone	388	140.613	210.330	0.45	873.193
Relative size of conflict zone	388	0.234	0.289	0.0001	1.000

Table 2. Determinants of conflict size and mobility

	Model 1 Absolute size	Model 2 Relative size	Model 3 Range	Model 4 Mobility
Constant	-125.82	704.24	-229.78	-70.93
	(96.37)	(79.66)	(125.14)	(87.56)
Territory	-163.67*	-164.92**	-217.59**	-150.14**
	(71.99)	(57.10)	(72.82)	(48.54)
In Country size	66.42*	-39.04*	94.64**	68.62*
	(26.15)	(16.31)	(33.93)	(27.91)
Duration	15.37**	11.62**	23.01**	16.18**
	(2.99)	(3.32)	(4.93)	(3.68)
In Population	-25.15	-33.36	-55.87*	-51.78
	(20.47)	(18.51)	(25.45)	(32.39)
GDP per capita	-8.02*	-4.08	-5.79	-7.92*
	(3.57)	(3.36)	(4.70)	(2.99)
N	119	119	119	119
R^2	0.32	0.35	0.34	0.34

*p<.05, **p<.01. Robust standard errors clustered by country in parentheses.