
Peering into the Fog of War: The Geography of the WikiLeaks Afghanistan War Logs, 2004–2009

John O'Loughlin, Frank D. W. Witmer, Andrew M. Linke,
and Nancy Thorwardson¹

Abstract: A team of U.S. political geographers analyzes the secret Afghanistan war logs released by WikiLeaks.org. They offer the chance to examine in detail the dynamics of the conflict in that country. Doing so in a spatial framework is possible because each of the 77,000 events has geographic coordinates and dates. Using cartographic and geostatistical tools, the authors map the changing distribution of the events and compare them to the well-known violent-events ACLED database (see O'Loughlin et al., 2010 in this issue). They conclude that ACLED comprises a representative set of the more comprehensive data in the released files. The released war logs show that the Afghan insurgency spread rapidly in 2008–2009, that the insurgency is moving out of its traditional Pashtun heartlands, and remains mostly rural in location. Hotspot and cluster analysis identifies the key locations of the current war, which indicate that it is relocating to new provinces in Afghanistan while intensifying in the eastern border regions and in the south. *Journal of Economic Literature*, Classification Numbers: H560, H770, O180. 14 figures, 2 tables, 30 references. Key words: Afghanistan, U.S. military, WikiLeaks, ACLED, Taliban, insurgency, civilian casualties, Pakistan, Al-Qaeda, violence hotspots.

The release on July 25, 2010 of almost 77,000 individual war logs from January 2004 to December 2009 of the events of the war in Afghanistan by WikiLeaks.org represents the possibility of a dramatic turn in the discussion and analysis of contemporary conflicts. Efforts to make sense of ongoing wars by most commentators (apart from individuals who have served in government agencies) chiefly rely on media accounts of the conflict; they spend immense amounts of time and resources trying to penetrate the texts of unclassified and declassified government documents, and official and non-governmental agency reports. As a result of the uncertain reliability of these sources, policy suggestions and academic analysis are always subject to the criticism that the data from which conclusions are drawn have been falsified or are biased, uneven in coverage, amnesiac about certain subjects, or exculpatory of government decisions. The unexpurgated war logs offer a ground-level view of the fighting, with each row in the enormous database corresponding to an event for which a report is filed.²

¹Respectively, College Professor of Distinction, Institute of Behavioral Science and Department of Geography, University of Colorado, Campus Box 487, Boulder, CO 80309 (johno@colorado.edu); Post-Doctoral Research Associate, Institute of Behavioral Science, University of Colorado, Campus Box 487, Boulder, CO 80309 (witmer@colorado.edu); Graduate Research Assistant, Institute of Behavioral Science, University of Colorado, Campus Box 487, Boulder, CO 80309 (andrew.m.linke@colorado.edu); and Professional Research Assistant, Computing and Research Services, Institute of Behavioral Science, University of Colorado, Boulder, Campus Box 483, CO 80309 (nancy.thorwardson@colorado.edu). The authors thank Clionadh Raleigh, Trinity College, Dublin, Ireland for allowing us access to the unreleased ACLED data for Afghanistan and Pakistan and Victor Winston for arranging to delay the publication of this issue of *Eurasian Geography and Economics* to include the paper. The interpretations and conclusions in this paper are solely those of the authors.

²In the paper's title, we use the phrase "fog of war," well-known from the Errol Morris documentary film about Robert McNamara, to characterize the mass of information. Penetrating its thickness is a task for a "data mining" approach, to identify trends and patterns that are not evident from a reading of innumerable individual records.

Thus, for the first time, researchers have access to the full account in all its confused renderings in the same way as contract researchers have had in recent conflicts.³

Immediately after we finished a companion paper for this issue of *Eurasian Geography and Economics* (O'Loughlin et al., 2010) on the contemporary wars in Afghanistan and Pakistan in 2008–2009, the release of the war logs by WikiLeaks.org offered the chance to compare our results and conclusions (which were based on the coding of media reports) to those derived from analysis of the “raw data,” the material upon which the media reports were based. We recognized the risks of this re-analysis in the atmosphere of an enormous outcry by government officials about the leaks and the significant public interest in the early (indeed preliminary) stories about the contents in the three news outlets (*The New York Times*, *The Guardian*, and *Der Spiegel*) that had access to them for a month before the data were put online. Each newspaper wrote its own stories and tended to emphasize different facets of the material, although the volume of the documents hindered measured consideration of the myriad of stories contained therein. While the *New York Times* (e.g., see *The War Logs*, 2010) highlighted the WikiLeaks events that had accounts of Pakistani ISI (Inter-Service Intelligence) agents meeting with the Taliban, the *Guardian* emphasized the level of (previously unreported) civilian casualties and *Der Spiegel* focused on the logs with contents of the German military contingent's participation in the actions. These newspapers and dozens of bloggers soon made interactive maps and graphs available of portions of the war logs' contents online, a task easily accommodated by the georeferences and date time-stamp on every event. The *Guardian* reported that its war logs website received over a million hits in the first few days of publication.

Officialdom was quick to condemn the person or persons who provided the leaked information and the organization (WikiLeaks) that placed the data on the web, but they did not deny the veracity of the information. President Obama, among others, stated that the information was old (2004–2009) and thus did not reflect current realities, including the changed U.S. posture toward reducing civilian casualties by more accurate use of force (instituted in July 2009) and the additional 30,000 U.S. troops that now (August 2010) bring the ISAF (International Security Assistance Force) in Afghanistan to about 120,000. Top U.S. officials, such as Secretary of Defense Robert Gates and Chairman of the Joint Chiefs of Staff, Mike Mullen, accused Julian Assange of WikiLeaks of risking the lives of Afghan informants and called for full investigation of the leak's source and paths of dissemination. Assange stated that WikiLeaks had received over 91,000 files but that he released only 76,911 initially because the remaining files contained individual names and other sensitive material.⁴

Unlike the Pentagon Papers (a classified history of U.S. involvement in Vietnam) leak to the *New York Times* by Daniel Ellsberg in 1971, to which it was initially compared, the WikiLeaks files are uncensored accounts of the daily nitty-gritty of war and do not contradict the official version of the origins of the conflict, as did the Pentagon Papers (e.g., see Rich, 2010).⁵ In the broader sense, the July 25, 2010 WikiLeaks files, though classified as “Secret,” simply provide a soldier's eye view of the escalating Afghanistan conflict and hold relatively few diplomatic or intelligence reports. What the files show, above all else, is the growing strength and activity of the Taliban over the six-year period, the geographic spread of violence from its longtime concentration in a few provinces in eastern and southern Afghanistan, and a dramatic rise in casualties in the last two years of the data series. Most observers of the

³Details of the release and the role of reporters and news media can be found in Hendler (2010).

⁴See <http://wardiary.wikileaks.org/>

⁵Quoting the WikiLeaks site, “The reports come from US Army with the exception of most Special Forces activities. The reports do not generally cover top-secret operations or European or other ISAF Forces operations.”

conflict and even a majority of the U.S. public had already reached similar conclusions. An August 3, 2010 Gallup poll showed that 43 percent of Americans “think that the war was a mistake,” while approval of President Obama’s handling of the war had dropped to 36 percent (Jones, 2010).

We put aside initial misgivings about re-analyzing the WikiLeaks war logs when we saw the numerous web cartographies of the war data and the often naïve examination of the trends. The unclassified December 2009 report by Major-General Michael Flynn (2009) plots and maps the data that are marked as “Significant Events” in the WikiLeaks files. Flynn’s report is generally a pessimistic one, as he documents the growth of “kinetic events” over time, their geographic diffusion, and the sophistication of the Taliban tactics (see Biddle et al., 2010). Unlike most circumstances faced by academic researchers embarking on a new project, we found ourselves confronted by a mountain of data that offered a different set of challenges to the usual data deficiency; primarily this entailed deciding which categories of events to analyze and which methods to use. Because of the remarkable significance of these war logs to public debate about the Afghanistan conflict in the U.S. and elsewhere, we decided to dispense with the usual academic format (literature review, research questions, hypotheses, elaboration of statistical methodology) and refer the reader to our companion paper (O’Loughlin et al., 2010) for these elements. Instead, we concentrate on a (carto)graphical presentation of the violent events (46,568) in the WikiLeaks data and present the interpretation in as accessible a manner as we can. We provide references only to essential supporting sources and, believing that visual presentation is more intuitive to the reader, keep the text to the minimum necessary for guiding the reader through the figures. Obviously, different readers will bring different positions about the war to this paper, but by presenting the huge volume of records in a manageable format, we hope that readers will become more informed about America’s “longest war,” even if the analysis does not conform to their predilections. All of our data, the companion paper on the wider war that involves Pakistan, supportive materials including R data analysis code, GIS shapefiles, and more detailed and specific graphs and maps are available on our project website: <http://www.colorado.edu/ibs/johnno/afpak/>. We believe strongly in research transparency and replication of results as an essential component of the academic enterprise.

THE WIKILEAKS DATA: IDENTIFYING VIOLENT EVENTS

There are 32 columns for the 76,911 rows (events) in the database released by WikiLeaks. Various columns include report numbers and keys, reporting units, tracking numbers, updates after the event occurred, command region of Afghanistan, and classification of the event by “enemy,” “friend,” “neutral,” or “other.” Critical to our analysis of the data were the georeferences for each event in latitude/longitude.⁶ Two fields contain (often) lengthy accounts of an event and, as many commentators have indicated, list names of individuals, commanders, informants, or politicians. Eight fields portray the casualty count for each event (killed and wounded)—for civilians, Afghan national army and police (titled “Host Nation” [ANSF in our classification for Afghan National Security Forces], “enemy” (which we title “insurgents” in our text and figures), and “friendly” (ISAF forces including NATO countries and the U.S. military). Another field lists the number of insurgents detained in an operation.

The key coding decision that we faced was to select a field that allowed us to clearly distinguish the two sides in the conflict. While some have used the column (“attack on”) for such a classification, examination of the text records associated with the “enemy” and “friend”

⁶We converted the coordinates to UTM, zone 41N for analysis and mapping.

categories suggests that this column is mis-labeled and should carry the reverse label, "attack by." We chose column C, "type," which classifies events as "friendly action," "enemy action," "non-combat event," and several additional types; our reading of a sample of the text files associated with the classifications attests to its accuracy. Like many of the other columns, numerous events are labeled as "unknown" or "neutral," and we eliminated these items from our analysis. We decided to adopt a conservative (perhaps even undercounting) approach to violent events. We summarized events therefore into two categories—events in which the Taliban and their allies were the instigators (labeled "Insurgent" hereafter) and those in which the allied forces (Afghan national forces and ISAF/NATO) were the actors (labeled "ISAF/ANSF" hereafter). Of the 46,568 violent events that we selected, 9,185 actions were taken by ISAF/ANSF and the insurgents conducted 37,383.

Separating violent event types from the other events reported in the database is not a simple task. While many events, such as improvised explosive device (IED) ambush or "non-combat events," can be categorized easily, others constituted a mix of violence and non-violence. We sampled a set of each subcategory and set a threshold of approximately 10 percent violence; if a subcategory contained more than 10 percent of the events with casualties, we included it in the violence category.⁷ In this manner, we selected 61.3 percent of the events in the WikiLeaks files, while 38.7 percent are excluded from the statistical and cartographic analysis. Because 953 events were geocoded for locations outside Afghanistan, we also dropped these items. Most (660) were missing coordinates, but some were incorrect coordinates (in one case, the city of Gardez was incorrectly located in Tajikistan), while others were records of meetings in adjacent countries with political and military figures.

Table 1 presents summary statistics on events and casualties. Although we selected only 61.3 percent of all events for analysis, these counted for 92.8 percent (22,255) of all deaths and 88.1 percent (23,213) of reported wounded. This high retention of violence in the events that we analyzed should mitigate concerns about preferences for excluded subcategories. Deaths and injuries in the events that we coded as non-violent tended to be the result of accidents or criminal activities. Breaking down the totals by the main category of participants, the ratio of events with casualties is consistently in the 80–95 percent range (Table 1). The lower proportion of injury and death for civilians is the outcome of the inclusion in the reports of many non-political events (especially accidents). Nearly half (47.2 percent) of those detained by ISAF/ANSF were captured in violent events, after a firefight; if the arrests of suspected insurgents did not result in casualties, we typically allocated them to the non-violent categories.

By summing the number of events for each type of casualty, and calculating a casualty rate for events, the average is about 2 casualties for all types of events, except for insurgents who are recorded as "killed," a mean value of 4.13 per engagement. We suspect that this higher figure suffers from two errors. First, the total number of civilians killed over the six-year period (3,175) is significantly smaller than the totals reported by non-governmental organizations. More than 6,000 civilian deaths are reported for 2006–2009 by Human Rights Watch (929 in 2006 and 1,633 in 2007) and UNAMA—the United Nations Assistance Mission in Afghanistan (2,118 in 2008 and 2,412 in 2009) (Civilian, 2010). Because it appears that the majority of the civilian casualties are caused by Taliban actions, it is likely that many go unreported in the ISAF records. Second, many of the deaths reported in the high casualty figures for insurgents (14,608 killed and 1,709 injured) are either uncertified estimates from reports of pilots engaged in the confusing air-ground fighting or partial body counts. The disproportionate presence of round numbers (10, 15, 20, 30, 40, 50, 100) in the casualty counts

⁷The full listing of violent and non-violent event subcategories is available from the project website.

Table 1. Summary of WikiLeaks Data Indicating Combatants, Casualties, and Events, 2004–2009

Category	Casualties	All Afghanistan events		Selected violent events in Afghanistan		Proportion with casualties
		<i>N</i> of casualties	<i>N</i> of events	<i>N</i> of casualties	<i>N</i> of events	
ANSF	Killed	3,754	1,629	3,449	1,487	91.9
	Wounded	8,394	3,510	7,482	3,145	89.1
ISAF	Killed	1,140	703	1,023	644	89.7
	Wounded	7,239	3,662	6,435	3,229	88.9
Civilians	Killed	3,943	1,425	3,175	1,152	80.5
	Wounded	8,916	2,943	7,587	2,521	85.1
Insurgents	Killed	15,152	3,704	14,608	3,532	96.4
	Wounded	1,786	793	1,709	743	95.7
	Detained	7,272	2,349	3,429	966	47.2
Total	Killed	23,989		22,255		92.8
	Wounded	26,335		23,213		88.1
Total events			75,958 ^a		46,568	61.3

^aThis number excludes the 953 events that fall outside Afghanistan's border or do not have a latitude/longitude tag.

lends credence to this assumption of inaccuracy. As Shachtman (2010b) writes, while the WikiLeaks data provide a reasonable synopsis account of the fighting, they do not capture all the nuances of the battles in their short reports. Civilian casualties from insurgent attacks and bombs began to increase in July 2009 (as evident in war logs to the end of 2009) and subsequently increased by 50 percent in 2010 (Nordland, 2010), which is consistent with the overall increase in civilian casualties in 2010 (*ibid.*).

The comparison of the geographic and temporal distribution of the included (violent) and excluded (non-violent) events clarifies the changing nature of the Afghan war. In Figure 1, we map the two sets of events by district (subunits of the provinces, totaling 328) and in Figure 2, we present the monthly comparison of violent and non-violent events.⁸ The geographic distributions of violent and excluded events are similar, with concentrations in the war's loci of violence (the provinces of Helmand and Kandahar in the south, and Kunar and Paktia provinces in the east along the Pakistan border). The comparative patterns do not suggest any selective regional bias in the distinction that we make between violence and non-violence. More broadly, plots and count analysis of the total events by command region (6 in Afghanistan) and by year by Drew Conway (2010) suggest that there is no evident selection bias in the data that were released.

The monthly plots of both series, though, clearly indicate a different story. Beginning in early 2008, violent events skyrocket while the reports of non-violent events continue at approximately the same rate as in 2006 before trending upward in 2009. What had previously

⁸The ACLED portion of this graph is discussed later.

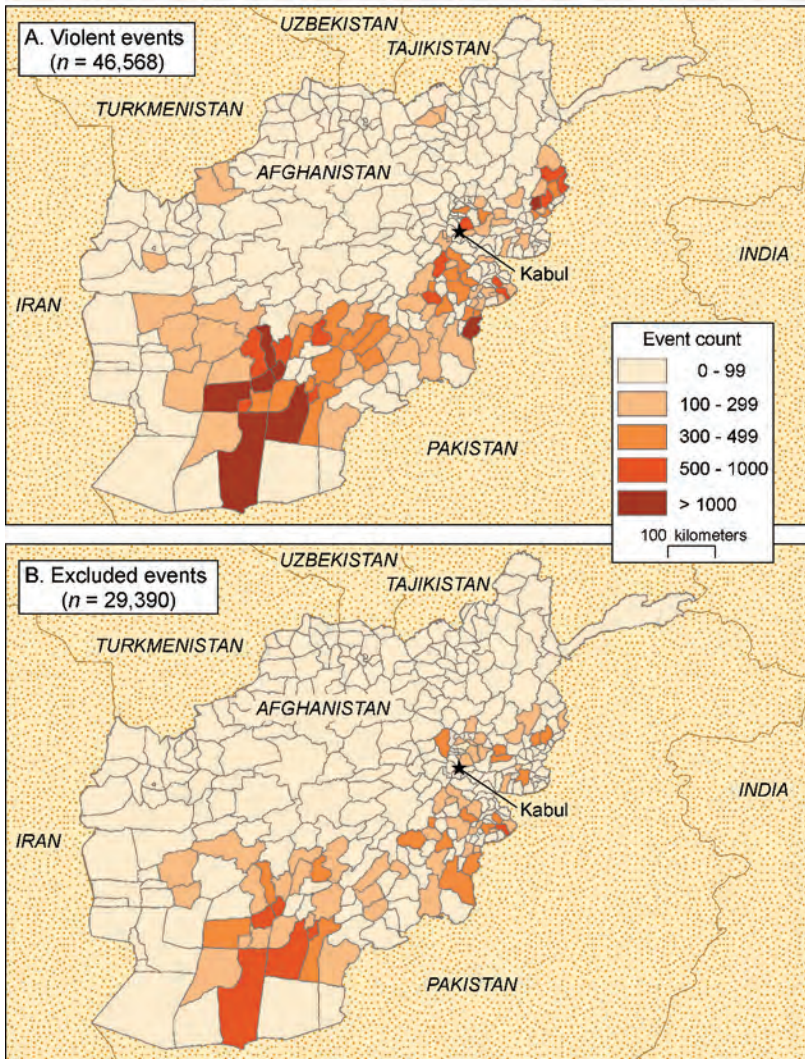


Fig. 1. Geographic distribution of violent and non-violent events 2004–2009 in the WikiLeaks data by district in Afghanistan.

been a balance in the logs between violent and non-violent events disappears in the last two years as the proportion of violent events more than triples. In 2004, for example, there are only 19 violent events counted for the ISAF/ANSF category compared to 537 that we placed in the “non-violent” category, all consisting of “cache found” or “patrol.” As the ISAF mission in Afghanistan expanded in scope, troop commitments, and geographic extent, the ratio of violent to non-events changed significantly so that by 2009, it was 20,804 violent events to 7,552 non-violent events.

Beginning about May 2006, there is an evident separation of the two trend lines for ISAF/ANSF and insurgent-initiated events (Fig. 3). While the Afghan government forces and their ISAF allies display hardly any increase in their activity over time, the insurgent events

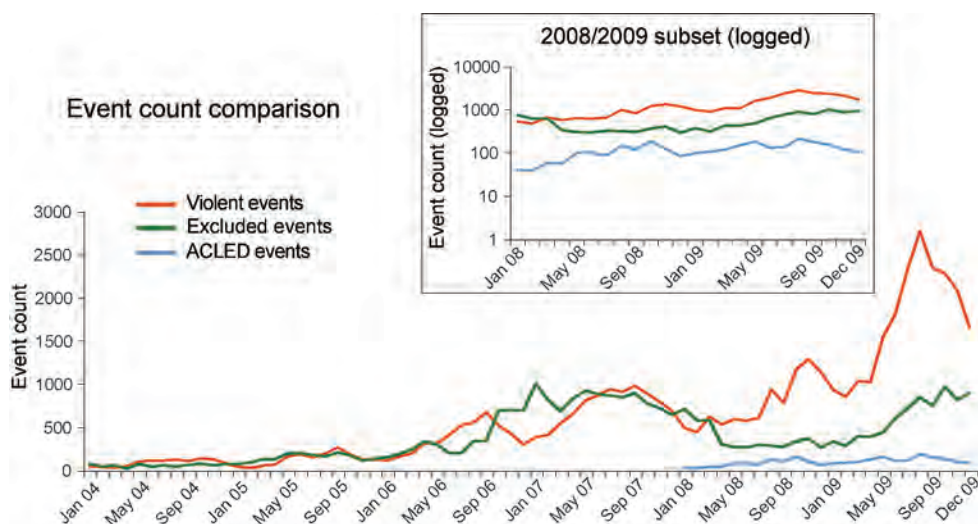


Fig. 2. Monthly distribution (2004–2009) of violent and non-violent events in the WikiLeaks data.

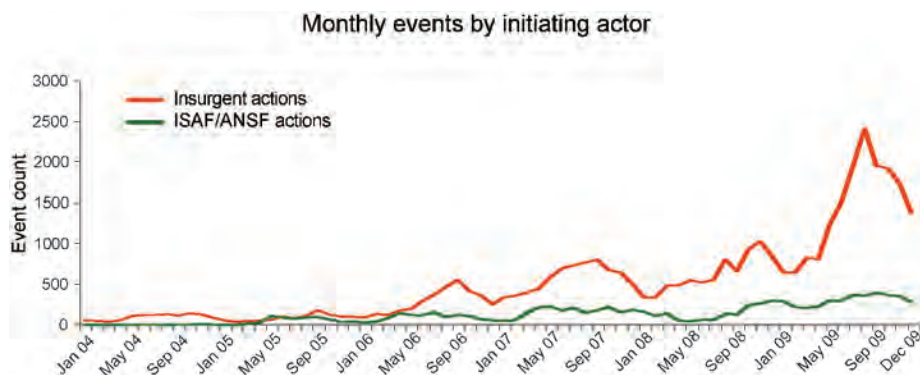


Fig. 3. Monthly distribution of violent events in the WikiLeaks data by actor (ISAF/ANSF or insurgents), 2004–2009.

increase dramatically from 200 in May 2006 to a peak of over 2,500 in August 2009 (the time of the Afghan presidential election campaign). While one might quibble with our categorization of the reports into violence and non-violence, there is little doubt from the trend lines in Figure 3 that the Taliban attacks surged dramatically in the last half of the data series, with summer peaks of activity clearly visible for each year after 2005.⁹

Although we cast doubts earlier on the accuracy of the reports in counting casualties, we present the monthly casualty data in Figure 4. The data on insurgent casualties stand in sharp relief to the other three series. As we noted, the insurgent casualty numbers tend to rely

⁹Flynn's 2009 plots of violence show that the Taliban attacks decreased during the Ramadan observance periods in each year.

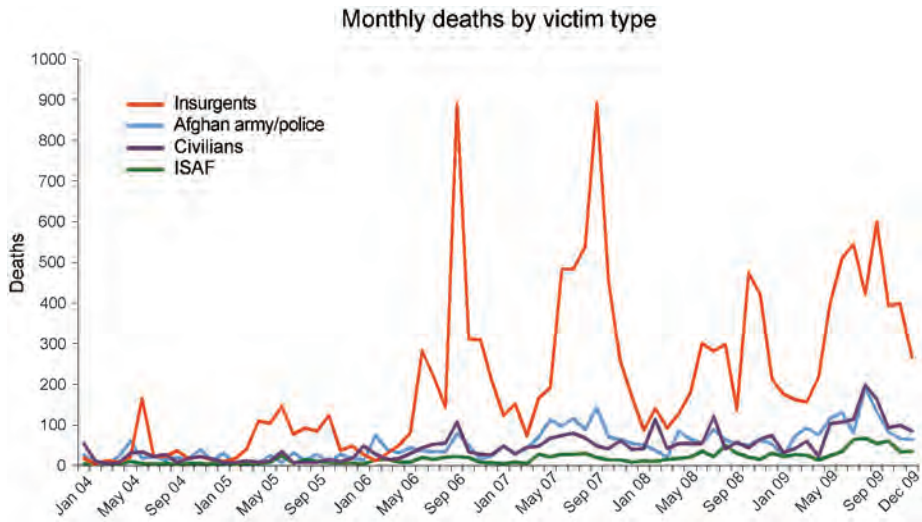


Fig. 4. Monthly deaths by victim type, Afghanistan 2004–2009, as reported in WikiLeaks data.

on imprecise estimates from ISAF and ANSF personnel in the field, who often have no reliable evidence, that is, corpses easily identifiable as civilian or insurgent. During the Vietnam War, the U.S. military command published weekly body counts for the Vietcong and North Vietnamese to indicate that the U.S. was winning the war, but as the war dragged on, the numbers became derided as fictional. Because the evening news broadcasts also announced the U.S. and South Vietnamese body counts, the overall impact of this practice over the many years of the war was both to dull the impact of the huge numbers reported and to draw attention to the substantial U.S. losses. As a result of this experience, the Pentagon no longer provides such public estimates; for example, no figures for Iraqi casualties were given for the Gulf war in 1991 or for the Iraq war after 2003 (Boettcher and Cobb, 2005). Though not disseminated to the media, it is now clear that the U.S. military continues to collect (unreliable) casualty data in Afghanistan. The four peaks in Figure 4 for insurgents killed correspond to major advances by ISAF. In late summer 2006, the peak corresponds to “Operation Medusa” in Kandahar Province, part of an expanded operation into southern Afghanistan. A year later, severe fighting near Kandahar was part of an effort to remove those insurgents who had returned to the area; the war logs report hundreds of Taliban killed in separate incidents during this period, including 181 in the action that ended on September 9, 2007. The smaller peaks in October 2008 and August 2009 correspond to augmented U.S. airstrikes in 2008 and the violent Presidential election period of August 2009 that also saw a sizeable jump in civilian and host-nation (Afghan army and police) deaths. We portray these numbers to show the yearly ebb and flow of the conflict, and that despite the lowered peaks for insurgents killed in 2008 and 2009, the overall casualty figures are rising for other participants in the conflict.

COMPARING WAR STUDIES—MILITARY LOGS AND MEDIA REPORTS OF THE AFGHANISTAN WAR

We accept the criticism by Shachtman (2010b) and spokespersons for the Pentagon and the U.S. Government that the WikiLeaks data do not constitute a complete picture of the

Afghanistan war, although governmental officials never portrayed the war logs as factually inaccurate. Indeed, no reporting mechanism is able to account for all developments in an even-handed way, inasmuch as both sides are unable to access all information for the other side. In a similar manner, academic efforts to build conflict databases are hampered by haphazard sources that replicate the biases in the official data (on one side) or are silent about the activities of the other (typically rebel) side. To compensate for this disparity, academics sometimes turn to rebel websites, although these sources also tend to overstate victorious engagements and government losses while minimizing their own setbacks. In the past two decades, civil war and political violence researchers have attempted to understand the dynamics of conflicts that tend to be highly localized, frequently personal, characteristically complex in the ideologies and motivations of participants, and rarely end in a clear defeat for one side. While historians and some political scientists have tended to examine singular cases, the recent turn to large-*N* studies of multiple wars and locations is motivated by a sense that these complex dynamics can be modeled and forecast.¹⁰

The WikiLeaks and other conflict-events datasets should correspond for one primary reason, violent events extracted from the complete file of released data are noted in local and (sometimes) international media. As such, coders read the account and record it according to established protocols. The major concerns of inconsistency, where it arises, revolve around the types of violence that are covered in the media, and the types excluded. Several factors may lead media to ignore events. When otherwise media-worthy violent events (e.g., a politically motivated assassination) take place in a clandestine manner, they are not likely to be reported. The WikiLeaks data, as an example of this, include “sniper operations” as a category. Additionally, researchers involved in large-scale data collection must recognize the fundamental distinction between active and passive monitoring of social activity. Active monitoring entails a decision to record incidents by a participant in the activity, as events transpire, for use in future analysis. Possible disparities that emerge between each variety of collection approach concern the limitations of record-keeping; in the world of conflict-event coding and analysis, the media are customarily the gate-keepers.

One of the major databases now being assembled for the analysis of civil war dynamics is ACLED (Armed Conflict Location Event Data) project (Raleigh et al., 2010a). We analyzed a forthcoming dataset from the ACLED project on the Afghanistan-Pakistan conflicts for 2008–2009 in our companion paper (O’Loughlin et al., 2010). Like others using such databases, we rely on the care of the (student) coders in collecting the relevant media reports of the conflicts, georeferencing the places of violence, and in separating political violence from reports of non-violent events and of other (criminal) violence. Though researchers have some level of confidence in the event data (why analyze them otherwise?), there remains uncertainty about their representativeness and completeness. What the release of the WikiLeaks data allows is the first formal comparison of event data collected from media reports to the internal military reports that document the day-to-day war developments.

In Figure 2, we present three timelines; the one for the ACLED data covers only 2008–2009. The other two lines are the respective violent and non-violent trends discussed earlier. The appropriate comparison of the ACLED trend is to the violent events line and, to ease the comparison, we also plot the logged values¹¹ for the past two years in the inset. Not surprisingly, given our knowledge of different monitoring frameworks and the limitations of media reporting, the ACLED values are a small fraction of the WikiLeaks event numbers,

¹⁰For a literature review, see O’Loughlin and Witmer (2010) and Raleigh et al. (2010b).

¹¹Logged values permit an easier assessment of values that differ widely in magnitude.

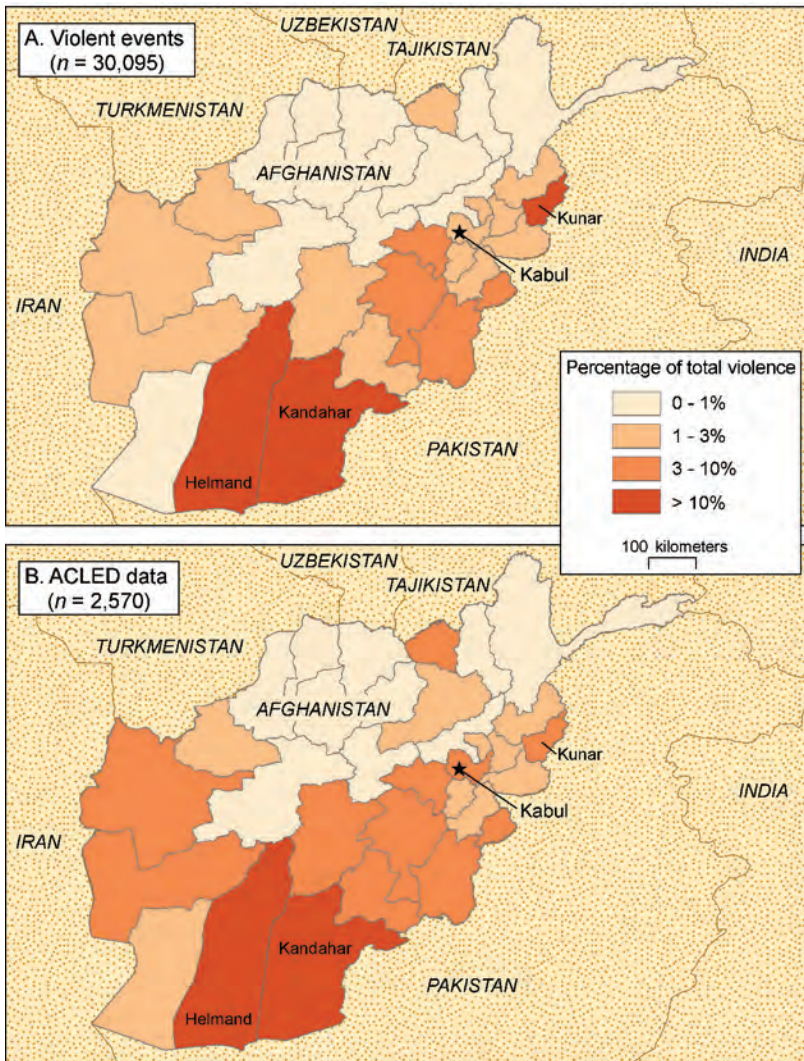


Fig. 5. Comparison of the ACLED and WikiLeaks violent-event data for 2008–2009 in Afghanistan (provincial ratios of respective country totals). Note that the WikiLeaks data on the map only covers the 2008–2009 period (and not the full 2004–2009 period that the war logs record).

only one-fiftieth to one-twentieth of their absolute values. Typically, only important military engagements with casualties (some killed or wounded) make it to the open-access conflict-events databases through the filters of the military public relations offices, journalistic norms and proclivities, and coder preferences. However, the logged data association demonstrates the utility of the ACLED events data; that line closely parallels the higher value for the WikiLeaks data in its general upward trend and its seasonal rise-fall (higher in summer and lower in winter).

A second, more demanding, comparison of the WikiLeaks and ACLED data is a map evaluation where the respective data are shown by the 32 provinces of Afghanistan. For this

exercise, all violent events for each province are summed for 2008 and 2009, and mapped as a ratio of a two-year total of all violent events across the country. Happily, the maps in Figure 5 confirm the consistency of the two datasets as the geographic distributions are highly correlated ($r = 0.9$). In the WikiLeaks violent events, Helmand Province accounts for 30.3 percent, Kandahar 11.2 percent, and Kunar 11 percent, while of the remaining 29 provinces, only Paktika (along the Pakistan border) contributes more than 5 percent. The ACLED conflict-events data are less concentrated in these provinces (Helmand 20 percent, Kandahar 13.2, Kunar 3.6, and Paktika 3.8 percent). With the exception of Farah and Herat provinces (bordering Iran in the west of the country), all of the remaining percentages are within two percentage points of each other.

While this comparison for conflict in Afghanistan does not assure that other databases drawn from media sources are equally as reflective of the civil war occurrences, it nevertheless offers an important confirmation of the value of this type of event collection and analysis. In our companion paper, we argued for a diffusion model over the border from Afghanistan to understand the nature of the evolving war in Afghanistan-Pakistan over the past two years on the basis of the ACLED data. Our comparison of the ACLED data for Afghanistan to reports in the war logs gives us further confidence in the conclusions of the wider war study.

THE GEOGRAPHIC DISTRIBUTION OF VIOLENCE IN AFGHANISTAN 2004–2009

To visualize the risk of experiencing violence over time, we use several mapping and graphical techniques to explore the spatial distribution of the WikiLeaks conflict-events data. We first map both the annual distribution of violence using an intensity measure and spatial distribution statistics before dissecting these trends by temporal (quarterly) graphics that present patterns by key geographical features.

The first intensity measure overlays a 20 km \times 20 km set of grid cells on Afghanistan and calculates the intensity of the violence based on the number of violent events in that cell and surrounding cells. Figure 6 shows the intensity estimates for all violent events using a three-dimensional surface where peaks in the surface (oranges and reds) indicate the highest intensity of violence. The maps in this figure have been rotated 180° to allow visualization of peaks hidden in the conventional “north-is-up” orientation. This figure shows in dramatic form the sharp increase in violence in 2009, especially for the Korangal valley (Kunar Province on the Pakistan border), for Kandahar, and for Lashkar Gah in the south (for locations, see Fig. 1 of O’Loughlin et al., 2010 on p. 439 of this issue). The Korangal valley, just west of Asadabad and north of Jalalabad (shown in Fig. 1 in O’Loughlin et al., 2010), emerges in 2005 as a center of fighting (the year U.S. troops arrived in that area to remove Al-Qaeda and the Taliban and their local allies). The fighting intensified over time until April 14, 2010, when the U.S. announced its withdrawal from the valley and the closure of the Korangal outpost in what many observers viewed as an ill-conceived mission and a signal of defeat for ISAF/ANSF (Jaffe, 2010; Rubin, 2010). The main impression of the 3-D maps is the dramatic upsurge in violence in many locations after 2006, what the U.S. military refers to as “kinetic events” (Shachtman, 2010a).

We further examine the spatial distribution of violence over time by plotting the mean center and standard ellipse by year for each side in the conflict. These two-dimensional geographic measures are analogous to the one-dimensional mean and standard deviation statistics. The mean center is calculated by averaging the x-coordinates and y-coordinates of events separately to yield the locational center of violence, while the standard ellipse is calculated

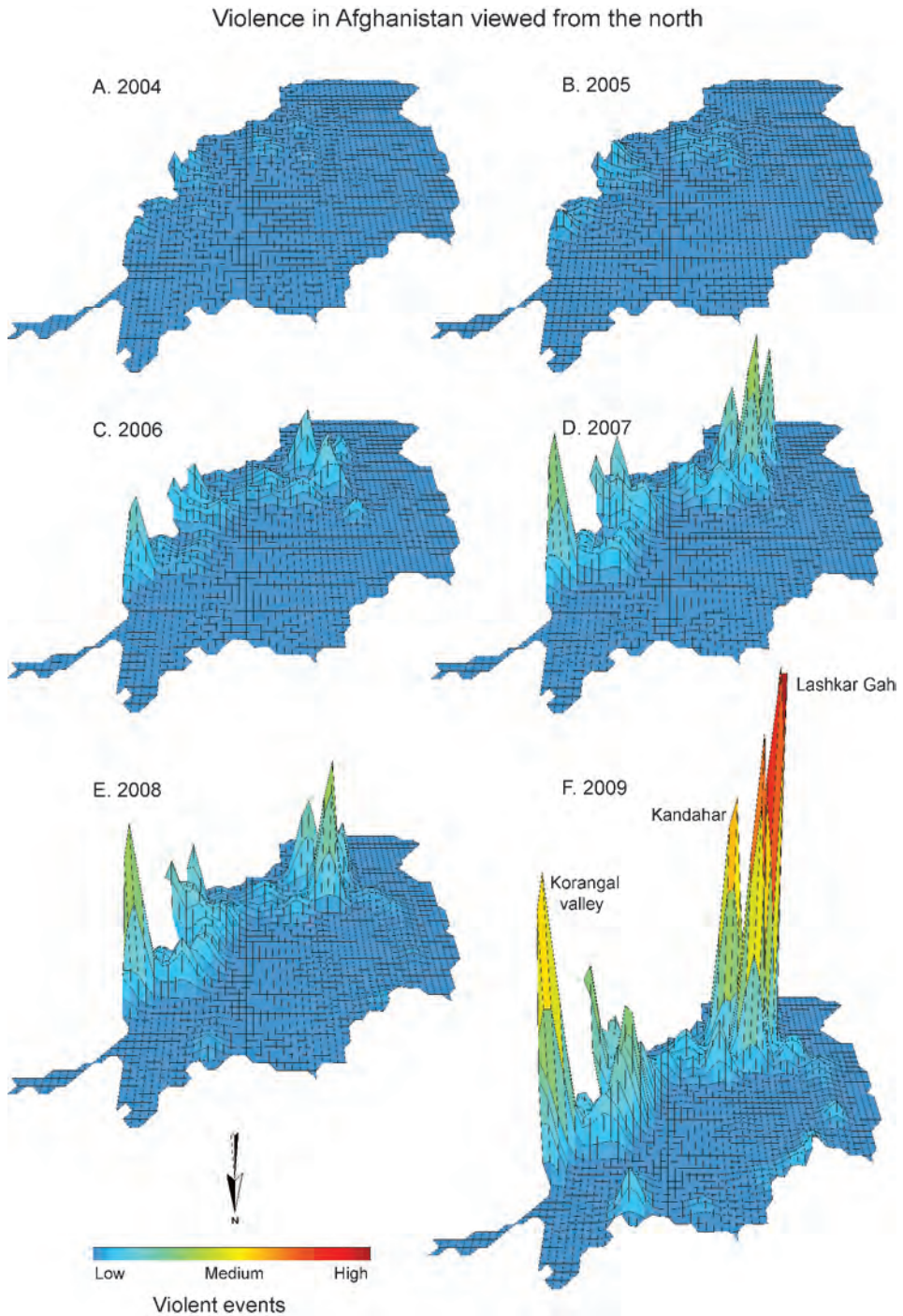


Fig. 6. Three-dimensional map of total violence in Afghanistan by 20 km grid cells and by year. The view from the north is from a view angle perspective of 30° . The density estimate calculation uses an isotropic Gaussian smoothing kernel with a standard deviation of 20 km to estimate the intensity of the point process that generated the observed conflict-event data (see Diggle, 1985).

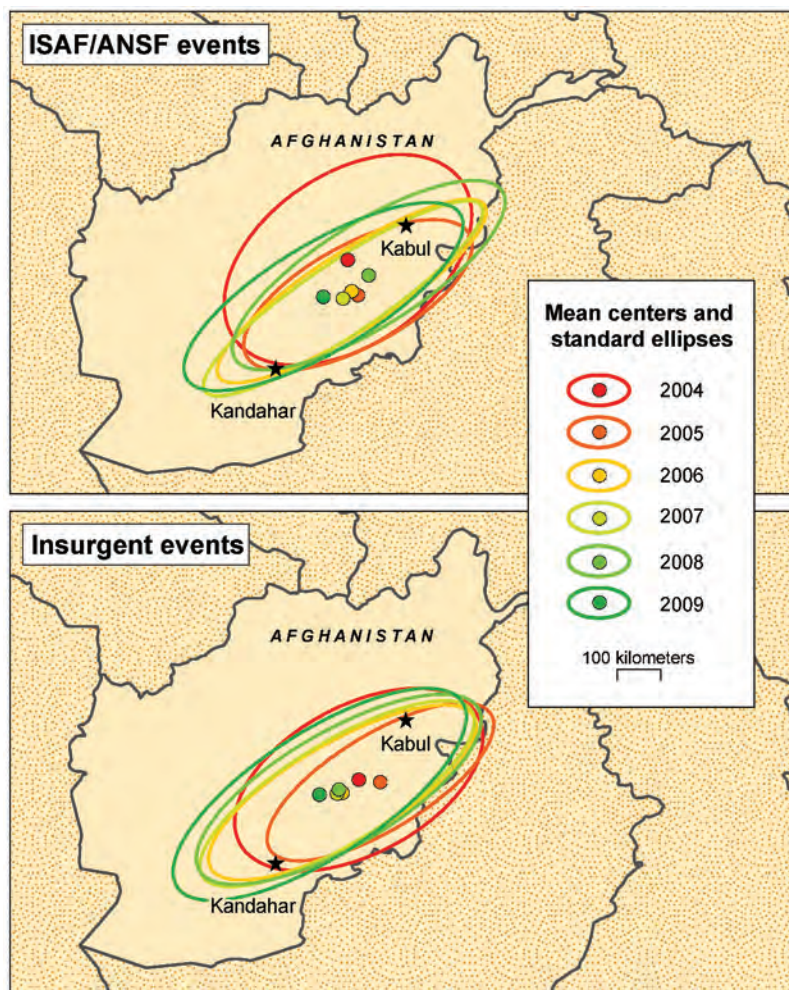


Fig. 7. Mean centers and standard ellipses for violence (ISAF/ANSF and insurgents) over 2004–2009.

so that two-thirds of the violence falls within the ellipse. Figure 7 shows the distributions for the two sides in the conflict, ISAF/ANSF and the insurgents. For 2004, with only 19 ISAF/ANSF events, the distributional statistics were more easily pulled to the center of Afghanistan. For both ISAF/ANSF and insurgent violence, the ellipses stretch from the Korangal valley to Lashkar Gah, with a clear shift over time toward the southwest, reflecting the sizeable increase in activity in Kandahar and Helmand provinces. In general, the ellipses are aligned along the border with Pakistan. Shifts over time in their mean centers, their size, and their orientation show lengthening and movement, especially for the insurgent events. Put simply, if the insurgency were being contained, these maps typically would show shrinking ellipses.

The intensity maps and distributional statistics provide effective visualizations of the overall violence trends, but they lack trend information by geographical categories. The next series of bar plots explores the quarterly violence trends for each side for several geographical factors that have been suggested as important for understanding the evolution of this war:

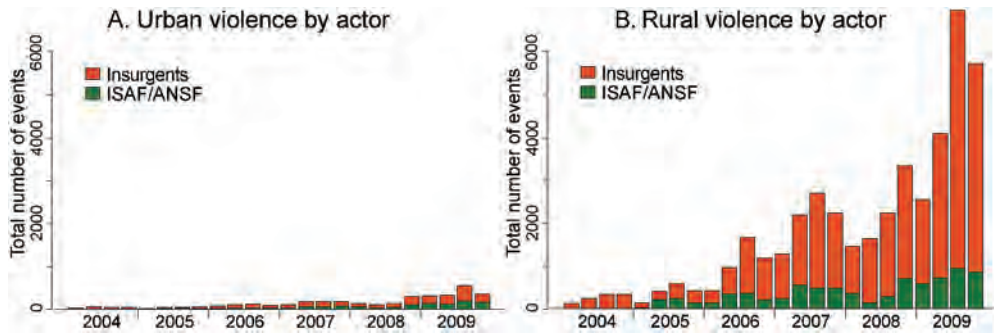


Fig. 8. Distribution of violence by actor and urban-rural locations in Afghanistan, 2004–2009 (by quarter).

rural/urban (Fig. 8), Pashtun/non-Pashtun (Fig. 9), proximity to the Pakistani border (Fig. 10), and ruggedness of terrain (Fig. 11). We include only a brief description of the data and methods here; full details can be found in the companion ACLED data article (O’Loughlin et al., 2010).

The rural/urban spatial boundaries were defined using nighttime lights imagery for 2008 (NOAA/USAF, 2009). We designated a low light-intensity threshold in defining urban areas to account for limited electrification, but only 0.5 percent of all land in Afghanistan is categorized as urban. Conflict-event data were then designated as either urban or rural and summarized in Figure 8. The predominantly rural landscape that dominates Afghanistan clearly shows in the distribution of violent events, with ISAF/ANSF forces favoring urban areas (12.8 percent of violent events) more than the insurgents (6.3 percent). Insurgency in Afghanistan is distinctly a rural phenomenon with the ratio of urban to rural context changing little over time. While most cities remain under the control of the Karzai regime, the countryside has seen an explosion in violence in a pattern that replicates the civil war experiences of dozens of countries in the past century.

We also examine the trends in violence based on the Pashtun ethnicity—the ethnic group that is often implicated in the violence and from whose heartland in southern Afghanistan the Taliban first emerged in the early 1990s as the opposition to the Soviet-installed regime then in power in Kabul. Pashtun-majority areas are designated using the Geo-Referencing of Ethnic Groups (GREG) data, a Soviet-era geographic dataset that was recently spatially digitized (Weidmann et al., 2010). Results from this aggregation (Fig. 9) show both insurgent and ISAF/ANSF violence concentrated from the start in Pashtun-majority areas (east and south of the country), with some of the 2009 increase in violence spilling into non-Pashtun majority areas (e.g., the area in and around Lashkar Gah is predominantly non-Pashtun). This trend has been noted also in an unclassified report for the U.S. military (Flynn, 2009) and should the Taliban and their allies be able to sustain the insurgency beyond their traditional ethnic havens, the implications for the Karzai regime and their Western allies will be dramatic. Before the U.S. invasion in October 2001, the Taliban were still locked in a civil war with militias based in non-Pashtun ethnic regions, especially with the Northern Alliance. If the Taliban can “nationalize” their opposition to the Karzai government and establish their credentials as a broad non-ethnic movement, a key assumption held by observers since 2001 (that the Taliban can only gain support among Pashtuns) will have been turned on its head.

To determine the role that the Afghanistan-Pakistan border plays in the spatial distribution of violence, we generated a 100 km buffer around the border, and use it to designate

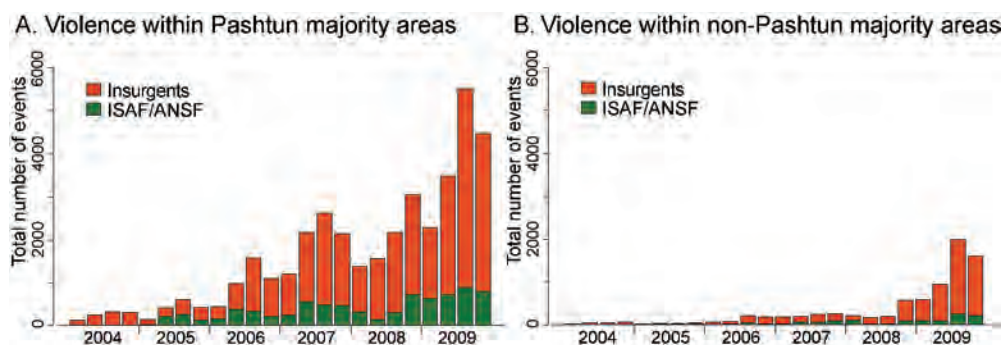


Fig. 9. Distribution of violence by actor and Pashtun/non-Pashtun majority locations in Afghanistan, 2004–2009 (by quarter).

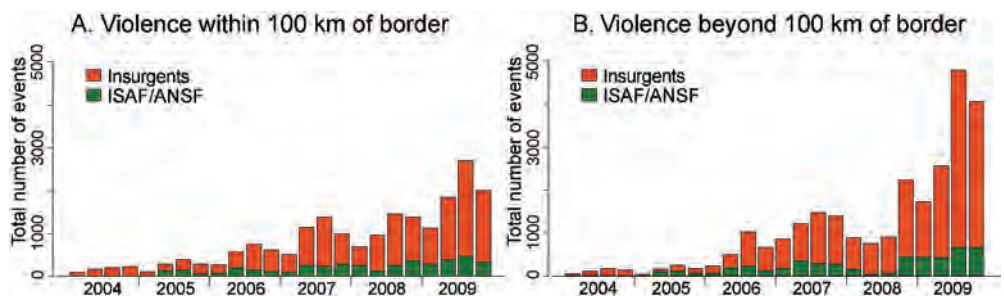


Fig. 10. Distribution of violence by actor and distance category relative to the Afghan-Pakistan border (within or greater than 100 km) in Afghanistan, 2004–2009 (by quarter).

violence as “near” or “far” from the border. The importance of the border for the insurgency is well established in the Afghanistan case (especially the use of safe havens in the Pashtun regions on the other side of the border in Pakistan) and in the literature on civil wars. Violence is similarly distributed with respect to the 100 km border buffer until 2009, when violence beyond 100 km of the border increased substantially, especially that related to insurgent activity (Fig. 10). The 100 km border buffer represents about one-quarter (26.9 percent) of the total area of Afghanistan, so if the violence were distributed randomly across the country, we would expect many fewer events in the “near” border category. Until the last two quarters of 2009 (a period coinciding with the Presidential election campaign and its aftermath), 46.8 percent of the violence was in border areas. This percentage dropped to 34.7 during the last two quarters of 2009. Much of the increase in late 2009 violence away from the border occurred in and around Lashkar Gah, which lies beyond the 100 km buffer (Kandahar and Kabul are just within the 100 km distance). This graph is a further illustration of the nationalization of the insurgency away from the traditional Taliban strongholds.

In our last graph of the series (Fig. 11), we explore the quarterly trends in violence according to terrain ruggedness, a topic that has received a good deal of attention in the conflict studies literature. This categorization enables us to address the role of inaccessible terrain as a safe-haven for insurgents to retreat to and strike from. Unlike O’Loughlin et al. (2010), we use just two terrain categories (flat and hilly/steep), because there were few differences in the earlier study when looking at the hilly and steep categories separately. To define these categories, we

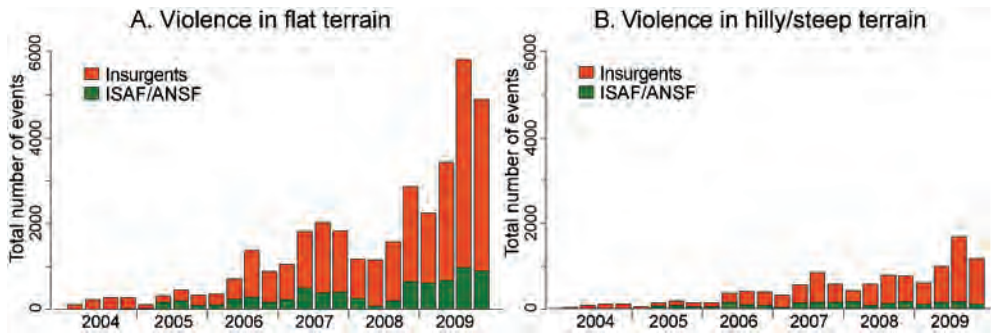


Fig. 11. Distribution of violence by actor and flat and hilly/steep terrain in Afghanistan, 2004–2009 (by quarter).

used slope data derived from the Shuttle Radar Topography Mission (USGS, 2004) and designated violent events whose nearby pixels had a mean of less than 4° in slope as flat, and all other rougher areas as hilly/steep.¹² Most of the violence occurs in flatter areas of Afghanistan, especially ISAF/ANSF actions in the later years (e.g., for 2009, 85 percent of ISAF/ANSF events and 77 percent of insurgent events were in flat terrain). Even in generally mountainous regions, most of the fighting takes place on or near roads that traverse the valleys.

Shachtman (2010a) and other commentators have noticed the tendency for violence to exhibit a circular pattern in Afghanistan that corresponds to the “ring road” that links most of the major regional capitals (see Fig. 1 of O’Loughlin et al., 2010 on p. 439 of this issue). We tested this notion by constructing a 5 km buffer around all the roads in the country; that buffer constitutes 48.3 percent of land area of which the “ring road” buffer accounts for 7.1 percent. ISAF/ANSF actions near the ring road total 11.6 percent of all of their actions over the six-year period, while insurgent actions account for 18 percent of their total. It is clear from this analysis that the ring road has become a disproportionately important target for IED actions as it remains a major transport artery for government and allied forces. It should also be remembered, however, that almost all violence takes place near a road, 88.5 percent for ISAF/ANSF and 85.9 percent for the insurgents.

IDENTIFYING CLUSTERS OF VIOLENCE IN AFGHANISTAN, 2004–2009

A standard measurement of spatial clustering in geographic analysis is the Getis-Ord statistic (Getis and Ord, 1992; Ord and Getis, 1995). This measurement of geographic distributions is one of many Local Indicators of Spatial Association, and is widely used to identify disease clusters and crime hotspots. Often referred to as “hotspot” analysis, the procedure identifies areas where certain aggregated count values are more highly concentrated than would be expected if the values were distributed randomly. There may be little reason to expect a completely random distribution of violence throughout Afghanistan, but we apply the procedure here to identify specific clusters of high values (hotspots) and low values (coldspots).¹³ In this study, the focus is on hotspots.

¹²Nearby pixels were selected using a 2.5 km buffer around each point. The mean slope for these pixels was calculated using Hawth’s Zonal Statistics ++ tool (<http://www.spatial ecology.com/htools/zonalstats.php>).

¹³For a technical description of the method, including formulas and commands, see www.colorado.edu/ibs/johno/apak.

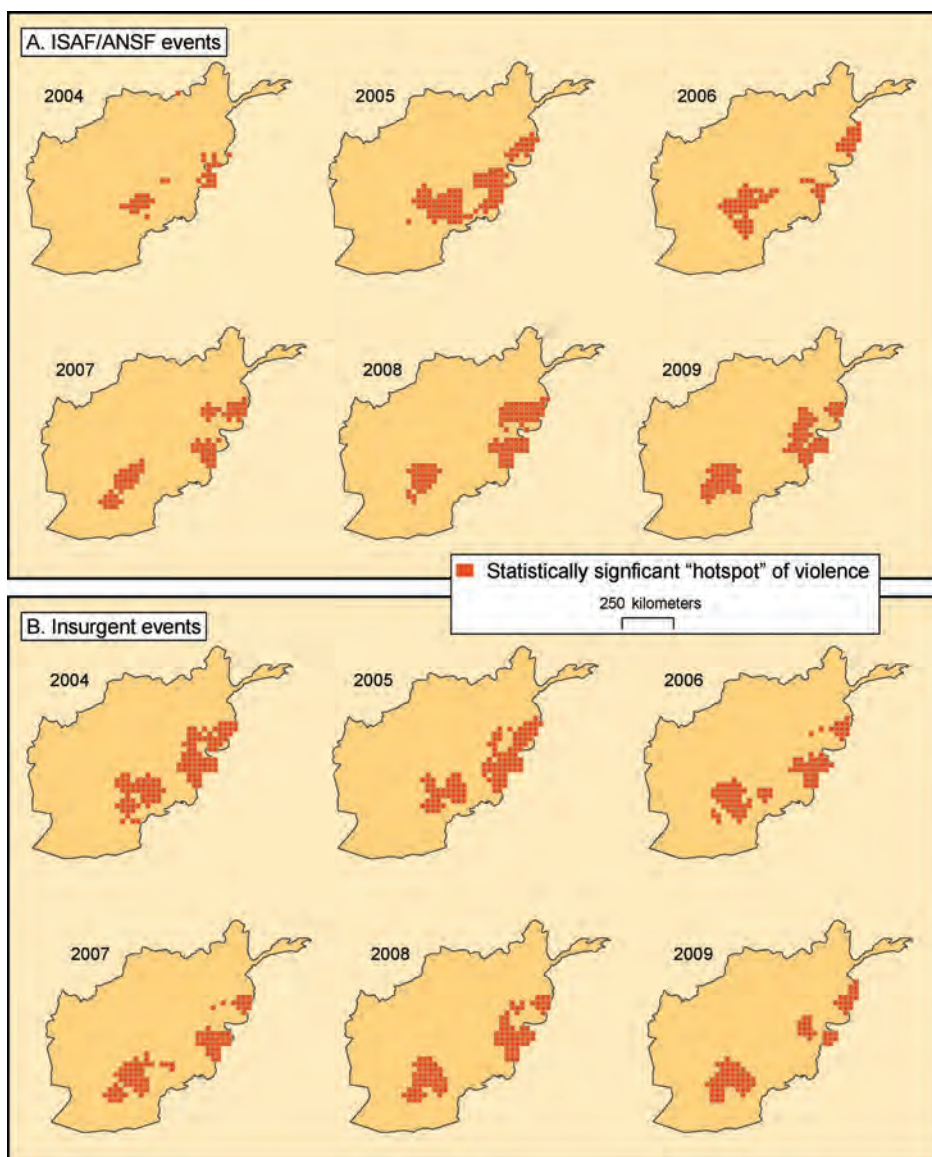


Fig. 12. Hotspots of violence in Afghanistan by year and actor.

Our unit of analysis is a 25 km grid cell ($n = 1,157$). We retain for analysis those cells that fall within or intersect Afghanistan's national border. The data delimiting Afghanistan's national border are from Global Administrative Areas (2010), version 1.0. and the Afghanistan Information Management Services (<http://www.aims.org.af/index.aspx>). Using the hotspot method, we analyze the total count of violent events that occurred within a grid cell during each year (2004–2009).¹⁴ We use a 50 km distance threshold from the center point of each grid cell as our definition of "neighbor," and this has the effect of including second-order

¹⁴We repeated the analysis for IED attacks and the maps from that analysis are available at the project website.

Table 2. Distribution of Afghanistan's Hotspots of Violent Activity by 25 km Grid Cell and Year (percentage of total grid cells)

Year	ISAF/ANSF	Insurgents
2004	3.3	10.9
2005	9.8	10.1
2006	7.1	8.4
2007	7.1	8.1
2008	8.9	9.3
2009	8.9	7.6

neighboring observations (the neighbor of a neighbor). To quantify our certainty that these clusters do not occur by chance, the procedure compares the observed distribution (actual violence in the data) to simulated values that are randomly assigned across the observation units. We retain statistically significant clusters of high levels of conflict (Z -score > 2).¹⁵

In Figure 12 and Table 2, we present the results of the Getis-Ord G_i^* analysis of ISAF/ANSF and insurgent activity. The eastern region of the country along Pakistan's border (including Asadabad, Gardez, Jalalabad, Kabul, Khost, Nuristan, and other cities) is an enlarged hotspot of both ISAF/ANSF and insurgent activity. A second hotspot is visible for south-central areas (including Lashkar Gah, Kandahar, Tirin Kot, Qalat, and other centers). Under more careful scrutiny, however, particular spatial relationships become evident. Over time, clusters of ISAF/ANSF violence migrate to the west and south, engulfing an area of the country (Helmand and Kandahar provinces) that was not as heavily affected during the early years of Operation Enduring Freedom. Media reports have emphasized the importance of the Afghanistan-Pakistan border area to the war's outcome and the WikiLeaks data certainly confirm this supposition. Table 2 presents the percentage of the country that is covered by a hotspot for each event type and each year.

One distinct trend in the clusters of violence is a convergence of event intensity surrounding key locations in the war. In 2007, ISAF/ANSF activity can be seen as concentrated in five distinct clusters or hotspots. By 2009, after shifts in policy at the onset of the Obama Administration, the distribution of violence constitutes three areas of statistically significant clustering. Additionally, for both event types, clusters in the south shifted farther west as time passed, which was largely associated with the U.S. effort (especially after Obama took office) to clear Helmand and Kandahar provinces of insurgents (the pattern is also visible in the ellipses). Table 2 shows further evidence of this congregating trend over time in the coverage of Getis-Ord G_i^* clusters across the country since 2004. For instance, insurgent activity was clustered on 10.9 percent of Afghanistan's territory during 2004, but only 7.6 percent in 2009.

A more sophisticated method for identifying temporal and spatial patterns in georeferenced data is a space-time scan statistic (Kulldorff et al., 2005; Kulldorff and Information Management Services, 2009). This method has the advantage of explicitly incorporating the temporal dimension as well as using the coordinates for each event within the WikiLeaks data. Though SaTScan can accommodate a control for the underlying population within a

¹⁵This is the default for the method in ArcGIS 9.3. To keep the paper accessible and to facilitate replication, we have not adjusted this threshold.

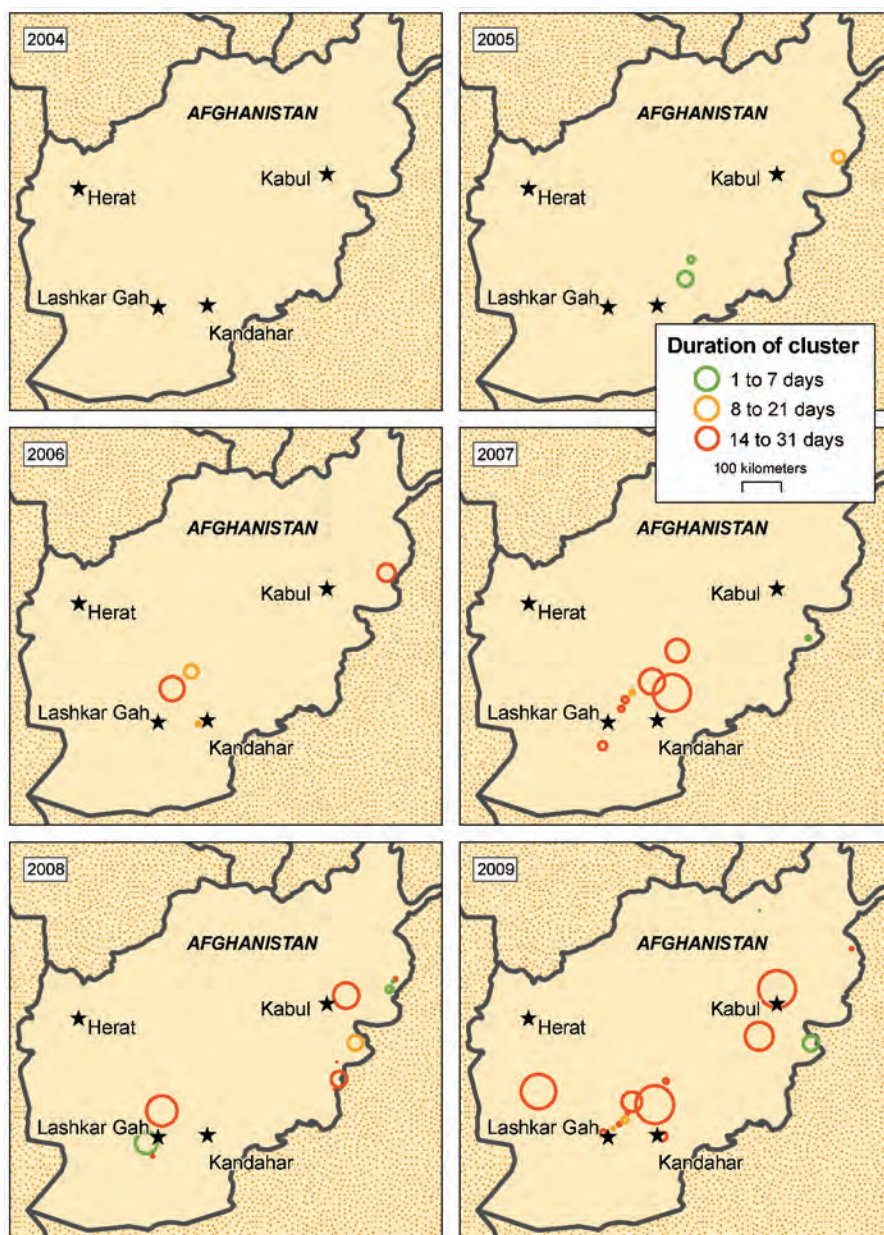


Fig. 13. Space-time scan statistic clusters of ISAF/ANSF activity, Afghanistan, 2004–2009.

study area, we decided not to “control” for population distribution in the absence of reliable local-area census data. The scan statistic methodology moves an analytical boundary, or window, through time or across space, or simultaneously in time-space mode. In the space-time version, the window is a three-dimensional cylinder with user-defined thresholds for length of time (height of cylinder) and width across space (radius of cylinder). In our analysis, the maximum cluster size is defined as 50 km, and the maximum temporal window is one month. The

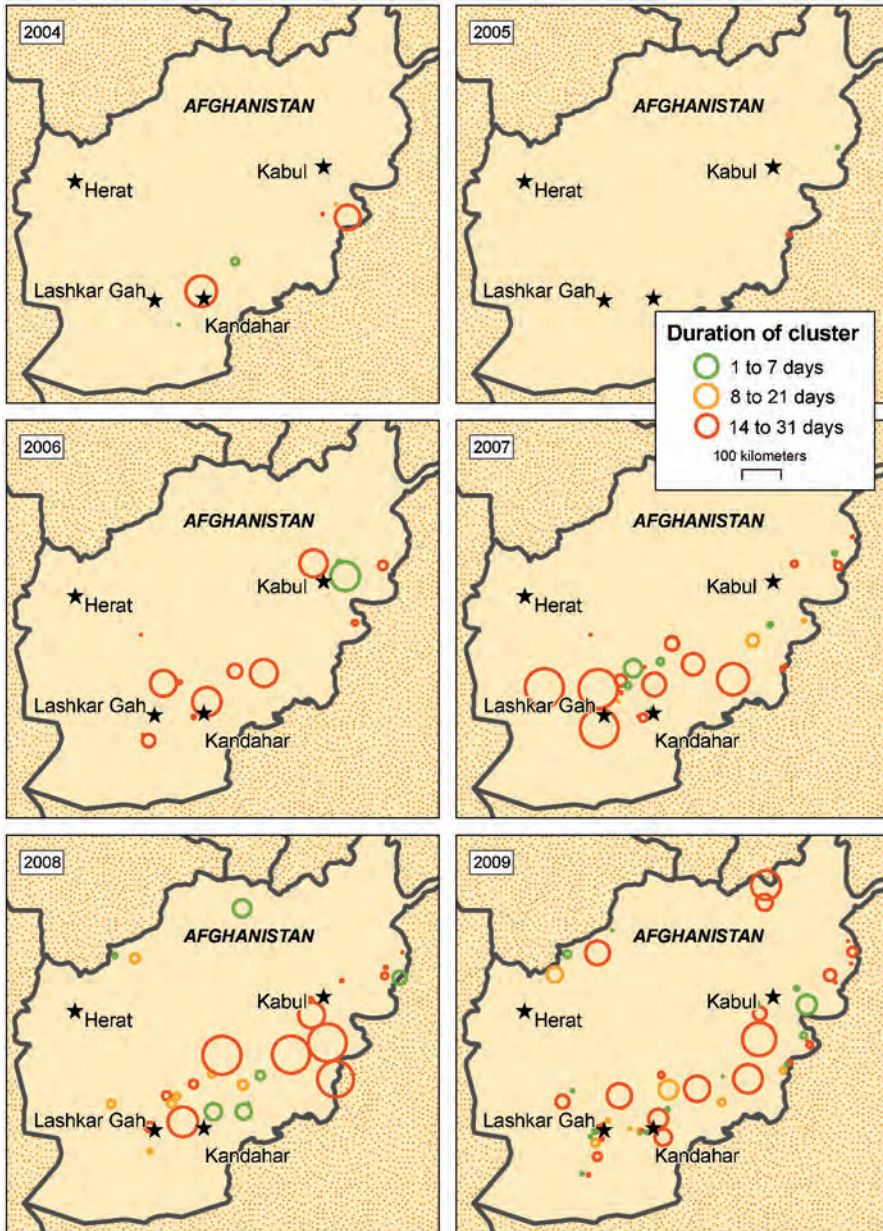


Fig. 14. Space-time scan statistic clusters of insurgent activity, Afghanistan, 2004–2009.

dimensions of the cylinder vary at each location in the data once the procedure begins, and expected (theoretical) and observed event counts are compared for each space-time window. The spatio-temporal clusters that are least likely to have appeared by chance¹⁶ are retained.

¹⁶P-values were assigned via 999 Monte Carlo simulations. Additional technical documentation is available www.colorado.edu/ibs/johno/afpak, and in the SaTScan user guide v9.0.1

By default, clusters may not overlap geographically.¹⁷ For display purposes, we used a Python script to generate shapefiles with a minimum cluster radius of 1 km (Figs. 13 and 14). We replicated our analysis for selected combinations of years and event types and found that the location, duration, and radii of clusters were identical. To mitigate the temporal boundary effect, we included a buffer of one week preceding and following each year. Because of the fine resolution of these data (in 2009, the insurgent category had 13,938 discrete locations), and the large number of records (17,331 events), the space-time scan statistic is computationally intensive, up to 48 hours on an Intel core TM i5 processor with 660 CPU, 3.33.GHz speed, and 4GB RAM.

Conceptually, one would expect some overlap between the clusters of the two types of action, from the ISAF/ANSF coalition and the Taliban and insurgent allies. As in previous cartographic displays of the Afghan violence, the clusters for both sides are found in Kandahar and Helmand provinces and in the eastern regions of Afghanistan along the border with Pakistan. Over time, the ISAF/ANSF significant clusters increase from 0 in 2004 to 17 in 2009, while the number of insurgent clusters grew from 6 to 59 (Figs. 13 and 14).

In 2006, several insurgent clusters of conflict appear north and east of Kabul along the Pakistan border. In 2006 also, insurgent clusters of conflict appear toward the south, especially to the region between Kandahar city and Lashkar Gah. During 2007–2009, the clustering of conflict becomes quite complex. Central and northern Helmand Province were characterized by numerous conflict clusters from 2007–2009. A large cluster of ISAF/ANSF actions, lasting for nearly three weeks, is located south of Herat (near the Iranian border in Farah Province) during December of 2009, a response to the spread of insurgent activity into western regions. By one account, in late 2009, “a small, tented camp was transformed into a giant operations hub” in Delaram, Farah Province (Starkey, 2009).

Clusters of ISAF/ANSF violence typically do not concentrate in major cities. For instance, during 2006–2007 and 2009, clusters of violence appear just outside (within 50 km) of Kandahar as ISAF/ANSF pushed into Taliban strongholds. A large cluster appears in 2009 just north of Kabul, the result of activity in and around the Bagram air base. These trends show that patrols (and other operations originating from Forward Operating Bases) often extend far into rural areas on their missions, suggesting that cities are used as staging points to strike out against the Taliban and allies.

Although the number of insurgent clusters increases almost tenfold between 2004 and 2009, many of the later clusters are quite small (this is especially clear in Helmand Province). These trends of insurgent-initiated activity visible in the space-time scan statistic reinforce pre- and post-WikiLeaks media reports of a revitalized insurgency. Whether the insurgency has become more active because increased troop numbers have presented insurgents with a greater number of targets, or whether the insurgency would have otherwise expanded cannot be known. Nevertheless, as the number of ISAF/ANSF clusters has increased steadily each year since 2004, the number of insurgent clusters has increased at a faster rate. Though sometimes small, insurgent clusters have appeared in relatively remote areas of the country, including the far north, near Kunduz, areas to the north and east of Herat, and near Asadabad north of Jalalabad.

Clusters of insurgent violence in Afghanistan’s south are relatively large in 2007, but become smaller during later years, suggesting a localizing trend. Clusters of conflict in 2009 in northwestern Bagdhis Province, along the border with Turkmenistan, also present a

¹⁷For the insurgent event category, there are two clusters near Kunduz, in the north of Afghanistan, that overlap in space, but do not overlap in time.

noteworthy deviation from previous years. These four clusters reflect the aftermath of a failed peace agreement between the Taliban and the Afghan government after July 2009, resulting in clashes immediately throughout that region of the north (Arnoldy, 2009). In late 2008, hundreds of families were driven from their homes in Moqor district by fighting, visible in a cluster that lasted from November 2 until November 13, 2008 with a radius of 11.5 km. Clustering analysis of the WikiLeaks data also highlights the insurgent response to the now well-known Kunduz airstrike, which took place on September 4, 2009 (Farrel and Opiel, 2009). The large cluster of insurgent events east of Kunduz (Fig. 14) lasted for nearly a month beginning on September 8, 2008.

The presence of dozens of small clusters near Lashkar Gah and other areas of central Helmand Province are also a notable shift from earlier years. Here, clusters of insurgent activity are small (about 3.5 km radius on average) and vary in length from five days to one month. One or more of the clusters occurs during each month of the year, showing quite persistent activity. Within about 50 km of Kandahar in any direction, there also exist several of these small clusters of insurgent violence. While slightly fewer in number, this clustering of activity dominates summer months and the late autumn, rather than consuming most of the year. Each cluster corresponds to a particular series of battles between coalition and Taliban forces as ISAF/ANSF pushed into territory previously held by the Taliban, and also destroyed opium stockpiles that provide much of the Taliban financing (Flynn, 2009). In addition to these localized clusters, geographically broader assemblages of violence remain in other regions of the country. Between May and the beginning of August 2008, the three large clusters along the Pakistani border in Paktia Province each have a radius of nearly 50 km, with the shortest lasting for a full two weeks.

The analysis and graphics of this paper present a view into the Afghan war violence showing where it is increasing, the changing nature of its spatial distribution, and the relative concentrations according to several geographic characteristics. Such synoptic analyses are not possible to achieve even for those who follow media reports closely because of the idiosyncratic nature of individual reports or from the sheer volume of the cumulative data, as evidenced in the massive WikiLeaks file. Only by analyzing large volumes of geographically detailed and temporally sequenced data can one begin to peer into the fog of war and pinpoint the overall spatial and temporal trends.

CONCLUSIONS

Do the WikiLeaks data provide any unique insight into the Afghanistan war? Qualitatively, there is little doubt that the contents of the reports bring to light the disturbing realities of the United States' "longest war." Detailed raw accounts that appear in the "summary" category in each row of the original file have been the focus of the political debate about the probity of the WikiLeaks action. We did not examine these texts, but instead focused our attention on the geographic and temporal elements of the files that are marked by the geographic and time stamps in the individual records. It remains indisputable that the war logs are only a brief account of most of the events, meaning that even the military commands are not fully informed about developments in the field (Shachtman, 2010b). Our comparative quantitative approach can contribute to debates about how (as well as when and where) the previously secret military data can change interpretations of the conflict that heretofore have been based largely on media reports. Both opponents and proponents of the U.S./ISAF involvement in Afghanistan noted the dramatic changes in 2008 and 2009 in the scope of mission, intensity, and geographic range of the conflict. The Obama Administration's response in late 2009 was

a surge of 30,000 troops and the result to date (August 2010) has been a rise in civilian, insurgent, and ISAF casualties. Because the expansion of the war coincided with the arrival of the “Great Recession” in the U.S., the public debate about the American war policy has been muted as political fortunes rest more on economic trends.

In this paper, we have only presented exploratory and descriptive measures of war dynamics over the past six years from a rich dataset whose legal use is still in question. The details in each record (event) allow for more intense analysis that can test hypotheses about the nature of the diffusion process in the conflict patterns, the identification of the leader (first actor) and target (responder) in a tit-for-tat analysis, the regional and local differences in the different weapons use (IEDs, small arms, etc.), and more incisive examination of the role of local (ethnic) communities in the fighting. Academic researchers rarely have access to timely and detailed data on wars, and it is in this regard that the WikiLeaks data are unique. Modern (geo)statistical tools can tease out some of the underlying trends and patterns that remain hidden from the public and the media by the understandable attention to individual personalities and organizations.

REFERENCES

- Arnoldy, Ben.** “Short-Lived Cease-Fire with Taliban Dims Prospects for Broader Deals,” *Christian Science Monitor*, July 27, 2007.
- Biddle, Stephen, Christia Fotini, and J. Alexander Their,** “Defining Success in Afghanistan,” *Foreign Affairs*, **89**, 4:48–51, 2010.
- Boettcher, William, III and Michael Cobb,** “Echoes of Vietnam? Body Counts and Public Perceptions of Success and Failure in the War in Iraq.” Paper presented at the annual meeting of the International Studies Association, Honolulu, Hawaii, March 5, 2005 [http://www.allacademic.com/meta/p69939_index.html], last accessed August 5, 2010.
- “Civilian Casualty Data,” Afghanistan Conflict Monitor (website), July 2010 [<http://www.afghanconflictmonitor.org/civilian.html>], last accessed August 11, 2010.
- Conway, Drew,** “Wikileaks Afghanistan Data,” July 29, 2010 [<http://www.drewconway.com/zia/>], last accessed August 3, 2010 (see also “Wikileaks Attack Data by Year and Type Projected on Afghanistan Regional Map,” August 7, 2010 [<http://www.drewconway.com/zia/>], last accessed August 11, 2010).
- Diggle, Peter,** “A Kernel Method for Smoothing Point Process Data,” *Journal of the Royal Statistical Society*, **34**, 2:138–147, 1985.
- Farrel, Stephen and Richard A. Opper, Jr.,** “NATO Strike Magnifies Divide on Afghan War”. *New York Times*, September 4, 2009.
- Flynn, Michael (Major-General),** *State of the Insurgency: Trends, Intentions and Objectives*. Unclassified Report, International Security Assistance Force, Afghanistan, December 22, 2009 [http://www.wired.com/images_blogs/dangerroom/2010/01/isaf-state-of-the-insurgency-231000-dec.ppt], last accessed August 8, 2010.
- Getis, Arthur and J. Keith Ord,** “The Analysis of Spatial Association by Use of Distance Statistics,” *Geographical Analysis*, **24**, 3:189–206, 1992.
- Global Administrative Areas,** *GADM Database of Global Administrative Areas*, 2010 [<http://www.gadm.org>], last accessed August 6, 2010.
- Hendler, Clint,** “The Story Behind the Publication of the WikiLeaks’ Afghanistan Logs: From Brussels, to a Bumkiet, to Blockbusters,” *Columbia Journalism Review*, July 28, 2010 [http://www.cjr.org/campaign_desk/the_story_behind_the_publicati.php], last accessed August 2, 2010.
- Jaffe, Greg,** “U.S. Retreat from Afghan Valley Marks Recognition of Blunder,” *The Washington Post*, April 15, 2010.
- Jones, Jeffrey M.,** “In U.S., New High of 43% Call Afghan War a ‘Mistake’,” Gallup website, August 3, 2010 [<http://www.gallup.com/poll/141716/>], last accessed August 13, 2010.

- Kulldorff, Martin, Richard Heffernan, Jessica Hartman, Renato Assuncao, and Farzad Mostashari**, "A Space-Time Permutation Scan Statistic for Disease Outbreak Detection," *PLoS Medicine*, **2**, 3:216–224, 2005.
- Kulldorff, Martin and Information Management Services**, *SaTScan: Software for the Spatial and Space–Time Scan Statistics*, v 9.0.1, 2009 [http://www.satscan.org/], last accessed August 5, 2010.
- NOAA/USAF (NOAA National Geophysical Data Center and U.S. Air Force Weather Agency)**, "DMSP-OLS Nighttime Lights Time Series, Version 4," 2009 [http://www.ngdc.noaa.gov/dmsp/downloadV4composites.html], last accessed May 24, 2010.
- Nordland, Rod**, "Rising Civilian Casualties in Afghanistan, UN Reports," *The New York Times*, August 10, 2010.
- O'Loughlin, John, and Frank D. W. Witmer**, "The Localized Geographies of Violence in the North Caucasus of Russia, 1999–2007," *Annals of the Association of American Geographers*, **100**, 4, 2010, (forthcoming).
- O'Loughlin, John, Frank D. W. Witmer, and Andrew M. Linke**, "The Afghanistan-Pakistan Wars, 2008–2009: Micro-Geographies, Conflict Diffusion, and Clusters of Violence," *Eurasian Geography and Economics*, **51**, 4: 437–471, 2010.
- Ord, J. Keith and Arthur Getis**, "Local Spatial Autocorrelation Statistics: Distributional Issues and Application," *Geographical Analysis*, **27**, 4:286–306, 1995.
- Raleigh, Clionadh, Andrew Linke, Håvard Hegre, and Joachim Carlson**, "Introducing ACLED: An Armed Conflict Event Location Dataset," *Journal of Peace Research*, **47**, 7, 2010a (forthcoming).
- Raleigh, Clionadh, Frank D. W. Witmer, and John O'Loughlin**, "The Spatial Analysis of War," in Robert A. Denemark, ed., *The International Studies Encyclopedia*, Vol. X. Oxford, UK: Wiley-Blackwell, 2010b, 6534–6553.
- Rich, Frank**, "Kiss This War Goodbye," *The New York Times*, August 1, 2010, WK8.
- Rubin, Alissa**, "U.S. Forces Close Post in Afghan 'Valley of Death'," *The New York Times*, April 14, 2010.
- Shachtman, Noah**, "Afghan Insurgency Can Sustain Itself Indefinitely': Top U.S. Intel Officer," *Wired.com*, January 8, 2010a [http://www.wired.com/dangerroom/2010/01/afghan-insurgency-can-sustain-itself-indefinitely-top-us-intel-officer/#ixzz0w9hdQ6wK], last accessed August 9, 2010.
- Shachtman, Noah**, "What I Saw at Moba Khan: The Military Reports Highlighted by WikiLeaks Don't Provide a Full Picture of the War," *Wall Street Journal*, July 28, 2010b.
- Starkey, Jerome**, "U.S. Army Builds Smalltown America in the Battlefield." *The Times (London)*, December 3, 2009.
- "The War Logs. The Afghan Struggle: A Secret Archive," *The New York Times*, July 26, 2010, A1, A8–A12.
- USGS (U.S. Geological Survey)**, "Shuttle Radar Topography Mission, 30 Arc Second Scene." College Park, MD: Global Land Cover Facility, University of Maryland, 2004.
- Weidmann, Nils B., Jan Ketil Rød, and Lars-Erik Cederman**, "Representing Ethnic Groups in Space: A New Dataset," *Journal of Peace Research*, **47**, 4:491–499, 2010.